

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Ahmad Islam

Entitled Theory and Characterization of Random Defect Formation and Its Implication in Variability of Nanoscale Transistors

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

M. A. Alam

Chair

M. S. Lundstrom

S. Datta

P. Ye

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): M. A. Alam

Approved by: V. Balakrishnan

Head of the Graduate Program

4/27/10

Date

**PURDUE UNIVERSITY
GRADUATE SCHOOL**

Research Integrity and Copyright Disclaimer

Title of Thesis/Dissertation: Theory and Characterization of Random Defect Formation and Its Implication in Variability of Nanoscale Transistors

For the degree of Doctor of Philosophy

I certify that in the preparation of this thesis, I have observed the provisions of *Purdue University Executive Memorandum No. C-22*, September 6, 1991, *Policy on Integrity in Research*.*

Further, I certify that this work is free of plagiarism and all materials appearing in this thesis/dissertation have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis/dissertation is in compliance with the United States' copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless Purdue University from any and all claims that may be asserted or that may arise from any copyright violation.

Ahmad Islam

Signature of Candidate

4/27/10

Date

*Located at http://www.purdue.edu/policies/pages/teach_res_outreach/c_22.html

THEORY AND CHARACTERIZATION OF RANDOM DEFECT FORMATION AND
ITS IMPLICATION IN VARIABILITY OF NANOSCALE TRANSISTORS

A Dissertation
Submitted to the Faculty
of
Purdue University
by
Ahmad Ehteshamul Islam

In Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy

May 2010
Purdue University
West Lafayette, Indiana

Dedicated to my parents
Amirun Nessa Begum & Ahmad Nurul Islam
and my (little) brother
Ahmad Neamul Islam

for their support, sacrifice and inspiration throughout the period of my study

ACKNOWLEDGEMENTS

This work reflects the inspiration and technical guidance from many people I meet up during my graduate years at Purdue University and also before that. I was fortunate to have Prof. Ashraf Alam as my graduate advisor, who is guiding me through this period with extreme professionalism. I admire his approach in selecting/solving a problem with great deal of preparation and future plan, which is complemented by his technical skills and diligence. To me he is an ideal mentor, whose guidance (I believe) can help me pass through many difficult situations in my academic, as well as professional life.

I gratefully acknowledge the research support we are receiving from Prof. Souvik Mahapatra and his students at IIT-Bombay (Gautam Gupta, Dr. Vrajesh D. Maheta, Shweta Deora, Hitesh Das, and Naresh Kumar). Without these experimental supports, most of the works performed in this thesis might have been impossible. In this regard, I would also like to thank Dr. Anthony Oates, Dr. J. C. Lin, Dr. J. H. Lee, and Dr. W. H. Wu of TSMC and Dr. Khaled Ahmed of Applied Materials for sharing us their confidential information and thus enable us to get a broad picture of the current (advanced) technology. I am also indebted to Prof. Mark Lundstrom and Changwook Jeong for sharing their views and providing us the time to discuss many important problems. It was also a pleasure to work with Charles Augustine, Dr. Kunhyuk Kang, and Prof. Kaushik Roy in the area of digital circuits and memories and I expect many significant outcomes from this collaborative effort. I also had many fruitful discussions with Dr. Anand Krishnan and Dr. Srikant Krishnan from Texas Instruments and thanks to them for providing the time within their busy schedule. Thanks to Prof. Supriyo Datta and Prof. Peide (Peter) Ye for serving my advisory committee and proving suggestions based on their own expertise.

I gratefully acknowledge the inspiration and feedback from Dr. Haldun Kufluoglu, who really helped me to understand our research area at the beginning period of my graduate studies. He, not only extended his help for settling down in a new research environment, but also provided me all sorts of personal supports, whenever needed. In this regard, I would also like to thank Dr. Dhanoop Varghese, Dr. Nauman Z. Butt, and Muhammad Masuduzzaman who helped me understand many new research topics and always provided time to resolve the experimental problems I faced over the years, while working in the lab.

I specially thank Abu Naser Zainuddin, Prof. Sayeef Salahuddin, and Dr. Lutfi Ahmed Siddiqui for their inspiration and all sorts of support (educational, as well as personal). I would also like to acknowledge my discussion with Dr. Sayed Hasan, Prof. Shaikh S. Ahmed, Dr. Pradeep Nair, Dr. Siyuranga Koswatta, Kurtis Cantley, Yang Liu, and Dr. Yangqing Wu on many important problems. In addition, I feel fortunate for being able to use the experimental facilities at Birck Nanotechnology Center and the computational facilities provided by Network for Computational Nanotechnology; without these any progress in my research would have been impossible. I also acknowledge the administrative helps from Cheryl Haines and Vicki Johnson, which I am receiving since coming to Purdue. Thanks also to the proof-readers of this thesis (Dr. Pradeep Nair, Dr. Dhanoop Varghese, Muhammad Masuduzzaman, Dhanya Varghese, Biswajit Ray, Sambit Palit, Sourabh Dongaonkar, and Ankit Zain) for their efforts in minimizing the typos.

Also, I would gratefully recognize contributions from Prof. Anisul Haque, who introduced me to the area of nanotechnology; which I still acknowledge as a stepping stone for my career. In addition, I am also indebted to Prof. Rezwana Khan, Prof. Kazi Mujibur Rahman, Prof. Mohammad Ali Chowdhury, and Prof. Enamul Basher for their guidance and teaching, which helped me to build the foundation in the area of electrical engineering during my undergraduate years at BUET.

No research works can be performed without proper funding. And here, I would like to recognize the funding from Taiwan Semiconductor Manufacturing Corporation

(TSMC), Applied Materials (AMAT), Nanoelectronic Research Initiative (NRI), 2008 IEEE Electron Device Society PhD Fellowship, and 2009-2010 Intel Foundation PhD Fellowship for supporting my research.

Finally, I would like to thank my entire family members for their inspiration. Here, the special thanks go to my mother Amirun Nessa Begum, to my father Ahmad Nurul Islam, and to my brother Ahmad Neamul Islam for their support and sacrifice whole through my life. My PhD life got a little twig after I got married in the final year, and in this regard I thank my wife Farzana Adil for her support and encouragement during that period. I feel proud for having such a family and without their support, I cannot imagine to come this far. And last, but not least, I thank Almighty for enabling me pass through many tough periods and thus providing me the mental strength to pursue higher education.

I apologize in advance to those, whose names I might have forgotten at this moment. But it doesn't change the fact that they helped me in one or many ways, which I should be able to recall at appropriate moments of my life.

TABLE OF CONTENTS

| | Page |
|---|-------|
| LIST OF TABLES..... | xvi |
| LIST OF FIGURES..... | xvii |
| ABSTRACT..... | xxxiv |
| 1. INTRODUCTION..... | 1 |
| 1.1. Background..... | 1 |
| 1.1.1. Process-induced Spatial Variability..... | 2 |
| 1.1.2. Defect-related Temporal Variability..... | 4 |
| 1.1.2.1. History of Defects..... | 4 |
| 1.1.2.2. Sources of Temporal Variability..... | 5 |
| 1.1.2.3. Optimizing the Effect of Defect..... | 6 |
| 1.1.2.4. Defects in Analogous Areas..... | 7 |
| 1.2. Objective of the Thesis..... | 7 |
| 1.3. Organization of Thesis..... | 8 |
| 2. ORIGIN OF DEFECTS IN TRANSISTORS..... | 10 |
| 2.1. Introduction..... | 10 |
| 2.2. Interface Defects: Si/SiO ₂ Structure..... | 11 |
| 2.2.1. Nature of Si/SiO ₂ Interface Defects..... | 11 |
| 2.2.2. Energetics of Si/SiO ₂ Interface Defects..... | 12 |
| 2.2.3. Theory of Donor/Acceptor Defects..... | 13 |
| 2.3. Interface Defects: Strained Si/SiO ₂ Structure..... | 15 |

| | |
|--|----|
| 2.4. Interface Defects: Si/SiON Structure | 16 |
| 2.5. Interface Defects: Si/high- κ Structure | 16 |
| 2.6. Oxide Defects in SiO ₂ | 17 |
| 2.7. Oxide Defects in High- κ Material | 18 |
| 2.8. Summary | 19 |
| 3. INTERFACE DEFECTS | 20 |
| 3.1. Introduction | 20 |
| 3.2. Existing Views of Interface Defect Formation..... | 21 |
| 3.3. First Reaction-Diffusion (R-D) Model..... | 21 |
| 3.4. Si-H Dissociation in R-D Model..... | 24 |
| 3.4.1. Hole-assisted Si-H Dissociation | 24 |
| 3.4.2. Dopant-activated Si-H Dissociation | 25 |
| 3.5. Diffusing Species in R-D Model..... | 27 |
| 3.5.1. Diffusion of Neutral Atomic Hydrogen (H)..... | 27 |
| 3.5.2. Drift of Charged Atomic Hydrogen (H ⁺) | 28 |
| 3.5.3. Diffusion of Molecular Hydrogen (H ₂) | 30 |
| 3.5.4. H-H ₂ Diffusion with Direct H-H ₂ Conversion | 32 |
| 3.5.5. H-H ₂ Diffusion with H ₂ Conversion at Poly/Oxide Interface | 38 |
| 3.5.6. Diffusing Species in R-D Model: Summary | 39 |
| 3.6. Recovery (or Reverse-Annealing) of Interface Defects | 40 |
| 3.6.1. Analytical Formalism | 40 |
| 3.6.2. N_{IT} Recovery: Direct H-H ₂ R-D Model..... | 42 |
| 3.6.3. N_{IT} Recovery: Poly H-H ₂ R-D Model..... | 44 |
| 3.6.4. N_{IT} Recovery: Combined (Direct+Poly) H-H ₂ R-D Model..... | 46 |

| | Page |
|---|------|
| 3.6.5. N_{IT} Recovery: R-D Theory vs. NBTI Experiment..... | 48 |
| 3.7. Field Dependence in H-H ₂ R-D Model | 50 |
| 3.7.1. Field vs. Voltage-Dependent N_{IT} Generation | 50 |
| 3.7.2. Hole-assisted, Field-enhanced, Thermal Si-H Dissociation..... | 51 |
| 3.7.3. Experimental Validation of Phenomenological $E\exp(\gamma E)$ Model | 53 |
| 3.7.4. Theoretical Estimation of a (Figure 3.18) | 56 |
| 3.7.5. Theoretical Estimation of γ_T | 57 |
| 3.7.6. Experimental Determination of γ_T and a | 60 |
| 3.7.7. Summary: Field Dependence of ΔN_{IT} | 62 |
| 3.8. Application of $E\exp(\gamma E)$ Model: Dielectric Material Dependence of ΔN_{IT} | 62 |
| 3.8.1. Determining $E\exp(\gamma E)$ Model Parameters..... | 62 |
| 3.8.2. Variation in $E\exp(\gamma E)$ Model Parameters with $\%N$ | 64 |
| 3.8.3. $E\exp(\gamma E)$ Model Parameters: Physical Origin of $\%N$ Variation | 65 |
| 3.9. Validation of $E\exp(\gamma E)$ Model: Strain Dependence of ΔN_{IT} | 68 |
| 3.9.1. ΔN_{IT} for Uniaxial Compressive Strain..... | 68 |
| 3.9.2. ΔN_{IT} for Biaxial Compressive Strain | 73 |
| 3.9.3. Summary: Strain Dependence of ΔN_{IT} | 74 |
| 3.10. Predicting Lifetime Using Field Model..... | 75 |
| 3.11. Summary | 76 |
| 4. STATISTICS OF INTERFACE DEFECT GENERATION | 77 |
| 4.1. Introduction | 77 |
| 4.2. MCMC Transition Matrix | 78 |
| 4.2.1. Transition Matrix: H Diffusion | 78 |

| | Page |
|--|------|
| 4.2.2. Transition Matrix: H-H ₂ Diffusion | 81 |
| 4.3. Statistics of R-D System: H Diffusion | 83 |
| 4.3.1. Analytical Calculation | 83 |
| 4.3.1.1. Statistics for $t_{STS}@N_{IT}$ | 84 |
| 4.3.1.2. Statistics for $N_{IT}@t_{STS}$ | 85 |
| 4.3.2. MCMC Simulation Results | 87 |
| 4.3.2.1. Consistency with Continuum Simulation | 87 |
| 4.3.2.2. Simulated Statistics of $t_{STS}@N_{IT}$ | 87 |
| 4.3.2.3. Simulated Statistics of $N_{IT}@t_{STS}$ | 89 |
| 4.3.2.4. Comparison Between Analytical and Numerical Results | 91 |
| 4.3.3. On The Existence of Correlated N_{IT} Creation | 94 |
| 4.4. Statistical Signatures of R-D: H-H ₂ Diffusion | 95 |
| 4.5. Summary | 97 |
| 5. PRE-EXISTING OXIDE DEFECTS | 99 |
| 5.1. Introduction | 99 |
| 5.2. Existing Views of N_{HT} Dynamics..... | 99 |
| 5.3. Tewksbury's Model [230]..... | 101 |
| 5.4. Effect of Gate Electrode in Hole Trapping | 102 |
| 5.4.1. Shockley-Read-Hall (SRH) Analysis..... | 102 |
| 5.4.2. Verification of Hole Trapping Model..... | 104 |
| 5.5. Detailed Hole Trapping Model..... | 107 |
| 5.5.1. Computational Framework..... | 107 |
| 5.5.2. Simulation Results..... | 109 |
| 5.6. Summary | 111 |

| | |
|--|-----|
| 6. DEFECT CHARACTERIZATION..... | 113 |
| 6.1. Introduction | 113 |
| 6.2. Defect Characterization Techniques | 114 |
| 6.3. Spin Resonance Techniques ($t_{meas} \sim$ minutes)..... | 115 |
| 6.3.1. Electron Spin Resonance (ESR)..... | 115 |
| 6.3.2. Electrically Detected Magnetic Resonance (EDMR)..... | 116 |
| 6.4. Measure-Stress-Measure (MSM) Technique ($t_{meas} \sim$ msec-sec) | 118 |
| 6.4.1. Determination of E_{eff} | 120 |
| 6.4.2. Determination of V_T , SS , m , μ_{eff} | 121 |
| 6.4.3. Variation in μ_{eff} Model Parameters..... | 122 |
| 6.4.4. Determining Physical Basis of $\Delta\mu_{eff}(N_{IT})$ | 125 |
| 6.5. Ultra-Fast V_G or UFV ($t_{meas} \sim 0.5 \mu s$) | 126 |
| 6.6. On-the-fly $I_{D,lin}$ or OTF- $I_{D,lin}$ ($t_{meas} \sim 0$)..... | 127 |
| 6.7. Interpreting Measurements..... | 128 |
| 6.7.1. Ultra-fast V_G (UFV): Body-effect Coefficient (m) Correction | 130 |
| 6.7.2. Ultra-fast V_G (UFV): Valence Band Electron Trapping | 132 |
| 6.7.2.1. Valence Band Electron Trapping: Process..... | 133 |
| 6.7.2.2. Valence Band Electron Trapping: Numerical Model | 136 |
| 6.7.2.3. Valence Band Electron Trapping: Simulation Results..... | 137 |
| 6.7.2.4. Valence Band Electron Trapping: Resolving UFV vs. OTF- $I_{D,lin}$ Relaxation | 139 |
| 6.7.3. On-the-fly $I_{D,lin}$: Time-zero (t_0)-delay Effect..... | 140 |
| 6.7.4. On-the-fly $I_{D,lin}$ (OTF- $I_{D,lin}$): Mobility Correction in ΔV_T estimation..... | 144 |
| 6.7.4.1. Algorithm for ΔV_T Estimation from ΔI_D | 144 |

| | Page |
|---|------|
| 6.7.4.2. Verifying $\Delta I_{D,lin} \rightarrow \Delta V_T$ Algorithm Using MSM | 145 |
| 6.7.4.3. Verifying $\Delta I_{D,lin} \rightarrow \Delta V_T$ algorithm using OTF- $I_{D,lin}$ | 147 |
| 6.7.4.4. Reliability Implications of Proper ΔV_T Estimation | 149 |
| 6.7.4.5. Compact ΔV_T Estimation Algorithm for Type-I Transistors..... | 152 |
| 6.7.5. N_{HT}/N_{IT} Separation for NBTI Degradation | 154 |
| 6.7.5.1. Basis for N_{HT}/N_{IT} Decomposition Method | 157 |
| 6.7.5.2. N_{HT}/N_{IT} Separation: NBTI Stress Phase | 163 |
| 6.7.5.3. N_{HT}/N_{IT} Separation: NBTI Relaxation Phase | 165 |
| 6.7.5.4. Implications for AC/DC Ratio | 169 |
| 6.7.5.5. Comparison with Existing N_{HT}/N_{IT} Separation Techniques | 172 |
| 6.8. Suggestions for NBTI Characterization | 172 |
| 6.9. Summary | 173 |
| 7. DEFECT FORMATION AND OFF-STATE POWER DISSIPATION (GATE LEAKAGE)..... | 174 |
| 7.1. Introduction | 174 |
| 7.2. Motivation | 175 |
| 7.3. Experimental Detail..... | 176 |
| 7.4. Gate Leakage in Plasma Oxynitride Transistors | 177 |
| 7.4.1. Estimation of device model parameters | 177 |
| 7.4.2. Variation of Parameters with $\%N$ | 178 |
| 7.4.3. Gate Leakage Variation | 183 |
| 7.5. NBTI Degradation..... | 184 |
| 7.5.1. Calculation of Electric Fields | 185 |
| 7.5.2. NBTI Model Parameters & Their $\%N$ Dependency..... | 186 |

| | Page |
|--|------|
| 7.5.3. Safe Operating Condition | 188 |
| 7.6. Optimization of gate leakage and NBTI..... | 190 |
| 7.7. Analytical Construction of Design Diagram | 191 |
| 7.7.1. Model for electric fields | 192 |
| 7.7.2. Analytical Model for Gate Leakage Current | 193 |
| 7.7.3. Calculation of V_{safe} | 193 |
| 7.8. Summary | 194 |
| 8. DEFECT FORMATION AND TRANSISTOR PERFORMANCE | 195 |
| 8.1. Introduction | 195 |
| 8.2. Defects on V_T and SS Variation | 195 |
| 8.3. Defects on Off Current Variation..... | 198 |
| 8.4. Defects on Mobility Variation..... | 198 |
| 8.4.1. Effect of N_{IT} (NBTI Stress) | 198 |
| 8.4.2. Effect of N_{ET} (PBTI Stress) | 201 |
| 8.4.3. Effect of N_{HT} (NBTI Stress) | 203 |
| 8.5. Effect of Defect on I_{ON} : Drift-diffusion Theory | 204 |
| 8.5.1. $I_{D,lin}$ Variation due to N_{IT} | 204 |
| 8.5.2. $I_{D,sat}$ Variation due to N_{IT} | 206 |
| 8.5.3. $I_{D,lin}$ and $I_{D,sat}$ Variation due to N_{ET} | 207 |
| 8.6. Effect of Defect on I_{on} : Scattering Theory | 207 |
| 8.7. Summary | 209 |
| 9. TOWARDS VARIATION RESILIENT TRANSISTORS AND CIRCUITS. 211 | |
| 9.1. Introduction | 211 |
| 9.2. Motivation | 211 |

| | Page |
|--|------|
| 9.3. Principle of Variation Resilience | 213 |
| 9.4. Increasing μ_{eff} - E_{eff} Steepness: Strain | 215 |
| 9.4.1. Qualitative Theory | 216 |
| 9.4.2. Quantitative Analysis of Self-Compensation | 217 |
| 9.4.3. Self-Compensation at Practical Strain Limit | 221 |
| 9.5. Increasing μ_{eff} - E_{eff} Steepness: Temperature..... | 222 |
| 9.6. Sources of V_T Fluctuation: Show-stopper for CMOS Scaling?..... | 223 |
| 9.7. Self-Compensation in $I_{D,lin}$ Fluctuation | 225 |
| 9.7.1. EOT/ Φ_{GS} Variation..... | 225 |
| 9.7.2. Random Dopant Variation..... | 227 |
| 9.7.3. Defect-induced ΔV_T | 229 |
| 9.7.4. 1/f Noise | 229 |
| 9.8. Self-Compensation: From Transistor to Digital Electronics..... | 231 |
| 9.8.1. Effect of N_{IT} on Single-Stage Inverter..... | 232 |
| 9.8.2. Effect of N_{IT} on SRAM performance | 233 |
| 9.8.3. Large-scale SPICE analysis: Issues with Unphysical ΔE_{eff} | 235 |
| 9.8.4. Circuit Analysis: Using Physics-based ΔE_{eff} | 237 |
| 9.9. Summary | 239 |
| 10. THESIS SUMMARY AND FUTURE WORK | 241 |
| 10.1. Thesis Summary | 241 |
| 10.1.1. Reaction-Diffusion Modeling of Interface Defect..... | 241 |
| 10.1.2. NBTI for Oxynitride Dielectric Transistors | 242 |
| 10.1.3. NBTI for Strained Transistors | 242 |
| 10.1.4. Interpretation of NBTI Measurements | 242 |

| | Page |
|---|------|
| 10.1.5. Concept of Variation-Resilient Transistor..... | 243 |
| 10.2. Future Work | 244 |
| 10.2.1. Statistical Distribution Si-H Bonds..... | 245 |
| 10.2.2. Research on Variation Resilience..... | 248 |
| 10.2.3. Extension/Implication of Hole Trapping Analysis..... | 250 |
| 10.3. Epilogue..... | 251 |
| LIST OF REFERENCES..... | 252 |
| A. DEFINITION OF PARAMETERS..... | 288 |
| A.1. Effective Oxide Thickness (EOT)..... | 288 |
| A.2. Lifetime (t_{life})..... | 288 |
| A.3. Voltage Acceleration (I_V) | 288 |
| A.4. Safe Operating Voltage (V_{safe})..... | 289 |
| A.5. Activation Energy (E_A)..... | 289 |
| B. ALTERNATE MODELS FOR N_{IT} FORMATION | 290 |
| B.1. R-D Model: Dispersive Drift/diffusion | 290 |
| B.2. Bond Dispersion Model | 294 |
| B.3. Statistical Mechanics Based Model..... | 295 |
| B.4. Energy-level Perspective of N_{IT} Formation..... | 297 |
| B.4.1. Energy-well Representation..... | 297 |
| B.4.2. Energy-well Representation of R-D Model | 299 |
| B.4.2.1. Energy-well Representation of R-D Model (H Diffusion) | 300 |
| B.4.2.2. Energy-well Representation of R-D Model (H^+ Drift) | 302 |
| B.4.3. Energy-level Perspective: Revisited | 304 |
| B.4.3.1. Double-well Model | 304 |

| | Page |
|---|------|
| B.4.3.2. Triple Well Model..... | 306 |
| C. DRIFT/DIFFUSION SYSTEM USING GENERATING FUNCTION | 308 |
| C.1. Calculation of Diffusion Flux: | 311 |
| C.2. Calculation of Drift/Diffusion Flux: | 313 |
| D. MOMENTS OF STATISTICAL DISTRIBUTION: USING GENERATING FUNCTION..... | 315 |
| E. PARAMETRIC CHARACTERIZATION OF SCATTERING THEORY .. | 317 |
| E.1. Experimental Details | 318 |
| E.2. Variation of Ballistic Efficiencies | 319 |
| E.3. Variation of λ | 321 |
| E.4. Variation of ℓ | 322 |
| E.5. Scattering Theory Analysis: Summary | 323 |
| VITA..... | 324 |

LIST OF TABLES

| Table | Page |
|--|------|
| 3.1: Effect of diffusion species and its (i.e. diffusing species) generation mechanism on N_{IT} formation power-law time exponent, n | 39 |
| 3.2: Band-structure information for strained transistors | 59 |
| 3.3: Comparing a determined by various methodologies | 60 |
| 3.4: Properties ($\%N$, T_{PHY} and EOT) of plasma oxynitride dielectric for the type-I transistors used in section 3.8. | 63 |
| 6.1: Properties of plasma oxynitride dielectric of the type-I transistors used in sections 6.4, 6.7.1, 6.7.3, and 6.7.4: N_2 Dose ($\%N$), physical oxide thickness (T_{PHY}), effective oxide thickness (EOT) | 120 |
| 6.2: Properties ($\%N$, T_{PHY} and EOT) of plasma oxynitride dielectric for the transistors used in section 6.7.5..... | 158 |
| 7.1: Model parameters used for characterizing MOS structures..... | 178 |
| 9.1: Capacitance values in Intel's recent CMOS technologies. | 225 |
| Appendix Table | |
| C.1: Transition probabilities from lattice site ' i ' in Figure C.1. | 309 |

LIST OF FIGURES

| Figure | Page |
|---|------|
| 1.1: Different types of variabilities in nanoscale transistors..... | 2 |
| 2.1: Structure of paramagnetic centers on Si wafers for (a) (111) and (b) (100) Si/SiO ₂ interfacial orientation. Taken from [69]. Structure of P _{b1} center in Fig. b is the one proposed in [78], whereas study in [80] indicates P _{b1} center also having three back-bonded Si atom, like P _b and P _{b0} centers..... | 12 |
| 2.2: (a) Formation energies and (b) concentration of dangling bonds (T_3) as a function of E_{FS} | 15 |
| 2.3: (a) Oxygen vacancy-free SiO ₂ configuration. (b-f) Possible types of oxide defects or E' centers, present in SiO ₂ dielectric. (b) Oxygen vacancy or Si-Si bridge is considered to be origin of many E' centers. (c) Capture of holes by oxygen vacancy leads to positively-charged E' center. Other schematics in this figure correspond to the illustration (d) for a neutral E' center, (e) for a E' center (74 G doublet), and (f) for a E' center (10.4 G doublet). These configurations are taken from [69, 81, 86]. | 18 |
| 3.1: (a) Schematic representation of Si-H bond dissociation and resulting diffusion of H through dielectric and poly-Si. (b) Diffusion front at time t reaches $\sqrt{D_H t}$ and the diffusion profile can be approximated using triangle, which is used in obtaining total hydrogen (hence, N_{IT}) concentration in equation (3.6)..... | 28 |
| 3.2: (a) Schematic representation of Si-H bond dissociation and resulting drift (dominates over diffusion) of H ⁺ through dielectric and poly-Si. (b) Drift front at time t reaches $\mu_{H^+} E_{ox} t$ and can be approximated using rectangle [125], which is used to obtain total hydrogen (hence, N_{IT}) concentration in equation (3.15). | 29 |
| 3.3: (a) Schematic representation of Si-H bond dissociation and resulting diffusion of H ₂ through dielectric and poly-Si. (b) Diffusion front at time t reaches $\sqrt{D_{H_2} t}$ and the diffusion profile can be approximated using triangle, which is used in obtaining total hydrogen (hence, N_{IT}) concentration in deriving equation (3.22). | 32 |
| 3.4: (a) Schematic representation of Si-H bond dissociation and resulting diffusion of hydrogen species (H and H ₂) through dielectric and poly-Si [143] during NBTI stress. (b) At long stress time, diffusion of H ₂ (concentration N_{H_2}) dominates and concentration of H (N_H) is small..... | 34 |

| Figure | Page |
|---|------|
| 3.5: (a) Simulation of generalized R-D model (H-H ₂ system) explains the experimental trends [140] at any stress time. R-D model parameters (consistent with literature [142, 143]): $k_F = 6 \times 10^{-3} E_c \exp(0.65 E_{ox}) \text{ sec}^{-1}$ (E_c : oxide electric field due to mobile carriers, E_{ox} : total oxide electric field), $N_0 = 5 \times 10^{12} \text{ cm}^{-2}$, $k_R = 3 \times 10^{-9} \text{ cm}^3 \text{ sec}^{-1}$, $D_H = 3 \times 10^{-13} \text{ cm}^2 \text{ sec}^{-1}$, $D_{H2} = 1.8 \times 10^{-14} \text{ cm}^2 \text{ sec}^{-1}$, $k_H = 1.4 \times 10^{-3} \text{ cm}^3 \text{ sec}^{-1}$, $k_{H2} = 95.4 \text{ sec}^{-1}$. (b) Comparison of numerical solution (lines) of equations (3.24)-(3.26) with analytical solution of equation (3.30) (lines with symbols)..... | 35 |
| 3.6: Comparison of numerical results (lines) from equations (3.24)-(3.26), with analytical results (lines with symbols) for $k_H N_{H(0)}^2$ and $k_{H2} N_{H2(0)}$. Although analytical $N_{H2(0)}$ prediction is incorrect at short stress time (whereas, $N_{H(0)}$ prediction is excellent), dominance of $k_H N_{H(0)}^2$ over $k_{H2} N_{H2(0)}$, still makes the analytical solution of N_{IT} appropriate for all stress time (see Figure 3.5b). | 37 |
| 3.7: Approximate hydrogen profile at the start ($t = t_{STS}$) and after a recovery time of t_{REC} (i.e., at $t = t_{STS} + t_{REC}$) of N_{IT} relaxation. See text for the definition of other parameters..... | 41 |
| 3.8: (a) Simulation of N_{IT} recovery using direct H-H ₂ R-D model for a transistor having 1.4 nm gate dielectric. N_{IT} is generated at $V_{STS} = -2.3\text{V}$ for $t_{STS} = 1000\text{sec}$ at 125°C , before it is allowed to recover at two different V_{REC} . Significant amount of relaxation starts around $t_{NIT,start} \sim 1\text{sec}$. (b) A trend towards complete N_{IT} recovery is observed at $V_{REC} = 0$ | 43 |
| 3.9: (a) H ₂ Profile at the end of stress for a transistor having 1.4nm gate dielectric after $t_{STS} = 1000 \text{ sec}$. (b) Percentage of H ₂ within distance 0 to z , for different z (i.e., distance from the interface). | 43 |
| 3.10: (a) Time exponent variation for poly H-H ₂ R-D model for a transistor having 2nm gate oxide. (b) Percentage of H ₂ within distance 0 to z , for different z (i.e., distance from the interface). | 45 |
| 3.11: (a) Comparison of N_{IT} recovery simulations for the direct H-H ₂ R-D model (solid line) and the poly H-H ₂ R-D model (dashed line). $t_{NIT,start}$ for poly H-H ₂ model is order of magnitude smaller than the one for direct H-H ₂ model. (b) There is no change in $t_{NIT,start}$ when V_T calculation incorporates defects at the oxide/poly interface, in addition to the defects at the substrate/oxide interface (poly R-D: Modified V_T)..... | 45 |
| 3.12: (a) There is discontinuity in the stress phase simulation for the combined (direct + poly) H-H ₂ R-D model (compared to direct H-H ₂ and poly H-H ₂ R-D model). (b) Here, the relaxation also happens very quickly compared to the other H-H ₂ version of R-D model..... | 47 |
| 3.13: (a) Ultra-fast V_T measurement [170] predicts $t_{REC,start} \sim 10^{-10} t_{STS}$, which is much smaller than $t_{NIT,start} \sim 10^{-4} t_{STS}$, predicted by poly H-H ₂ R-D model. (b) Similarly, on-the-fly $I_{D,lin}$ predicts a recovery voltage (V_{rec}) dependent relaxation and $t_{REC,start} \sim (10^{-9}-10^{-6}) t_{STS}$. However, consideration of valence band electron trapping (see section 6.7.2) and hole detrapping from oxide defects (see section | |

| Figure | Page |
|---|------|
| 6.7.5), even for type-I transistors, can significantly reduce the difference between $t_{REC,start}$ and $t_{NIT,start}$ | 49 |
| 3.14: Presence of non-zero saturation during N_{IT} relaxation (figure taken from [171]) is a feature sometimes observed in NBTI experiments. | 49 |
| 3.15: Universal view of N_{IT} generation under NBTI stress in strained/unstrained transistors. (a) Schematic of PMOS, showing the application of uniaxial compressive strain using $Si_{1-y}Ge_y$ Source/Drain or biaxial tensile strain using $Si_{1-x}Ge_x$ epi-layer. (b) Dissociation mechanism for Si-H bond (see Figure 3.16 for details). (c) Band diagram during NBTI stress, along the transverse direction, for unstrained transistor shows the degeneracy of electron (Δ_2 and Δ_4) and hole valleys (LH and HH). (d) Band diagram for uniaxially strained device shows the strain induced splitting of electron valleys (Δ_2 and Δ_4) and hole valleys (LH and HH). (e) Band diagram for biaxially strained device shows an extra band discontinuity at the strained-Si/ $Si_{1-x}Ge_x$ interface [179]..... | 54 |
| 3.16: The proposed model for k_F assumes that inversion layer holes (concentration $p_h \sim E_c$) near the Si/dielectric interface tunnel into and is captured by Si-H bonds $\sim 1.5\text{\AA}$ [69] (Si-H bond length) away from the interface – leading to a hole-assisted, field-enhanced, thermal generation of interface defects. | 55 |
| 3.17: γ and γ_V variation with EOT . As predicted for field-dependent NBTI model, γ remains almost constant (dashed) and γ_V keeps around $1/EOT$ line (solid). For ΔV_T measurements, necessary delay corrections have been accounted for. | 56 |
| 3.18: Polarization properties of Si-H bond provides an estimation of $a = p\epsilon_{int}$. Here, the polarization vector is opposite in direction of the applied field E_{ox} under NBTI stress. | 57 |
| 3.19: Tunneling probability (T_H) at 1.5\AA (for clarity, the valence band-diagram in the inset of Fig. a is drawn upside-down) away from the MOS channel, calculated at different E_{ox} either using self-consistent Schrödinger-Poisson solution of MOS electrostatics [178] (in Fig. a) or using modified WKB approximation [183] (in Fig. b,c). Field dependence of T_H provides an estimation of γ_T for type-I transistors, having – (a) different %N within the dielectric, (a) variation of uniaxial compressive strain within the channel (a) variation of biaxial tensile strain within the channel. Theoretically extracted values of γ_T show negligible variation with any process technology. | 58 |
| 3.20: (a-d) Fitting $\Delta V_T \sim \Delta V_{IT}$ (measured by OTF- $I_{D,lin}$) vs. t (log-log scale) at different T_{STS} and V_{STS} by equation (3.2) gives A , γ , n as fitting parameters. Here $n = 0.15\sim 0.16$ and N_2 dose: 8%. Data is obtained from Dr. Anand Krishnan of Texas Instruments. (e) Obtained γ is plotted against $1/k_B T$, resulting $\gamma_T = 0.27\pm 0.05$ dec/MV/cm as intercept at $T_{STS}\rightarrow\infty$ and $a = 0.8 \pm 0.15$ q \AA as slope. (f) Determining E_A at different electric field enables estimation of a , hence γ_T using equation (3.3). | 61 |

| Figure | Page |
|--|------|
| 3.21: Voltage dependent $\Delta V_T \sim \Delta V_{IT} \sim \Delta N_{IT}$ stress data (taken at 125 °C) fitted using equation (3.2) enables one to estimate A, γ, n for a device ($E_{D1} \sim 0.9\text{eV}$ assumed, see section 3.8.2). Lifetimes (t_1, t_2, t_3, t_4) at different stress voltages (V_1, V_2, V_3, V_4) and safe operating voltage (V_{safe}) can also be determined, as discussed in Appendix A.4..... | 64 |
| 3.22: Variation of $Eexp(\gamma E)$ model parameters (A_{IT}, γ, γ_T) with $\%N$ ($a \sim 0.8 \text{ q}\text{\AA}$ used) for transistor in Table 3.4. The error margins in calculations are negligible. | 65 |
| 3.23: Variation of γ (where, ΔV_T is estimated either using classical method of equation (6.14), i.e., $ \Delta V_T \sim (V_G - V_{T0}) \Delta I_D / I_{D0} $ and also using Figure 6.18) for type-I transistors as a function of N_{int} (lines are guide to eye only). See Figure 6.23e for the definition of N_{int} | 67 |
| 3.24: (a-c) Comparison of ΔV_T and body-effect co-efficient degradation (Δm) indicates negligible hole trapping in the uniaxial devices under study. (d,e) Power-law time exponents of ΔV_T in on-the-fly $I_{D,lin}$ measurements (time-zero delay, $t_0 \sim 1\text{ms}$, see section 6.7.3 for a discussion on the importance of t_0 in NBTI measurements) are close to $\sim 1/6$, again indicating negligible hole trapping [119, 129, 138, 154] in all the uniaxial compressively strained transistors under study..... | 69 |
| 3.25: Uniaxial compressive strain causes: (a) negligible change in NBTI activation energy ($E_{A,IT} \sim 0.09 \text{ eV}$), thus indicating similar diffusing species, as well as negligible hole trapping [119, 129, 138, 154]; and (b) negligible variation for γ . Insets show the calculation procedure for $E_{A,IT}, \gamma$ | 70 |
| 3.26: (a) Fitting μ_{eff} vs. $(V_G - V_T)$ using well-known empirical mobility expression, we estimate μ_0 and θ at different stress intervals. (b) Estimated degradation in μ_0 ($\Delta\mu_0$) and θ ($\Delta\theta$) correlate well with ΔN_{IT} and can be used to calculate P_1, P_2 . (c) Variation of P_1 and P_2 are independent of uniaxial compressive strain, within the error margin of estimation, which is an independent signature of the invariance in wavefunction interaction (γ_f) with strain (see section 6.7.4.5 and [203])...... | 71 |
| 3.27: Though uniaxial compressive strain does not alter field acceleration parameter γ (Figure 3.25b), (a) it decreases ΔV_T , measured at similar electric field (or, $V_G - V_T$). This is shown here with three different strained devices, measured using on-the-fly $I_{D,lin}$ ($t_0 \sim 1\text{ms}$) at two different $V_G - V_T$. Decrease in ΔV_T with strain, for similar γ implies decrease in $N_0 P_T$ in equation (3.5). Similarly, (b) decrease in gate leakage (at same $V_G - V_T$) for similar devices signifies a reduction in P_T , which should be dominating over any increase in N_0 [95, 96] to give reduced $N_0 P_T$ with increase in compressive strain. | 73 |
| 3.28: Biaxial tensile strain increases in ΔV_T (with respect to that for unstrained devices, measured at similar electric field); line is guide to eye only. | 74 |
| 3.29: Comparison between exponential and power law fit of ΔV_T data for obtaining safe operating condition using stress data. Exponential fit underestimates and power law fit overestimates lifetime..... | 75 |

| Figure | Page |
|---|------|
| 3.30: Decrease in long-term NBTI exponent results from the reduction of stress field (for constant voltage stress) with time. Here, thick lines consider such reduction in electric field over time, whereas thin lines do not consider such “self-saturation” effect..... | 76 |
| 4.1: Markov model for probabilistic motion in an R-D system having H diffusion. The model is based on equations (4.4)-(4.6). The model considers reflecting boundary at the interface of poly/SiN cap (at $j = N$). Such relection leads to saturation of N_{IT} and is only expected to occur at very long stress time. | 79 |
| 4.2: Markov model for probabilistic motion in an R-D system having H-H ₂ diffusion. The model is based on equations (4.8)-(4.10). | 82 |
| 4.3: Schematic illustrating the idea that a log-normal distribution for $t_{STS}@N_{IT}$ guarantees a log-normal distribution for $N_{IT}@t_{STS}$ with contant σ for the distribution expressed in the form of equation (4.17)..... | 86 |
| 4.4: Time dependence and PDF of interface defect generation for (a) large transistor ($L = 100\text{nm}$, $W = 1\mu\text{m}$) and (b) small transistor ($L = 50\text{nm}$, $W = 100\text{nm}$) using $k_F = 3 \text{ sec}^{-1}$, $k_R = 2 \times 10^{-16} \text{ cm}^{-3}\text{sec}^{-1}$, and $D_H = 10^{-13} \text{ cm}^2\text{-sec}^{-1}$. For the large transistor, the PDF evolves from skewed normal (in the reaction phase) to a normal (in the diffusion phase) distribution. However, for the small transistor, the distribution stays skewed normal, even in the diffusion phase, which is commonly captured within the measurement window of NBTI measurement scheme (like MSM). | 88 |
| 4.5: Distribution of t_{STS} at different levels of N_{IT} obtained from 1000 MCMC simulation in the small transistor under study. The distribution is clearly log-normal, having cut-off at $t_{STS} = 1 \text{ msec}$ (sampling time) and $t_{STS} = 1 \text{ sec}$ (maximum length of MCMC simulation). The distribution also has constant standard deviation in the semilog-x plot, as expected from analytical calculation. .. | 89 |
| 4.6: (a) Though the mean (μ_{IT}) and standard deviation (σ_{IT}) increases with t_{STS} with a power-law time exponent (~ 0.25 for μ_{IT} and ~ 0.1 for σ_{IT}), skewness (γ_{IT}) reduces with t_{STS} . Such reduction in γ_{IT} indicates that the distribution is approaching towards a Gaussain one at longer t_{STS} . (b) The normal probability plot also confirms the existence of skewness at $t_{STS} = 1\text{sec}$ for Figure 4.4b. The PDF has a shape similar to the one reported in [207, 208, 210]. | 92 |
| 4.7: Both PDF and CDF of $N_{IT}@t_{STS} = 1\text{sec}$ can consistently be explained using the skewed normal distribution of equations (4.18)-(4.19). Other statistics (Gaussian, Poisson, and log-normal) are fitted with the simulated data using maximum likelihood estimation (MLE). SSE for different distribution functions are: 4.6×10^{-5} for skew-normal, 4×10^{-4} for Gaussian, 11.4×10^{-3} for Poisson, and 2×10^{-3} for log-normal..... | 92 |
| 4.8: (a,b) PDF and CDF for a large transistor in the diffusion regime showing no distinction between a log-normal and normal distribution. (c,d) Log-normal was expected for the same transistor in the reaction regime; however, in this reaction | |

| Figure | Page |
|--|------|
| regime, the distribution gets constraint by $N_{IT} > 0$ and gets skewed towards positive side of N_{IT} as in Figure 4.7. | 93 |
| 4.9: (a) PDF and (b) CDF of N_{IT} -induced ΔV_T for a small transistor having pre-existing interface defect. The distribution can explain the observation of negative ΔV_T tails in the experiments of [210]. | 95 |
| 4.10: (a) Time dependence and PDF of interface defect generation for a small transistor ($L = 50\text{nm}$, $W = 100\text{nm}$) using $k_F = 30 \text{ sec}^{-1}$, $k_R = 2 \times 10^{-15} \text{ cm}^{-3} \text{ sec}^{-1}$, $D_H = D_{H2} = 10^{-13} \text{ cm}^2 \text{ sec}^{-1}$, $k_H = 9 \times 10^{-15} \text{ cm}^{-3} \text{ sec}^{-1}$, and $k_{H2} = 600 \text{ sec}^{-1}$. Similar to the observation in Figure 4.4b, the distribution stays skewed normal, unless the magnitude of mean N_{IT} (μ_{IT}) is quite large. (b) Both mean and standard deviation increases with t_{STS} with a power-law time exponent (~ 0.17 for μ_{IT} and ~ 0.04 for σ_{IT}), however, positive skewness γ_{IT} reduced with t_{STS} | 96 |
| 4.11: Distribution of t_{STS} at different levels of N_{IT} obtained from 1000 MCMC simulation (for R-D system with H-H ₂ diffusion) in the small transistor under study. The distribution is clearly log-normal, having cut-off at $t_{STS} = 1\text{msec}$ (sampling time) and $t_{STS} = 1\text{sec}$ (maximum length of MCMC simulation). | 97 |
| 5.1: Three possible mechanisms for trapping/ detrapping of carriers into/ out of oxide defects considered in [230]. (a) elastic tunneling of carriers from the conduction band (<i>cb</i>) and valence band (<i>vb</i>), (ii) interfacial defect (<i>it</i>) assisted trapping/detrapping of carriers, and (iii) phonon-assisted (through lattice relaxation, <i>lr</i>) trapping/ detrapping of carriers. | 101 |
| 5.2: Simulated time evolution of ΔV_{IT} , ΔV_{HT} and total ΔV_T for a transistor, having thick oxide with uniform distribution of defects. Only elastic hole trapping is considered in calculation of ΔV_{HT} . As a result, hole-trapping controls total degradation at short stress time and saturates at t_{sat} . It also results in decrease in n at intermediate stress time. Inset shows tunneling probabilities T_1 (into/out of defects from/towards substrate) and T_2 (into/out of defects from/towards gate) for elastic hole trapping/ detrapping model considered in the text. | 103 |
| 5.3: (a) Typical nitrogen profile in Plasma and Thermal nitrided gate dielectric [238, 239]. Trapped hole profile for (b) type-I plasma and (b) type-II thermal nitrided transistors, having similar %N ₂ dose. Thermal nitrided oxides trap more holes compared to plasma nitrided oxides. Only elastic traps are considered in simulation. | 105 |
| 5.4: a) Experimental thickness dependence of average power law time exponents obtained from OTF- $I_{D,lin}$ in type-I plasma-nitrided transistor and type-II thermal-nitrided transistor (lines are guide to eye only). b) Simulated (involving elastic hole traps) thickness dependence under same condition. In Fig. b, it is assumed that N_{IT} -induced ΔV_{IT} in type-I transistor has $n \sim 0.145$, as observed in Fig. a. | 106 |
| 5.5: (a) EOT dependence of $E_{A,VT}$ obtained from OTF- $I_{D,lin}$ in type-I plasma-nitrided and type-II thermal-nitrided transistors (lines are guide to eye only). (b) Simulated EOT dependence of $E_{A,VT}$ for similar cases. In Fig. b, it is assumed | |

| Figure | Page |
|--|------|
| that N_{IT} -induced ΔV_{IT} in type-I transistor has $E_A \sim 0.095$, as observed in experiments..... | 106 |
| 5.6: Carrier fluxes considered in the hole trapping model of equation (5.3). Related additional variables, over the one used in equation (5.1), are also defined. | 109 |
| 5.7: (a) Timing diagram used in the hole trapping simulation using the framework of section 5.5.1. (b) Occupancy of defects or, hole trapping sites (assumed to have existence within the colored region only), after the transistor is kept at $V_G = -2.1V$ for 1000 sec. (c) Occupancy of hole trapping sites after the transistor is switched to $V_G = -1.0V$ at 1000 sec and kept at that voltage up to 2000 sec..... | 110 |
| 5.8: (a) Voltage and (b) Temperature dependent simulation of hole trapping. Although, hole trapping shows significant voltage or electric field dependency, it is independent of temperature. | 111 |
| 6.1: Basic principles for obtaining resonance in an ESR/EPR experiment. Device under study is placed in a slowly varying magnetic field (B), along with an electromagnetic signal (having frequency, ν) applied perpendicular to B . At resonance, there is significant absorption of electromagnetic signal. From the peak position of the absorption line (zero-crossing point of the first derivative) one estimates g , thus the nature of defect. This figure is taken from http://images.google.com | 116 |
| 6.2: (a) Schematic of PMOS connected in gated-diode configuration for EDMR or SDR measurement. Here, a forward bias (V_F) is applied to shorted source/drain and a change in substrate current (I_{sub}) at resonance (i.e. at $h\nu = g\mu_B B$) is monitored. (b,c) Band diagrams are drawn respectively along gate/oxide/substrate and drain/substrate/substrate-contact directions. | 118 |
| 6.3: (a) Timing diagram for MSM experiment on a PMOS transistor for characterizing parametric degradation during NBTI stress. After taking initial $I_{D,lin}-V_G$ at $t_{STS,0}$, stress voltage (V_{STS}) is applied. The stress is later interrupted at $t_{STS,1}$, $t_{STS,2}$, ..., etc. time intervals for taking subsequent $I_{D,lin}-V_G$ sweeps. (b) Under NBTI stress, $I_{D,lin}-V_G$ shifts in negative direction, which indicates an increase in $ V_T $, a decrease in $I_{D,lin}$, and also change in other parameters e.g., μ_{eff} , SS , m | 119 |
| 6.4: Uniform generation of donor-type N_{IT} causes left-shift of C_G-V_G . Pre-stress C_G-V_G is calibrated with experimental data. (b) Relationship between ψ_S and V_G before and after stress. Interface defect causes ψ_S to decrease at constant V_G (movement along vertical line). Increase in V_G during NBTI degradation (moving along horizontal line) can keep ψ_S constant (needed in determining Δm) with degradation. In this figure, we use device #3 of Table 6.1 for illustrating the impact of N_{IT} generation. | 121 |
| 6.5: Modeling μ_{eff} vs. (V_G-V_T) using empirical mobility model of equation (6.5) with $\eta = 1$ before ($N_{IT} = 0$) and after ($N_{IT} > 0$) NBTI degradation. Here, device #5 of Table 6.1 is used. | 123 |

| Figure | Page |
|---|------|
| 6.6: Variation of mobility parameters ($\Delta\mu_0$ and $\Delta\theta$), obtained using equation (6.5) with $\eta = 1$ (Figure 6.5), and Δm have good correlation with ΔV_T . Here, device #5 of Table 6.1 is used..... | 123 |
| 6.7: μ_{eff} (a) increases for device #5 of Table 6.1 (having $\theta_0 = 1.34V^{-1}$) or (b) decreases for device #2 of Table 6.1 (having $\theta_0 = 0.8V^{-1}$) during NBTI when measured at constant gate voltage (V_{meas}), although it shows (c) a decreasing trend when measured at constant E_{eff} | 124 |
| 6.8: Modeling μ_{eff} vs. E_{eff} using equations (6.6)-(6.7) in both pre- and post- stress conditions for device #5 of Table 6.1. This modeling approach is used for identifying the physical basis of $\Delta\mu_{eff}$ (ΔN_{IT})..... | 126 |
| 6.9: (a) Timing diagram for UFV experiment applied to a PMOS transistor for estimating ΔV_T during NBTI stress [140]. (b) In UFV, one chooses a particular current level, $I_{D,lin(UFV)}$ and then estimate V_{Gi} required for having $I_{D,lin(UFV)}$ by interrupting V_{STS} at different stress interval. Generation of interface/bulk trap during NBTI increases $ V_{Gi} $ within time. Later, $\Delta V_T(i)$ is assumed to be equal to $\Delta V_G(i) = V_{Gi} - V_{G0}$. Recovery voltage (V_{REC}) in UFV can have any magnitude and polarity. | 127 |
| 6.10: (a) Timing diagram for OTF- $I_{D,lin}$ measurement. (b) Hypothetical $I_{D,lin}$ - V_G sweeps are shown at different time steps during OTF- $I_{D,lin}$. In OTF- $I_{D,lin}$, only the initial short (V_G)-range $I_{D,lin}$ - V_G is measured for estimating V_{T0} . Later, V_{STS} is applied and $I_{D,lin0}$ is measured after t_0 . Finally, ΔV_T is indirectly estimated using $\Delta I_{D,lin}$. Moreover, $ V_{REC} $ is always above threshold and $ V_{meas} $ is same as $ V_{STS} $ or $ V_{REC} $ | 128 |
| 6.11: (a) During a $I_{D,lin}$ - V_G measurement, extremely popular in μ_{eff} calculation, presence of high I_G (in MOSFETs with ultra-thin gate dielectric) will assist I_S and will oppose $I_{D,lin}$. (b) Different current components measured during the $I_{D,lin}$ - V_G measurement in such a p-channel transistor. Excessive I_G reduces $I_{D,lin}$ (at high V_G), thus corrupts the extracted μ_{eff} at higher E_{eff} (inset). Here, I_{D0} represents the corrected drain current, using the algorithm proposed in [273]. | 130 |
| 6.12: (a) $ \Delta V_G \neq \Delta V_T $ for UFV, as assumed in [140]; at sub-threshold, $ \Delta V_G < \Delta V_T $ and at slightly above threshold, $ \Delta V_G > \Delta V_T $. (b) Appropriate Δm correction using equation (6.10) and Figure 6.6 enables us to have a refined estimation of ΔV_T (using ΔV_G), which becomes more consistent with the one obtained by drawing slopes at $g_{m,max}$ (section 6.4.2). | 132 |
| 6.13: NBTI relaxation measured using OTF- $I_{D,lin}$ at $V_{REC} < V_{T0}$ predicts measured ΔV_T recovery to start later compared to the one measured using UFV at $V_{REC} = 0$ | 133 |
| 6.14: Energy band diagram for a PMOS transistor biased in (a) NBTI stress ($V_G = V_{STS}$, substrate inverted), (b) NBTI measurement ($V_G = V_{meas} \sim V_T$) and (c) NBTI recovery ($V_G = V_{REC} = 0$, substrate moderately depleted) condition for the UFV scheme. E_{FS} is substrate Fermi level and $E_{C(bulk)}$, $E_{V(bulk)}$ are conduction and | |

| Figure | Page |
|--|------|
| valence band level in bulk of the substrate. Arrow indicates valence band electron trapping during recovery and measurement conditions. (d) Movement of Quasi-Fermi level as the transistor is switched from stress to recovery, then to the measurement mode. Energetically suppressed substrate electron capturing makes the trap filling a $\log-t$ process from $t_{REC} = \sim\mu\text{s}$ to $\sim\text{ms}$ | 135 |
| 6.15: Simulation of fractional recovery ($t_{STS} = 10^5$ sec) in UFV measurement scheme ($t_{meas} \sim \mu\text{s}$) on a PMOS transistor having parameters similar to the one used in [140]. Our simulation indicates $E_{FS,rec}$ (@ $V_{REC} = 0\text{V}$) to be $\sim 0.45\text{eV}$ above the valence band edge at the interface. In (a), $E_{FS,meas}$ is considered to be same as $E_{FS,rec}$ (<i>i.e.</i> , $V_{meas} = V_{REC}$); whereas in (b) $E_{FS,meas}$ (@ $V_{meas} = V_{T0} \sim -0.4\text{V}$) \sim is considered to be 0.2eV above valence band edge at the interface. Figure (b) also indicates actual ΔV_T recovery (blue line), measured ΔV_T recovery (red line) and actual ΔN_{IT} recovery (black line: predicted from classical H-H ₂ R-D model). | 139 |
| 6.16: Consideration of valence band electron trapping in UFV makes start of recovery consistent with OTF-I _{D,lin} measurement. | 140 |
| 6.17: (a) Effect of t_0 -delay on OTF-I _{D,lin} compared with UFV, where OTF-I _{D,lin} is performed on transistors from different technologies. Here, ΔV_T is scaled along Y-axis for clarity. Next, in (b) the possibility of technology to technology variation is avoided by studying the effect of t_0 on similar transistors. In both cases, measurement with lowest t_0 is fitted using classical H-H ₂ R-D model (section 3.5) and other measurements require no additional fitting parameter, except the consideration of t_0 . These measurements indicate that for higher t_0 , there is an increase in n at a particular ($t_{STS} - t_0$) mainly in the short-term stress.... | 143 |
| 6.18: Algorithm for ΔV_T estimation in OTF-I _{D,lin} from $\Delta I_{D,lin}/I_{D,lin0}$ | 145 |
| 6.19: (a) Estimation of $\Delta\mu_{eff}$ and Δm enables one to evaluate ΔV_T using Figure 6.18; which is in good agreement (closer to solid line representing 1:1 relation) with ΔV_T obtained from $g_{m,max}$. (b) (c) Classical estimate of $ \Delta V_T \sim \Delta I_D/I_{D0}(V_G - V_{T0}) $ neglects $\Delta\mu_{eff}$ and Δm , thus under/over-estimates $ \Delta V_T $. (d) Neglecting contribution from m does not affect ΔV_T estimation significantly. Information for the devices used in this figure can be found in Table 6.1..... | 147 |
| 6.20: (a) Estimated ΔV_T using algorithm in Figure 6.18 provides similar ΔV_T at start of recovery for the same V_{STS} , T_{STS} , but different V_{REC} ; whereas (b) equation (6.14) incorrectly estimates different ΔV_T at start of recovery. Information for the device used in this figure can be found in Table 6.1. | 148 |
| 6.21: Although power-law time exponent remains unaffected, when $\Delta\mu_{eff}$ and Δm are neglected in ΔV_T estimation (for OTF-I _{D,lin}) compared to the case, where $\Delta\mu_{eff}$ and Δm are considered, lifetime is (a) over-estimated or (c) under-estimated. As a result, voltage acceleration (b), (d) is always under-estimated and V_{safe} is over- or under-estimated in classical estimation. Information for the devices used in this figure can be found in Table 6.1. | 150 |

| Figure | Page |
|---|------|
| 6.22: (a-c) Correlations for $\Delta\mu_0$, $\Delta\theta$ and Δm with ΔV_T are independent of T_{STS} . Hence, no change in NBTI activation energy (E_A) is observed before and after considering $\Delta\mu_0$, $\Delta\theta$ and Δm in ΔV_T estimation for OTF- $I_{D,lin}$ (d). | 151 |
| 6.23: EOT and N_D independent parameters, $P_1 = d(\Delta\mu_0/\mu_{0(0)})/d(\Delta N_{IT})$ and $P_2 = d(\Delta\theta/\theta_0)/d(\Delta N_{IT})$, obtained using equation (6.5) and Figure 6.6 show better correlation with (c, d) N_{int} , compared to the one with (a, b) $\%N$. Lines in (c, d) are guide to eye only. (e) Nitrogen profile in a type-I transistor, having plasma nitrided dielectric, shows nitrogen concentration $[N]$ to drop exponentially from poly/oxide interface towards oxide/substrate interface [238]. This information helps us to estimate N_{int} from $\%N$ and T_{PHY} using equation (6.18). | 153 |
| 6.24: Variation of β_{IT} for type-I transistors (lines are guide to eye only) indicates a decrease, similar to the decrease of P_1 and P_2 , observed in Figure 6.23c,d..... | 154 |
| 6.25: (a) $t_{REC,start}$ in UFV measurement [140], commences ~ 5 decades earlier in time compared to the prediction of H-H ₂ R-D theory. Our OTF- $I_{D,lin}$ (with $t_0 = 1\mu s$) measurement also shows similar relaxation trend. The slight difference between UFV and OTF- $I_{D,lin}$ can be attributed to the valence band electron trapping phenomena, explained in section 6.7.2. (b) Resolution between the measured ultra-fast ΔV_T relaxation and comparatively slower N_{IT} relaxation becomes possible, after decomposing ΔV_T (solid line) into ΔV_{IT} (dotted line), ΔV_{HT} (dashed line) components; each having separate time-dynamics. | 156 |
| 6.26: (a-c) Time evolution of degradation under NBTI stress at identical E_{ox} , but different T_{STS} , for plasma nitrided transistors having similar T_{PHY} ($\sim 24 \text{ \AA}$) but different $\%N$ (hence, different N_{int}) and EOT . (d) Power-law NBTI time exponent (n) decreases with increase in $\%N$, indicating the presence of ΔV_{HT} for higher $\%N$. In addition, n shows an increment with increase in t_0 -delay (see section 6.7.3). (d) Similar temperature independence was also reported for lightly dosed nitrided transistors, using UFV [140]. Moreover, the effect of un-optimized PNA (leading to significant ΔV_{HT} , even at low $\%N$ [138]) is also evident in Fig. e. | 161 |
| 6.27: (a) Though $t_{REC,start}$ is considered to be $\sim \mu s$ in several reports [168, 278, 293], our OTF- $I_{D,lin}$ measurements on nitrided transistors, at different V_{REC} , demonstrate that $t_{REC,start}$ depends on $\%N$ and $(V_{STS} - V_{REC})$. (b) Similar to NBTI stress phase (Figure 6.26), early relaxation phase (up to $\sim ms$) is approximately temperature independent..... | 162 |
| 6.28: (a) NBTI relaxation at different V_{REC} show significantly different $t_{REC,start}$. Since relaxation for $V_{REC} \leq -1.8V$ is very close the R-D theory, there is an additional relaxation mechanism for $V_{REC} > -1.8V$. Our observation suggests hole detrapping to be the additional mechanism. (b) Schematic (based on simulation within a Shockley-Read-Hall trapping-detrapping framework [119]) for expected hole trapping sites (hatched region) at V_{STS} . (c) When gate bias is | |

| Figure | Page |
|--|------|
| switched from V_{STS} to V_{REC} , the hatched region will detrapp the captured hole in a temperature independent manner..... | 162 |
| 6.29: (a) Measured ΔV_T and extracted ΔV_{IT} and ΔV_{HT} components for device #1 in Table 6.2. ΔV_{IT} and ΔV_{HT} are extracted according to the decomposition method proposed in 6.7.5.2. Here, the effect of mobility [165] and electric field-reduction [119] is taken into account in estimating ΔV_T . (b) Temperature dependence and corresponding activation energies (E_A) of ΔV_T , ΔV_{IT} and ΔV_{HT} components for the same transistor. (c) Estimated $\Delta V_{HT}/\Delta V_T$ and $\Delta V_{IT}/\Delta V_T$ at different temperature indicate an increase (decrease) in $\Delta V_{IT}/\Delta V_T$ ($\Delta V_{HT}/\Delta V_T$) at higher temperature, which is primarily due to the temperature independence of the ΔV_{HT} component. | 164 |
| 6.30: (a) Activation energy for extracted ΔV_{IT} and ΔV_{HT} components indicates negligible $\%N$ dependence. (b) $\%N$ Dependence of measured ΔV_T and extracted ΔV_{IT} and ΔV_{HT} components for a particular stress time, temperature and electric field. Lines are guide to the eye only..... | 165 |
| 6.31: Time evolution of NBTI relaxation at different temperature for both type-I (device #3 in Table 6.2) and type-II (device #5 in Table 6.2) transistors. Initial relaxation, up to $t_{REC} \sim 10$ ms in type-I and ~ 1 s in type-II transistor, is temperature independent, thus indicating a hole detrapping process. | 166 |
| 6.32: (a) Considering total hole detrapping at $V_{REC} = -1.3$ V in the early phase of NBTI relaxation brings $t_{NIT,start}$ at all V_{REC} closer to the relaxation, predicted by R-D theory. The error bar along the x-axis for $V_{REC} = -1.3$ V is due to the $I_{D,lin0}$ error in calculating $\Delta V_{IT}/\Delta V_T$ (see Figure 6.29). (b) Furthermore, if one assumes negligible amount of oxide defect generation during NBTI stress (with $\Delta V_{OT} \sim 3$ mV at $t_{STS} = 10^3$ sec, having the time dependence similar to Figure 6.33c), then consideration of hole detrapping from these generated oxide defect brings $t_{NIT,start}$ more closer to the predictions of R-D theory..... | 167 |
| 6.33: (a) Time exponent of NBTI degradation (fitted solid lines) increases at $V_{STS} \geq V_{G4}$ [175], thus indicating clear presence of N_{OT} (or ΔV_{OT}) at these V_{STS} . By assuming negligible ΔV_{OT} at $V_{STS} \leq V_{G3}$, we can estimate ΔV_{IT} at $V_{STS} \geq V_{G4}$ (dashed lines) and hence calculate $\Delta V_{OT} (= \Delta V_T - \Delta V_{IT})$ at $V_{STS} \geq V_{G4}$. (b) Similar calculation of ΔV_{OT} for the measurements in [175]. (c) Estimated ΔV_{OT} at different V_{STS} scales universally (similar to the observation in [114]) and can be fitted using equation (6.21)..... | 168 |
| 6.34: (a) Since hole trapping and detrapping happens at similar time-scale, we expect total hole detrapping, at the end of OFF-state during AC NBTI stress. Therefore, (b) AC/DC ratio (with ΔV_T for AC NBTI measured at the end of AC cycles) will show consistency with R-D theory only for transistor having lower ΔV_{HT} or lower $\%N$ | 170 |
| 6.35: (a) AC/DC ratio (when ΔV_T is measured at the end of AC cycles) vs. duty cycle plot for different nitrided transistors. The experiments show significant consistency with the prediction of R-D theory (solid line) for low $\%N$. (b) | |

| Figure | Page |
|---|------|
| Though AC/DC ratio for any %N is always frequency independent, there is significant %N dependency due to the presence of ΔV_{HT} in DC NBTI stress. | 171 |
| 7.1: Experimental (a) C_G - V_G and (b) J_G - V_G curves (in inversion region) for a few SiON devices fitted using QM simulation [178]. | 179 |
| 7.2: Variation in $ \Delta V_{FB} $ for PMOS SiON samples, neglecting the boron penetration effect. | 180 |
| 7.3: Quadratic variation of (a) ϵ_{di} , m_{di} , (b) $E_{g,di}$, ϕ_{be} with %N (0% for SiO ₂ , 57.1% for Si ₃ N ₄) for samples having $EOT \sim 1.3$ nm. Here, ϵ_0 is free space dielectric constant and m_0 is mass of electron. The error bar shows the variation in model parameters expected due to $\pm 0.5 \text{ \AA}$ error in measurement of T_{PHY} by XPS. Dashed line in (a) plots the variation of ϵ_{di} vs. %N, based on equation (7.3). | 182 |
| 7.4: $J_G @ V_G$ vs. %N for NMOS devices at different EOT (Symbols: QM simulation, lines: analytical approach). The error bar @ $EOT = 1.2$ nm indicates J_G variation expected due to $\pm 0.5 \text{ \AA}$ error in T_{PHY} , which is negligible up to 30% of N ₂ dose. .. | 184 |
| 7.5: Flowchart for electric fields calculation (from $ V_G $) in PMOS inversion. Reverse calculation (electric field to $ V_G $) is also possible. Similar calculation can also be done for NMOS devices. The notations used here are described in the text. | 187 |
| 7.6: Electric field vs. gate voltage relation obtained using analytical approach of Figure 7.5 (symbols) are in agreement with that obtained from QM simulation (lines). | 188 |
| 7.7: V_{safe} vs. %N at different EOT ($T = 125 \text{ }^\circ\text{C}$; Failure criteria: $\Delta V_T = 60$ mV, $t_{life} = 5$ years for CMOS DC operation). Dotted line for $EOT = 1.6$ nm indicates the improvement in AC condition and with $T_{stress} = 85 \text{ }^\circ\text{C}$ | 189 |
| 7.8: Design diagram for SiON devices, obtained using Figure 7.4 ($J_G @ V_G = 0.8, 1.0, 1.2$ V are considered) and Figure 7.7. Solid black lines use J_G calculated by QM simulation and dotted red lines use J_G calculated from equations (7.4)-(7.8). Note the universality (<i>i.e.</i> , curves at different V_G fall on top of each other) of the figure when J_G calculated by QM simulation. Operating point A ($V_{safe} = 1.1$ V and $J_G = 10 \text{ A/cm}^2$) suggests, for example, the use of $EOT \sim 1.4$ nm and $\sim 20\%$ N ₂ dose. | 191 |
| 8.1: Variation of Δm for Type-I (circles) and Type-II (squares and triangles) devices. Type-II transistors have negligible Δm at small ΔV_T , thus showing signatures of hole trapping. | 197 |
| 8.2: Variation of body-effect coefficient (m) or SS is negligible and within the error margin of SS estimation during PBTI stress. Here, the measurement is performed on a number of NMOS transistors, and each of them indicates that ΔSS is (on average) independent of the level of PBTI-induced N_{ET} | 197 |
| 8.3: (a) A schematic representation of the effect of N_{IT} generation on μ_{eff} - E_{eff} plot. The functional dependencies of different mobility components are provided in section 6.4.4. N_{IT} induces extra coulomb scattering and thus reduces μ_{eff} , | |

| Figure | Page |
|---|------|
| <p>especially at low E_{eff}. (b) μ_{eff} vs. E_{eff} in both pre- and post- N_{IT} generation conditions for different steepness of μ_{eff} vs. E_{eff} characteristics. Although μ_{eff} decreases at constant operating E_{eff} (A \rightarrow B); at constant V_G, E_{eff} decreases with degradation (B \rightarrow C) and thus gives rise to increase/decrease in μ_{eff} (A \rightarrow C), depending on the steepness (θ) of the μ_{eff} vs. E_{eff} plot. Physics of E_{eff} reduction: (c-d) At constant V_G, there should be equal amount of gate and substrate charges in order to maintain the electrostatics. As a result, effect of N_{IT} generation, when measured at constant V_G, should consider a corresponding reduction in substrate charge (c\rightarrowd). Since in an inverted MOS, Q_{dep} remains invariant of N_{IT} generation, $E_{eff} \sim Q_{dep} + 1/3Q_{inv}$ reduces with N_{IT}.</p> | 200 |
| 8.4: (a) μ_{eff} vs. E_{eff} for the uniaxially strained PMOS devices under study, having channel length of $L = 0.13, 0.26$ and $1.00 \mu\text{m}$. Short channel devices have more strain and steeper curve compared to long channel devices. $\Delta\mu_{eff}$ for (b) $L=1.00\mu\text{m}$ transistor is negative, whereas for (d) $L=0.13\mu\text{m}$ transistor it is always positive. For $L=0.26 \mu\text{m}$ transistor in (c), $\Delta\mu_{eff}$ is initially negative at low $ V_G $ and becomes positive at high $ V_G $. Variation of $\Delta\mu_{eff}/\mu_{eff0}$ from negative to positive is evident as L is reduced, or strain is increased. (d) All these transistors show degradation in $\Delta\mu_{eff}/\mu_{eff0}$, when measured at constant E_{eff} (here shown for $ \Delta V_T \sim 37\text{mV}$). | 201 |
| 8.5: (a) Electron trapping into oxide defects during PBTI stress causes negligible change in $\Delta g_{m,max}$. Whereas, (b) N_{IT} generation decreases $\Delta g_{m,max}$ significantly irrespective of the amount of strain, which is similar to the decrease in $\Delta\mu_{eff}@E_{eff}$ shown in Figure 8.4e. This indicates that although there is decrease in $\Delta\mu_{eff}@E_{eff}$ due to N_{IT} generation, electron trapping has negligible impact on $\Delta\mu_{eff}@E_{eff}$ | 202 |
| 8.6: (a) $\Delta\mu_{eff}$ at different (V_G-V_{T0}) for the highly strained PMOS transistor in Figure 8.4. (b) Comparison of θ between the high- κ NMOS transistor under study and the highly strained PMOS transistor of Figure 8.4. (b) Although NMOS transistor has smaller θ , presence of negligible $\Delta\mu_{eff}@E_{eff}$ and E_{eff} reduction at constant V_G stress provide positive $\Delta\mu_{eff}@V_G$ due to N_{ET} | 204 |
| 8.7: $\Delta I_{D,lin}$ for (a) $1.00\mu\text{m}$, (b) $0.26\mu\text{m}$, and (c) $0.13\mu\text{m}$ transistors, measured at various V_G . Note that, although $0.13\mu\text{m}$ transistor is stressed more compared to the other transistors, it has the lowest $\Delta I_{D,lin}$ at $ V_G > 1.1\text{V}$ | 205 |
| 8.8: $\Delta I_{D,sat}$ for (a) $1.00\mu\text{m}$ and (b) $0.13\mu\text{m}$ transistors measured at various V_G . Similar to the observation involving $\Delta I_{D,lin}$ in Figure 8.7, $L = 0.13\mu\text{m}$ transistor has negligible $\Delta I_{D,sat}$ compared to $L = 1.00\mu\text{m}$ transistor. | 206 |
| 8.9: (a) $\Delta I_{D,lin}$ and (b) $\Delta I_{D,sat}$ at different V_G-V_{T0} for the high- κ NMOS transistor, having $\Delta\mu_{eff}@V_G$ values of Figure 8.6c. | 207 |
| 9.1: (a) Linear transfer characteristics of a classical MOS transistor, where any sources of ΔV_T lead to $\Delta I_{D,lin}$. The effect of ΔV_T can be minimized by using guard-band voltage ($V_{DD}-V_1$). (b) Comparison of mobility-field ($\mu_{eff} - E_{eff}$) | |

| Figure | Page |
|--|------|
| relationship in classical and variation-resilient transistors. An increase in V_T always results in an increase in μ_{eff} , except for N_{IT} -related ΔV_T . (c) Transfer characteristics of the proposed variation-resilient MOS transistor, where ΔV_T is not reflected in $\Delta I_{D,lin}$. Thus variation-resilient transistor can operate with reduced guard-band. | 214 |
| 9.2: Measured μ_{eff} - E_{eff} at different levels of strain (ε) are fitted using equations (9.2)-(9.5). Here, we estimate ε by using Fig. 11 of [193]. The measurement at 0% strain is consistent with the universal mobility measurement (dashed line) of [267]..... | 219 |
| 9.3: (a) Calculation of $\Delta_{H,\varepsilon}$ within single sub-band, Airy-function approximation shows an increase in $\Delta_{H,\varepsilon}$ with E_{eff} , as also observed in [187]. (b) $\Delta E_{LH-HH} = \Delta_{H,\varepsilon} - \Delta_{H,\varepsilon=0}$ at different E_{eff} is close to its magnitude of 6.426eV at zero E_{eff} , where 6.426eV is the deformation potential..... | 221 |
| 9.4: (a) μ_{eff} - E_{eff} curves for different levels of strain, estimated using strain-dependent mobility model of equations (9.2)-(9.5). (b) Corresponding $I_{D,lin}$ - V_G characteristics at $V_{DS} = 50\text{mV}$ for a simulated transistor having $L = 100\mu\text{m}$, $EOT = 1.4\text{nm}$, and $N_{sub} = 3 \times 10^{17} \text{ cm}^{-3}$ | 222 |
| 9.5: (a) Reduction in temperature increases the $\mu_{eff} - E_{eff}$ steepness for the highly-strained PMOS transistor of Figure 8.7c. (b) The transistor, operated at reduced temperature, is less sensitive to N_{IT} -induced $\Delta I_{D,lin}$ at $ V_G - V_{T0} = 0.8\text{V}$ | 223 |
| 9.6: (a) Comparison of $\mu_{eff} - E_{eff}$ steepness for unstrained ($\varepsilon = 0\%$) and strained ($\varepsilon = 2.5\%$) transistors. (b, d) Unstrained transistor having less $\mu_{eff} - E_{eff}$ steepness suffers from EOT and Φ_{GS} variations. Whereas, (c, e) strained transistor with higher $\mu_{eff} - E_{eff}$ steepness or small $g_{m,ON}$ is less sensitive to EOT and Φ_{GS} variation at operating V_G | 227 |
| 9.7: (a) Unstrained transistor suffers from $\pm 20\%$ N_{sub} variation. (b) Though $g_{m,ON} \sim 0$ at $V_{GS} \sim 1.2\text{-}1.5\text{V}$ for 2.5% strained transistor, dopant fluctuation remains uncompensated around that range of V_{GS} | 228 |
| 9.8: (a) Trapping/detrapping of channel carriers with $\sim k_B T$ energy range around the Fermi-level gives rise to I_D noise having distinct frequency. (b) Noise from defects at different distances away from the interface leads to $1/f^n$ signal. | 230 |
| 9.9: (a) Configuration of the inverter used in section 9.8.1. In stress phase ($V_{IN} = 0$), there is NBTI degradation in PMOS transistor. (b) The impact of NBTI degradation (<i>i.e.</i> , change in rise τ_R and fall delay τ_F) is monitored by applying a test pulse as V_{IN} and calculating V_{OUT} using equation (9.12). (c) Self-compensation reduces the rise delay degradation by 50% – thus can increase NBTI lifetime by ~ 64 times (assuming time exponent for delay degradation, $n \sim 1/6$, as obtained in [64])..... | 233 |
| 9.10: (a) Schematic of a 6T-SRAM Cell. PL and PR are PMOS transistors within the storage framework; whereas NL and NR are corresponding NMOS transistors. AXL and AXR are NMOS transistors used for accessing (reading or writing) | |

| Figure | Page |
|--|------|
| stored data. (b) Static Noise Margin (SNM) calculation using voltage-transfer characteristics of back-to-back inverters (blue lines: before N_{IT} generation; red lines: after N_{IT} generation). (c) Schematic of voltage transfer characteristics for PR-NR inverter. (d) Schematic of I_D - V_D characteristics for the NMOS/PMOS transistors of PR-NR inverter. Similar modes of operation for NMOS/PMOS transistors are also marked in Fig. c,d. Fig. a is re-drawn from [379] and Fig. c-d is taken from the Purdue's ECE 612 lecture (available at [383]). | 235 |
| 9.11: Simulated I_D - V_G characteristics at $ V_{DS} = 0.1V$ for (a) unstrained and (b) 2.5% uniaxial strained transistor. Effect of N_{IT} -induced ΔV_T in NAND gate and 5-stage RO circuits for (c) unstrained and (d) 2.5% uniaxial strained transistor. The difference between classical SPICE simulation based on equations (9.13)-(9.14) (solid symbols) and circuit simulation based on equations (9.2)-(9.5) (open symbols) is clearly evident in strained CMOS transistor. | 238 |
| 10.1: Schematic representation of the Si-SiO ₂ interface, seen from the top of a MOS transistor. Most of the dangling Si- bonds (total density $\sim 10^{14} \text{ cm}^{-2}$) are terminated by O atoms during oxidation of Si substrate. (a) Remaining Si- bonds (density $\sim 10^{12} \text{ cm}^{-2}$) are annealed using hydrogen. These Si-H bonds are spatially distributed at the interface. (b) After most of the Si-H bonds are dissociated due to the application of NBTI stress, the resultant dangling Si-bonds can be sparsely distributed at the Si/SiO ₂ interface. | 246 |
| 10.2: Statistical R-D simulation considers distribution of interface defects at the Si/SiO ₂ interface. This increases the dimensionality of hydrogen diffusion at smaller time. As such, time exponent of N_{IT} generation at smaller time will be considerably higher than the one obtained from continuum simulation having one-dimensional hydrogen diffusion (Figure 3.1). | 247 |
| 10.3: (a) During Hot-Carrier-Injection (HCI) stress, Si-H bonds are broken only near the drain end by energetic carriers. Resultant hydrogen species have a 2D diffusion front. (b) R-D simulation and (c) experiment both indicates negligible amount of HCI relaxation or annealing of Si- bonds due to difficulty in the back-diffusion of hydrogen species. Taken from [391]. | 248 |
| 10.4: Flowchart for designing transistors for variation-resilient circuit operation. | 249 |

Appendix Figure

| | |
|---|-----|
| B.1: Schematic representation of the N_{IT} generation model with (a) classical hydrogen transport having single energy level and (b) dispersive hydrogen transport through exponentially distributed multiple energy levels, $g(E) \sim \exp(-E/E_0)$. Adapted from [33]. | 291 |
| B.2: Simulation results of R-D framework (equations (B.3)-(B.6)) with dispersive hydrogen transport. Simulation parameters: $\nu_0 = 10^{12} \text{ Hz}$, $E_0 = 47\text{meV}$, $N_T = 5 \times 10^{21} \text{ cm}^{-3}$, $N_C = 10^{19} \text{ cm}^{-3}$, and $E_{min} = 0.47\text{eV}$. (a) Hydrogen distribution near | |

| Appendix Figure | Page |
|--|------|
| the oxide/substrate interface at $t = 0.1\mu\text{s}$. (b) Evolution of average trap energy level E_{avg} at different temperature. | 293 |
| B.3: Dispersive H^+ drift/diffusion transport simulation, implemented using Scharfetter-Gummel discretization, within the R-D framework agrees well with experimental data from [33], albeit using unphysical parameters like: $\nu_0 = 6.33 \times 10^{-4}$ Hz, $E_0 = 105\text{meV}$, and $N_C/N_T = 5 \times 10^2$ | 294 |
| B.4: Energy-well representation of N_{IT} generation (taken from [115]). Here, hydrogen is released from the Si-H bond (equilibrium state) thermally over a barrier, or tunnel through the barrier at $T < 1\text{K}$, into transport states. Dissociation of Si-H leaves dangling Si- bonds at the oxide/substrate interface. | 297 |
| B.5: Energy-well representation of stretched exponential relaxation in $a\text{-Si:H}$ samples through dissociation of trapped hydrogen from T states (taken from [422]). Here, R refers to the ground state or reservoir, and I refers to the interstitial hydrogen states for diffusion through the crystal. | 298 |
| B.6: Analogy between equations (B.16) and (3.2), with $X \equiv H$ | 299 |
| B.7: Schematic for McKelvey's flux method for carrier transport. Here, f and b denote forward and backward fluxes, and α and α' represent characteristic decay length. | 300 |
| B.8: Analogy between (a) energy-well representation with R and I states having same energy and (b) R-D system, having H diffusion. Analogous parameters are: $N_T \equiv N_0/\delta$; $[T] \equiv (N_0 - N_{IT})/\delta$; $[I] \equiv N_{H(0)}$; $N_T - [T] \equiv N_{IT}/\delta$; $k_T \equiv k_F$; $\sigma_T \equiv k_R$; and $D_H \equiv \nu_T/2\alpha = \delta^2 \sigma_H N_I$ | 301 |
| B.9: When the transport species is charged hydrogen (H^+), then application oxide electric field under NBTI stress makes the transport from I to R (towards gate) easier compared to the transport from R to I (towards substrate)..... | 302 |
| B.10: Analogy between (a) energy-well representation with R and I states having same energy and (b) R-D system, having H^+ drift/diffusion. Analogous parameters are: $N_T \equiv N_0/\delta$; $[T] \equiv (N_0 - N_{IT})/\delta$; $[I] \equiv N_{H(0)}$; $N_T - [T] \equiv N_{IT}/\delta$; $k_T \equiv k_F$; $\sigma_T \equiv k_R$; $D_{H^+} \equiv \nu_T/2\alpha = \delta^2 \sigma_H N_I$; and $\mu_{H^+} E_{ox} \equiv 2D_{H^+} \Delta/\delta k_B T$ | 304 |
| B.11: Analogy between double-well model and B-D model with H^+ generation. | 305 |
| B.12: (a) Triple-well representation for explaining N_{IT} formation has unsubstantiated amount of dispersion in Si-H bond dissociation and hydrogen transport [115]. (b) As discussed in section B.4.2, such a triple-well model is nothing but an alternate representation of R-D model..... | 307 |
| C.1: Random walks in 1D lattice having $N+1$ sites ($j = 0, 1, 2, \dots, N$). Probability of right and left jumps at site j is represented by p_j and q_j respectively. Hence, $(1 - p_j - q_j)$ represents the staying probability at grid j . Taken from [214]. | 308 |
| C.2: One dimensional random network with transition rates between neighboring sites. E_0 represents the site energy at $i = 0$ | 312 |

- E.1: (a) Experimental analysis of scattering theory can be performed either by varying μ_{eff} using V_G and T or using strain. This work broadens the application of scattering theory, by varying μ_{eff} through defects, as well as V_G . (b) Flowchart for the experimental steps performed in validating scattering theory, where λ is varied by creation of additional coulomb scattering..... 318
- E.2: (a) Creation of interfacial defects (marked using $\Delta V_{T,lin}$) for a PMOS strained transistor increases λ (Figure 8.4d). Resultant change in λ effects both $I_{D,lin}$ and $I_{D,sat}$, which is monitored here using $(I-B_{lin})\Delta\lambda/\lambda_0$ and $(I-B_{sat})\Delta\lambda/\lambda_0$. (b) Estimated average $(I-B_{sat})/(I-B_{lin})$ for PMOS (due to interfacial defects) and NMOS (due to oxide defects) transistors (having L of 0.13~1 μ m) decreases with decrease in L , which indicates an increase in ballisticity with reduced L 320
- E.3: Interfacial defects in a PMOS transistor increases both B_{lin} and B_{sat} , when measured at fixed V_G , relationship between B_{lin} and B_{sat} remains approximately invariant of defects (a) Using Maxwell-Boltzmann approximation for Fermi integrals, *i.e.*, equations (8.10)-(8.11); (b) Using equations (8.8)-(8.9). Both the figures prove that for a particular transistor B_{lin} and B_{sat} depends mainly on $|V_G - V_T|$ 321
- E.4: Variation of λ extracted using (a) Maxwell-Boltzmann approximation for Fermi Integrals, *i.e.*, equations (8.10)-(8.11) or (b) exact expressions of equations (8.8)-(8.9). Figure (b) also shows the underestimation of λ induced by using $\lambda = 2k_B T \mu_{eff} / q v_{inj}$ [341]. 322
- E.5: (a) Extracted ℓ vs. $|V_G|$, using Maxwell-Boltzmann approximation (*i.e.*, equations (8.10)-(8.11)), shows an initial decrease, then saturation, when measured at $|V_{DS,sat}| = 1.1V$. Whereas, (b) variation of ℓ indicates a turnaround when measured at lower $|V_{DS,sat}| = 0.6V$. (c) Turn-around in ℓ variation starts to show up even at $|V_{DS,sat}| = 1.1V$, when exact expressions for drain currents or equations (8.8)-(8.9) are used. 323

ABSTRACT

Islam, Ahmad Ehteshamul. Ph. D., Purdue University, May, 2010. Theory and Characterization of Random Defect Formation and its Implication in Variability of Nanoscale Transistors. Major Professor: Muhammad Ashraful Alam.

Over the last 50 years, carrier transport has been the central research topic in the semiconductor area. The outcome was a dramatic improvement in the performance of a transistor, which is one of the basic building blocks in almost all the modern electronic devices. However, nanoscale dimensions of current transistor following Moore's law have shifted the spotlight from carrier transport towards the reliability and variability constraints. Modern transistors operate at a high electric field. They also use small metal gates and high- κ gate dielectric. Therefore, these transistors regularly suffer from process variations due to statistical variation in metal grain orientations at the gate, number of dopants in the substrate, thickness of the dielectric, *etc.* In addition to these 'time-zero' variation sources, presence of high oxide electric field and use of high- κ materials as dielectric (like silicon oxynitride and hafnium-based materials) strains the chemical bonds in the bulk and interface of the amorphous dielectric. As a result, defects are formed within a transistor, which leads to 'time-dependent' variation. Taken together, these 'time-zero' and 'time-dependent' phenomenon cause variation in transistor parameters (*e.g.*, threshold voltage, mobility, sub-threshold slope, drain current) – which eventually lead to the IC failure, when the variation goes beyond a certain pre-defined limit.

In this thesis, a physical model is developed to understand the defect formation at the dielectric/substrate interface of a transistor (a phenomenon, generally known as Negative Bias Temperature Instability), which is one of the major scaling concerns in current transistors. The time dynamics of interface defect generation is captured within a

Reaction-Diffusion framework and hence compared with the characteristic experimental signatures measured over a wide range of supply voltage, temperature, materials within the dielectric, and channel strain. This comprehensive analysis further establishes the subtleties in interface defect characterization using modern techniques and also explains the intricacies for analyzing the impact of defect generation at circuit level. More importantly, the study with interface defects has identified the presence of self-compensation in advanced CMOS technology. Later, such self-compensation is shown to be generally applicable to many sources of ‘time-dependent’ and ‘time-zero’ variabilities. Design of such variation-resilient transistor may reshape how circuits are designed and evaluated currently for handling process-induced (time-zero) and temporal (time-dependent) variations – which is one of the grand challenges for continuing transistor scaling following Moore’s law.

1. INTRODUCTION

1.1. Background

The invention of transistors in late 1940s has revolutionized semiconductor industry. In 1947, Bell Labs scientists William Shockley, John Bardeen, and Walter Brattain provided the first demonstration of a transistor (later known as a bipolar junction transistor or BJT) by placing two metal-semiconductor (point) contacts very close to each other [1, 2]. Here, the current through one point contact (passed at a particular direction) was used to control the current through the second contact. Later in the early 1960s, a current-controlled BJT was replaced with a voltage-controlled metal-oxide-semiconductor field-effect transistor (MOSFET), which significantly reduced the power consumption of the gate circuits. At the same time (1958-1959), Jack Kilby at the Texas Instruments and Robert Noyce at the Fairchild Semiconductor pioneered the concept of integrated circuits (IC), which allowed designers to place large number of transistors on a single chip in a much simpler, but efficient way. Since then, semiconductor industry has gone through significant scaling in transistor dimensions. Such scaling has not only made the IC's more compact and dense, but also enhanced its performance. Though the vision of Gordon Moore in 1965 that the complexities of an IC will double every one and half years [3] seemed to be a dream at that time, it came true over the next four decades – thanks to the relentless efforts from thousands of physicists, chemists and engineers during this period. At present, in the year 2010, we have more than a billion transistors in one single “Intel Atom” processor.

Although there were claims of transistor industry hitting a “red brick” wall at 100nm technology node in 1998 [4], scientists are currently working towards developing transistors even for the 22nm technology node. In such small dimensions, variability of transistor parameters is becoming increasingly important [5-7]. The sources of these

variabilities, present in nanoscale transistors, are shown in Figure 1.1, which can be generally classified into two broad types: spatial variability and temporal (or time dependent) variability.

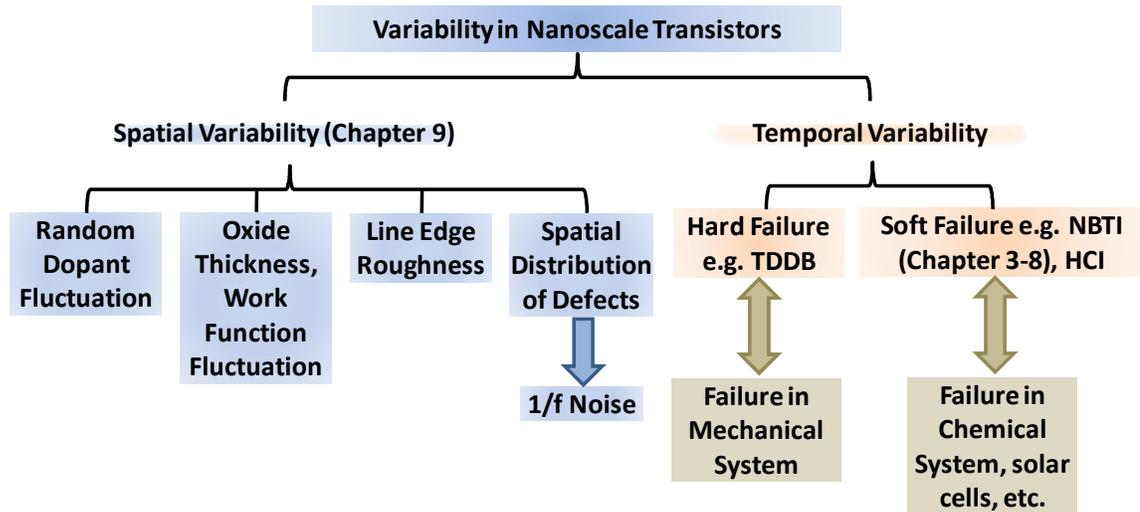


Figure 1.1: Different types of variabilities in nanoscale transistors.

1.1.1. Process-induced Spatial Variability

Aggressive scaling of MOS dimensions has made statistical (device to device, die to die, wafer to wafer, *etc.*) fluctuation of transistor parameters increasingly important [5, 8-11]. For example, the ratio of fluctuation in a MOS transistor parameter (having magnitude N) with respect to the magnitude of the parameter itself can be expressed as, $\sim \sqrt{N}/N = 1/\sqrt{N}$ [12]; which obviously increases with decrease in N . Moreover, fluctuations in nano-dimensions, could have been easily ignored in micron-size transistors of 1990s, but cannot be avoided in nanoscale transistors studied in current MOS technology.

In MOS transistors, the sources of statistical fluctuations include the ones associated with substrate dopant (commonly known as random dopant fluctuation), oxide thickness, gate workfunction, surface roughness, line edge roughness, *etc.* *Random dopant fluctuation (RDF)* arises from statistical variation in the number and the position of the

dopant atoms in the active region of MOSFETs. This makes the transistors microscopically different and as a result causes significant variations in drive current from device to device [8, 13]. Also, *gate oxide thickness fluctuation* in the order of 1-2 atomic layers [5, 14] is typically expected in the 5-6 atomic layers thick oxide, currently being used in MOSFET. Such oxide thickness fluctuation introduces variations in *MOS electrostatics*, *surface roughness limited mobility* and *gate tunneling*. In addition, *gate workfunction fluctuation* is especially a major variability concern in modern MOSFET involving metal gate/high- κ stack [15]. Such fluctuation arises from the statistical variation of metal grain orientations¹ and sizes over the transistor gate area. Moreover, statistical variation in the incident photon count during the lithography exposure, photon absorption rate, chemical reactivity, and molecular composition of the photo resist introduces unavoidable *line edge roughness (LER)* in the gate pattern (or channel length) definition among various transistors [17].

In additional to the statistical fluctuation from the above sources, *spatial distribution of defects* within oxide gives rise to flicker (1/f) noise in MOS structures, which has been extensively studied since 1950s [18-23]. Capture/emission of channel electrons/holes to/from an oxide defect give rise to noise, characterized by a distinct frequency, in the transistor's source-drain current. When the defects are spatially distributed within the oxide, the one located further away from the channel will give rise to low-frequency noise; whereas the one located near the channel will give rise to high-frequency noise. Subsequently, adding up the noise contribution from different defects will give rise to a flicker (1/f) noise in the source-drain current of MOS transistors.

¹ Metal workfunction depends on its surface orientation, because the surface atomic density that governs the workfunction of a metal depends on its surface orientation [16]. Metal orientation, having higher (lower) surface atomic density, has larger (smaller) electron wavefunction penetration across the metal surface, and hence has higher (lower) surface dipole moment, which has more (less) attraction on the surface electrons. Therefore, metal surfaces having higher (lower) surface atomic density have higher (lower) workfunction.

1.1.2. Defect-related Temporal Variability

1.1.2.1. History of Defects

Though defects are considered here as a source of temporal variability in nanoscale transistors, defect formation within materials have always been a concern in different aspects of condensed matter physics. In basic crystallographic courses [24, 25], one learns about the thermodynamics and equilibrium concentrations of the Schottky defects, the Frenkel defects, dislocations, stacking faults, grain boundaries, *etc.* during the growth of crystal structures. Analogously, amorphous materials (*e.g.*, calcogenide glasses, organic polymers) also suffer from topological defects (*e.g.*, dangling bonds, vacancies) mostly due to deviations from the Mott's 8-n rule [26], where n is the number of valence electrons in the atoms forming the amorphous structure. In addition, foreign materials or impurities are also present (due to intentional doping or unintentional consequence of material synthesis) in both crystalline and amorphous structures, thus modifying their properties. Although, these foreign materials are often advantageous for improving the performance of materials (*e.g.*, phosphorous doping in a crystalline silicon increases the electron conductivity, presence of impurities in glass reduces its melting temperature), in general, defects have adverse effect on carrier transport properties of a material used in transistors.

Moreover, defects are formed not only *within the bulk* of the crystalline or amorphous structures, but also *at the interfaces* among any two materials (either amorphous or crystalline). This is mainly due to the lattice mismatch between two dissimilar materials combined together; *e.g.*, dangling bonds at the Si/SiO₂ interface of a MOS transistor falls in this category. However, when two similar materials having slightly different properties are joined (*e.g.*, interfaces of p-n junction), interfacial defects result due to the non-equilibrium placement of impurities and dislocations at the interface of the materials.

Historically, these bulk and interface defects were always a problem in semiconductor industry. For reducing the amount of bulk defects from the dielectric

oxide layers (and also from the semiconducting substrate) during IC manufacturing, a semiconductor wafer and the subsequent ICs built on it go through a series of purifying process (gettering) steps. However, interfacial defects cannot be removed or reduced, like bulk defects, through any sort of purification. At the Si-SiO₂ interface of a MOS transistor, interfacial defects (dangling Si- bonds) are terminated or passivated (Si- + H → Si-H) by using hydrogen compounds like silane (SiH₄)², which only provides a time-zero solution from interface defects. These interface defects again starts to show up during the operation of a transistor, as discussed next.

1.1.2.2. Sources of Temporal Variability

After a period of transistor operation, the interfacial Si-H bond starts to dissociate or depassivate (Si-H → Si- + H) and the interface defects begin to reappear. Formation of these interface defects due to Si-H depassivation leads to time-dependent instabilities (or *soft-failure*) in the transistor parameters (*e.g.*, threshold or ON/OFF switching voltage, drivability), which is one of the major reliability concerns in current MOS technology. Such interface defect related instability is dominant in p-type MOS (PMOS) transistors and the associated phenomenon is known as Negative Bias Temperature Instability (NBTI). As the name suggests, NBTI indicates a temperature accelerated degradation in PMOS transistors, when it is stressed with negative gate voltage. Although NBTI was identified as a reliability concern in integrated circuits since mid 1960s [29], its significance became particularly important below the 130 nm technology node [6, 30-34] due to the – (a) use of increasingly higher oxide electric field for better transistor

² Historically, the Si-SiO₂ interface defects were the main problem behind the commercialization of the MOSFET before the BJT, though the concept of MOSFET [27, 28] predates BJT [1, 2] by two decades. The problem persisted till 1970s, when researchers from Fairchild Semiconductor followed the lead of amorphous-Si process trick of using hydrogen-passivation for reducing the Si-SiO₂ interface defects in a MOSFET, thus enabling its commercialization.

performance [35], and (b) use of oxynitride gate dielectric to reduce gate leakage and boron penetration effects [36-38].

In addition to the instability concerns from the interface defects, introduction of new materials within MOS structure (*e.g.*, oxynitrides and different Hf-based high- κ gate dielectrics for reducing gate leakage [38, 39], metal gate to avoid poly-depletion effects and Fermi-level pinning [40-42]) have generated additional concerns related to defects within the oxide material during NBTI stress. Thus, hole trapping into oxide defects has become an additional NBTI mechanism and electron trapping into oxide defects of n-type MOS (NMOS) transistors also become an extra soft-failure mechanism, which is termed as the positive bias temperature instability (PBTI) [43, 44].

Like NBTI/PBTI, defect formation is not always uniform. Pretty often, the carriers near the drain (or collector) end of MOS transistors cause non-uniform generation of defects near the drain and thus lead to Hot-Carrier Injection or HCI, as an additional form of device instability [45, 46]. Moreover, the size of current MOS transistors necessitates the consideration of statistical fluctuation in defect formation at the interfaces and also within the oxide [47-50], in addition to the time-dependent mean shift discussed above.

In addition to the defect related transistor instabilities discussed above, formation of oxide defects and their subsequent alignment can increase the gate leakage current to such an extent that the dielectric film will no longer be effective in insulating the gate of a MOS structure – thus causing dielectric breakdown or *hard failure* of transistors. Such time-dependent dielectric breakdown (TDDB) is a key transistor reliability issue in modern MOS transistor, which often limits its operating lifetime [51-57].

1.1.2.3. Optimizing the Effect of Defect

In modern transistors, defect formation is becoming more and more important, because a small change in transistor performance over the period of operation can bring a transistor out of the window of design tolerance and thus make the IC useless. Device designers often feel that processing modifications can do little to address these problems.

Hence, a number of efforts have been undertaken to alleviate its effect on circuit level. For example, a standard option is to operate the transistors with an extra guard band voltage, over and above the voltage required for operation [58, 59], so that the circuit remains functional despite the defect formation. Similarly, there are proposals for using adaptive body bias [60-62] or adaptive power supply (*i.e.*, circuit sleeping) [63] for minimizing the impact of degradation. Moreover, algorithms involving area-resizing of transistors over the critical path in a digital circuit have also been proposed for optimizing the circuit-level degradation [64].

1.1.2.4. Defects in Analogous Areas

Most of the variabilities in MOS transistors have their analogous (*i.e.*, have similar origin) forms in many other systems. For example, TDDB in thick oxides of MOS transistors and crack propagation in mechanical systems [65] can analogously be explained using Weibull statistics. Similarly, interface defect formation during NBTI involves dissociation of Si-H bond and is based on similar physics as the dissociation of Si-H bonds in amorphous-Si solar cells [66]. Moreover, since chemical sensors often use MOS-like structures [67], the implication of defects in those systems can also be explained in an analogous form.

1.2. Objective of the Thesis

This thesis mainly focuses on developing a theoretical framework and characterization techniques for Negative Bias Temperature Instability (NBTI) in PMOS transistors, by analyzing its functional dependencies over time, supply voltage, temperature, materials within the dielectric, and channel strain. Such comprehensive analysis also allows us to identify different features in modern NBTI characteristics. Here, we recognize both interface and oxide defects as contributing factors for NBTI in current CMOS technology. In some transistors, NBTI is dominated by interface defects and the mechanics of NBTI is modeled within an analytically tractable Reaction-

Diffusion framework and compared with the characteristic experimental signatures. In other type of transistors, hole trapping into pre-existing oxide defects dominates and generalized Shockley-Read-Hall framework is used to model the kinetics of hole trapping in these transistors. Taken together, this comprehensive theoretical framework – supported by detailed experiments – lead to a truly exciting possibility of the design of a NBTI-resilient transistor technology for CMOS architecture. Finally, we generalize the extent of NBTI-resilient transistor design for resilience against other sources of temporal and spatial variations. The proposed concept suggests the design a variation-resilient CMOS transistor, a possibility – if demonstrated and adopted – may reshape how CMOS circuits are currently designed for handling major sources of CMOS variabilities.

1.3. Organization of Thesis

The thesis covers one important aspect of MOS transistor’s temporal variability and then extends the discussion into other analogous aspects of transistor variability. In addition to this introductory chapter, the thesis contains nine (9) additional chapters. *Chapter 2* summarizes the existing views on the types and energetics of defects, formed at the oxide/substrate interface (interface defects), as well as within the bulk of the oxide (oxide defects), of a MOS transistor. *Chapter 3* analyzes the formation of interface defects (N_{IT}) through the dissociation of interfacial Si-H bonds in a wide variety of nitrided and strained transistors, and thus develops a consistent model for predicting N_{IT} formation in CMOS transistors. Later in *Chapter 4*, we use the same formalism developed in chapter 3 to study the statistical distribution interface defect generation in nanoscale transistor. Then we move on to *Chapter 5* and cover the dynamics of hole trapping into pre-existing oxide defects (N_{HT}), which is especially important for a particular type of PMOS transistors, having high- κ dielectric. Next in *Chapter 6*, we analyze the techniques that are recently being used for characterizing interface/oxide defects of a transistor and identify different limitations of these techniques. Later, our understanding of defect formation is used in *Chapter 7* to perform a quantitative analysis of defect formation vs. off-state power consumption (gate leakage). Similarly, *Chapter 8*

studies the impact of defect generation on the transistor performance parameters, like sub-threshold slope, mobility, drain current, *etc.* Here, we also describe how the interplay between mobility and carrier number fluctuation leads to a (degradation-free) transistor having negligible drain current variation due to defect generation. *Chapter 9* extends the above concept of optimizing temporal fluctuation (degradation-free transistor) and show how a similar concept can be used as well for optimizing different sources of spatial fluctuations, thus leading to the idea of a variation-resilient transistor. Finally, in *chapter 10*, we summarize the main findings of our research and propose future works for utilizing the understanding for resolving some emerging issues.

2. ORIGIN OF DEFECTS IN TRANSISTORS

2.1. Introduction

Electrically active defects are atomic configurations, which creates new electronic states in the band gap of crystalline or amorphous materials. This chapter summarizes the existing views on the types and energetics of defects, formed at the oxide/substrate interface (interface defects), as well as within the bulk of the oxide (oxide defects), in a metal-oxide-semiconductor (MOS) structure. Since its introduction in 1960s, semiconductor/ insulator materials, used in MOS structure, have evolved considerably with scaling and technology evolution. MOS technology initially started with Si/SiO₂ as the semiconductor/insulator interface and gradually incorporated high- κ dielectrics (incorporating silicon oxynitrides and Hf-based materials) for reducing gate leakage. In addition, below 90nm MOS technology Si- channel is also strained to improve transistor performance. To cover all these semiconductor/ insulator interfaces, we begin with a discussion on interface defects in classical Si/SiO₂ system (section 2.2). Then we move on to the strained Si/SiO₂ system and discuss the associated changes in interface properties due to the incorporation of strain (section 2.3). Next in sections 2.4 and 2.5, we consider the types of interface defects, present in Si/SiON and Si/high- κ /Interfacial layer (IL) structures.

Following the review on interface defects, we discuss the standard views of oxygen vacancy related oxide defects in the Si/SiO₂ system (section 2.6), which are mainly responsible for time-dependent dielectric breakdown studies [51-57]. However, recent introduction high- κ dielectric within the MOS structure made these oxide defects relevant even for bias temperature instability studies. Section 2.7 presents the current understanding of these oxide defects in high- κ MOS structures.

2.2. Interface Defects: Si/SiO₂ Structure

Interface defects at the Si/SiO₂ interface involve dangling Si- bonds, which are left behind after oxidation of bare Si wafer and subsequent treatment using forming gas anneal [29, 68-70]. To estimate the concentration of dangling bonds at the Si/SiO₂ interface, note that Si-Si bond length in crystalline-Si (c-Si) is $\sim 5\text{\AA}$. Assuming that average distance between two dangling Si- bonds is also same³, bandgap of Si is $\sim 1\text{eV}$ [73], and interfacial thickness is on the order of an \AA [69], total number of dangling Si- bonds at the bare c-Si interface would be, $D_{IT} \sim (1/5\text{\AA}^3) \times 1\text{\AA} / 1\text{eV} \sim 10^{14} \text{ cm}^{-2}\text{-eV}^{-1}$. After oxidation (before hydrogen treatment) some of the dangling Si- bonds are terminated by O atoms, leaving a D_{IT} of $\sim 10^{12} \text{ cm}^{-2}\text{-eV}^{-1}$ [69, 74, 75]. Then subsequent annealing of the Si/SiO₂ structure using hydrogen compounds (*e.g.*, H₂, silane SiH₄) at 200-500 °C reduces D_{IT} to $\sim 10^{10} \text{ cm}^{-2}\text{-eV}^{-1}$ or even below [6, 69, 74, 75]. We will use these three numbers, *i.e.*, $D_{IT} \sim 10^{10} \text{ cm}^{-2}\text{-eV}^{-1}$ for passivated Si/SiO₂ interface, $D_{IT} \sim 10^{12} \text{ cm}^{-2}\text{-eV}^{-1}$ for fully de-passivated Si/SiO₂ interface, and $D_{IT} \sim 10^{14} \text{ cm}^{-2}\text{-eV}^{-1}$ for the number of dangling Si- bonds before oxidation, in many places in subsequent discussions.

2.2.1. Nature of Si/SiO₂ Interface Defects

Nature of interface defects in Si/SiO₂ structure are mostly characterized by using Electron Spin Resonance (ESR), or its electrically detectable version, such as Spin Dependent Recombination (SDR) experiments (see section 6.3 for details on this characterization technique). Such experiment for characterizing (111) Si/SiO₂ interface was first performed in [76]. Later, many research groups [68, 77-79] have repeated and standardized the experimental procedures on both (111) and (100) Si/SiO₂ interfaces. These studies have identified the presence of P_b center in (111) interface and P_{b0}, P_{b1} centers in (100) Si/SiO₂ interface. P_b and P_{b0} centers also show similar magnetic resonance properties [78]. Thus, P_b and P_{b0} centers are (electrically) equivalent dangling bonds and generally believed to have three back-bonded Si atoms (Figure 2.1a). On the

³ Reconstruction of silicon surfaces [71, 72] does not change the essential argument.

other hand, detailed structure of P_{b1} center is not so well understood. Though it was initially proposed [78] to have two back-bonded Si and one back-bonded O atom (Figure 2.1b), later experiments of [80] proved that all P_b centers should have same neighbors, *i.e.*, three back-bonded Si atoms. However, it is well accepted [81] that P_{b1} center is clearly a different (compared to P_b and P_{b0}) type of dangling bonds at the interface.

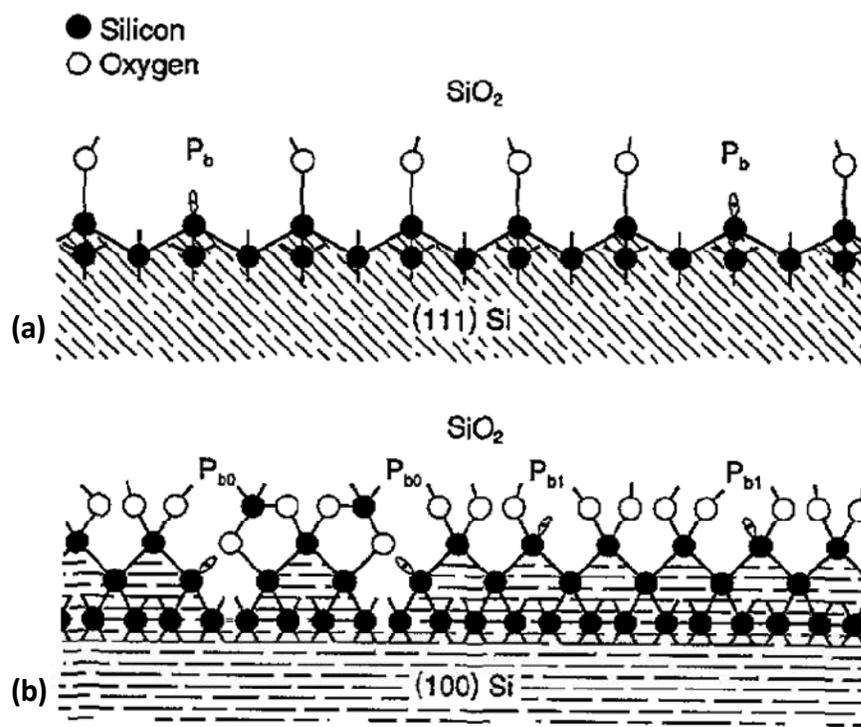


Figure 2.1: Structure of paramagnetic centers on Si wafers for (a) (111) and (b) (100) Si/SiO₂ interfacial orientation. Taken from [69]. Structure of P_{b1} center in Fig. b is the one proposed in [78], whereas study in [80] indicates P_{b1} center also having three back-bonded Si atom, like P_b and P_{b0} centers.

2.2.2. Energetics of Si/SiO₂ Interface Defects

Energy distribution of P_b centers within the bandgap has also been studied in both (111) and (100) Si/SiO₂ interfaces [77, 82-85]. Here also, P_b center in (111) and P_{b0} center in (100) are widely accepted to have distributed donor-like (neutral when filled

with electron and positively charged after donating electron) and acceptor-like (neutral when empty of electron and negatively charged after accepting electron) states near the valence and conduction band edges within the Si-bandgap, respectively. However, P_{b1} center in (100) are either reported to have a wide distribution similar to P_b or P_{b0} centers, according to some reports [85, 86], or shown to have a narrow distribution near the Si midgap with a skew towards the valence band, according to others [87]. Overall, the total interface defects (sum of all P_b centers) are uniformly distributed throughout the entire bandgap [69, 88].

Although the above studies discuss the existence of both donor and acceptor type of interface defects, recent capacitance-voltage measurements on PMOS transistors subjected to negative bias stress (also known as NBTI stress) [89, 90] have indicated the existence of only donor-type defects at the Si/SiO₂ interface. This discrepancy between ESR experiment (showing both donors and acceptors) and C-V measurement (showing only donors) are not well understood, although it could be related to the process conditions. In any case, the recognition that the existing defect density is relatively low and all new defect generation (due to NBTI stress) is primarily donor-type will be sufficient for our purposes, as discussed in subsequent chapters.

2.2.3. Theory of Donor/Acceptor Defects

Following the analysis in [91, 92], let us consider the energy levels for donor/acceptor type of defects from thermodynamic point of view. Three-fold co-ordinated dangling Si- bonds are considered to have three charge states: T_3^0 , T_3^- , and T_3^+ . The formation energy (defined as the change in total energy of an ideal random network when the defect is introduced) of uncharged defect T_3^0 is denoted by $F[T_3^0]$, which is independent of Fermi-level at the Si/SiO₂ interface (E_{FS}). Hence, the formation energies of the two charged states of T_3 are:

$$F[T_3^+, E_{FS}] = F[T_3^0] + E_{FS} - E(0/+) \quad (2.1)$$

and

$$F[T_3^-, E_{FS}] = F[T_3^0] + E(-/0) - E_{FS}, \quad (2.2)$$

where $E(r/s)$ represents the thermodynamic transition energy that is required to move from charge state r to s . The last two terms in equation (2.1) represent the energy expended to ionize the neutral defect, thus completely relax the resulting positive defect, and then deposit the electron at E_{FS} . Similarly, the last two terms in equation (2.2) represent the energy expended to remove an electron from E_{FS} , deposit it on the neutral defect, and thus completely relax the resulting negative defect. At temperature T , the concentration of any T_3 defect will be –

$$n(T_3^q, E_{FS}) = n_{Si} \exp\left[-F(T_3^q, E_{FS})/k_B T\right], \quad (2.3)$$

where n_{Si} is the number of Si atoms within the silicon matrix at which a defect can form ($\sim 10^{12} \text{ cm}^{-2} \text{ eV}^{-1}$ following the discussion in section 2.2).

Figure 2.2a exhibits both formation energy and the *log* concentration of T_3 defects as a function of E_{FS} . As of equations (2.1) and (2.2), the formation energies of the positively (negatively) charged dangling bond increases (decreases) linearly with E_{FS} . Also, according to equation (2.3), defect concentration increases with decrease in formation energy, which is also shown using the right-hand axis of Figure 2.2a. Here, defects are assumed to have two sharp thermodynamic transitions within the Si bandgap such that effective correlation energy,

$$U_{eff} = E(-/0) - E(0/+) \quad (2.4)$$

is positive. At a given value of E_{FS} , the defect with the lowest value of F is found in the greatest concentration. Thus, for a system having $U_{eff} > 0$ (necessary to interpret the existence of neutral dangling bond in ESR experiments [93, 94]), donor type interfacial defects are favorable near the valence band and acceptor type defects are favorable near the conduction band, as shown in Figure 2.2b. However, $U_{eff} < 0$ is sometimes obtaining from optical-absorption experiments on doped amorphous silicon structures [92].

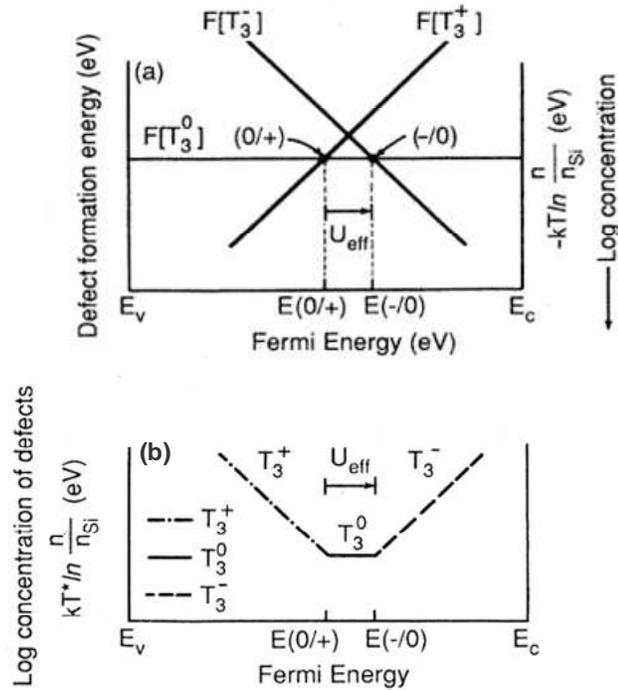


Figure 2.2: (a) Formation energies and (b) concentration of dangling bonds (T_3) as a function of E_{FS} .

2.3. Interface Defects: Strained Si/SiO₂ Structure

Interface defects in a strained-Si/SiO₂ system have been studied in only a few reports. In one study [95], *in-situ* tensile/compressive strain on (111) Si/SiO₂ structure is applied using mechanical-bending apparatus to show that the type of interface defects (*i.e.*, magnetic resonance properties of P_b centers) does not change with strain. However, defect density increases (decreases) due to the application of compressive (tensile) strain. To explain this, it is hypothesized that compressive (tensile) strain decreases (increases) Si-Si bond length within the substrate and thus increases (decreases) the lattice mismatch between Si and SiO₂ layer – causing P_b signal to increase (decrease) with compressive (tensile) strain. In another study [96], biaxial tensile strain of silicon channel is varied in an (100) Si/SiO₂ structure, by varying Ge content (1-x) of the Si_xGe_{1-x} epi-layer placed underneath the Si-channel. Thus, it is shown that tensile strain – applied using epi-layer –

can indeed reduce P_b center density, which is again consistent with the mechanical strain dependent study on (111) Si/SiO₂ structure [95].

2.4. Interface Defects: Si/SiON Structure

Recently, spin dependent recombination (SDR) experiments (which mainly probes interfacial region) are performed on MOS transistors, having (plasma-dosed) silicon oxynitride (SiON) dielectric [88, 97]. Such study reveals that interface defects in Si/SiON structure has resonance properties, which are completely different compared to the one present in Si/SiO₂ structure. Contrary to the presence of anisotropy in g -values (which is a spin orientation dependent parameter and is affected by spin-orbit coupling; see section 6.3 for details) for the interfacial P_b centers, in Si/SiO₂ structures, g -values for interface defects in Si/SiON structure is observed to have no directional dependency or isotropic [88, 97]. Such isotropic nature of g -values indicates that the interface defects in Si/SiON structures have no preferential orientation, similar to the oxygen vacancy related E' center, present in bulk SiO₂ (see section 2.6). Moreover, the g -values in Si/SiON structures, its directional independence, centerline width of resonance signal, and hyperfine splitting parameter suggest the similarity of Si/SiON interface defects with the K centers [88, 97], observed in Si₃N₄ dielectrics [98], which have three back-bonded nitrogen atoms. Finally, (forward p-n junction) bias dependent study in SDR experiments on Si/SiON structure suggests the localization of defects near the Si midgap. However, the presence of K centers and energetic localization near mid-gap in Si/SiON structures are yet to be confirmed by other research groups.

2.5. Interface Defects: Si/high- κ Structure

P_b -like interface defects are expected also in Si/high- κ structures, because of the existence of thin (~ 1 nm) SiO₂ interfacial layer on these transistors [39]. ESR studies on Si/HfO₂ structure [99] (where, HfO₂ are grown through atomic layer deposition [100]) identified the presence of P_b -like interface defects, similar to the (111) P_b centers and

(100) P_{b0} centers, but with small shifts in the resonance conditions. Similarly, another recent ESR study [101] have also reported observation of P_b like centers, along with resonance shifts. On the other hand, while a separate study [102] confirmed the existence of P_{b0} and P_{b1} centers in HfO_2 -based device structures, however they did not observe any shifts in resonance conditions compared to the values reported for conventional Si/SiO₂ interfaces. In general, regardless of the dielectric material, forming gas anneal (H_2/N_2) is observed to reduce the density of P_b or P_b -like centers [99, 102] universally.

2.6. Oxide Defects in SiO₂

A number of point defects such as the oxygen vacancy or silicon-rich sites, peroxy radicals, non-bonding oxygen centers (NBOC), and hydrogenic species have been detected in amorphous SiO₂ [103]. However, in SiO₂ constituting the dielectric layer in the MOS devices, majority of the detectable effects appear due to the oxygen vacancy centers (Figure 2.3a), referred to as the E' centers (first proposed in [104]) in the literature.

E' centers generally involve an unpaired electron localized on a silicon atom back-bonded to three oxygens. However, ESR experiments [105-107] also suggest the existence of hydrogen-related E' centers (Figure 2.3d-e), which are identified using doublets or two sidebands below and above the main resonance signal. The configurations of Figure 2.3 are only reasonable first-order approximations of E' centers, commonly discussed in literature [69, 81, 86]. The roles of E' centers in high quality thermally grown oxides are fairly well understood, but more work is needed for unambiguous specification. In general, these defects are expected to have the following properties [81], though details structure of the E' variants are yet to be identified:

- They can act as hole traps,
- These defect levels play a significant role in trap-assisted tunneling phenomena in electrically stressed oxides; for example, stress-induced leakage currents,

- ESR signals from the E' centers show no directional dependence (*i.e.*, has isotropic g values).

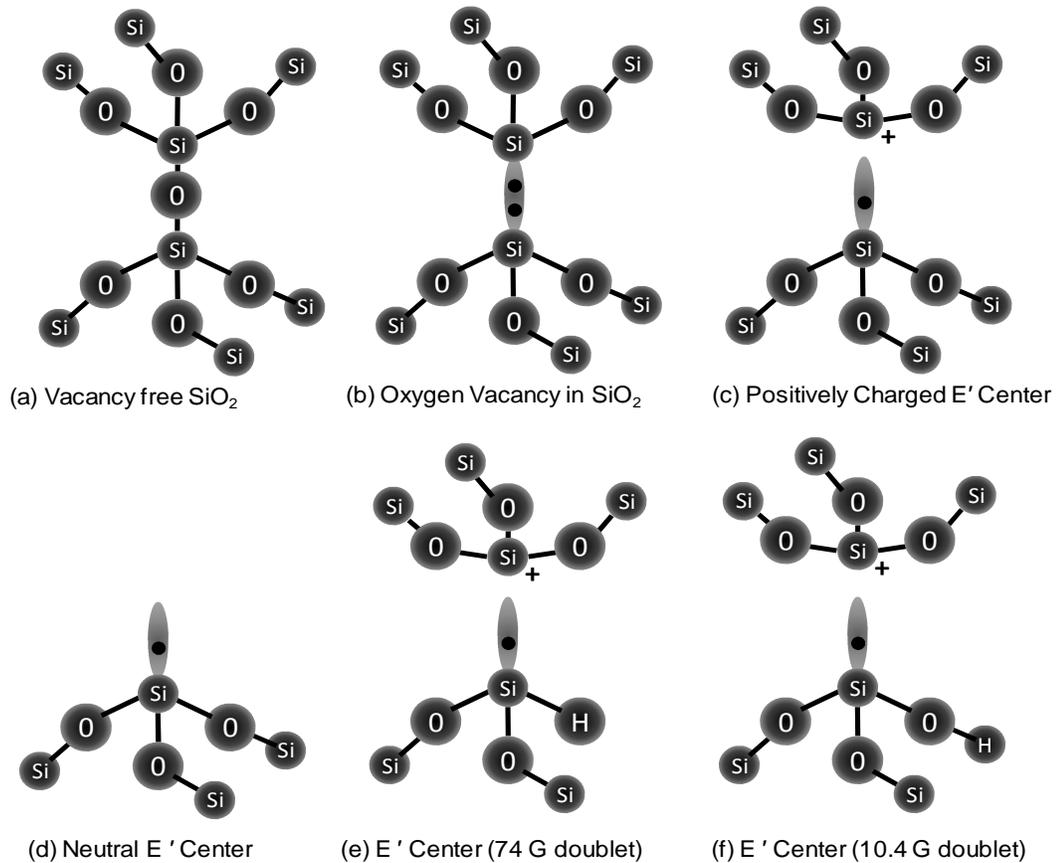


Figure 2.3: (a) Oxygen vacancy-free SiO_2 configuration. (b-f) Possible types of oxide defects or E' centers, present in SiO_2 dielectric. (b) Oxygen vacancy or Si-Si bridge is considered to be origin of many E' centers. (c) Capture of holes by oxygen vacancy leads to positively-charged E' center. Other schematics in this figure correspond to the illustration (d) for a neutral E' center, (e) for a E' center (74 G doublet), and (f) for a E' center (10.4 G doublet). These configurations are taken from [69, 81, 86].

2.7. Oxide Defects in High- κ Material

In general, high- κ dielectrics are more prone to the formation of oxide defects [39, 108-110] compared to SiO_2 dielectric. Fundamentally, this is because SiO_2 has a low

coordination number (or number of nearest neighbors), so that its bonding can relax and re-bond any broken bonds at possible defect sites [108]. Any remaining broken bonds or defects can easily be passivated through hydrogen treatment. On the other hand, high- κ oxides have higher coordination numbers and they have more ionic bonding compared to SiO_2 [108]. Thus, they are poor glass formers (*i.e.*, have lower crystallization temperature) and their bonding cannot relax as easily as SiO_2 [108-110]; hence they have large intrinsic defect concentration. Much of the present-day engineering of these oxides consists of pragmatic strategies of trying to reduce defect densities by processing control and annealing.

Similar to SiO_2 , oxide defects in HfO_2 -based dielectrics can also have E' centers. As reported in [111], there can be two types of E' centers of quite high density (up to 10^{19} cm^{-3}) within the interfacial layer of the high- κ dielectric structure. Moreover, density of E' centers depends quite strongly upon the ALD deposition conditions, and the presence of a metal gate greatly enhances E' generation resulting from the post deposition anneals [81].

2.8. Summary

Nature and energetics of interfacial and oxide defects have been studied over a long time, mostly using electron spin resonance (ESR) experiments. Except for recent reports on Si/SiON structure [88, 97], all interface defects have anisotropy or directional dependence in the measured g -values. On the other hand, oxide defects (in general) have isotropic g -values. Although a comprehensive understanding of these defects are yet to be established, their overall importance as the interface defects (P_b) and oxide deep level centers (E') are quite clear. In the next two chapters (chapters 3 and 4), we will discuss the formation interface defects, which will be followed by a discussion on pre-existing oxide defects in chapter 5. Later, different electrical characterization techniques for defects, in addition to ESR, are discussed in chapter 6. Thus in chapter 6, we highlight why ESR experiment is not enough for defect characterization. Subsequently, we discuss the effect of defect formation in transistors performance in remaining chapters.

3. INTERFACE DEFECTS

3.1. Introduction

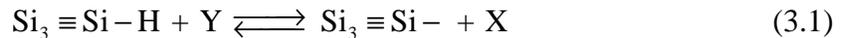
In conventional CMOS technology, having SiO₂ or high- κ materials as gate dielectric, the oxide/substrate interfacial layer consists of both silicon-oxygen (Si-O) and silicon-hydrogen (Si-H) bonds. Of these two types of bonds, Si-O bonds have higher activation energy and needs interaction with energetic carriers, normally present in off-state hot-carrier injection study [112-114], to dissociate. Such dissociation generates dangling bonds (Si-) or defects at the interface. On the other hand, Si-H bonds gets dissociated and forms dangling bonds in presence of cold holes near the interface, even at normal operating condition [30, 115-119]. In this chapter, we analyze the dissociation of these Si-H bonds in a wide variety of nitrated and strained transistors in order to develop a consistent model for predicting interface defect formation. We start by discussing the existing views of Si-H bond dissociation (resultant dangling bond or interface defect) in section 3.2. Later, we present different versions of reaction-diffusion (R-D) model in predicting time dynamics of Si-H bond dissociation (sections 3.3-3.5) and reverse annealing of dangling Si- bonds (section 3.6). Our analysis explains NBTI experiments measured over a wide range of time (sections 3.5.4, 3.5.5; H-H₂ R-D model), electric field (section 3.7), temperature (sections 3.5 and 3.7), nitrogen within the dielectric (section 3.8), and channel strain (section 3.9). Later, we discuss how our proposed field model (section 3.7) can be used for extracting transistor lifetime and hence can predict long term stress behavior for reliability analysis (section 3.10). Finally, in relation to the contents of this chapter, we discuss different alternate models and their limitations in Appendix B, which have recently been proposed for interpreting Si-H bond dissociation. More interestingly, we show that some of these models (see section B.4.3) are different representations of the classical R-D models.

3.2. Existing Views of Interface Defect Formation

It has been known since 1970s that the presence of hydrogen during oxidation of silicon surface can passivate the dangling Si– bonds at the oxide/substrate interface [29, 76, 82, 120]. Later works from different research groups [68, 121, 122] confirmed the efficiency of atomic hydrogen (H) in passivating the surface dangling bonds. These passivated dangling bonds (or, Si-H bonds) became the major source of device instability under the application of negative gate bias V_{STS} (or, negative oxide electric field E_{ox}) and at elevated temperature [6, 29, 30, 117, 123, 124]; a phenomena commonly known as Negative Bias Temperature Instability (NBTI). As such, the dangling Si- bonds reappear as interface defects (*reaction*) and generated hydrogen species move away from the interface (*diffusion*). The entire phenomenon of interface defect (N_{IT}) formation is generally captured within a Reaction-Diffusion (R-D) framework, which has several variants in existing literature. Variation within the existing R-D models mainly arise due to the differences in explaining (a) the Si-H bond dissociation, (b) the nature of generated hydrogen species and (c) the mechanism of hydrogen species getting away from the interface. In the following sections, we start with the first proposal of R-D model by Jeppson *et al.* [117] (section 3.3) and later explain the main features of the existing variants of this original proposal.

3.3. First Reaction-Diffusion (R-D) Model

A reaction-diffusion model for interface defect generation was first proposed by Jeppson et al. [117] with the dissociation reaction having the following format:



where, a particular species Y reaches the interface, reacts with the Si-H bond, creates N_{IT} and a second species X diffuses away from the interface. Diffusing species in [117] is considered as neutral. Thus, the governing differential equation corresponding to the reaction and diffusion parts can be expressed as –

$$\frac{\partial N_{IT}}{\partial t} = k_F (N_0 - N_{IT}) - k_R N_{IT} N_X^{(0)}, \quad (3.2)$$

$$\frac{\partial N_X}{\partial t} = D_X \frac{\partial^2 N_X}{\partial z^2}, \quad (3.3)$$

where, k_F is the Si-H dissociation rate (influence of Y as a catalyst is considered within this factor), k_R is the annealing or backward reaction rate, N_0 is the initial ($t = 0$) concentration of Si-H bonds, N_X is the concentration of diffusing species X along the z -axis ($z = 0$ is considered as the interface), and D_X is the diffusion co-efficient for species X within the diffusing media. In addition to equations (3.2) and (3.3), we also need to satisfy the following boundary condition near the interface between reaction and diffusion fluxes –

$$\frac{\partial}{\partial t} \left[\frac{1}{2} \delta N_X^{(0)} \right] = \frac{\partial N_{IT}}{\partial t} - D_X \frac{\partial N_X^{(0)}}{\partial z}, \quad (3.4)$$

where, δ is the interfacial layer thickness, which is approximately equal to the length of Si-H bond of 1.5 Å [69], and superscript (0) indicates near-interface quantities. Equation (3.4) designates that the concentration of species X at the interface is increased by Si-H dissociation (through equation (3.2)) and is decreased by outward diffusion flux (represented using equation (3.3)).

For obtaining a solution of equations (3.2)-(3.4) at $t \gg 0$, the system is considered to be diffusion-limited, such that the surface reaction equation (3.2) is in quasi-equilibrium, *i.e.*, $\partial N_{IT}/\partial t \sim 0$ or –

$$k_F N_0 \cong k_R N_X^{(0)} N_{IT}. \quad (3.5)$$

Moreover, since solution of equation (3.3) is an error function $\sim \operatorname{erfc} \left[z / \sqrt{D_X t} \right]$, having significant values of N_X from $z = 0$ to $z \sim \sqrt{D_X t}$, it can be approximated using a triangular profile [30, 125, 126]. Therefore, the total concentration of species X from $z = 0$ to $z \sim \sqrt{D_X t}$ (*i.e.*, N_{IT}) can be expressed as –

$$N_{IT} = \int_{z=0}^{z=f(D_X, t)} N_X(z, t) dz \sim \frac{1}{2} N_X^{(0)} \sqrt{D_X t}. \quad (3.6)$$

Combining equations (3.5) and (3.6), we have –

$$N_{IT} \sim \left(\frac{k_F N_0}{k_R} \right)^{1/2} (D_X t)^{1/4}. \quad (3.7)$$

Thus, Jeppson's R-D model has one-to-one correlation between N_{IT} and diffusing species X, according to equation (3.1), and hence predicts N_{IT} formation to have a power-law time exponent, $n \sim 1/4$. Moreover, considering both forward and reverse reaction, as well as diffusion processes, are Arrhenius activated (*i.e.*, have the form of $\sim \exp[-E_A/kT]$, where E_A is the activation energy for the process under consideration), equation (3.7) suggests the following expression for activation energy associated with N_{IT} formation:

$$E_{A,IT} = \frac{1}{2} (E_{A,F} - E_{A,R} - aE_{ox}) + \frac{1}{4} E_{A,X} \quad (3.8)$$

where, $E_{A,F}$, $E_{A,R}$, and $E_{A,X}$ are activation energies for k_F , k_R , and D_X , respectively; and aE_{ox} represents the field-induced barrier reduction with a being the effective dipole moment of Si-H bond (see section 3.7 for detailed discussion on this E_{ox} dependence).

Although Jeppson *et al.* proposed the first version of R-D model, they did not clarify the following aspects in their model:

- The exact mechanism of Si-H bond dissociation is unspecified in this model; an improvement in this respect is discussed in section 3.4.
- Diffusing species is also unidentified here, which we discuss in section 3.5. More specifically, it has been shown in sections 3.5.6 that power-law time exponent (stated as $\sim 1/4$ in equation (3.7)) and activation energy for N_{IT} formation depends on the diffusing species, as well as the associated diffusion mechanism.
- Jeppson *et al.* used standard measure-stress-measure setup (see section 6.4) in measuring N_{IT} formation, which obviously suffers from significant limitation (see chapter 6 for details). Significant refinement in N_{IT} measurement (mainly reduction in t_{meas} , see section 6.2) suggests power-law time exponent, $n \sim 1/6$ [119, 125, 127-

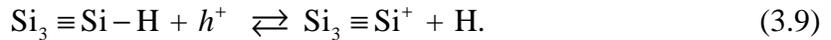
130], $E_{A,IT} \sim 0.1\text{eV}$ [119, 129] and $a \sim 0.8 \text{ q}\text{\AA}$ [119, 128], which is inconsistent with Jeppson's measurement and subsequent analysis. We address these limitations in subsequent sections of this chapter.

3.4. Si-H Dissociation in R-D Model

Although the exact mechanism of Si-H bond dissociation is unspecified in the original R-D model [117], two proposals for Si-H bond dissociation can be found in recent literature.

3.4.1. Hole-assisted Si-H Dissociation

This version of R-D model considers direct dissociation of Si-H bond due to the existence of high negative electric field at the interface [30, 31, 118, 119, 125, 131]. Since the dissociation is enhanced by the existence of holes near the interface (during PMOS inversion or NMOS accumulation), one presumes the following reaction to weaken the Si-H bonds at the interface (which replaces Y with holes h^+ and X with atomic hydrogen H in equation (3.1)) –



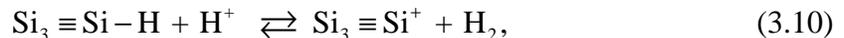
The resultant hydrogen species diffuses away from the interface and depending on the nature of diffusion, time dependence of N_{IT} formation can have different values, as discussed in section 3.5.

This model also allows the possibility of dimerization for generated H atom into H_2 molecule [131, 132], thereby explaining the observed $\sim 1/6$ power-law time exponent in recent ΔN_{IT} measurements [118, 119, 125, 127-130]. Moreover, the model also provides a consistent explanation for nitrogen (within the dielectric) and strain (within the MOS channel) dependency of N_{IT} formation [119, 128, 133], a feature, which cannot be predicted by the competing dopant-activated Si-H dissociation model, discussed in

section 3.4.2. Our field dependent study over wide range of voltages and temperatures is also consistent with this Si-H dissociation model, as discussed in section 3.7.

3.4.2. Dopant-activated Si-H Dissociation

Some of the recent first-principle calculation [134, 135] suggests that direct dissociation of Si-H bond (discussed in section 3.4.1) may be difficult, as the process has a dissociation energy barrier, $E_F \sim 2.5$ eV (which in presence of hole reduces to ~ 2.3 eV). Moreover, the reaction is strongly endothermic (*i.e.*, has positive reaction energy, $\Delta E \sim 2$ eV). As such, Si-H bond dissociation must be initiated by the availability of additional agents like positive hydrogen ion (H^+) at the interface, rather than only holes, as in the hole-assisted model (section 3.4.1). Capture of H^+ reduces E_F down to ~ 1.0 - 1.1 eV, which in presence of hole further reduces to ~ 0.5 eV. Here, passivated dopants like phosphorous-hydrogen (P-H) complex, present in PMOS transistor with n-type substrate, acts as the supply of H (through dissociation). These dissociated H become positively charged by capturing holes present in the inverted PMOS substrate. In addition, H stored in Si-O-Si bonds (which normally act as H trapping sites) next to the Si-H entity can also act as a source of H^+ . The reaction in dopant-activated Si-H dissociation model takes the following form (by replacing Y and X in equation (3.1) with H^+ and H_2 respectively):



which, thus results in a dangling bond and hydrogen molecule.

Dopant-activated model of equation (3.10) can explain the reduced amount of N_{IT} formation in a NMOS transistor compared to a PMOS transistor [90, 136], when both of them are subjected to similar negative V_{STS} . The explanation here involves the observation of higher dissociation energy (~ 1.28 eV) for boron-hydrogen (B-H) complex, compared to the dissociation energy (~ 0.3 eV) for P-H complex in an inverted PMOS [134]. Nevertheless, dopant-activated model has the following drawbacks:

- Although dopant-activated Si-H dissociation model provides an interpretation for the difference among measured ΔV_T in NMOS and PMOS transistor, when both

are subjected to similar negative V_{stress} , as reported in [90, 136]; several alternate explanations are also possible for the observed phenomena:

- 1) Defects at the oxide/substrate interface are either reported to be acceptor-like in the upper-half and donor-like in the lower half of the band gap [77, 82-85] or donor-like in both halves of the band gap [89, 90]. When N_{IT} distribution has only donors [89, 90], ΔV_T from interface defects are indeed expected to be smaller in NMOS compared to PMOS [124]. However, when N_{IT} distribution has both donors and acceptors [82], contribution from hole trapping may be required to explain higher ΔV_T in PMOS [124].
 - 2) The differences among NMOS and PMOS ΔV_T 's can also result from poly-gate work-function (χ_{PG}) difference between NMOS and PMOS transistors, which approximately equals the bandgap of Si ($\sim 1\text{eV}$) [30]. So, the same negative E_{ox} corresponds to different negative V_{STS} in NMOS and PMOS transistors and this difference is $\sim 1\text{V}$. In other words, same negative V_{STS} can correspond to a reduced E_{ox} in a accumulated NMOS transistor compared to the E_{ox} in an inverted PMOS transistor [30]. Thus, same negative V_{STS} will cause reduced ΔV_T in NMOS compared to that in PMOS⁴.
- Dopant-activated Si-H dissociation has one to one correlation between N_{IT} and diffusing species H_2 (*i.e.*, for one N_{IT} we have one H_2 molecule). Thus, N_{IT} generation still follows the 1/4 power-law time exponent, predicted by [117] in equation (3.7) and also showed through R-D simulation in [131]. Such 1/4 power-law time exponent is not supported by N_{IT} measurement using recent characterization techniques [129, 137-141].

⁴ However, calculation of V_{STS} for same E_{ox} in practical transistors, having V_T of $\pm 0.4\text{V}$ (positive for NMOS and negative for PMOS), suggests that the voltage difference for same E_{ox} is only 0.2V [124]; which is much less than the χ_{PG} difference between NMOS and PMOS transistors. So the difference among ΔV_T 's for NMOS and PMOS at same negative V_{STS} may not be entirely due to the χ_{PG} difference, as proposed in [30].

- The calculated activation energy $E_{A,IT}$ in this model is ~ 0.36 eV [134], which is inconsistent with experimental observations of ~ 0.1 eV [127-130, 140, 142, 143].

3.5. Diffusing Species in R-D Model

Diffusing species mainly governs the value of power-law time exponent (n) expected from R-D framework. In addition, presence of dispersion in the diffusion mechanism can modify over and above the value set by the diffusing species (see Appendix B). As hydrogen is generated from Si-H bond dissociation, the diffusing species can be: (a) neutral or charged atomic hydrogen (H or H^+), (b) molecular hydrogen (H_2), and (c) both H and H_2 .

3.5.1. Diffusion of Neutral Atomic Hydrogen (H)

Early versions of R-D model considered diffusion of H generated in equation (3.9) [30, 31] (Figure 3.1). As such, N_{IT} generation can be expressed using equation (3.7), with X being replaced by H. Thus H-diffusion based R-D model has $n \sim 1/4$, which supported experimental observations, measured before the year 2003 [31, 144]. Later, presence of $n \sim 1/4$ was conclusively shown to be an artifact of measurement delay or t_{meas} [130, 139, 145].

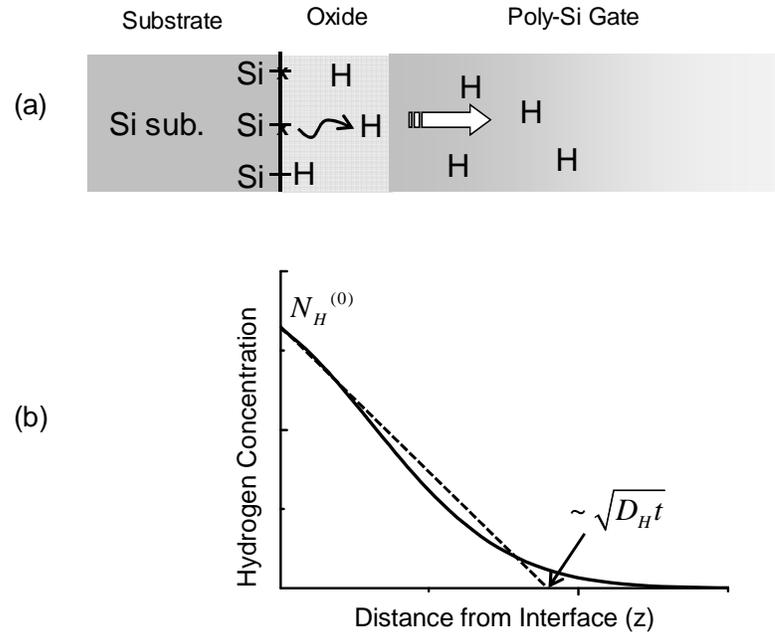


Figure 3.1: (a) Schematic representation of Si-H bond dissociation and resulting diffusion of H through dielectric and poly-Si. (b) Diffusion front at time t reaches $\sqrt{D_H t}$ and the diffusion profile can be approximated using triangle, which is used in obtaining total hydrogen (hence, N_H) concentration in equation (3.6).

3.5.2. Drift of Charged Atomic Hydrogen (H^+)

Though H was used as diffusing species in pre-2003 literature (supporting the experimental observation at that time), first-principle calculation [146, 147] always predicted positively-charged atomic hydrogen or proton (H^+) to be the only stable form of atomic hydrogen at the Si/SiO₂ interface. This motivated researchers [33, 148] to use H^+ as diffusing species, along with dispersive diffusion (see Appendix B.1), for explaining the signatures observed in their experiments. However, the model with dispersive H^+ diffusion was later explained in [130] as an artifact of t_{meas} .

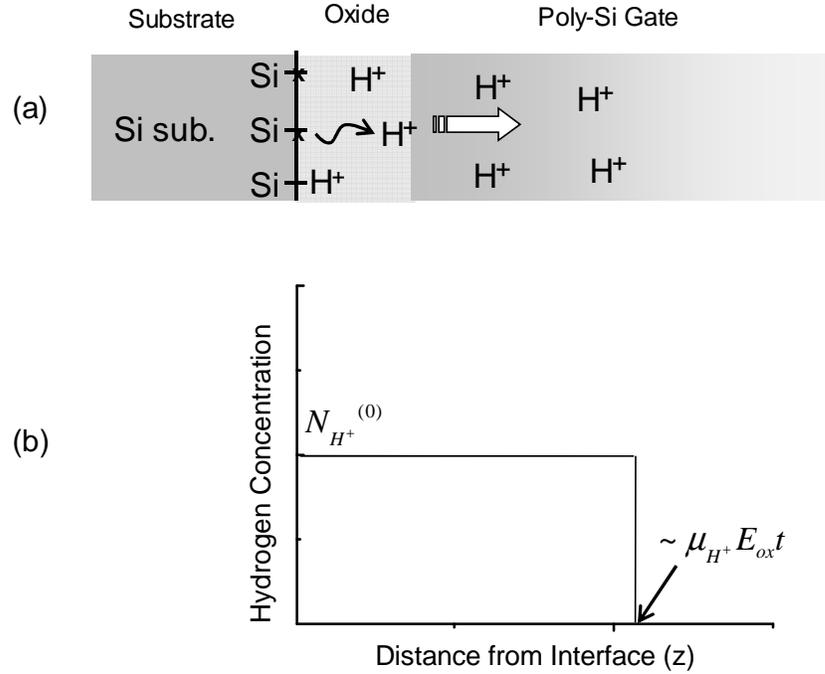
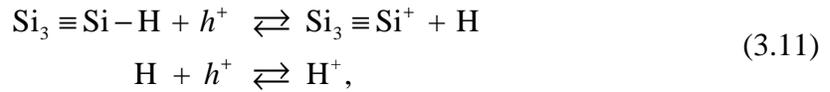


Figure 3.2: (a) Schematic representation of Si-H bond dissociation and resulting drift (dominates over diffusion) of H^+ through dielectric and poly-Si. (b) Drift front at time t reaches $\mu_{H^+} E_{ox} t$ and can be approximated using rectangle [125], which is used to obtain total hydrogen (hence, N_{IT}) concentration in equation (3.15).

The Si-H bond dissociation in such model (Figure 3.2) has the following form:



and the governing differential equation is similar to equation (3.2), except X replaced using H^+ . Hence for $t \gg 0$, *i.e.*, when the surface reaction equation (3.2) reaches quasi-equilibrium, we can write

$$k_F N_0 \equiv k_R N_{H^+}^{(0)} N_{IT}. \quad (3.12)$$

Considering the diffusion/drift of proton to follow an Arrhenius behavior (*i.e.*, governed by hopping of H^+ through single energy level), we can replace equations (3.3) and (3.4) using –

$$\frac{\partial N_{H^+}}{\partial t} = D_{H^+} \frac{\partial^2 N_{H^+}}{\partial z^2} + \mu_{H^+} E_{ox} \frac{\partial N_{H^+}}{\partial z}, \quad (3.13)$$

$$\frac{\partial}{\partial t} \left[\frac{1}{2} \delta N_{H^+}^{(0)} \right] = \frac{\partial N_{IT}}{\partial t} - D_{H^+} \frac{\partial N_{H^+}^{(0)}}{\partial z} - \mu_{H^+} E_{ox} N_{H^+}^{(0)}. \quad (3.14)$$

Now, as drift term will dominate over the diffusion term in equations (3.13) and (3.14), for charged particles like protons, we can estimate the amount of N_{IT} at $t \gg 0$ using [125] –

$$N_{IT} = \int_{z=0}^{z=f(\mu_{H^+}, t)} N_{H^+}(z, t) dz \sim N_{H^+}^{(0)} \mu_{H^+} E_{ox} t, \quad (3.15)$$

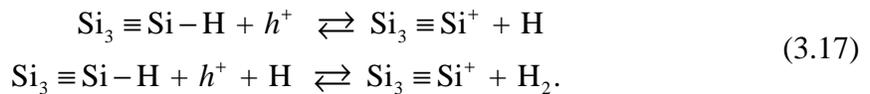
where, $\mu_{H^+} E_{ox}$ is the drift velocity of H^+ , moving away from the interface. Combining equations (3.12) and (3.15), we can write –

$$N_{IT} \sim \left(\frac{k_F N_0}{k_R} \mu_{H^+} E_{ox} \right)^{1/2} t^{1/2}. \quad (3.16)$$

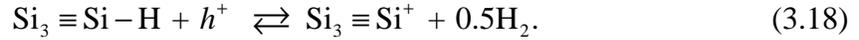
Thus, R-D model with Arrhenius-activated H^+ drift-diffusion predicts $n \sim 1/2$, which is not yet observed in any N_{IT} generation experiments, unless some dispersion (*i.e.*, drift-diffusion is governed by hopping through multiple trap energy levels [149-153]) is considered within the framework for reducing n [33, 148] (see Appendix B for details).

3.5.3. Diffusion of Molecular Hydrogen (H_2)

Direct dissociation of Si-H bond (section 3.4.1) creates atomic hydrogen (H), which can react again with another Si-H bond to form molecular hydrogen (H_2) through the following chemical reaction:



Overall reaction can also be expressed as,



and the governing differential equation for this chemical reaction will take the following form:

$$\frac{\partial N_{IT}}{\partial t} = k_F (N_0 - N_{IT}) - k_R N_{IT} \sqrt{N_{H_2}^{(0)}}. \quad (3.19)$$

Considering the Arrhenius diffusion of H_2 , we replace equations (3.3) and (3.4) using –

$$\frac{\partial N_{H_2}}{\partial t} = D_{H_2} \frac{\partial^2 N_{H_2}}{\partial z^2}, \quad (3.20)$$

$$\frac{\partial}{\partial t} [\delta N_{H_2}^{(0)}] = \frac{\partial N_{IT}}{\partial t} - 2D_{H_2} \frac{\partial N_{H_2}^{(0)}}{\partial z}. \quad (3.21)$$

Following analysis similar to the one we used in original R-D model (see section 3.3; equations (3.5) - (3.7)), we have –

$$N_{IT} \sim \left(\frac{k_F N_0}{k_R} \right)^{2/3} (D_{H_2} t)^{1/6}. \quad (3.22)$$

So R-D model with H_2 diffusion (Figure 3.3) predicts a power-law time exponent, $n \sim 1/6$, which is more consistent with recent ΔN_{IT} measurements (or ΔV_T measurements) on type-I transistors [119, 129, 138]⁵. Moreover, considering k_F , k_R , and D_{H_2} are Arrhenius activated, equation (3.22) suggests the following expression for activation energy associated with N_{IT} formation:

$$E_{A,IT} = \frac{2}{3} (E_{A,F} - E_{A,R} - aE_{ox}) + \frac{1}{6} E_{A,H_2}, \quad (3.23)$$

⁵ Here, type-I transistors are defined as those transistors, where the NBTI-induced ΔV_T is dominated by N_{IT} generation [119, 129, 154]. In addition, there are also type-II transistors, where hole trapping (detrapping) to (from) pre-existing defects also plays significant role in NBTI degradation [129, 141, 155]. See section 6.7.5 for further discussion.

where, $E_{A,F}$, $E_{A,R}$, and E_{A,H_2} are activation energies for k_F , k_R , and D_{H_2} , respectively (see section 3.7 for detailed discussion on aE_{ox} term). Now, generalized scaling (see Appendix A.4 for a discussion on generalized scaling) consistently suggested $E_{A,IT} \sim E_{A,H_2}/6$ [30, 127, 130, 156], indicating contribution from first term in equation (3.23) is negligible. Hence, extracted $E_{A,H_2} \sim 6E_{A,IT}$ has a value of $\sim 0.5-0.6\text{eV}$ [30, 127, 130, 156], which is consistent with the E_{A,H_2} reported in [157, 158]. Hence, N_{IT} formation at long stress time can only be explained using R-D model with H_2 diffusion.

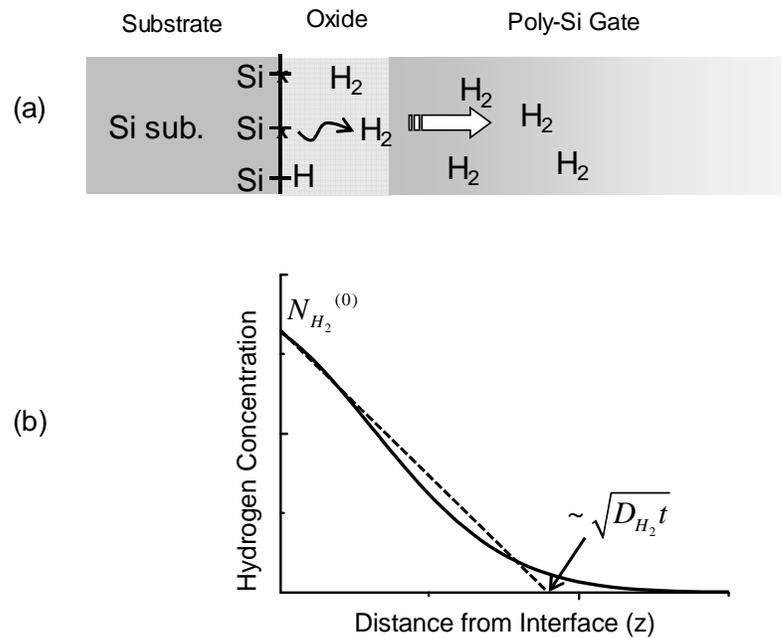


Figure 3.3: (a) Schematic representation of Si-H bond dissociation and resulting diffusion of H_2 through dielectric and poly-Si. (b) Diffusion front at time t reaches $\sqrt{D_{H_2}t}$ and the diffusion profile can be approximated using triangle, which is used in obtaining total hydrogen (hence, N_{IT}) concentration in deriving equation (3.22).

3.5.4. H- H_2 Diffusion with Direct H- H_2 Conversion

Chemical reaction of equation (3.17) requires generated H from first Si-H dissociation to dissociate another Si-H bond, thus form H_2 . As H is associated with only

1% (see section 2.2) of total dangling Si- bonds at the interface (*i.e.*, sources of H are sparsely distributed at the oxide/silicon interface), it will be difficult for H to complete the second chemical reaction of equation (3.17) to form H₂ molecule and then diffuse away from the interface. Rather H, from the dissociation of two Si-H bonds can dimerize to form H₂ and resultant hydrogen species (*i.e.*, both H and H₂) can diffuse into the oxide/gate structure (Figure 3.4). Although neutral charge state may not be a stable form of atomic hydrogen, its transient formation is indeed possible [159]. Therefore, diffusion of both H and H₂, as well as, H \leftrightarrow H₂ conversion are explicitly incorporated in the generalized R-D framework by the following equations [119, 160, 161]:

$$\frac{dN_{IT}}{dt} = k_F (N_0 - N_{IT}) - k_R N_{IT} N_H^{(0)}, \quad (3.24)$$

$$\begin{aligned} \frac{\delta}{2} \frac{dN_H^{(0)}}{dt} &= D_H \frac{dN_H^{(0)}}{dz} + \frac{dN_{IT}}{dt} - \delta k_H [N_H^{(0)}]^2 + \delta k_{H_2} N_{H_2}^{(0)}, \\ \frac{\delta}{2} \frac{dN_{H_2}^{(0)}}{dt} &= D_{H_2} \frac{dN_{H_2}^{(0)}}{dz} + \frac{\delta}{2} k_H [N_H^{(0)}]^2 - \frac{\delta}{2} k_{H_2} N_{H_2}^{(0)}, \end{aligned} \quad (3.25)$$

$$\begin{aligned} \frac{dN_H}{dt} &= D_H \frac{d^2 N_H}{dz^2} - k_H N_H^2 + k_{H_2} N_{H_2}, \\ \frac{dN_{H_2}}{dt} &= D_{H_2} \frac{d^2 N_{H_2}}{dz^2} + \frac{1}{2} k_H N_H^2 - \frac{1}{2} k_{H_2} N_{H_2}. \end{aligned} \quad (3.26)$$

Equation (3.24) represents passivation/de-passivation effects of Si-H bond, where, k_F , k_R , N_0 , N_{IT} , $N_H^{(0)}$ are defined as Si-H bond-breaking rate, Si-H bond-annealing rate, initial bond density available before stress, interface defect density and atomic hydrogen density at the Si/dielectric interface, respectively. Equation (3.25) corresponds to the conservation of fluxes of diffusing hydrogen species (H and H₂) near the interface (along x axis), whereas equation (3.26) describes diffusion (along x axis) of H and H₂. $k_H N_H^2$ and $k_{H_2} N_{H_2}$ terms in equation (3.25) incorporates the H-H₂ conversion within the generalized R-D framework. Among the symbols used in equations (3.25) and (3.26), k_H , k_{H_2} represent generation and dissociation rates of H₂; D_H , D_{H_2} represent diffusion coefficients for H and H₂; N_H , N_{H_2} represent the concentration of atomic and molecular hydrogen and δ is the interfacial thickness (~ 1.5 Å [69]).

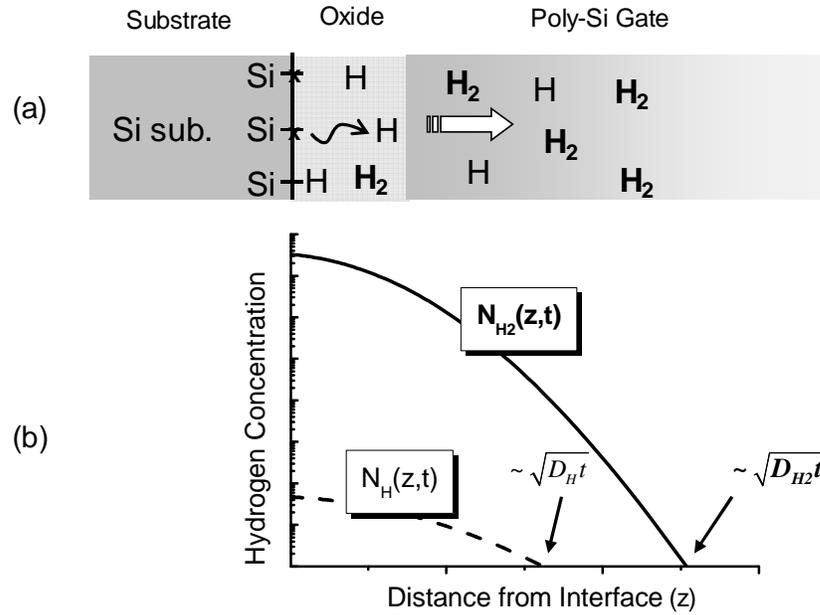


Figure 3.4: (a) Schematic representation of Si-H bond dissociation and resulting diffusion of hydrogen species (H and H₂) through dielectric and poly-Si [143] during NBTI stress. (b) At long stress time, diffusion of H₂ (concentration N_{H_2}) dominates and concentration of H (N_H) is small.

The solid lines in Figure 3.5a show that the numerical solution of equations (3.24)-(3.26) interprets the data from [140] very well. The solution reflects the dominance of $H \leftrightarrow H_2$ conversion during short-term stress ($< 10s$) extends the transition between reaction-limited region ($n \sim 1$ [30, 142, 162]) and H₂ diffusion limited region ($n \sim 1/6$ [34, 142]) by several orders of time scale, enabling intuitive interpretation of short-term NBTI degradation. Figure 3.6 shows the time-dependent formation of H and H₂ at the Si/SiO₂ interface, and explicitly supports the notion that H – H₂ transformation governs NBTI generation at short time scales. Later, when $H \leftrightarrow H_2$ conversion reaches quasi-equilibrium (see Figure 3.6), dominance of H₂ diffusion results in $n \sim 1/6$. We use field-dependence of k_f (see section 3.7) to interpret the voltage-dependent data in Figure 3.5.

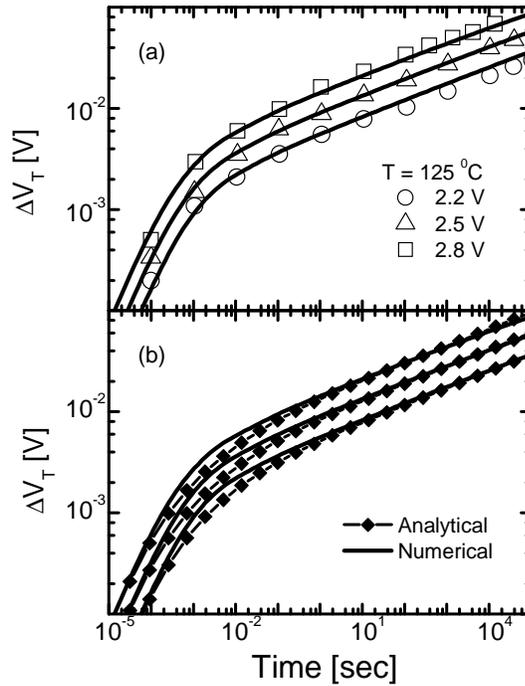


Figure 3.5: (a) Simulation of generalized R-D model (H-H₂ system) explains the experimental trends [140] at any stress time. R-D model parameters (consistent with literature [142, 143]): $k_F = 6 \times 10^{-3} E_c \exp(0.65 E_{ox}) \text{ sec}^{-1}$ (E_c : oxide electric field due to mobile carriers, E_{ox} : total oxide electric field), $N_0 = 5 \times 10^{12} \text{ cm}^{-2}$, $k_R = 3 \times 10^{-9} \text{ cm}^3 \text{ sec}^{-1}$, $D_H = 3 \times 10^{-13} \text{ cm}^2 \text{ sec}^{-1}$, $D_{H_2} = 1.8 \times 10^{-14} \text{ cm}^2 \text{ sec}^{-1}$, $k_H = 1.4 \times 10^{-3} \text{ cm}^3 \text{ sec}^{-1}$, $k_{H_2} = 95.4 \text{ sec}^{-1}$. (b) Comparison of numerical solution (lines) of equations (3.24)-(3.26) with analytical solution of equation (3.30) (lines with symbols).

Analytical Solution of H-H₂ R-D Model

Although the numerical model, *i.e.*, equations (3.24)-(3.26), interprets the experimental data adequately, an analytical model may be easier to use and provide additional insight. In earlier studies, analytical expressions for R-D model were provided in reaction regime [30, 142, 162], H₂ diffusion regime [34, 142], and H-H₂ transition regime [160]. Here, *we provide a complete analytical solution*, which captures the

dynamics of H-H₂ R-D model in all the three regimes (see equation (3.30)). Assuming $N_0 \gg N_{IT}$ [30, 142] and $dN_{IT}/dt \sim N_{IT}/t$ [162], equation (3.24) simplifies to,

$$N_H^{(0)} = \frac{k_F N_0 - N_{IT}/t}{k_R N_{IT}}. \quad (3.27)$$

Moreover, the numerical solutions indicate that for continuous NBTI stress, $dN_H^{(0)}/dt$ and diffusion of H is negligible at all stress time, so that H₂ diffusion part in equation (3.25) reduces to,

$$\frac{N_{IT}}{t} = \delta k_H N_H^{(0)2} - \delta k_{H2} N_{H2}^{(0)}. \quad (3.28)$$

As stated earlier, H-H₂ diffusion reaches steady state at long stress time, when H₂ diffusion dominates. Under that condition, simulation shows –

$$N_{IT} \approx N_{H2}^{(0)} \sqrt{6D_{H2}t}. \quad (3.29)$$

Equation (3.29) requires that the extent of diffusion profile is larger than $\sqrt{D_H t}$, commonly used for approximating the complementary error function solution of the diffusion equation [30, 125, 126]; which is more consistent with the use of $\sqrt{16D_H t/\pi}$ in [153, 163]. Now, by eliminating $N_H^{(0)}$ and $N_{H2}^{(0)}$ from equations (3.27)-(3.29), we have –

$$\frac{N_{IT}}{t} - \frac{\delta k_H (k_F N_0 - N_{IT}/t)^2}{k_R^2 N_{IT}^2} + \frac{\delta k_{H2} N_{IT}}{\sqrt{6D_{H2}t}} = 0. \quad (3.30)$$

Equation (3.30) is the analytical solution of H-H₂ R-D model, presented in equations (3.24)-(3.26). A comparison of the numerical and analytical solutions, presented in Figure 3.5b and Figure 3.6, shows excellent matching. Equation (3.30) reduces to (i) reaction-limited solution [$N_{IT} = k_F N_0 t$], when first and third term becomes negligible (compared to second term) at very short stress time, (ii) it reduces to H₂ diffusion limited solution [$N_{IT} = (k_H/k_{H2})^{1/3} (k_F N_0/k_R)^{2/3} (6D_{H2}t)^{1/6}$], when N_{IT}/t becomes negligible at long stress time, and (iii) at intermediate stress time, when diffusive term (third term) in equation (3.30) is negligible and $k_F N_0 \gg N_{IT}/t$, yields a solution of [$N_{IT} = (k_F N_0/k_R)^{2/3} (\delta k_H t)^{1/3}$]; which is a signature of dominant H-H₂ conversion [160]. Indeed,

Figure 3.5b suggests that the analytical solution of equation (3.30) is good enough to reproduce the numerical solution of equations (3.24)-(3.26) over ~ 10 decades in stress time (maximum error of 20% in the transition region)! Moreover, very good agreement between numerical solutions (solid lines) and analytic solutions (lines with symbols) in Figure 3.6 suggests that equations (3.27) and (3.28) capture the dynamics of $N_H(x=0,t)$ and $N_{H_2}(x=0,t)$ very well. The inaccuracy of N_{H_2} at very early times is inconsequential, because N_{IT} generation in this time-scale is dictated by generation and transformation of N_H .

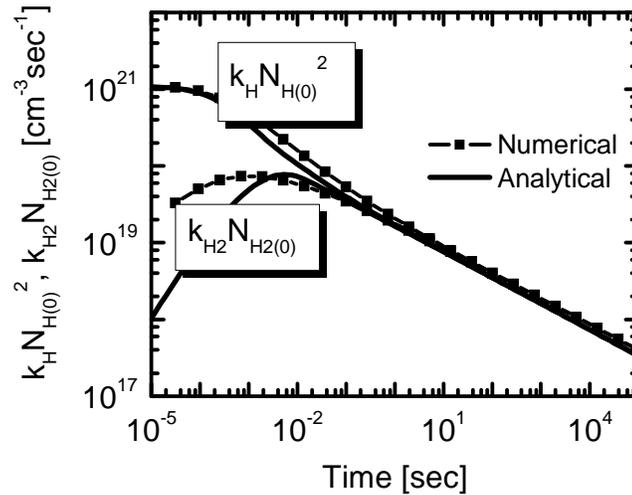
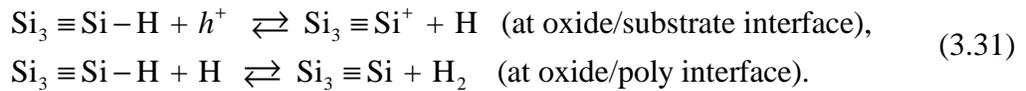


Figure 3.6: Comparison of numerical results (lines) from equations (3.24)-(3.26), with analytical results (lines with symbols) for $k_H N_{H(0)}^2$ and $k_{H_2} N_{H_2(0)}$. Although analytical $N_{H_2(0)}$ prediction is incorrect at short stress time (whereas, $N_{H(0)}$ prediction is excellent), dominance of $k_H N_{H(0)}^2$ over $k_{H_2} N_{H_2(0)}$, still makes the analytical solution of N_{IT} appropriate for all stress time (see Figure 3.5b).

3.5.5. H-H₂ Diffusion with H₂ Conversion at Poly/Oxide Interface

Generation of H₂ at $t \gg 0$ (hence, $n \sim 1/6$) can also result from reactions involving both the oxide/gate and oxide/substrate interfaces [131, 143]. Here, the creation of H through Si-H bond dissociation at the oxide/substrate interface is followed by a second reaction between generated H and Si-H bond at the oxide/gate interface. Thus, contrary to equation (3.17), here the second Si-H bond dissociation is considered at the oxide/poly interface, rather than at the same oxide/substrate interface. The overall chemical reaction takes the following form:



Hence, the chemical process in equation (3.31) indicates presence of dangling bonds at the oxide/poly interface, which has been experimentally proven through low-voltage stress induced leakage current (LV-SILC) experiment in [143, 164]. Resultant H/H₂ diffuses away from poly/dielectric interface. Overall chemical reaction at the two interfaces are provided in equation (3.31), which can be expressed using the following differential equations:

$$\frac{dN_{IT(s)}}{dt} = k_{F(s)} (N_{0(s)} - N_{IT(s)}) - k_{R(s)} N_{IT(s)} N_H^{(s)} \quad (3.32)$$

$$\frac{dN_{IT(p)}}{dt} = k_{F(p)} (N_{0(p)} - N_{IT(p)}) N_H^{(p)} - k_{R(p)} N_{IT(p)} N_{H_2}^{(p)} \quad (3.33)$$

where, variables with (s) and (p) represent the quantities at the oxide/substrate and oxide/poly interfaces, respectively. This H-H₂ R-D model has similar long-term stress-phase signatures (*i.e.*, time exponent of $\sim 1/6$ [143, 164]) as the H-H₂ R-D model having direct H-H₂ conversion.

3.5.6. Diffusing Species in R-D Model: Summary

Based on the discussion in sections 3.4-3.5, we can generate a table of power-law time exponents (n), shown in Table 3.1. Such variation in n mainly results from the proposed variations in hydrogen species undergoing drift or diffusion, as well as the mechanism by which these hydrogen species are generated. Recent measurement of $n \sim 1/6$ for N_{IT} generation (which is dominant in type-I transistors) [119, 128-130, 138, 154, 165, 166] strongly suggests the existence of H_2 diffusion at long stress time. This is further supported by an observed $E_{A,IT} \sim 0.1\text{eV}$, again expected from H_2 diffusion [118, 119, 129].

Table 3.1: Effect of diffusion species and its (i.e. diffusing species) generation mechanism on N_{IT} formation power-law time exponent, n

| Sections | Reaction | Mobile Species | n ($t \gg 0$) |
|--------------|--|----------------|----------------------|
| 3.4.1, 3.5.1 | $Si_3 \equiv Si-H + h^+ \rightleftharpoons Si_3 \equiv Si^+ + H$ | H | 1/4 |
| 3.4.1, 3.5.2 | $Si_3 \equiv Si-H + h^+ \rightleftharpoons Si_3 \equiv Si^+ + H$ $H + h^+ \rightleftharpoons H^+$ | H^+ | 1/2 |
| 3.4.1, 3.5.3 | $Si_3 \equiv Si-H + h^+ \rightleftharpoons Si_3 \equiv Si^+ + H$ $Si_3 \equiv Si-H + h^+ + H \rightleftharpoons Si_3 \equiv Si^+ + H_2$ | H_2 | 1/6 |
| 3.4.1, 3.5.4 | $Si_3 \equiv Si-H + h^+ \rightleftharpoons Si_3 \equiv Si^+ + H$ $H + H \rightleftharpoons H_2$ | H, H_2 | 1/6 |
| 3.4.1, 3.5.5 | Oxide/Subs: $Si_3 \equiv Si-H + h^+ \rightleftharpoons Si_3 \equiv Si^+ + H$ Oxide/Poly: $Si_3 \equiv Si-H + H \rightleftharpoons Si_3 \equiv Si + H_2$ | H, H_2 | 1/6 |
| 3.4.2 | $Si_3 \equiv Si-H + H^+ \rightleftharpoons Si_3 \equiv Si^+ + H_2$ | H_2 | 1/4 |

3.6. Recovery (or Reverse-Annealing) of Interface Defects

In sections 3.2-3.5, we have shown how R-D model can explain the time dynamics of interface defect generation in MOS transistors. However, one important feature of interfacial defects is its reverse-annealing, once the stress voltage is removed [131, 139, 140, 167-169]. Such existence of N_{IT} recovery makes NBTI different from other transistor reliability concerns, such as TDDB (where, the generated defects do not recover [51-57]) and HCI (where, recovery is negligible [112-114]). The R-D model discussed above can also capture the dynamics of N_{IT} recovery. In R-D model, N_{IT} relaxation is attributed to the reverse-annealing of dangling Si- bonds by the near-interface hydrogen species. These hydrogen species (needed for reverse-annealing) were generated during NBTI stress-phase, which starts to diffuse back towards the interface during the relaxation-phase. In the following sections, we study the current status of different versions of H-H₂ R-D model, which best describes the dynamics of N_{IT} generation, in explaining NBTI relaxation experiments.

3.6.1. Analytical Formalism

The following analysis of N_{IT} recovery was originally developed in [167], and is in general valid for any R-D formalism, irrespective of whether H or H₂ is the dominant diffusing species. In an R-D formalism, N_{IT} at the end of stress time ($t = t_{STS}$) can be expressed as –

$$N_{IT}(t_{STS}) \approx N_X^{(0)} \sqrt{D_X t_{STS}}, \quad (3.34)$$

where, $N_X^{(0)}$ represents the concentration of hydrogen (H or H₂) species near the interface and D_X represents the diffusion co-efficient for the hydrogen species under consideration. At time t_{REC} after the end of stress, a part of hydrogen profile, present near the interface (hatched region in Figure 3.7), will diffuse back and anneal (re-passivate) N_{IT} . This amount of annealed N_{IT} (represented here as N_{IT}^*) can be expressed as –

$$N_{IT}^*(t_{STS} + t_{REC}) \approx N_X^* \sqrt{\xi D_X t_{REC}}, \quad (3.35)$$

where, $\xi = 0.5$ for the two-sided diffusion present in this case. Next, we approximate the amount of N_{IT} at t_{REC} using –

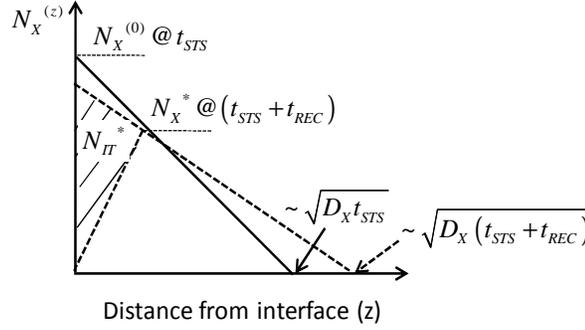


Figure 3.7: Approximate hydrogen profile at the start ($t = t_{STS}$) and after a recovery time of t_{REC} (*i.e.*, at $t = t_{STS} + t_{REC}$) of N_{IT} relaxation. See text for the definition of other parameters.

$$\begin{aligned} N_{IT}(t_{STS} + t_{REC}) &= N_{IT}(t_{STS}) - N_{IT}^*(t_{STS} + t_{REC}) \\ &\approx N_X^* \sqrt{D_X (t_{STS} + t_{REC})}. \end{aligned} \quad (3.36)$$

Using equations (3.34)-(3.36), one can show that [126, 167] –

$$N_{IT}(t_{STS} + t_{REC}) \approx N_{IT}(t_{STS}) \left/ \left[1 + \sqrt{\frac{\xi t_{REC}}{t_{STS} + t_{REC}}} \right] \right. \quad (3.37)$$

As shown in [167], one requires $\xi \sim 0.58$ to match up with detailed simulation. Now, the following features of N_{IT} recovery (within the R-D framework) are evident from equation (3.37) –

- R-D relaxation feature-1: If we define the start-time of N_{IT} recovery $t_{NIT,start}$ as the time, when $N_{IT}(t_{STS} + t_{REC}) / N_{IT}(t_{STS}) \sim 95\%$ (*i.e.*, $t_{NIT,start}$ corresponds to 5% N_{IT} recovery), then according to equation (3.37), $t_{NIT,start} \sim 0.005 t_{STS}$.
- R-D relaxation feature-2: At long $t_{REC} \gg t_{STS}$, N_{IT} is expected recover completely (*i.e.*, $N_{IT} \sim 0$ for $t_{REC} \rightarrow \infty$).

In the following sections (3.6.2-3.6.4), we show that the two relaxation features mentioned above are the general property of any versions of H-H₂ R-D model, except $t_{NIT,start}$ is somewhat smaller for poly H-H₂ R-D model. Later, in section 3.6.5, we compare the N_{IT} recovery predicted by the H-H₂ R-D model with the NBTI relaxation experiments and thus identify the attributes of NBTI relaxation experiments that are still to be captured.

3.6.2. N_{IT} Recovery: Direct H-H₂ R-D Model

Direct H-H₂ R-D model (sections 3.5.4) considers depassivation of Si-H bonds at the Si/dielectric interface and resultant diffusion of hydrogen species into gate dielectric and poly-Si, as the main cause for N_{IT} generation. Here, we study N_{IT} recovery in a transistor structure, having 1.4nm gate dielectric, for the simulation. Before the initiation of the N_{IT} recovery phase, the transistor is stressed at $V_{STS} = -2.3V$, $T = 125$ °C for $t_{STS} = 1000$ sec. During the recovery phase (*i.e.*, when gate bias V_{STS} is reduced to V_{REC}), hydrogen species starts to diffuse back towards the interface and, therefore, anneals N_{IT} . Figure 3.8a shows such N_{IT} recovery for two different V_{REC} , which suggests $t_{NIT,start}$ of ~ 1 sec ($= 10^{-3} t_{STS}$), which is similar to the one predicted using analytical calculations in equation (3.37). Moreover, Figure 3.8b shows a trend towards complete relaxation of N_{IT} , which is also expected from equation (3.37).

To understand why N_{IT} relaxation in direct H-H₂ R-D simulation starts at ~ 1 sec, we plot the H₂ profile for the same transistor (used in Figure 3.8) at the end of $t_{STS} = 1000$ sec. Figure 3.9 indicates that most of the hydrogen has diffused beyond the oxide and its integrated density is predominantly within the poly-Si. The percentage of hydrogen within different distances from the interface is plotted in Figure 3.9b (*e.g.*, only 2% H₂ are within oxide, 6% H₂ within 4nm away from the interface, 50% H₂ within 35nm away from the interface). Therefore, in order to have $\sim 6\%$ recovery, one needs to wait $t_{REC} = z^2 / 4D = 2.2$ sec (for $D_{H_2} \sim 1.8 \times 10^{-14}$ cm²-sec⁻¹ [143] used in the simulation). Thus, it is not surprising that direct H-H₂ R-D model predicts $t_{NIT,start} \sim 1$ sec for a stress time of 1000 sec (for any recovery bias), as shown in Figure 3.8. Moreover, since most of the

hydrogen is within the poly-Si structure at the end of stress, enhanced diffusion within the oxide is not expected to reduce $t_{NIT,start}$ significantly.

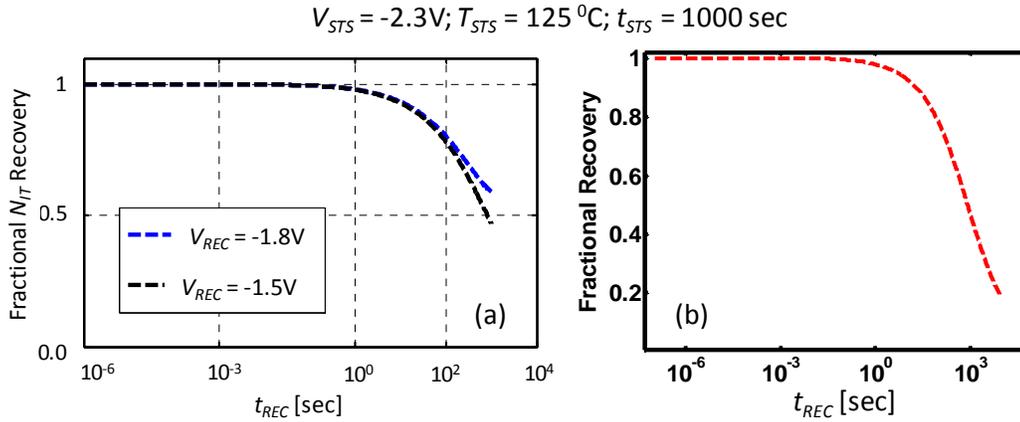


Figure 3.8: (a) Simulation of N_{IT} recovery using direct H-H₂ R-D model for a transistor having 1.4 nm gate dielectric. N_{IT} is generated at $V_{STS} = -2.3V$ for $t_{STS} = 1000\text{sec}$ at $125\text{ }^{\circ}C$, before it is allowed to recover at two different V_{REC} . Significant amount of relaxation starts around $t_{NIT,start} \sim 1\text{sec}$. (b) A trend towards complete N_{IT} recovery is observed at $V_{REC} = 0$.

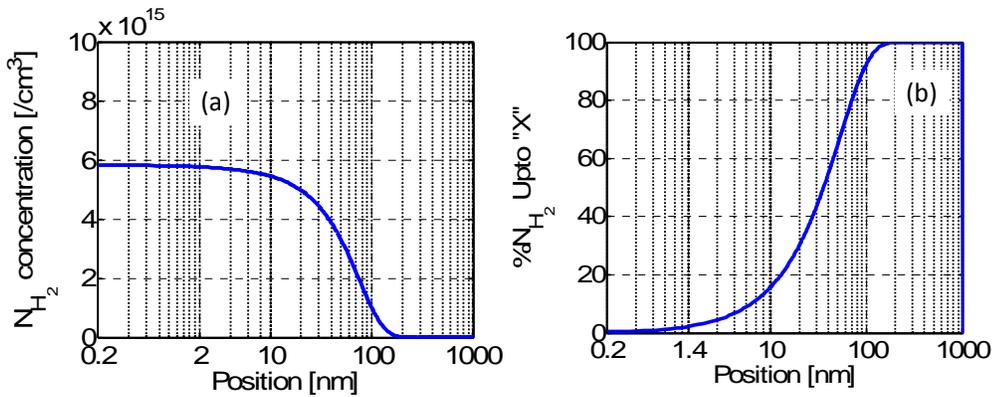


Figure 3.9: (a) H₂ Profile at the end of stress for a transistor having 1.4nm gate dielectric after $t_{STS} = 1000\text{ sec}$. (b) Percentage of H₂ within distance 0 to z , for different z (*i.e.*, distance from the interface).

3.6.3. N_{IT} Recovery: Poly H-H₂ R-D Model

This form of H-H₂ R-D model (poly H-H₂ R-D Model) was first proposed in [31, 131] and is elaborated in section 3.5.5. Here, no explicit H-H₂ conversion is employed within the R-D framework. Rather, the atomic hydrogen (H) generated from the substrate/oxide interface diffuses towards oxide/poly interface and reacts there with the Si-H bond. As summarized in Table 3.1, poly H-H₂ R-D model has similar long-term stress-phase signatures (*i.e.*, time exponent of $\sim 1/6$) as the direct H-H₂ R-D model. However, when oxide is considered to have enhanced diffusion compared to poly [143], there is an intermediate time exponent of $\sim 1/4$ (similar to R-D model with H diffusion, see section 3.5.1), representing initial H diffusion phase within oxide (Figure 3.10a). Figure 3.10b plots the percentage of hydrogen within different distances from the interface for a transistor having 2nm gate oxide stressed up to $t_{STS} = 10^5$ sec (*e.g.*, about $>10\%$ H₂ are within oxide, 50% H₂ within 10nm from the interface).

For N_{IT} recovery simulation, Figure 3.11a suggests $t_{NIT,start} \sim 10^{-4} t_{STS}$, which is one order of magnitude smaller than the same for direct H-H₂ R-D model (which has $t_{NIT,start} \sim 10^{-3} t_{STS}$, see Figure 3.8). This is because poly H-H₂ model has H-H₂ conversion only at the oxide/poly interface, whereas direct H-H₂ model has H-H₂ conversion over the entire oxide/poly region. Therefore, the amount of H₂ remained within the oxide (at the end of t_{STS}) for poly H-H₂ model is much larger than the amount of H₂ within the oxide for direct H-H₂ model. As a result, during recovery phase of the poly H-H₂ model, the H₂ within the oxide diffuses quickly towards the oxide/poly interface (as the hydrogen diffusion coefficient within oxide is larger compared to that in poly-Si [143]) and converts into H. The generated H preferentially diffuses back towards the substrate/oxide interface, again due to the existence of higher hydrogen diffusion coefficient for the oxide. As a result, $t_{NIT,start}$ for poly H-H₂ R-D model is smaller compared to $t_{NIT,start}$ in direct H-H₂ R-D model.

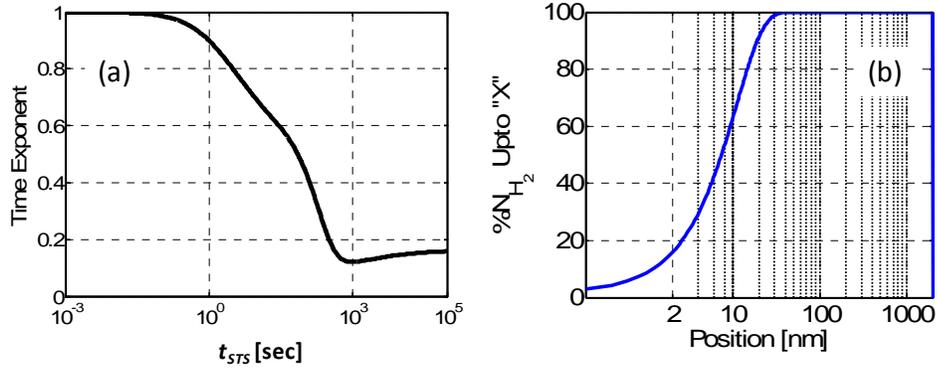


Figure 3.10: (a) Time exponent variation for poly H-H₂ R-D model for a transistor having 2nm gate oxide. (b) Percentage of H₂ within distance 0 to z , for different z (*i.e.*, distance from the interface).

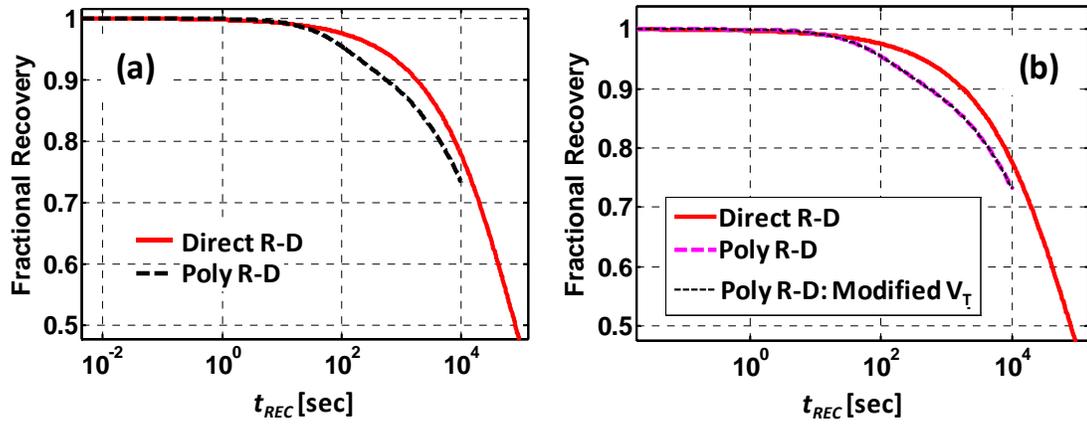


Figure 3.11: (a) Comparison of N_{IT} recovery simulations for the direct H-H₂ R-D model (solid line) and the poly H-H₂ R-D model (dashed line). $t_{NIT,start}$ for poly H-H₂ model is order of magnitude smaller than the one for direct H-H₂ model. (b) There is no change in $t_{NIT,start}$ when V_T calculation incorporates defects at the oxide/poly interface, in addition to the defects at the substrate/oxide interface (poly R-D: Modified V_T).

Moreover, one might anticipate further improvements in terms of $t_{NIT,start}$ by including the effect of poly interface (which is closer to the diffusion front) in the V_T calculation; *i.e.*, by using $\Delta V_T = q\Delta N_{IT(s)}/C_{ox} + q\Delta N_{IT(p)}/C_{poly}$, where C_{poly} is the poly-

depletion capacitance. However, Figure 3.11b shows that such improvement is negligible, because most of the hydrogen is far away from the oxide/poly interface. For example, it requires at least $t = z^2 / 4D_{H_2} = 1000 \text{ sec}$ (for $D_{H_2, \text{poly}} \sim 1.8 \times 10^{-17} \text{ cm}^2\text{-sec}^{-1}$ used in the simulation) for hydrogen to reach oxide/poly interface from $z \sim 2 \text{ nm}$. Here, the value for $D_{H_2, \text{poly}}$ is quite small to maintain sufficient H_2 within oxide. If that is changed, one will indeed have different % H_2 profile (compared to the one shown in Figure 3.10b). However, there will be no change in the fractional recovery plot for the poly H- H_2 R-D model (Figure 3.11b).

Thus, we conclude that poly H- H_2 R-D model (with H_2 generation at the poly interface) shows an order of magnitude improvements in terms of $t_{NIT, \text{start}}$ over direct H- H_2 R-D model. However, incorporation of poly interface in ΔV_T calculation does not change the fractional recovery picture (*i.e.*, both interfaces are observed to recover within the same timeframe).

3.6.4. N_{IT} Recovery: Combined (Direct+Poly) H- H_2 R-D Model

The model for H to H_2 conversion at the poly-interface (discussed in sections 3.5.5 and 3.6.3) avoids any direct conversion between H and H_2 (discussed in sections 3.5.4 and 3.6.2). Here, we consider the effect of combining both the H_2 generation mechanisms, thus merging the two models (direct H- H_2 R-D and poly H- H_2 R-D) together. This combined H- H_2 model affects the stress phase by inserting discontinuity (inconsistent with NBTI stress phase experiments) at the point when H/ H_2 front reaches (and gets absorbed at) the poly interface to initiate the reaction. Nonetheless, it goes back to 1/6 at long stress time, when diffusion front enters poly region (Figure 3.12a).

During recovery phase, N_{IT} anneals in an uncontrollable manner (Figure 3.12b), which can be explained by considering the followings:

- As we have explicit H- H_2 conversion within the framework, when the forward reaction at the substrate interface stops (for $V_{REC} = 0$), it creates a non-equilibrium state for all other reactions in the system.

- The H (although small) within the oxide has higher diffusion coefficient and moves back to substrate interface (initiating negligible recovery). This creates non-equilibrium between H and H₂, thus initiating H₂ to H conversion at the poly interface. Moreover, the non-equilibrium between H and H₂ at poly interface also creates non-equilibrium for the reaction (SiH + H ↔ Si + H₂) at the poly interface. So, there is a feedback loop happening between these two reactions at the poly interface (*i.e.*, H ↔ H₂ and SiH + H ↔ Si + H₂). Overall, we have a steady supply of H at the poly interface, a part of which continuously diffuses back to substrate interface causing complete N_{IT} relaxation. To verify this, we reduce the rate of either or both of the H ↔ H₂ and SiH + H ↔ Si + H₂ reactions during relaxation phase. This is observed to slow down the substrate interface recovery significantly.

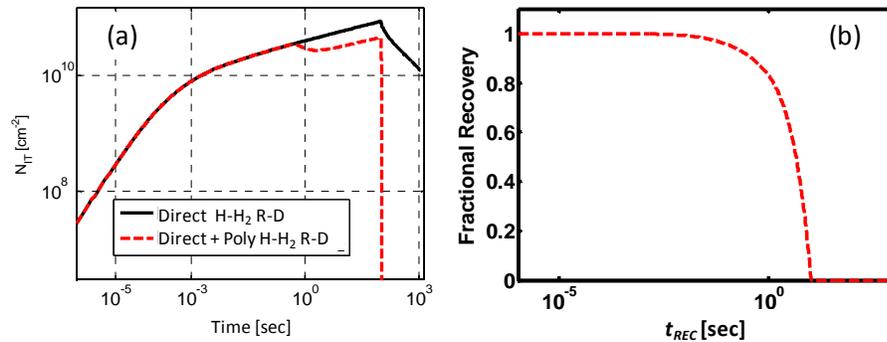


Figure 3.12: (a) There is discontinuity in the stress phase simulation for the combined (direct + poly) H-H₂ R-D model (compared to direct H-H₂ and poly H-H₂ R-D model). (b) Here, the relaxation also happens very quickly compared to the other H-H₂ version of R-D model.

Moreover, incorporation of bond dispersion at both substrate and poly interface (in k_f and k_r terms) was expected to slow down the relaxation at substrate interface. However, bond dispersion only becomes effective, when the supply of hydrogen is constant, *i.e.*, k_R is the limiting factor in N_{IT} recovery. Since for any R-D model, hydrogen diffusion is the main limiting factor for N_{IT} recovery, incorporation of bond dispersion in combined H-H₂

model show insignificant improvement. Therefore, the combined H-H₂ R-D model has smaller $t_{NIT,start}$ compared to the direct H-H₂ and poly H-H₂ model.

3.6.5. N_{IT} Recovery: R-D Theory vs. NBTI Experiment

Based on the analysis in sections 3.6.2-3.6.4, H-H₂ R-D theory, whose stress-phase time dynamics are quite consistent with NBTI experiment, suggests $t_{NIT,start}$ of $\sim 10^{-4}t_{STS}$ for poly H-H₂ model⁶. However, NBTI relaxation experiment suggests a start of NBTI relaxation (designated using $t_{REC,start}$) much less than $10^{-4}t_{STS}$ (Figure 3.13). This indicates that N_{IT} recovery is not the only component of NBTI-induced ΔV_T relaxation. As such, NBTI measurements require appropriate correction before comparing with R-D theory. As discussed in section 6.7, consideration of valence band electron trapping for the ultra-fast V_T experiment (see section 6.7.2) and hole trapping into oxide defects both for ultra-fast V_T and OTF- $I_{D,lin}$ experiments (section 6.7.5) can significantly reduce the gap between $t_{NIT,start} \sim 10^{-4}t_{STS}$ and $t_{REC,start} \sim 10^{-10}t_{STS}$ for Figure 3.13a and $t_{REC,start} \sim (10^{-9}-10^{-6})t_{STS}$ for Figure 3.13b.

Another prediction R-D model is complete N_{IT} recovery for $t_{REC} \rightarrow \infty$. However, some of the NBTI relaxation experiments (Figure 3.13a and Figure 3.14) indicate the presence of non-zero N_{IT} recovery at larger t_{REC} – a feature which requires careful study.

⁶ Although, combined H-H₂ model has the smallest $t_{NIT,start}$, its stress-phase signatures (Figure 3.12a) and the existence of very fast N_{IT} recovery (Figure 3.12b) is inconsistent with NBTI experiments.

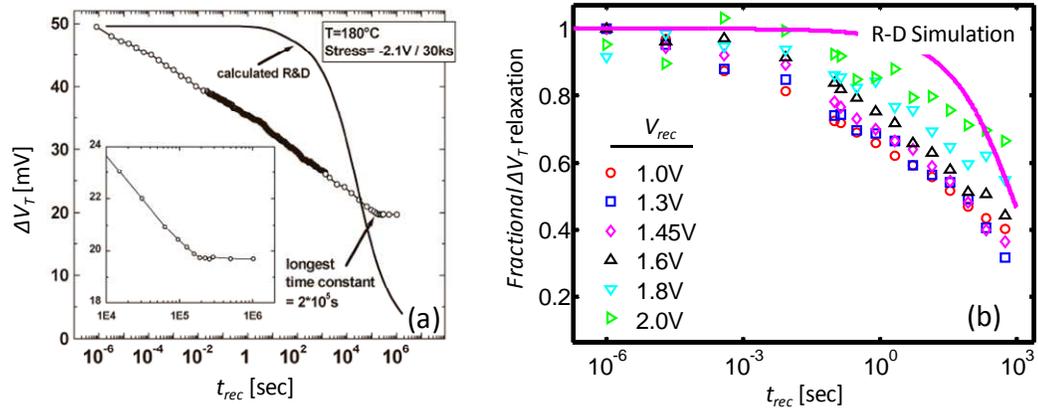


Figure 3.13: (a) Ultra-fast V_T measurement [170] predicts $t_{REC,start} \sim 10^{-10} t_{STS}$, which is much smaller than $t_{NIT,start} \sim 10^{-4} t_{STS}$, predicted by poly H-H₂ R-D model. (b) Similarly, on-the-fly $I_{D,lin}$ predicts a recovery voltage (V_{rec}) dependent relaxation and $t_{REC,start} \sim (10^9-10^6) t_{STS}$. However, consideration of valence band electron trapping (see section 6.7.2) and hole detrapping from oxide defects (see section 6.7.5), even for type-I transistors, can significantly reduce the difference between $t_{REC,start}$ and $t_{NIT,start}$.

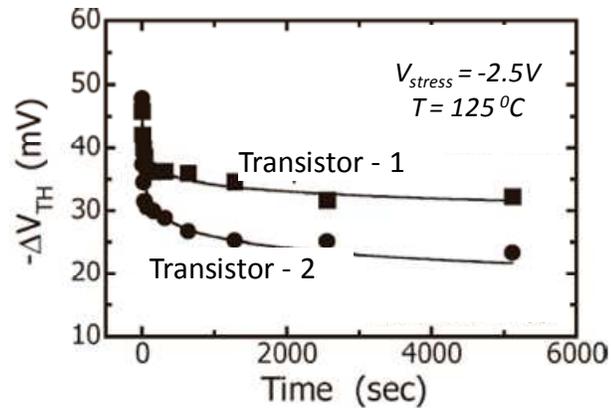


Figure 3.14: Presence of non-zero saturation during N_{IT} relaxation (figure taken from [171]) is a feature sometimes observed in NBTI experiments.

3.7. Field Dependence in H-H₂ R-D Model

So far, we have discussed the time dynamics of N_{IT} during NBTI stress-phase and relaxation-phase. However, time dynamics is not the only (though the major one) for reliability analysis. Reliability experiments are often accelerated at higher stress voltage V_{STS} to observe the significant amount of N_{IT} generation (thus identify the dynamics with better noise margin) within the experimentally feasible measurement window of several hours to some weeks. Based on this accelerated testing at higher voltage, one needs to predict the dynamics at operating condition (which is much less than V_{STS}) upto the lifetime (~ 10 years) of a transistor. Moreover, transistors at different parts of an integrated circuit operate at different voltages. Therefore, one needs to have a physical model for N_{IT} generation, so that experiment performed in a set of V_{STS} can be used to predict the level of N_{IT} at any V_{STS} .

In this section, we develop (and validate) such model for hole-assisted Si-H bond dissociation, which was briefly presented in 3.4.1. As such, we present the physical basis of the phenomenological field dependence within R-D model, first demonstrated in [30]. In addition, we discuss the consistency of the proposed field model by analyzing N_{IT} generation (ΔN_{IT}) experiments on wide variety of type-I transistors (having dominance of N_{IT} generation during NBTI) involving nitrogen and strain [119, 128, 133]. Here, we use H-H₂ R-D model of section 3.5.4 for interpreting N_{IT} experiments, though any other R-D models having $n = 1/6$ (see Table 3.1) at long stress time could have served the purpose.

3.7.1. Field vs. Voltage-Dependent N_{IT} Generation

In order to explain how ΔN_{IT} or NBTI in type-I transistors varies with V_{STS} , we first need to identify whether NBTI is a field dependent or a voltage dependent phenomena. This has been studied by different research groups [30, 90, 136, 172-174] in a number of ways. By studying N_{IT} generation on transistors having different effective oxide thickness EOT (see Appendix A for definition), it has been shown [172] that at same V_{STS} , NBTI is lower for higher EOT transistors (due to lower E_{ox}); whereas at same E_{ox} , NBTI is lower for lower EOT transistors (as $\Delta V_T \sim \Delta V_{IT} \sim q\Delta N_{IT}/C_{di} \sim EOT$ for same E_{ox} or ΔN_{IT} ; where

$C_{di} \sim \epsilon_{SiO_2}/EOT$ is the dielectric capacitance and ϵ_{SiO_2} is the dielectric constant of SiO_2). Similar study is also performed in [174] to conclude that ΔN_{IT} is EOT independent, when measured at same E_{ox} . Moreover, by applying same negative V_{STS} to bias a PMOS in inversion and NMOS in accumulation, it has been shown [90, 136, 173] that PMOS transistor, having higher electric field at same negative V_{STS} , degrades more compared to NMOS transistor. Also, the difference between NMOS and PMOS degradation gets reduced, when negative V_{STS} for NMOS is increased [173] to ensure that both PMOS and NMOS degrade at similar surface electric field. Therefore, ΔN_{IT} is expected to be a field-dependent phenomenon. See section 3.7.3 for more signatures on this field dependence.

Such field dependence of interface trap generation, hence NBTI for N_{IT} -dominated type-I transistors, has been considered using several empirical field models, *e.g.*, exponential model: $\Delta V_T \propto \exp(\gamma E)$ [172, 175], power-law model $\Delta V_T \propto (E)^\nu$ [176], mixed model: $\Delta V_T \propto E \exp(\gamma E)$ [30, 128, 130, 142], *etc.* Although a study in [176] shows power-law model to give better results compared to the exponential model, their experimental data has saturating trends, a possible artifact due to recovery [130]. In the next section, a physically based field dependent model is developed by considering the dipolar effects of Si-H bond, as well as field-dependent hole tunneling. It is shown that $E \exp(\gamma E)$ model (developed on a more physical basis [30, 128]) gives a consistent explanation of the field dependence in NBTI. Such model avoids the necessity of NBTI field dependence coming purely from dipolar dissociation (as suggested in [117]) and anticipates that neither power law nor exponential model gives proper estimation of NBTI lifetime; rather they misinterpret low-voltage NBTI degradation by over- or underestimating the lifetime, respectively as shown in section 3.10.

3.7.2. Hole-assisted, Field-enhanced, Thermal Si-H Dissociation

As shown in section 3.5.4 (see discussion below equation (3.30)), ‘H-H₂ R-D’ model anticipates that for type-I transistors at long stress time,

$$N_{IT} = \left(\frac{k_H}{k_{H2}} \right)^{1/3} \left(\frac{k_F N_0}{k_R} \right)^{2/3} (6D_{H2}t)^{1/6}. \quad (3.1)$$

Now, the field-dependence of N_{IT} generation must reflect the field-dependencies of k_F , k_R , D_{H2} , k_H and k_{H2} . As the diffusing hydrogen species is neutral [30, 128], k_R , D_{H2} , k_H and k_{H2} are field independent. As such, the field-dependence must be encapsulated in k_F and is explained by Figure 3.15b and Figure 3.16. The proposed model for k_F assumes that inversion layer holes (concentration $p_h \sim E_c$) near the Si/dielectric interface tunnel into and is captured by the Si-H bonds $\sim 1.5\text{\AA}$ (Si-H bond length) [69] away from the interface – leading to a hole-assisted, field-enhanced, thermal generation of interface defects. Here, E_c represents the electric field due to mobile carriers, which excludes the depletion charge contribution from total electric field E_{ox} in inversion and equals E_{ox} in accumulation; hole tunneling probability is expressed as: $T_H \sim P_T \exp(\gamma_T E_{ox})$, where $P_T \sim \exp(-\sqrt{m_{ox}\phi_{bh}})$ (see equation (3.10)) is field-independent pre-factor (m_{ox} : oxide effective mass and ϕ_{bh} : barrier height for hole tunneling) and $\exp(\gamma_T E_{ox})$ is field-dependent factor with field acceleration γ_T ; and the rate of field-enhanced, thermal dissociation of Si-H bond can be expressed as, $B \propto \exp[-(E_{A,F} - aE_{ox})/k_B T]$, where aE_{ox} represents thermal barrier lowering due to positioning of polar Si-H bond in the electric field, with a being the effective dipole moment [177]. In sum (Figure 3.16), $k_F \sim p_h T_H B$, so that –

$$\Delta V_{IT} = \frac{q\Delta N_{IT}}{C_{di}} = A_{IT} * EOT * (E_c)^{2/3} \exp\left(\frac{2\gamma_T E_{ox}}{3}\right) \exp\left(-\frac{nE_{D1}}{k_B T}\right) t^n, \quad (3.2)$$

where,

$$\gamma = \gamma_T + a/k_B T, \quad (3.3)$$

$$nE_{D1} = nE_{A,H2} + \frac{2}{3}(E_{A,F} - E_{A,R}), \quad (3.4)$$

and the pre-factor

$$A_{IT} \sim (N_0 P_T)^{2/3} \sim \left[N_0 \exp\left(-\sqrt{m_{ox}\phi_{bh}}\right) \right]^{2/3}. \quad (3.5)$$

For equation (3.2), C_{di} is dielectric capacitance and EOT is obtained from a fit of experimental C-V using simulators like [178] that includes the effects of multi-subband electron/hole quantization, poly-depletion, *etc*; whereas for equation (3.4), $E_{A,R}$, E_D are respectively activation energy for k_R and combined activation energy for D_{H2} , k_H , k_{H2} . The overall activation energy for ΔV_{IT} from equation (3.2) can be written as,

$$E_{A,IT} \equiv nE_{A,H_2} + \frac{2}{3}(E_{A,F} - E_{A,R} - aE_{ox}) = nE_{D1} - \frac{2}{3}aE_{ox}. \quad (3.6)$$

3.7.3. Experimental Validation of Phenomenological $\text{Exp}(\gamma E)$ Model

Figure 3.17 shows the field acceleration parameter (γ) and voltage acceleration factor (γ_V ; see Appendix A.3) as a function of EOT from a wide range of experiments, done by several groups (also compared with some of our results). Here, we use equation (3.2) for interpreting the experimental results and found that as EOT changes, γ remains almost constant ($\sim 0.6 \pm 0.05$) as expected for a field dependent degradation, while γ_V varies as $\sim 1/EOT$. To understand the γ_V vs. EOT dependence (for constant γ), let us consider two oxides of same EOT . If they are subjected to stress voltages V_1 (electric field, E_1) and V_2 (E_2) for producing same ΔV_T after time t_1 and t_2 , respectively, using equation (3.2), we have –

$$E_1^{2/3} \exp(2\gamma E_1/3)t_1^n = E_2^{2/3} \exp(2\gamma E_2/3)t_2^n. \quad (3.7)$$

Now, using $E \sim V/EOT$ (where, units of E , V , EOT are MV/cm, V, Å respectively) and $n \sim 1/6$, we obtain –

$$\gamma_V \text{ (in dec/V)} \approx \frac{\log_{10}\left(\frac{t_1}{t_2}\right)}{(V_2 - V_1)} \sim \frac{40}{0.23(EOT)} \left(\gamma + \frac{\ln(E_1/E_2)}{E_1 - E_2} \right). \quad (3.8)$$

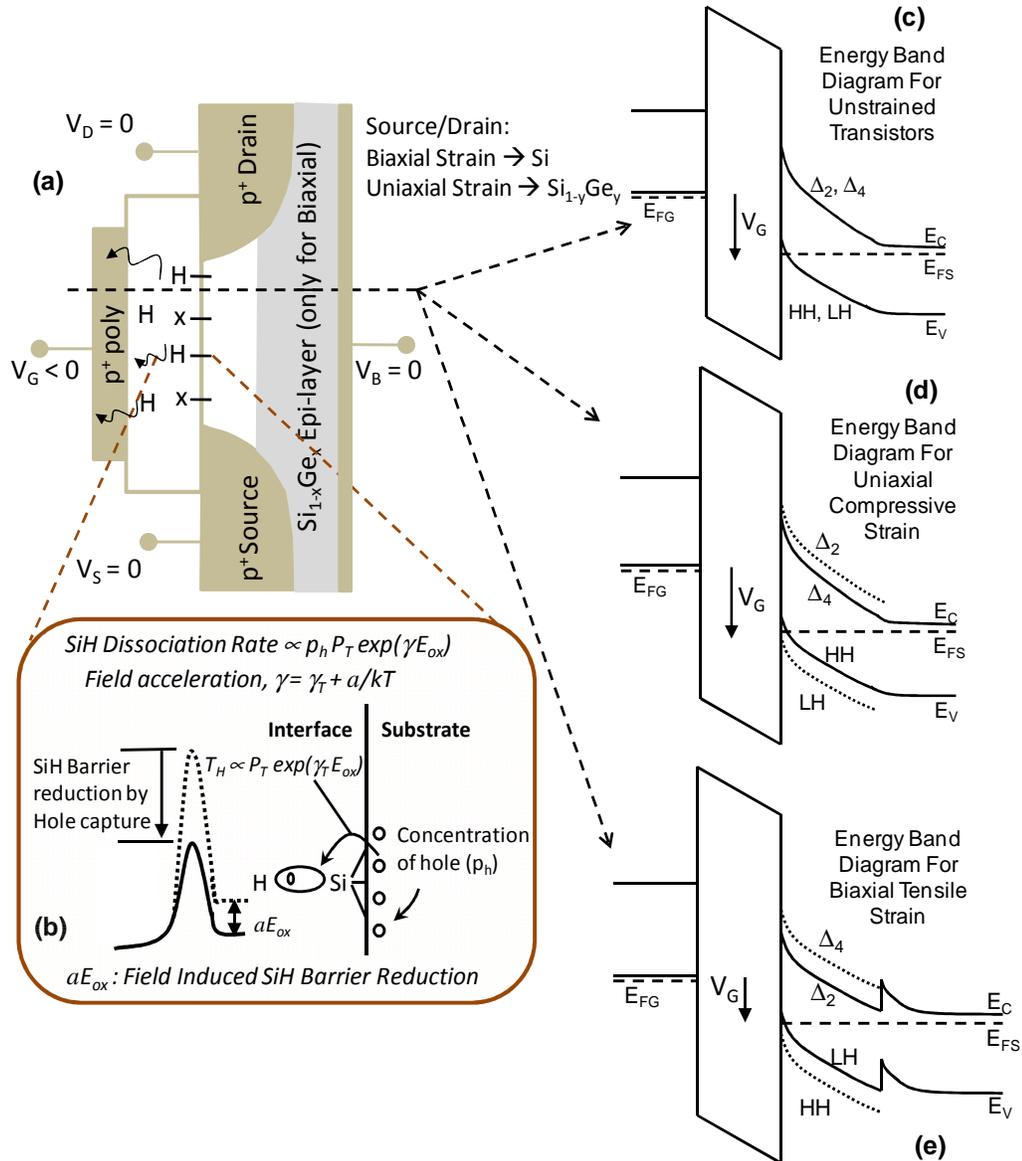


Figure 3.15: Universal view of N_{IT} generation under NBTI stress in strained/unstrained transistors. (a) Schematic of PMOS, showing the application of uniaxial compressive strain using $\text{Si}_{1-y}\text{Ge}_y$ Source/Drain or biaxial tensile strain using $\text{Si}_{1-x}\text{Ge}_x$ epi-layer. (b) Dissociation mechanism for Si-H bond (see Figure 3.16 for details). (c) Band diagram during NBTI stress, along the transverse direction, for unstrained transistor shows the degeneracy of electron (Δ_2 and Δ_4) and hole valleys (LH and HH). (d) Band diagram for uniaxially strained device shows the strain induced splitting of electron valleys (Δ_2 and Δ_4) and hole valleys (LH and HH). (e) Band diagram for biaxially strained device shows an extra band discontinuity at the strained-Si/ $\text{Si}_{1-x}\text{Ge}_x$ interface [179].

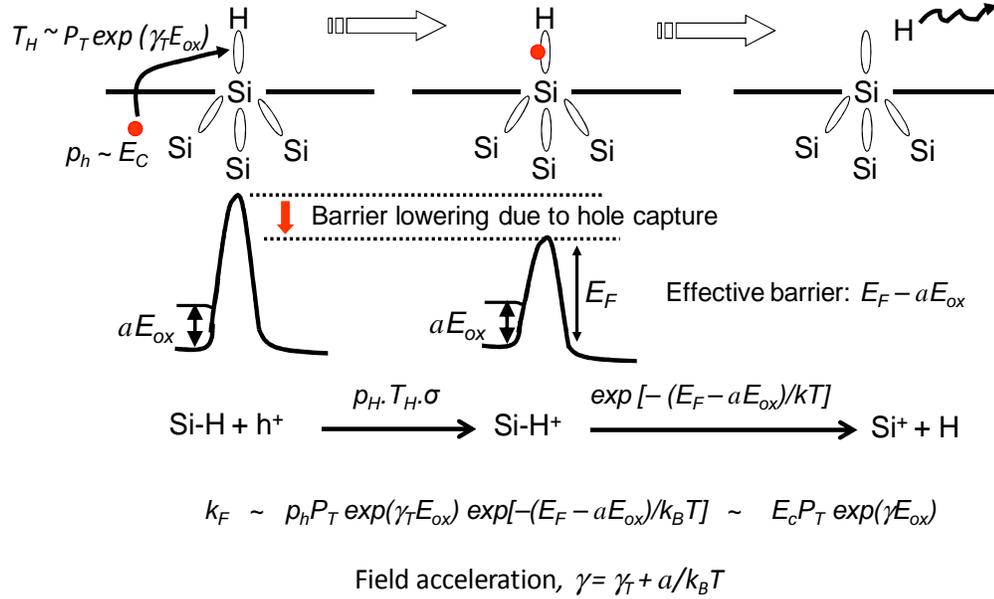


Figure 3.16: The proposed model for k_F assumes that inversion layer holes (concentration $p_h \sim E_c$) near the Si/dielectric interface tunnel into and is captured by Si-H bonds $\sim 1.5\text{\AA}$ [69] (Si-H bond length) away from the interface – leading to a hole-assisted, field-enhanced, thermal generation of interface defects.

Equation (3.8) suggests that an inverse relationship between γ_V and EOT (see Figure 3.17 and [30]) for constant field acceleration parameter (γ), but the constant of proportionality depends on electric field used to extract the data. This electric field-dependence is clearly seen in the deviation from inverse- EOT relationship in Figure 3.17 due to the use of ‘different E_1 and E_2 ’ for different devices. For example, $(E_1 - E_2)$ are higher for $EOT = 13\text{\AA}$ device compared to $EOT = 50\text{\AA}$ device. As a result, when $\sim 1/EOT$ line is drawn in Fig. 6 through $EOT = 13\text{\AA}$ data point, the drawn line falls below $EOT = 50\text{\AA}$ data point. As a result, γ_V values are on and above the solid line drawn in Fig. 6 for the respective devices considered⁷. Our analysis of degradation measurements performed

⁷ As electric fields for transistors (taken from references) considered in Figure 3.17 is estimated, a distinction between $E \exp(\gamma E)$ and $\exp(\gamma E)$ models may appear to be tentative.

by several groups validates the universality of the phenomenological $Eexp(\gamma E)$ field model.

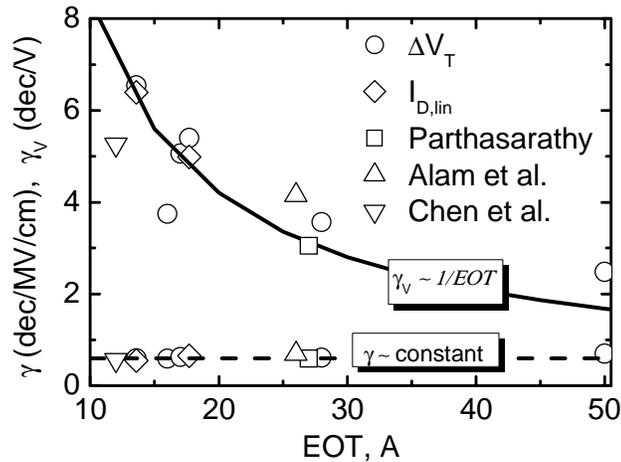


Figure 3.17: γ and γ_V variation with EOT . As predicted for field-dependent NBTI model, γ remains almost constant (dashed) and γ_V keeps around $1/EOT$ line (solid). For ΔV_T measurements, necessary delay corrections have been accounted for.

3.7.4. Theoretical Estimation of a (Figure 3.18)

According to Pauling scale [180], electronegativity difference between Si ($x_{Si} = 1.8$) and H ($x_H = 2.1$) is 0.3. The corresponding ionic bond energy is, $U_{H-Si} = 1.3(x_{Si} - x_H)^2 = 0.1$ eV allows one to estimate the effective charge transfer within Si-H dipole, i.e., $\sqrt{4\pi\epsilon_0 r U_{H-Si}} = 0.1q$ [177, 180], where q is the electron charge and r is Si-H Bond length ($\sim 1.5\text{\AA}$)⁸. Resultant dipole moment in Si-H bond is, $p = 0.15q\text{\AA}$. Considering that Si-H

However, based on the discussion in section III-C and the prediction of deviation in Figure 3.17 by equation (3.8), $Eexp(\gamma E)$ model appears more plausible.

⁸ Similar charge transfer is also calculated in [181], based on Sanderson's scale.

dipole resembles a dipolar orientation normal to a thin slab, local electric field [24] is given by, $E_{loc} = (1+L\chi_{int})E_{ox} = \epsilon_{int}E_{ox} = (3.9-8)E_{ox}$; where, $L=1$ (depolarization factor [24]), χ_{int} is dielectric susceptibility at interface and ϵ_{int} is dielectric permittivity (relative) at interface having values approximately 3.9 (SiO₂) to 8 (average of SiO₂ and Si) [182]. When Si-H bond is placed in an electric field, applied opposite to the polarization vector (which is the field polarity for NBTI stress), its energy will be increased by $pE_{loc} = p\epsilon_{int}E_{ox} = aE_{ox}$. Therefore, $a = p\epsilon_{int} = 0.6\sim 1.2$ qÅ.

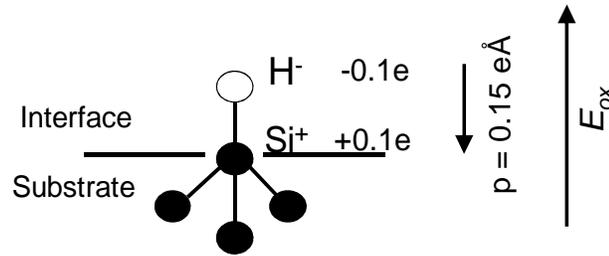


Figure 3.18: Polarization properties of Si-H bond provides an estimation of $a = p\epsilon_{int}$. Here, the polarization vector is opposite in direction of the applied field E_{ox} under NBTI stress.

3.7.5. Theoretical Estimation of γ_T

To estimate γ_T in equation (3.3) theoretically for *unstrained transistors*, first we determine T_H using self-consistent Schrödinger-Poisson solution of MOS electrostatics [178], which provides an energy band diagram similar to the one shown in Figure 3.15c. Electric field dependency of T_H gives γ_T in the range of 0.16-0.18, approximately independent of the nitrogen concentration in the dielectric (Figure 3.19a).

Next, we simulate γ_T for different *uniaxially compressive strained transistors* (Figure 3.19b), using the modified WKB approximation [183]. This involves calculation of p_h using self-consistent Schrodinger-Poisson solver [178, 184] and the energy band profile of Figure 3.15d, with the effective mass and band-splitting information from Table 3.2. We further repeat the procedure (compute γ_T) for different *biaxially tensile strained*

transistors (Figure 3.19c), using energy band diagram of Figure 3.15e with effective mass and band-splitting information from Table 3.2. As observed in unstrained transistors (Figure 3.19a), theoretically extracted values of γ_T also show negligible variation with any type of strain.

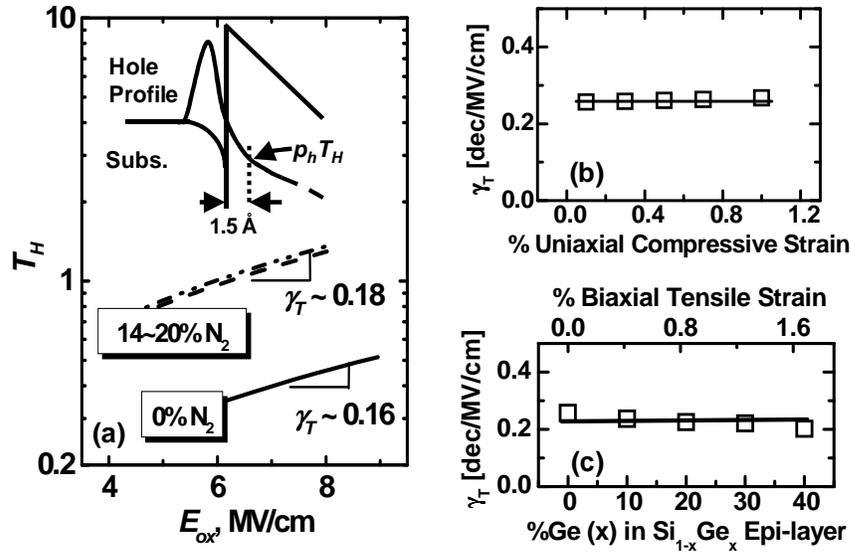


Figure 3.19: Tunneling probability (T_H) at 1.5Å (for clarity, the valence band-diagram in the inset of Fig. a is drawn upside-down) away from the MOS channel, calculated at different E_{ox} either using self-consistent Schrödinger-Poisson solution of MOS electrostatics [178] (in Fig. a) or using modified WKB approximation [183] (in Fig. b,c). Field dependence of T_H provides an estimation of γ_T for type-I transistors, having – (a) different %N within the dielectric, (a) variation of uniaxial compressive strain within the channel (a) variation of biaxial tensile strain within the channel. Theoretically extracted values of γ_T show negligible variation with any process technology.

Table 3.2: Band-structure information for strained transistors

| Strain Type | Quantity of interest | Expression | Reference |
|--|--|--|-----------|
| [110] Uniaxial compressive on (001) substrate (as a function of strain, ϵ) | Band gap [eV] | $1.08 - 4.35\epsilon$ | [185] |
| | Band splitting: electron [eV] | $\Delta_2: -140\epsilon^2 - 11.86\epsilon$ $\Delta_4: -5.6\epsilon^2 - 1.77\epsilon$ | [186] |
| | Band splitting: hole [eV] | $\Delta_{LH-HH} = 6.426 \epsilon$ | |
| | Effective mass: electron [$*m_0$]; where, m_0 is free electron mass | $\Delta_2: 0.19 + 4\epsilon(m_{t,\perp});$ $0.19 - 2\epsilon(m_{t,\parallel}); 0.916$ $(m_z);$ $\Delta_4: 0.916(m_l); 0.19(m_t);$ $0.916(m_z)$ | |
| | Effective mass: hole [$*m_0$] | $m_{HH} = 0.214 + 1.36\epsilon;$ $m_{LH} = 0.26 + 0.29\sqrt{\epsilon} -$ 0.54ϵ | [187] |
| Biaxial tensile strain using SiGe (as a function of Ge content, x in $Si_{1-x}Ge_x$) | Affinity of strained-Si | $4.05 + 0.58x$ | [179] |
| | Bandgap of strained-Si | $1.084 - x(0.31 + 0.53x)$ | |
| | Affinity of $Si_{1-x}Ge_x$ | $4.05 - 0.05x$ | |
| | Bandgap of $Si_{1-x}Ge_x$ | $1.084 - 0.42x$ | |
| | Dielectric constant of $Si_{1-x}Ge_x$ | $11.9 + 4.1x$ | |
| | Band-splitting between s-Si and $Si_{1-x}Ge_x$ | $\Delta_{Ec}: 0.63x; \Delta_{Ev}: x(0.74$ $- 0.53x)$ | |
| | Band splitting in s-Si | $\Delta_2 - \Delta_4: 0.67x; LH - HH:$ $0.4x$ | [188] |
| | Band splitting in $Si_{1-x}Ge_x$ | $HH-SO: 0.238x + 0.044$ | [189] |
| | Effective mass in s-Si | \sim strain invariant | [186] |
| | Effective mass in $Si_{1-x}Ge_x$ | Electron: strain invariant Hole: Using Vegart's law and Luttinger parameter | [189] |

3.7.6. Experimental Determination of γ_T and a

In determining γ_T and a experimentally on type-I transistors, having $\Delta V_T \sim \Delta V_{IT}$, we use equation (3.2) to fit a wide range of experimental ΔV_T vs. t data measured at different voltages and temperatures, as shown in Figure 3.20(a-d), which gives A_{IT} , γ , n as fitting parameters for each temperature (effect of E_{DI} can be included in A_{IT}). According to (3.3), a plot of γ vs. $1/k_B T$ (Figure 3.20e) gives tunneling parameter, γ_T , as intercept at $T_{STS} \rightarrow \infty$, and the effective dipole moment, a , as slope. Alternately, as equation (3.6) suggests a change in $E_{A,IT}$ with E_{ox} , $0.25 dE_{A,IT}/dE_{ox}$ can also be used to determine a (Figure 3.20f). A statistical summary of a , obtained using different experimental and theoretical approaches, is presented in Table 3.3, which shows remarkable consistency between the values obtained by our theory and measurements.

Table 3.3: Comparing a determined by various methodologies

| a [qÅ] | Obtained From | |
|-------------|--|--------------|
| 0.8 | $\gamma = \gamma_T + a/k_B T_{STS}$ | Figure 3.20e |
| 0.98 (avg.) | $a = 0.25 dE_{A,IT}/dE_{ox}$ | Figure 3.20f |
| 1.47 | $a = (\gamma - \gamma_T)k_B T_{STS}$ * | [117] |
| 0.6~1.2 | Si-H polarization | Figure 3.18 |

* γ obtained after delay correction and using $\gamma_T \sim 0.27$ (from Figure 3.20e). In [117], $a = 3.2$ qÅ.

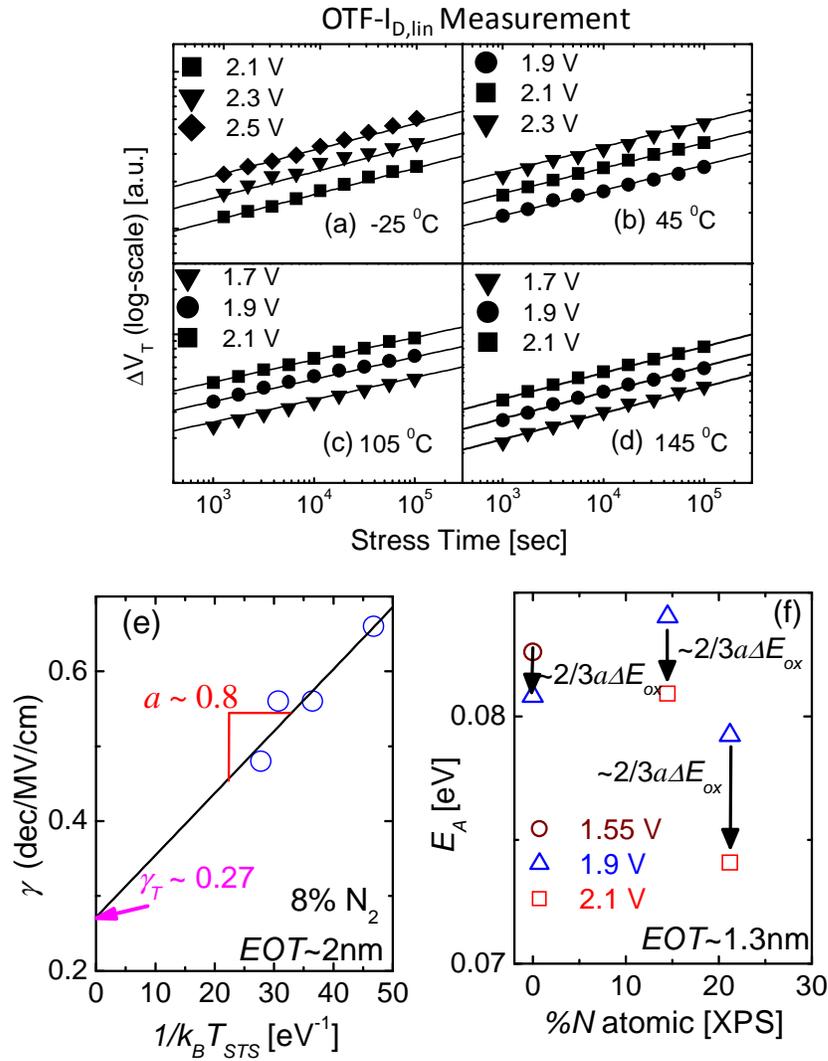


Figure 3.20: (a-d) Fitting $\Delta V_T \sim \Delta V_{IT}$ (measured by OTF- $I_{D,lin}$) vs. t (log-log scale) at different T_{STS} and V_{STS} by equation (3.2) gives A , γ , n as fitting parameters. Here $n = 0.15\sim 0.16$ and N_2 dose: 8%. Data is obtained from Dr. Anand Krishnan of Texas Instruments. (e) Obtained γ is plotted against $1/k_B T$, resulting $\gamma_T = 0.27 \pm 0.05$ dec/MV/cm as intercept at $T_{STS} \rightarrow \infty$ and $a = 0.8 \pm 0.15$ qÅ as slope. (f) Determining E_A at different electric field enables estimation of a , hence γ_T using equation (3.3).

3.7.7. Summary: Field Dependence of ΔN_{IT}

In summary, we observe that part of the field-dependencies of ΔN_{IT} arises from E_{ox} -dependence of T_H and the other part, from E_{ox} -induced thermal barrier lowering. Thus, we have provided the first consistent interpretation of field-dependence of ΔN_{IT} within the R-D framework. Previous analysis of delay-contaminated NBTI data [117] suggested experimental values of a that was 3~4 times larger than theoretically expected values. It appears that similar to the resolution of puzzles involving time-exponent ($n=0.25 \rightarrow 0.16$) and temperature dispersion parameter ($d = 1 - k_B T_{STS} / E_0 = 0.6 \rightarrow 0$, where E_0 is the characteristic width for density-of-states of hydrogen trapping sites in dispersive transport [33, 130]), zero-delay measurements and partitioning γ_T into $\gamma_T + a/k_B T$ (previous work incorrectly assigns $\gamma = a/k_B T_{STS}$ [117] or to $\gamma = \gamma_T$ [30] exclusively) hold the key for consistent estimation of field-acceleration of ΔN_{IT} .

In the next two sections, we further validate the proposed field model (section 3.7.2) of ΔN_{IT} by studying transistors having different dielectric material and channel strain.

3.8. Application of $E_{exp}(\gamma E)$ Model: Dielectric Material Dependence of ΔN_{IT}

In this section, we use our ΔN_{IT} generation model to explain the observed nitrogen ($\%N$) dependency of $\Delta V_T \sim \Delta V_{IT}$ in type-I transistors. As such, we identify the variation of model parameters with $\%N$ and provide physical explanation for the observed variations.

3.8.1. Determining $E_{exp}(\gamma E)$ Model Parameters

We use equations (3.2)-(3.6) in estimating $E_{exp}(\gamma E)$ model parameters ($A_{IT}, \gamma, \gamma_T, a, E_{DI}, n$) for type-I transistors (Table 3.4) by fitting ΔV_T measured at various V_{STS} and T_{STS} . For example, using equation (3.2) one can use voltage dependent stress data (@ fixed T_{STS}) to obtain A_{IT}, γ, n (Figure 3.21) or temperature dependent stress data (@ fixed V_{STS}) to obtain $n, E_A @ V_{STS}$ [30, 130] (therefore, E_{DI} using equation (3.6)). Determination of the

remaining two parameters (γ_T , a) are done by analyzing both voltage and temperature dependent data and using either equation (3.3) and/or (3.6) [190, 191]. For example, an estimation of γ at different T_{STS} and a plot of γ vs. $1/k_B T_{STS}$ (Figure 3.20e) gives γ_T as intercept at $T_{STS} \rightarrow \infty$ and a as slope, according to equation (3.3). Alternately, determining E_A at different electric fields (Figure 3.20f⁹) and using equation (3.6) enables estimation of a . Later using previously obtained γ and equation (3.3), γ_T can be estimated.

Table 3.4: Properties ($\%N$, T_{PHY} and EOT) of plasma oxynitride dielectric for the type-I transistors used in section 3.8.

| Device | N ₂ Dose [atomic %] | T _{PHY} [Å] | EOT [Å] |
|--------|-----------------------------------|----------------------|---------|
| #1 | 0 | 13.6 | 13.6 |
| #2 | 14.45 | 15.05 | 13.0 |
| #3 | 19.00 | 16.27 | 12.5 |
| #4 | 21.24 | 16.69 | 13.0 |

⁹ Although, negligible variation of E_A vs. $\%N$ in Figure 3.20f (@ 1.9V) contradicts with the E_A vs. $\%N$ variation reported in [37, 192], we identify this as a signature of negligible hole trapping in the plasma oxynitride devices studied here, which have lower N₂ concentration near substrate interface [129, 190], compared to the samples in [37, 192].

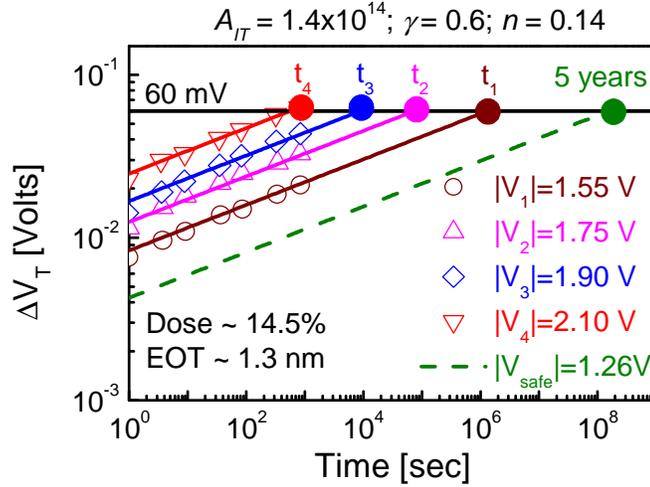


Figure 3.21: Voltage dependent $\Delta V_T \sim \Delta V_{IT} \sim \Delta N_{IT}$ stress data (taken at 125 °C) fitted using equation (3.2) enables one to estimate A, γ, n for a device ($E_{DI} \sim 0.9\text{eV}$ assumed, see section 3.8.2). Lifetimes (t_1, t_2, t_3, t_4) at different stress voltages (V_1, V_2, V_3, V_4) and safe operating voltage (V_{Safe}) can also be determined, as discussed in Appendix A.4.

3.8.2. Variation in $E_{exp}(\gamma E)$ Model Parameters with %N

Among four independent $E_{exp}(\gamma E)$ model parameters ($A_{IT}, \gamma, n, E_{DI}$), $n \sim 0.14$ is found to be constant (independent of %N) for the transistors under study. The parameters A_{IT} and γ , however, show systematic variation with %N (Figure 3.22). The source of variation in γ can be traced to variation in a and γ_T (see equation (3.3)). We find that the variation of a with %N is small (within the error margin of both methods of estimations in Figure 3.20). Therefore, we use $a = 0.8 \text{ q}\text{\AA}$ (Figure 3.20e) for all devices and we estimate γ_T using (3.3) for different %N (Figure 3.22); in other words, the %N-dependence of γ is actually reflected in %N-dependence of γ_T .

Finally, we use $E_{DI} \sim 0.9\text{eV}$, which is calculated (see eqn. (3.6)) using average value of $E_A (= 0.08\text{eV})$ for devices tested at $E_{ox} = 7\sim 10 \text{ MV/cm}$ and $27\sim 125 \text{ }^\circ\text{C}$, which is considered as E_A for all type-I transistors at an average E_{ox} of 8.5 MV/cm . Note that constancy of n with %N reflects dominance of interface defect generation and diffusion of hydrogen species within poly-Si [143] as the main cause of ΔN_{IT} for such transistors.

But constancy of E_{DI} with $\%N$ does not reflect any fundamental requirement, except that any variation is possibly embedded within the error-margin of the data. Thus ΔN_{IT} with $\%N$ variation is attributed to two factors: A_{IT} , γ_T .

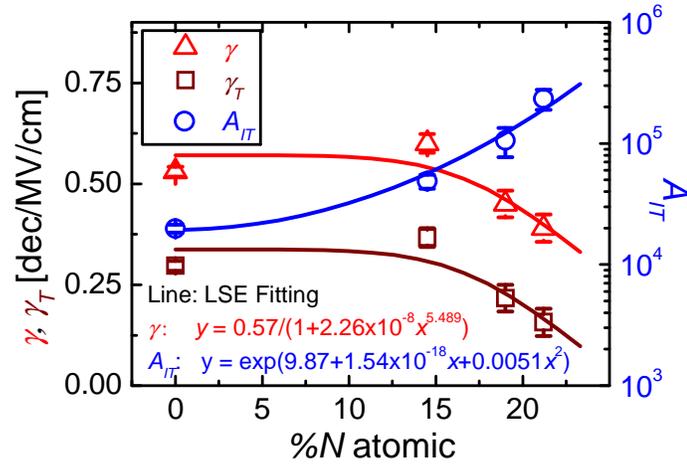


Figure 3.22: Variation of $Eexp(\gamma E)$ model parameters (A_{IT} , γ , γ_T) with $\%N$ ($a \sim 0.8$ qÅ used) for transistor in Table 3.4. The error margins in calculations are negligible.

3.8.3. Eexp(γE) Model Parameters: Physical Origin of $\%N$ Variation

In order to understand the physical origin of variations in A_{IT} and γ_T , we first calculate the theoretical estimates of these quantities. Both these quantities depend on tunneling probability (T_H) and within WKB approximation it can be expressed as –

$$T_H \sim \exp \left[-2 \int_0^{t_{int}} dx \sqrt{2qm_{ox} (\varphi_{bh} - xE_{ox})} \right], \quad (3.9)$$

where t_{int} is the interfacial layer thickness (approximately equals Si-H bond length of 1.5Å). By performing the integration in equation (3.9) and hence expanding the resultant terms in Taylor series, we have –

$$\ln T_H \sim -\frac{2\sqrt{2qm_{ox}\phi_{bh}}}{\hbar}t_{int} + \sqrt{\frac{m_{ox}}{2q\phi_{bh}}}\frac{qt_{int}^2E_{ox}}{\hbar} - \dots \dots \dots, \quad (3.10)$$

which indicates $P_T \sim \exp(-\sqrt{m_{ox}\phi_{bh}})$. This leads to $A_{IT} \sim [N_0 \exp(-\sqrt{m_{ox}\phi_{bh}})]^{2/3}$, as given in equation (3.5). Now the extracted values of oxide effective mass (m_{ox}) and barrier height for hole tunneling into Si-H bond (ϕ_{bh}), obtained from our gate leakage analysis, decrease with increase in %N (see section 7.4.2 for details). Therefore, according to equation (3.5), a decrease in m_{ox} and ϕ_{bh} with an increase in %N, is expected to increase A_{IT} , which is consistent with the experimental observation in Figure 3.22. Note that it is also possible to have an increase in A_{IT} due to an increase in N_0 ; but we can't distinctly identify such variation from our experimental observations.

We confirm the variation γ , observed in Figure 3.22, by estimating the same using a second set of type-I transistors, having a broader range of %N (using transistors of Table 6.1). Again, we observe a consistent decrease in γ (using both approximate or correct form of ΔV_T estimation, see section 6.7.4) with increase in interfacial nitrogen concentration, N_{int} (Figure 3.23). Moreover, our study of mobility parameters (P_1 , P_2 in Figure 6.23c,d and β_{IT} in Figure 6.24), which are signatures of hole wavefunction interaction with N_{IT} , indicates a decrease in these quantities (hence γ_T) with increase in N_{int} . This further indicates that the observed reduction in γ is mostly due to a reduction γ_T (as we used in Figure 3.22), *i.e.*, the change in a due to change in %N or N_{int} is negligible. Thus, from all these experimental signatures, we conclude that hole wavefunction interaction with N_{IT} (estimated either using γ_T or P_1 or P_2 or β_{IT}) decreases with increase in N_{int} (*i.e.*, with increase in %N for transistors having similar T_{PHY}).

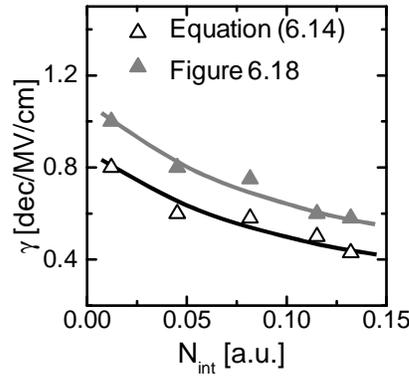


Figure 3.23: Variation of γ (where, ΔV_T is estimated either using classical method of equation (6.14), i.e., $|\Delta V_T| \sim |(V_G - V_{T0}) \Delta I_D / I_{D0}|$ and also using Figure 6.18) for type-I transistors as a function of N_{int} (lines are guide to eye only). See Figure 6.23e for the definition of N_{int} .

To explain the reduction of γ or γ_T with N_{int} , we go back to the equation being used for extracting γ (i.e., equation (3.2)), where we want to highlight the multiplication between γ and E_{ox} terms. As a reminder, E_{ox} represents an equivalent oxide electric field, referred to SiO_2 dielectric. Now the dielectric in type-I transistors are nitrided and hence the interfacial dielectric constant ϵ_{int} should have a value higher than SiO_2 (see Figure 7.3a). Therefore, the actual interfacial electric field, $E_{int} = \epsilon_{\text{SiO}_2} E_{ox} / \epsilon_{int}$, will be less than the field (i.e., E_{ox}) used for calculating hole tunneling in our Si-H bond dissociation model of Figure 3.16. As such, the reduction in γ or γ_T with N_{int} can easily be associated with the reduction in E_{int} (due to increase in ϵ_{int}) will N_{int} . In addition, there are also reports of Si-H bonds getting deeper into the dielectric, as N_{int} is increased [88, 97], which can also be partly responsible for the observed reduction in γ or γ_T for the type-I transistors under study.

3.9. Validation of $E_{exp}(\gamma E)$ Model: Strain Dependence of ΔN_{IT}

Strain technology has been extensively used over the last three CMOS generations [193]. ESR experiments [95, 96] on these transistors confirm that regardless of strain, these transistors are initially characterized by the same type of N_{IT} or P_b centers. However, it is still unclear whether the mechanism and the rate of NBTI-induced ΔN_{IT} in these strained transistors are, in any way, different from that of unstrained ones. Empirically, even though NBTI activation energy is strain-independent [194-197], reported ΔV_T due to ΔN_{IT} , for similar E_{ox} , are erratic (reported to be slightly decreasing/strain-invariant in [195, 198] or increasing with strain in [194, 196, 197, 199-201]). So, the question is: “If the type and mechanism/ activation energy of ΔN_{IT} is strain-independent, why does $\Delta N_{IT}@E_{ox}$ (i.e. rate of NBTI-induced ΔN_{IT}) depend on strain?”

In this section, we answer the aforementioned question by applying our $E_{exp}(\gamma E)$ model (Figure 3.16) to a wide-variety of uniaxial-compressive and biaxial-tensile strained transistors. Also, by explaining the variation of γ , $E_{A,IT}$ and hole wave function interaction with N_{IT} (i.e. γf), we justify the universality of our $E_{exp}(\gamma E)$ model (Figure 3.16).

3.9.1. ΔN_{IT} for Uniaxial Compressive Strain

Before starting our analysis, we need to verify the dominance of ΔN_{IT} in observed ΔV_T for these transistors, so that correction related to hole trapping (see section 6.7.5) can be safely ignored. For that, we measure ΔV_T using MSM setup (see section 6.4) for various uniaxial-compressively strained transistors having lightly dosed plasma nitrided dielectric ($EOT \sim 1.4\text{nm}$). Hence, we estimate N_{IT} 's contribution to ΔV_T or ΔV_{IT} using $\Delta V_{IT} = \Delta m * E_{G,Si}$, which assumes uniform N_{IT} generation within the Si bandgap ($E_{G,Si} \sim 1.1\text{eV}$). As sub-threshold slope is not effected by hole trapping (see Figure 8.1), presence of hole trapping is expected to make $\Delta m * E_G < \Delta V_T$ [202], which is not observed for strained/unstrained transistors under study (Figure 3.24a-c). To reaffirm this conclusion, we perform on-the-fly $I_{D,lin}$ measurement, with 1ms time-zero delay and appropriate mobility correction (section 6.7.4), for both unstrained (Figure 3.24d) and

strained (Figure 3.24e) transistors. We find power-law time exponent (n) of $\sim 1/6$ at wide range of voltages and temperatures and an activation energy of $\sim 0.09\text{eV}$ (Figure 3.25a), again suggesting ΔN_{IT} dominates ΔV_T [119, 129, 138, 154] in these transistors.

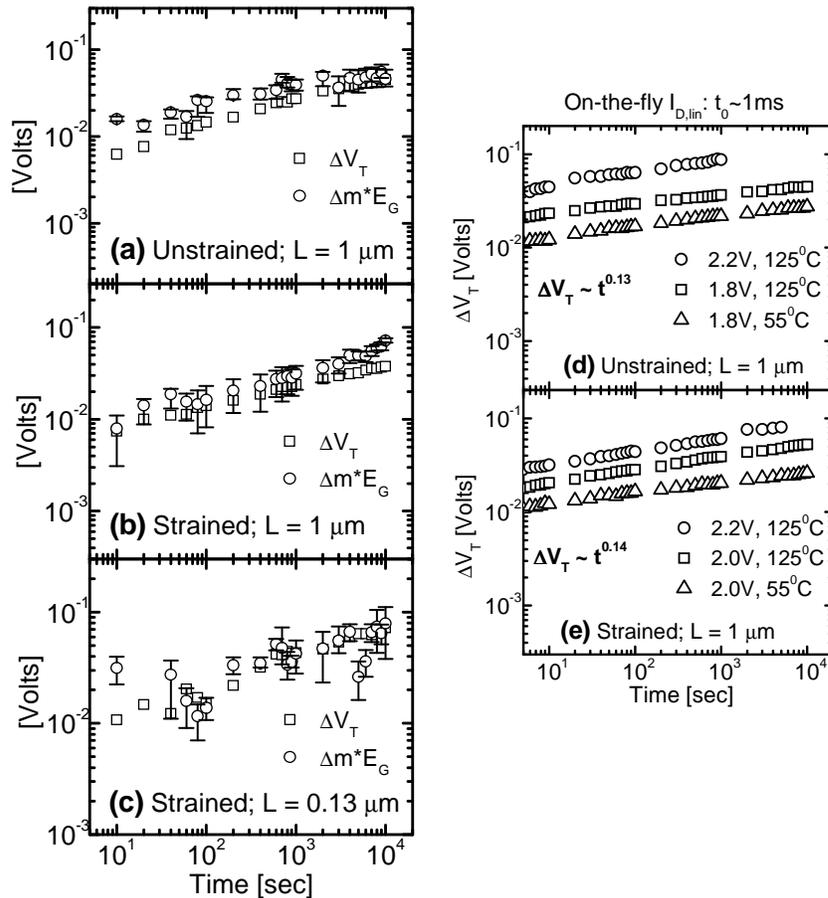


Figure 3.24: (a-c) Comparison of ΔV_T and body-effect co-efficient degradation (Δm) indicates negligible hole trapping in the uniaxial devices under study. (d,e) Power-law time exponents of ΔV_T in on-the-fly $I_{D,lin}$ measurements (time-zero delay, $t_0 \sim 1\text{ms}$, see section 6.7.3 for a discussion on the importance of t_0 in NBTI measurements) are close to $\sim 1/6$, again indicating negligible hole trapping [119, 129, 138, 154] in all the uniaxial compressively strained transistors under study.

Next, we fit ΔV_T using our field model (Figure 3.16) to determine γ as a function of uniaxial compressive strain (Figure 3.25b). In addition, we observe in section 3.7.5 and Figure 3.19b that γ_T is approximately invariant with uniaxial compressive strain. Therefore, according to equation (3.3) there would also be negligible variation in a , as well. Hence, *SiH bond dissociation mechanism (i.e., a , γ_T) is independent of uniaxial compressive strain.*

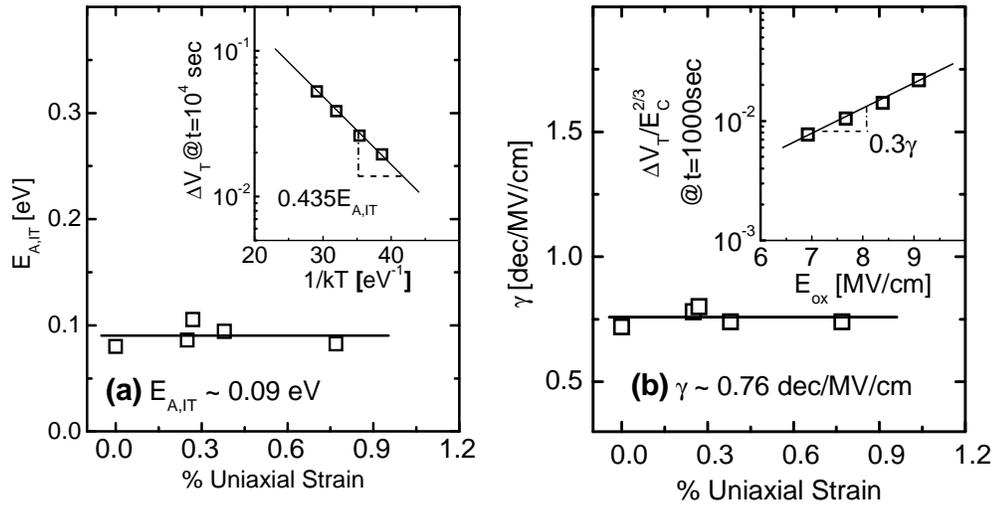


Figure 3.25: Uniaxial compressive strain causes: (a) negligible change in NBTI activation energy ($E_{A,IT} \sim 0.09$ eV), thus indicating similar diffusing species, as well as negligible hole trapping [119, 129, 138, 154]; and (b) negligible variation for γ . Insets show the calculation procedure for $E_{A,IT}$, γ .

Similar to the analysis in section 3.8.3, we verify the independence of γ_T (i.e., hole wave-function interaction) with *uniaxial* strain by studying the effective mobility (μ_{eff}) variation in Figure 3.26. Later, empirical mobility model ($\mu_{eff} = \mu_0 / [1 + \theta(V_G - V_T)^n]$; see section 6.4 for details) is used to extract variation in μ_0 ($\Delta\mu_0$) and θ ($\Delta\theta$) parameters using μ_{eff} vs. $(V_G - V_T)$ data obtained before and after NBTI stress (Figure 3.26a). We then estimate parameters P_1 and P_2 using equation (6.17) (Figure 3.26b), which provide a

reliable signature of hole wave function interaction with N_{IT} or γ_T (see section 6.7.4.5 and [203]). We find P_1 and P_2 essentially constant (Figure 3.26c), thus validating our conclusion that variation of γ_T with strain is negligible.

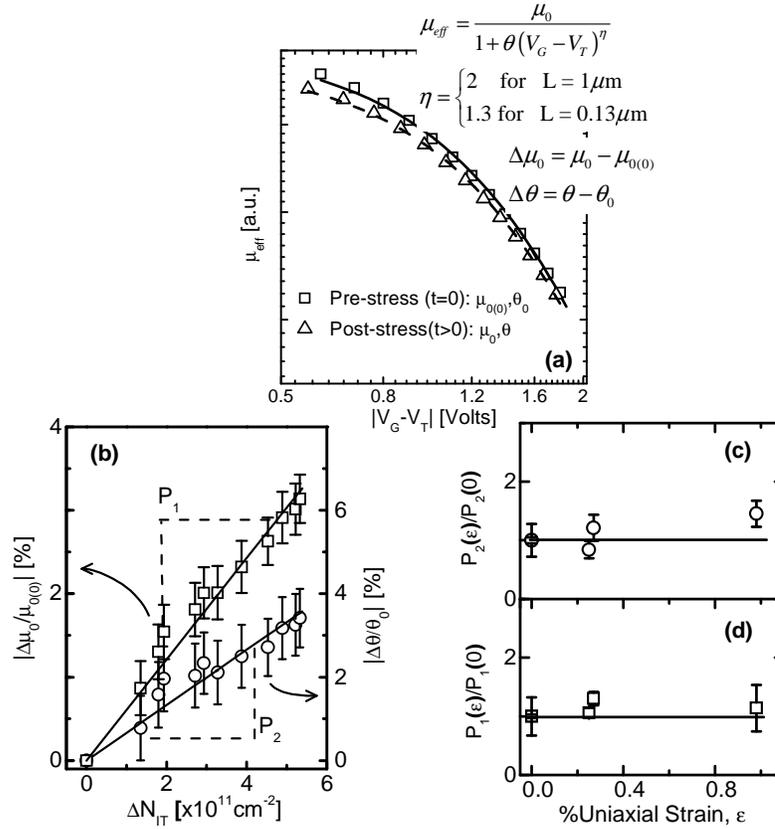


Figure 3.26: (a) Fitting μ_{eff} vs. $(V_G - V_T)$ using well-known empirical mobility expression, we estimate μ_0 and θ at different stress intervals. (b) Estimated degradation in μ_0 ($\Delta\mu_0$) and θ ($\Delta\theta$) correlate well with ΔN_{IT} and can be used to calculate P_1 , P_2 . (c) Variation of P_1 and P_2 are independent of uniaxial compressive strain, within the error margin of estimation, which is an independent signature of the invariance in wavefunction interaction (γ_T) with strain (see section 6.7.4.5 and [203]).

Given the independence in γ_T and a (hence γ) with uniaxial compressive strain, the observation that ΔV_T at constant E_{ox} ($\sim V_G - V_T$) decreases with increase in strain (Figure 3.27a) can only be explained if the factor $A_{IT} \sim [N_0 P_T]^{2/3} \sim [N_0 \exp(-\sqrt{m_{ox}} \phi_{bh})]^{2/3}$ in equation (3.5) decreases with compressive strain. We know that under compressive strain sub-bands with higher confinement effective mass (m_z), *i.e.*, heavy hole (HH) bands, form the top of the valence band (Figure 3.15c). Also, these HH bands are well-separated from the light hole (LH) bands. Hence, carrier occupancy in HH bands increases with strain and these carriers face higher ϕ_{bh} , due to higher m_z [193, 204]. As a result, A_{IT} or P_T decreases with increase in strain, which is consistent with the observed reduction in PMOS gate leakage (at fixed E_{ox}) for these devices (Figure 3.27b, also see [193]). Moreover, compressive strain also enhances the lattice mismatch between Si substrate and oxide [95, 96], thus increases N_0 . In sum, since $N_0 P_T$ (thus $\Delta V_T @ E_{ox}$ for fixed γ) in our compressively strained transistors decreases with strain (Figure 3.27a), we presume that the decrease in P_T supersedes any increase in N_0 for these transistors. Also, there are reports of transistors having increased $\Delta V_T @ E_{ox}$ (hence, higher $N_0 P_T$ for same γ) with compressive strain [197, 199, 200]. This is due to an increase in N_0 by incorporation of enhanced H during SiN deposition. Such increase in N_0 , exceeding the one expected from lattice mismatch [95, 96], is probably larger than the decrease in P_T (Figure 3.27b), thus increasing $\Delta V_T @ E_{ox}$ with strain for the devices in [197, 199, 200].

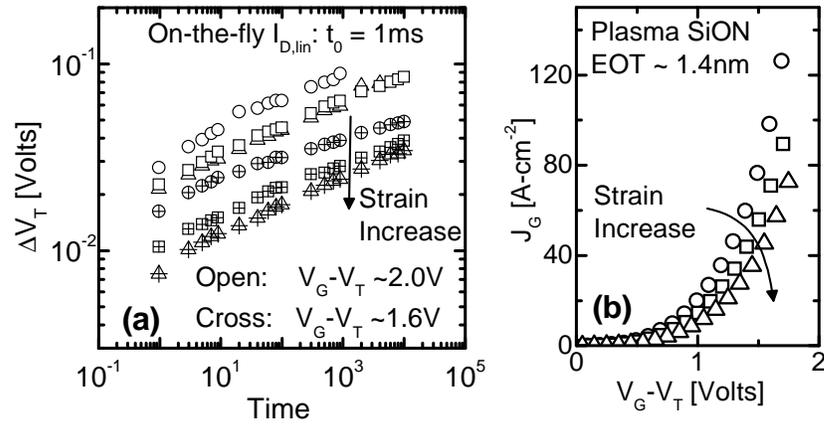


Figure 3.27: Though uniaxial compressive strain does not alter field acceleration parameter γ (Figure 3.25b), (a) it decreases ΔV_T , measured at similar electric field (or, $V_G - V_T$). This is shown here with three different strained devices, measured using on-the-fly $I_{D,lin}$ ($t_0 \sim 1ms$) at two different $V_G - V_T$. Decrease in ΔV_T with strain, for similar γ , implies decrease in $N_0 P_T$ in equation (3.5). Similarly, (b) decrease in gate leakage (at same $V_G - V_T$) for similar devices signifies a reduction in P_T , which should be dominating over any increase in N_0 [95, 96] to give reduced $N_0 P_T$ with increase in compressive strain.

3.9.2. ΔN_{IT} for Biaxial Compressive Strain

In section 3.7.5 and Figure 3.19c, we observed γ is approximately invariant with biaxial tensile strain. Since field acceleration (γ) is also independent of biaxial strain [194, 196], effective dipole moment for SiH bond (*i.e.*, a) should be invariant with biaxial tensile strain, as well as uniaxial compressive strain (as discussed in section 3.9.1).

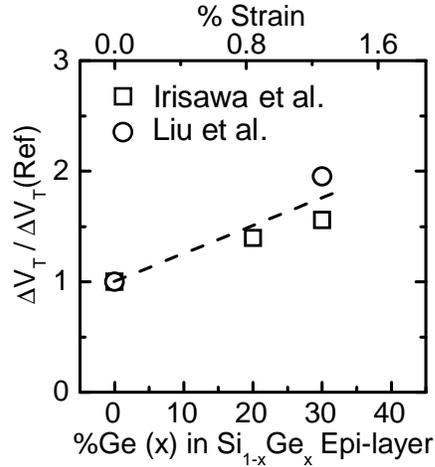


Figure 3.28: Biaxial tensile strain increases in ΔV_T (with respect to that for unstrained devices, measured at similar electric field); line is guide to eye only.

Moreover, contrary to compressive strain, tensile strain forces LH to form the top of the valence band (Figure 3.15d) [204]. Hence, occupancy in the LH band increases with tensile strain, which enhances gate leakage at fixed E_{ox} [193, 196] (due to lower m_z , hence reduced effective hole tunneling barrier for carriers in the LH band) – signifying an increase of P_T in our field model (Figure 3.16). Tensile strain also reduces N_0 due to reduced lattice mismatch between Si substrate and oxide [95, 96]. Hence, increased $\Delta V_T @ E_{ox}$, observed in [194, 196, 201] (Figure 3.28), should have resulted from a dominant increase in P_T with tensile strain, compared to the reduction in N_0 .

3.9.3. Summary: Strain Dependence of ΔN_{IT}

Thus by critically examining the impact on NBTI-induced ΔN_{IT} due to uniaxial/biaxial and compressive/tensile strain, we identify physical model parameter ($N_0 P_T$) as the single parameter responsible for the observed strain dependency over a wide range of experiments. We also can consistently explain the increase/decrease of $N_0 P_T$ with tensile/compressive strain for the transistors under study.

3.10. Predicting Lifetime Using Field Model

The field-dependent model of section 3.7 can be used to calculate the safe operating voltage for a given lifetime (say, 5 years) and failure criteria (*e.g.*, 60mV of ΔV_T) (Figure 3.21). Our calculation suggests that, neither exponential nor power law fit of ΔV_T data (taken at higher stress condition) provide exact estimations of safe operating voltage (Figure 3.29). Exponential fit underestimates and power law fit overestimates the lifetime for the transistor being considered. The power law fit can provide good estimate if low stress field data is used in lifetime projection (which is time-consuming).

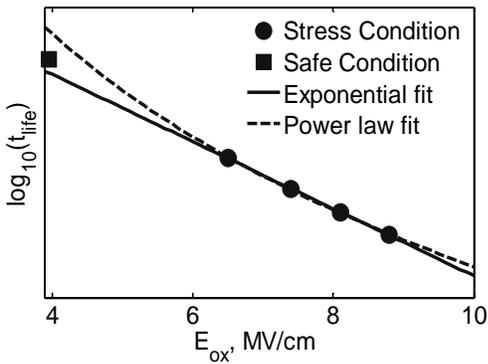


Figure 3.29: Comparison between exponential and power law fit of ΔV_T data for obtaining safe operating condition using stress data. Exponential fit underestimates and power law fit overestimates lifetime.

Finally, we consider the soft-saturation of ΔV_T at long stress time (*i.e.*, reduction of power-law time exponent n at $t_{\text{stress}} > 0$), which is very important for predicting NBTI lifetime. Because a reduction in n can significantly increase the lifetime of a transistor at a certain operating condition, thus relaxes the reliability criteria for transistor designers. Figure 3.30 shows measured time exponent as a function of time for various stress bias. Note that the time exponent reduces from the robust $n \sim 1/6$ value at long stress time (*e.g.*, $\delta n \sim 0.005/\text{dec}$), the reduction being higher for higher stress V_G . This reduction arises from the decrease in oxide electric field at constant voltage stress due to increase in ΔV_T over

time. The figure predicts a decrease in n to ~ 0.145 for stress condition @ 2.8V, 125 °C in Figure 3.5 and such exponent is consistently observed in experiments for type-I transistors [127, 128, 130, 140].

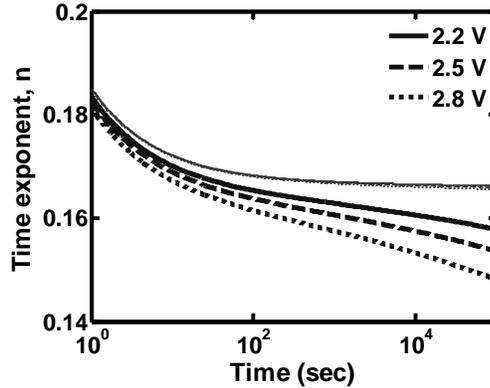


Figure 3.30: Decrease in long-term NBTI exponent results from the reduction of stress field (for constant voltage stress) with time. Here, thick lines consider such reduction in electric field over time, whereas thin lines do not consider such “self-saturation” effect.

3.11. Summary

In this chapter, we have systematically explored the time-dependent N_{IT} generation for various types of transistors having wide variation in nitrogen concentration within the dielectric film, as well as strain within the channel. Our analyses over different varieties of Reaction-Diffusion (R-D) theory enable us to identify H-H₂ R-D model to capture major features of N_{IT} experiments. In the process, we develop the first physical explanation of field-dependence within the R-D framework for describing N_{IT} generation. The model has been validated for type-I transistors (having dominance of N_{IT} generation during NBTI stress), which has different amount of nitrogen within the dielectric and strain within the channel. Finally, the relevance of this analysis is explored in predicting the lifetime of a transistor.

4. STATISTICS OF INTERFACE DEFECT GENERATION

4.1. Introduction

Dimension of modern transistors are extremely small, *e.g.* transistors used in 6T-SRAM for 45nm CMOS technologies have a gate area of $0.0048 \mu\text{m}^2$ [48]. As a result, a typical Si-H bond density of $\sim 10^{12} \text{ cm}^{-2}$ leads only to 48 Si-H bonds for the 45nm CMOS technology. Since Si-H bonds are the pre-cursor for interface defect N_{IT} , NBTI-induced N_{IT} will also be small (less than 10, for $N_{IT} \sim 10^{11} \text{ cm}^{-2}$). This necessitates a designer to consider the statistical distribution of generated interface defects at a particular stress time t_{STS} , in addition to the median N_{IT} (discussed in chapter 3). Here, one needs to estimate the spread of N_{IT} -induced ΔV_T up to transistor's lifetime and thus ensure that the tails of ΔV_T distribution do not cross the pre-defined reliability criteria.

The importance of such ΔV_T spread in circuit operation has already been analyzed in several publications over the last few years [49, 50, 205-209]. Nevertheless, (as we will see) the statistics of NBTI degradation, on which the circuit analysis is performed, requires better theoretical understanding. Though Poisson statistics was initially used (empirically) to interpret the spread of N_{IT} -induced ΔV_T [49], statistics like Gaussian or normal [48, 50, 205, 206], Skellam [207, 210], *etc.* have also been considered in subsequent years. Recently, negative ΔV_T tails are also observed in small transistors [207, 210], which has again been empirically explained by using the statistics of two uncorrelated Poisson-random process during N_{IT} generation [207], namely the dissociation of Si-H bond during the *creation process* and annealing of broken bonds during the *destruction process*. However, no systematic study has so far analyzed the *physical* origin of N_{IT} distribution in small transistors.

In this chapter, we obtain the statistics of N_{IT} -induced ΔV_T by performing Markov Chain Monte-Carlo (MCMC) analysis of Reaction-Diffusion (R-D) model both by analytical (section 4.3.1) and numerical (sections 4.3.2 and 4.4) calculations. We find that in small transistors, the distribution of $t_{STS}@N_{IT}$ follows a log-normal distribution and the distribution of $N_{IT}@t_{STS}$ in small transistors follows a skew normal distribution. However, at very long stress time or for large size transistors, both $N_{IT}@t_{STS}$ and $t_{STS}@N_{IT}$ statistics do revert back to the expected normal distribution. Therefore, we show that the statistics obtained from MCMC is consistent with the one obtained using analytical calculation. Moreover, contrary to the analysis in [207, 210], our MCMC R-D simulation establishes that the creation and destruction of N_{IT} is a correlated process, hence N_{IT} -induced ΔV_T should always be positive (section 4.3.3).

4.2. MCMC Transition Matrix

In this section, we set up the Markov Chain transition matrix using the R-D systems of section 3.5 with atomic hydrogen (section 4.2.1) and both atomic and molecular hydrogen (section 4.2.2) as diffusing species. These simulation frameworks allow us to obtain the N_{IT} distribution at different levels of t_{STS} and N_{IT} .

4.2.1. Transition Matrix: H Diffusion

To create the transition matrix for H-diffusion based R-D system, we first discretize equations (3.2)-(3.4) for constant time step Δt and constant grid size along the diffusion direction Δx . This results the following discretized version of R-D equations:

$$\left[N_{IT}(i+1) - N_{IT}(i) \right] / \Delta t = k_F \left[N_0 - N_{IT}(i) \right] - k_R N_{IT}(i) N_H^{(j=0)}, \quad (4.1)$$

$$\left[N_H^j(i+1) - N_H^j(i) \right] / \Delta t = D_H \Delta x^{-2} \left[N_H^{j+1}(i) + N_H^{j-1}(i) - 2N_H^j(i) \right], \quad (4.2)$$

$$0.5 \left[N_H^{j=0}(i+1) - N_H^{j=0}(i) \right] \Delta x / \Delta t = D_H \Delta x \left[N_H^{j=1}(i) - N_H^{j=0}(i) \right] + k_F \left[N_0 - N_{IT}(i) \right] - k_R N_{IT}(i) N_H^{(j=0)}, \quad (4.3)$$

where i and j indicate the time index and the space index, respectively ($j = 0$ indicates the Si/dielectric interface). Rearranging the terms in equations (4.1)-(4.3), we have –

$$\left[N_0 - N_{IT}(i+1) \right] / \Delta x = (1 - p_F) \left[N_0 - N_{IT}(i) \right] / \Delta x + p_R N_H^{(j=0)}, \quad (4.4)$$

$$N_H^j(i+1) = (1 - 2p_{DH}) N_H^j(i) + p_{DH} \left[N_H^{j+1}(i) + N_H^{j-1}(i) \right], \quad (4.5)$$

$$N_H^{j=0}(i+1) = 2p_F \left[N_0 - N_{IT}(i) \right] / \Delta x + \left[1 - 2p_R - 2p_{DH} \right] N_H^{(j=0)} + 2p_{DH} N_H^{j=1}(i), \quad (4.6)$$

where $p_F = k_F \Delta t$, $p_R = k_R N_{IT} \Delta t / \Delta x$, and $p_{DH} = D_H \Delta t \Delta x^{-2}$. Equations (4.4)-(4.6) enable us to define the Markov model of Figure 4.1 for the R-D system of equations (3.2)-(3.4). We can also express equations (4.4)-(4.6) in the following matrix form:

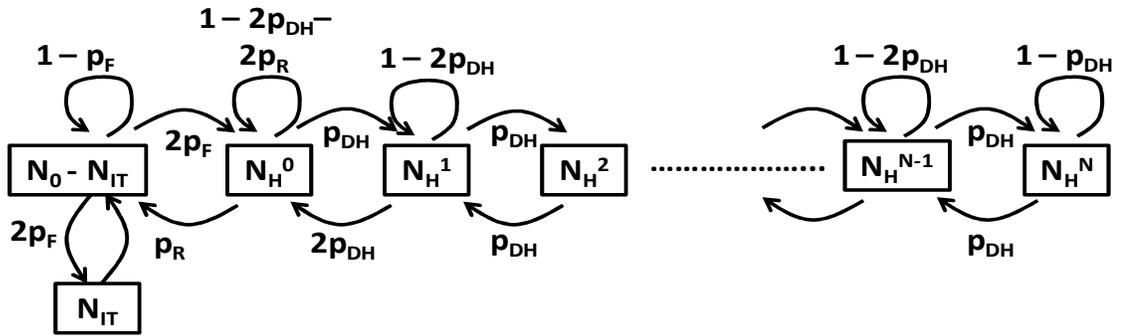


Figure 4.1: Markov model for probabilistic motion in an R-D system having H diffusion. The model is based on equations (4.4)-(4.6). The model considers reflecting boundary at the interface of poly/SiN cap (at $j = N$). Such reflection leads to saturation of N_{IT} and is only expected to occur at very long stress time.

4.2.2. Transition Matrix: H-H₂ Diffusion

To create the transition matrix for H-diffusion based R-D system, we first discretize equations (3.24)-(3.26) for constant time step Δt and constant grid size along the diffusion direction Δx . This results the following discretized (and rearranged) version of R-D equations:

$$\left[N_0 - N_{IT}(i+1) \right] / \Delta x = (1 - p_F) \left[N_0 - N_{IT}(i) \right] / \Delta x + p_R N_H^{(j=0)}, \quad (4.8)$$

$$\begin{aligned} N_H^{j=0}(i+1) = & 2p_F \left[N_0 - N_{IT}(i) \right] / \Delta x + \left[1 - 2p_R - 2p_{D_H} - 2p_{H,0} \right] N_H^{j=0}(i) \\ & + 2p_{D_H} N_H^{j=1}(i) + 2p_{H_2,0} N_{H_2}^{j=0}(i), \end{aligned} \quad (4.9)$$

$$N_{H_2}^{j=0}(i+1) = p_{H,0} N_H^{j=0}(i) + \left[1 - 2p_{D_{H_2}} - p_{H_2,0} \right] N_{H_2}^{j=0}(i) + 2p_{D_H} N_{H_2}^{j=1}(i),$$

$$\begin{aligned} N_H^j(i+1) = & p_{D_H} N_H^{j-1}(i) + \left(1 - 2p_{D_H} - p_{H,j} \right) N_H^j(i) \\ & + p_{D_H} N_H^{j+1}(i) + p_{H_2,j} N_{H_2}^j(i), \\ N_{H_2}^j(i+1) = & p_{D_{H_2}} N_{H_2}^{j-1}(i) + \left(1 - 2p_{D_{H_2}} - 0.5p_{H_2,j} \right) N_{H_2}^j(i) \\ & + p_{D_{H_2}} N_{H_2}^{j+1}(i) + 0.5p_{H,j} N_H^j(i), \end{aligned} \quad (4.10)$$

where $p_{D_{H_2}} = D_{H_2} \Delta t \Delta x^{-2}$, $p_{H,j} = k_H N_H^j \Delta t$, $p_{H_2,j} = k_{H_2} \Delta t$, D_{H_2} is the molecular hydrogen diffusion co-efficient, k_H is the conversion rate from H to H₂, and k_{H_2} is the conversion rate from H₂ to H. Equations (4.8)-(4.10) enable us to define the Markov model of Figure 4.2 for the R-D system of equations (3.24)-(3.26), having H-H₂ diffusion. We can also express equations (4.8)-(4.10) in the matrix form of equation (4.11).

The definition of matrix elements in equation (4.11) are similar to the one in equation (4.7), except equation (4.11) has some extra terms related to H₂ diffusion and H-H₂ inter-conversion. Also, the steps for performing monte-carlo simulation are similar to the one stated at the end of section 4.2.1.

4.3. Statistics of R-D System: H Diffusion

In this section, we use the transition matrix, defined in equation (4.7), to obtain the statistics of N_{IT} with H as the diffusing species. Here, one has to note that N_{IT} generation under NBTI stress is actually governed by H₂ diffusion (having power-law time exponent at long stress time, $n \sim 1/6$), rather than H diffusion (having $n \sim 1/4$) considered here. However, the statistical nature of R-D system does not depend on H or H₂ or H-H₂ diffusion and R-D system with H diffusion is easier to track analytically using generating function (see section 4.3.1). Therefore, we use MCMC analysis with H diffusion to draw the main conclusions about the statistical nature of R-D system and later show that R-D system with H-H₂ diffusion (see section 4.4) also has similar signature. We also artificially increase the magnitude of k_F (beyond the one reported in Figure 3.5) to observe N_{IT} statistics in the diffusion phase (commonly observed in experiment, having $N_{IT} \sim 10^{11} \text{ cm}^{-2}$) within the simulation time-frame of 1 sec. For these parameter sets, we observe no reflection from the poly/SiN cap interface within the simulation time-frame.

4.3.1. Analytical Calculation

Before performing MCMC simulation, we obtain the statistics of H-diffusion based R-D system through an analytical calculation of the moments of the distribution. This calculation enables us to interpret the simulation results, shown in section 4.3.2.

4.3.1.1. Statistics for $t_{STS}@N_{IT}$

Power series generating functions [212] are well adapted to calculate the moments of $t_{STS}@N_{IT}$ distribution by using the transition and staying probabilities of the markov transition matrix of equation (4.7). It can be shown using generating function that the mean t_{STS} for a particular level of N_{IT} is [213-215]:

$$\mu_{0,S} = \left[\sum_{k=0}^{S-1} \frac{1}{P_k} + \sum_{k=0}^{S-2} \frac{1}{P_k} \sum_{i=k+1}^{S-1} \prod_{j=k+1}^i \frac{q_j}{P_j} \right] \Delta t, \quad (4.12)$$

where p 's and q 's are the right (moving away from the Si/dielectric interface in the R-D system under study) and left (moving towards the Si/dielectric interface) transition probabilities, respectively; S is the average number of monte-carlo (or random walk) steps taken for generated N_{IT} ; for example, $S = 0$ means no Si-H bond dissociation, $S = 1$ means Si-H bond dissociation with insignificant diffusion, and $S > 1$ means that the dissociated H has diffused (on average) a distance $(S-1)\Delta x$ away from the interface. For equation (4.7), $p_0 \sim 2p_F = 2k_F\Delta t$, $q_1 \sim p_R = k_R N_{IT}\Delta t/\Delta x$, and $p_1 \sim p_2 = p_3 = \dots = p_{S-1} = q_2 = q_3 = \dots = q_S = p_{D_H} = D_H\Delta t\Delta x^{-2}$. Hence, using equation (4.12), we calculate –

$$\mu_{0,S} = \left[\frac{1}{2k_F\Delta t} + \frac{S-1}{D_H\Delta t\Delta x^{-2}} \left(\frac{S}{2} + \frac{k_R N_{IT}/\Delta x}{2k_F} \right) \right] \Delta t. \quad (4.13)$$

Similarly, using generating function we can also calculate (see Appendix D for details) variance of $t_{STS}@N_{IT}$ distribution:

$$\sigma_{0,S}^2 = \left[\begin{aligned} & \frac{1}{(2k_F\Delta t)^2} - \frac{1}{2k_F\Delta t} + \frac{S(S+1)(S^2+S+1)}{6(D_H\Delta t\Delta x^{-2})^2} + \frac{S}{D_H\Delta t\Delta x^{-2}} \\ & \left(\frac{2k_R N_{IT}/\Delta x}{(2k_F\Delta t)^2} - \frac{2k_R N_{IT}/\Delta x}{2k_F\Delta t} - \frac{S+1}{2} \right) \\ & + \frac{S(S+1)(2S+1)}{3(D_H\Delta t\Delta x^{-2})^2} \frac{k_R N_{IT}/\Delta x}{2k_F\Delta t} + \frac{S^2}{(D_H\Delta t\Delta x^{-2})^2} \left(\frac{k_R N_{IT}/\Delta x}{2k_F\Delta t} \right)^2 \end{aligned} \right] (\Delta t)^2. \quad (4.14)$$

Now, in the reaction (when, $S \sim 0$) and diffusion (when, $S \gg 1$) limits of the H-diffusion based R-D system, equations (4.13)-(4.14) reduces to –

$$\mu_{0,s} \approx \frac{1}{2k_F}; \quad \sigma_{0,s}^2 \approx \frac{1}{4k_F^2}, \quad (\text{Reaction regime}) \quad (4.15)$$

$$\mu_{0,s} \approx \frac{S^2}{2D_H \Delta x^{-2}}; \quad \sigma_{0,s}^2 \approx \frac{S^4}{6(D_H \Delta x^{-2})^2}. \quad (\text{Diffusion regime}) \quad (4.16)$$

Here, equation (4.15) reflects that if there were no diffusion or annealing (*i.e.*, both D_H and k_R are negligible), Si-H bond dissociation will occur after a mean time of $\sim 1/2k_F$, having a standard deviation of the same order. Interestingly, equations (4.15) and (4.16) suggest that both in the reaction and diffusion limits of the R-D system, $\mu_{0,s} \sim \sigma_{0,s}$. Remarkably, this is a general characteristics for a log-normal distribution, where $\sqrt{\exp \sigma^2 - 1} \sim 1$ in the following expression for probability distribution function (PDF):

$$PDF = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \quad (4.17)$$

For the $t_{STS}@N_{IT}$ distribution, we calculate $\mu = \ln \mu_{0,s} - 0.5 \ln(1 + \sigma_{0,s}^2 / \mu_{0,s}^2) \sim \ln \mu_{0,s}$ and $\sigma = \sqrt{\ln(1 + \sigma_{0,s}^2 / \mu_{0,s}^2)} \sim \sqrt{\ln 2}$ (or 0.362 in the base-10 log-scale). Thus $\sqrt{\exp \sigma^2 - 1} \sim 1.16$ for our case, such that, $\mu_{0,s} \sim \sigma_{0,s} \sim \exp(\mu + \sigma^2 / 2)$. Therefore, the generating function calculation anticipates a log-normal distribution for $t_{STS}@N_{IT}$, at any stress time, with an approximately constant σ . Obviously, for large t_{STS} or for large-area transistors, where $\mu_{0,s}$ is very large, log-normal distribution of equation (4.17) reverts back to the well-known Gaussian or normal distribution.

4.3.1.2. Statistics for $N_{IT}@t_{STS}$

Though we have developed an analytical formalism for predicting the distribution of $t_{STS}@N_{IT}$ in section 4.3.1.1, measurements are always performed to obtain the distribution of $N_{IT}@t_{STS}$. To interpret such measurements, we use our analytical results of $t_{STS}@N_{IT}$ distribution and hence graphically predict the $N_{IT}@t_{STS}$ distribution using Figure 4.3.

We know that the R-D system with H-diffusion has a power-law time exponent $n \sim 1$ in the reaction regime and $n \sim 1/4$ in the diffusion regime. Such power-law behavior suggests that when the $t_{STS}@N_{IT}$ distribution is log-normal (following equation (4.17)) with constant σ , the $N_{IT}@t_{STS}$ distribution should also be log-normal with constant σ . Now, the σ for the $N_{IT}@t_{STS}$ distribution will be n times the σ for the $t_{STS}@N_{IT}$ distribution (as graphically illustrated in Figure 4.3). Thus, the $N_{IT}@t_{STS}$ distribution will have a constant $\sigma \sim 0.362$ (in base-10 log-scale) in reaction regime and $\sigma \sim 0.0905$ in the H-diffusion regime. Our MCMC simulation also suggests that one can readily observe a log-normal distribution for $t_{STS}@N_{IT}$ (see Figure 4.5). However, the constraints of $N_{IT} > 0$ and presence of small σ results in a skew-normal or normal distribution for $N_{IT}@t_{STS}$ (see Figure 4.4 and Figure 4.7), as discussed later.

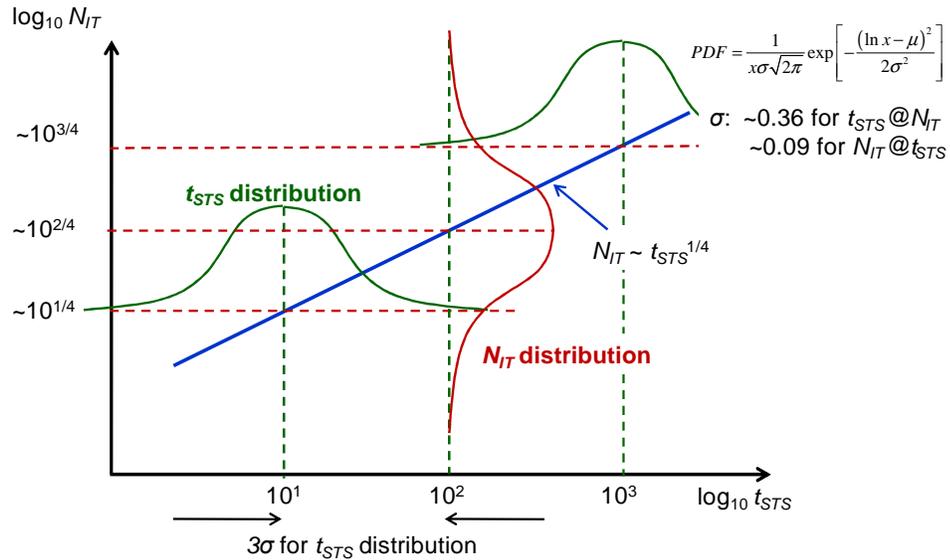


Figure 4.3: Schematic illustrating the idea that a log-normal distribution for $t_{STS}@N_{IT}$ guarantees a log-normal distribution for $N_{IT}@t_{STS}$ with constant σ for the distribution expressed in the form of equation (4.17).

4.3.2. MCMC Simulation Results

4.3.2.1. Consistency with Continuum Simulation

Before studying the statistics of N_{IT} generation, we first confirm that our MCMC results (simulated using equation (4.7)) is consistent with the continuum simulation results (obtained by numerically solving equations (3.2)-(3.4)). As shown in Figure 4.4, the *average* of about 100 MCMC simulations has remarkable consistency with the continuum simulation for a large ($L = 100\text{nm}$; $W = 1\mu\text{m}$) transistor (Figure 4.4a), as well as for a small ($L = 50\text{nm}$; $W = 100\text{nm}$) transistor (Figure 4.4b). However, average N_{IT} is not just the quantity that we care about in small transistors; we need to know the statistics. In the subsequent sections, we simulate the statistics of N_{IT} generation using equation (4.7), both analytically and numerically.

4.3.2.2. Simulated Statistics of $t_{STS}@N_{IT}$

After checking the consistency of our MCMC simulation with the continuum simulation, we use the MCMC simulation to obtain the distribution of $t_{STS}@N_{IT}$. As shown in Figure 4.5, such distribution is symmetric around the median value in a semilog-x plot, while only being cut-off at the two ends by the sampling time and the simulation length (1 ms and 1 sec, respectively for Figure 4.5). As such, it is evident that the $t_{STS}@N_{IT}$ distribution follows a log-normal behavior (as predicted from our analytical calculation in section 4.3.1.1), having its median point being traced by the continuum simulation. More interestingly, the log-normal distributions of Figure 4.5 has a constant $\sigma \sim 0.5$ (at all t_{STS}), which is a little larger than the one obtained from the approximate analysis in section 4.3.1.1.

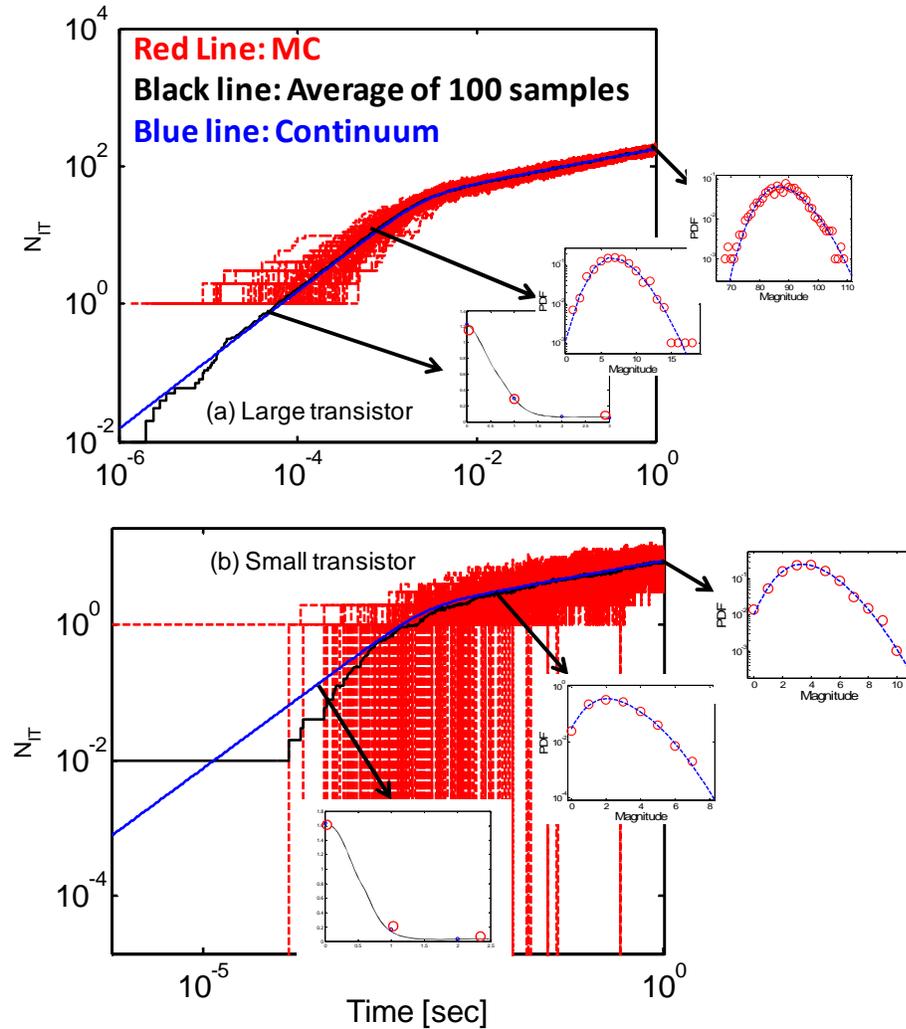


Figure 4.4: Time dependence and PDF of interface defect generation for (a) large transistor ($L = 100\text{nm}$, $W = 1\mu\text{m}$) and (b) small transistor ($L = 50\text{nm}$, $W = 100\text{nm}$) using $k_F = 3 \text{ sec}^{-1}$, $k_R = 2 \times 10^{-16} \text{ cm}^{-3}\text{sec}^{-1}$, and $D_H = 10^{-13} \text{ cm}^2\text{-sec}^{-1}$. For the large transistor, the PDF evolves from skewed normal (in the reaction phase) to a normal (in the diffusion phase) distribution. However, for the small transistor, the distribution stays skewed normal, even in the diffusion phase, which is commonly captured within the measurement window of NBTI measurement scheme (like MSM).

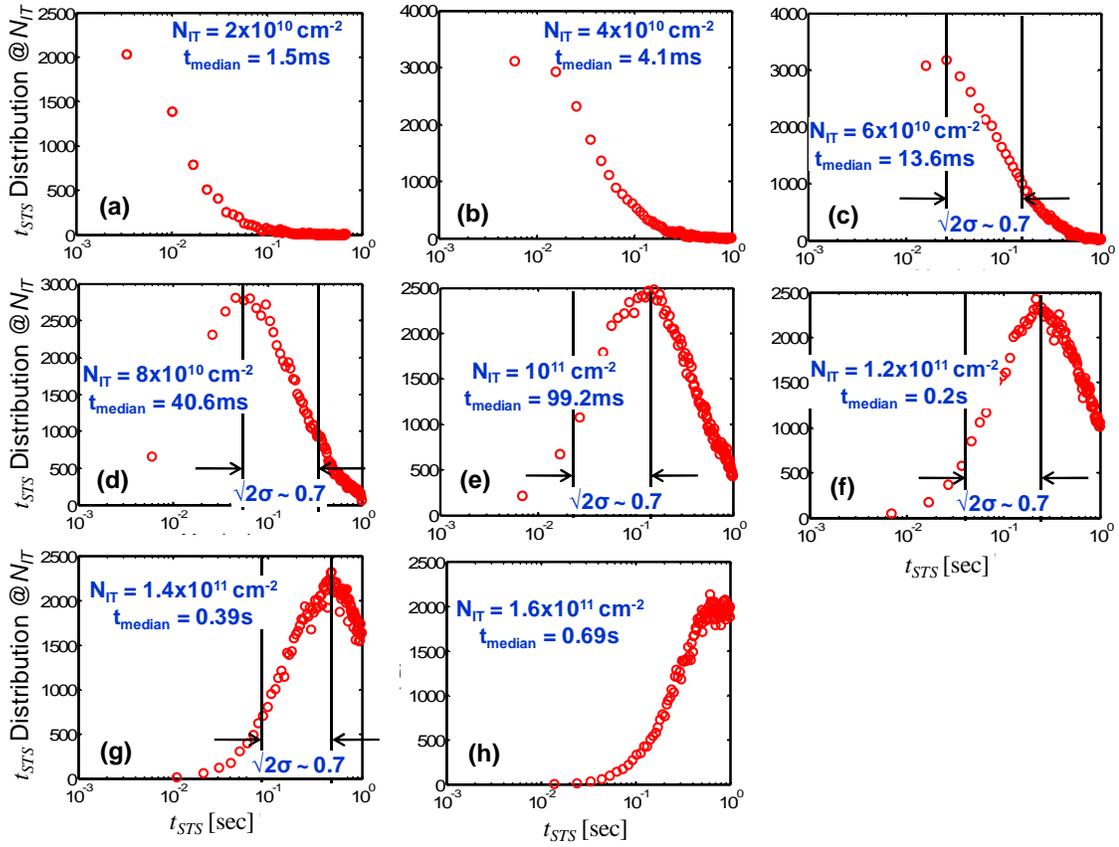


Figure 4.5: Distribution of t_{STS} at different levels of N_{IT} obtained from 1000 MCMC simulation in the small transistor under study. The distribution is clearly log-normal, having cut-off at $t_{STS} = 1$ msec (sampling time) and $t_{STS} = 1$ sec (maximum length of MCMC simulation). The distribution also has constant standard deviation in the semilog-x plot, as expected from analytical calculation.

4.3.2.3. Simulated Statistics of $N_{IT}@t_{STS}$

Given that the $t_{STS}@N_{IT}$ distribution is log-normal, let us study the distribution of $N_{IT}@t_{STS}$ (commonly measured in NBTI experiments) for the *large transistor* ($L = 100\text{nm}$, $W = 1\mu\text{m}$). Though the distribution is clearly Gaussian at longer simulation time, it shows significant skewness at shorter time (insets of Figure 4.4a). However, such skewness is only observed at a mean N_{IT} , $\mu_{IT} \ll 10^{10} \text{ cm}^{-2}$ (or $\Delta V_{T,mean} < 1\text{mV}$ for a transistor with $EOT \sim 1\text{nm}$), which is seldom reported in conventional NBTI statistics measurement. On the other hand, for *small transistor* ($L = 50\text{nm}$, $W = 100 \text{ nm}$), $N_{IT}@t_{STS}$

distribution do not become Gaussian, even for a $\mu_{IT} \sim 2 \times 10^{11} \text{ cm}^{-2}$ (insets of Figure 4.4b), commonly measured in NBTI experiments. The standard deviation of the N_{IT} distribution (σ_{IT}) always increases with mean μ_{IT} (see Figure 4.6a), which is also consistent with experimental observations [48, 49, 206, 208]. More importantly, $N_{IT}@t_{STS}$ distribution in small transistor has significant positive skewness γ_{IT} (see Figure 4.6), also reported in many NBTI measurements [207, 208, 210]. Therefore, the following expression of skewed Gaussian distribution [216] best fits the $N_{IT}@t_{STS}$ distribution in small-size transistors:

$$f(N_{IT}) = 2\phi\left(\frac{N_{IT} - \xi_{IT}}{\omega_{IT}}\right) \Phi\left(\alpha_{IT} \frac{N_{IT} - \xi_{IT}}{\omega_{IT}}\right), \quad (4.18)$$

where $\phi(x)$ and $\Phi(x)$ indicate the probability distribution function (PDF) and cumulative distribution function (CDF) of a Gaussian distribution, and ξ_{IT} , ω_{IT} , α_{IT} have the following relationship with the μ_{IT} , σ_{IT} , and γ_{IT} of the $N_{IT}@t_{STS}$ distribution [216]:

$$\begin{aligned} \mu_{IT} &= \xi_{IT} + \omega_{IT} \delta_{IT} \sqrt{2/\pi}, \\ \sigma_{IT} &= \omega_{IT} \sqrt{1 - 2\delta_{IT}^2/\pi}, \\ \delta_{IT} &= \alpha_{IT} / \sqrt{1 + \alpha_{IT}^2}, \\ \gamma_{IT} &= \frac{4 - \pi}{2} \frac{(\delta_{IT} \sqrt{2/\pi})^3}{(1 - 2\delta_{IT}^2/\pi)^{3/2}}. \end{aligned} \quad (4.19)$$

Now for fitting the $N_{IT}@t_{STS}$ statistics using equation (4.18), we first estimate μ_{IT} , σ_{IT} , and γ_{IT} of the $N_{IT}@t_{STS}$ distribution. Then, we estimate the skew-normal distribution parameters ξ_{IT} , ω_{IT} , and α_{IT} using equation (4.19) and hence check the goodness of equation (4.18) in matching the $N_{IT}@t_{STS}$ distribution. Figure 4.7 shows the $N_{IT}@t_{STS}$ distribution, obtained from 10000 MCMC simulations, in the small transistor ($L = 50\text{nm}$; $W = 100\text{nm}$) under study at $\mu_{IT} \sim 10^{11} \text{ cm}^{-2}$ and skew-normal distribution (equations (4.18) -(4.19)) indeed shows the best match compared to other distribution functions. Sum of squares due to error for skew-normal distribution is $\sim 4.6 \times 10^{-5}$, which is at least an order of magnitude smaller than the same obtained for other distribution functions.

4.3.2.4. Comparison Between Analytical and Numerical Results

Thus, the $N_{IT}@t_{STS}$ distribution follows either a skew-normal or normal distribution, depending on the size of the transistor. However, the interpretation of Figure 4.3 suggests that both $N_{IT}@t_{STS}$ and $t_{STS}@N_{IT}$ statistics should follow a log-normal distribution. Whereas, MCMC simulation suggests a log-normal behavior only for the $t_{STS}@N_{IT}$ distribution (Figure 4.5), but not for the $N_{IT}@t_{STS}$ distribution (see insets of Figure 4.4 and Figure 4.7). This is because σ of equation (4.17) for the $N_{IT}@t_{STS}$ distribution is quite small compared to μ , which was not the case for the $t_{STS}@N_{IT}$ distribution. As explained in Figure 4.5, σ for $N_{IT}@t_{STS}$ in the diffusion regime is 4 times smaller than the σ for $t_{STS}@N_{IT}$. As a result of this small σ for the $N_{IT}@t_{STS}$ distribution in the diffusion regime, the log-normal distribution for the $N_{IT}@t_{STS}$ statistics approaches a normal distribution, as observed for the large transistor in Figure 4.8a.

It would have been possible to observe log-normal $N_{IT}@t_{STS}$ distribution in the reaction-regime for large transistor (where, σ for the $N_{IT}@t_{STS}$ distribution is comparatively larger), however, in that regime the $N_{IT}@t_{STS}$ distribution is skewed towards the positive side (see Figure 4.8b) due to the constraints of $N_{IT} > 0$ (as is also the case in small transistor, see Figure 4.7). Therefore, though log-normal distribution is physically expected for the $N_{IT}@t_{STS}$ statistics, due to the constraints of small σ (for equation (4.17)) the $N_{IT}@t_{STS}$ statistics approaches a normal distribution at comparatively larger μ_{IT} . Also, for the constraints of $N_{IT} > 0$, the $N_{IT}@t_{STS}$ statistics for small μ_{IT} (as the case in small transistors) approaches a skew-normal distribution.

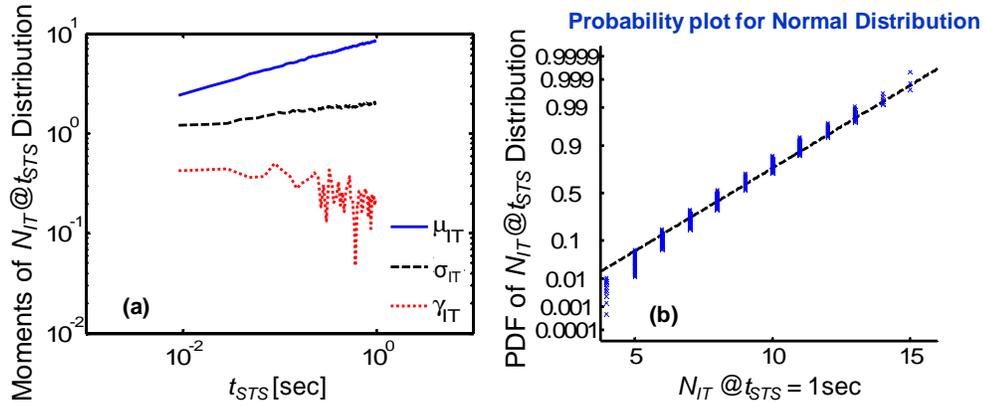


Figure 4.6: (a) Though the mean (μ_{IT}) and standard deviation (σ_{IT}) increases with t_{STS} with a power-law time exponent (~ 0.25 for μ_{IT} and ~ 0.1 for σ_{IT}), skewness (γ_{IT}) reduces with t_{STS} . Such reduction in γ_{IT} indicates that the distribution is approaching towards a Gaussian one at longer t_{STS} . (b) The normal probability plot also confirms the existence of skewness at $t_{STS} = 1$ sec for Figure 4.4b. The PDF has a shape similar to the one reported in [207, 208, 210].

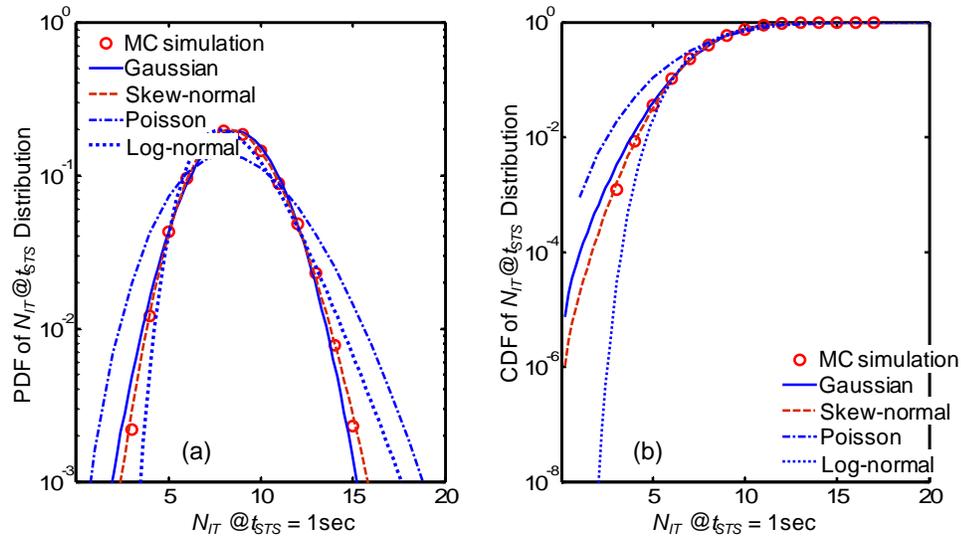


Figure 4.7: Both PDF and CDF of $N_{IT}@t_{STS} = 1$ sec can consistently be explained using the skewed normal distribution of equations (4.18)-(4.19). Other statistics (Gaussian, Poisson, and log-normal) are fitted with the simulated data using maximum likelihood estimation (MLE). SSE for different distribution functions are: 4.6×10^{-5} for skew-normal, 4×10^{-4} for Gaussian, 11.4×10^{-3} for Poisson, and 2×10^{-3} for log-normal.

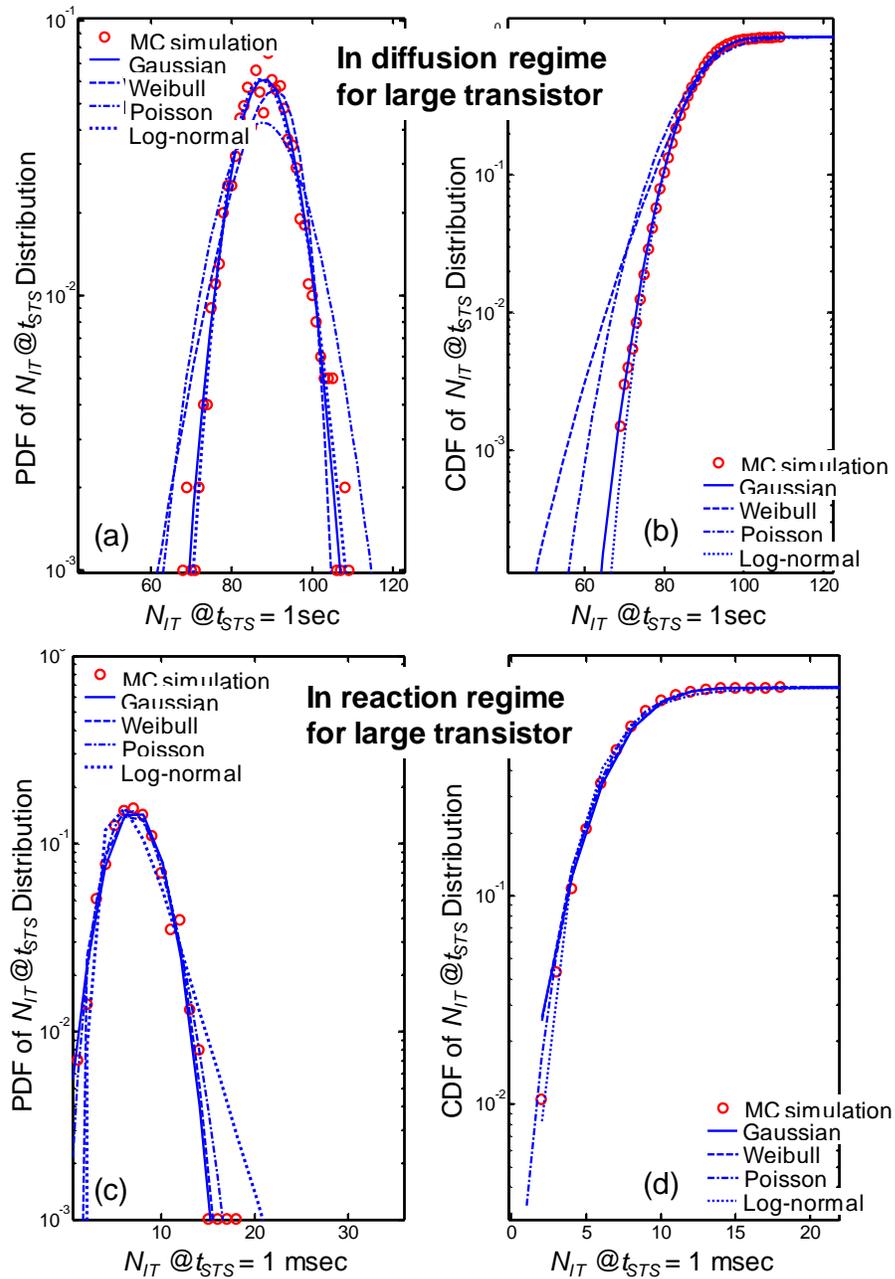


Figure 4.8: (a,b) PDF and CDF for a large transistor in the diffusion regime showing no distinction between a log-normal and normal distribution. (c,d) Log-normal was expected for the same transistor in the reaction regime; however, in this reaction regime, the distribution gets constraint by $N_{IT} > 0$ and gets skewed towards positive side of N_{IT} as in Figure 4.7.

4.3.3. On The Existence of Correlated N_{IT} Creation

As evident from our discussions so far, the statistics of a R-D system follows a log-normal distribution of equation (4.17), unless it is modified to a normal distribution for small σ , or is modified to a skew-normal distribution due to the $N_{IT} > 0$ constraints. Let us now discuss the correlation between the creation (Si-H bond dissociation) and destruction (Si- bond annealing) processes of N_{IT} . Such correlation was neglected to interpret the recently observed negative ΔV_T tails in NBTI experiments [207, 210].

In all our MCMC R-D simulations, the creation and destruction of N_{IT} , as well as diffusion of H, are considered as separate random processes. However, destruction of N_{IT} , through the second term in right-hand side of equation (3.2), is only possible for $N_{IT} > 0$. This dictates that the creation and destruction processes in R-D model can not be uncorrelated (as presumed in [207]) and consequently generated N_{IT} (hence ΔV_T) can not be negative. In support of this argument, none of our MCMC R-D simulation show negative N_{IT} -induced ΔV_T , suggesting a correlation between the creation and destruction processes in the R-D system.¹¹ Thus, the observation of negative ΔV_T tails in [207, 210] may be a consequence of N_{IT} relaxation in transistors having pre-existing N_{IT} or higher V_T , before the application of NBTI stress. As a result, the pre-existing N_{IT} of these high V_T transistors gets annealed during the NBTI stress and results negative ΔV_T (see Figure 4.9).

¹¹ Similarly, we also expect a correlation between the hole trapping and detrapping processes and, therefore, a non-negative ΔV_T due to the overall trapping-detrapping mechanism.

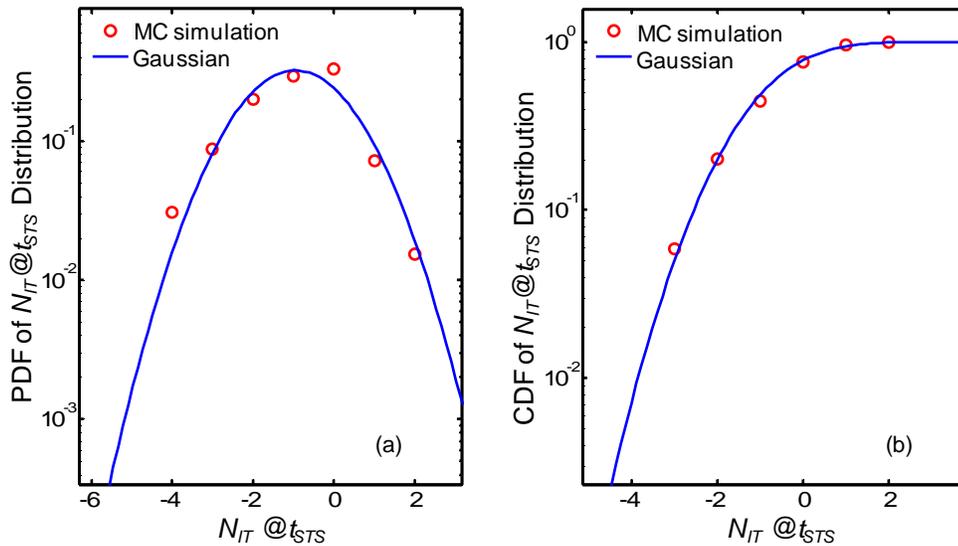


Figure 4.9: (a) PDF and (b) CDF of N_{IT} -induced ΔV_T for a small transistor having pre-existing interface defect. The distribution can explain the observation of negative ΔV_T tails in the experiments of [210].

4.4. Statistical Signatures of R-D: H-H₂ Diffusion

In the last section, we have analyzed the statistical signature of R-D system having atomic hydrogen as the diffusing species. Here, we repeat the same steps for the R-D system having both atomic and molecular hydrogen as the diffusing species. Such H-H₂ R-D system results in a power-law time exponent, $n \sim 1/6$, and hence more relevant for N_{IT} generation under NBTI stress. As we did for the H-based R-D system, we use a transition matrix (this time, we use the $[M]_i$ of equation (4.11)) to simulate the statistics of N_{IT} . The R-D parameters chosen in the MC simulation ensures that we observe the N_{IT} statistics in the diffusion phase (commonly observed in experiment, with $N_{IT} \sim 10^{11} \text{ cm}^{-2}$) within the simulation time-frame of 1 sec.

After setting the simulation framework, we check the consistency of our MCMC simulation (obtained using equation (4.11)) with the continuum simulation (obtained by numerically solving equations (3.24)-(3.26)). As shown in Figure 4.10a, the average of about 1000 MCMC simulations has remarkable consistency with the continuum

simulation results for a small ($L = 50\text{nm}$; $W = 100\text{nm}$) transistor. Next, we study the distribution of N_{IT} for the *small transistor* at different time steps. Again, similar to the H-based R-D system, N_{IT} distribution only reach a normal distribution for higher density of N_{IT} (insets of Figure 4.10a). Moreover, an analysis of the moments of the $N_{IT}@t_{STS}$ distribution shows an increase in the standard deviation of the distribution (σ_{IT}) with the increase in mean (μ_{IT}) (see Figure 4.10b). The distribution also has positive skewness at lower μ_{IT} (see Figure 4.10b), which reduces with increase in μ_{IT} .

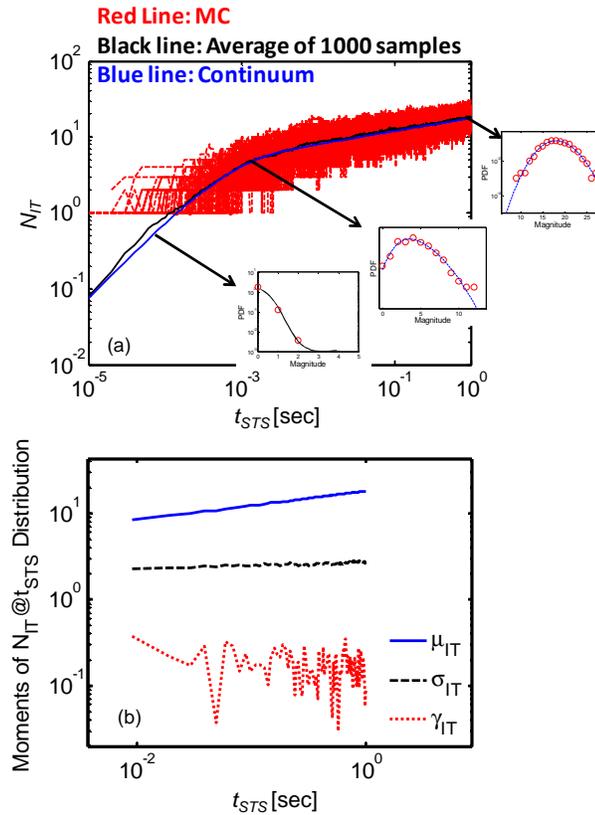


Figure 4.10: (a) Time dependence and PDF of interface defect generation for a small transistor ($L = 50\text{nm}$, $W = 100\text{nm}$) using $k_F = 30 \text{ sec}^{-1}$, $k_R = 2 \times 10^{-15} \text{ cm}^{-3} \text{ sec}^{-1}$, $D_H = D_{H2} = 10^{-13} \text{ cm}^2 \text{ sec}^{-1}$, $k_H = 9 \times 10^{-15} \text{ cm}^{-3} \text{ sec}^{-1}$, and $k_{H2} = 600 \text{ sec}^{-1}$. Similar to the observation in Figure 4.4b, the distribution stays skewed normal, unless the magnitude of mean N_{IT} (μ_{IT}) is quite large. (b) Both mean and standard deviation increases with t_{STS} with a power-law time exponent (~ 0.17 for μ_{IT} and ~ 0.04 for σ_{IT}), however, positive skewness γ_{IT} reduced with t_{STS} .

Now, let us review the $t_{STS}@N_{IT}$ statistics for the H-H₂ R-D system. Similar to the case for H diffusion based R-D system (Figure 4.5), H-H₂ R-D system also follows a log-normal distribution (Figure 4.11) and is curtailed at the two ends by the sampling time and the simulation length. Thus, we conclude that R-D system with H-H₂ diffusion has similar statistical nature as the R-D system with H diffusion.

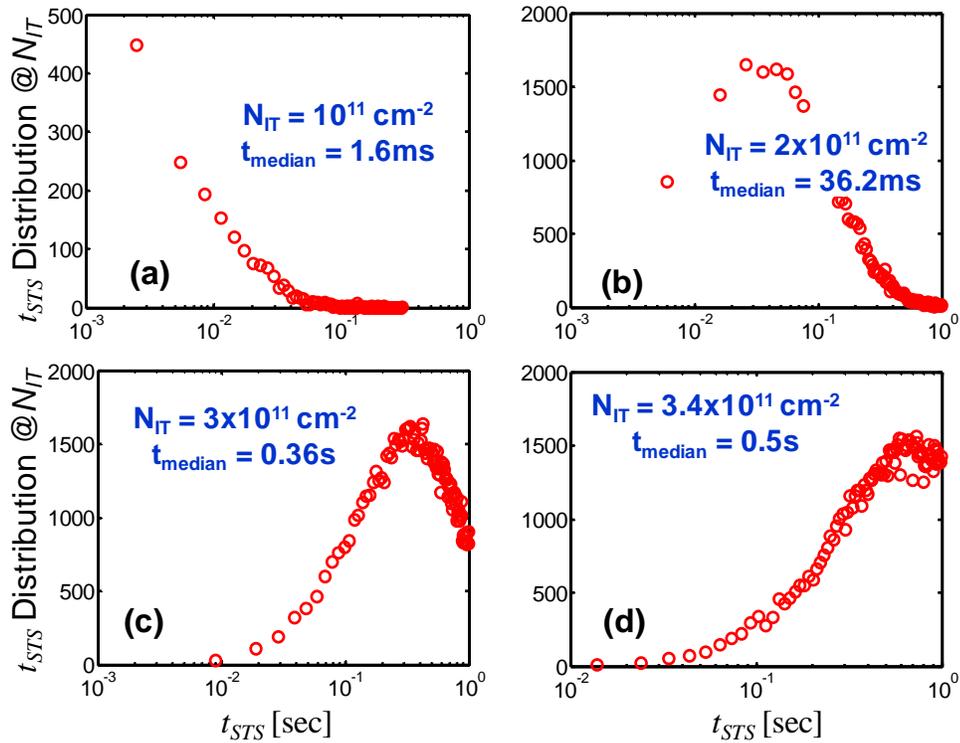


Figure 4.11: Distribution of t_{STS} at different levels of N_{IT} obtained from 1000 MCMC simulation (for R-D system with H-H₂ diffusion) in the small transistor under study. The distribution is clearly log-normal, having cut-off at $t_{STS} = 1\text{msec}$ (sampling time) and $t_{STS} = 1\text{sec}$ (maximum length of MCMC simulation).

4.5. Summary

In summary, we perform MCMC analysis of the R-D system in small transistors and show that both the $t_{STS}@N_{IT}$ and $N_{IT}@t_{STS}$ distributions should follow log-normal

behavior. One can readily observe such log-normal distribution for the $t_{STS}@N_{IT}$ distribution, which is only cut-off by the sampling time at lower end and simulation length (or equivalently measurement window) at the higher end. However, due to the constraints of $N_{IT} > 0$ and smaller (log-scale) standard deviation, log-normal behavior is seldom seen in $N_{IT}@t_{STS}$ distribution. The distribution is either skewed towards the higher N_{IT} side (from constraints of $N_{IT} > 0$) or approaches a Gaussian distribution (due to smaller log-scale standard deviation). Such behavior is universally observed in the R-D system, irrespective of the diffusing species being H or H₂. Moreover, our analysis also suggests that $N_{IT}@t_{STS}$ distribution should always have non-negative values, unless it is contaminated by pre-existing N_{IT} in certain statistically rare transistors.

The statistical distribution, presented in this chapter, considers Si-H bond dissociation and resultant hydrogen diffusion as an 1D problem; *i.e.*, we have a point source of hydrogen at the interface and the hydrogen is diffusing along an 1D line. However, Si-H bonds (and generated N_{IT}) are distributed along the 2D surface at the interface and generated hydrogen species diffuses in a 3D (gate and poly-Si) region. So for the same amount of N_{IT} , the corresponding ΔV_T will have its own distribution, depending on the spatial distribution of N_{IT} over the entire transistor structure [217, 218]. Because the number of dopants in modern nanoscale transistor is so small that different channel region of the transistor can have different V_T , which depends on whether that channel region has a dopant or not. As a result, if N_{IT} is created right on top of a dopant (*i.e.*, in the channel region having higher V_T), then that region will be completely depleted of inversion carriers, giving rise to large ΔV_T [kaczer]. However, if the same N_{IT} is created in a place away from the dopant, then ΔV_T can be significantly smaller. This aspect of spatial distribution of N_{IT} and its effect on ΔV_T can be a part of future study.

5. PRE-EXISTING OXIDE DEFECTS

5.1. Introduction

This chapter will cover dynamics of hole trapping into pre-existing oxide defects N_{HT} , which is especially important for type-II transistors having dominance of hole trapping during NBTI (see section 6.7.5). As usual, we first start by summarizing the existing views of N_{HT} dynamics in section 5.2, which is followed by a discussion in section 5.3 on the suitability of these models in explaining hole trapping experiments. Following these, in sections 5.4 and 5.5, we develop a hole trapping model within Shockley-Read-Hall trapping/detrapping framework and explained the importance of considering carrier fluxes from gate electrodes. The developed model is well-supported by experiments on transistors having a wide range of plasma and thermal nitridation within the dielectric.

5.2. Existing Views of N_{HT} Dynamics

Dynamics of oxide defects have been studied since 1970s' [219-221], when Metal-Nitride-Oxide-Semiconductor (MNOS) structure became a popular memory element, because of its simple structure and nonvolatile nature. Detrapping of carriers from pre-existing charged defects in the nitride layers was identified as the main mechanism for limiting the retention time in these memories. Moreover, discharge time was observed to increase with the increase in oxide thickness of the oxide layer, which proves that tunneling of carriers from charged defects within the nitride layer is getting reduced with increase in oxide thickness [219].

Similarly, pre-existing oxide defects was also considered as a source of threshold voltage instabilities in MOS transistors during the same period [222-224]. However, the

concept become quite popular in terms of hole trapping for explaining ΔV_T during NBTI stress, after the incorporation of nitrogen within the dielectric of MOS structures [136, 140, 141, 169, 202, 225-227]. Presence of similar trapping (sometimes with additional mechanisms [44, 228, 229]) of both electrons and holes, during NMOS PBTI (Positive Bias Temperature Instability) and PMOS NBTI respectively, is also reported in the recently used high- κ MOS transistors.

In terms of modeling of charge trapping (detrapping) into (out of) pre-existing oxide defects, three possible mechanisms are considered by Tewksbury *et al.* in [230]; which are (Figure 5.1): (a) elastic tunneling of carriers from the conduction band and valence band, (b) interfacial defect assisted trapping/ detrapping of carriers, and (c) phonon-assisted trapping/ detrapping of carriers. All these mechanisms consider carrier injection and emission only from the substrate into and out of the defects, respectively. Similar mechanism has also been used in [136, 140, 169, 225, 231, 232] for modeling trapping in MOS transistors, having both oxynitride and high- κ dielectric. In addition to this, the literature is replete of trapping/ detrapping models utilizing the concept of negative-U defects [228], hopping of carriers between the defects [44], defects having variable capture cross-section [230, 233], defects having energy-level shift [229], *etc.* Suitability of these models, mostly applied for MOSFET's having high- κ dielectric, is yet to be verified and is a topic of future research. However, in MOS structures having oxynitride dielectric studied in this chapter, only Tewksbury's approach is commonly used for explaining hole trapping (during NBTI stress) involving pre-existing oxide defects [136, 140, 169, 225].

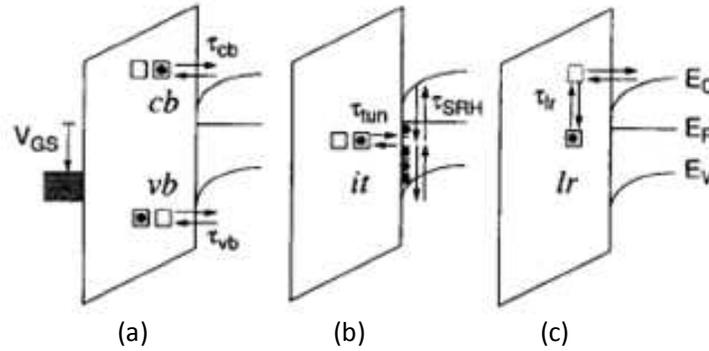


Figure 5.1: Three possible mechanisms for trapping/ detrapping of carriers into/ out of oxide defects considered in [230]. (a) elastic tunneling of carriers from the conduction band (cb) and valence band (vb), (ii) interfacial defect (it) assisted trapping/detrapping of carriers, and (iii) phonon-assisted (through lattice relaxation, lr) trapping/ detrapping of carriers.

5.3. Tewksbury's Model [230]

As stated earlier, this model considers the contribution from substrate carrier fluxes (through the three mechanism presented in Figure 5.1) in calculating trapping dynamics. This worked well for the thick dielectric ($T_{PHY} = 25\sim 75\text{nm}$) structures for which it was originally proposed. However, transferring the same concept in case of ultra-scale MOS structures having quite thin dielectric ($T_{PHY} = 1.5\sim 3\text{nm}$) becomes questionable [119, 129, 234]. In these thin dielectric structures, carriers from gate can also interact with the defects, thus making trapping dynamics a bit complicated. In the following section, we discuss the importance of incorporating carrier fluxes from gate electrode, within Tewksbury's trapping/detrapping model. Using such modified version of Shockley-Read-Hall trapping dynamics [235, 236], we model hole trapping (N_{HT}) in type-II transistors (section 5.4.1). Thus, we explain the effect of N_{HT} in NBTI degradation, in terms of power-law time-exponents and activation energies (section 5.4.2).

5.4. Effect of Gate Electrode in Hole Trapping

5.4.1. Shockley-Read-Hall (SRH) Analysis

To investigate the effect of N_{HT} in type-II transistors, we use a simple *elastic* trapping/de-trapping model, in which holes from the inversion layer are first trapped in the pre-existing defects within the oxide and then these trapped holes detrapp towards the empty states in poly-Si and channel. Trap filling probability (f_T) evolves with time according to the following modified version of Shockley-Read-Hall (SRH) equation:

$$\frac{df_T}{dt} = \sigma v_{th} \left[p_h T_1 (1 - f_T) - n_s T_1 f_T - n_G T_2 f_T \right], \quad (5.1)$$

where p_h is the inversion layer hole density, v_{th} is thermal velocity; n_s , n_G are concentration of detrapping states at substrate and poly-Si respectively; T_1 , T_2 are tunneling probability (obtained using WKB approximation) of holes from Si/Dielectric interface to the defect and defect to poly. Let us again emphasize that the role of detrapping towards the poly-gate – while negligible in thick films (as for the case in [219, 230]) – are fundamentally important in the trapping dynamics of thin films. The detrapping process in thin oxides limits the possibility of hole trapping in sites located near the gate – an important consideration that remains poorly appreciated in NBTI literature.

Analytical solution of (5.1) gives,

$$f_T(t) = \frac{T_1 \left[1 - \exp\left(-\sigma v_{th} (p_h T_1 + n_s T_1 + n_G T_2) t\right) \right]}{(1 + n_s/p_h) T_1 + n_G T_2 / p_h}. \quad (5.2)$$

We estimate maximum occupancy, $f_{T(\max)} = T_1 / \left[(1 + n_s/p_h) T_1 + n_G T_2 / p_h \right]$ at different trapping sites, involving elastic tunneling of carriers into and out of defects (named as *elastic traps*) with tunneling probability T_1 for trap/substrate transition and T_2 for trap/gate transition (inset of Figure 5.2), for a nitrated sample. Here, we assume $\sigma = 10^{-17} \text{ cm}^2$ (which is within the range of $\sim 10^{-13} \text{ cm}^2$ [237] and $\sim 10^{-20} \text{ cm}^2$ [233] used in literature

for bulk trapping), $p_h = 10^{20} \text{ cm}^{-3}$ slightly above the substrate Fermi level at the Si/dielectric interface (typical value in an inverted channel, verified using [178]), $n_S \sim 0$ (*i.e.*, no electrons in the substrate valence band slightly above E_{FS} , because these states are already occupied by holes), and $n_G = 10^{20} \text{ cm}^{-3}$ (approximate electron concentration in the valence band of poly within an energy range of $\sim k_B T$, which can act as elastic hole-trapping sites). Using $T_{PHY} = 37 \text{ \AA}$ (corresponding to $EOT = 24 \text{ \AA}$ for nitrated oxide having $\sim 25\%$ N_2 dose [128]), calculated $f_{T(\max)}$ for hole traps located at 20 \AA is ~ 4 order of magnitude lower compared to that for site located at 10 \AA . Hence, the time constant (or trapping saturation time, t_{sat}) for the hole trapping component of ΔV_T (ΔV_{HT}) will be dominated by the time constant of individual defects, $\tau = 1/\sigma v_{th} (p_h T_1 + n_s T_1 + n_G T_2)$, located close to the substrate.

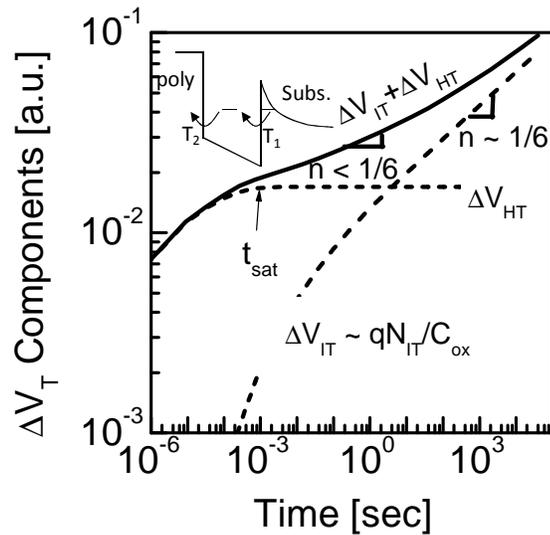


Figure 5.2: Simulated time evolution of ΔV_{IT} , ΔV_{HT} and total ΔV_T for a transistor, having thick oxide with uniform distribution of defects. Only elastic hole trapping is considered in calculation of ΔV_{HT} . As a result, hole-trapping controls total degradation at short stress time and saturates at t_{sat} . It also results in decrease in n at intermediate stress time. Inset shows tunneling probabilities T_i (into/out of defects from/towards

substrate) and T_2 (into/out of defects from/towards gate) for elastic hole trapping/detrapping model considered in the text.

5.4.2. Verification of Hole Trapping Model

We now verify the hypothesis of hole trapping discussed above by systematically exploring the thickness dependent time-exponents n (within the range of 1-10³ s) and activation energies E_A of ΔV_T for oxides with low- (type-I transistors having plasma nitrated dielectric [238, 239]) and high- (type-II transistors having thermal nitrated dielectric [239]) interfacial nitrogen. We assume that hole traps are distributed proportionally to the (presumed) N-profile (see Figure 5.3a) within the dielectric. This simple model involving elastic hole traps results saturation of ΔV_{HT} within $\sim 1\text{ms}$ (*i.e.*, $t_{sat} \sim 1\text{ms}$ in Figure 5.2). Thus overall ΔV_T shows $n < 1/6$ due to additional contribution from ΔV_{HT} within the time-scale of $\sim 1\text{-}1000$ sec. Our calculations show that *both* type-I plasma nitrated (see Figure 5.3b) and type-II thermal nitrated (see Figure 5.3c) transistors have contributions from hole trapping, but of different magnitudes. The profile decays from Si/oxide interface towards oxide/poly-Si interface because of hole detrapping into empty states of poly. For thicker oxides, the number of trapped holes increases, because detrapping to empty-states in poly is difficult. Moreover, since the centroids of the trapped charges shift away from the poly/oxide interface, in general thicker oxides have larger contributions to ΔV_T from ΔV_{HT} . However, the N-profile in the type-I plasma nitrated transistors, having peak near oxide/poly-Si interface (Figure 5.3a), makes the contribution from hole trapping in type-I transistors approximately two orders of magnitude smaller than that in type-II thermal nitrated transistors, with peak near oxide/substrate interface (Figure 5.3a). As such, the n and $E_{A,VT}$ for type-I transistors continue to be dominated by ΔV_{IT} and remains essentially unchanged with thickness (see Figure 5.4 and Figure 5.5). However, the contributions from hole trapping in type-II transistors is significant enough (Figure 5.3c) to reduce the overall ΔV_T time-exponent (Figure 5.4), activation energy (Figure 5.5), and one anticipates a reduction of both n and $E_{A,VT}$ with increasing oxide thickness. In short, the role of hole trapping in the transistors

studied is significant for any thick film with significant interfacial N concentration and must be accounted for as an additional contribution to overall NBTI degradation. So, $n < 0.1$ reported in [240] can be attributed to the existence of increased amount of defects within the oxide. In addition, this can also be due to the suppression of relative contribution from N_{IT} to ΔV_T in room-temperature.

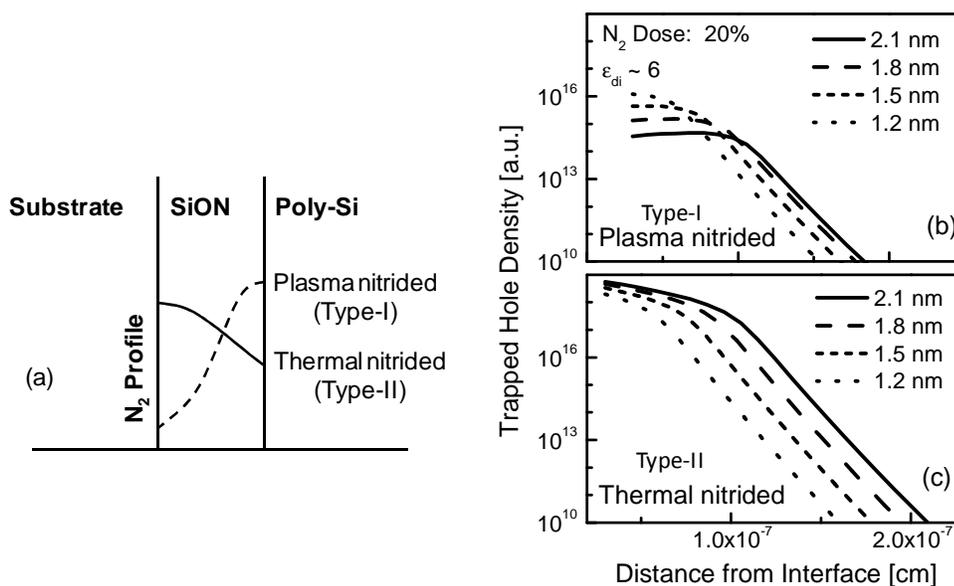


Figure 5.3: (a) Typical nitrogen profile in Plasma and Thermal nitrided gate dielectric [238, 239]. Trapped hole profile for (b) type-I plasma and (b) type-II thermal nitrided transistors, having similar %N₂ dose. Thermal nitrided oxides trap more holes compared to plasma nitrided oxides. Only elastic traps are considered in simulation.

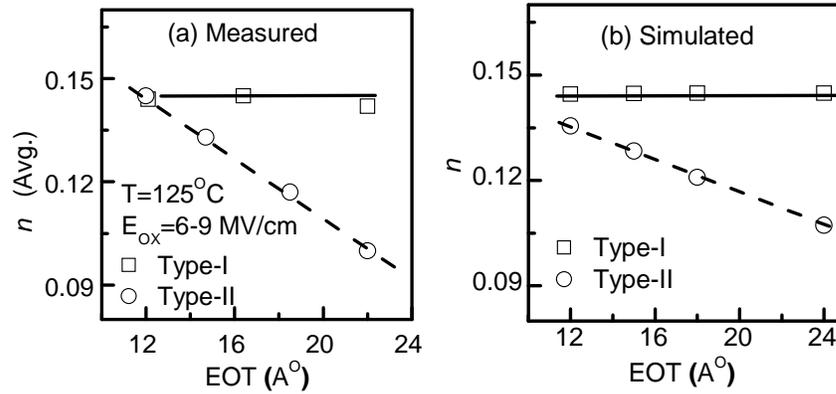


Figure 5.4: a) Experimental thickness dependence of average power law time exponents obtained from OTF- $I_{D,\text{lin}}$ in type-I plasma-nitrided transistor and type-II thermal-nitrided transistor (lines are guide to eye only). b) Simulated (involving elastic hole traps) thickness dependence under same condition. In Fig. b, it is assumed that N_{IT} -induced ΔV_{IT} in type-I transistor has $n \sim 0.145$, as observed in Fig. a.

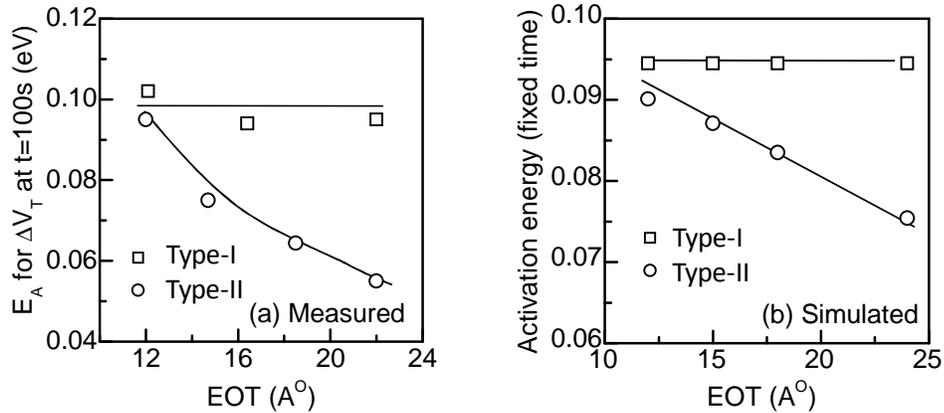


Figure 5.5: (a) EOT dependence of $E_{A,VT}$ obtained from OTF- $I_{D,\text{lin}}$ in type-I plasma-nitrided and type-II thermal-nitrided transistors (lines are guide to eye only). (b) Simulated EOT dependence of $E_{A,VT}$ for similar cases. In Fig. b, it is assumed that N_{IT} -induced ΔV_{IT} in type-I transistor has $E_A \sim 0.095$, as observed in experiments.

5.5. Detailed Hole Trapping Model

In the last section, we presented a simple *elastic* hole trapping model, which explained the thickness dependent experiments on type-I and type-II transistors. Obviously, hole trapping can also have some inelastic components [230], as shown in Figure 5.1. Here, we present a detailed hole trapping model, which contains the inelastic components, along with the contributions from gate and substrate fluxes considered in section 5.4.

5.5.1. Computational Framework

When inelastic components are added in a hole trapping-detrapping model, the time dynamics of f_T can be obtained using:

$$\frac{df_T}{dt} = F_1(1 - f_T) - F_2 f_T - F_3 f_T + F_4(1 - f_T) + F_5(1 - f_T) \quad (5.3)$$

Equation (5.3) considers contributions from hole emission flux towards substrate valence band (F_3), hole capture flux from substrate conduction band (F_4), and gate conduction band (F_5); in addition to the F_1 and F_2 fluxes, considered in equation (5.1). Moreover, hole emission towards conduction bands of substrate and gate are considered negligible. Using information related to energy band profile and Eigen-energy states (E_{ij}) from a QM-CV simulator [178], we can calculate the above fluxes using –

$$F_1 = \begin{cases} \sum_j \sigma v_{th} T_{ij} \frac{m_j^* kT}{\pi \hbar^2} \ln \left[1 + \exp \frac{E_{FS} - E_{ij}}{k_B T} \right] & \text{for } E_T > E_{ij} \\ \sum_j \sigma v_{th} T_{ij} \frac{m_j^* kT}{\pi \hbar^2} \ln \left[1 + \exp \frac{E_{FS} - E_{ij}}{k_B T} \right] \exp \frac{-(E_T - E_{ij})}{k_B T} & \text{for } E_T < E_{ij} \end{cases} \quad (5.4)$$

$$F_2 = \begin{cases} \sigma v_{th} T_2 n_G & \text{for } E_T < E_{VG} \\ \sigma v_{th} T_2 n_G \exp \left[-(E_T - E_{VG}) / k_B T \right] & \text{for } E_T > E_{VG} \end{cases} \quad (5.5)$$

$$\mathbf{F}_3 = \begin{cases} \sum_j \sigma v_{th} T_{ij} \frac{m_j^* kT}{\pi \hbar^2} \ln \left[1 + \exp \frac{E_{ij} - E_{FS}}{k_B T} \right] \exp \frac{-(2E_{ij} - E_{FS} - E_T)}{k_B T} & \text{for } E_t > E_{ij} \\ \sum_j \sigma v_{th} T_{ij} \frac{m_j^* kT}{\pi \hbar^2} \left[\ln \left(\frac{1 + \exp \frac{E_{FS} - E_t}{k_B T}}{1 + \exp \frac{E_{FS} - E_{ij}}{k_B T}} \right) + \ln \left(\frac{1 + \exp \frac{E_T - E_{FS}}{k_B T}}{1 + \exp \frac{E_T - E_{FS}}{k_B T}} \right) \exp \frac{-(E_T - E_{FS})}{k_B T} \right] & \text{for } E_t < E_{ij} \end{cases} \quad (5.6)$$

$$\mathbf{F}_4 = \begin{cases} \sigma v_{th} T_1 p_S & \text{for } E_T > E_{CS} \\ \sigma v_{th} T_1 p_S \exp \left[(E_T - E_{CS}) / k_B T \right] & \text{for } E_T < E_{CS} \end{cases} \quad (5.7)$$

$$\mathbf{F}_5 = \begin{cases} \sigma v_{th} T_2 p_G & \text{for } E_T \geq E_{CG} \\ \sigma v_{th} T_2 p_G \exp \left[(E_T - E_{CG}) / k_B T \right] & \text{for } E_T < E_{CG} \end{cases} \quad (5.8)$$

where the related additional variables (*i.e.*, addition to the ones used in equation (5.1)) are defined in Figure 5.6. In addition, T_{ij} indicates tunneling co-efficients calculated from E_{ij} energy levels (for the j -th sub-band in the i -th valley), m_j^* is the in-plane effective mass of the j -th subband, p_S is the hole concentration in the substrate conduction band, and p_G is the hole concentration in the gate conduction band.

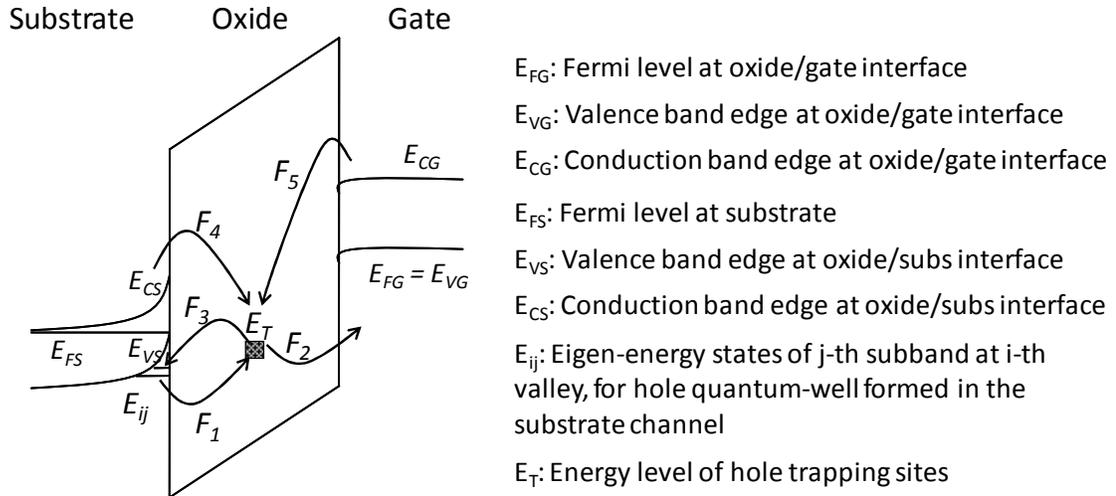


Figure 5.6: Carrier fluxes considered in the hole trapping model of equation (5.3). Related additional variables, over the one used in equation (5.1), are also defined.

5.5.2. Simulation Results

Using equations (5.3)-(5.8), we simulate hole trapping in a type-II transistor (thermal nitrided dielectric; $T_{PHY} = 22.97\text{\AA}$; $EOT = 13.5\text{\AA}$; %N = 39.22 using XPS). Pre-existing oxide defects are considered within a limited region of the dielectric (see Figure 5.7b, c) and outside these regions defects are observed to be insensitive to any variations in voltage or temperature. Now for simulating hole trapping dynamics, we first match experimental C_G-V_G for this transistor using QM-CV simulator [178], thus calculate the carrier fluxes using equations (5.4)-(5.8). Next we solve equation (5.3) for estimating occupancy of hole trapping sites at different time steps of the timing diagram in Figure 5.7a. The trap occupancy at two different time intervals, *i.e.*, after the application of $V_G = -2.1\text{V}$ (Figure 5.7b) and -1.0V (Figure 5.7c) for 1000sec, we observe the following main features of trap occupancy (f_T):

- f_T near the substrate and gate is dominated by E_{FS} and E_{FG} , respectively.
- In the intermediate region (*i.e.*, for $x = 0$ to T_{PHY}) and at a certain energy (E_T), f_T is limited by tunneling of carriers into and out of the trapping site

- Defects at energies much lower than E_{VS} or much higher than E_{CS} are either completely empty of holes (*i.e.*, filled with electrons) or completely filled with holes (*i.e.*, empty of electrons), respectively.

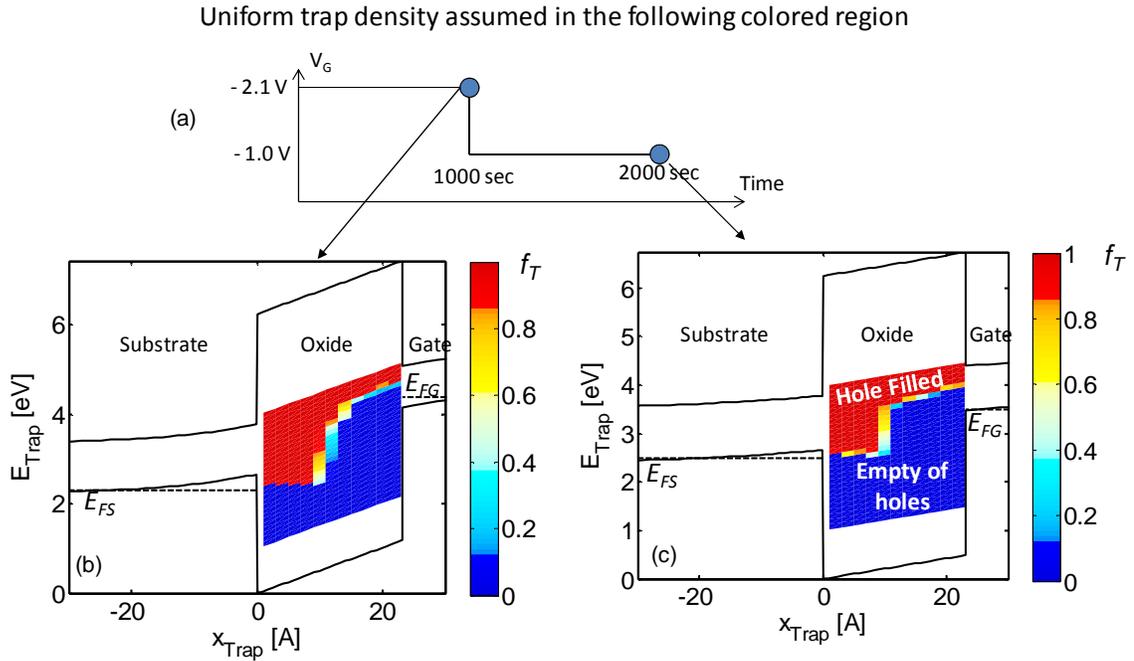


Figure 5.7: (a) Timing diagram used in the hole trapping simulation using the framework of section 5.5.1. (b) Occupancy of defects or, hole trapping sites (assumed to have existence within the colored region only), after the transistor is kept at $V_G = -2.1\text{ V}$ for 1000 sec . (c) Occupancy of hole trapping sites after the transistor is switched to $V_G = -1.0\text{ V}$ at 1000 sec and kept at that voltage up to 2000 sec .

Next, we study the voltage and temperature dependencies of the hole trapping process. Our simulation shows that although hole trapping is voltage (or electric field) dependent (Figure 5.8a), it is entirely temperature independent (Figure 5.8b). Voltage (negative) or field dependency of hole trapping arises from the fact that when $|E_{ox}|$ is increased, more trapping sites will go above E_{FS} (see the change of trapping sites f_T values from Figure 5.7b to Figure 5.7a), thus contribute to ΔV_{HT} . Moreover, the

observation of temperature invariant of hole trapping is consistent with our experimental study (see section 6.7.5), which is also observed experimentally by several other groups [140, 241]. However, there are reports of temperature dependent electron trapping in PBTI experiments [47, 242], the origin of which is still unknown. The final important outcome of our simulation is the fact that response of hole trapping sites (*i.e.*, the increase of ΔV_H with time) is extremely fast and it saturates within $\sim \mu\text{s}$, at any voltage and temperature used in simulation (Figure 5.8).

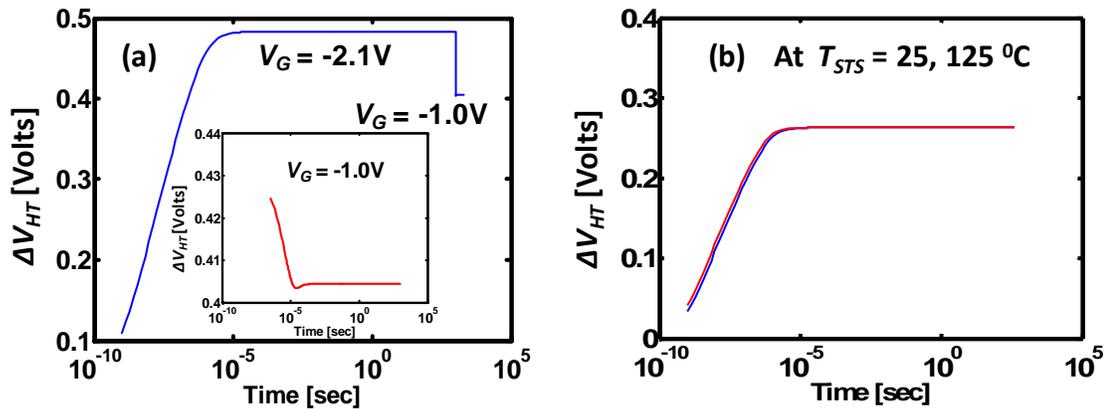


Figure 5.8: (a) Voltage and (b) Temperature dependent simulation of hole trapping. Although, hole trapping shows significant voltage or electric field dependency, it is independent of temperature.

5.6. Summary

Our simulation of hole trapping, which involves an integration of QM-CV simulator [178] and modified SRH modeling, enables us to make the following remarks:

- Hole trapping within a ultra-scaled thin dielectric used in today's CMOS technology is fundamentally different from those studied previously for thick dielectric.

- Trapping into pre-existing defects in case of thin dielectric transistors should saturate quickly within $\sim \mu\text{s}$.
- Our study of trapping clearly verifies the voltage dependence and temperature independence of hole trapping, observed in experiments.

6. DEFECT CHARACTERIZATION

6.1. Introduction

In this chapter, we analyze the techniques that are recently being used for characterizing interface/bulk defects of a transistor. In section 6.2, we present the importance of measurement duration in the defect characterization techniques and thus identify their suitability for different purposes. Later in sections 6.3-6.6, we explain the measurement procedure and commonly extracted parameters in these techniques, currently used for characterizing interface/bulk defects of a transistor. Then in section 6.7, we identify the correction steps needed for proper interpretation of measured quantities in different defect characterization techniques. As such, we can extract information related to time dynamics of defect generation, its physical nature and its impact on observables. More specifically, we highlight the importance of considering corrections related to (i) body effect co-efficient (m), (ii) valence band electron trapping in Ultra-fast V_G (UFV) measurement [140], and (iii) time-zero (t_0) delay, (iv) effective mobility variation ($\Delta\mu_{eff}$) in On-the-fly $I_{D,lin}$ (OTF- $I_{D,lin}$) measurement [129, 130, 139, 141, 143]. Later in section 6.7.5, we present a novel methodology for separating contributions of interface defects and oxide defects by studying defect generation over a wide variety of transistors having plasma-nitrided dielectric. We explain how the proposed separation methodology can significantly resolve the issues raised in section 3.6.5 for explaining NBTI relaxation experiments using R-D theory. And finally in section 6.8, we use our understanding to propose a guideline for appropriate defect characterization that can be extended for future technologies.

6.2. Defect Characterization Techniques

Measurement duration (t_{meas}) is an important criterion in defect characterization. Since generated defects relaxes or recovers within a short time (see section 3.6 for relaxation of interfacial defects, N_{IT} and section 5.5 for relaxation of oxide hole trapping, N_{HT}), capturing the time dynamics of defect formation requires one to use a characterization technique, having $t_{meas} \sim 0$. However, for the identification of the nature/properties of defects and finding its impact on other transistor parameters like sub-threshold slope (SS), threshold voltage (V_T), effective mobility (μ_{eff}), *etc.* t_{meas} does not need to be that small. In general, t_{meas} for different defect characterization techniques vary widely from micro-seconds to several minutes. For example,

Electron Spin Resonance (ESR) technique [76, 86] and its electrically detectable version (EDMR: Electrically Detected Magnetic Resonance) [88, 243-248] allow one to identify the physical and chemical nature of defects present at the oxide/semiconductor interface or within the oxide bulk. Since, measurement steps in ESR experiments are quite sophisticated, t_{meas} is normally several minutes [88] and study of exact time dynamics of defect generation (which relaxes quickly after removal of stress, like the case for interface defects, see section 3.6) is not possible using this technique.

Linear I_D - V_G characterization [249] enables one to determine parameters like SS , V_T , and μ_{eff} and requires sufficient time (order of seconds) for stable measurements. Taking such $I_{D,lin}$ - V_G sweeps at different stress levels [89, 203, 250], known as **Measure-Stress-Measure or, MSM** technique enables one to estimate the impact of defect generation on SS , V_T , and μ_{eff} . Measurement duration in MSM can be reduced to orders of milliseconds by limiting the sweep interval around V_T (hence, avoiding the determination of SS and μ_{eff}) [33]. Alternately, one can also employ an **ultra-fast $I_{D,lin}$ - V_G (UFIV)** scheme [141, 251] of taking $I_{D,lin}$ - V_G sweeps at or above V_T within $t_{meas} \leq 1\mu s$, thus avoiding the determination of SS in such process.

Ultra-fast V_G measurement technique, having $t_{meas} \sim \mu s$, has recently been proposed [140] for characterizing defect generation through an approximate

measurement of V_T degradation (ΔV_T) by monitoring shift of $I_{D,lin}$ - V_G plot (*i.e.*, degradation in V_G) at constant I_D .

On-the-fly (OTF) defect characterization are performed, through an indirect estimation, without any interruption of stress (*i.e.*, $t_{meas} \sim 0$). OTF techniques avoid the use of time-consuming I_D - V_G sweeps for transistor's V_T characterization and indirectly estimate ΔV_T (or the signature of defect formation) by monitoring – (a) $I_{D,lin}$ degradation ($\Delta I_{D,lin}$) at constant stress V_G [129, 130, 139, 141, 143], or (b) $g_m/I_{D,lin}$ degradation at constant stress V_G [137, 252]. If the ΔV_T estimation is perfect, these OTF techniques are ideal for obtaining time dynamics of defect's generation and recovery.

In the next few sections (6.3-6.6), we explain the measurement procedure and commonly extracted parameters in the above techniques. Then in section 6.7, we identify the correction steps needed, in general, for proper interpretation of measured quantities.

6.3. Spin Resonance Techniques ($t_{meas} \sim$ minutes)

6.3.1. Electron Spin Resonance (ESR)

ESR (also called Electron Paramagnetic Resonance, EPR) experiment was first reported in [76] for characterizing metal-oxide-semiconductor structure. Subsequently, different research groups [68, 253-256] experimentally demonstrated the existence of paramagnetic defects near the Si/SiO₂ interface (known as P_b centers), as well as within the bulk of oxide (known as E' centers). In conventional ESR measurement (Figure 6.1), the transistor under study is placed between a large magnet at low temperature, whose magnetic field (B) is varied at a comparatively slower rate. Application of B splits the normally (for $B = 0$) degenerate energy levels of unpaired electron within the defect by $g\mu_B B$; where, g is spectroscopic splitting factor [24], $\mu_B = qh/4\pi m_e$ is Bohr magneton, q is the electron charge, h is the Planck's constant and m_e is the electron mass. Here, electron having spin, $S_e = -1/2$ constitutes the lower energy state and can readily absorb an

electromagnetic signal (usually have microwave frequency, $\nu \sim 10$'s of GHz and is applied in a direction perpendicular to B), provided $h\nu = g\mu_B B$ (*i.e.*, resonance) condition is satisfied. Therefore, the transistor under test absorbs larger amount of electromagnetic signal at resonance, where the first derivative of absorption spectrum indicates a zero-crossing or peak in the absorption line. Finally, knowledge of ν and B at resonance allows one to estimate g , whose value for different types of defects are tabulated in [86].

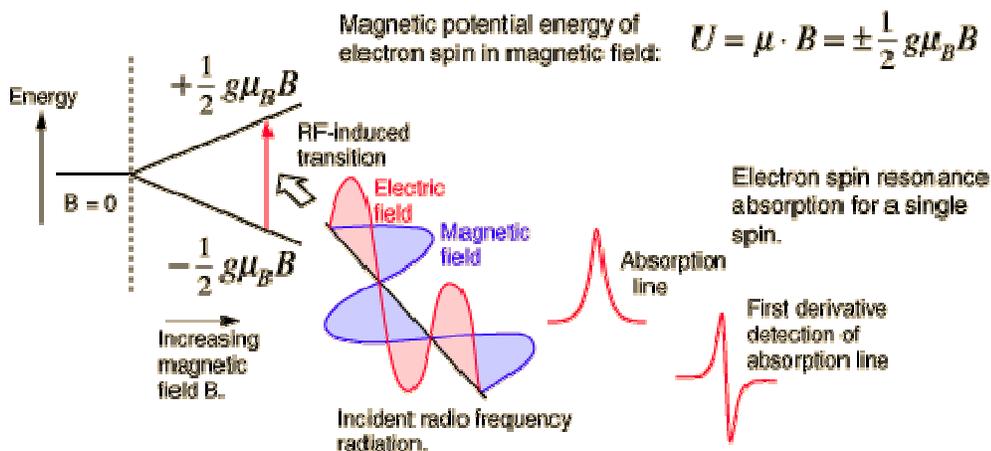


Figure 6.1: Basic principles for obtaining resonance in an ESR/EPR experiment. Device under study is placed in a slowly varying magnetic field (B), along with an electromagnetic signal (having frequency, ν) applied perpendicular to B . At resonance, there is significant absorption of electromagnetic signal. From the peak position of the absorption line (zero-crossing point of the first derivative) one estimates g , thus the nature of defect. This figure is taken from <http://images.google.com>.

6.3.2. Electrically Detected Magnetic Resonance (EDMR)

Conventional ESR spectrometers are sensitive to defect density of $\geq 10^{11}/\text{cm}^2$ [257] and also requires large sample area ($\sim 1 \text{ cm}^2$) for obtaining detectable absorption of the electromagnetic signal [258]. Thus, ordinary ESR measurement on industry standard MOSFET's, which normally have defect density less than $10^{11}/\text{cm}^2$ and effective transistor area in the range of $\sim \mu\text{m}^2$, becomes impossible. Optically Detected Magnetic

Resonance (ODMR) [259] and Electrically Detected Magnetic Resonance (EDMR) [88, 244, 257] overcome this problem by providing higher spin sensitivities compared to conventional ESR by detecting electronic transition at resonance either optically (in ODMR through luminescence) or electrically (in EDMR through source/drain currents of a transistor). Among these two techniques, EDMR has recently been extensively used for characterizing interface/bulk defects in MOSFET structures, subjected to Bias Temperature Instability (BTI) [88, 260], Hot Carrier Injection (HCI) [261] and radiation damage [257, 262].

In EDMR experiment, a MOSFET is connected as gate-controlled diode configuration [263] (*i.e.*, source/drain shorted together, see Figure 6.2a). The p-n junction of the gated-diode is slightly forward biased ($V_F > 0$), which splits the quasi-Fermi levels of channel electrons and holes (F_N and F_P respectively) near the source/drain (Figure 6.2b,c). As a result, electrons coming from n⁺ substrate contact and holes coming from p⁺ source/drain recombine through interface/near-interface defects. The recombination current, measured at the substrate terminal (I_{sub}), varies with applied gate bias (V_G) and reaches maximum, when the concentration of electron and hole near the recombination center are similar (which is the case during depletion mode of a MOSFET, see Figure 6.2b). Application of a slowly varying magnetic field (B), along with electromagnetic radiation having microwave frequency (ν), increases I_{sub} significantly at resonance (*i.e.*, for $h\nu = g\mu_B B$). The exact mechanism behind the increase in the spin-dependent I_{sub} at resonance is still debated [243-245, 248, 264]. However, the explanation involving singlet-triplet formation during the recombination and magnetic field dependency of the triplet recombination is generally used for explaining EDMR. Moreover, since EDMR measures the recombination current of a MOSFET in gate-diode configuration and this current depends on spin configurations of electrons, holes and traps taking part in recombination, EDMR is also widely known as Spin Dependent Recombination (SDR) technique.

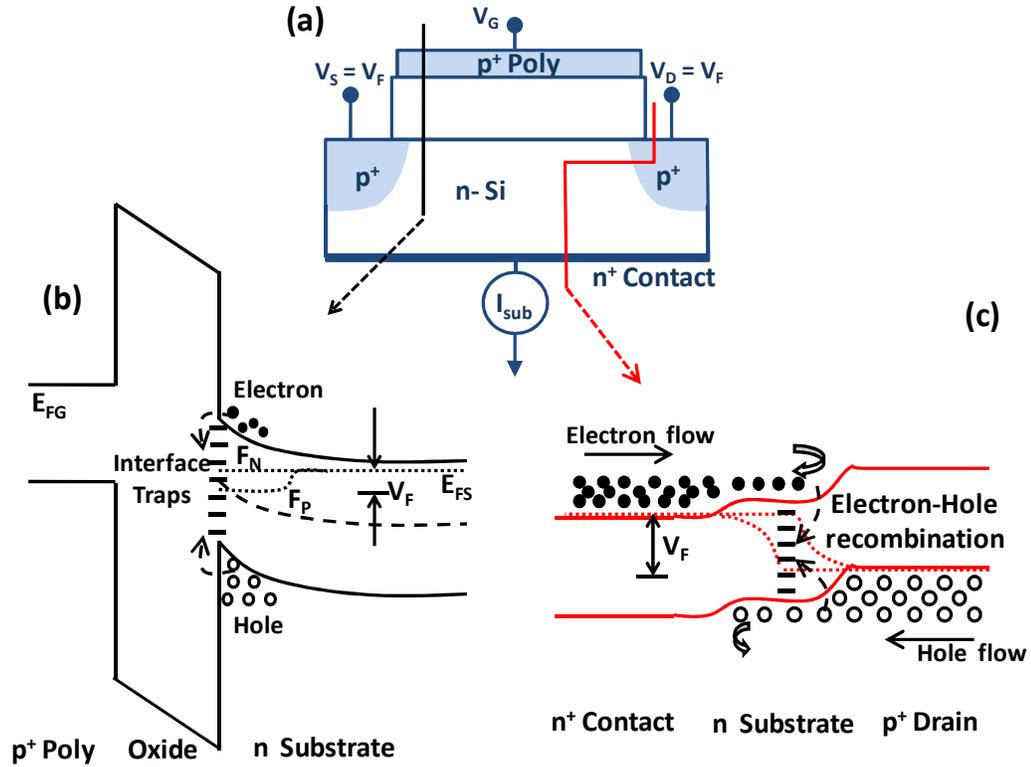


Figure 6.2: (a) Schematic of PMOS connected in gated-diode configuration for EDMR or SDR measurement. Here, a forward bias (V_F) is applied to shorted source/drain and a change in substrate current (I_{sub}) at resonance (i.e. at $h\nu = g\mu_B B$) is monitored. (b,c) Band diagrams are drawn respectively along gate/oxide/substrate and drain/substrate/substrate-contact directions.

6.4. Measure-Stress-Measure (MSM) Technique ($t_{meas} \sim \text{msec-sec}$)

Defect characterization using MSM technique starts by taking $I_{D,lin}(0)-V_G(0)$ sweep ($t_{meas} \sim 2$ sec in our MSM setup), and then stressing the transistor for a pre-defined duration ($t_{STS,i}$). Subsequently, the stress is interrupted for t_{meas} to take another $I_{D,lin}(i)-V_G(i)$ sweep. So, the measurement of $I_{D,lin}-V_G$ and application of stress continues for the desired period of defect characterization. Later on, the $I_{D,lin}(i)-V_G(i)$ sweeps taken at different stress time, $t_{STS,i}$ (where, $i = 0, 1, 2, 3, \dots$ etc., with $t_{STS,0} = 0$), are used for determining $V_T(i)$, $SS(i)$ (hence, $m(i)$) and $\mu_{eff}(i)@V_G$, using the procedures mentioned below. Changes in these parameters illustrate the impact of defect formation on

transistor's performance. An empirical mobility modeling of μ_{eff} [89, 165, 249, 265] and hence determining the changes in the model parameters are also useful for better understanding of μ_{eff} variation due to defect generation. Moreover, a physical characterization of μ_{eff} mandates one to estimate effective electric field (E_{eff}) and thus study changes in the physical model [266] parameters.

The timing diagram for MSM technique is shown in Figure 6.3a. Here, we apply MSM for observing the impact of N_{IT} formation during NBTI stress in type-I PMOS transistors (where, $\Delta V_T \approx \Delta V_{IT} \sim q\Delta N_{IT}/C_{di}$), having lightly dosed plasma-oxynitride (PNO) dielectric. Table 6.1 provides the list of type-I transistors used in this study. As schematically shown in Figure 6.3b, generation of N_{IT} shifts the $I_{D,lin}-V_G$ characteristics in negative direction, which indicates an increase in $|V_T|$, a decrease in $I_{D,lin}$, and also change in other parameters *e.g.*, μ_{eff} , SS , m .

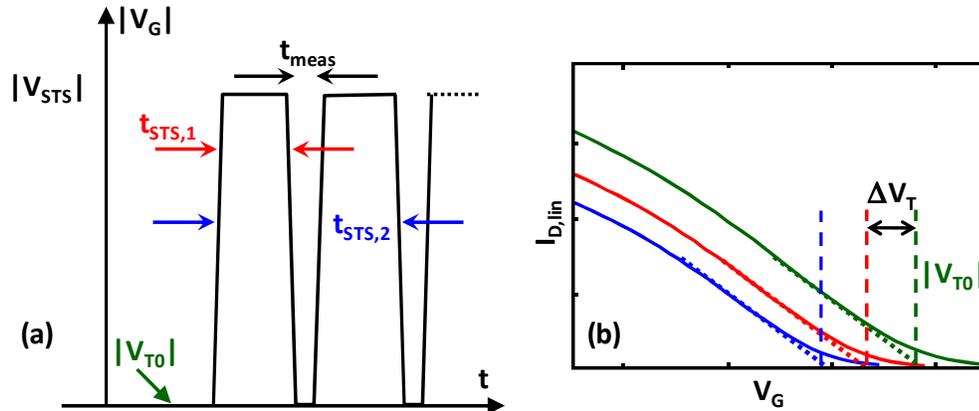


Figure 6.3: (a) Timing diagram for MSM experiment on a PMOS transistor for characterizing parametric degradation during NBTI stress. After taking initial $I_{D,lin}-V_G$ at $t_{STS,0}$, stress voltage (V_{STS}) is applied. The stress is later interrupted at $t_{STS,1}$, $t_{STS,2}$, ..., *etc.* time intervals for taking subsequent $I_{D,lin}-V_G$ sweeps. (b) Under NBTI stress, $I_{D,lin}-V_G$ shifts in negative direction, which indicates an increase in $|V_T|$, a decrease in $I_{D,lin}$, and also change in other parameters *e.g.*, μ_{eff} , SS , m .

Table 6.1: Properties of plasma oxynitride dielectric of the type-I transistors used in sections 6.4, 6.7.1, 6.7.3, and 6.7.4: N₂ Dose (%N), physical oxide thickness (T_{PHY}), effective oxide thickness (EOT)

| Device | N ₂ Dose [%] | T _{PHY} [Å] | EOT [Å] |
|--------|-------------------------|----------------------|---------|
| #1 | 22.64 | 18.48 | 14.0 |
| #2 | 19.45 | 22.34 | 17.7 |
| #3 | 16.67 | 28.13 | 23.5 |
| #4 | 10.22 | 18.60 | 15.1 |
| #5 | 16.55 | 18.60 | 14.6 |

6.4.1. Determination of E_{eff}

Characterization of μ_{eff} and its changes with N_{IT} require us to determine $E_{eff}(V_G)$ at different time interval $t_{STS,i}$ by using [249, 267],

$$E_{eff} = (Q_{dep} + Q_{inv} / 3) / \epsilon_{Si} \sim \frac{V_T - V_{FB} - 2\psi_B}{3T_{ox}} + \frac{(V_G - V_T)}{9T_{ox}}, \quad (6.1)$$

where Q_{dep} is the depletion charge, Q_{inv} is the inversion charge (or holes), ϵ_{Si} is relative dielectric constant for Si, and ψ_B is the difference between the substrate Fermi level and the intrinsic Fermi level of silicon. Determination of E_{eff} in pre-stress ($t_{STS,0}$) condition involves calibration of QM-CV simulator [178] with pre-stress C_G - V_G data (Figure 6.4a) and then using this calibrated simulator to estimate inversion ($Q_{inv(t=0)}$) and depletion ($Q_{dep(t=0)}$) charges. Later, E_{eff} in post-stress ($t_{STS} > 0$) condition is estimated by incorporating the effect of uniform N_{IT} generation within the band-gap. Presence of donor-type N_{IT} due to NBTI stress over the entire silicon bandgap [90] shifts CV curve towards left (Figure 6.4a), thereby reduces surface potential, ψ_S (Figure 6.4b) (with corresponding reduction in Q_{dep} , Q_{inv} and E_{eff}) at a particular V_G .

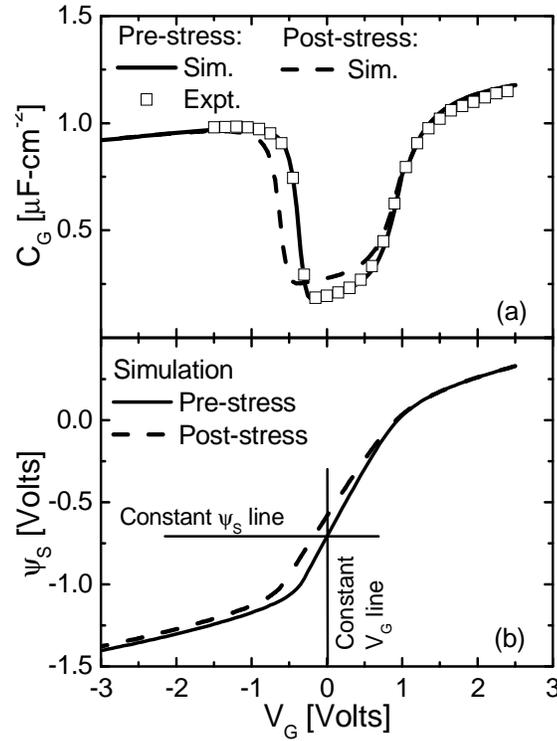


Figure 6.4: Uniform generation of donor-type N_{IT} causes left-shift of C_G - V_G . Pre-stress C_G - V_G is calibrated with experimental data. (b) Relationship between ψ_s and V_G before and after stress. Interface defect causes ψ_s to decrease at constant V_G (movement along vertical line). Increase in V_G during NBTI degradation (moving along horizontal line) can keep ψ_s constant (needed in determining Δm) with degradation. In this figure, we use device #3 of Table 6.1 for illustrating the impact of N_{IT} generation.

6.4.2. Determination of V_T , SS , m , μ_{eff}

We use $I_{D,lin}(i)$ - $V_G(i)$ characteristics, obtained at different $t_{STS,i}$ during MSM experiment, to determine $V_T(i)$, $SS(i)$, $m(i)$, $\mu_{eff}(i)$ @ V_G , and its changes with N_{IT} generation. First, we estimate $m(i)$ at specific V_G (or $V_{G,m}$) by using –

$$SS(i) = 2.3m(i)k_B T / q. \quad (6.2)$$

Now, assuming uniform N_{IT} generation within the bandgap, we have

$$m(\psi_s) = 1 + \left[C_D(\psi_s) + C_{IT} \right] / C_{ox}, \quad (6.3)$$

where C_{ox} , C_D and C_{IT} are oxide, depletion and interface trap capacitances respectively. As we desire changes in m with N_{IT} only (through C_{IT}), possible variation in C_D is suppressed by estimating m at constant ψ_s (horizontal line in Figure 6.4b). Next, $V_T(i)$ is calculated from the x-axis intercept (V_X) of the slope on $I_{D,lin}(i)$ - $V_G(i)$ curve, drawn at maximum transconductance ($g_{m,max}$) point (Figure 6.3b) and using $|V_T| = |V_X| + m|V_{DS}|/2$ [249]. Estimated $V_T(i)$ and $m(i)$ ¹² enable us to determine $\mu_{eff}(i)$ using the following expression, commonly used for modeling $I_{D,lin}$ - V_G at $|V_G| \gg |V_T|$:

$$|I_{D,lin}| = \mu_{eff} C_G W/L \left(V_G - V_T - 0.5mV_{DS,lin} \right) V_{DS,lin}. \quad (6.4)$$

Using $m(i)$, we estimate $\Delta m = m(i) - m(0)$, which increases with NBTI and shows good correlation with ΔV_T for all the transistors used in this study (Figure 6.6).

6.4.3. Variation in μ_{eff} Model Parameters

To investigate the effect of N_{IT} on μ_{eff} , we interpret the measured μ_{eff} (section 6.4.2) through a well-known empirical mobility model. This mobility model

$$\mu_{eff} = \mu_0 \left[1 + \theta (V_G - V_T)^\eta \right], \quad (6.5)$$

(where μ_0 , θ , η are empirical constants) is appropriate for $|V_G| \gg |V_T|$ and has been widely used in circuit simulations [89, 249, 265, 268]. We de-embed the constants μ_0 and θ by fitting μ_{eff} vs. $(V_G - V_T)$ in both pre- and post- NBTI stress conditions (Figure 6.5). Thus, we determine $\Delta\mu_0 = \mu_0(t>0) - \mu_0(t=0)$ and $\Delta\theta = \theta(t>0) - \theta(t=0)$, which show systematic variation with ΔV_T for all the transistors studied here (see Figure 6.6).

¹² Note that we have used a constant m , calculated from sub-threshold characteristics, in estimating μ_{eff} above threshold using equation (6.4). Thus, any variation in m from sub-threshold to super-threshold (according to equation (6.3)) is ignored here.

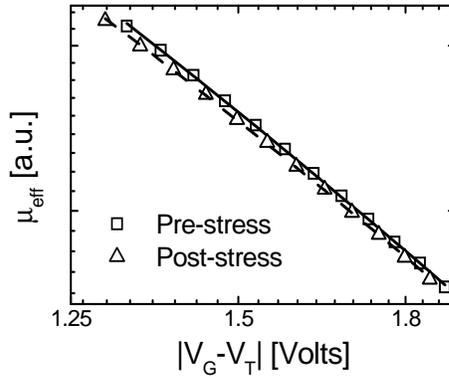


Figure 6.5: Modeling μ_{eff} vs. $(V_G - V_T)$ using empirical mobility model of equation (6.5) with $\eta = 1$ before ($N_{IT} = 0$) and after ($N_{IT} > 0$) NBTI degradation. Here, device #5 of Table 6.1 is used.

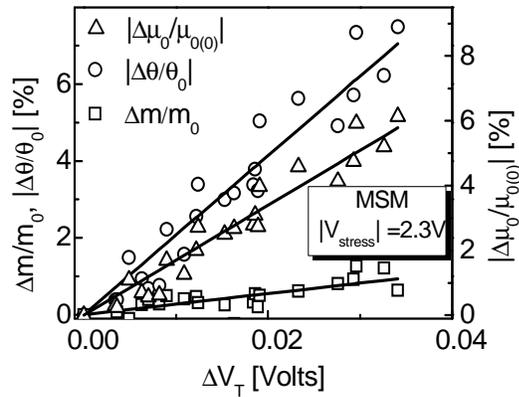


Figure 6.6: Variation of mobility parameters ($\Delta\mu_0$ and $\Delta\theta$), obtained using equation (6.5) with $\eta = 1$ (Figure 6.5), and Δm have good correlation with ΔV_T . Here, device #5 of Table 6.1 is used.

Next, estimated variation of $\Delta\mu_0(\Delta V_T)$ and $\Delta\theta(\Delta V_T)$ are used to evaluate $\Delta\mu_{eff}(\Delta V_T)$ using equation (6.5) at different gate voltage (V_{meas}) (shown for transistors having $\theta_0 = 1.34 \text{ V}^{-1}$ and $\theta_0 = 0.8 \text{ V}^{-1}$ in Figure 6.7a and Figure 6.7b respectively). Though interface defect always causes a reduction of μ_{eff} due to enhanced coulomb scattering [89], when measured at constant E_{eff} (Figure 6.7c), we observe both an increase/decrease in μ_{eff} at

constant V_G , depending on the θ_0 of the transistor under study (see Figure 6.7a, b). The puzzling increase of μ_{eff} with ΔV_T for Figure 6.7a is easily explained (using equation (6.5)) by realizing that μ_0 degradation due to N_{IT} ($\Delta\mu_0$) is over-compensated by reduced θ ($\Delta\theta$) and increased ΔV_T (*i.e.*, decreased E_{eff} [165, 269]). On the other hand, such compensation is not present in transistors with lower θ_0 (Figure 6.7b) – which shows a degradation of μ_{eff} at constant V_{meas} . In sum, although $\Delta\mu_{eff}$ could either be positive (Figure 6.7a) or negative (Figure 6.7b) when measured at constant V_G (V_{meas}), it is always negative [89] when measured at constant E_{eff} (Figure 6.7c).

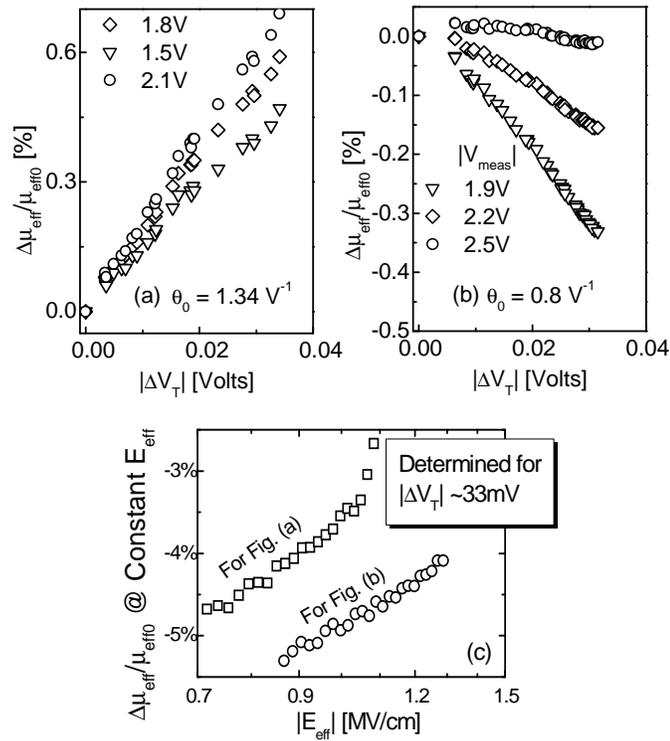


Figure 6.7: μ_{eff} (a) increases for device #5 of Table 6.1 (having $\theta_0 = 1.34\text{V}^{-1}$) or (b) decreases for device #2 of Table 6.1 (having $\theta_0 = 0.8\text{V}^{-1}$) during NBTI when measured at constant gate voltage (V_{meas}), although it shows (c) a decreasing trend when measured at constant E_{eff} .

6.4.4. Determining Physical Basis of $\Delta\mu_{eff}(N_{IT})$

MSM experiment not only allows us to analyze the influence of interfacial defects (generated during NBTI stress on type-I transistors) on the empirical mobility model parameters of equation (6.5), it also helps us to obtain a physical understanding for the μ_{eff} variation. To explore the physical basis of $\Delta\mu_{eff}(N_{IT})$, here we use a comparatively more rigorous mobility model [266, 267, 270, 271], where three dominant components of carrier scattering – phonon scattering, surface roughness scattering, and coulomb scattering both due to positively charged ionized donors and interface defects – are explicitly accounted for. Using Mathiessen’s rule¹³, μ_{eff} is then expressed as –

$$\frac{1}{\mu_{eff}} = \frac{1}{\mu_{coul}} + \frac{1}{\mu_{ph}} + \frac{1}{\mu_{sr}} \quad (6.6)$$

where

$$\mu_{coul} = \mu_1 Q_{inv}^{\alpha_1} / (Q_{dep} + \beta_{IT} Q_{inv} N_{IT}), \quad (6.7)$$

$$\mu_{ph} = \mu_2 E_{eff}^{\alpha_2}, \text{ and } \mu_{sr} = \mu_3 E_{eff}^{\alpha_3}.$$

We use $\alpha_1 = 1$, $\alpha_2 = -0.3$ [267, 270], $\alpha_3 = -1.5$ and the remaining coefficients (μ_1 , μ_2 , μ_3 , β_{IT}) as fitting parameters. Pre-stress μ_{eff} vs. E_{eff} relationship (solid line in Figure 6.8) is used to estimate μ_1 , μ_2 , and μ_3 . Later, β_{IT} (proportional to the square of the hole wavefunction at the location of interface states [266]) is determined by fitting post-NBTI stress μ_{eff} vs. E_{eff} data (dashed line in Figure 6.8). Hence, by using a more physically based mobility model (*i.e.*, using equations (6.6)-(6.7)), we can characterize the effects of interface defects on $\Delta\mu_{eff}(\Delta N_{IT})$ using a single physical parameter, β_{IT} .

¹³ Validity of Mathiessen’s rule is based on the assumption that independent scattering mechanisms have similar energy dependence, which is seldom justified in practice [272]. Nevertheless, it is often used in literature [266, 267, 270, 271] for self-consistent mobility calculation for different scattering mechanisms, which is also appropriate in our case.

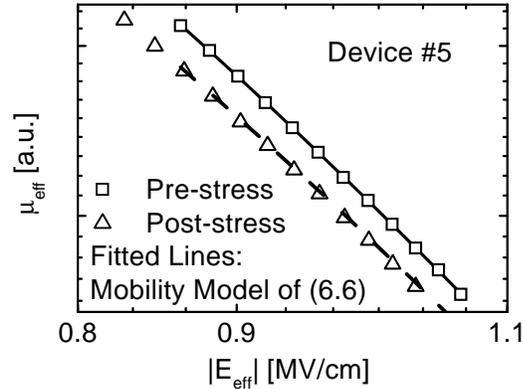


Figure 6.8: Modeling μ_{eff} vs. E_{eff} using equations (6.6)-(6.7) in both pre- and post- stress conditions for device #5 of Table 6.1. This modeling approach is used for identifying the physical basis of $\Delta\mu_{eff}(\Delta N_{IT})$.

6.5. Ultra-Fast V_G or UFV ($t_{meas} \sim 0.5 \mu s$)

UFV technique is recently proposed [140] for characterizing defect generation through an approximate measurement of ΔV_T by monitoring shift of $I_{D,lin}-V_G$ plot at constant I_D . Timing diagram for UFV experiment applied to a PMOS transistor for estimating ΔV_T during NBTI stress is shown in Figure 6.9a. After choosing a particular current level, $I_{D,lin(UFV)} = 10^{-5}$ A in Figure 6.9b, one estimates V_{Gi} required for having $I_{D,lin(UFV)}$ by interrupting V_{STS} (for $t_{meas} \sim \mu s$ [140]) at different stress interval ($t_{STS,i}$; where, $i = 0, 1, 2, 3, \dots$ etc. with $t_{STS,0} = 0$). Later, $\Delta V_T(i)$ is assumed to be equal to $\Delta V_G(i) = V_{Gi} - V_{G0}$ [140]. Moreover, UFV technique supports NBTI relaxation measurement for any recovery voltage (V_{REC}), provided appropriate corrections are made (see sections 6.7.1 and 6.7.2), thus offers an advantage over the OTF- $I_{D,lin}$ technique (discussed in section 6.6) in relaxation measurements, which only allows NBTI relaxation measurement for $|V_{REC}| > |V_T|$.

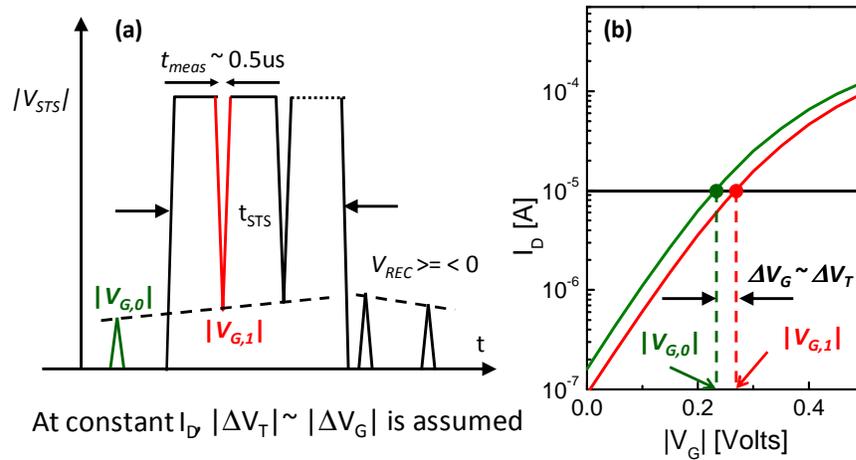


Figure 6.9: (a) Timing diagram for UFV experiment applied to a PMOS transistor for estimating ΔV_T during NBTI stress [140]. (b) In UFV, one chooses a particular current level, $I_{D,lin(UFV)}$ and then estimate V_{Gi} required for having $I_{D,lin(UFV)}$ by interrupting V_{STS} at different stress interval. Generation of interface/bulk trap during NBTI increases $|V_{Gi}|$ within time. Later, $\Delta V_T(i)$ is assumed to be equal to $\Delta V_G(i) = V_{Gi} - V_{G0}$. Recovery voltage (V_{REC}) in UFV can have any magnitude and polarity.

6.6. On-the-fly $I_{D,lin}$ or OTF- $I_{D,lin}$ ($t_{meas} \sim 0$)

OTF- $I_{D,lin}$ is one of the most widely used defect characterization technique now-a-days [128-130, 141, 143, 154, 169], because it eliminates the possibility of undesirable recovery during characterization, which is crucial requirement for obtaining exact time dynamics of defects (see Figure 6.10 for the timing diagram). OTF- $I_{D,lin}$ starts by taking an initial short-range $I_{D,lin}-V_G$ (i.e., V_G is chosen to have a range just around V_T), enough for estimating V_{T0} by drawing slope at maximum transconductance point. Later, V_{STS} is applied and $I_{D,lin0}$ is measured after t_0 (≥ 1 ms for standard OTF [128-130, 141, 143, 154, 169] and $\sim 1\mu$ s for fast-OTF [154]). Finally, through a continuously monitoring of $I_{D,lin}$ at constant V_G [139], ΔV_T is indirectly estimated using $\Delta I_{D,lin} = I_{D,lin} - I_{D,lin0}$ through appropriate $\Delta I_{D,lin} \rightarrow \Delta V_T$ conversion technique (see section 6.7.4 for a discussion on this

conversion issues). Moreover, for OTF- $I_{D,lin}$, recovery voltage $|V_{REC}|$ is always above threshold and measurement voltage, $|V_{meas}|$ is same as $|V_{STS}|$ or $|V_{REC}|$.

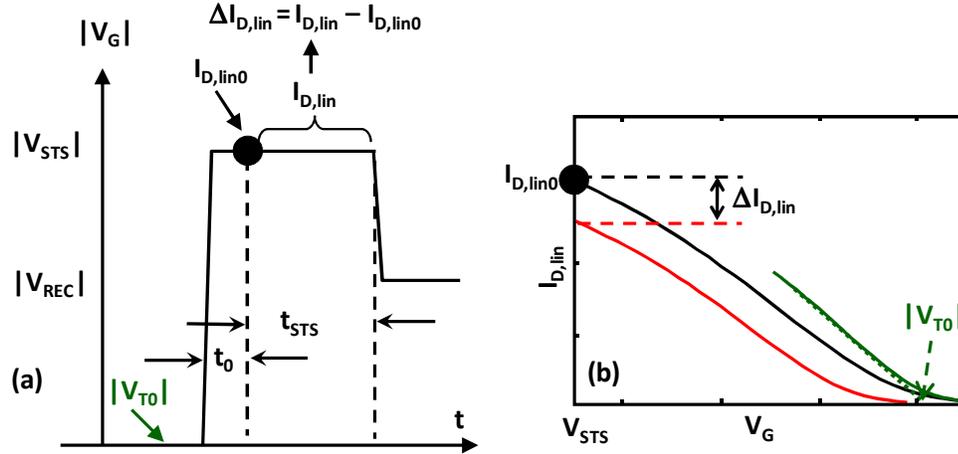


Figure 6.10: (a) Timing diagram for OTF- $I_{D,lin}$ measurement. (b) Hypothetical $I_{D,lin}$ - V_G sweeps are shown at different time steps during OTF- $I_{D,lin}$. In OTF- $I_{D,lin}$, only the initial short (V_G)-range $I_{D,lin}$ - V_G is measured for estimating V_{T0} . Later, V_{STS} is applied and $I_{D,lin0}$ is measured after t_0 . Finally, ΔV_T is indirectly estimated using $\Delta I_{D,lin}$. Moreover, $|V_{REC}|$ is always above threshold and $|V_{meas}|$ is same as $|V_{STS}|$ or $|V_{REC}|$.

6.7. Interpreting Measurements

Measurement of a quantity generally involves monitoring certain observables (*e.g.*, absorption of electromagnetic signal in ESR, increase in recombination current in SDR, change in $I_{D,lin}$ - V_G in MSM, change in V_G for constant $I_{D,lin}$ in UFV, change in $I_{D,lin}$ at constant V_G in OTF- $I_{D,lin}$, *etc.*), which are easily accessible through the apparatus being used. In most of the cases, these observables only provide an indirect estimate of the desired quantity and sometimes get corrupted due to different measurement artifacts. Therefore, proper interpretation of the observables and de-embedding the desired information is a challenge one needs to overcome during any measurement. For example,

- EDMR technique involves the measurement of recombination current through the substrate of a MOSFET, connected in gate-diode configuration (Figure 6.2a). Here, the microwave (electromagnetic) radiation, used to satisfy the resonance condition (Figure 6.1), can induce some unintentional forward-bias between the source/drain to substrate diode [88]. Therefore, the effect of this unintentional forward-bias needs to be corrected by the application of a counter bias using external supply. Otherwise, significant confusion may rise about the origin of spin-dependent current being measured, *i.e.*, whether it is coming from spin dependent tunneling (SDT) or spin-dependent recombination (SDR) [88].
- Mobility measurement using $I_{D,lin}-V_G$ in a MOSFET, having high gate leakage, is another challenging issue [273-275]. As shown in Figure 6.11, presence of high I_G in MOSFETs with ultra-thin dielectric opposes the flow of $I_{D,lin}$. Therefore, measured $I_{D,lin}$ at high V_G decreases and extracted μ_{eff} indicates an unphysically low magnitude at high E_{eff} . Such effect is more dominant in transistors with long-channel length, which has comparatively higher gate leakage due to larger gate area. Different correction techniques [273-275] are available in literature to handle this problem; among them the one proposed in [274] is the simplest. As I_G is unaffected by small lateral field [276], corrected $I_{D,lin}$ can be calculated using –

$$I_{D,lin(corrected)} = I_{D,lin} @ (V_{DS} = 0.01 + V_{DS,meas}) - I_D @ (V_{DS} = 0.01) \quad (6.8)$$

where $V_{DS,meas}$ is the V_{DS} at which $I_{D,lin}-V_G$ sweep is intended to be measured. We followed the same procedure in measuring $I_{D,lin}-V_G$ sweeps during our MSM experiments.

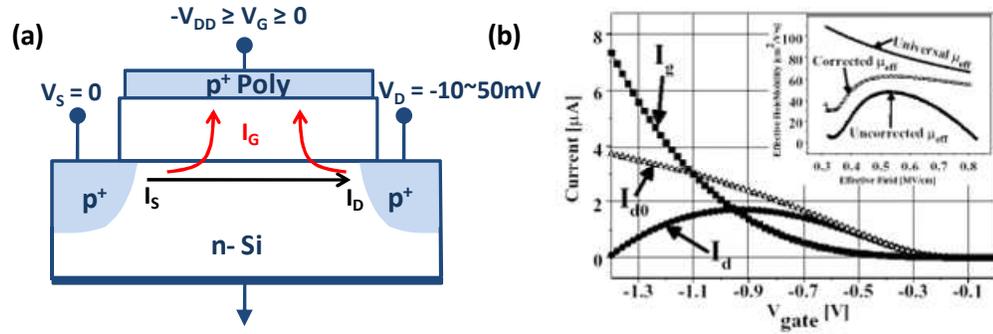


Figure 6.11: (a) During a $I_{D,lin}-V_G$ measurement, extremely popular in μ_{eff} calculation, presence of high I_G (in MOSFETs with ultra-thin gate dielectric) will assist I_S and will oppose $I_{D,lin}$. (b) Different current components measured during the $I_{D,lin}-V_G$ measurement in such a p-channel transistor. Excessive I_G reduces $I_{D,lin}$ (at high V_G), thus corrupts the extracted μ_{eff} at higher E_{eff} (inset). Here, I_{D0} represents the corrected drain current, using the algorithm proposed in [273].

- Significant corrections steps are also necessary for the interpretation of UFV and OTF- $I_{D,lin}$ measurements, which will be discussed in the following sections. Initially, we address the corrections related to body-effect coefficient (section 6.7.1) and valence band electron trapping (section 6.7.2) for UFV. Later, we discuss the effect of time-zero (t_0)-delay (section 6.7.3) and the need for $\Delta\mu_{eff}$ corrections (section 6.7.4) in OTF- $I_{D,lin}$ measurements.

6.7.1. Ultra-fast V_G (UFV): Body-effect Coefficient (m) Correction

UFV measures change in V_G (ΔV_G) required for maintaining constant I_D (within $t_{meas} \sim \mu s$) and hence estimate ΔV_T using, $\Delta V_T \sim \Delta V_G$ [140]. As shown in Figure 6.12a, such approximation significantly under-estimates $|\Delta V_T|$, when $I_{D,lin(UFV)}$ is chosen in the sub-threshold regime [165, 277]. Similarly, when $I_{D,lin(UFV)}$ is chosen with $|V_{meas}|$ slightly above threshold, the use of $|\Delta V_T| \sim |\Delta V_G|$ over-estimates $|\Delta V_T|$ [165].

Limitation of using $\Delta V_T \sim \Delta V_G$ in sub-threshold region can be explained by starting with the expression for $I_{D,lin}$ commonly used in this region, which is –

$$|I_{D,lin}| \sim \exp\left[q|V_G - V_T|/mk_B T\right]. \quad (6.9)$$

As $I_{D,lin}$ is kept constant (at $I_{D,lin(UFV)}$) during UFV, we need to differentiate equation (6.9) for constant $I_{D,lin}$ for obtaining the following relationship between ΔV_T and ΔV_G :

$$|\Delta V_G| - |\Delta V_T| = -|V_G - V_T| \Delta m / m. \quad (6.10)$$

Equation (6.10) clearly indicates that for positive Δm (Figure 8.1), $|\Delta V_G|$ in sub-threshold regime will always be less than $|\Delta V_T|$ (Figure 6.12a). Thus, estimation of ΔV_T using UFV in sub-threshold regime should involve appropriate Δm correction according to equation (6.10), only when the estimated ΔV_T becomes consistent with the one (section 6.4.2) measured by drawing slopes at $g_{m,max}$ (Figure 6.12b).

Similarly, a qualitative explanation for $|\Delta V_G| > |\Delta V_T|$ (at constant $I_{D,lin}$) above threshold can be obtained by differentiating equation (6.4) for constant $I_{D,lin}$ (*i.e.*, for constant E_{eff}). Hence, we obtain –

$$|\Delta V_G| - |\Delta V_T| = -\frac{\Delta \mu_{eff}}{\mu_{eff}} \bigg|_{E_{eff}} \left| V_G - V_T - m \frac{V_{DS}}{2} \right| + \Delta m \left| \frac{V_{DS}}{2} \right|. \quad (6.11)$$

Now, for $I_{D,lin} = I_{D,lin(UFV)}$ or constant E_{eff} , $\Delta \mu_{eff}$ is negative (Figure 6.8) and Δm is positive (Figure 6.6). Hence, according to equation (6.11), $|\Delta V_G|$ is always greater than $|\Delta V_T|$ for above threshold UFV measurement (Figure 6.12a). Lack of appropriate mobility models near threshold restricts us in obtaining a quantitative correction above threshold, as it is possible in the sub-threshold region (Figure 6.12b).

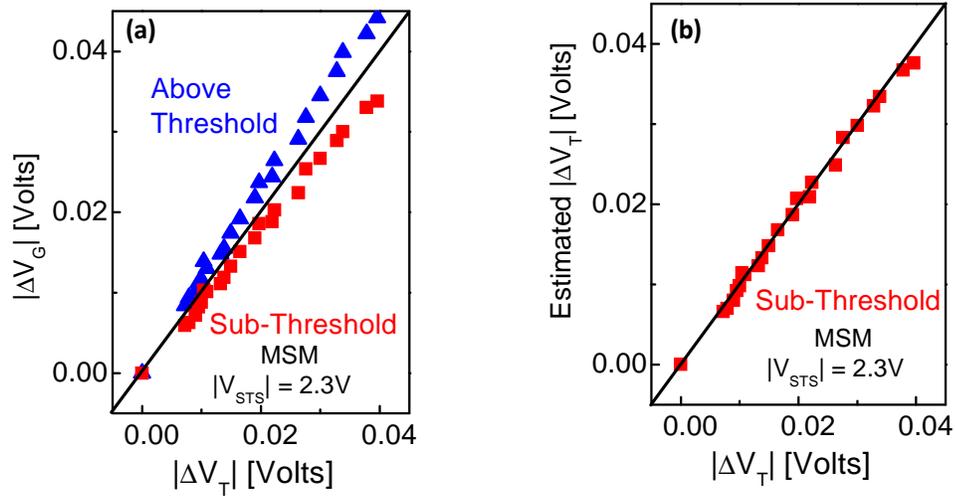


Figure 6.12: (a) $|\Delta V_G| \neq |\Delta V_T|$ for UFV, as assumed in [140]; at sub-threshold, $|\Delta V_G| < |\Delta V_T|$ and at slightly above threshold, $|\Delta V_G| > |\Delta V_T|$. (b) Appropriate Δm correction using equation (6.10) and Figure 6.6 enables us to have a refined estimation of ΔV_T (using ΔV_G), which becomes more consistent with the one obtained by drawing slopes at $g_{m,max}$ (section 6.4.2).

6.7.2. Ultra-fast V_G (UFV): Valence Band Electron Trapping

One of the advantages of UFV over OTF- $I_{D,lin}$ is its capability of measuring ΔV_G relaxation for any V_{REC} within $t_{meas} \sim \mu s$ (Figure 6.9a). Later ΔV_T can be estimated using appropriate correction, discussed in section 6.7.1. Thus UFV allows NBTI relaxation measurement at $V_{REC} < V_{T0}$ and the measured relaxation, in most of the transistors, starts from a recovery time (t_{REC}) of $\sim \mu s$ and has a *distinct logarithmic nature* [140]. This motivated many researchers [115, 140, 168, 278] to believe that N_{IT} relaxation in type-I transistors, where $\Delta V_T \sim \Delta V_{IT} \sim q\Delta N_{IT}/C_{ox}$ (due to the dominance of ΔN_{IT} over hole trapping during NBTI stress), starts from $\sim \mu s$. However, we observe that N_{IT} relaxation does not always start at $\sim 1\mu s$, as measured using UFV at $V_{REC} \sim 0V$. When relaxation is measured using OTF- $I_{D,lin}$ at $V_{REC} < V_{T0}$ in a PMOS transistor, the start of relaxation can be as large as 10ms for $t_{STS} = 10^3$ sec (Figure 6.13).

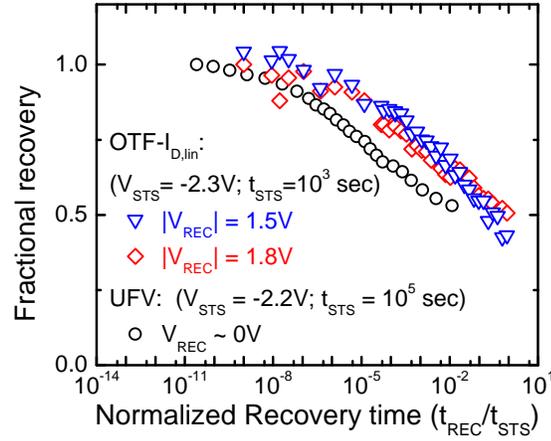


Figure 6.13: NBTI relaxation measured using OTF- $I_{D,lin}$ at $V_{REC} < V_{TO}$ predicts measured ΔV_T recovery to start later compared to the one measured using UFV at $V_{REC} = 0$.

To resolve the discrepancy between UFV and OTF- $I_{D,lin}$ measurements in Figure 6.13, we identify that there is V_{REC} dependency of *measured* ΔV_T relaxation. As such the relaxation measured in UFV at $V_{REC} = 0$, and starting at $\sim \mu\text{s}$, may not be the actual N_{IT} relaxation. Because for the ranges of V_{REC} used in UFV, the ultra-fast ΔV_T relaxation can result from the neutralization of dangling bonds by electron capture ($\text{Si}^+ + e^- \rightarrow \text{Si}$) from valence band. As a result, in spite of having negligible amount of ΔV_T relaxation by hydrogenated re-passivation of dangling bonds (N_{IT} recovery) within $1\mu\text{s}$ - 10ms time range, these neutralized dangling bonds (denoted by N_e) can also give rise to measured ΔV_T recovery.

6.7.2.1. Valance Band Electron Trapping: Process

In order to explain the process of valence band electron trapping for the fast-transient relaxation observed in UFV, we first need to identify the Fermi-level within the substrate (E_{FS}) at the bias conditions of UFV (Figure 6.9a). This will help us to identify the fraction of donor-type N_{IT} going below E_{FS} , as the transistor is switched from V_{STS} to V_{REC} . These

donor-type N_{IT} going below E_{FS} during relaxation will contribute to ΔV_T relaxation (with no N_{IT} relaxation) by capturing electron from the valence band. As shown in Figure 6.14, the following cases need to be considered for explaining such valence band electron trapping:

- In stress phase, *i.e.*, $V_G = V_{STS}$, E_{FS} is well below valence band edge, considering that more than $2\phi_F$ surface potential is needed to invert the substrate due to QM effects [178, 279-282] (where ϕ_F is the separation between E_{FS} and the intrinsic level within the bulk of substrate). So the generated N_{IT} 's over the entire energy range within the bandgap will be positively charged (or empty of electrons, see Figure 6.14a) and thus will contribute towards V_T .
- In measurement phase, *i.e.*, $V_G = V_{meas}$ (normally $\leq V_{REC}$), pushes part of the N_{IT} 's below E_{FS} (Figure 6.14b). In the meantime, carriers (electrons/holes) in valence band of the PMOSFET's inverted channel can quickly equilibrate with source/drain within majority carrier diffusion time. However, electron trapping in the donor-type N_{IT} 's, which has gone below E_{FS} , will require some time. Among these N_{IT} 's, the ones closest to the valence band can capture electrons and become neutral within $\sim\mu\text{s}$. Thus these (near valence band) N_{IT} 's cannot be measured by UFV. As such, even in stress-phase measurement, we will be able to estimate ΔV_T (only) due to those N_{IT} 's which have not captured electrons within $\sim\mu\text{s}$. This phenomena of capturing only a fraction of generated N_{IT} 's during measurement has also been discussed in [168].
- In recovery phase, *i.e.*, $V_G = V_{REC}$, E_{FS} is either pushed further up towards the conduction band (for $V_{meas} < V_{REC}$, Figure 6.14c) or stays the same as in measurement condition (for $V_{meas} = V_{REC}$, Figure 6.14b). In both cases, N_{IT} 's below E_{FS} gets more time (*i.e.*, the entire recovery phase) to become neutral, thus contribute to ΔV_T relaxation. In addition, when V_{meas} is less than V_{REC} (*i.e.*, E_{FS} is closer to valence band for $V_G = V_{meas}$ compared to that for $V_G = V_{REC}$), a fraction of N_{IT} 's (already in neutral state at $V_G = V_{REC}$) may release

the captured electron within $t_{meas} \sim \mu\text{s}$. These N_{IT} 's, which release captured electrons during the switching of transistor from V_{REC} to V_{meas} , will not contribute to ΔV_T relaxation.

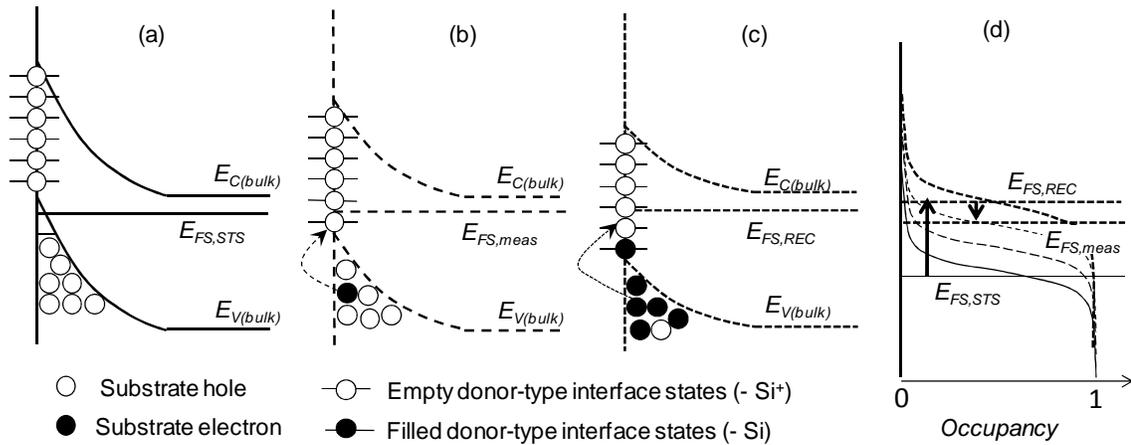


Figure 6.14: Energy band diagram for a PMOS transistor biased in (a) NBTI stress ($V_G = V_{STS}$, substrate inverted), (b) NBTI measurement ($V_G = V_{meas} \sim V_T$) and (c) NBTI recovery ($V_G = V_{REC} = 0$, substrate moderately depleted) condition for the UFV scheme. E_{FS} is substrate Fermi level and $E_{C(bulk)}$, $E_{V(bulk)}$ are conduction and valence band level in bulk of the substrate. Arrow indicates valence band electron trapping during recovery and measurement conditions. (d) Movement of Quasi-Fermi level as the transistor is switched from stress to recovery, then to the measurement mode. Energetically suppressed substrate electron capturing makes the trap filling a $\log-t$ process from $t_{REC} = \sim \mu\text{s}$ to $\sim \text{ms}$.

Therefore, it is evident that some of the N_{IT} 's, located at energy levels closer to the valence band edge and also below E_{FS} , can capture valence band electrons at $V_G = V_{REC}$ and keep neutralized for $t_{meas} \sim \mu\text{s}$ without releasing the captured electron at $V_G = V_{meas}$. With the increase in t_{REC} , N_{IT} 's located away from valence band edge starts to capture electrons (Figure 6.14d). Overall, valence band electron trapping into N_{IT} 's, going below E_{FS} at $V_G = V_{REC}$, can have a wide spread of trapping time constant, thus resulting in

logarithmic time ($\log-t$) dependence of NBTI recovery. Evidently, the fraction of N_{IT} 's contributing to the $\log-t$ dependence or, the fraction of defects neutralized through electron trapping will depend on E_{FS} positions at $V_G = V_{STS}, V_{REC}, V_{meas}$; which ultimately depends on doping density in the substrate and poly-gate, EOT , work-function for metal-gate, *etc.*

6.7.2.2. Valence Band Electron Trapping: Numerical Model

In order to verify the significance of valence band electron trapping into N_{IT} 's in explaining ultra-fast $\log-t$ relaxation, measured in UFV, we consider that donor-type N_{IT} 's are generated uniformly throughout the band gap. Here, we model the generation and recovery of N_{IT} within a direct H-H₂ Reaction-Diffusion (R-D) framework (see sections 3.4.1 and 3.5 for details). Let us also define n_{IT} ($= D_{IT}\Delta E$) as the concentration of interface traps within an energy range of ΔE generated at time t . Therefore, total interface trap density can be expressed as: $N_{IT} = \sum_{E_g} n_{IT}$. At stress condition, all the generated interface states will be positively charged ($-\text{Si}^+$) (Figure 6.14a), and thus contribute to measured ΔV_T . As the transistor is switched to recovery mode ($V_G = V_{REC}$), ΔV_T relaxation can occur in the following two ways:

- Charged interface traps ($-\text{Si}^+$) can capture atomic hydrogen that diffuses back towards the interface. In this case, recovered H- Si^+ bond go to an energy state well below the valence band edge [147] and hence can easily capture an electron to complete the passivation process ($-\text{Si}^+ + \text{H} + e^- \rightarrow \text{H-Si}$). Such ΔV_T relaxation process directly impacts ΔN_{IT} by reducing n_{IT} at appropriate energy levels closer to the valence band edge.
- Alternately, $-\text{Si}^+$ can capture an electron from valence band and thus form a neutral donor state ($-\text{Si}^+ + e^- \rightarrow -\text{Si}$) with a concentration of n_e . In such case, these neutralized donor interface states will reflect as ΔV_T recovery in UFV measurement. Moreover, neutralized defects n_e have their energy levels within Si band gap and may release the captured electron during ΔV_T measurement

(when $V_{meas} < V_{REC}$) or show up as a fast ΔV_T generation in subsequent stress phase. These neutralized dangling bonds (n_e) can also anneal through re-passivation by the back-diffused hydrogen species.

Thus, valence band electron-capture process splits interface defects within the band-gap into two parts: empty/positively-charged n_{IT} and filled/neutral n_e . The time dynamics of these defects at certain energy E_t are governed by:

$$\begin{aligned} \frac{dn_{IT}(E_t)}{dt} &= k_F [n_0 - n_{IT}(E_t)] - k_R n_{IT}(E_t) N_H^{(0)} \\ &\quad - c_{IT} \exp[(E_S - E_t)/k_B T] n_S n_{IT}(E_t) + c_{IT} p_S n_e(E_t), \end{aligned} \quad (6.12)$$

$$\begin{aligned} \frac{dn_e(E_t)}{dt} &= -k_R n_e(E_t) N_H^{(0)} \\ &\quad + c_{IT} \exp[(E_S - E_t)/k_B T] n_S n_{IT}(E_t) - c_{IT} p_S n_e(E_t), \end{aligned} \quad (6.13)$$

where k_F , k_R , $N_H^{(0)}$ are Si-H bond-breaking rate, Si-H bond-annealing rate and hydrogen density at the Si/dielectric interface, respectively; n_0 is the initial annealed dangling bond at the beginning of stress within energy range ΔE (in other words, total Si-H bond density available before stress is $N_0 = \sum_{E_g} n_0$); c_{IT} is the capture coefficient for valence band electron trapping into and out of n_{IT} ; n_S (p_S) is surface electron (hole) density contributing to trapping (detrapping) process; and E_S is the energy level for surface electrons. Here, it is important to note that k_F , k_R , $N_H^{(0)}$ and n_0 are terms commonly used in classical R-D model (see chapter 3 for details), whereas the remaining terms take care of extra contribution from valence band electron trapping.

6.7.2.3. Valence Band Electron Trapping: Simulation Results

We perform QM-CV simulation [178] to calculate n_S , p_S and E_S for the transistor under consideration at different V_G , which are fed into equations (6.12) and (6.13), along with the hydrogen diffusion terms. This enables us to determine $n_{IT}(E_t)$ and $n_e(E_t)$ at different stress/recovery time intervals. In the simulated stress-phase, we apply $V_{STS} = -2.2V$ and $T_{STS} = 125^0 C$ on a PMOS transistor having device parameters ($V_{T0} \sim -0.4V$,

$EOT \sim 2.2\text{nm}$, substrate doping density $\sim 10^{18}\text{ cm}^{-3}$, etc.) similar to the one used in [140]. After $t_{STS} = 10^5$ sec, the transistor is allowed to recover at $V_{REC} = 0\text{V}$. Both $\Delta V_T = q \sum_{E_g} n_{IT} / C_{ox}$ and $\Delta N_{IT} = q \sum_{E_g} (n_{IT} + n_e) / C_{ox}$ are monitored at pre-defined stress/recovery intervals. In addition, capture co-efficient (c_{IT}) for valence band electron trapping simulation considers capture cross-section, $\sigma_{IT} \sim 10^{-18}\text{ cm}^2$ (which is similar to the value reported in [283]) and thermal velocity, $v_{th} \sim 10^7\text{ cm/sec}$. Also valence band electrons from quantized energy levels are considered to tunnel an interfacial layer thickness of $\sim 1.5\text{ \AA}$ [69] before getting captured by interface defects.

Comparison of ΔV_T and ΔN_{IT} relaxation in Figure 6.15a shows that due to the extra contribution from valence band electron trapping component, fractional ΔV_T recovery is much larger than ΔN_{IT} recovery. More importantly, though ΔN_{IT} starts to relax by back-diffused hydrogen species at $\sim 10^2$ s, relaxation of measured ΔV_T starts much earlier. Moreover, since $V_{meas} (\sim -0.4\text{V}) < V_{REC} (\sim 0\text{V})$ in our simulation, some N_{IT} 's can release the captured electrons within $t_{meas} \sim \mu\text{s}$. As such, this electron release process during t_{meas} will result reduced ΔV_T recovery in the relaxation measurement (see Figure 6.15b).¹⁴ In both cases irrespective of V_{meas} , we get a $\log-t$ dependence in measurable ΔV_T relaxation, which happens before the N_{IT} 's achieve total re-passivation (*i.e.*, N_{IT} recovery in classical R-D model) through the back-diffusion of hydrogen species.

¹⁴ This also proves that if t_{meas} were significantly higher or V_{meas} were smaller than V_{T0} (as is the case for ultra-fast I_D-V_G setup [141]), no effect from valence band electron trapping will be observed in measured ΔV_T relaxation.

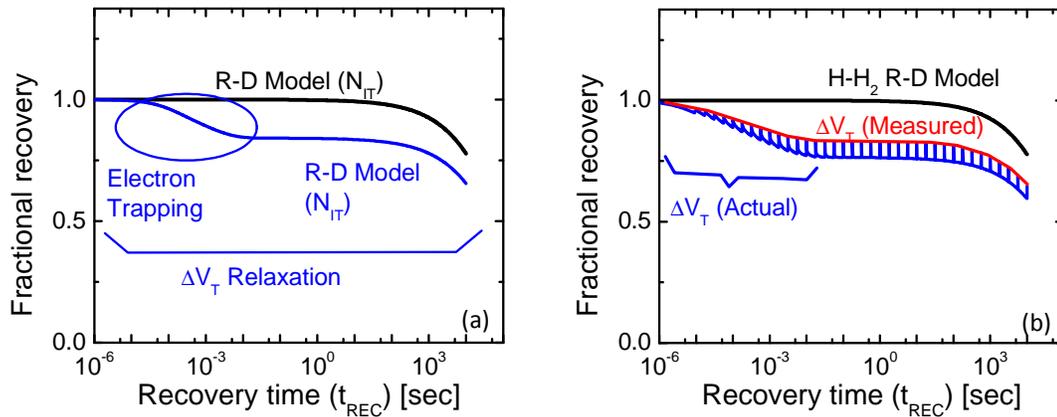


Figure 6.15: Simulation of fractional recovery ($t_{STS} = 10^5$ sec) in UFV measurement scheme ($t_{meas} \sim \mu\text{s}$) on a PMOS transistor having parameters similar to the one used in [140]. Our simulation indicates $E_{FS,rec}$ (@ $V_{REC} = 0\text{V}$) to be $\sim 0.45\text{eV}$ above the valence band edge at the interface. In (a), $E_{FS,meas}$ is considered to be same as $E_{FS,rec}$ (*i.e.*, $V_{meas} = V_{REC}$); whereas in (b) $E_{FS,meas}$ (@ $V_{meas} = V_{T0} \sim -0.4\text{V}$) \sim is considered to be 0.2eV above valence band edge at the interface. Figure (b) also indicates actual ΔV_T recovery (blue line), measured ΔV_T recovery (red line) and actual ΔN_{IT} recovery (black line: predicted from classical H-H₂ R-D model).

6.7.2.4. Valence Band Electron Trapping: Resolving UFV vs. OTF- $I_{D,lin}$ Relaxation

For UFV relaxation measurement [140], Figure 6.15a indicates that valence band electron trapping can contribute to 20% ΔV_T relaxation from $1\mu\text{s}$ - 10ms . Considering this information, we can subtract the contributions of electron trapping from ΔV_T relaxation data. The corrected ΔV_T relaxation for UFV at $V_{REC} \sim 0\text{V}$ shows remarkable agreement with the one obtained from OTF- $I_{D,lin}$ measurement at $V_{REC} < V_{T0}$ (Figure 6.16). This proves the necessity of valence band electron trapping in explaining ultra-fast NBTI relaxation measurements at $V_{REC} > V_{T0}$ and confirms that actual N_{IT} relaxation does not start at $\sim \mu\text{s}$, when considerable amount of stress is applied. Thus for $t_{STS} = 10^3$ sec, we predict the N_{IT} relaxation to start at $\sim 10\text{ms}$, which is confirmed by two different types of experiments in Figure 6.16. Moreover, Figure 6.16 also shows significant differences

between the experimentally measured start of relaxation $t_{REC,start}$ and the one calculated $t_{NIT,start}$ using Reaction-Diffusion (R-D) model. This difference is attributed to the existence of hole detrapping from oxide defects, as discussed in section 6.7.5.

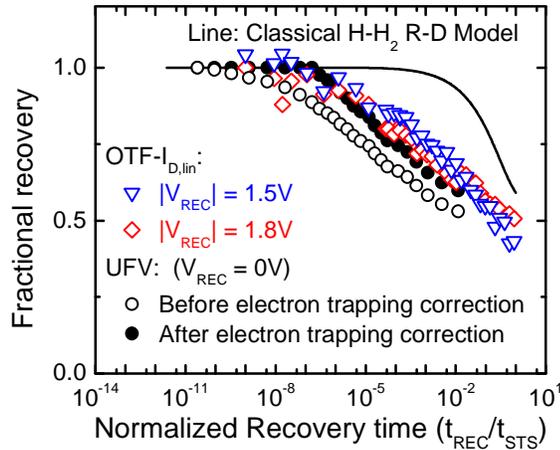


Figure 6.16: Consideration of valence band electron trapping in UFV makes start of recovery consistent with OTF- $I_{D,lin}$ measurement.

6.7.3. On-the-fly $I_{D,lin}$: Time-zero (t_0)-delay Effect

As discussed in section 6.6, OTF- $I_{D,lin}$ primarily measures V_{T0} and $\Delta I_{D,lin}$ and thus make a zeroth-order estimate of ΔV_T using [127, 129, 130, 139, 169, 284, 285]¹⁵ –

$$|\Delta V_T(t)| \cong \left| \frac{[I_{D,lin}(t) - I_{D,lin}(t_0)]}{I_{D,lin}(t_0)} (V_G - V_{T0}) \right| = \left| \frac{\Delta I_{D,lin}}{I_{D,lin0}} (V_G - V_{T0}) \right|. \quad (6.14)$$

A crucial point that is not always obvious in OTF- $I_{D,lin}$ measurement, hence ignored, is the use of t_0 in the above calculation. Exact ΔV_T measurement using OTF- $I_{D,lin}$ mandates the use of $t_0 = 0$. Yet OTF- $I_{D,lin}$ setups require some time (≥ 1 ms for standard OTF [128-

¹⁵ Please see section 6.7.4 for a better estimate of ΔV_T through proper mobility correction.

130, 141, 143, 154, 169] and $\sim 1\mu\text{s}$ for fast-OTF [154]) before measuring $I_{dlin}(t_0)$, as a result t_0 (termed as *time-zero* delay) $\neq 0$ in such measurements. Therefore, while solving the issue of unintentional recovery at each measurement step present in any other measurement technique (*e.g.*, MSM, SDR, etc.), OTF- $I_{D,lin}$ introduces an error in degradation estimates due to non-zero t_0 , the effect of which will be particularly important at short stress times.¹⁶ To observe the effect of *time-zero* (t_0)-delay in the OTF- $I_{D,lin}$ results, we have used the experimental data from [140] (measured at $V_{STS} = -2.2\text{V}$, $T_{STS} = 125^\circ\text{C}$), with a specified $t_{meas} \sim \mu\text{s}$ (hence very small recovery) and $t_0 \sim 0$ (similar to MSM), as reference. Once correction due to t_0 is accounted for, we can explain n for all OTF- $I_{D,lin}$ data in Figure 6.17a, having various $t_0 \neq 0$. Thus, Figure 6.17a indicates that the use of $t_0 \neq 0$ in calculating $\Delta V_T(t)$ for OTF- $I_{D,lin}$, results in an “apparent and artificial” increase of power-law time exponent, n (when ΔV_T is expressed as, $\Delta V_T \sim t^n$) mainly in short-term NBTI degradation. Indeed, if t_0 is very high (> 1 sec), even the longer-term stress data will be affected by t_0 -delay. Thus, t_0 -delay can explain the discrepancy of OTF- $I_{D,lin}$ results from different experimental groups [130, 139, 286].

To be more convincing, we vary t_0 -delay in OTF- $I_{D,lin}$ experiments, performed on type-I transistors having dominance of N_{IT} over N_{HT} in NBTI degradation, *i.e.*, $\Delta V_T \sim N_{IT}$ [119, 129] (see Figure 6.17b). We fit the lowest t_0 ($= 1\mu\text{s}$; hence assumed to be negligible) measurement using classical H-H₂ R-D model (see section 3.5 for details). The other measurements having higher t_0 nicely follows the simulation trend, once the effect of t_0 is considered using $\Delta V_T(t - t_0) \sim [N_{IT}(t) - N_{IT}(t_0)]/N_{IT}(t_0)$ with no additional fitting parameters. Thus, one can estimate power-law time exponent n (when the degradation is plotted in *log-log* format) for any value of t_0 by differentiating $\log(t^n - t_0^n)$ with respect to $\log(t - t_0)$, which results –

¹⁶ Such t_0 -error is not present either in MSM (which is erroneous due to recovery issue, discussed earlier) or in the UFV measurement, because in these cases V_{T0} is obtained before the device is stressed.

$$n = n_0 \frac{(t_{STS} - t_0)}{(t_{STS}^{n_0} - t_0^{n_0})} t_{STS}^{n_0-1}, \quad (6.15)$$

where n_0 is the power-law time exponent for $t_0 = 0$. Figure 6.17b also shows that although n at a particular t_{STS} increases with t_0 , all the degradation curves tend to merge at long t_{STS} . Thus as a matter of practical interest, if such t_0 -delay correction in OTF- $I_{D,lin}$ is difficult, our analysis suggests that lifetime prediction [128, 175] and projection back to operating condition [128] in type-I transistors must be carried out using long-term NBTI stress data, obtained at $t_{STS} \geq 10$ sec with $t_0 \leq 1$ ms; because higher-exponents in short-term degradation can result in unnecessarily pessimistic projections. Whereas, in case of type-II transistor (having dominance of N_{HT} over N_{IT}), t_0 needs to be as small as $1\mu\text{s}$ for proper characterization of ultra-fast N_{HT} part. Overall, t_0 -delay effect can obviously be minimized by using small t_0 [141, 154, 161, 287].

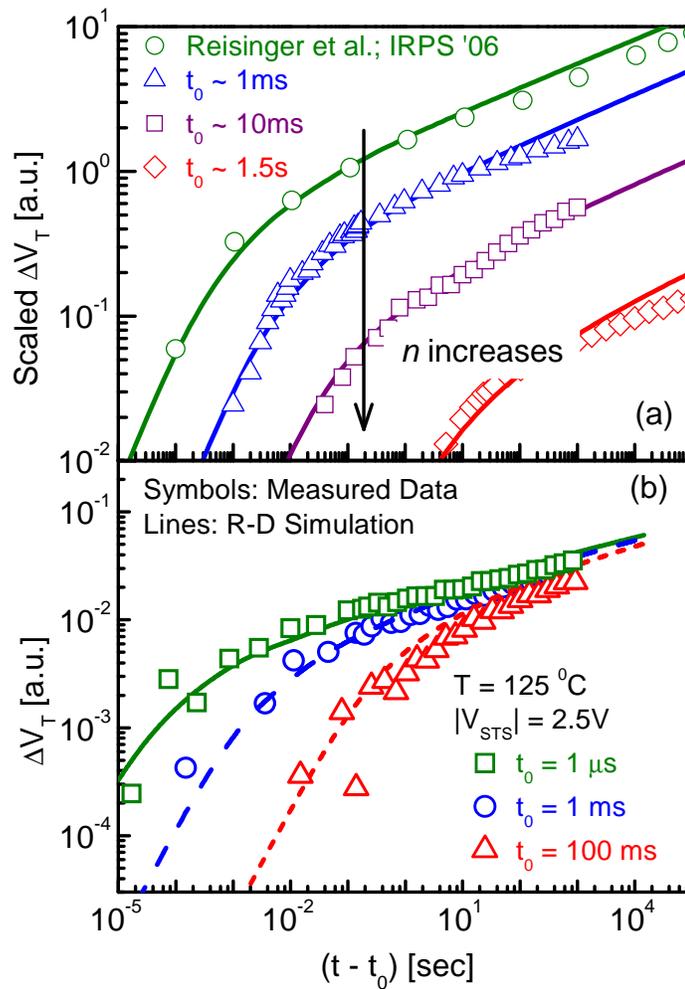


Figure 6.17: (a) Effect of t_0 -delay on OTF- $I_{D,\text{lin}}$ compared with UFV, where OTF- $I_{D,\text{lin}}$ is performed on transistors from different technologies. Here, ΔV_T is scaled along Y-axis for clarity. Next, in (b) the possibility of technology to technology variation is avoided by studying the effect of t_0 on similar transistors. In both cases, measurement with lowest t_0 is fitted using classical H-H₂ R-D model (section 3.5) and other measurements require no additional fitting parameter, except the consideration of t_0 . These measurements indicate that for higher t_0 , there is an increase in n at a particular $(t_{\text{STS}} - t_0)$ mainly in the short-term stress.

6.7.4. On-the-fly $I_{D,lin}$ (OTF- $I_{D,lin}$): Mobility Correction in ΔV_T estimation

This is one of the more challenging problem in OTF- $I_{D,lin}$, because OTF- $I_{D,lin}$ needs to perform an indirect estimation of ΔV_T using the measured quantities like $\Delta I_{D,lin}$ and V_{T0} [127, 129, 130, 139, 169, 284, 285]. In the following discussions, we show that these two quantities are not sufficient for estimating ΔV_T and hence establish the importance of considering $\Delta\mu_{eff}$ in such estimation. We also show that although Δm has significant contribution in ΔV_T estimation for UFV (see section 6.7.1), it has negligible impact on ΔV_T estimation for OTF- $I_{D,lin}$. In determination of $\Delta\mu_{eff}$ (which is generally not available in OTF- $I_{D,lin}$ measurements), we use a theoretically-justified and well-calibrated $\Delta\mu_{eff}(N_{IT})$ model, which has been developed for type-I PMOS transistors (having dominant N_{IT} generation during NBTI) in section 6.4 using MSM experiment. Consequently, we propose an algorithm for estimating ΔV_T in OTF- $I_{D,lin}$ using prior knowledge of V_T , m and μ_{eff} in pre-stress ($t_{STS} = 0$) condition (section 6.7.4.1). The proposed algorithm is a significant improvement over the one presented in [169], where each device needs individual characterization to obtain the mapping between $\Delta I_{D,lin}$ and ΔV_T . We verify the algorithm using both MSM and OTF- $I_{D,lin}$ experiments (sections 6.7.4.2 and 6.7.4.3) and hence show the importance of this algorithm in robust lifetime projection using OTF- $I_{D,lin}$ measurement (section 6.7.4.4). Finally, we generalize this algorithm for type-I transistors having wide range of nitrogen concentration (section 6.7.4.5).

6.7.4.1. Algorithm for ΔV_T Estimation from ΔI_D

For developing the ΔV_T estimation algorithm (from $\Delta I_{D,lin}$), first we differentiate (6.4) to obtain –

$$-\left| \frac{\Delta I_{D,lin}}{I_{D,lin0}} \right| = \frac{\Delta\mu_{eff}}{\mu_{eff0}} \frac{V_G - V_T - \frac{m}{2} V_{DS,lin}}{V_G - V_{T0} - \frac{m_0}{2} V_{DS,lin}} + \frac{\mu_{eff}}{\mu_{eff0}} \frac{-|\Delta V_T| - \frac{1}{2} V_{DS,lin} |\Delta m|}{V_G - V_{T0} - \frac{m_0}{2} V_{DS,lin}} = f(\Delta V_T). \quad (6.16)$$

Relationship between $\Delta\mu_{eff}$ and Δm with ΔV_T (obtained in sections 6.4.2 and 6.4.3) – along with the prior knowledge of $\mu_{0(0)}$, θ_0 , and m_0 – are used in equation (6.16) to calculate *correct- ΔV_T* ¹⁷ from measured $\Delta I_{D,lin}/I_{D,lin0}$. In sum (see Figure 6.18), MSM study of section 6.4 enables us to calculate $\Delta m(\Delta V_T)$, $\Delta\mu_0(\Delta V_T)$ and $\Delta\theta(\Delta V_T)$ (Figure 6.6), thus $\Delta\mu_{eff}(\Delta V_T)$ using (6.5) (Figure 6.7). These values are subsequently used in equation (6.16) to obtain a relationship between $\Delta I_{D,lin}$ and ΔV_T – which can be used in ΔV_T estimation for OTF- $I_{D,lin}$.

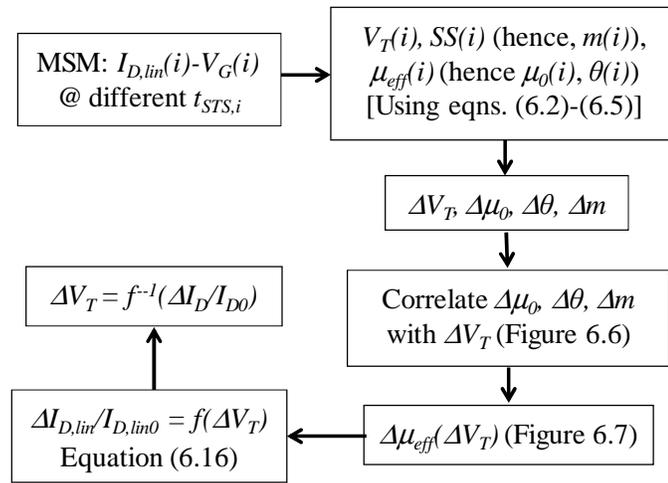


Figure 6.18: Algorithm for ΔV_T estimation in OTF- $I_{D,lin}$ from $\Delta I_{D,lin}/I_{D,lin0}$.

6.7.4.2. Verifying $\Delta I_{D,lin} \rightarrow \Delta V_T$ Algorithm Using MSM

The proposed algorithm is used to estimate ΔV_T from $\Delta I_{D,lin}/I_{D,lin0}$, measured at different V_{meas} using MSM experiments. Estimated ΔV_T (using information from Figure 6.6, Figure 6.7 into Figure 6.18) correlates very well with ΔV_T calculated by drawing

¹⁷ V_T calculation by determining the intercept for the line drawn at $g_{m,max}$, is widely used in device characterization [33, 89, 90, 168, 169, 249, 288, 289]. Thus, correcting OTF- $I_{D,lin}$ using the ΔV_T calculated from $g_{m,max}$ technique is considered appropriate.

slopes at $g_{m,max}$ (Figure 6.19a) – suggesting that the algorithm is unique and self-consistent for $|V_{meas}| \gg |V_T|$ ¹⁸. We verify this uniqueness and consistency for all the type-I transistors under study (see Table 6.1 for device details). This is an improvement over popular classical estimate, where both $\Delta\mu_{eff}$ and Δm are neglected and ΔV_T is estimated using equation (6.14). Figure 6.19b proves that classical calculation is under-estimating ΔV_T [89, 169, 290] by neglecting contribution from $\Delta\mu_{eff}$ and Δm (estimation is poorer at higher V_{meas}), thereby making theoretical interpretation of OTF- $I_{D,lin}$ experiments difficult. Similar calculation involving devices, where $\Delta\mu_{eff}@V_G$ degrades (Figure 6.7b) indicates classical calculation to over-estimate ΔV_T (Figure 6.19c). Such over-estimation can also be corrected using our algorithm of Figure 6.18.

To observe the relative significance of $\Delta\mu_{eff}$ and Δm in ΔV_T estimation from $\Delta I_{D,lin}$, we now ignore the contribution from Δm in estimating ΔV_T . Figure 6.19d shows that estimated ΔV_T in such condition compares pretty well with ΔV_T obtained at $g_{m,max}$. Thus ΔV_T estimation algorithm for OTF- $I_{D,lin}$ (Figure 6.18) seems to work well without considering sub-threshold slope (SS) degradation. This is because for small $V_{DS,lin}$ all the terms in equation (6.16) related to SS (and m) becomes negligible.

¹⁸ Despite its self-consistency, the proposed algorithm should not be used for low $|V_{meas}|$, because the mobility model (*i.e.*, equation (6.5)) will not be valid at that condition. Instead, one could use higher order empirical mobility models (suggested in [268]) to reconstruct Figure 6.18 at lower $|V_{meas}|$.

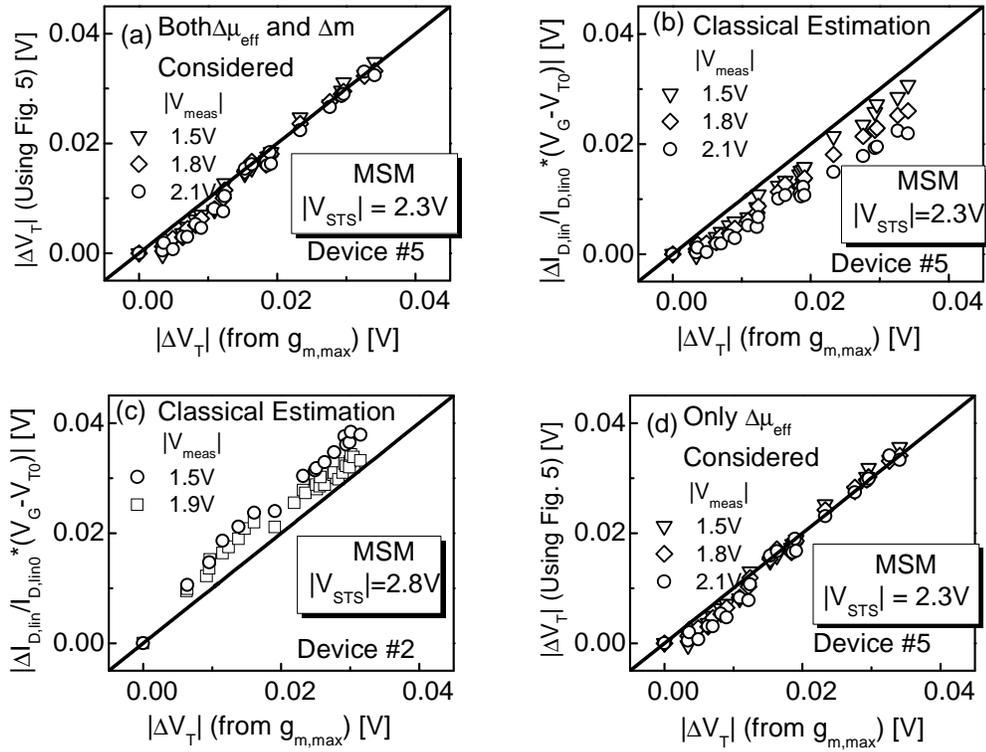


Figure 6.19: (a) Estimation of $\Delta\mu_{eff}$ and Δm enables one to evaluate ΔV_T using Figure 6.18; which is in good agreement (closer to solid line representing 1:1 relation) with ΔV_T obtained from $g_{m,max}$. (b) (c) Classical estimate of $|\Delta V_T| \sim |\Delta I_D / I_{D0} (V_G - V_{T0})|$ neglects $\Delta\mu_{eff}$ and Δm , thus under/over-estimates $|\Delta V_T|$. (d) Neglecting contribution from m does not affect ΔV_T estimation significantly. Information for the devices used in this figure can be found in Table 6.1.

6.7.4.3. Verifying $\Delta I_{D,lin} \rightarrow \Delta V_T$ algorithm using OTF- $I_{D,lin}$

Effectiveness in ΔV_T estimation using algorithm in Figure 6.18 might lead one to presume that the proposed methodology is only self-consistent, but not universal. In other words, as we have used the same MSM experiment in determining correlation between $\Delta\mu_{eff}$, Δm and ΔV_T (Figure 6.6 and Figure 6.7) and later used these information in equation (6.16) to estimate ΔV_T itself (section 6.7.4.2), the entire process is expected to show consistency (as obtained in Figure 6.19a). Thus effectiveness of Figure 6.18 might

appear to be valid only for this MSM experiment, not for OTF- $I_{D,lin}$ in general! *To disprove this*, we use information in Figure 6.6 and Figure 6.7 in doing ΔV_T estimation for NBTI recovery experiments, measured at different recovery voltage (V_{rec}) after stressing the transistors for same duration (1000 sec) under same stress condition ($V_{stress} = -2.3V$; $T_{stress} = 125$ °C). Figure 6.20 suggests that use of algorithm in Figure 6.18 in ΔV_T estimation correctly predicts similar start of relaxation, whereas classical ΔV_T estimation in equation (6.14) incorrectly shows that the relaxation is getting started at different point. This further proves the robustness of our ΔV_T estimation algorithm, which we will later generalize for the whole range of type-I transistors (section 6.7.4.5).

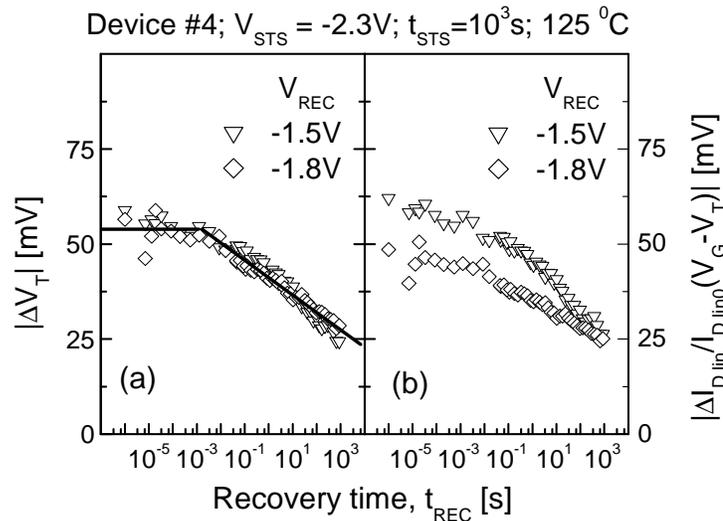


Figure 6.20: (a) Estimated ΔV_T using algorithm in Figure 6.18 provides similar ΔV_T at start of recovery for the same V_{STS} , T_{STS} , but different V_{REC} ; whereas (b) equation (6.14) incorrectly estimates different ΔV_T at start of recovery. Information for the device used in this figure can be found in Table 6.1.

Moreover, we have used parameters extracted from MSM experiment (Figure 6.6 and Figure 6.7) to correct OTF- $I_{D,lin}$ in Figure 6.20a. As MSM has inherent delay during its measurement, one might wonder if MSM and OTF- $I_{D,lin}$ (having no delay) probe the

same type of defects for these type-I transistors. And as such, one might question the *validity of using extracted MSM parameters in correcting OTD- $I_{D,lin}$* . Figure 6.20a responds to such concern by showing that in spite of having different delay, parameters from MSM, when used in OTF- $I_{D,lin}$, predicts similar start of recovery. Thus, we believe that the type of defects probed by MSM and OTF- $I_{D,lin}$ are indeed similar. In other words, *measurement delay in MSM is not affecting the ΔV_T estimation algorithm for type-I devices*.

6.7.4.4. Reliability Implications of Proper ΔV_T Estimation

To illustrate the importance of making proper ΔV_T estimation in OTF- $I_{D,lin}$ (using Figure 6.18) over its classical counterparts, *i.e.*, equation (6.14), we compare reliability parameters, *e.g.*, power-law time-exponent (n), lifetime (t_{life}), voltage acceleration (Γ_V), safe operating voltage (V_{safe}), and activation energy (E_A)¹⁹ obtained using both methods. Figure 6.21a (Figure 6.21c) suggest that while the algorithm in Figure 6.18 estimates higher (lower) ΔV_T compared to the classical calculation of equation (6.14), it causes negligible change in power-law time-exponents ($\Delta V_T \sim t^n$). Thus, by under (over)-estimating ΔV_T , classical prediction over (under)-estimates t_{life} (Figure 6.21a,c). Moreover, since this over (under)-estimation of t_{life} is more (less) dominant at higher V_{meas} (as evident from Figure 6.19b,c), classical ΔV_T calculation under-estimates Γ_V compared to the value obtained using Figure 6.18 (see Figure 6.21b,d). Hence, determination of V_{safe} is either lower/higher for classical calculation, depending on the specified lifetime criteria (Figure 6.21b,d).

¹⁹ See Appendix A for the definition/determination of parameters like t_{life} , Γ_V , V_{safe} .

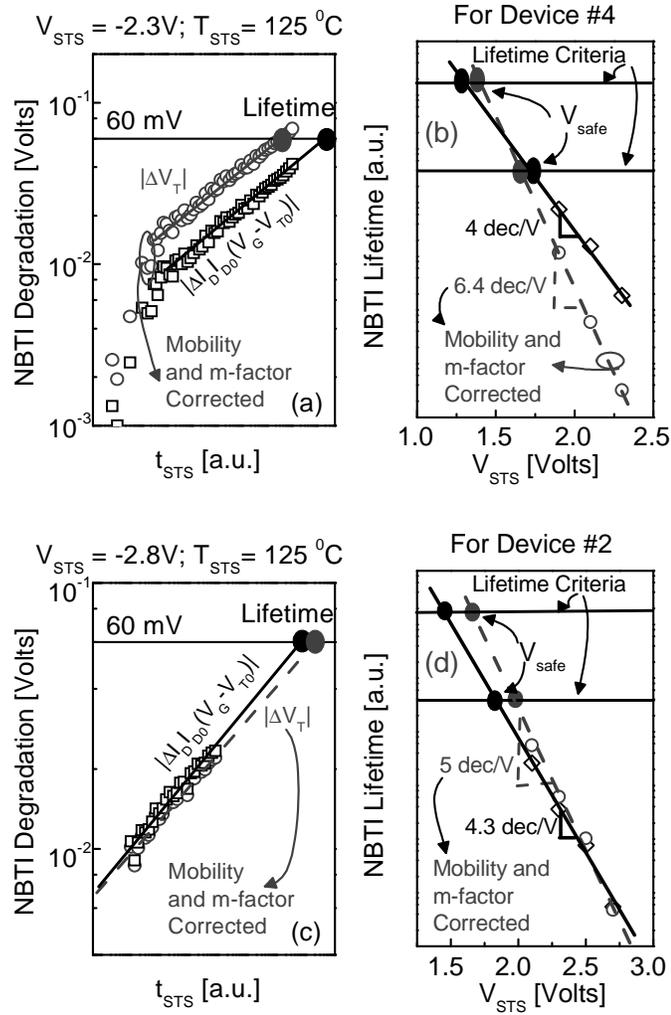


Figure 6.21: Although power-law time exponent remains unaffected, when $\Delta\mu_{eff}$ and Δm are neglected in ΔV_T estimation (for OTF- $I_{D,lin}$) compared to the case, where $\Delta\mu_{eff}$ and Δm are considered, lifetime is (a) over-estimated or (c) under-estimated. As a result, voltage acceleration (b), (d) is always under-estimated and V_{safe} is over- or under-estimated in classical estimation. Information for the devices used in this figure can be found in Table 6.1.

We also perform MSM at wide range of temperatures (55°C - 175°C) to investigate the effect of ignoring $\Delta\mu_{eff}$ and Δm on measured NBTI activation energy. Figure 6.22a-c

indicates that the correlation for mobility model parameters ($\Delta\mu_0$, $\Delta\theta$) and Δm with ΔV_T is unaffected due to change in temperature. Thus, the difference between estimated ΔV_T using Figure 6.18 and equation (6.14) is observed to be same at all temperatures, which implies that the measured E_A will be independent of $\Delta\mu_{eff}$ and Δm corrections (see Figure 6.22d).

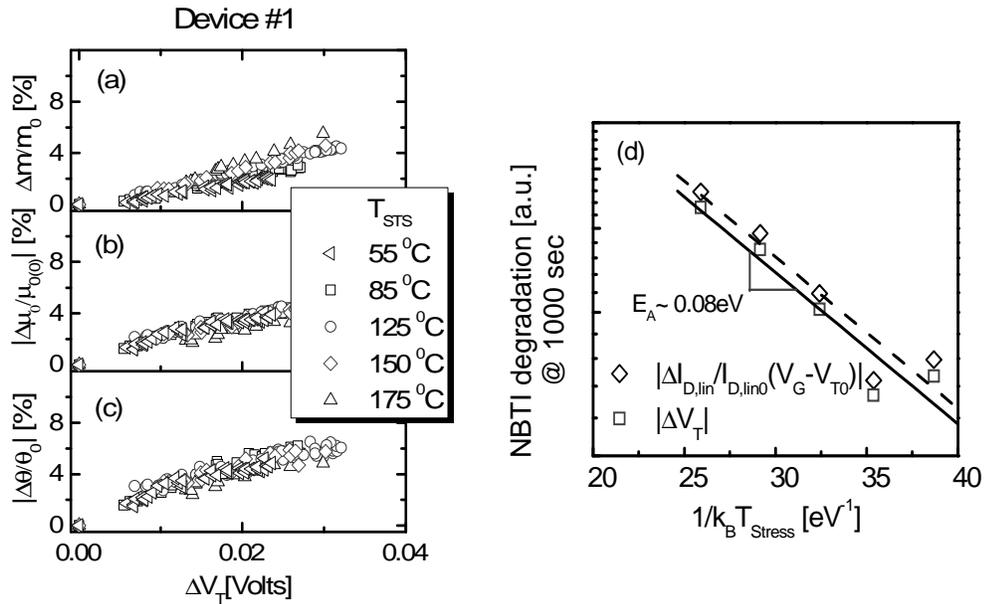


Figure 6.22: (a-c) Correlations for $\Delta\mu_0$, $\Delta\theta$ and Δm with ΔV_T are independent of T_{STS} . Hence, no change in NBTI activation energy (E_A) is observed before and after considering $\Delta\mu_0$, $\Delta\theta$ and Δm in ΔV_T estimation for OTF- $I_{D,lin}$ (d).

Therefore, we conclude that $\Delta\mu_{eff}$ (also Δm to some negligible extent) are playing an important role in ΔV_T estimation for OTF- $I_{D,lin}$. Ignoring these effects will significantly impact the estimated lifetime, voltage acceleration and safe operating condition and keeps power-law time exponent and activation energy the same.

6.7.4.5. Compact ΔV_T Estimation Algorithm for Type-I Transistors

So far we have discussed the methodologies for analyzing $\Delta\mu_{\text{eff}}(\Delta N_{IT})$ for transistors having specific EOT and N_2 dose (see Table 6.1). Utilizing this information, we proposed an algorithm for estimating ΔV_T from $\Delta I_{D,\text{lin}}$, which is one of the major challenges in OTF- $I_{D,\text{lin}}$ technique. However, the question is, “Can this algorithm be generalized so that we can eliminate the steps related to MSM experiments?”. In other words, “Can we avoid measurement of Figure 6.7 (which is obviously found to be different for devices having different EOT and N_2 dose) and replace it with a theoretical formula?”, thus simplifying the process of obtaining ΔV_T in OTF- $I_{D,\text{lin}}$ measurements. Obviously, predicting the results of Figure 6.7 will make the algorithm for ΔV_T estimation from $\Delta I_{D,\text{lin}}$ (Figure 6.18) more compact.

To develop this compact algorithm, first we generalize $\Delta\mu_{\text{eff}}(\Delta N_{IT})$ variation for type-I transistors. Similar to the analysis in Figure 6.6, we determine $\Delta\mu_0/\mu_{0(0)}$ and $\Delta\theta/\theta_0$, and hence –

$$\begin{aligned} P_1 &= d(\Delta\mu_0/\mu_{0(0)})/d(\Delta N_{IT}), \\ P_2 &= d(\Delta\theta/\theta_0)/d(\Delta N_{IT}), \end{aligned} \quad (6.17)$$

for all the type-I transistors (see Table 6.1) under consideration. Comparison of these quantities (*i.e.*, P_1 and P_2) for different type-I transistors eliminates any EOT and substrate doping (N_D) dependency of μ_0 and θ [249, 265] – hence P_1 and P_2 only contain information related to nitrogen dependence within the dielectric of these transistors. Figure 6.23a,b show insignificant correlation for P_1 and P_2 with total N_2 dose ($\%N$) for different type-I transistors. On closer examination, it appears that this discrepancy can be traced to differences in T_{PHY} for these transistors. As nitrogen profile within plasma oxynitride dielectric decreases exponentially [238, 239] (Figure 6.23e), interfacial nitrogen concentration (N_{int}) for devices having same $\%N$, but different T_{PHY} , will be different (for example, among device #3 and #5 in Table 6.1, device #3 is expected to have less N_{int} compared to device 5). So, a better correlation for P_1 and P_2 is expected for type-I transistors if N_{int} is estimated using –

$$N_{int} \sim (\%N)/T_N [\exp(T_{PHY}/T_N) - 1], \quad (6.18)$$

where, T_N is the characteristic decay length of nitrogen profile within the plasma oxynitride film ($\sim 5\text{\AA}$ [238]). Figure 6.23c,d show that both P_1 and P_2 have consistent decreasing trend with increase in N_{int} . In other words, for same amount of ΔN_{IT} , transistors having higher N_{int} will have reduced change in mobility parameters, μ_0 and θ – thus reduced change in μ_{eff} . This further proves that N_{int} (rather than $\%N$) governs N_{IT} dynamics for type-I transistors.

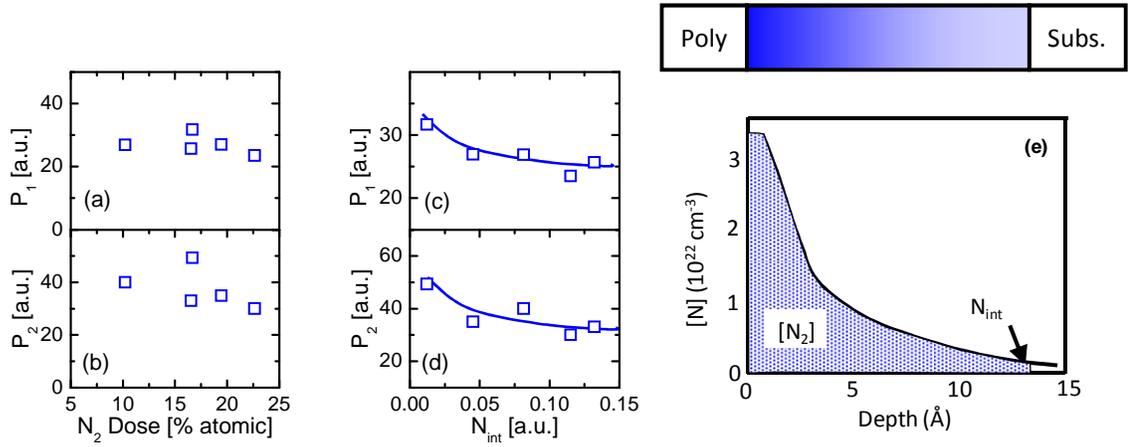


Figure 6.23: EOT and N_D independent parameters, $P_1 = d(\Delta\mu_0/\mu_{0(0)})/d(\Delta N_{IT})$ and $P_2 = d(\Delta\theta/\theta_0)/d(\Delta N_{IT})$, obtained using equation (6.5) and Figure 6.6 show better correlation with (c, d) N_{int} , compared to the one with (a, b) $\%N$. Lines in (c, d) are guide to eye only. (e) Nitrogen profile in a type-I transistor, having plasma nitrided dielectric, shows nitrogen concentration $[N]$ to drop exponentially from poly/oxide interface towards oxide/substrate interface [238]. This information helps us to estimate N_{int} from $\%N$ and T_{PHY} using equation (6.18).

In order to understand the physical basis of the variation in P_1 and P_2 , we estimate the variation of β_{IT} (see section 6.4.4) with N_{int} . Figure 6.24 indicates β_{IT} for type-I transistors also reduces with N_{int} , which signifies that interaction of holes with interface

states [266] is indeed getting reduced with increase in N_{int} . Now, using the systematic variation of P_1 and P_2 in Figure 6.23c, d and neglecting the effect due to SS or, m (which is shown in Figure 6.19d to have negligible effect in ΔV_T estimation from $\Delta I_{D,lin}$), we can estimate $\Delta\mu_0$, $\Delta\theta$ (hence, $\Delta\mu_{eff}$) vs. ΔN_{IT} (similar to Figure 6.6 and Figure 6.7) for transistors having any N_{int} (hence, $\%N$ using equation (6.18)), N_D and EOT . Thus the algorithm proposed in Figure 6.18 can indeed be made more compact obviating the need for any MSM measurement. The resultant algorithm, therefore, becomes generally applicable for any type-I transistors.

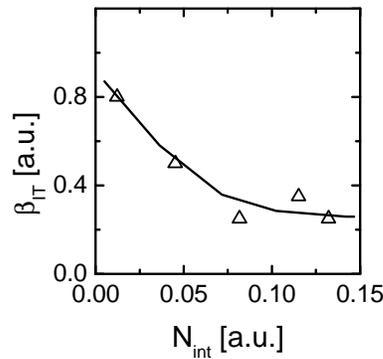


Figure 6.24: Variation of β_{IT} for type-I transistors (lines are guide to eye only) indicates a decrease, similar to the decrease of P_1 and P_2 , observed in Figure 6.23c,d.

6.7.5. N_{HT}/N_{IT} Separation for NBTI Degradation

There is an intense debate in recent literature regarding the magnitude of hole trapping in pre-existing oxide defects (N_{HT}) during NBTI degradation. Some consider N_{IT} generation as the only component in NBTI degradation [33, 115, 128, 130, 143]; whereas others have reported dominance of N_{HT} over N_{IT} in their transistors [32, 141, 202, 291]. Our recent study on transistors having wide range of nitrogen concentration [119, 129, 138, 154, 166] reveals that contribution of N_{HT} in NBTI degradation mainly depends on the amount of nitrogen concentration at the oxide/substrate interface and also on the physical thickness of the dielectric (T_{PHY}). In type-I transistors, which includes transistors

having: a) SiO₂ dielectric, b) dielectric with less than ~20% plasma N₂ dose, and c) thin (*i.e.*, small T_{PHY}) dielectric with thermal N₂ dose, NBTI degradation is dominated by N_{IT} generation with negligible amount of N_{HT} . Whereas in type-II transistors, which includes transistors having: a) dielectrics with more than ~20% plasma N₂ dose, b) thick dielectric with thermal N₂ dose, and c) high- κ dielectric, N_{HT} plays the dominant role in NBTI degradation. In general terms, both N_{IT} and N_{HT} play critical role in NBTI degradation (one dominating over the other, depending on the type of transistors under study) and corresponding ΔV_T can be expressed as:

$$\Delta V_T = \Delta V_{IT} + \Delta V_{HT} = \frac{q\Delta N_{IT}(t)}{C_{di}} + \frac{\int_0^{T_{PHY}} \int_{\langle E \rangle} x \rho_{HT}(x, E, t) dE dx}{C_{di} T_{PHY}}, \quad (6.19)$$

where ΔV_{IT} and ΔV_{HT} refer to the contributions to ΔV_T from N_{IT} and N_{HT} , respectively; $\rho_{HT}(x, E, t)$ represents trapped holes at location x (measured into the oxide from the poly/oxide interface) and at energy E_T ; and $\langle E \rangle$ represents the range of trap energy levels within the dielectric bandgap.

It is very important to separate out components of N_{IT} and N_{HT} from ΔV_T , especially for type-II transistors. Otherwise, it will be difficult to comprehend and theoretically model the dynamics of ΔV_T (and its components), measured during NBTI stress. Without performing such decomposition of N_{IT} and N_{HT} components, recently several efforts [141, 170, 292, 293] have been made to understand the differences between NBTI-induced ΔV_T relaxation (measured using ultra-fast methods [140, 141, 294]) and the N_{IT} relaxation predicted by the R-D theory. Though, R-D model explains many experimental NBTI signatures for p-MOSFETs having lightly nitrided oxides (including stress-phase time exponent, activation energy, field acceleration, frequency independence, *etc.* [119, 129, 167]), recent reports of ultra-fast NBTI relaxation that initiates at $\sim \mu\text{s}$ time-scale is inconsistent with N_{IT} dynamics, predicted by the R-D model (Figure 6.25a), where H-H₂ conversion takes place at poly-Si/dielectric interface and long-term diffusion of H₂ occurs in the poly-Si gate [31] (see section 3.5.5 for details). This inconsistency has raised questions regarding the general validity of the R-D theory [31]. Many alternative theories

[293, 295] have also been proposed to explain ultra-fast NBTI relaxation, although their ability in predicting the features of NBTI stress remains questionable.

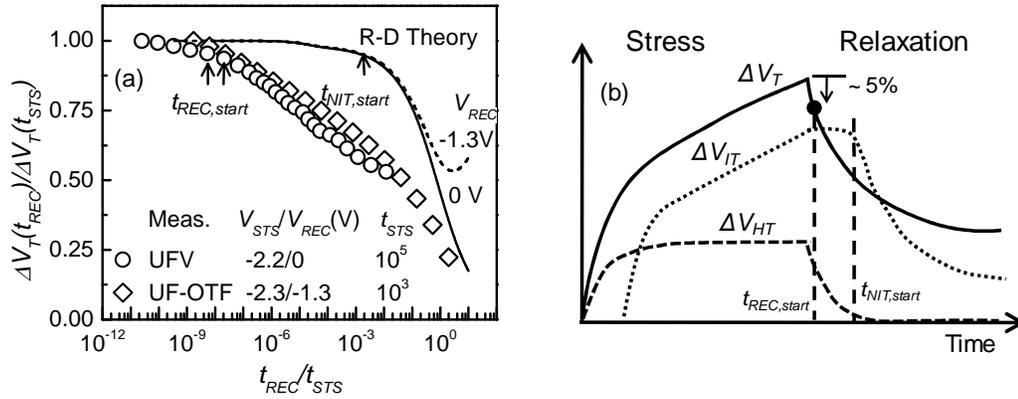


Figure 6.25: (a) $t_{REC, start}$ in UFV measurement [140], commences ~ 5 decades earlier in time compared to the prediction of H-H₂ R-D theory. Our OTF-I_{D,lin} (with $t_0 = 1\mu\text{s}$) measurement also shows similar relaxation trend. The slight difference between UFV and OTF-I_{D,lin} can be attributed to the valence band electron trapping phenomena, explained in section 6.7.2. (b) Resolution between the measured ultra-fast ΔV_T relaxation and comparatively slower N_{IT} relaxation becomes possible, after decomposing ΔV_T (solid line) into ΔV_{IT} (dotted line), ΔV_{HT} (dashed line) components; each having separate time-dynamics.

In this section, we present a phenomenological decomposition method [166, 296] for isolating N_{IT} and N_{HT} components from ΔV_T measured during NBTI stress and relaxation. To establish the decomposition procedure, we measure NBTI stress and relaxation on transistors having a wide range of thickness (EOT of 1.2-2.2 nm) and plasma N₂ dose (16-42 %N). We identify the signatures of NBTI (section 6.7.5.1), measured during the stress and relaxation phase. Applying such decomposition in NBTI stress-phase experiments (section 6.7.5.2), it is shown that once N_{HT} has been subtracted out, the time and temperature dependence of the extracted N_{IT} part are fully consistent with the R-D model having non-dispersive molecular H₂ diffusion (see sections 3.5 and 3.7). Later, the same decomposition scheme is applied to explain NBTI relaxation experiments (section 6.7.5.3). We show that N_{IT} relaxes slower compared to overall NBTI and this N_{IT}

relaxation is consistent with the predictions of R-D theory. Thus, the decomposition enables us to explain the difference between the start of overall NBTI relaxation ($t_{REC,start} \sim \mu\text{s}$) and N_{IT} relaxation ($t_{NIT,start} \sim \text{sec}$). The proposed method anticipates the duty-cycle/frequency dependencies of AC NBTI stress (section 6.7.5.4) and establishes the AC NBTI dependencies on nitrogen content ($\%N$), temperature (T_{STS}), and stress time (t_{STS}). Finally, we conclude the discussion on N_{IT}/N_{HT} decomposition by comparing our proposed technique with the one available in literature [140, 208, 241, 297] (see section 6.7.5.5).

6.7.5.1. Basis for N_{HT}/N_{IT} Decomposition Method

The dielectric properties of PMOS transistors (W/L=15/0.16 μm) used for establishing this decomposition method are shown in Table 6.2. On these transistors, we measure NBTI stress and relaxation using OTF- $I_{D,lin}$, having $t_0 \sim 1\mu\text{s}$ [154], and later use equation (6.14) for estimating ΔV_T . Mobility correction (discussed in 6.7.4) is difficult, especially for the type-II transistors (device #2,4,5,7,8 in Table 6.2) used in this section, hence ignored in the ΔV_T estimation. As time exponent (n) and activation energy (E_A) are the main parameters for establishing the separation algorithm and since both of them remain unaffected by mobility correction (see section 6.7.4.4), mobility correction should have no impact on the conclusions obtained in the subsequent discussions.

Table 6.2: Properties ($\%N$, T_{PHY} and EOT) of plasma oxynitride dielectric for the transistors used in section 6.7.5.

| Device | N ₂ Dose [atomic %], $\%N$ | T_{PHY} [Å] | EOT [Å] | Type |
|--------|---|---------------|-----------|------|
| #1 | 22.64 | 18.48 | 14.0 | I |
| #2 | 41.27 | 21.12 | 12.3 | II |
| #3 | 19.45 | 22.34 | 17.7 | I |
| #4 | 34.58 | 23.16 | 15.6 | II |
| #5 | 42.48 | 24.37 | 14.6 | II |
| #6 | 16.67 | 28.13 | 23.5 | I |
| #7 | 29.42 | 28.53 | 21.4 | II |
| #8 | 35.92 | 28.85 | 19.9 | II |

Stress Phase Signatures

Figure 6.26a-c show the measured ΔV_T at two different T_{STS} and identical E_{ox} on transistors, having similar T_{PHY} (~ 24 Å) but different $\%N$. As $\%N$ is increased, transistors tend to show more ΔV_T at similar E_{ox} , especially at short t_{STS} .²⁰ Moreover, temperature dependence of NBTI degradation at a particular t_{STS} (or $E_A@t_{STS}$) reveal the existence of a short-time, temperature-independent fast component (that saturates within \sim ms) in transistors having high $\%N$ plasma nitrided oxide (Figure 6.26c). Decrease in $\%N$ within the dielectric reduces the contribution from this fast component (Figure 6.26a) and shows the existence of temperature-activated slow component, even at short t_{STS} . Moreover, at long t_{STS} (>1 s), all transistors of Figure 6.26 show some degree of temperature activation, which is less in higher $\%N$ transistor. Since ΔV_{HT} is fast and T_{STS} independent compared

²⁰ Similar results are also presented in [138] using different types of nitridation process, where it is shown that increase in N_{int} plays the major role in the observed behavior.

to slow and T_{STS} dependent ΔV_{IT} [119, 129], the NBTI stress experiments on nitrided transistors indicate the dominance of ΔV_{HT} at short t_{STS} and the dominance of ΔV_{IT} at long t_{STS} . At short t_{STS} (<ms), the dominance of ΔV_{HT} over ΔV_{IT} (more for higher %N transistors) can explain the observed temperature insensitivity in ΔV_T . On the other hand, at long t_{STS} , dominance of ΔV_{IT} results significant temperature dependence for ΔV_T . Moreover, the increase in ΔV_{HT} with the increase in %N is also consistent with the decrease in power-law time exponent (n for $\Delta V_T \sim t^n$; see Figure 6.26d) for higher %N transistors [129]. Similar temperature independence at short t_{STS} was also observed in [Reisinger]. However, corresponding %N for the transistor in [140] (Figure 6.26d) is much lower than the one in Figure 6.26c, which has similar temperature independence at short t_{STS} . To explain this, note that in addition to %N, the quality of post-nitridation anneal (PNA) plays significant role for the presence of ΔV_{HT} [138], which leads to reduced n , even for low %N transistors (see Figure 6.26e). Possible presence of such low-quality PNA can explain the observation of temperature independent ΔV_T at short t_{STS} in [140], even for smaller %N.

Thus, Figure 6.26 allows us to conclude that there is significant amount of ΔV_{HT} in ΔV_T for type-II transistors and $\Delta V_{HT}/\Delta V_T$ increases with the increase in %N or N_{int} . Presence of this weakly temperature dependent ΔV_{HT} component is responsible for the observed reduction in E_A (with increase in %N) for $\Delta V_T (= \Delta V_{HT} + \Delta V_{IT})$ in these transistors, especially at short t_{STS} . Moreover, the reduction in long- t_{STS} n with the increase in %N indicates that the weakly temperature dependent ΔV_{HT} component has saturated at long- t_{STS} (presumably before $t_{STS} \sim 1$ sec, see Figure 6.25b). Such saturation of hole trapping at long- t_{STS} is also supported by our modified Shockley-Read-Hall (SRH) analysis [119, 129] of hole trapping, presented in section 5.4. Based on these, we can approximate equation (6.19) with –

$$\Delta V_T = \Delta V_{IT} + \Delta V_{HT} = A_{IT} t^{1/6} + B, \quad (6.20)$$

where, interface trap component of ΔV_T (*i.e.*, ΔV_{IT}) is assumed to follow power law dependence with $n \sim 1/6$ as observed by very long- t_{stress} measurements [128, 175, 298]²¹, which is less affected by saturated ΔV_{HT} , and which is also supported by a classical H-H₂ R-D model (see section 3.5.6).

Relaxation Phase Signatures

Similar to the stress phase NBTI experiments, our NBTI relaxation measurements also indicate the co-existence of N_{IT} and N_{HT} . As shown in Figure 6.27a, OTF- $I_{D,lin}$ measurement of fractional NBTI relaxation indicates a dependence on $\%N$ of the dielectric, as well as on the difference between stress and recovery voltages ($V_{STS}-V_{REC}$). Start of significant NBTI relaxation ($t_{REC,start}$) is larger (\sim ms) for low $\%N$ and smaller ($V_{STS}-V_{REC}$). This is due to the existence of different (N_{HT} and N_{IT}) species during stress [299], having very different dynamics during NBTI relaxation. We can easily identify the presence of such separate relaxation dynamics for N_{IT} and N_{HT} from temperature dependent OTF- $I_{D,lin}$ measurement of Figure 6.27b. These relaxation measurements clearly indicate weak temperature dependence up to $t_{REC} \sim$ ms, which indicates the presence of hole detrapping from pre-existing oxide defects as the predominant mechanism upto $t_{REC} \sim$ ms at $V_{REC} = -1.3V$. Therefore, the start of significant amount of temperature dependent N_{IT} relaxation ($t_{NIT,start}$, defined in Figure 6.25b) is also \sim ms for the transistor in Figure 6.27b. Moreover, OTF- $I_{D,lin}$ at $V_{REC} \leq -1.8V$ shows insignificant relaxation up to \sim ms (Figure 6.28a) and also indicates significant consistency with the relaxation, predicted by R-D theory. This enables us to conclude that hole detrapping occurs predominantly for $V_{REC} > -1.8V$, and that the amount of hole detrapping is similar

²¹ Reduction of electric field at constant V_G stress (see section 3.10) will definitely reduce long- t_{stress} time exponent below 1/6 [119], especially for type-II transistors having large ΔV_{HT} . So, without using appropriate correction due to this electric field reduction, the N_{HT}/N_{IT} separation algorithm presented in 6.7.5.2 will underestimate ΔV_{HT} or overestimate ΔV_{IT} .

from $V_{REC} = -1.3V$ to $-1.6V$ (Figure 6.28a). Thus, trapping sites within the quasi-Fermi levels at $V_{REC} = -1.8V$ and $-1.6V$ (shown schematically by the hatched region in Figure 6.28c) will detrapp all the holes that were captured during stress within a timescale of $\sim ms$.

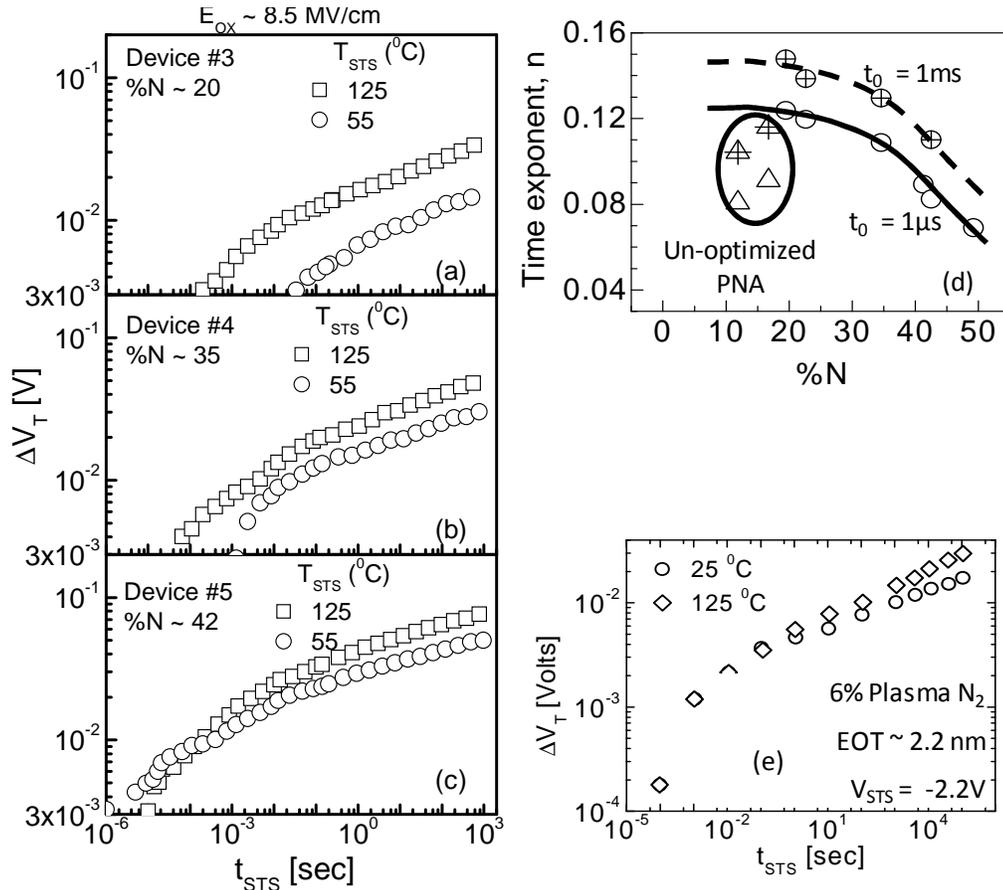


Figure 6.26: (a-c) Time evolution of degradation under NBTI stress at identical E_{ox} , but different T_{STS} , for plasma nitrided transistors having similar T_{PHY} ($\sim 24 \text{ \AA}$) but different $\%N$ (hence, different N_{int}) and EOT . (d) Power-law NBTI time exponent (n) decreases with increase in $\%N$, indicating the presence of ΔV_{HT} for higher $\%N$. In addition, n shows an increment with increase in t_0 -delay (see section 6.7.3). (d) Similar temperature independence was also reported for lightly dosed nitrided transistors, using UFV [140]. Moreover, the effect of un-optimized PNA (leading to significant ΔV_{HT} , even at low $\%N$ [138]) is also evident in Fig. e.

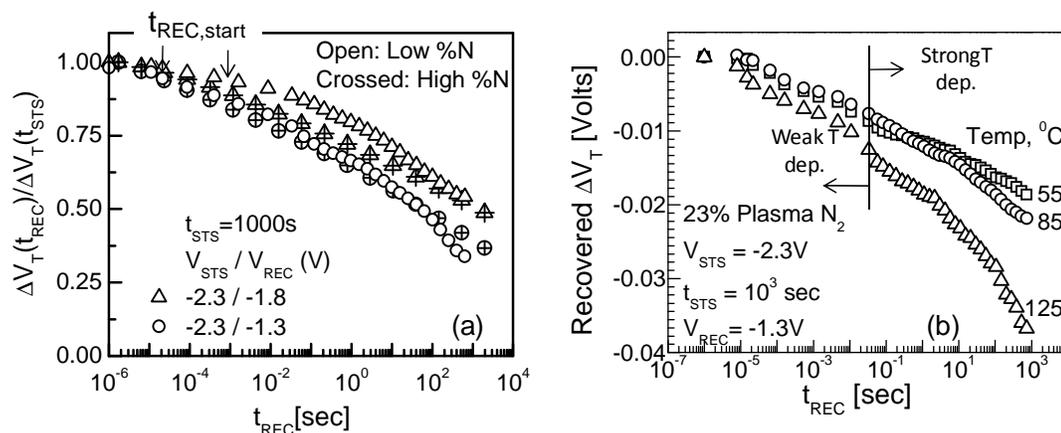


Figure 6.27: (a) Though $t_{REC,start}$ is considered to be $\sim \mu s$ in several reports [168, 278, 293], our OTF- $I_{D,lin}$ measurements on nitrided transistors, at different V_{REC} , demonstrate that $t_{REC,start}$ depends on %N and $(V_{STS} - V_{REC})$. (b) Similar to NBTI stress phase (Figure 6.26), early relaxation phase (up to $\sim ms$) is approximately temperature independent.

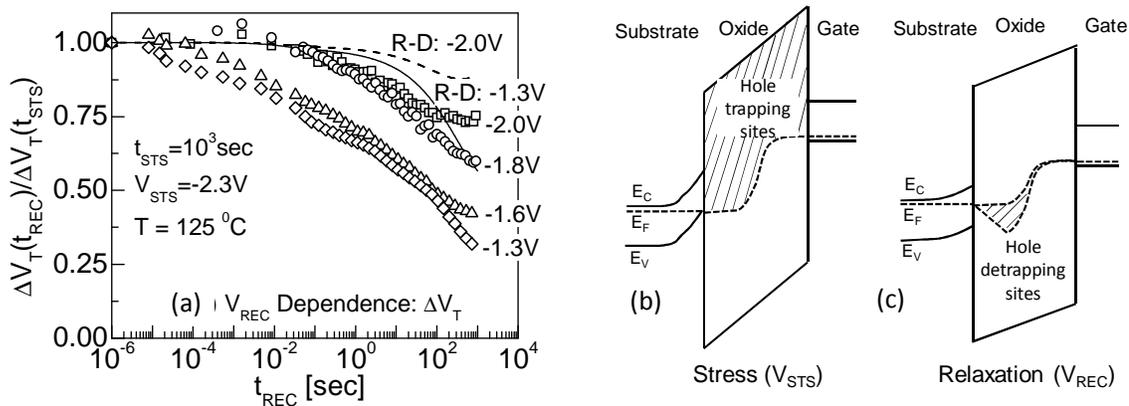


Figure 6.28: (a) NBTI relaxation at different V_{REC} show significantly different $t_{REC,start}$. Since relaxation for $V_{REC} \leq -1.8V$ is very close the R-D theory, there is an additional relaxation mechanism for $V_{REC} > -1.8V$. Our observation suggests hole detrapping to be the additional mechanism. (b) Schematic (based on simulation within a Shockley-Read-Hall trapping-detrapping framework [119]) for expected hole trapping sites (hatched region) at V_{STS} . (c) When gate bias is switched from V_{STS} to V_{REC} , the hatched region will detrapp the captured hole in a temperature independent manner.

Similar temperature independence at short t_{REC} and larger ($V_{STS}-V_{REC}$) is also observed in transistors of Table 6.2. As shown in [300], $t_{NIT,start}$ demonstrates $\%N$ dependence and is generally larger for transistors with higher $\%N$.

6.7.5.2. N_{HT}/N_{IT} Separation: NBTI Stress Phase

Based on the observations in section 6.7.5.1, we use equation (6.20) and estimate ΔV_{IT} by subtracting a constant (saturated) ΔV_{HT} from ΔV_T for $t_{STS} > 1$ s in such a way that it provides time exponent $n \sim 1/6$ for ΔV_{IT} at t_{STS} of 1-1000sec (Figure 6.29a). Our temperature dependent extraction at constant E_{ox} (Figure 6.29b) shows that E_A (at fixed t_{STS}) for ΔV_{IT} or, $E_{A,IT}$ is ~ 0.094 eV; which is expected for R-D model based interface defect generation with H_2 diffusion [119, 129, 299]. On the other hand, E_A for ΔV_{HT} or $E_{A,HT}$ is ~ 0.04 eV, which is typically expected in any hole trapping process, involving elastic tunneling (see section 5.5.2). The signature of $E_{A,IT} > E_{A,HT}$ is also evident from Figure 6.29c, which indicates an increase in the extracted $\Delta V_{IT}/\Delta V_T$ (*i.e.* decrease in $\Delta V_{HT}/\Delta V_T$) with increase in temperature, at fixed t_{STS} . The error bars in Figure 6.29 result from noise in $I_{D,lin0}$ measurement for OTF- $I_{D,lin}$ [294], which causes a ± 0.005 error in n for ΔV_T and a ± 1 mV error in estimated ΔV_{HT} .

Identical procedure is followed to isolate ΔV_{IT} and ΔV_{HT} for the transistors of Table 6.2, at all voltages and temperatures. The extracted $E_{A,HT}$ (supported by R-D theory of section 3.5.3) and $E_{A,IT}$ (supported by hole trapping theory of section 5.5) is similar for these transistors (see Figure 6.30a). Moreover, extracted ΔV_{HT} for these nitrided transistors at a particular E_{ox} , T_{STS} and t_{STS} (Figure 6.30b) is observed to increase significantly with the increase in $\%N$ (with a very rapid increase seen for $\%N > 30$), which is indeed a signature of higher hole trapping for higher $\%N$. In addition, ΔV_{IT} is also observed to increase slightly with $\%N$. Therefore, we observe clear signatures of $\%N$ dependent ΔV_{HT} in the transistors under study. Based on these observations, we decompose the N_{HT} and N_{IT} components for NBTI stress experiments and show that the extracted parameters for both N_{HT} and N_{IT} are consistent with the theoretical predictions.

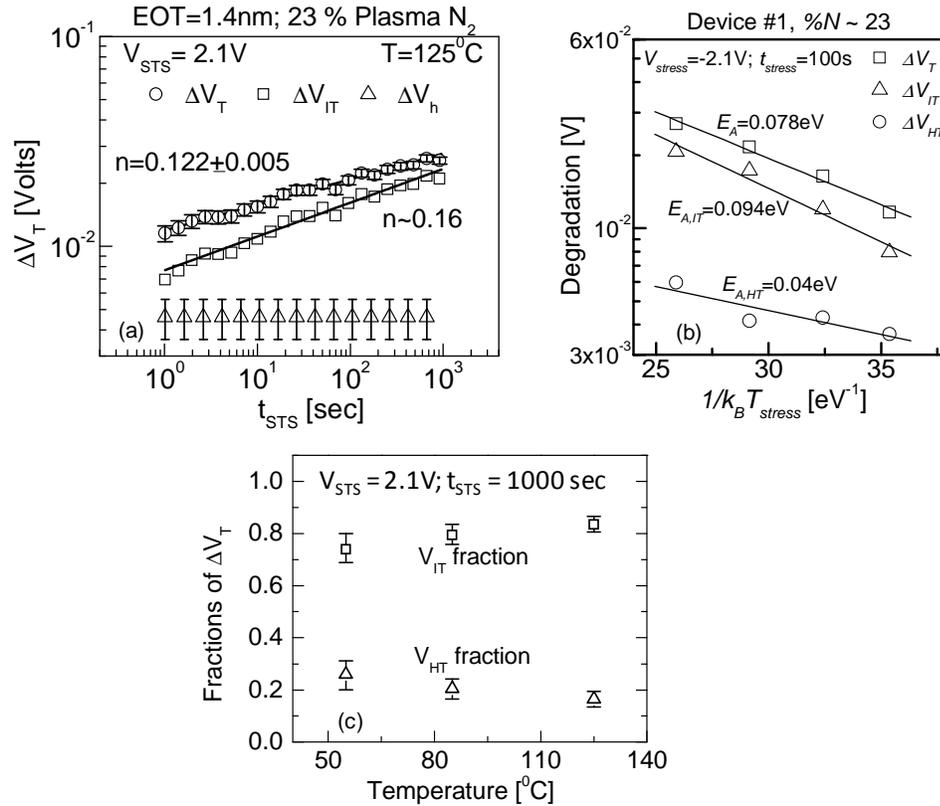


Figure 6.29: (a) Measured ΔV_T and extracted ΔV_{IT} and ΔV_{HT} components for device #1 in Table 6.2. ΔV_{IT} and ΔV_{HT} are extracted according to the decomposition method proposed in 6.7.5.2. Here, the effect of mobility [165] and electric field-reduction [119] is taken into account in estimating ΔV_T . (b) Temperature dependence and corresponding activation energies (E_A) of ΔV_T , ΔV_{IT} and ΔV_{HT} components for the same transistor. (c) Estimated $\Delta V_{HT}/\Delta V_T$ and $\Delta V_{IT}/\Delta V_T$ at different temperature indicate an increase (decrease) in $\Delta V_{IT}/\Delta V_T$ ($\Delta V_{HT}/\Delta V_T$) at higher temperature, which is primarily due to the temperature independence of the ΔV_{HT} component.

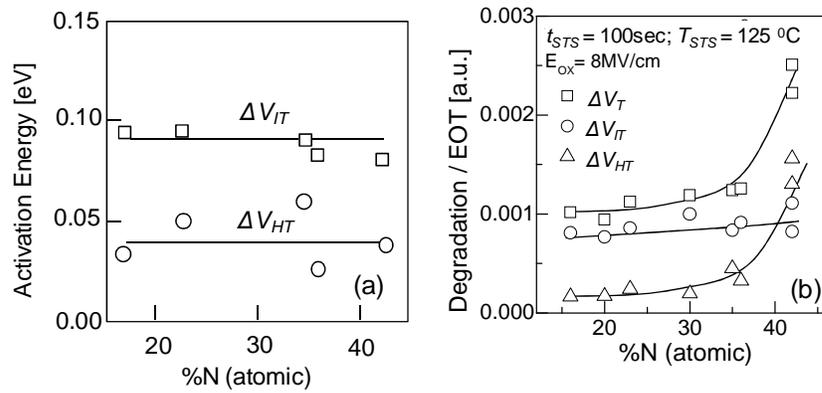


Figure 6.30: (a) Activation energy for extracted ΔV_{IT} and ΔV_{HT} components indicates negligible %N dependence. (b) %N Dependence of measured ΔV_T and extracted ΔV_{IT} and ΔV_{HT} components for a particular stress time, temperature and electric field. Lines are guide to the eye only.

6.7.5.3. N_{HT}/N_{IT} Separation: NBTI Relaxation Phase

So far we have used a phenomenological N_{HT}/N_{IT} separation technique during NBTI stress phase for obtaining individual defect components. Temperature independence of short-term t_{STS} NBTI measurement is the main empirical feature, used in this separation technique. Similar temperature independence is also observed in the early part of NBTI relaxation for both type-I (up to $t_{REC} \sim \text{ms}$ in Figure 6.27b and Figure 6.31) and type-II (up to $t_{REC} \sim 1\text{sec}$ in Figure 6.31) transistors.

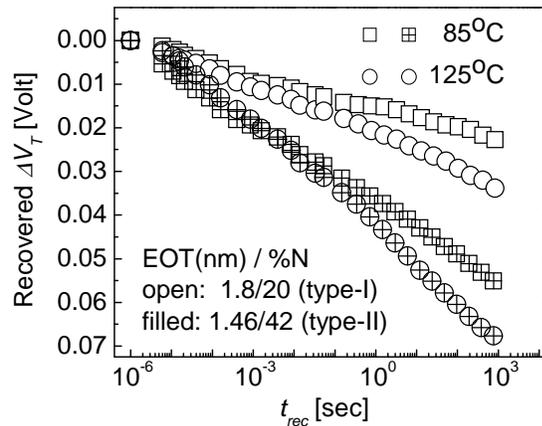


Figure 6.31: Time evolution of NBTI relaxation at different temperature for both type-I (device #3 in Table 6.2) and type-II (device #5 in Table 6.2) transistors. Initial relaxation, up to $t_{REC} \sim 10\text{ms}$ in type-I and $\sim 1\text{s}$ in type-II transistor, is temperature independent, thus indicating a hole detrapping process.

Now, let us apply the N_{HT}/N_{IT} decomposition in relaxation measurements on the type-I transistor (device #1 in Table 6.2), where we have also performed a stress-phase decomposition (see Figure 6.29). Moreover, as discussed in Figure 6.28, for $V_{STS} = -2.3\text{V}$, N_{HT} starts to show significant amount of temperature independent relaxation for $V_{REC} > -1.6\text{V}$. Considering such total hole detrapping at $V_{REC} = -1.3\text{V}$, $t_{NIT,start} \sim \text{ms}$ is obtained at all V_{REC} , which is comparable with the relaxation predicted by R-D theory (Figure 6.32a).

Remaining theory-experiment gap ($\times 10$ in t_{STS} , or 10% in N_{IT} relaxation) is reduced (Figure 6.32b), if it is realized that NBTI is also TDDB stress on p-MOSFETs [156]. Hence, in addition to pre-existing oxide defects, small amount of oxide defect generation is also possible during NBTI stress. As shown in Figure 6.33a, time exponent of NBTI degradation increases at higher V_{STS} , thus indicating generation of oxide defect [156] or N_{OT} (and associated ΔV_{OT}) at higher V_{STS} . Considering negligible ΔV_{OT} at lower V_{STS} (*i.e.*,

$\Delta V_T \sim \Delta V_{IT}$)²² and calculating field acceleration for ΔV_{IT} , we can estimate ΔV_{IT} and hence calculate $\Delta V_{OT} = \Delta V_T - \Delta V_{IT}$ at higher V_{STS} (Figure 6.33b). The calculated ΔV_{OT} for the measurements in [15] shows a decrease in time exponent for ΔV_{OT} at higher V_{STS} , as also observed in [114]. Interestingly, the estimated ΔV_{OT} at different V_{STS} scales universally and can be fitted (see Figure 6.33c) using [114, 301] –

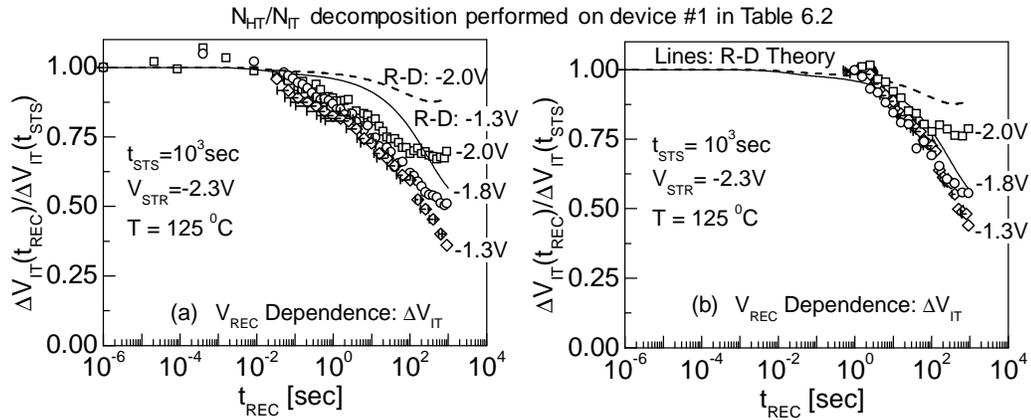


Figure 6.32: (a) Considering total hole detrapping at $V_{REC} = -1.3V$ in the early phase of NBTI relaxation brings $t_{NIT,start}$ at all V_{REC} closer to the relaxation, predicted by R-D theory. The error bar along the x-axis for $V_{REC} = -1.3V$ is due to the $I_{D,lin0}$ error in calculating $\Delta V_{IT}/\Delta V_T$ (see Figure 6.29). (b) Furthermore, if one assumes negligible amount of oxide defect generation during NBTI stress (with $\Delta V_{OT} \sim 3mV$ at $t_{STS} = 10^3$ sec, having the time dependence similar to Figure 6.33c), then consideration of hole detrapping from these generated oxide defect brings $t_{NIT,start}$ more closer to the predictions of R-D theory.

²² Note that these are slow measurements (like MSM), hence should not have any ΔV_{HT} component in ΔV_T .

$$\begin{aligned}
 N_{OT} &= \int n_{OT}(E) dE, \\
 n_{OT}(E) &= g(E) [1 - \exp(-k_F(E) * t)], \\
 k_F(E) &= k_{F0} \exp[-(E - E_0) / kT], \\
 g(E) &\sim \frac{1}{\sigma_{OT}} \frac{\exp[(E - E_0) / \sigma_{OT}]}{[1 + \exp((E - E_0) / \sigma_{OT})]^2}
 \end{aligned} \tag{6.21}$$

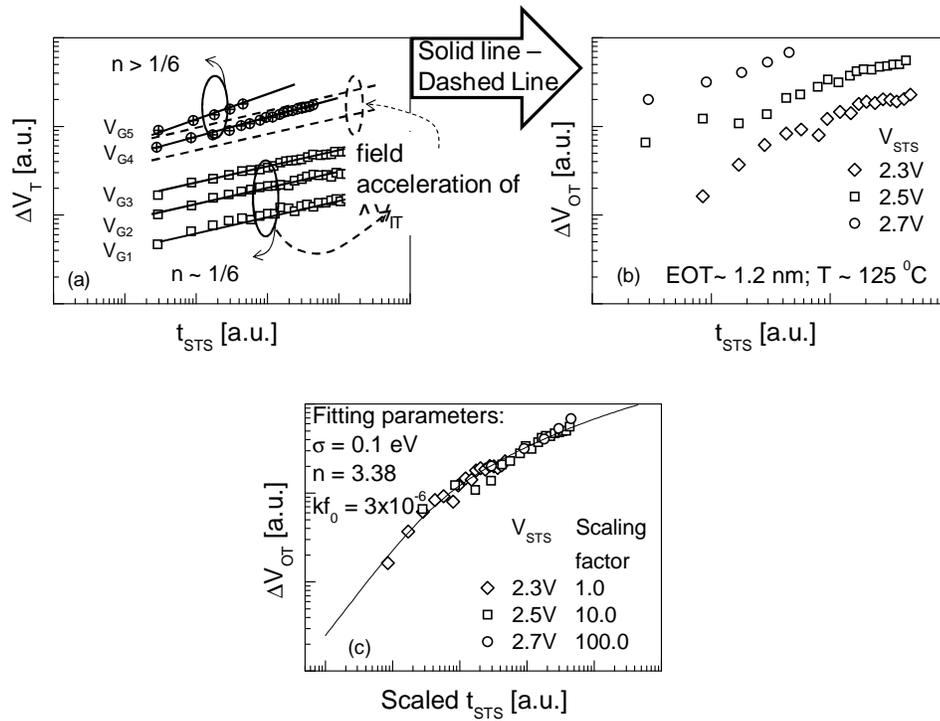


Figure 6.33: (a) Time exponent of NBTI degradation (fitted solid lines) increases at $V_{STS} \geq V_{G4}$ [175], thus indicating clear presence of N_{OT} (or ΔV_{OT}) at these V_{STS} . By assuming negligible ΔV_{OT} at $V_{STS} \leq V_{G3}$, we can estimate ΔV_{IT} at $V_{STS} \geq V_{G4}$ (dashed lines) and hence calculate $\Delta V_{OT} (= \Delta V_T - \Delta V_{IT})$ at $V_{STS} \geq V_{G4}$. (b) Similar calculation of ΔV_{OT} for the measurements in [175]. (c) Estimated ΔV_{OT} at different V_{STS} scales universally (similar to the observation in [114]) and can be fitted using equation (6.21).

where E_0 is the average bond-dissociation energy and σ_{OT} is its standard deviation. Based on this, let us assume a negligible amount of N_{OT} -induced $\Delta V_{OT} \sim 3$ mV at $t_{STS} = 10^3$ sec

for the transistor of Figure 6.32 (having the time dependence similar to Figure 6.33c). Considering hole trapping/ detrapping from these generated oxide defects, Figure 6.32c suggests that $t_{NIT,start}$ will come more closer to the relaxation, predicted by R-D theory. Therefore, decomposition of contributions from interface and oxide defects enables us to explain the NBTI relaxation features in a theoretically consistent way.

6.7.5.4. Implications for AC/DC Ratio

So far, we have applied the decomposition procedure to identify the N_{HT} and N_{IT} components of NBTI-induced ΔV_T . Our analysis reveals that hole trapping and detrapping occurs at similar time-scales (for example, compare time-scale of the temperature independent hole trapping component in Figure 6.26c, and hole detrapping component in Figure 6.31). Therefore, we expect total hole detrapping for a AC NBTI stress (measured at the end of OFF-state) with $\leq 50\%$ duty cycle (Figure 6.34a). In other words, AC/DC ratio for $\leq 50\%$ duty cycle in high $\%N$ transistors will measure $\Delta V_{T,AC}/\Delta V_{T,DC} \sim \Delta V_{IT,AC}/[\Delta V_{HT,DC} + \Delta V_{IT,DC}]$ and hence will always be less than the contribution from N_{IT} 's component, $AC/DC(N_{IT}) = \Delta V_{IT,AC}/\Delta V_{IT,DC}$, predicted by R-D theory [167] (Figure 6.34b). Moreover, as ΔV_{HT} decreases for smaller $\%N$, low $\%N$ transistors will have $\Delta V_{T,AC}/\Delta V_{T,DC} \sim \Delta V_{IT,AC}/\Delta V_{IT,DC}$, and thus the measured AC/DC ratio should be consistent with the one, obtained from R-D theory (Figure 6.34b).

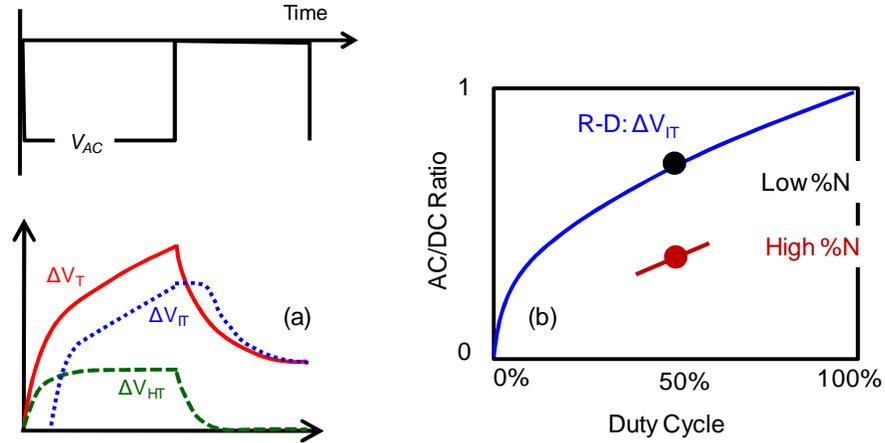


Figure 6.34: (a) Since hole trapping and detrapping happens at similar time-scale, we expect total hole detrapping, at the end of OFF-state during AC NBTI stress. Therefore, (b) AC/DC ratio (with ΔV_T for AC NBTI measured at the end of AC cycles) will show consistency with R-D theory only for transistor having lower ΔV_{HT} or lower $\%N$.

Indeed, our $\Delta V_{T,AC}/\Delta V_{T,DC}$ measurement on nitrided transistor show significant consistency with the predictions from R-D theory for low $\%N$ (Figure 6.35a), where $\Delta V_T \sim \Delta V_{IT}$. Consideration of ΔV_{HT} for DC NBTI stress, in addition to ΔV_{IT} , can also explain the cases for high $\%N$ transistors. The shape of $\Delta V_{T,AC}/\Delta V_{T,DC}$ vs. duty cycle plot is similar to the R-D's prediction for high $\%N$ transistors, even up to $\sim 80\%$ duty cycle; which indicates that time scale for hole-detrapping is much less than the time scale for hole trapping for this type-II transistor. Moreover, consideration of total hole detrapping happens at $\sim 50\%$ duty cycle can also explain why R-D theory can explain the widely observed frequency independence in NBTI measurements, while overestimating the measured $\Delta V_{T,AC}/\Delta V_{T,DC}$ at different frequencies. Since ΔV_{HT} will be absent for AC NBTI stress with $\sim 50\%$ duty cycle, the frequency independence of $\Delta V_{IT,AC}/\Delta V_{IT,DC}$ (following R-D theory [167]) will also result in frequency independence of $\Delta V_{T,AC}/\Delta V_{T,DC}$; except, the magnitude of $\Delta V_{T,AC}/\Delta V_{T,DC}$ will be lower than $\Delta V_{IT,AC}/\Delta V_{IT,DC}$ due to the presence of ΔV_{HT} in $\Delta V_{T,DC}$. Therefore, $\Delta V_{T,AC}/\Delta V_{T,DC}$ measurement on nitrided transistor show frequency independence, irrespective of $\%N$ (Figure 6.35b); while the magnitude of $\Delta V_{T,AC}/\Delta V_{T,DC}$ is only consistent with R-D theory for low $\%N$ transistors. Thus the co-

existence of N_{HT} and N_{IT} and their decomposition can explain both the duty cycle and frequency dependent NBTI experiments on nitrided transistors. Moreover, as the effect of ΔV_{HT} is higher at lower temperature (see Figure 6.29) or at lower t_{STS} (Figure 6.26), AC/DC ratio (or $\Delta V_{T,AC}/\Delta V_{T,DC}$) for lower temperature or lower t_{STS} will always be less than $AC/DC(N_{IT}) = \Delta V_{IT,AC}/\Delta V_{IT,DC}$ in high %N transistors.

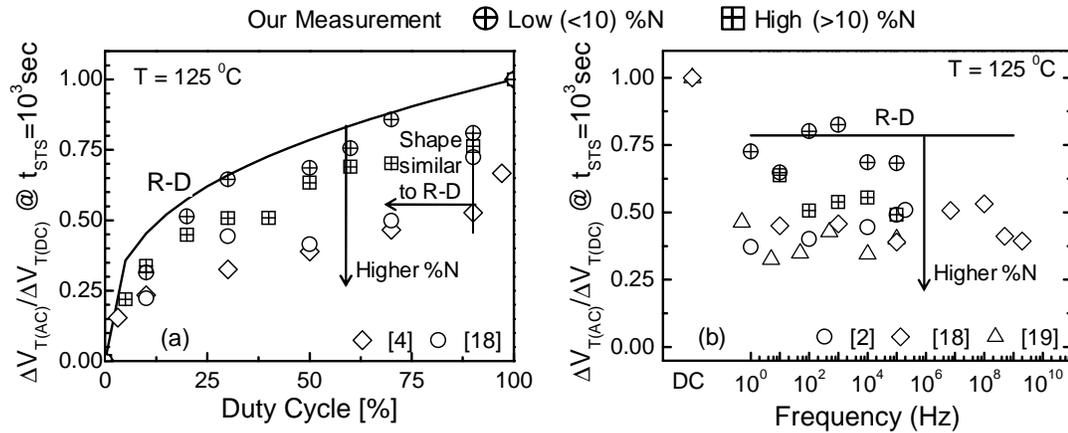


Figure 6.35: (a) AC/DC ratio (when ΔV_T is measured at the end of AC cycles) vs. duty cycle plot for different nitrided transistors. The experiments show significant consistency with the prediction of R-D theory (solid line) for low %N. (b) Though AC/DC ratio for any %N is always frequency independent, there is significant %N dependency due to the presence of ΔV_{HT} in DC NBTI stress.

In general, Figure 6.35 can be used to roughly estimate ΔV_{HT} , which appears to be significant for the transistor reported in [278, 292, 293]. This is again consistent with $n \sim 0.11$ for $t_0 \sim 1$ ms [293] in this transistor; such small n is expected for high %N or unoptimized PNA dielectrics (see Figure 6.26d). Thus R-D theory – appropriately augmented with the effects of hole trapping/detrapping – can explain both the duty-cycle and the frequency dependent NBTI degradation in any nitrided oxide.

6.7.5.5. Comparison with Existing N_{HT}/N_{IT} Separation Techniques

Finally, we compare/contrast the proposed N_{HT}/N_{IT} separation technique with the ones already available in literature. Separation technique proposed in [140, 208, 297] (technique-1) uses long- t_{STS} time exponent of $\sim 1/4$ for the ΔV_{IT} component, as predicted by classical R-D model with H diffusion (see section 3.5 for details on this model and its inconsistency with experimental data). We repeat the procedure of technique-1 and observe that the extracted $E_{A(IT)}$ values (in the range of 0.075-0.1eV as $\%N$ is varied) are not consistent with the one expected from classical H-diffusion based R-D model, which predicts $4E_{A(IT)}$ (activation energy for H diffusion) in the range of 0.1-0.2eV [302]. Moreover, technique-1 extracts $E_{A(HT)}$ of $\sim 0.06-0.07$ eV, which cannot be physically justified without assuming a strong phonon-assisted hole trapping process.

On the other hand, N_{HT}/N_{IT} separation technique (technique-2) proposed in [241] assumes ΔV_{HT} to be equal to ΔV_T obtained at $t_{STS} = 1$ sec. As such, extracted $\Delta V_{IT}(t_{STS}) = \Delta V_T(t_{STS}) - \Delta V_T(t_{STS} = 1\text{sec})$ shows power-law time exponent of ~ 0.19 , which is higher than the one obtained from long-term stress data [128, 175, 298]. Repetition of technique-2 in our transistors also provides $E_{A(HT)}$ of > 0.06 eV, which again cannot be physically justified without assuming a strong phonon-assisted hole trapping process.

6.8. Suggestions for NBTI Characterization

Based on the above analysis of NBTI characterization techniques, we provide the following guidelines for characterizing interface trap/oxide defects, present in MOS transistors during NBTI stress (which can be extended/ modified for other reliability phenomena, as well):

- Spin Dependent Recombination (SDR) experiment is suitable for identifying the physical nature of defects. Though it cannot be used for obtaining the time dynamics of defects, hence for lifetime estimation for NBTI.
- Measure-Stress-Measure (MSM) experiment identifies the effect of defect generation on parameters like effective transistor mobility (μ_{eff}), sub-threshold

slope (SS), *etc.*, which can later be used in indirect estimation of parametric degradation in UFV and OTF- $I_{D,lin}$ techniques. However, similar to SDR, MSM cannot be used for lifetime estimation of NBTI.

- UFV and OTF- $I_{D,lin}$ are ideally suitable for lifetime prediction during NBTI stress, provided appropriate corrections are made in the ΔV_T estimation. Later, MSM experiment can be used to estimate the time dynamics of changes in μ_{eff} , SS , *etc.*

Finally, one needs to separate out the components of interface and hole trapping to obtain the correct dynamics of these components, which can be used to compare with the theoretical models for these two components.

6.9. Summary

In this chapter, we have discussed different time-critical defect characterization techniques that are recently being used in literature. We illustrated the measurement of different transistor parameters, which are performed directly or indirectly using these techniques. As such, we identified the limitations of indirect parameter estimation in on-the-fly (OTF)- $I_{D,lin}$ and Ultra-fast V_G (UFV) measurement techniques. Moreover, since it is impossible (at present) to directly measure the time dynamics of interface and hole trapping components, we proposed a phenomenological technique for separating out these components, which has been verified using a wide set of experiments.

7. DEFECT FORMATION AND OFF-STATE POWER DISSIPATION (GATE LEAKAGE)

7.1. Introduction

In chapters 3-5, we have discussed models that can predict the rate of defect generation within a transistor at given stress voltage/electric field and temperature. Though some part of the models still requires some work (see discussions in section 3.6.5), we can still use the predictability of the classical H-H₂ R-D model (in terms of electric field dependence, temperature dependence, and frequency independence) for resolving a very important problems in recent CMOS technology. In section 7.2, we present the importance of studying the optimization between defect generation and static-power dissipation (gate leakage) in current-day CMOS transistors, which uses plasma oxynitride materials as gate dielectric (similar to type-I transistors discussed earlier). To achieve this objective, we

- measure gate leakage (current density, J_G) and NBTI (using OTF $I_{D,lin}$) over broad range of $\%N$ in type-I transistors (section 7.3),
- model gate leakage as a function of EOT and $\%N$ by calibrating measurements with detailed numerical simulation [178] and physically based analytical expression (section 7.4),
- model NBTI degradation in these type-I transistors, which has dominance of N_{IT} generation over hole trapping, within a theoretically consistent framework of R-D model [190, 191], by taking into account the voltage, temperature and frequency dependencies of NBTI (section 7.5),
- establish a compact analytical approach for predicting transverse electric field in MOS structures, useful for analyzing NBTI degradation (section 7.5.1), and

- finally, construct a design diagram for co-optimization of J_G and NBTI in type-I transistor technology for arbitrary %N and EOT combinations (section 7.6).

Based on this study [128, 184], we conclude that although there is no optimum %N for plasma oxynitride CMOS technology that can simultaneously reduce both gate leakage and NBTI, the reduction in J_G at NBTI-limited %N can be significant and would reduce power dissipation without affecting NBTI-margin.

7.2. Motivation

Over the last few decades, aggressive scaling in transistor dimensions has resulted in dramatic improvement in IC performance [35]. Before the introduction of 130 nm technology node, SiO₂ (with little or no nitrogen) was used as a gate dielectric and such pure SiO₂ offered excellent interfacial properties with Si-substrate [35, 38, 39]. However, as the dielectric thickness were scaled below 2nm, higher static power dissipation (due to increase in gate leakage [35, 38, 39, 303]) and increased flat-band voltage (due to boron penetration from p+ poly-Si to PMOS substrate [38, 39, 303]) necessitated the use of dielectrics other than pure SiO₂. Different high- κ dielectrics were proposed to replace SiO₂ [39]. Of these, oxynitride dielectrics appeared to be the near-term alternative (which some industries intend to use in transistors down to 28nm CMOS technology node), offering convenient processing and better reliability compared to other high- κ counterparts [38, 39, 303].

However, the advantages of oxynitride gate dielectric (namely, lower gate leakage and resistance to boron penetration) are counterbalanced by its poorer reliability characteristics, especially NBTI, which is observed to get enhanced by the inclusion of nitrogen within gate dielectrics [6, 34, 37, 129, 191, 303-305]. In general, numerous studies have established the importance of the nitrogen spatial profile in dictating the boron penetration and NBTI characteristics [303, 304, 306, 307]. Such studies show that reduction of both NBTI and boron penetration requires a nitrogen profile having (comparatively) lower nitrogen near substrate-dielectric interface (for optimum NBTI)

and higher nitrogen near gate-dielectric interface (for reduced boron penetration). Plasma nitridation with decreasing nitrogen concentration from poly-Si/dielectric interface to substrate/dielectric interface, therefore, became an ideal choice [238, 304, 306, 307] for oxynitride gate dielectrics. Adjustment of nitrogen profile within the dielectric of devices having plasma nitrided oxide (PNO) ensures improved device performance in terms of boron penetration and NBTI, even down to an EOT (effective oxide thickness) of 1.1nm [306, 307].

Thus among the three design considerations (gate leakage, boron penetration, and NBTI) for PNO devices, several studies reported the gate leakage vs. boron penetration [38, 39, 303] and boron penetration vs. NBTI [303, 304, 306, 307] issues. However, the remaining, and perhaps equally important optimization of gate leakage and NBTI in current CMOS technology (the topic of this chapter) has never been considered. Indeed, reduction of gate leakage under certain limit necessitates a minimal nitrogen concentration [38, 308-311], but whether such nitrogen concentration keeps NBTI within a desired limit (so that device lifetime meets the required criteria) has never been investigated. In the subsequent sections of this chapter, we perform a quantitative analysis of leakage/NBTI trade-off (as a function of nitrogen concentration) to address this question: whether or not co-optimization of gate leakage/NBTI is possible at any nitrogen concentration.²³

7.3. Experimental Detail

Sample preparation for devices used in this study are discussed in detail in [307, 308] and will not be repeated here. Physical thickness (T_{PHY}) and %N for the PNO samples are determined using X-ray Photoelectron Spectroscopy (XPS) [239] with repeatability better than 3% and 2% for T_{PHY} and %N respectively [308]. For each sample, we measure the gate capacitance (C_G) and J_G at different gate voltage (V_G). This is followed by NBTI

²³ We have ignored boron penetration in our study, which is shown to be negligible for $EOT \geq 1.1\text{nm}$ [306, 307].

degradation measurement at various stress condition (voltage and temperature) using OTF- $I_{D,lin}$ technique (section 6.6), with t_0 of ~ 1 ms, where ΔV_T is estimated using (6.14). Thus, estimated ΔV_T ignores the necessary mobility correction (see section 6.7.4). So, the reported lifetime (t_{life}), field acceleration and safe operating condition (see section 7.5) should be interpreted as a qualitative illustration for NBTI-leakage optimization.

7.4. Gate Leakage in Plasma Oxynitride Transistors

7.4.1. Estimation of device model parameters

Leakage analysis involving MOS structures first necessitates the determination of approximately nine model parameters that characterizes the gate dielectric. These parameters and the measurements (C_G-V_G , J_G-V_G) necessary for extracting them are summarized in Table 7.1. We use quantum mechanical (QM) analysis (multi-subband electron/hole quantization, wave-function penetration within the dielectric, and poly-depletion features are used in this study) [178] of experimental C_G-V_G and J_G-V_G characteristics to extract these parameters. Excellent agreement between experiment and QM simulation similar to that shown in Figure 7.1 is obtained for each sample used in this work.

Table 7.1: Model parameters used for characterizing MOS structures.

| Symbol | Definition | Extracted From |
|-----------------|---|----------------|
| EOT | Effective Oxide Thickness | C_G-V_G |
| V_{FB} | Flat-band voltage | C_G-V_G |
| ϵ_{di} | Relative dielectric constant for the dielectric ($= \epsilon_{SiO_2} T_{PHY} / EOT$; where ϵ_{SiO_2} is the relative dielectric constant for SiO_2) | C_G-V_G |
| N_{sub} | Substrate doping (assumed uniform) | C_G-V_G |
| N_{poly} | Poly-gate doping (assumed uniform) | C_G-V_G |
| m_{di} | Carrier effective mass within dielectric | J_G-V_G |
| $E_{g,di}$ | Dielectric band-gap | J_G-V_G |
| ϕ_{be} | Electron barrier height at the substrate-dielectric interface | J_G-V_G |
| ϕ_{bh} | Hole barrier height at the substrate-dielectric interface | J_G-V_G |

7.4.2. Variation of Parameters with %N

The procedure of parameter extraction discussed in the previous section can now be repeated to extract the nine model parameters for oxides with various %N. Devices considered in this study have T_{PHY} of 1.6-1.7 nm, EOT of ~ 1.3 nm and N_{sub} of $\sim 3 \times 10^{17}$ cm^{-3} . Nitrogen dose for PNO samples is varied from $\sim 1.25 \times 10^{15}$ to $\sim 2.5 \times 10^{15}$ cm^{-2} . For all the NMOS samples, N_{poly} is $\sim 8.5 \times 10^{19}$ cm^{-3} and V_{FB} is ~ -1.0 V. For PMOS PNO samples, N_{poly} is $\sim (5.5 \pm 0.5) \times 10^{19}$ cm^{-3} and V_{FB} is $\sim 0.9 \pm 0.1$ V; both N_{poly} , V_{FB} are found to decrease with increasing %N (consistent with the assessment in [312]). Finally, due to boron penetration effect in pure SiO_2 samples, higher V_{FB} (~ 1.21 V) is estimated. Our analysis shows that the boron penetration effect also reduces N_{poly} for pure SiO_2 samples to $\sim 10^{19}$ cm^{-3} near poly-dielectric interface from its bulk value of $\sim 10^{20}$ cm^{-3} . As we are

neglecting the boron penetration effect for simplicity, we can approximate $|V_{FB}|$ for the samples as –

$$V_{FB} = \pm \frac{k_B T}{q} \ln \frac{N_{sub} N_{poly}}{n_i^2} + V_{FB(cor)} + \Delta V_{FB} \quad (7.1)$$

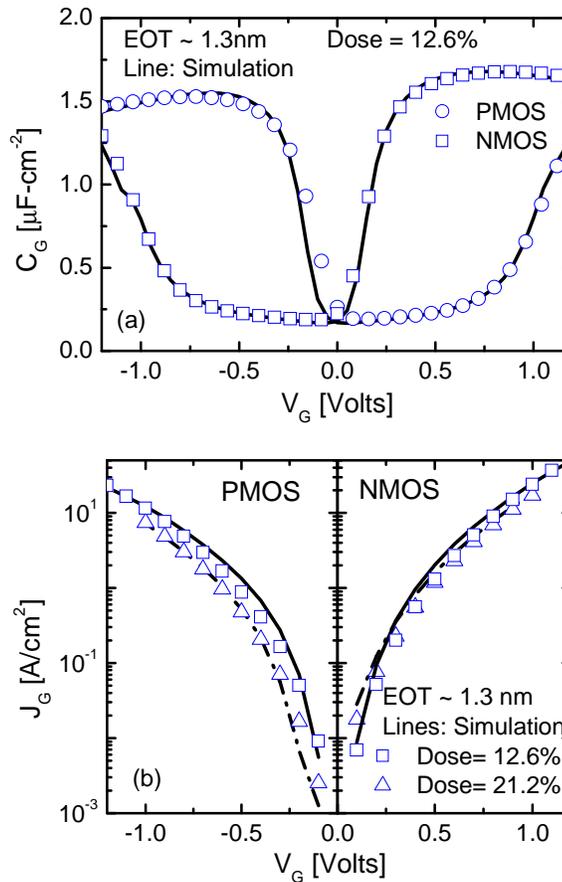


Figure 7.1: Experimental (a) C_G - V_G and (b) J_G - V_G curves (in inversion region) for a few SiON devices fitted using QM simulation [178].

The first term in equation (7.1) represents the standard semi-classical approximation of flat-band voltage for poly-Si gate NMOS/PMOS (-/+) structures (where, q and n_i are

electron charge and intrinsic carrier concentration respectively), previously used in [313]. Second term ($V_{FB(cor)} = -0.04V$ and $0.05V$ for NMOS and PMOS devices respectively) indicates the correction needed to match with V_{FB} obtained from QM simulation [178]. And the last term ΔV_{FB} takes into account the contribution from interfacial charges at the poly-Si/ oxynitride dielectric interface [312]. For NMOS samples ΔV_{FB} is $\sim 0.057V$; whereas for PMOS samples, ΔV_{FB} is negative and $|\Delta V_{FB}|$ increases with $\%N$ (see Figure 7.2).

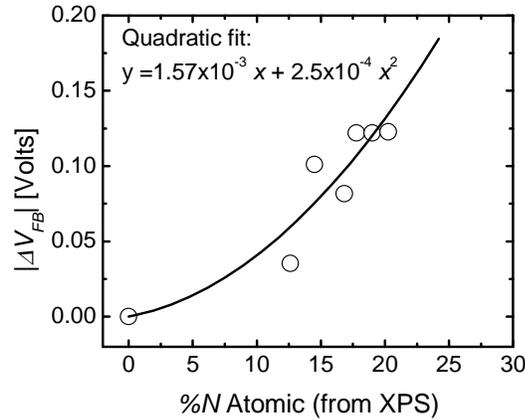


Figure 7.2: Variation in $|\Delta V_{FB}|$ for PMOS SiON samples, neglecting the boron penetration effect.

The variation of remaining device parameters ϵ_{di} , m_{di} , $E_{g,di}$, ϕ_{be} are shown in Figure 7.3. Here, the parameters for Si_3N_4 (57.1% N_2 dose with $\phi_{be} = 2.1$ eV, $E_{g,di} = 5.1$ eV, $\epsilon_{di} = 7.6$, $m_{di} = 0.23\sim 0.28$) are obtained from [38, 309-311]. The error bars indicate the variation in model parameters expected due to $\pm 0.5\text{\AA}$ error in T_{PHY} measurement by XPS. These parameters show an approximately quadratic dependence on $\%N$ (ϵ_{di} , m_{di} , $E_{g,di}$, $\phi_{be} \propto a + b(\%N) + c(\%N)^2$); where the constants a , b , c for each parameter are shown in Figure 7.3). This quadratic trend contradicts with the linear variation typically used in the literature [38, 308-311] and requires some discussion. The origin of the “linear

approximation” can be traced to the work of Brown *et al.* [314], where they measured ϵ_{di} for different oxynitride devices (along with devices having SiO₂ and Si₃N₄ dielectric) and later estimated %N by *assuming* linear variation of ϵ_{di} from SiO₂ to Si₃N₄. Referring (directly or indirectly) to this work by Brown *et. al.*, the authors in [38, 308-311] used linear dependency of ϵ_{di} and other parameters without any additional proof.

For a physical interpretation of the variation of ϵ_{di} , $E_{g,di}$, ϕ_{be} with %N, we consider the oxynitride as a uniform pseudo-binary alloy, i.e. SiO_aN_b = (Si₃N₄)_x(SiO₂)_{1-x} [315, 316] that satisfies the stoichiometry (i.e. $2a+3b = 4$) requirement. Thus, percentage atomic concentration (uniform) of [O] and [N] can be expressed as –

$$\%N = \frac{4x}{3+4x} \times 100; \quad \%O = \frac{2-3x}{3+4x} \times 100 = \frac{400-7(\%N)}{6}. \quad (7.2)$$

Using equation (7.2) and following similar steps as reported in [316], we obtain –

$$\epsilon_{di} = \epsilon_{SiO_2} \frac{4-7n_a}{4+2n_a} + \epsilon_{Si_3N_4} \frac{9n_a}{4+2n_a}. \quad (7.3)$$

where $n_a = \%N / 100$ is the atomic fraction of [N] within the oxynitride. Dashed line in Figure 7.3a plots of ϵ_{di} vs. uniform %N based on equation (7.3). Observed non-linearity of this plot contradicts the linear variation used in [38, 308-311] even for uniformly dosed nitrided dielectric. Now, the plasma nitrided samples studied in this article have [N] peak near poly-gate [239, 307, 308]. The XPS measurement of %N for these samples, therefore, over-estimates %N in the film [239]. Such over-estimation explains the observed discrepancy between the dashed line in Figure 7.3a (based on equation (7.3) for uniformly dosed sample) and our experimental values (fitted solid line in Figure 7.3a for ϵ_{di} vs. %N). Similar over-estimation of %N, by assuming %N equals peak [N] value, also gave identical shape for ϵ_{di} vs. %N in [317]. Moreover, as $E_{g,di}$ (hence, ϕ_{be}) $\propto 1/\epsilon_{di}$, quadratic relation for ϵ_{di} vs. %N (from Figure 7.3a) also results similar relation for $E_{g,di}$ (hence, ϕ_{be}) vs. %N (in Figure 7.3b). Thus, consideration of oxynitride as a pseudo-binary alloy, along with over-estimation of %N for PNO samples using XPS setup, explains the observed “quadratic dependence” of parameters in Figure 7.3. If over-estimation of %N

were a bit smaller (i.e. for a more uniform profile compared to PNO), one could also have observed an approximately “linear dependence” of parameters with $\%N$. Hence, we identify variation (linear or quadratic) of ϵ_{di} , m_{di} , $E_{g,di}$, ϕ_{be} to be sample-specific, which also depends on the measurement setup being used.

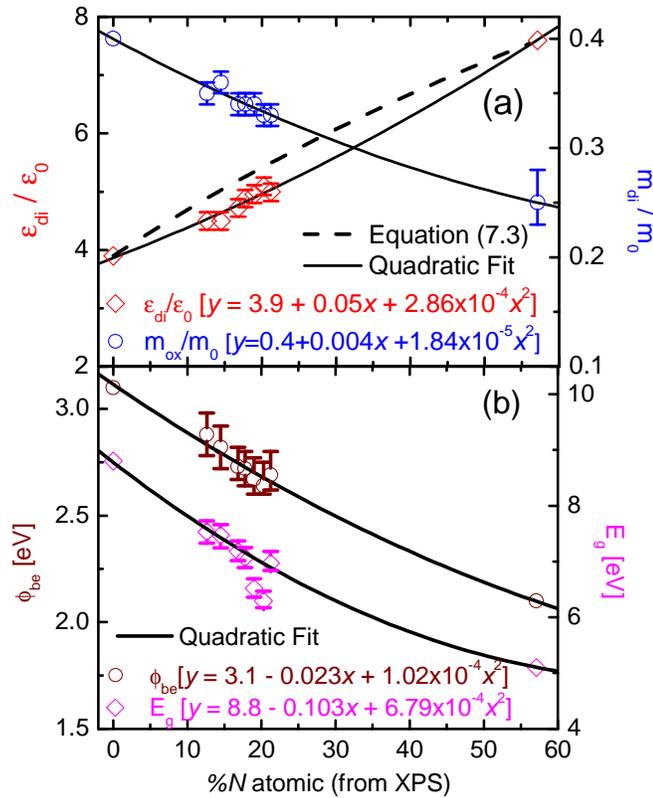


Figure 7.3: Quadratic variation of (a) ϵ_{di} , m_{di} , (b) $E_{g,di}$, ϕ_{be} with $\%N$ (0% for SiO_2 , 57.1% for Si_3N_4) for samples having $EOT \sim 1.3\text{nm}$. Here, ϵ_0 is free space dielectric constant and m_0 is mass of electron. The error bar shows the variation in model parameters expected due to $\pm 0.5\text{\AA}$ error in measurement of T_{PHY} by XPS. Dashed line in (a) plots the variation of ϵ_{di} vs. $\%N$, based on equation (7.3).

7.4.3. Gate Leakage Variation

Using the explicit dependencies of device parameters on %N (from previous section), we can predict the gate leakage variation (at a particular V_G) both for NMOS and PMOS devices at different EOT . Assuming that change in EOT does not affect characteristics of the PNO film (*i.e.*, PNO with any N_2 dose has similar properties irrespective of EOT)²⁴, we have done QM simulation [178] to obtain gate leakage variation (*e.g.*, Figure 7.4 shows J_G at $V_{DD,ITRS}$ [35] for different EOT) as a function of %N and EOT for NMOS transistors. We also observe that for similar absolute gate voltage ($|V_G|$), PMOS gate leakage is lower than that of NMOS for N_2 dose less than ~20-25%, consistent with the findings in [308]. The error bar in Figure 7.4 @ $EOT = 1.2$ nm indicates J_G variation expected due to ± 0.5 Å error in T_{PHY} , which is observed to be negligible up to 30% N_2 dose. Furthermore, Figure 7.4 does not indicate a minima in gate leakage as %N is varied, in contradiction to the predictions in [308, 309, 311]. As stated in [309] and also verified by us, existence of minima is a direct consequence of the value of the device parameters (ϵ_{di} , m_{di} , $E_{g,di}$, ϕ_{be}) used for Si_3N_4 , hence should not be considered as a property for oxynitride samples.

In sum, the gate leakage study performed in this section indicates a well-known reduction in leakage with increase in %N for devices having similar EOT , with no existence of minima in the variation. To make this gate leakage study more realistic, one should refine it using appropriate nitrogen profile information (like [304]), thus taking into account the EOT dependence of parametric variation, presented in Figure 7.4.

²⁴ This is a very crude assumption and reflects a lack of knowledge of detailed %N profile within the film. With more information regarding nitrogen profile for all PNO samples (like [304]) becomes available, our results can be refined. Therefore, we believe that due to such assumptions involved in this analysis, our approach should be used as a guideline; not an absolute reference for design optimization.

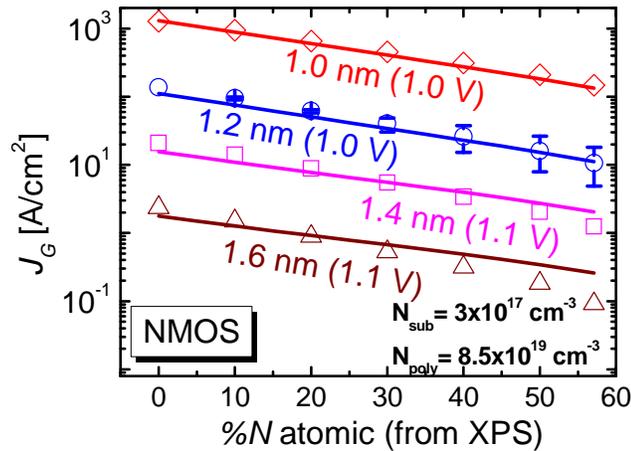


Figure 7.4: $J_G @ V_G$ vs. $\%N$ for NMOS devices at different EOT (Symbols: QM simulation, lines: analytical approach). The error bar @ $EOT = 1.2$ nm indicates J_G variation expected due to ± 0.5 Å error in T_{PHY} , which is negligible up to 30% of N_2 dose.

7.5. NBTI Degradation

We now consider the second aspect of our optimization problem, namely the NBTI degradation, which has become an important reliability concern for technology nodes using EOT below 2 nm [6, 30, 31, 34, 190, 192, 318]. The origin of NBTI degradation in type-I transistors (*i.e.*, N_{IT} formation) is discussed in sections 3.5.4 and 3.7, which has the following main features:

- (1) NBTI degradation results mainly from depassivation of Si-H bonds at the Si/dielectric interface (leading N_{IT} generation) and resultant diffusion of hydrogen species into gate dielectric and poly-Si. Reaction-Diffusion (R-D) model is used to interpret such NBTI degradation in type-I transistors [30, 31, 34, 129, 130, 142, 143, 190, 191].
- (2) At long stress time ($t > 10$ sec), interface trap generation (ΔN_{IT}) is governed by the classical (Arrhenius-activated) diffusion of molecular hydrogen (H_2) [31, 34, 129, 130, 142, 143, 190, 191] and can be estimated using (3.2)-(3.6).

- (3) NBTI degradation is independent of frequency for AC NBTI degradation, which is shown both theoretically [126, 167] and experimentally for wide frequency range [289, 319].

In addition, there are type-II transistors, where hole trapping (detrapping) to (from) pre-existing defects also plays significant role in NBTI degradation [129, 141, 155] (see section 6.7.5 and chapter 5 for details). These transistors, mostly having thick dielectric with high N_2 near interface [129], also show frequency dependence in measured AC NBTI degradation [141, 155]. As we focus our analysis on PNO transistors having less N_2 near interface, we ignore such hole trapping/detrapping in our analysis. In principle, the presented optimization scheme can be more generalized for any oxynitride transistors by treating hole trapping/ detrapping using appropriate models (*e.g.*, [129, 141, 155, 190] and section 5.5).

7.5.1. Calculation of Electric Fields

As evident from equations (3.2)-(3.6), estimation of electric fields (E_{ox} and E_c) are critical in NBTI degradation analysis for various EOT and $\%N$. Experimentally, one could integrate measured C_G - V_G from V_{FB} to V_G to determine total charge within the substrate (Q_{sub}) and hence, calculate E_{ox} (@ V_G) = Q_{sub}/ϵ_{SiO_2} , which equals E_c for NMOS in accumulation; while for PMOS in inversion, E_c could be calculated using $C_G(V_G - V_T)/\epsilon_{SiO_2}$. However, this direct method is difficult to apply in oxides with EOT less than 2.0 nm that have high gate leakage, where one must either use specialized test structures for measuring C_G - V_G [320] or use a reconstruction algorithm to correct the leakage-contaminated C_G - V_G [321] data. Moreover, high gate leakage for oxynitrides with EOT lower than ~ 1.2 nm loads the C-V analyzer and makes the C_G measurement more difficult.

For such thin oxides, therefore, an analytical approach is better suited in predicting E_{ox} and E_c for a particular technology node using known values of EOT , N_{sub} , N_{poly} , and $\%N$. A flowchart for such calculation (for inverted PMOS substrate with negative V_G) is shown in Figure 7.5, where valence-band tip at the substrate/dielectric interface is used as energy reference and charge centroids are determined from the same interface. Our

approach is similar to the algorithm proposed in [280]. The notations used in Figure 7.5 are given as follows (QM notations are discussed extensively in [280]): ϵ_{Si} is the dielectric constant for Si; $\phi_{F(bulk)}$ is the Fermi-level position in the bulk of the substrate (with respect to the midgap); ϕ_S is the surface potential; ϕ_{dep} is the depletion charge contribution to ϕ_S ; $|(E_F - E_V)_{bulk}|$ is the difference between the valence band and the Fermi level in the substrate bulk; V_{poly} is the voltage drop in poly-Si; N_D^+ , N_s , N_{inv} , and N_{dep} are the concentrations of the ionized donor, substrate charge, inversion charge, and depletion charge, respectively; E_{FS} is the substrate Fermi level; E_{ij} is the eigen-energy for the j -th subband in the i -th valley within the potential well formed near the substrate/dielectric interface (for PMOS device in inversion, only the subbands of the heavy hole valleys are considered); N_{ij} and z_{ij} are the charge concentration and charge centroid in eigen-state E_{ij} ; $m_{dos,i}$ is the density-of-state effective mass for the i -th valley; and z_{avg} is the average inversion layer thickness.

The results using such fast algorithm is in excellent agreement with more detailed QM C-V simulation [178] (Figure 7.6). Both analytical and QM simulation also predicts negligible change in E_{ox} under change in temperature at a particular V_G . Therefore, electric field calculated at a particular temperature can be used for reliability predictions at any other temperature.

7.5.2. NBTI Model Parameters & Their %N Dependency

Determination of device information and electric field enables us to use equations (3.2)-(3.6) to fit ΔV_T measured at various $V_{G,STS}$, T_{STS} , thus estimate NBTI model parameters (A_{IT} , γ , γ_f , a , E_{DI} , n) for type-I transistors in Table 3.4. As discussed in sections 3.8.1 and 3.8.2, we can use $n \sim 0.14$, $a = 0.8 q\text{\AA}$, and $E_{DI} \sim 0.9\text{eV}$ for these transistors and thus express the %N dependency in terms of variations in A_{IT} and γ_f (Figure 3.22).

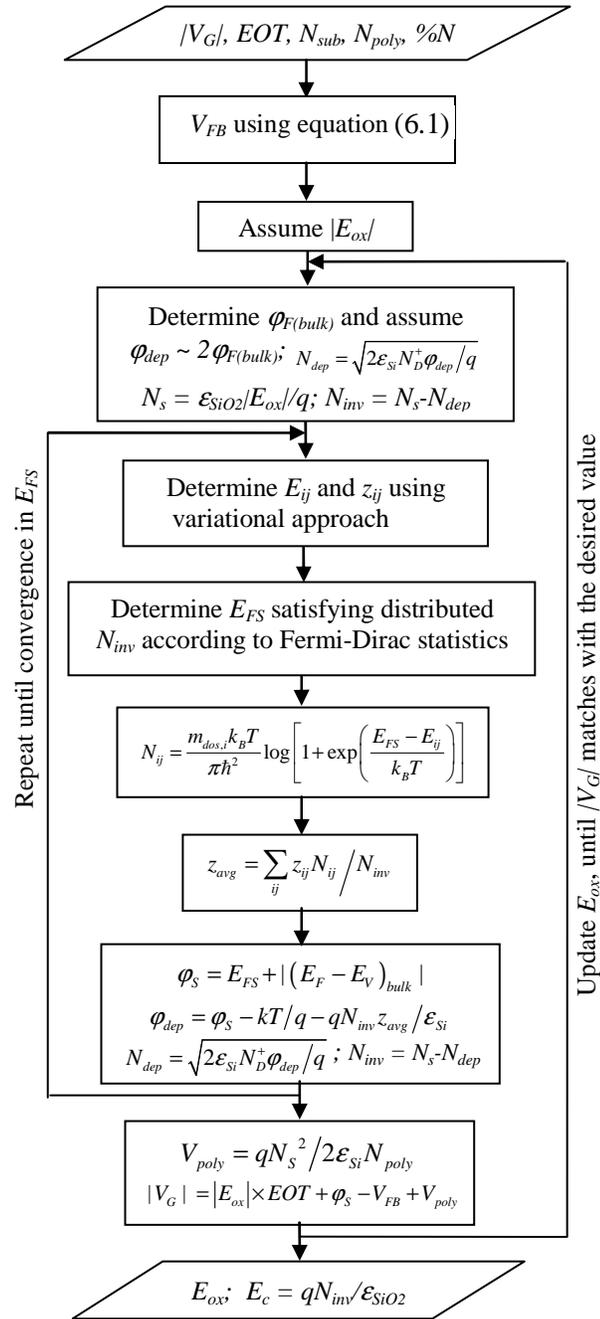


Figure 7.5: Flowchart for electric fields calculation (from $|V_G|$) in PMOS inversion. Reverse calculation (electric field to $|V_G|$) is also possible. Similar calculation can also be done for NMOS devices. The notations used here are described in the text.

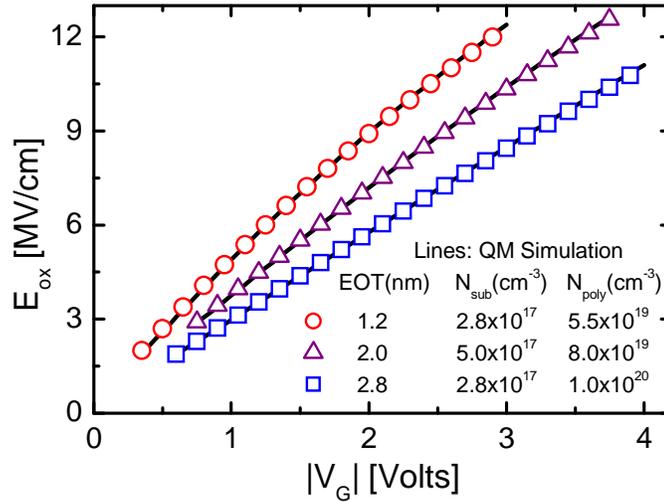


Figure 7.6: Electric field vs. gate voltage relation obtained using analytical approach of Figure 7.5 (symbols) are in agreement with that obtained from QM simulation (lines).

7.5.3. Safe Operating Condition

The variation of NBTI performance for different oxynitride devices is monitored by V_{safe} (see section A.4 for definition). The parameters $A_{IT}(N)$, $\gamma_T(N)$ (derived by fitting the NBTI data, Figure 3.22), $V_{FB}(N)$ (obtained from C-V analysis, Figure 7.2) and other (approximately constant) PMOS parameters ($N_{poly} \sim 6 \times 10^{19} \text{ cm}^{-3}$, $N_{sub} \sim 3 \times 10^{17} \text{ cm}^{-3}$, $a \sim 0.8 \text{ q}\text{\AA}$, $E_{DI} \sim 0.9 \text{ eV}$) are used to calculate $V_{safe}(N, EOT)$ for any combinations of %N and EOT (Figure 7.7)²⁵ by means of equations (3.2)-(3.6). Here, we use the analytical

²⁵ Although Figure 7.7 suggests similar V_{safe} for different EOT at N_2 dose $\sim 20\%$, it should not be taken as absolute number. As PNO samples have nitrogen profile with [N] peaking near poly [238, 239], devices having same %N (measured by XPS [239]), but higher EOT , will comparatively have lower N_2 near substrate interface – hence less NBTI degradation [129, 190]. Thus, V_{safe} for higher EOT devices is expected to be higher than the values shown in Figure 7.7. Such correction again requires one to have accurate

approach presented in Figure 7.5 for calculating electric fields. NBTI performance for DC operation is calculated for failure criteria of $\Delta V_{T(max)} = 60$ mV, $t_{life} = 5$ years (equivalent to $t_{life} = 10$ years in CMOS operation, assuming 50% activity of PMOS transistors). Use of NBTI aware circuit design [322] will also help in further reduction in required lifetime. A lifetime improvement of $2^{1/2n}$ times (based on R-D model simulation for AC degradation having 50% duty cycle) is used for illustrating AC effects in NBTI degradation, although there are reports for $2^{1/n}$ times improvement (corresponding to AC NBTI degradation being $\sim 50\%$ of DC degradation [289, 319]). Thus we predict an improvement in $V_{safe}(N, EOT)$ with inclusion of AC effects and also with reduction in temperature.

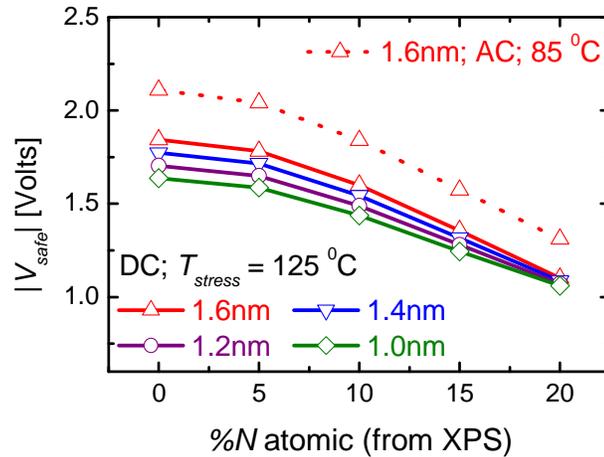


Figure 7.7: V_{safe} vs. %N at different EOT ($T = 125$ °C; Failure criteria: $\Delta V_T = 60$ mV, $t_{life} = 5$ years for CMOS DC operation). Dotted line for $EOT = 1.6$ nm indicates the improvement in AC condition and with $T_{stress} = 85$ °C.

knowledge of N_2 profile within the PNO film. Once such information is available, the optimization process can be readily updated.

7.6. Optimization of gate leakage and NBTI

So far we have discussed the variation of gate leakage (obtained using QM simulation, see Figure 7.4) and NBTI degradation (resultant change in V_{safe} , Figure 7.7; which uses results from Figure 7.2, Figure 7.5 and Figure 3.22) for plasma oxynitride devices with $\%N$ and EOT . Merging the plots of $J_G(N, EOT)$ at different V_G and $V_{safe}(N, EOT)$, we obtain a design diagram for plasma oxynitride technology (Figure 7.8), which can be used for $V_{safe} = V_G = 0.8\sim 1.2V$. Note that the curves obtained at different V_G fall on top of each other (solid black lines in Figure 7.8). This makes Figure 7.8 to be applicable for optimization in the range of $V_{safe} = 0.8\sim 1.2V$. The range of V_{safe} can be easily extended by considering $J_G(N, EOT)$ at more V_G in the design diagram (hence, adding more y-axis with appropriate current scales). Moreover, change in operating temperature and inclusion of AC effects (thus changing V_{safe}) can shift the diagram along x-axis (as gate leakage has negligible change with temperature). Resultant design diagram can then be used to calculate any pair of the variables (V_{safe} , J_G , $\%N$, EOT) when the other pair is given. For example, if an IC design requires $V_{DD} = 1.1V$ and $J_G = 10A/cm^2$, one needs to draw a vertical line from x-axis @ 1.1V and a horizontal line from y-axis @ 1.1V, $10A/cm^2$. The two lines intersect at point A (see Figure 7.8). Thus design diagram suggests the use of $EOT \sim 1.4nm$ and $\%N \cong 20\%$.

If the optimization anticipated by the conservative design diagram is unacceptable, the optimization at higher $\%N$ and lower EOT will be possible by reconstructing a new diagram with reduced operating temperature, including AC effects and increasing the leakage constraint. Finally, since PMOS leakage exceeds NMOS leakage for N_2 dose exceeding 20-25% (see section 7.4.3), one should consider PMOS leakage/PMOS NBTI rather than the NMOS leakage/PMOS NBTI (discussed in this chapter) for optimization above 20-25% N_2 dose.

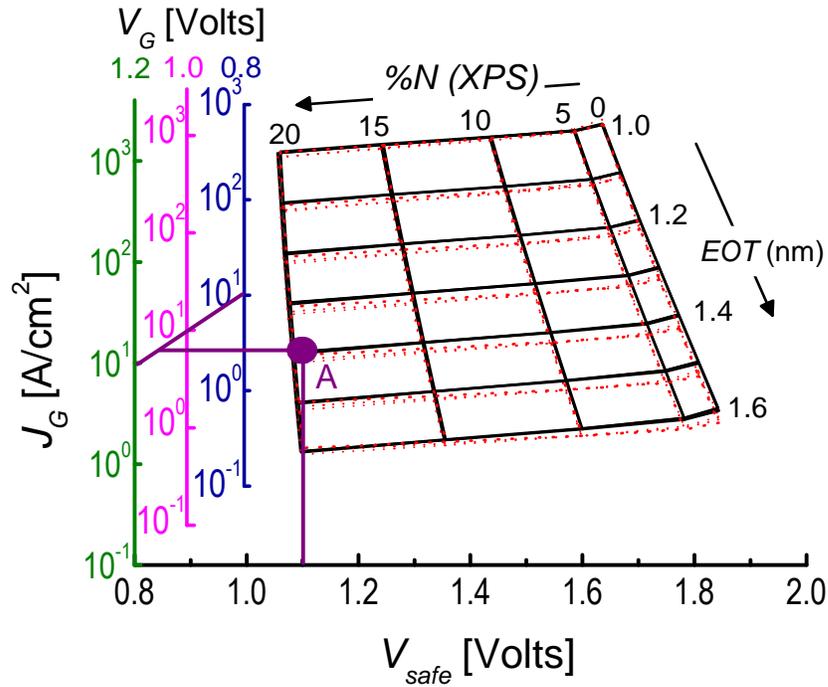


Figure 7.8: Design diagram for SiON devices, obtained using Figure 7.4 (J_G @ $V_G = 0.8, 1.0, 1.2\text{V}$ are considered) and Figure 7.7. Solid black lines use J_G calculated by QM simulation and dotted red lines use J_G calculated from equations (7.4)-(7.8). Note the universality (*i.e.*, curves at different V_G fall on top of each other) of the figure when J_G calculated by QM simulation. Operating point A ($V_{safe} = 1.1\text{V}$ and $J_G = 10\text{A/cm}^2$) suggests, for example, the use of $EOT \sim 1.4\text{nm}$ and $\sim 20\%$ N_2 dose.

7.7. Analytical Construction of Design Diagram

In the last few sections, we have provided a physically rigorous procedure to optimize gate current and NBTI degradation based on a set of comprehensive measurements. However, the simulation tools we used may not be readily available to many research groups. Therefore, we develop analytical expressions for calculating electric fields, leakage current and safe operating voltage (which is an indicator of NBTI degradation), which would enable one to rapidly construct the design diagram in Figure 7.8 with reasonably good accuracy.

7.7.1. Model for electric fields

Noting the universal quadratic relationship between E_{ox} and $|V_G|$ in Figure 7.6 irrespective of the values for EOT , N_{sub} , N_{poly} , $\%N$ (for larger EOT samples quadratic behavior appears at larger $|V_G|$ compared to smaller EOT samples), we can estimate E_{ox} using-

$$E_{ox} = P_1 |V_{G,eff}|^2 + P_2 |V_{G,eff}| + P_3, \quad (7.4)$$

where $|V_{G,eff}| = |V_G| - |\Delta V_{FB}|$ and the coefficients are given by equations (7.5) and (7.6) for NMOS and PMOS devices respectively, *i.e.*,

$$\begin{aligned} P_1 &= (-0.97 + 0.7EOT - 0.15EOT^2) \begin{pmatrix} 1 + 0.014N_{poly} - 3.22 \times 10^{-3} N_{poly}^2 \\ 1 - 1.77 \times 10^{-3} N_{sub} - 5.08 \times 10^{-4} N_{sub}^2 \end{pmatrix}, \\ P_2 &= (8.56 - 3.66EOT + 0.62EOT^2) \begin{pmatrix} 0.67 + 0.084N_{poly} - 3.78 \times 10^{-3} N_{poly}^2 \\ 1 - 6.73 \times 10^{-4} N_{sub} - 1.51 \times 10^{-4} N_{sub}^2 \end{pmatrix}, \\ P_3 &= (0.93 - 0.87EOT + 0.24EOT^2) \begin{pmatrix} 1.98 - 0.27N_{poly} + 0.016N_{poly}^2 \\ 0.94 + 0.01N_{sub} - 2.7 \times 10^{-3} N_{sub}^2 \end{pmatrix}, \end{aligned} \quad (7.5)$$

$$\begin{aligned} P_1 &= (-0.93 + 0.68EOT - 0.15EOT^2) \begin{pmatrix} 1.06 + 0.013N_{poly} - 2.7 \times 10^{-3} N_{poly}^2 \\ 1 + 2.6 \times 10^{-3} N_{sub} - 4.43 \times 10^{-4} N_{sub}^2 \end{pmatrix}, \\ P_2 &= (8.4 - 3.65EOT + 0.644EOT^2) \begin{pmatrix} 0.69 + 0.08N_{poly} - 3.64 \times 10^{-3} N_{poly}^2 \\ 1 - 6.5 \times 10^{-4} N_{sub} - 1.47 \times 10^{-4} N_{sub}^2 \end{pmatrix}, \\ P_3 &= (0.78 - 0.73EOT + 0.2EOT^2) \begin{pmatrix} 2 - 0.29N_{poly} + 0.018N_{poly}^2 \\ 0.9 + 0.027N_{sub} - 2.33 \times 10^{-3} N_{sub}^2 \end{pmatrix}, \end{aligned} \quad (7.6)$$

EOT , N_{sub} and N_{poly} in equations (7.5) and (7.6) need to be expressed in units of nm, 10^{17} cm^{-3} , and 10^{19} cm^{-3} respectively. After calculating E_{ox} using equation (7.4),

$E_c = E_{ox} - E_{dep} = E_{ox} - \sqrt{2q\epsilon_{Si}N_D^+\phi_{dep}} / \epsilon_{SiO_2}$ can be calculated by using, $\phi_{dep} \sim 2\phi_{F(bulk)} + 6kT/q$ [323]. Such estimations of E_{ox} and E_c are in excellent agreement with the results from both QM simulation and analytical formalism of section 7.5.1.

7.7.2. Analytical Model for Gate Leakage Current

Next, we develop a physically based analytical approach to estimate NMOS gate leakage variation in inversion by using $J_G \sim N_{inv} f \bar{T}$ [183]. We use $N_{inv} \sim E_{ox}$ (considering $N_{dep} \ll N_{inv}$), electron impact frequency (f) proportional to $E_{ox}^{0.6}$ [183], and estimate mean tunneling probability $\bar{T} \sim \exp(-2\alpha T_{PHY})$ by replacing the triangular energy barrier for electron tunneling using an effective square barrier of height $\phi_{be} - E_{ox} T_{PHY}/2$ [308]; hence, $\alpha \equiv \sqrt{2m_{ox}q(\phi_{be} - E_{ox} T_{PHY}/2)}/\hbar$. Based on these, we fit J_G vs. E_{ox} (calculated by QM simulation) using an expression,

$$J_G = A E_{ox}^{1.6} \exp\left(-B\sqrt{1 - C E_{ox}}\right). \quad (7.7)$$

Fitting of J_G vs. E_{ox} at different T_{PHY} , ϕ_{be} and m_{ox} enables us to determine the coefficients A , B , C such that-

$$\begin{aligned} A &= 2 \times 10^9, \\ B &= (2.73 + 17.8 T_{PHY} - 2.13 T_{PHY}^2)(0.72 + 0.7 m_{ox})(0.6 + 0.13 \phi_{be}), \\ C &= (-3.56 \times 10^{-3} + 3.71 \times 10^{-2} T_{PHY} - 7.1 \times 10^{-3} T_{PHY}^2)(1.82 - 0.266 \phi_{be}). \end{aligned} \quad (7.8)$$

Thus, we observe an increase of B with increase in T_{PHY} , ϕ_{be} , m_{ox} , and an increase of C with increase in T_{PHY} and/or decrease in ϕ_{be} ; which is consistent with the semi-classical intuition presented above. Using the quadratic dependency of ϵ_{di} , ϕ_{be} , m_{ox} from Figure 7.3 and E_{ox} calculated using equation (7.4), we obtain leakage current variation with %N using equation (7.8). The obtained variation is remarkably consistent with QM simulation results (Figure 7.4).

7.7.3. Calculation of V_{safe}

Next, we calculate V_{safe} by first determining safe operating electric field (E_{safe}) for specific %N and EOT . This is easily done by inverting equation (3.2) such that –

$$E_{safe} = \frac{3}{2\gamma} \ln \left[\Delta V_{T(max)} / \left\{ A * EOT * \exp(-0.126/k_B T_{stress}) (E_{safe} - E_{dep})^{2/3} t_{life}^{0.14} \right\} \right] \quad (7.9)$$

where (as used in Figure 7.8) $\Delta V_{T(max)} = 60$ mV, $t_{life} = 5$ years, $E_{D1} = 0.9$ eV, $n = 0.14$, $E_{dep} = \sqrt{2q\epsilon_{Si}N_D^+\phi_{dep}}/\epsilon_{SiO_2}$, $\phi_{dep} \sim 2\phi_{F(bulk)+} + 6kT/q$ [323], and A_T , γ (γ_T) variation are obtained from Figure 3.22. V_{safe} is later calculated using equation (7.4).

Thus, we recalculate the design diagram (dotted red lines in Figure 7.8) using analytical expressions in equations (7.4)-(7.9), and compare it with the procedure presented in main paper. Both approaches agree with each other reasonably well (though curves at different V_G do not fall on top of each other, due to the approximations involved in the obtained analytical expressions), motivating the use these simple equations over sophisticated QM simulation in constructing the design diagram.

7.8. Summary

We have analyzed an extensive set of leakage and degradation data to construct a co-optimization scheme of J_G and NBTI for arbitrary %N and EOT combinations. In such process, we highlight the importance of gate leakage vs. NBTI study for transistors having plasma oxynitride dielectric. A design diagram is proposed based on such co-optimization, which enables one to establish the DC (AC)-NBTI limited upper limit of %N for core CMOS technologies. Using the procedure described as a guideline and incorporating detailed nitrogen profile information, with effect of boron penetration, should enable one to reliably (up to a desired NBTI lifetime) operate a transistor with optimum static power dissipation. We anticipate that such analysis would have broad impact for optimization of %N content in sub-2nm gate dielectrics.

8. DEFECT FORMATION AND TRANSISTOR PERFORMANCE

8.1. Introduction

We have so far presented models for analyzing defects, present at the oxide/substrate interface in chapter 3 and 4. This is followed by a study on the dynamics of oxide defects in chapter 5. Later in chapter 6, we set up a methodology for proper characterization of defect formation in nanoscale transistors. Our defect modeling efforts are also utilized in chapter 7 for developing a scheme for co-optimizing (both) gate leakage and NBTI in transistors having nitrated dielectric. In this chapter, we further use our understanding of defect formation (chapters 3-6) or threshold voltage variation (ΔV_T) for studying the impact of defect generation on transistor parameters like sub-threshold slope and mobility. The analysis on variations of sub-threshold slope (SS) and V_T enables us to study off-current (I_{OFF}) variation in transistors, where off-current is dominated by thermionic emission. Similarly, we also investigate the effects of defect formation on on-current (I_{ON} : $I_{D,lin}$ and $I_{D,sat}$) and transconductance (g_m) of a transistor. This enable us to propose a variation-resilient transistor technology within existing CMOS architecture.

8.2. Defects on V_T and SS Variation

It has been shown in earlier chapters that under the application of negative gate bias or NBTI stress, there will be generation of interfacial defects (N_{IT}) and hole trapping into pre-existing oxide defects (N_{HT}). These will result in a decrease in V_T of a PMOS transistor or an increase in $|V_T|$ (see Figure 6.3b). Moreover, comparison of Δm vs. ΔV_T for type-I and type-II transistors (see Figure 8.1) reveals that Δm (hence, change in sub-threshold slope) ~ 0 for type-II devices at small ΔV_T and Δm only starts to increase after

ΔV_T is considerably higher²⁶. This undoubtedly shows that ΔV_T for these type-II transistors is initially dominated by N_{HT} (thus, does not affect SS or, m) and generation of N_{IT} is observed only at higher ΔV_T . Similar presence of dominant N_{HT} in type-II transistors at early stress is also validated by time and temperature dependencies of NBTI degradation [129, 154], as discussed in section 6.7.5. Therefore, only interfacial defect during NBTI stress predominantly changes SS . In such analysis involving SS change (ΔSS), one should be careful about the uncertainty associated with ΔSS estimation, which comes from the non-uniformity of N_{IT} distribution within the silicon bandgap. As $SS \sim m$ is a function of ψ_S (see equation (6.3)), determination of ΔSS will depend on ΔC_{IT} at that ψ_S , hence on the V_G at which it is determined. For a particular amount of $|\Delta V_T|$, ΔSS can be expressed as,

$$\Delta SS (@ \psi_S) = 2.3k_B T / q \Delta m = 2.3k_B T \frac{\Delta D_{IT} (@ E_{FS} |_{\psi_S})}{C_{ox}} \quad (8.1)$$

where E_{FS}/ψ_S indicates the position of E_{FS} at the ψ_S under consideration. Thus depending on the V_G at which SS is estimated, there can be significant uncertainty in ΔSS estimation, as observed in Figure 3.24a-c for three different devices. However, one can always obtain an average value of ΔSS by determining it using a wide V_G range, such that

$$\Delta SS_{avg} = 2.3k_B T / q \Delta m_{avg} = 2.3k_B T \frac{\Delta V_T}{qE_{g,Si}} \quad (8.2)$$

where $E_{g,Si}$ is the bandgap of silicon in eV.

On the other hand, presence of electron trapping (N_{ET}) causes an increase in V_T , during positive bias stress of a NMOS transistor (especially the one with high- κ dielectric [43]). This phenomenon is commonly known as positive bias temperature instability (PBTI). Since, interface trap or C_{IT} is not affected due to electron trapping into oxide

²⁶ Similar behavior (Δm vs. ΔV_T) is also reported in [324] for devices having higher interfacial nitrogen concentration, thus higher hole trapping.

defects (which are located sufficiently away from the interface [44, 234, 325-327]), ΔSS is negligible for PBTI experiments (Figure 8.2).

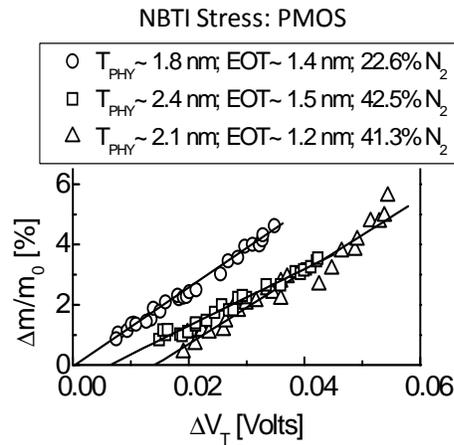


Figure 8.1: Variation of Δm for Type-I (circles) and Type-II (squares and triangles) devices. Type-II transistors have negligible Δm at small ΔV_T , thus showing signatures of hole trapping.

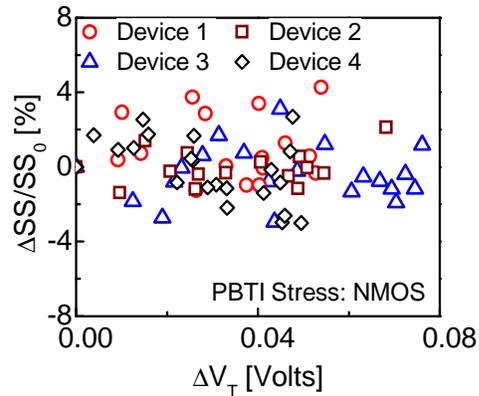


Figure 8.2: Variation of body-effect coefficient (m) or SS is negligible and within the error margin of SS estimation during PBTI stress. Here, the measurement is performed on a number of NMOS transistors, and each of them indicates that ΔSS is (on average) independent of the level of PBTI-induced N_{ET} .

8.3. Defects on Off Current Variation

Since I_{OFF} for a transistor is dominated by thermionic emission [282] (when band-to-band tunneling is negligible²⁷), it can be expressed as (*c.f.*, equation (6.9)),

$$|I_{OFF}| \sim \exp\left[-q|V_T|/mk_B T\right] \quad (8.3)$$

Differentiating equation (8.3) in sub-threshold regime ($V_G < V_T$), we obtain –

$$\frac{\Delta I_{OFF}}{I_{OFF}} = \frac{1}{mk_B T} \left(-|\Delta V_T| + \frac{\Delta m}{m} |V_T| \right) \quad (8.4)$$

where, m is the body-effect co-efficient, as defined in equation (6.3). Interface trap generation during NBTI stress increases $|\Delta V_T|$, as well as, Δm or ΔSS . According to Figure 8.1, in a type-I transistor, $|\Delta V_T| \sim 30\text{mV}$ causes $\Delta m/m < 0.05$. Thus for a present day MOS transistor, having a $|V_T|$ of $\sim 0.3\text{V}$ [7], $|\Delta V_T|$ is greater than $\Delta m/m|V_T|$; *i.e.*, I_{OFF} increases due to N_{IT} generation. On the other hand, hole trapping (during NBTI) or electron trapping (during PBTI) only increases $|\Delta V_T|$, thus always causes a decrease in I_{OFF} .

8.4. Defects on Mobility Variation

8.4.1. Effect of N_{IT} (NBTI Stress)

As presented in section 6.4.3, when measured at constant E_{eff} , generation of N_{IT} always decreases μ_{eff} ; an observation broadly consistent with the literature [89, 267].

²⁷ If I_{OFF} is dominated by band-to-band tunneling (as the case for short-channel transistors having low bandgap Ge substrates [328]), any change in V_T will directly reflect in a corresponding change in I_{OFF} . Hence, self-compensation for I_{OFF} is also unexpected in such short-channel transistors.

However, when measured at constant V_G , μ_{eff} is observed to either increase (Figure 6.7a) or decrease (Figure 6.7b), depending on the steepness of the μ_{eff} - E_{eff} relationship.

The variation in μ_{eff} - E_{eff} and μ_{eff} - V_G relationships before and after N_{IT} generation can be explained, when we consider the well-known fact that an increase in N_{IT} always acts as an extra coulomb scattering center [89, 203, 266, 267] (Figure 8.3a). As a result, when measured at constant E_{eff} , μ_{eff} always decreases due to N_{IT} generation (A \rightarrow B in Figure 8.3b). However, if μ_{eff} is measured at constant V_G , we need to consider contributions from E_{eff} reduction as well (B \rightarrow C in Figure 8.3b). Because N_{IT} generation not only reduces $\mu_{eff}@E_{eff}$, it also reduces the amount of substrate charges (mostly Q_{inv} , when the transistor is biased in strong inversion) at constant V_G (Figure 8.3c-d). As such, depending on the steepness (θ) of the μ_{eff} vs. E_{eff} relationship, transistors with larger θ will have dominance from B \rightarrow C movement and thus show an overall increase in $\mu_{eff}@V_G$; whereas those with smaller θ will have dominance from A \rightarrow C movement (Figure 8.3b) and thus show an overall decrease in $\mu_{eff}@V_G$.

To experimentally verify our hypothesis, we choose three types of uniaxially strained transistors (which are compressively strained using SiGe source/drain and contact etch stop layer) having $L = 0.13\mu\text{m}$, $0.26\mu\text{m}$ and $1\mu\text{m}$ ($EOT \sim 1.4\text{nm}$). For similar etch depth and contact etch stop layer, short-channel transistors are expected to have more strain compared to long channel transistors [329], which is evident from the μ_{eff} vs. $|E_{eff}|$ curve for these three transistors (Figure 8.4a). Moreover, we observe an increase in steepness of μ_{eff} vs. $|E_{eff}|$ curve with reduced L (*i.e.*, increasing strain). As a result, we observe $\Delta\mu_{eff}/\mu_{eff0}$ to be always negative for $1.00\mu\text{m}$ transistor (Figure 8.4b) and always positive for $0.13\mu\text{m}$ transistor (Figure 8.4d). Similarly, Figure 8.4c indicates that $0.26\mu\text{m}$ transistor (which have intermediate steepness for μ_{eff} vs. $|E_{eff}|$ curves) have negative $\Delta\mu_{eff}/\mu_{eff0}$ at lower V_G (similar to $1\mu\text{m}$ transistor), but it becomes positive at higher V_G (similar to $0.13\mu\text{m}$ transistor). All these transistors show degradation in $\Delta\mu_{eff}/\mu_{eff0}$, when measured at constant E_{eff} , irrespective of the steepness in μ_{eff} vs. $|E_{eff}|$ relationship (Figure 8.4e) – a fact that is well known in classical reliability literature [89, 124, 165, 203].

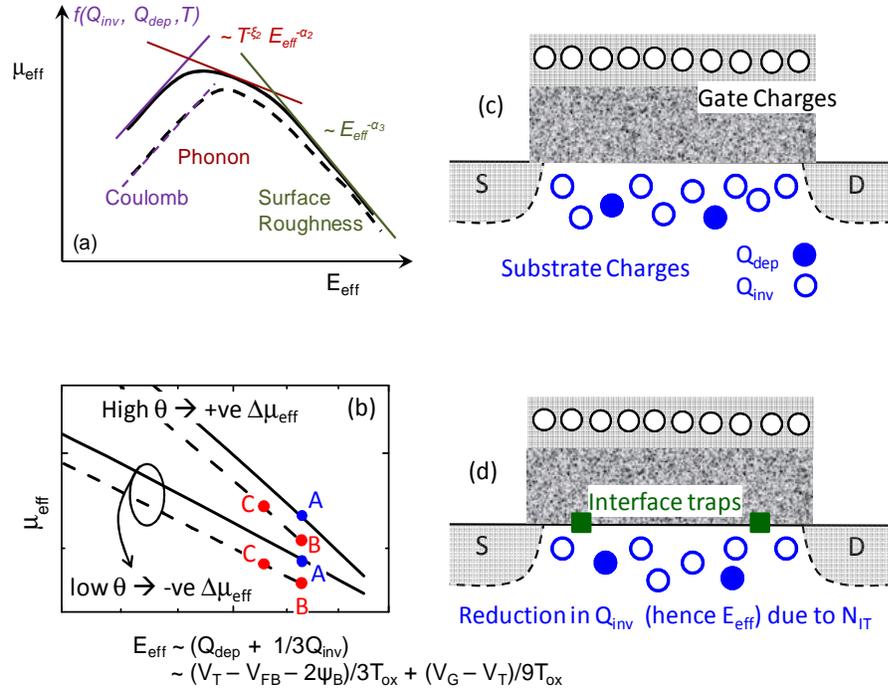


Figure 8.3: (a) A schematic representation of the effect of N_{IT} generation on μ_{eff} - E_{eff} plot. The functional dependencies of different mobility components are provided in section 6.4.4. N_{IT} induces extra coulomb scattering and thus reduces μ_{eff} , especially at low E_{eff} . (b) μ_{eff} vs. E_{eff} in both pre- and post- N_{IT} generation conditions for different steepness of μ_{eff} vs. E_{eff} characteristics. Although μ_{eff} decreases at constant operating E_{eff} ($A \rightarrow B$); at constant V_G , E_{eff} decreases with degradation ($B \rightarrow C$) and thus gives rise to increase/decrease in μ_{eff} ($A \rightarrow C$), depending on the steepness (θ) of the μ_{eff} vs. E_{eff} plot. Physics of E_{eff} reduction: (c-d) At constant V_G , there should be equal amount of gate and substrate charges in order to maintain the electrostatics. As a result, effect of N_{IT} generation, when measured at constant V_G , should consider a corresponding reduction in substrate charge (c \rightarrow d). Since in an inverted MOS, Q_{dep} remains invariant of N_{IT} generation, $E_{eff} \sim Q_{dep} + 1/3Q_{inv}$ reduces with N_{IT} .

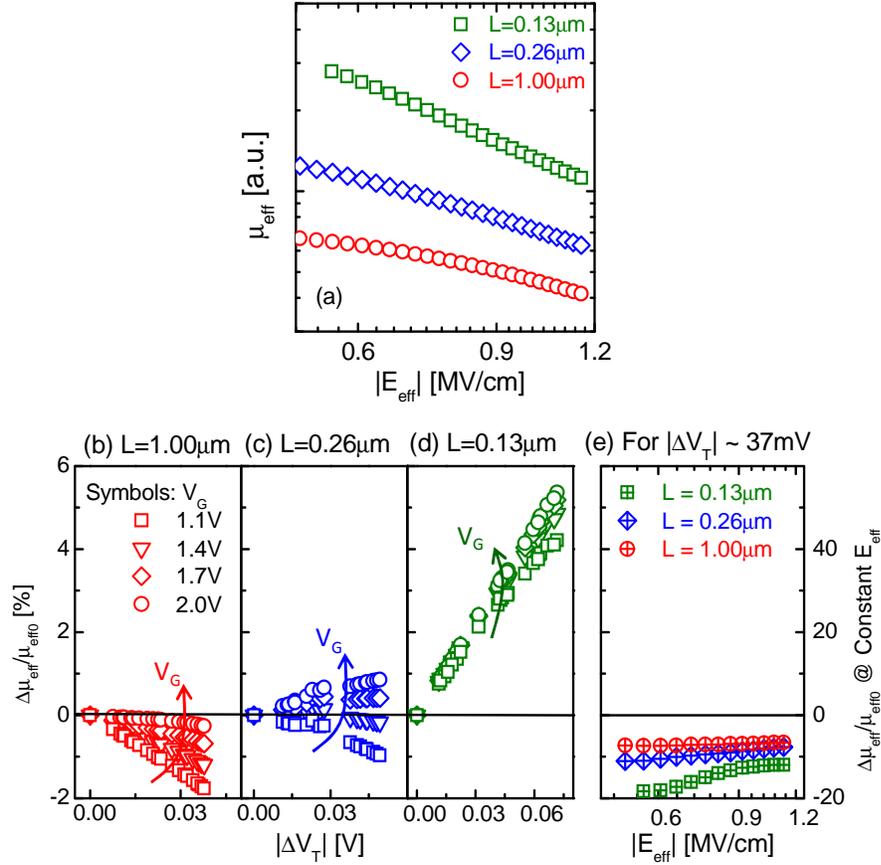


Figure 8.4: (a) μ_{eff} vs. E_{eff} for the uniaxially strained PMOS devices under study, having channel length of $L = 0.13, 0.26$ and $1.00 \mu\text{m}$. Short channel devices have more strain and steeper curve compared to long channel devices. $\Delta\mu_{eff}$ for (b) $L=1.00\mu\text{m}$ transistor is negative, whereas for (d) $L=0.13\mu\text{m}$ transistor it is always positive. For $L=0.26 \mu\text{m}$ transistor in (c), $\Delta\mu_{eff}$ is initially negative at low $|V_G|$ and becomes positive at high $|V_G|$. Variation of $\Delta\mu_{eff}/\mu_{eff0}$ from negative to positive is evident as L is reduced, or strain is increased. (d) All these transistors show degradation in $\Delta\mu_{eff}/\mu_{eff0}$, when measured at constant E_{eff} (here shown for $|\Delta V_T| \sim 37\text{mV}$).

8.4.2. Effect of N_{ET} (PBTI Stress)

To study the variation in μ_{eff} due to electron trapping (N_{ET}), within high- κ dielectric of NMOS transistors, we measure change in maximum transconductance ($\Delta g_{m,max}$) under the application of PBTI stress. Here, $\Delta g_{m,max}$ is used as a signature of $\Delta\mu_{eff}@E_{eff}$, which

can be understood by differentiating the expression for $I_{D,lin}$ (see equation (6.4)) with respect to V_G and then taking its maximum. Such operation leads to –

$$g_{m,max} = \max\left(\frac{\partial I_{D,lin}}{\partial V_G}\right) = \mu_0 C_{ox} \frac{W}{L} |V_{DS,lin}|. \quad (8.5)$$

Equation (8.5) considers the fact that beyond $g_{m,max}$, decrease in g_m is mostly due to negative slope of the μ_{eff} vs. V_G or μ_{eff} vs. E_{eff} relationship, which in turns comes from the increase in phonon and surface roughness scattering [272]. Our experiment indicates that even a significant amount of electron trapping or corresponding ΔV_T , during PBTI stress on NMOS transistors, causes negligible change in $\Delta g_{m,max}$ (Figure 8.5a). Similar invariance of $\Delta g_{m,max}$ due to N_{ET} are also reported in [47].

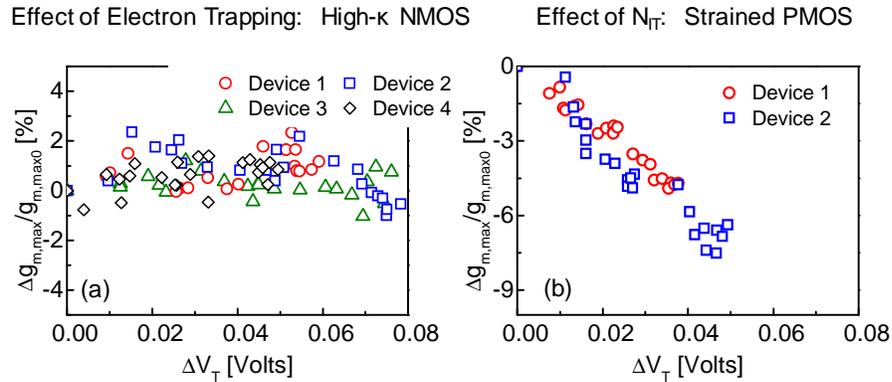


Figure 8.5: (a) Electron trapping into oxide defects during PBTI stress causes negligible change in $\Delta g_{m,max}$. Whereas, (b) N_{IT} generation decreases $\Delta g_{m,max}$ significantly irrespective of the amount of strain, which is similar to the decrease in $\Delta\mu_{eff}@E_{eff}$ shown in Figure 8.4e. This indicates that although there is decrease in $\Delta\mu_{eff}@E_{eff}$ due to N_{IT} generation, electron trapping has negligible impact on $\Delta\mu_{eff}@E_{eff}$.

Next, we compare $\Delta g_{m,max}$ due to N_{ET} (Figure 8.5a) and N_{IT} (Figure 8.5b). We observe that considerable decrease in $\Delta\mu_{eff}@E_{eff}$ during N_{IT} generation (Figure 8.4e) leads to an equivalent decrease of $\Delta g_{m,max}$ (Figure 8.5b). On the other hand, invariance of $\Delta g_{m,max}$ due to N_{ET} (Figure 8.5a) means negligible change in $\Delta\mu_{eff}@E_{eff}$ during PBTI

stress. Thus, Figure 8.5a indicates that there is negligible amount of remote coulomb scattering [330-332] from the electron trapping centers in the NMOS transistors under study. Now that we know SS , $g_{m,max}$, and $\mu_{eff}@E_{eff}$ are relatively insensitive to trapping in oxide defects, let us study how $\mu_{eff}@V_G$ is affected by such trapping.²⁸ Since $\mu_{eff}@E_{eff}$ remains unchanged due to N_{ET} (*i.e.*, $A \rightarrow B$ movement of Figure 8.3b is negligible) and E_{eff} always reduces due to N_{ET} at constant V_G (since $B \rightarrow C$ movement in Figure 8.3b is present for any defects), $\Delta\mu_{eff}@V_G$ for N_{ET} (or, in general, due to oxide defects) will *always* be positive (see Figure 8.6c). Therefore, we obtain comparable values of positive $\Delta\mu_{eff}@V_G$ for N_{ET} in high- κ NMOS transistor and for N_{IT} in the highly strained PMOS transistor of Figure 8.4a (also compare Figure 8.6c with Figure 8.6a), even though $\mu_{eff}-E_{eff}$ steepness for the high- κ NMOS transistor is smaller (see Figure 8.6b) than that of highly strained PMOS transistor of Figure 8.4a. This is because the presence of N_{IT} causes negative $\Delta\mu_{eff}@E_{eff}$ (Figure 8.5b), which dictates that $\mu_{eff}-E_{eff}$ steepness be enhanced for N_{IT} to achieve overall positive $\Delta\mu_{eff}@V_G$. On the other hand, $\Delta\mu_{eff}@E_{eff}$ is negligible for N_{ET} (Figure 8.5a), thus N_{ET} -related $\Delta\mu_{eff}@V_G$ is always positive, irrespective of $\mu_{eff}-E_{eff}$ steepness.

8.4.3. Effect of N_{HT} (NBTI Stress)

Our study on electron trapping reveals that the $B \rightarrow C$ movement in Figure 8.3b dominates for oxide defects, when remote coulomb scattering from these defects (*i.e.*, $A \rightarrow B$ movement in Figure 8.3b) are negligible. Since hole trapping also happens within the oxide, we expect hole trapping during NBTI stress to show similar behavior.

²⁸ As shown in section 8.5, knowledge of $\mu_{eff}@E_{eff}$ can not be sufficient to understand the effect of defect on transistor performance, hence can not be sufficient to perform SPICE analysis, as well.

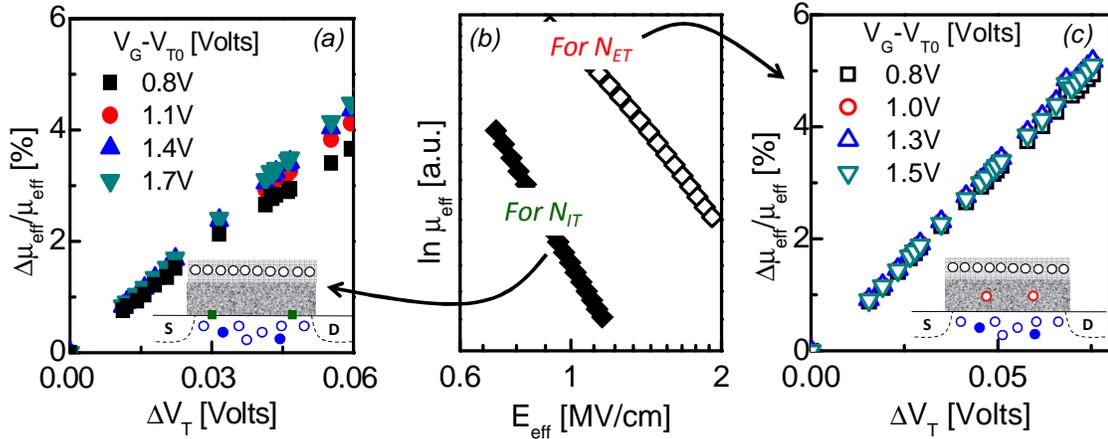


Figure 8.6: (a) $\Delta\mu_{eff}$ at different (V_G-V_{T0}) for the highly strained PMOS transistor in Figure 8.4. (b) Comparison of θ between the high- κ NMOS transistor under study and the highly strained PMOS transistor of Figure 8.4. (c) Although NMOS transistor has smaller θ , presence of negligible $\Delta\mu_{eff}@E_{eff}$ and E_{eff} reduction at constant V_G stress provide positive $\Delta\mu_{eff}@V_G$ due to N_{ET} .

8.5. Effect of Defect on I_{ON} : Drift-diffusion Theory

So far we have explained how threshold voltage (hence, amount of carriers within the channel) and mobility of MOS transistor varies due to interfacial or oxide defects. Here, we will use this information related to carrier number variation and mobility variation to study the impact of defect formation on I_{on} using drift-diffusion theory. As shown in Figure 6.19, variation in SS has negligible impact on I_{on} .

8.5.1. $I_{D,lin}$ Variation due to N_{IT}

For estimating variations in $I_{D,lin}$ within the framework of drift-diffusion theory, we use equation (6.16) for estimating $\Delta I_{D,lin}@V_G$. For explaining the major features of $\Delta I_{D,lin}@V_G$, we can simplify equation (6.16) as –

$$\frac{\Delta I_{D,lin}}{I_{D,lin0}} \Big|_{V_G} = \frac{\Delta\mu_{eff}}{\mu_{eff0}} \Big|_{V_G} - \frac{|\Delta V_T|}{V_G - V_{T0}}. \quad (8.6)$$

As we know by now, $|\Delta V_T|$ always increases for any form of defects; whereas $\Delta\mu_{eff}@V_G$ sometimes increases or decreases with N_{IT} generation (Figure 8.4) and always increases with hole trapping into oxides (Figure 8.6b). As a result, $\Delta I_{D,lin}@V_G$ depends on the type of transistor under study.

For example, let us consider the case where N_{IT} generation dominates (Figure 8.4), as in type-I PMOS transistor during NBTI stress (see section 6.7.5 for a discussion on type-I transistors). There the transistors with positive $\Delta\mu_{eff}@V_G$ (i.e., $L = 0.13 \mu\text{m}$ transistor in Figure 8.4d) can *self-compensate* the effect from $|\Delta V_T|$, thus resulting reduced variation in $\Delta I_{D,lin}$ (Figure 8.7c). Such invariance in $\Delta I_{D,lin}$ is more effective at higher $|V_G|$, where the second term in equation (8.6) is less negative. Under similar operating condition, $\Delta\mu_{eff}/\mu_{eff0}$ is negative for $L = 1 \mu\text{m}$ transistor (Figure 8.4b), which then adds up with the second term in (8.6) and hence $\Delta I_{D,lin}$ always degrades for this transistor (Figure 8.7a). Similarly, $L = 0.26 \mu\text{m}$ transistor having both positive/negative $\Delta\mu_{eff}/\mu_{eff0}$ (Figure 8.4c) shows an intermediate behavior in terms of $\Delta I_{D,lin}$ (Figure 8.7b).

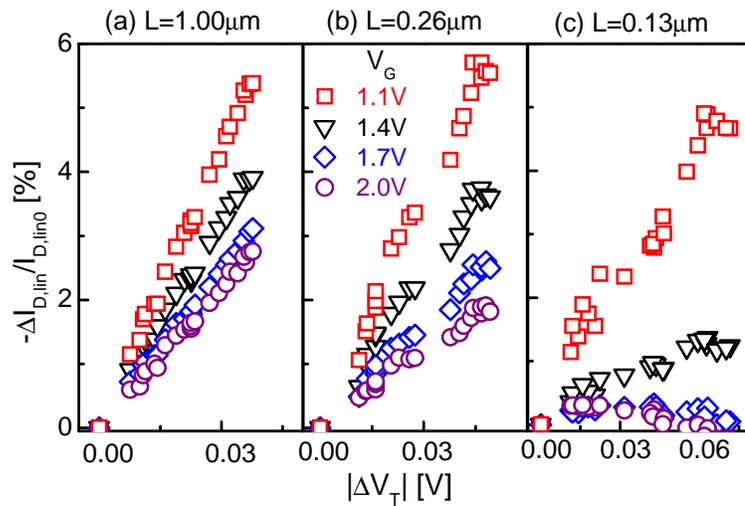


Figure 8.7: $\Delta I_{D,lin}$ for (a) $1.00\mu\text{m}$, (b) $0.26\mu\text{m}$, and (c) $0.13\mu\text{m}$ transistors, measured at various V_G . Note that, although $0.13\mu\text{m}$ transistor is stressed more compared to the other transistors, it has the lowest $\Delta I_{D,lin}$ at $|V_G| > 1.1\text{V}$.

8.5.2. $I_{D,sat}$ Variation due to N_{IT}

For analyzing the impact of defect formation on $I_{D,sat}$, we characterize $I_{D,sat}$ using

$$I_{D,sat} = A_{sat} (V_G - V_{T,sat})^{\alpha_{sat}} \quad (8.7)$$

where A_{sat} and α_{sat} are complicated functions of μ_{eff} , saturation velocity, *etc.* [282]. We observe that similar to the self-compensation between μ_{eff} and V_T variations in linear region for $L = 0.13 \mu m$ transistor (Figure 8.7c), compensation also exists between the variations in A and V_T (Figure 8.8b). Whereas, $L = 1.00 \mu m$ transistor has no such compensation (Figure 8.8a), which again supports its behavior in linear region (Figure 8.7a). Due to complicated dependency of factor A_{sat} and α_{sat} on other physical parameters, it becomes difficult to perform any further analysis using drift-diffusion theory. Such limitation motivated us to use scattering theory [333] to have a physical insight on $I_{D,sat}$ variation (see section 8.6 and Appendix E).

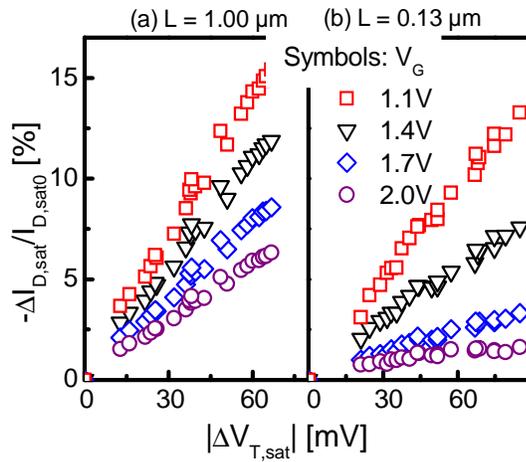


Figure 8.8: $\Delta I_{D,sat}$ for (a) $1.00 \mu m$ and (b) $0.13 \mu m$ transistors measured at various V_G . Similar to the observation involving $\Delta I_{D,lin}$ in Figure 8.7, $L = 0.13 \mu m$ transistor has negligible $\Delta I_{D,sat}$ compared to $L = 1.00 \mu m$ transistor.

8.5.3. $I_{D,lin}$ and $I_{D,sat}$ Variation due to N_{ET}

Figure 8.7c indicates the presence of self-compensation between $\Delta\mu_{eff}$ and ΔV_T terms of (8.6), leading to negligible $I_{D,lin}$. Since $\Delta\mu_{eff}@V_G$ is always positive for oxide defects like N_{ET} (see section 8.4.2), $\Delta I_{D,lin}$ can indeed be totally compensated at higher $(V_G - V_{T0})$ and partially compensated at lower $(V_G - V_{T0})$ through mobility improvement (see Figure 8.9a). However, for the same transistors at similar $(V_G - V_{T0})$, we observe reduced self-compensation in the saturation regime (see Figure 8.9b). Nevertheless, in both cases of Figure 8.8b and Figure 8.9b, the improvement through partial self-compensation (or reduction in $\Delta I_{D,sat}$) is still significant and relevant for reduction in design margins.

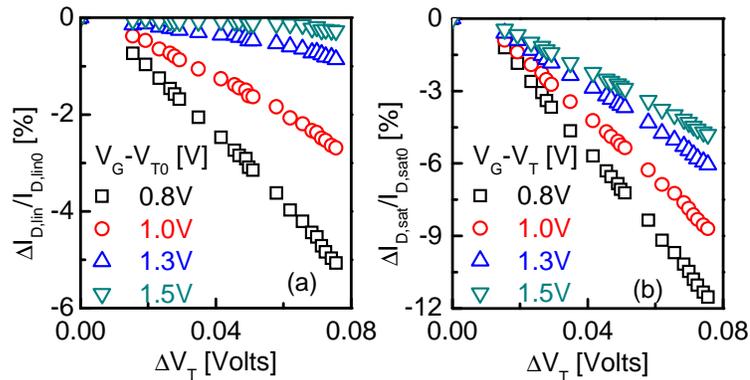


Figure 8.9: (a) $\Delta I_{D,lin}$ and (b) $\Delta I_{D,sat}$ at different $V_G - V_{T0}$ for the high- κ NMOS transistor, having $\Delta\mu_{eff}@V_G$ values of Figure 8.6c.

8.6. Effect of Defect on I_{on} : Scattering Theory

As the channel length of nanoscale MOS transistors is comparable to the carrier mean-free path [333-340], quasi-ballistic transport becomes important in characterizing them. In such quasi-ballistic transport regime, scattering theory [333] suggests that classical equation (obtained from drift-diffusion and velocity saturation) for I_D of a transistor (width W , channel length L), should be replaced by [341]:

$$I_D / W = Q_{inv} \frac{1-r}{1+r} \left\{ v_T \frac{\mathfrak{S}_{1/2}(\eta_F)}{\mathfrak{S}_0(\eta_F)} \right\} \frac{1 - \mathfrak{S}_{1/2}(\eta_F - U_D) / \mathfrak{S}_{1/2}(\eta_F)}{1 + \left[(1-r)/(1+r) \right] \left[\mathfrak{S}_0(\eta_F - U_D) / \mathfrak{S}_0(\eta_F) \right]} \quad (8.8)$$

where $Q_{inv} = C_G(V_G - V_T)$, r is the back-scattering co-efficient, v_T is the thermal velocity, η_F is the location of Fermi-level near the top-of-the-barrier at the source-channel junction (normalized to thermal voltage $k_B T/q$), U_D is drain voltage (normalized to $k_B T/q$) and $\mathfrak{S}_{1/2}$ and \mathfrak{S}_0 are Fermi integrals of order $1/2$ and 0 respectively. Equation (8.8) can be used for extracting parameters in linear region, but presence of high V_{DS} in saturation region leads to –

$$I_{D,sat} / W = C_G (V_G - V_{T,sat}) v_{inj} \frac{(1 - r_{sat})}{(1 + r_{sat})}, \quad (8.9)$$

where v_{inj} is the injection velocity, $r_{sat} = \ell/(\ell + \lambda)$, $\ell \sim (1 \sim 3)kT/q / \mathcal{E}_{0(+)}$ [342, 343], λ is the mean-free path, $\mathcal{E}_{0(+)}$ is the electric field along the transport direction near the top-of-the-barrier at the source-channel junction, and ℓ is the distance along the transport direction over which there is approximately $(1 \sim 3)kT/q$ [342, 343] voltage drop. However, use of Maxwell-Boltzmann approximation for $\mathfrak{S}_{1/2}$ and \mathfrak{S}_0 leads to –

$$I_{D,lin} / W = C_G (V_G - V_{T,lin}) \frac{v_T}{2kT/q} (1 - r_{lin}) V_{DS,lin}, \quad (8.10)$$

$$I_{D,sat} / W = C_G (V_G - V_{T,sat}) v_T \frac{(1 - r_{sat})}{(1 + r_{sat})}, \quad (8.11)$$

for linear (_{lin}) and saturation (_{sat}) regions, respectively, where, $r_{lin} = L/(L + \lambda)$. Scattering theory presumes that the transport near the top-of-the-barrier (virtual source) dictates the current flow and offers a simple correlation between $I_{D,lin}$ and $I_{D,sat}$, *i.e.*, from equations (8.10) and (8.11), we have, $I_{D,lin}/I_{D,sat} = V_{DS,lin}/(2kT/q) * (\lambda + 2\ell)/(\lambda + L) \sim B_{lin}/B_{sat}$; where, $B_{lin} = \lambda/(\lambda + L)$ and $B_{sat} = \lambda/(\lambda + 2\ell)$ are ballistic efficiencies of a MOS transistor in linear and saturation regions, respectively.

Equations (8.8)-(8.11) can be used to characterize the variation of scattering theory parameters before and after defect generation. Such analysis (see Appendix E for details)

suggests that in linear regime of source-to-drain transport, source-injected carriers are continuously scattered over the entire channel length (L) and hence a change in mobility in all parts of the channel will affect the current transport. On the other hand, in the saturation regime, electric field along the channel is very high. As a result, for calculating current transport, we only need to consider scattering of source-injected carriers over a characteristic distance (ℓ) away from the top of the energy barrier, located near the source end. Carriers that have crossed ℓ indeed face scattering, but these scattered carriers can not overcome the thermal barrier to reach the source side and always gets collected at the drain end. Hence, $\Delta\mu_{eff}$ will have reduced impact on $I_{D,sat}$ (having effective scattering length of ℓ), compared to $I_{D,lin}$ (having scattering length of L). Thus, the overall change in ΔI_D due to BTI degradation in linear and saturation regions can be approximated as,

$$\left. \frac{\Delta I_{D,lin(sat)}}{I_{D,lin(sat)0}} \right|_{V_G} = (1 - B_{lin(sat)}) \left. \frac{\Delta\mu_{eff}}{\mu_{eff0}} \right|_{V_G} - \frac{|\Delta V_{T,lin(sat)}|}{V_G - V_{T,lin(sat)0}}. \quad (8.12)$$

Now, the ballistic coefficients in the linear regime (B_{lin}) is ~ 0 , in the saturation regime (B_{sat}) ranges from $0 \ll B_{sat} < 1$ [333, 344]. Therefore, one expects self-compensation between the $\Delta\mu_{eff}$ and ΔV_T terms of equation (8.12) to be more effective for $\Delta I_{D,lin}$ compared to $\Delta I_{D,sat}$. This is exactly what we have observed in sections 8.5.2 and 8.5.3 for interface and oxide defects, respectively.

8.7. Summary

Throughout this chapter, we have explained the effect of interface trap generation and electron/hole trapping on different device parameters, used for characterizing MOS transistors. As such, we have identified several non-intuitive variations in transistor parameters. For example,

- We have identified the uncertainty involved in sub-threshold slope measurement and explain how it is mainly affected by interfacial defects.
- We have presented the first report and consistent explanation for the increase in effective mobility when measured at constant gate voltage.

- We also have provided the example of transistors having negligible fluctuation in drain current due to defect generation through a compensation between opposite signs of mobility fluctuation and carrier number fluctuation (see next chapter for further analysis on this topic).
- Finally, in relation to this chapter, Appendix E broadens the application of scattering theory with time-dependent parameter shifts and shows that – despite dramatic changes in the device parameters – the relationship between the ballistic efficiencies are essentially invariant with respect to its time-zero value, once the effect of threshold voltage shift is taken into account.

9. TOWARDS VARIATION RESILIENT TRANSISTORS AND CIRCUITS

9.1. Introduction

In chapter 8, we have described how and interplay between mobility and carrier number fluctuation can lead to a ‘degradation-free’ transistor, having negligible drain current variation. In this chapter, we will extend this concept of optimizing temporal V_T fluctuation and show how a similar concept can also be used for self-compensating different sources of spatial V_T fluctuations, like oxide thickness fluctuation, metal-work function fluctuation, $1/f$ noise fluctuation originating from random placement of defects, *etc.* As such, we propose the concept of variation-resilient transistor, which may relax the requirement of circuit and system level optimization, currently being widely used to handle these variations. In section 9.2, we start this chapter by discussing the need for studying variation resilience within CMOS architecture. Then in section 9.3, we discuss the transistor design principles for achieving variation resilience. Later, in sections 9.4-9.7, we demonstrate the validity of our proposed principle through simulation and experiment. Finally, we conclude in section 9.8 by highlighting the circuit level implications of our idea of self-compensation.

9.2. Motivation

Until now, the research on transistor’s parametric variations has focused on correlating the fluctuation in parameters to variation in device characteristics [9, 119, 168], as well as mitigating these variations through the use of extra processing steps (*e.g.*, use of monolayer doping for reducing dopant fluctuation [345], plasma nitridation over thermal nitridation in SiON CMOS technology for reducing time-dependent V_T variation

[129]). In spite of these efforts, device designers often feel that process modifications alone can not adequately address the variation problem without unacceptable loss in device performance. So, a standard option is to operate the transistors with an extra guard band voltage, over and above the voltage that is required for nominal operation [58, 59], so that the circuit remains functional despite significant parametric variation. Similarly, there are proposals for using adaptive body bias [60-62], adaptive power supply (*i.e.*, circuit sleeping) [63], (area-)resizing of transistors along the critical path [64], *etc.* for minimizing the impact of temporal variability or reliability. Such circuit level reliability optimization requires one to monitor the degradation level of IC at different stages of operation using specialized circuits like ‘Silicon Odometers’ [58] (based on the measurement of quiescent leakage current [346], for example).

In chapter 8, we have discussed the possibility of self-compensation of NBTI in strained PMOS transistors (see sections 8.5.1 and 8.5.2) and of PBTI in high- κ NMOS transistors (see section 8.5.3). Sufficient negative steepness in the effective mobility (μ_{eff}) of the inversion carrier vs. effective vertical electric field (E_{eff}) characteristic, achieved through uniaxial strain in Figure 8.4, has been the key requirement for designing such *self-compensated transistors*. A steep μ_{eff} - E_{eff} characteristic ensures that NBTI and PBTI in these transistors result in an opposing variation in μ_{eff} and carrier density (N_{inv}) within the inverted PMOS channel. As a result, the drivability (*i.e.*, drain current $I_D \sim \mu_{eff}N_{inv}$) of a variation-resilient transistor becomes less sensitive to NBTI and PBTI (hence, BTI as a whole).

In the subsequent sections 9.3-9.8, we broaden the discussion of self-compensation in four important ways. *First*, we show that the principle of self-compensation is also relevant to different process variations and can reduce the effect of variations in different oxide and gate parameters (*e.g.*, EOT , Φ_{MS}); as such, the ‘degradation-compensated’ strained transistor proposed in [347] can also be considered as a ‘variation-resilient’ transistor. *Second*, we characterize the μ_{eff} - E_{eff} relationship as a function of uniaxial strain and identify the *practical* limits of uniaxial strain technology in achieving self-compensation. *Third*, we highlight the *limitations* of achieving self-compensation in modern short channel transistors by showing its inability to reduce the effects of channel

length variation, random dopant fluctuation (RDF), as well as variation in off-state leakage current. And *finally*, we verify the consequence of self-compensation in a set of digital circuits and show that the delay degradation in digital circuits can indeed be reduced by considering physical changes in transistor parameters (like V_T , E_{eff} , μ_{eff} , *etc.*) during circuit simulation.²⁹ Therefore, we illustrate how variation resilience may ease the burden of designers in guard-banding against some of the variability issues (like BTI, EOT -fluctuation, Φ_{MS} -fluctuation) that the nanoscale CMOS technology is currently struggling to handle.

9.3. Principle of Variation Resilience

Let us first review the design principle involving self-compensation in a variation-resilient transistor and compare its characteristics with a classical CMOS transistor. For this, we need to understand the effect of process-related and stress-induced parametric variation in transistors drain current. Drain current in linear region ($I_{D,lin}$) can be expressed as $I_{D,lin} \sim \mu_{eff}N_{inv}$; where, $N_{inv} \sim (V_G - V_T)$. Thus, the fluctuation of $I_{D,lin}$ ($\Delta I_{D,lin}$) can be expressed as:

$$\Delta I_{D,lin} / I_{D,lin} \sim \Delta \mu_{eff} / \mu_{eff} + \Delta N_{inv} / N_{inv} \sim \Delta \mu_{eff} / \mu_{eff} - \Delta V_T / (V_G - V_T). \quad (9.1)$$

According to equation (9.1), a transistor can be resilient to $I_{D,lin}$ variation (*i.e.*, have $\Delta I_{D,lin} \sim 0$) *only if* an increase (decrease) in μ_{eff} compensates the corresponding increase (decrease) in V_T . All MOS transistors have negative differential steepness in its $\mu_{eff} - E_{eff}$ characteristics (Figure 9.1b), mainly due to effect of surface roughness [267]; and this ‘negative steepness’ dictates that any increase in V_T is always accompanied by an increase in $\mu_{eff}@V_G$ and vice versa.³⁰

²⁹ Note that conventional SPICE simulation will fail to capture the effect of self-compensation, due to the reasons discussed in section 9.8.3.

³⁰ However, in case of interface trap (N_{IT})-related ΔV_T , one needs to consider the decrease in $\mu_{eff}@E_{eff}$ due to additional coulomb scattering introduced by N_{IT} [267]. As a

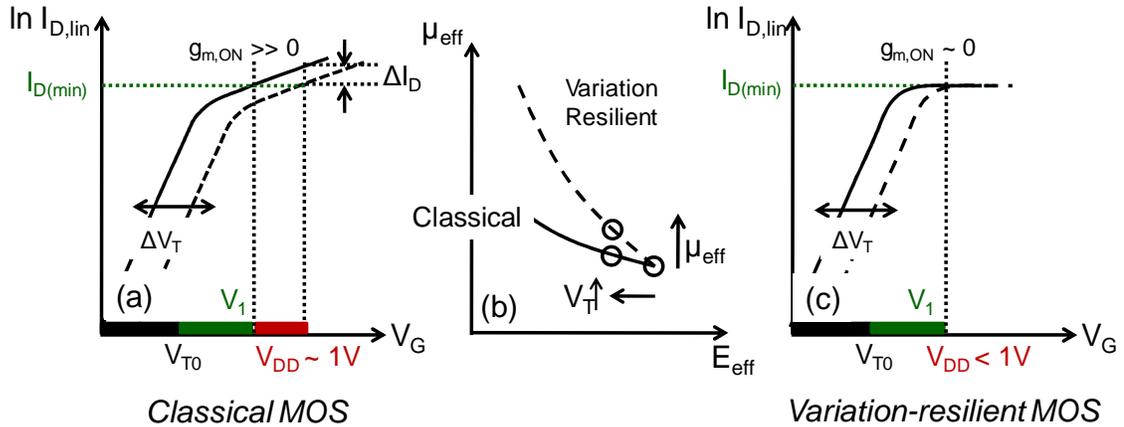


Figure 9.1: (a) Linear transfer characteristics of a classical MOS transistor, where any sources of ΔV_T lead to $\Delta I_{D,lin}$. The effect of ΔV_T can be minimized by using guard-band voltage ($V_{DD}-V_I$). (b) Comparison of mobility-field ($\mu_{eff} - E_{eff}$) relationship in classical and variation-resilient transistors. An increase in V_T always results in an increase in μ_{eff} , except for N_{IT} -related ΔV_T . (c) Transfer characteristics of the proposed variation-resilient MOS transistor, where ΔV_T is not reflected in $\Delta I_{D,lin}$. Thus variation-resilient transistor can operate with reduced guard-band.

In classical MOS transistor, the steepness of the $\mu_{eff} - E_{eff}$ or $\mu_{eff} - V_G$ characteristic is generally insufficient to fully compensate the linear increase in N_{inv} with V_G . As a result, transfer (e.g., $I_{D,lin}-V_G$) characteristic of a classical transistor is dominated by the increase in N_{inv} with V_G ; in other words, a classical transistor has finite ON-state transconductance $g_{m,ON} \equiv \partial I_D / \partial V_G|_{V_G \gg V_T}$ (Figure 9.1a). In addition, relatively shallow $\mu_{eff} - E_{eff}$ steepness in classical transistors ensures negligible change in μ_{eff} ($\Delta \mu_{eff}$) due to ΔV_T . Thus according to equation (9.1), $\Delta I_{D,lin}$ is dominated by the ΔV_T term, which makes $I_{D,lin}$ of a classical transistor sensitive to various sources of ΔV_T . To mitigate the effect of ΔV_T in these transistors, one can use a fraction of the supply voltage ($V_{DD} - V_I$ in Figure 9.1a) to

result, overall change in $\mu_{eff}@V_G$ can be either positive or negative depending on the $\mu_{eff} - E_{eff}$ steepness [347].

guard-band against ΔV_T [58, 59] – so that even the slowest transistor (at the end of product lifetime) has sufficient drive current ($I_{D,min}$) for a functional integrated circuit.

It is possible, however, to increase the negative steepness of $\mu_{eff} - E_{eff}$ relationship by various techniques, *e.g.*, by introducing uniaxial strain (thus reducing phonon-scattering, while keeping surface roughness essentially unchanged) [347], by reducing operating temperature (thus making temperature-independent surface roughness scattering dominant over the temperature-dependent phonon-scattering) [270, 348]. And with suitably high negative steepness for the $\mu_{eff} - E_{eff}$ characteristic, positive $\Delta\mu_{eff}@V_G$ due to positive ΔV_T (and vice versa) can be large enough to balance the two terms in equation (9.1), thereby making the variation-resilient transistor insensitive to ΔV_T (in terms of $I_{D,lin}$). Moreover, since an increase in the negative steepness of $\mu_{eff} - E_{eff}$ or $\mu_{eff} - V_G$ can match the positive steepness of $N_{inv} - V_G$, variation-resilient transistors have flatter $I_{D,lin} - V_G$ characteristics (smaller $g_{m,ON}$) compared to the classical ones (Figure 9.1c). Thus consideration of self-compensation between $\Delta\mu_{eff}$ and ΔN_{inv} is eventually reflected in the flat transfer characteristics above threshold. And it is intuitively clear from Figure 9.1c that with such flat transfer characteristics, $I_{D,lin}$ will not be affected by ΔV_T . Therefore, the presence of self-compensation suggests the possibility of reducing the guard-band voltage, required for $I_{D,min}$.

Next, let us see whether the presence of $I_{D,lin}$ self-compensation ensures corresponding behavior in the saturation regime, as well. As theoretically analyzed in section 8.6 and experimentally characterized in Appendix E, modern transistors operate in quasi-ballistic regime and ΔI_D in linear and saturation regions can be approximated using equation (8.12). Therefore, one expects self-compensation to be more effective for $\Delta I_{D,lin}$ compared to $\Delta I_{D,sat}$.

9.4. Increasing $\mu_{eff} - E_{eff}$ Steepness: Strain

We have now established larger $\mu_{eff} - E_{eff}$ steepness or smaller $g_{m,ON}$ as a requirement for designing self-compensated transistors. Such increase in $\mu_{eff} - E_{eff}$ steepness can be achieved by having a larger increase in μ_{eff} at low E_{eff} compared to the μ_{eff} at high E_{eff} .

Within the picture of universal mobility, μ_{eff} can be expressed as [267] (*c.f.*, equation (6.6)):

$$\mu_{eff}^{-1} = \mu_{coul}^{-1} + \mu_{ph}^{-1} + \mu_{sr}^{-1} \quad (9.2)$$

where, μ_{coul} , μ_{ph} , and μ_{sr} are components of μ_{eff} due to coulomb scattering, phonon scattering and surface roughness scattering, respectively. As shown in [267], μ_{coul} , μ_{ph} are the dominant components at low E_{eff} and μ_{sr} is the dominant component at high E_{eff} . Thus to achieve self-compensation by increasing μ_{eff} - E_{eff} steepness, one needs to improve μ_{coul} , μ_{ph} and keep similar μ_{sr} by increasing uniaxial strain within the channel [347] or by reducing operating temperature [348]. In the following, we qualitatively (section 9.4.1) and quantitatively (section 9.4.2) explain the variation in μ_{eff} - E_{eff} steepness as a function of uniaxial strain, thus identify the extent of uniaxial strain necessary for self-compensation of ΔV_T (section 9.4.3).

9.4.1. Qualitative Theory

As it is widely used for transistor's performance improvement [193], uniaxial strain (rather than biaxial strain) is also preferable for achieving self-compensation. Biaxial tensile strain though shows performance or μ_{eff} improvement for NMOS, compared to that for PMOS [193], it will not be suitable for obtaining self-compensation in NMOS transistors. This is because biaxial tensile strain increases μ_{sr} (as well as μ_{ph}) for NMOS by reducing the roughness amplitude [349, 350]. Thus, biaxial tensile strain in NMOS improves μ_{eff} at all E_{eff} , and hence is not expected to increase the μ_{eff} - E_{eff} steepness by that much. On the other hand, one employs process induced uniaxial strain by source/drain engineering, after the gate insulator growth. Thus, it does not change the Si and gate insulator interface properties and hence results negligible change in surface roughness or μ_{sr} [351]; therefore, uniaxial strain satisfies the first requirement (*i.e.*, reduced increase in μ_{eff} at higher E_{eff}) of self-compensation.

Let us now understand whether the requirements of enhanced μ_{coul} and μ_{ph} (for self-compensation) are also satisfied in uniaxial strain technology. Here, we consider the

consequence of using tensile strain and $\langle 110 \rangle$ compressive strain³¹ for NMOS and PMOS, respectively; because these uniaxial strains improve transistor performance significantly [193]. *Tensile strain* splits the Δ_2 - Δ_4 electron valleys and hence reduces inter-valley phonon scattering, *i.e.*, tensile strain increases μ_{ph} for electron transport. In addition, since Δ_2 constitutes the lower electron valley under tensile strain, there is an increase in electron occupancy in the Δ_2 valley, which also has lower transport effective mass. As such, tensile strain also increases μ_{coul} for electron transport [352]. As a result, self-compensation in NMOS transistor can be achieved through tensile strain. On the other hand, *compressive strain* splits the heavy hole-light hole (HH-LH) valleys and hence reduces inter-valley phonon scattering or increases μ_{ph} for hole transport. In addition, since HH valley has the higher occupancy of holes under compressive strain (which also has lower transport effective mass in the $\langle 110 \rangle$ direction), $\langle 110 \rangle$ compressive strain is also expected to increase μ_{coul} for hole transport. As a result, self-compensation in PMOS transistor can be achieved through compressive strain.

9.4.2. Quantitative Analysis of Self-Compensation

In this section, we experimentally demonstrate the consequence of $\langle 110 \rangle$ compressive uniaxial strain in achieving self-compensation for PMOS transistor. Since the theory of self-compensation is similar for NMOS and PMOS transistors (see section 9.4.1), we expect similar improvement in NMOS transistor, as well. The PMOS transistors under study have $\langle 110 \rangle$ compressive uniaxial strain applied through SiGe

³¹ Band structure calculation in [351] suggests a reduction in μ_{sr} , due to changes in scattering rate and effective mass, for $\langle 110 \rangle$ compressive strain on (100) substrate of PMOS. Consideration of such μ_{sr} reduction with uniaxial strain (compared to our presumption of strain independent μ_{sr}) can further increase the μ_{eff} - E_{eff} steepness for same amount of strain. However, due to the lack of experimental evidence for the uniaxial strain dependence of μ_{sr} (equivalent to the one performed for biaxial strain in [349, 350]), we presume μ_{sr} to be constant in our analysis.

Source/Drain and contact etch stop layer (therefore, a reduction of L increases strain [329]); the transistors also have a doping density of $N_{sub} \sim 3 \times 10^{17} \text{ cm}^{-3}$ and $EOT \sim 1.4 \text{ nm}$. Now, to estimate the effect of uniaxial strain (ε) on $\mu_{eff} - E_{eff}$, we determine μ_{eff} and E_{eff} for unstrained and strained transistors having different channel length (L), using the measurement procedure discussed in sections 6.4.1 and 6.4.2. Then, we calculate the mobility enhancement factor $\mu_{eff,\varepsilon}/\mu_{eff,\varepsilon=0}$ as a function of E_{eff} at different L (or equivalently at different ε).³² Finally, by using $\mu_{eff,\varepsilon}/\mu_{eff,\varepsilon=0}$ and $\mu_{eff,\varepsilon=0}$ for longest L transistor (thus avoiding the extra mechanism for shorter L transistor [353]), we estimate $\mu_{eff}@E_{eff}$ for different amount of uniaxial compressive strain (Figure 9.2). Our measurement of μ_{eff} vs. E_{eff} for unstrained $L = 10 \mu\text{m}$ transistor is consistent with the universal mobility curve [267]; here, the only difference at low E_{eff} is due to presence of higher substrate doping ($N_{sub} \sim 3 \times 10^{17} \text{ cm}^{-3}$) in our transistors. Figure 9.2 also indicates an increase in $\mu_{eff}@E_{eff}$ with strain, as well as an increase in the negative steepness of the $\mu_{eff} - E_{eff}$ characteristics, as expected from the qualitative analysis in section 9.4.1.

³² Here, we could have estimated $\mu_{eff,\varepsilon}/\mu_{eff,\varepsilon=0}$ by measuring μ_{eff} on transistors having different L (*i.e.*, μ_{eff} for longest L transistor will serve as $\mu_{eff,\varepsilon=0}$ and μ_{eff} for shorter L transistors will serve as $\mu_{eff,\varepsilon}$). However, such procedure requires explicit decomposition between the effect of strain and the effect of extra scattering from source/drain halo regions [353] for shorter channel transistors. Therefore, we estimate $\mu_{eff,\varepsilon}/\mu_{eff,\varepsilon=0}$ at a particular ε (or L), by comparing $\mu_{eff}@E_{eff}$ of unstrained and strained transistors having same channel length.

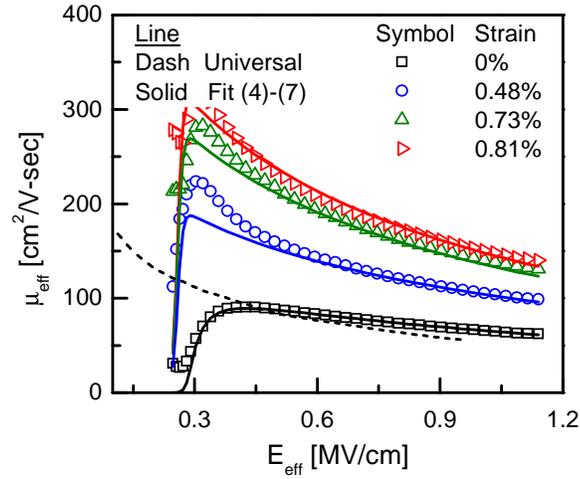


Figure 9.2: Measured $\mu_{eff}E_{eff}$ at different levels of strain (ε) are fitted using equations (9.2)-(9.5). Here, we estimate ε by using Fig. 11 of [193]. The measurement at 0% strain is consistent with the universal mobility measurement (dashed line) of [267].

To capture the strain dependence in a quantitative mobility model, we add two features in the universal mobility relationship. Based on the discussion in section 9.4.1, we know that compressive uniaxial strain increases μ_{coul} by having higher occupancy in the lower transport effective mass HH valley and also increases μ_{ph} due to HH-LH band splitting. To capture this, we use the following expressions for μ_{coul} , μ_{ph} , and μ_{sr} for fitting the $\mu_{eff}E_{eff}$ characteristics in Figure 9.2:

$$\mu_{coul} = \frac{\mu_1 \mu_{1str}}{\left[\frac{N_{Dop}}{10^{18}} + \beta_{IT} \frac{N_{IT}}{10^{12}} \frac{N_{inv}}{10^{12}} \right]} \left(\frac{N_{inv}}{10^{12}} \right)^{\alpha_1} \quad (9.3)$$

$$\mu_{ph} = \mu_2 E_{eff}^{\alpha_2} \left[1 + B \left(\exp \frac{\Delta E_{LH-HH}}{kT} - 1 \right) \right] \quad (9.4)$$

$$\mu_{sr} = \mu_3 E_{eff}^{\alpha_3} \quad (9.5)$$

Note that the phenomenological expression for strain-induced phonon mobility enhancement for μ_{ph} in equation (9.4) was first introduced in [354] and later used by

many groups [355, 356]. Using $\mu_{eff}\text{-}E_{eff}$ characteristics of unstrained transistor, we first estimate the parameters: $\mu_1 = 97.3 \text{ cm}^2/\text{V}\cdot\text{sec}$, $\mu_2 = 82.52 \text{ cm}^2/\text{V}\cdot\text{sec}$, and $\mu_3 = 357.5 \text{ cm}^2/\text{V}\cdot\text{sec}$. The remaining parameters are consistent with literature; *vis.* $\alpha_1 = 1$, $\alpha_2 = -0.32$ are the same as used in [267] and $\alpha_3 = -1.6$ is the same as extracted from the μ_{sr} components reported in [357]. Next, band-structure information of strained transistors are used to estimate the effective energy difference of HH-LH valleys in strained ($\Delta_{H,\varepsilon}$) transistors using single sub-band and Airy-function approximation. Similar to the findings in [187], our calculation of $\Delta_{H,\varepsilon}$ shows an increase in $\Delta_{H,\varepsilon}$ with E_{eff} (see Figure 9.3a). Then, by comparing $\Delta_{H,\varepsilon}$ for unstrained and strained transistors, we calculate $\Delta E_{LH-HH} = \Delta_{H,\varepsilon} - \Delta_{H,\varepsilon=0}$ at different E_{eff} (see Figure 9.3b). Interestingly, ΔE_{LH-HH} at higher E_{eff} has small deviation from its value of $6.426\varepsilon \text{ eV}$ at zero E_{eff} (where, $6.426\varepsilon \text{ eV}$ is the deformation potential [187]). Using the information of ΔE_{LH-HH} from Figure 9.3b, we fit $\mu_{eff}\text{-}E_{eff}$ characteristics (Figure 9.2) and hence estimate $\mu_{1str} = (1+9 \times 10^4 \varepsilon)$ and $B = 0.27$. Finally, for studying the effect of N_{IT} in section 9.8.4, we use $\beta_{IT} \sim 0.04$ for the wavefunction interaction parameter [266], which has been obtained by fitting $\mu_{eff}\text{-}E_{eff}$ characteristics before and after N_{IT} generation (see section 6.4.4), and is assumed to be strain independent (as observed in Figure 3.26c).

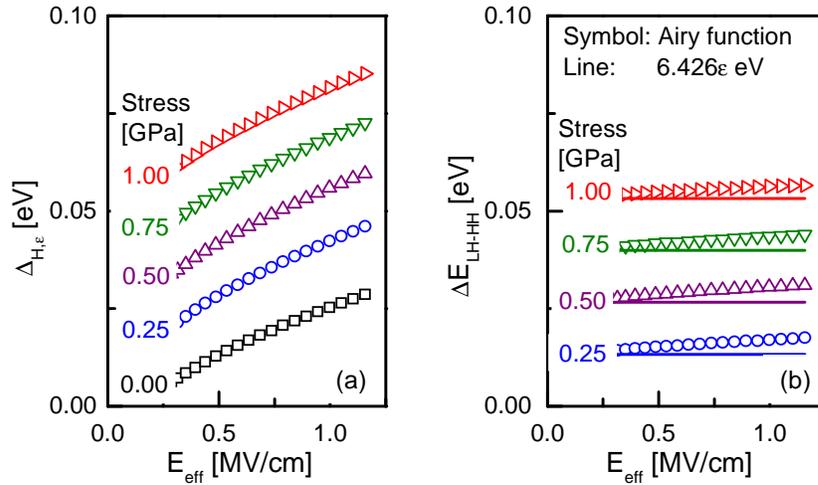


Figure 9.3: (a) Calculation of $\Delta_{H,\epsilon}$ within single sub-band, Airy-function approximation shows an increase in $\Delta_{H,\epsilon}$ with E_{eff} , as also observed in [187]. (b) $\Delta E_{LH-HH} = \Delta_{H,\epsilon} - \Delta_{H,\epsilon=0}$ at different E_{eff} is close to its magnitude of 6.426ϵ eV at zero E_{eff} , where 6.426ϵ eV is the deformation potential.

9.4.3. Self-Compensation at Practical Strain Limit

Using the strain-dependent mobility model of equations (9.2)-(9.5), we estimate μ_{eff} - E_{eff} for different levels of strain (Figure 9.4a), which suggests that for higher strain (*i.e.*, for $\epsilon > 2.5\%$ in Figure 9.4a) mobility enhancement gets limited by the strain independent μ_{sr} component.³³ Next in Figure 9.4b, we calculate $I_{D,lin}$ - V_G for $L = 100\text{nm}$ using μ_{eff} - E_{eff} of Figure 9.4a. Figure 9.4b suggests that with $\epsilon = 2.5\%$ flatness or total self-compensation in $I_{D,lin}$ - V_G characteristics can be achieved at $V_{GS} \sim 1.2\text{-}1.5\text{V}$.³⁴ Thus, we conclude that

³³ The observation of mobility enhancement saturation through uniaxial strain is also consistent with [193].

³⁴ However, for a practical transistor, one needs to consider the contribution from additional scattering mechanisms near source/drain halo regions [353], which may modify this estimate of ϵ in achieving self-compensation.

complete self-compensation ($\Delta I_{D,lin} \sim 0$ at all V_G) may not be possible within practical strain limits. However, as discussed in section 9.8.4, *partial* self-compensation (through positive $\Delta\mu_{eff}$) is indeed a reality in uniaxial strained transistors and circuits.

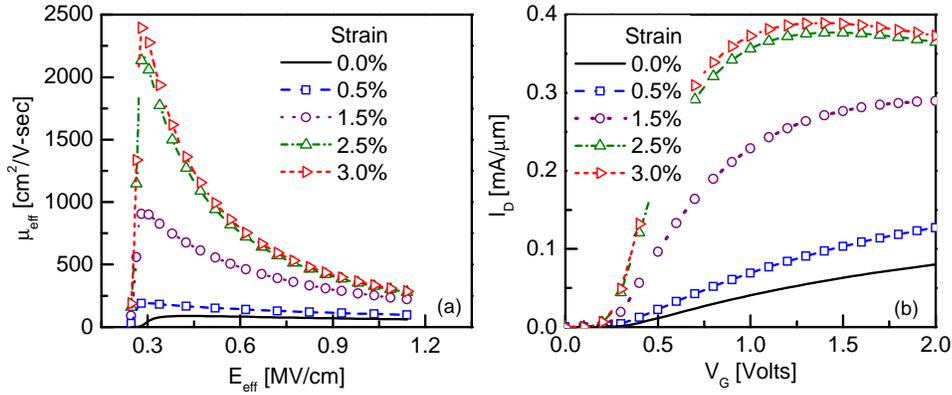


Figure 9.4: (a) μ_{eff} - E_{eff} curves for different levels of strain, estimated using strain-dependent mobility model of equations (9.2)-(9.5). (b) Corresponding $I_{D,lin}$ - V_G characteristics at $V_{DS} = 50\text{mV}$ for a simulated transistor having $L = 100\mu\text{m}$, $EOT = 1.4\text{nm}$, and $N_{sub} = 3 \times 10^{17} \text{cm}^{-3}$.

9.5. Increasing μ_{eff} - E_{eff} Steepness: Temperature

In section 9.4, we illustrate how one can achieve self-compensation by increasing the steepness of $\mu_{eff} - E_{eff}$ characteristics through uniaxial strain. Here, the principle was to reduce the inter-valley phonon scattering and thus make surface roughness scattering (having larger μ - E steepness [267] compared to the phonon component) the dominant component in the $\mu_{eff} - E_{eff}$ characteristics. However, one might presume that the presence of self-compensation is a property of strained technology and may not be applicable, in general. To refute such presumption, in this section, we study self-compensation in one particular strained transistor (the one with highest strain in Figure 8.7c) at different temperature. Since a transistor operated at low temperature has reduced phonon scattering (like the case for higher uniaxial strain), one can have higher $\mu_{eff} - E_{eff}$ steepness at lower temperature. We experimentally observe such an increase in the $\mu_{eff} - E_{eff}$ steepness with

reduction in temperature (see Figure 9.5a).³⁵ And consequently, higher $\mu_{eff} - E_{eff}$ steepness gets reflected in the reduced $\Delta I_{D,lin}$ at lower temperature (see Figure 9.5b).

Therefore, higher $\mu_{eff} - E_{eff}$ steepness or small $g_{m,ON}$ is indeed required for designing a self-compensated transistor. Encouragingly, the recent reports of $I_D - V_G$ characteristics in Intel's 45-nm strained MOS technology [47] and III-V transistors [358] indicate the presence of higher $\mu_{eff} - E_{eff}$ steepness or small $g_{m,ON}$. Thus, advanced substrates involving strained-Si and III-V show the possibility of self-compensating ΔV_T with reduced use of guard-band or any complicated circuit-level optimization.

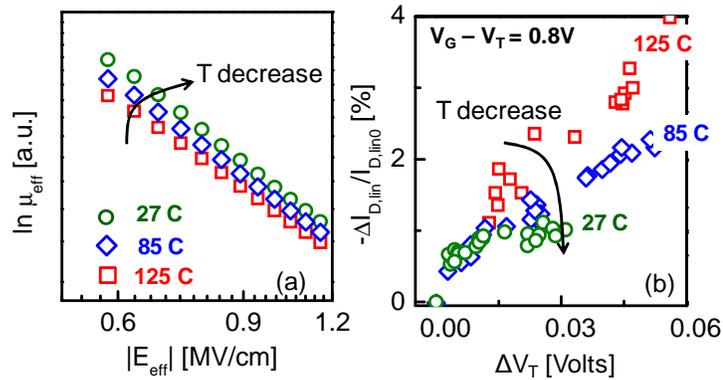


Figure 9.5: (a) Reduction in temperature increases the $\mu_{eff} - E_{eff}$ steepness for the highly-strained PMOS transistor of Figure 8.7c. (b) The transistor, operated at reduced temperature, is less sensitive to N_{IT} -induced $\Delta I_{D,lin}$ at $|V_G - V_{T0}| = 0.8V$.

9.6. Sources of V_T Fluctuation: Show-stopper for CMOS Scaling?

From our discussions in sections 9.4 and 9.5, we know that it is possible to achieve partial self-compensation through uniaxial strain. In this section, using the following

³⁵ Similar increase in steepness is already shown for both PMOS and NMOS transistors in [270].

expression for V_T in bulk CMOS [282], we identify the sources of ΔV_T that one needs to self-compensate:

$$V_T = \Phi_{GS} + \frac{EOT}{\epsilon_{SiO_2}} \sqrt{4k_B T \epsilon_{Si} N_{sub} \ln \frac{N_{sub}}{n_i}} + \frac{2k_B T}{q} \ln \frac{N_{sub}}{n_i} + q \frac{N_T(t)}{C_{di}}. \quad (9.6)$$

In equation (9.6), $N_T(t)$ indicates the time-dependent contribution from defects. These defects can be present within the oxide (oxide defects N_{OT} ; see chapter 5), which are observed during BTI stress on PMOS/NMOS transistors [43]. Defects can also be present at the oxide/substrate interface (interface defects N_{IT} ; see chapter 3) are observed during NBTI stress on PMOS transistor [119]. Therefore, equation (9.6) indicates that ΔV_T can arise from fluctuation in process parameters like Φ_{GS} , EOT , N_{sub} or in reliability parameters like $N_T(t)$. As shown in [48, 359], process-related fluctuation in Intel's 45 nm technology corresponds to a 3σ of ~ 150 mV, whereas temporal fluctuation has a mean shift of ~ 60 mV with a 3σ of ~ 30 mV.

In sum, during transistor design, which considers no self-compensation [346, 360] (just like the case for a classical transistor of Figure 9.1a), one needs to handle a $\Delta V_{T,max} \sim 240$ mV (above V_{T0}) either by guard-banding the transistor (or by operating it at a voltage higher than V_I in Figure 9.1a) [58, 59] or using complex algorithms [60-64]. The use of guard-band voltage ($V_{DD} - V_I$ in Figure 9.1a) limits the scalability of V_{DD} in nanoscale transistors, which is ~ 1 V since 65 nm CMOS technology [7]. Such un-scalability of V_{DD} is considered as the main reason for continuous increase in CMOS power consumption, which is expressed as –

$$P \sim f C_L V_{DD}^2 + V_{DD} I_{Leakage}, \quad (9.7)$$

where f is the operating frequency, C_L is the load capacitance, and $I_{Leakage}$ is the leakage current at V_{DD} . As transistor scaling induces an increase in f [7], an increase in $I_{Leakage}$ (which can be suppressed by using high- κ dielectrics [39, 361]), and negligible variation in C_L (see Table 9.1 for Intel's recent CMOS technology values), un-scalability of V_{DD} has been identified as the main show-stopper for future scaling [362-365]. This initiated research on emerging technologies [366] involving spinFETs [367-369], magnetically

coupled spin-torque devices [370], magnetic quantum cellular automata [371], *etc.* However, we will show in the next few sections that there are still some promises within CMOS technology in terms of V_{DD} scaling. And indeed a simple idea of variation resilience (even to some partial extent) might help reduce the guard-band voltage, thus may allow a reduction in CMOS power consumption.

Table 9.1: Capacitance values in Intel’s recent CMOS technologies.

| Intel’s Technology Node | RO Delay per stage @ V_{DD} [ps] | Ref. | Output capacitance (each RO stage) C_L [fF/ μm] | Intrinsic output capacitance, $C_{intrinsic}$ [fF/ μm] | $C_{intrinsic}/C_L$ [%] |
|-------------------------|------------------------------------|-------|---|--|-------------------------|
| 180 | 10.6 (1.5V) | [372] | 1.56 | 1.15 | 73.7 |
| 130 | 6.2 (1.4V) | [373] | 1.66 | 0.94 | 56.7 |
| 90 | 5.5 (1.2V) | [374] | 2.00 | 0.78 | 39 |
| 65 | 4.25 (1.0V) | [375] | 1.9 | 0.6 | 31.6 |
| 45 | 2.7 (1.0V) | [361] | 1.62 | 0.86 | 53.4 |

9.7. Self-Compensation in $I_{D,lin}$ Fluctuation

Given the principle of $I_{D,lin}$ self-compensation discussed in section 9.3, let us explore in this section how an appropriately designed variation-resilient transistor might lead to self-compensation in $I_{D,lin}$ fluctuation against different sources of ΔV_T .

9.7.1. EOT/ Φ_{GS} Variation

As discussed in chapter 1, use of thin oxides and metal gate in current CMOS technology raises concerns regarding I_D fluctuation due to variations in EOT [11] and Φ_{GS} [15]. In this section, we explore how an appropriately designed variation-resilient transistor (through uniaxial strain) might lead to $I_{D,lin}$ self-compensation against process

fluctuations related to Φ_{GS} and EOT . Our approach involves the calculation of V_T using equation (9.6) and V_G , E_{eff} , $I_{D,lin}$ using the following equations [282]:

$$V_G = \Phi_{GS} + \left| \sqrt{2\epsilon_{Si}k_B T N_{sub}} \left[\frac{q\psi_S}{k_B T} + \frac{n_i^2}{N_{sub}^2} \left(e^{q\psi_S/k_B T} - \frac{q\psi_S}{k_B T} - 1 \right) \right] \right| / C_{di} + \psi_S, \quad (9.8)$$

$$E_{eff} = \frac{1}{\epsilon_{Si}} \left[(1-\eta) \sqrt{4k_B T \epsilon_{Si} N_{sub} \ln \frac{N_{sub}}{n_i}} - \eta C_{di} \frac{2k_B T}{q} \ln \frac{N_{sub}}{n_i} + \eta \frac{\epsilon_{SiO_2}}{EOT} (V_G - \Phi_{MS}) \right], \quad (9.9)$$

$$I_{D,lin} = \mu_{eff} C_{di} \frac{W}{L} V_{DS} \frac{\frac{2mkT}{q} \ln \left[1 + \exp \frac{q(V_G - V_T)}{2mkT} \right]}{1 + \frac{2m}{m-1} \exp \frac{q(V_T - V_G)}{2mkT}}; \quad (9.10)$$

where the symbols have their standard definitions, as in [282]. Note that equation (9.10) leads to the usual expressions for $I_{D,lin}$ in sub-threshold (for $V_G \ll V_T$) and super-threshold (for $V_G \gg V_T$) regions. Using equations (9.2)-(9.10), we simulate the transfer characteristics for a PMOS transistor. Thus, we observe that an unstrained transistor (having $L = 130\text{nm}$, $EOT \sim 2.0\text{nm}$) is prone to $\pm 0.2\text{nm}$ EOT variation (Figure 9.6b) and $\pm 10\%$ Φ_{GS} variation (Figure 9.6d). However, increase in strain to 2.5% results in smaller $g_{m,ON}$ near $V_{GS} \sim 1.2-1.5\text{V}$ (Figure 9.6c, Figure 9.6e); therefore, ΔV_T gives rise to negligible $\Delta I_{D,lin}$ in that voltage range, as explained earlier.

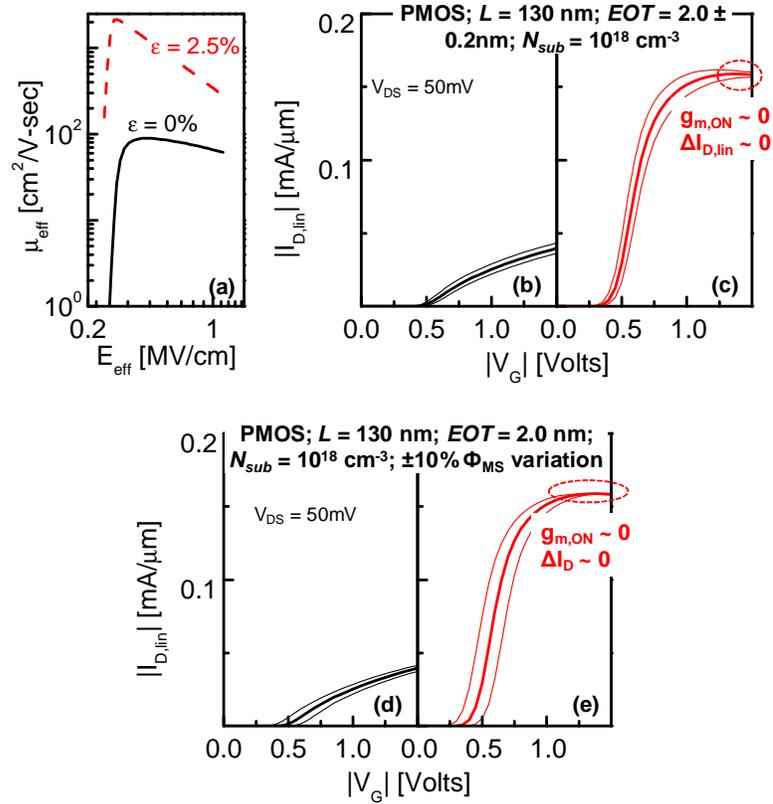


Figure 9.6: (a) Comparison of $\mu_{eff} - E_{eff}$ steepness for unstrained ($\epsilon = 0\%$) and strained ($\epsilon = 2.5\%$) transistors. (b, d) Unstrained transistor having less $\mu_{eff} - E_{eff}$ steepness suffers from EOT and Φ_{GS} variations. Whereas, (c, e) strained transistor with higher $\mu_{eff} - E_{eff}$ steepness or small $g_{m,ON}$ is less sensitive to EOT and Φ_{GS} variation at operating V_G .

9.7.2. Random Dopant Variation

Another major source of variation in CMOS technology is the dopant fluctuation. As the devices are scaled down following Moore's law, the volume of the device structure has become so small that total number of dopants inside the channel, *i.e.*, $N_{D,total}$ is only about ~ 100 or so [8]. In such small structures, statistical fluctuation of dopants ($\sim \sqrt{N_{D,total}}$) among the transistors on the same wafer is significant, giving rise to the phenomena commonly known as Random Dopant Fluctuation (RDF). Resultant fluctuation in V_T (according to equation (9.6)) due to RDF gives rise to fluctuation in I_D (according to

equation (9.1)), as well. It would be interesting to study whether the variation resilience concept, presented in section 9.3, works for RDF as well.

Unfortunately, the concept is not at all applicable for RDF compensation. Our simulations suggest that $I_{D,lin}$ remains sensitive to RDF, irrespective of the $\mu_{eff} - E_{eff}$ steepness (see Figure 9.7). To understand this result, consider the expression for E_{eff} in equation (9.9). Variation in N_{sub} mainly effects the first term in the right-hand side of equation (9.9). Therefore, any increase (decrease) in N_{sub} , over and above the mean value, raises (lowers) both V_T according to equation (9.6), as well as E_{eff} according to equation (9.9). Therefore, dopant fluctuation decreases (increases) $\mu_{eff}@V_G$ with an increase (decrease) in V_T – and as such can not satisfy the requirement for variation resilience based on equation (9.1). However, the proposed approach remains relevant as the technology evolves towards fully-depleted silicon-on-insulator with no issues related to RDF. Similarly, higher steepness in $\mu_{eff} - E_{eff}$ characteristics can not compensate the effects of channel length fluctuation (LER) in uniaxial strained transistor. This is because, changes in L above (below) the mean value decreases (increases) strain and μ_{eff} , as such causes $I_{D,lin}$ to decrease (increase), according to equation (9.10), irrespective of $\mu_{eff} - E_{eff}$ steepness.

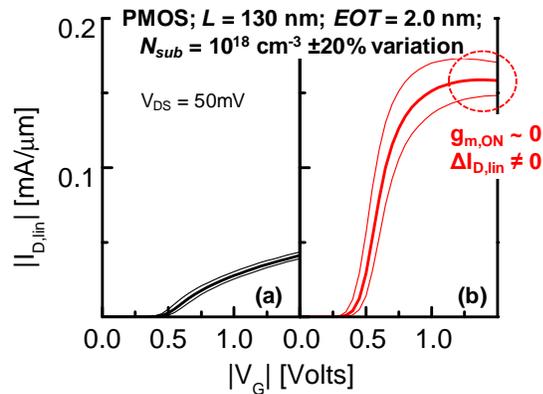


Figure 9.7: (a) Unstrained transistor suffers from $\pm 20\%$ N_{sub} variation. (b) Though $g_{m,ON} \sim 0$ at $V_{GS} \sim 1.2\text{-}1.5 \text{ V}$ for 2.5% strained transistor, dopant fluctuation remains uncompensated around that range of V_{GS} .

9.7.3. Defect-induced ΔV_T

Similar to the process-related V_T fluctuations, we can also compensate N_{OT} or N_{IT} - related time-dependent V_T fluctuations by increasing the $\mu_{eff} - E_{eff}$ steepness [347, 348, 376]. As discussed in sections 8.2, 8.4.2, and 8.5.3, N_{OT} -induced ΔV_T will reduce $E_{eff}@V_G$, hence improve $\mu_{eff}@V_G$ and can partially compensate ΔV_T , according to equation (8.12). Similarly, N_{IT} -induced ΔV_T will also reduce $E_{eff}@V_G$ and consequently increase $\mu_{eff}@V_G$ (see sections 8.2, 8.4.1, and 8.5). However, contrary to the case for N_{OT} , N_{IT} introduces additional coulomb scattering centers [267] and hence reduces $\mu_{eff}@E_{eff}$. As a result, $\Delta\mu_{eff}@V_G$ due to N_{IT} -induced ΔV_T is not always positive (as the case for N_{OT} -induced ΔV_T) and depends on the $\mu_{eff} - E_{eff}$ steepness (Figure 8.7 and Figure 8.8). Only for the highly strained transistors having larger $\mu_{eff} - E_{eff}$ steepness, μ_{eff} improvement through E_{eff} reduction dominates over the downward shift of the mobility-field curve (*i.e.*, negative $\Delta\mu_{eff}@E_{eff}$) – resulting positive $\Delta\mu_{eff}@V_G$ for N_{IT} -induced ΔV_T . Consequently, according to equation (8.12), positive $\Delta\mu_{eff}/\mu_{eff,0}@V_G$ for the highly-strained transistors can easily compensate the negative contribution from $-\Delta V_T/(V_G - V_{T0})$, resulting negligible $\Delta I_{D,lin}$ for the time-dependent degradation related to N_{IT} .

9.7.4. 1/f Noise

Interfacial and oxide defects not only create shift in V_T , spatial distribution of oxide defects can also lead to 1/f noise signal, which is routinely observed in transistors involving high- κ dielectric [377, 378]. Figure 9.8 illustrates the origin of 1/f noise in high- κ dielectric, as presented in [377].

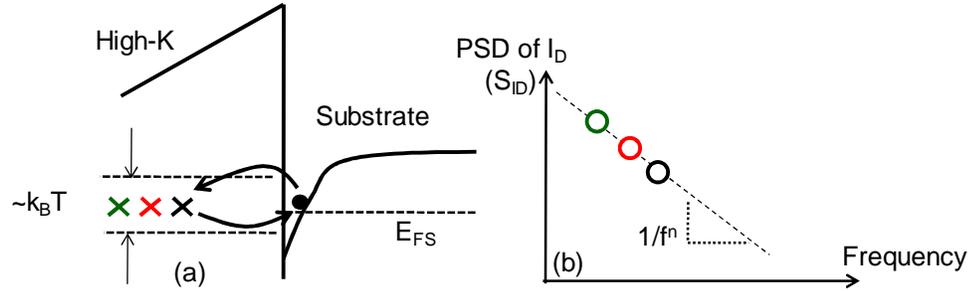


Figure 9.8: (a) Trapping/detrapping of channel carriers with $\sim k_B T$ energy range around the Fermi-level gives rise to I_D noise having distinct frequency. (b) Noise from defects at different distances away from the interface leads to $1/f^n$ signal.

The model in Figure 9.8 assumes that there is a distribution of defects within the high- κ dielectric. The channel carriers can tunnel into and out of a defect at a particular frequency (f), which is determined by the distance of that defect from the interface. Obviously, defects at a distance far away from the interface will have lower carrier tunneling rate (or low f) and vice versa. Moreover, trapping/detrapping of carriers requires both a filled state (corresponding Fermi function, F) and empty state (corresponding Fermi function, $1-F$) within the channel. Hence, tunneling into/out of defects will happen within an energy range $\sim k_B T$ around the substrate Fermi-level (where, $F(1-F)$ is maximum).

Now, due to trapping of carriers into a defect, there will be a reduction in $Q_{inv} \sim V_G - V_T$, as well as a change in μ_{eff} . As such, each defect will give rise to a reduction in I_D (due to trapping of carriers into the defect) and a current-reversal (due to subsequent detrapping of carriers back into the channel). Such decrease and increase in I_D will give rise to a distinct-frequency noise in I_D , where the noise frequency will depend on the distance of that defect from the interface (*e.g.*, defects that are far away from the interface will have lower f). Summing the response from all the defects within the oxide normally shows a $1/f^n$ trend (Figure 9.8), which is commonly termed as “1/f noise” in semiconductor literature. Now, following the analysis similar to the one presented in [377], power spectral density of drain current noise (S_{I_D}) for a defect having a trapping/detrapping frequency of f can be expressed as:

$$S_{I_D} \Big|_{f, V_G} \sim |\Delta I_D @ f|^2 k_B T \sim [I_D @ f, V_G]^2 \left[\frac{\Delta \mu_{eff}}{\mu_{eff0}} \Big|_{f, V_G} - \frac{|\Delta V_T(f)|}{V_G - V_{T0}} \right]^2 k_B T. \quad (9.11)$$

Since equation (9.11) has similarity with equation (8.6), it is expected that the degradation-free transistor [133, 347], which has negligible drain current fluctuation from electron/hole trapping, will also have resilience to 1/f noise.

9.8. Self-Compensation: From Transistor to Digital Electronics

So far, we have explored the implication of self-compensation on individual transistor parameters and demonstrated that (partial) self-compensation can be routinely achieved under various conditions in modern CMOS transistors. However, we have shown in sections 8.5 and 8.6 that self-compensation is not equally effective at all bias conditions. Indeed, during circuit operation, a transistor goes through a variety of bias conditions, therefore the obvious question is: ‘Without perfect self-compensation at all bias conditions, will there be any significant effect of this phenomena in actual circuit operation?’. In this section, we answer this question by studying the consequence of self-compensation in simple logic and memory circuits. First, we use measured I-V characteristics before and after N_{IT} generation to obtain performance response of two basic circuits, *i.e.*, single-stage inverter (section 9.8.1), and SRAM (section 9.8.2). Thus, we show how self-compensation improves the circuit performance significantly. Next, we move onto the study of self-compensation in larger circuits through SPICE simulation and explain why the current approach of SPICE simulation gives *unphysical* ΔE_{eff} due to ΔV_T and hence fails to capture the effect of self-compensation (section 9.8.3). Finally, by considering appropriate changes in E_{eff} due to ΔV_T , we illustrate how self-compensation can reduce N_{IT} -induced delay degradation for NAND and ring oscillator circuits (section 9.8.4), once physical change in E_{eff} is taken into account. Though we restrict our analysis to time-dependent N_{IT} variations, similar analysis should be valid for other sources of ΔV_T variation, as well.

9.8.1. Effect of N_{IT} on Single-Stage Inverter

Let us consider the effect of N_{IT} self-compensation in an inverter driving a constant load $C_L \sim 6\text{fF}$ (Figure 9.9a). Since C_L is typically dominated by a comparatively large interconnect capacitance (over transistor's intrinsic capacitance), delay in other digital circuits (*e.g.*, ring oscillator, combinatorial gates, *etc.*) would follow similar trends. For the inverter under study, only the PMOS transistor is subjected to NBTI-induced $\Delta V_T \sim 70\text{mV}$ due to N_{IT} generation. On the other hand, the NMOS transistor has matched I-V characteristics (with PMOS) in pre-NBTI stress condition. Later, for a particular input voltage V_{IN} (with frequency =1GHz and rise/fall time =1ps), we calculate the output voltage V_{OUT} transient using –

$$\frac{dV_{out}}{dt} = \frac{I_P(V_{in}, V_{out}) - I_N(V_{in}, V_{out})}{C_{out}(V_{out})}, \quad (9.12)$$

where I_P and I_N are currents through the PMOS and NMOS transistors of the inverter, respectively, as shown in Figure 9.9a. After simulating the V_{OUT} transients, we compare V_{IN} with V_{OUT} and estimate the rise (τ_R) and fall (τ_F) delay of the inverter (see Figure 9.9b for a schematic). Since we only consider NBTI degradation, the fall delay remains (approximately) unaffected due to stress. On the other hand, the rise delay for a self-compensated transistor (having $\Delta I_{D,lin} \sim 0$ for $|V_G - V_{T0}| > 1.1\text{V}$) increases by $\sim 5\%$ (*i.e.*, $\Delta\tau_R \sim 5\%$), which is approximately half of the rise delay degradation for a non-compensated transistor (having $\Delta\tau_R \sim 11\%$ for similar NBTI-induced ΔV_T). Thus, considering that delay degradation due to NBTI for a circuit has similar time-exponent ($n \sim 1/6$) as an individual transistor [64], even the presence of partial self-compensation may provide ~ 64 times increase in NBTI lifetime (Figure 9.9c) – possibly eliminating the NBTI degradation as a CMOS reliability concern!

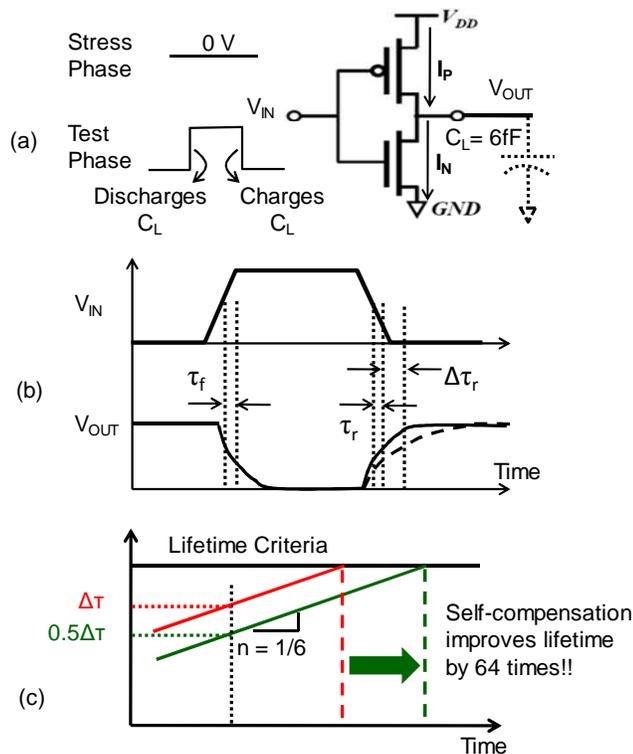


Figure 9.9: (a) Configuration of the inverter used in section 9.8.1. In stress phase ($V_{IN} = 0$), there is NBTI degradation in PMOS transistor. (b) The impact of NBTI degradation (*i.e.*, change in rise τ_r and fall delay τ_f) is monitored by applying a test pulse as V_{IN} and calculating V_{OUT} using equation (9.12). (c) Self-compensation reduces the rise delay degradation by 50% – thus can increase NBTI lifetime by ~64 times (assuming time exponent for delay degradation, $n \sim 1/6$, as obtained in [64]).

9.8.2. Effect of N_{IT} on SRAM performance

Next, we use a conventional 6T SRAM cell (Figure 9.10a) for explaining the effect of BTI degradation on SRAM cells. One common feature of NBTI degradation in SRAM cell is the change in static noise margin (SNM, defined in Figure 9.10b) [379-382]. For example, for a particular bit storage, if $V_L = '0'$ and $V_R = '1'$, then transistor PR will be under NBTI degradation and transistor NL will be under PBTI degradation. As a result, $|V_T|$ for PR (due to NBTI) will increase and, therefore, shift the PR-NR of Figure 9.10b towards left, thus reduces SNM_1 and increases SNM_2 . Thus, the storage condition of $V_L =$

‘0’ and $V_R = ‘1’$, always reduces SNM_1 and increases SNM_2 ; whereas, the opposite happens for the storage condition of $V_L = ‘1’$ and $V_R = ‘0’$.³⁶

Now to understand the effect of N_{IT} self-compensation, we compare the SNM variation for the two cases: in one case we use a self-compensated transistor for PR having $\Delta I_{D,lin}$ and $\Delta I_{D,sat}$ of Figure 8.7c and Figure 8.8b, respectively, and in the second case we use a non-compensated transistor for PR. We observe that self-compensation reduces SNM_1 degradation from 20mV (or 2.75% for the non-compensated transistor) to 7.2mV (or 2% for the compensated transistor) and increases SNM_2 improvement from 19.5mV (or 2.68% for the non-compensated transistor) to 23.7mV (or 6.64% for the compensated transistor). To explain the reduction of SNM_1 (the limiting factor for SRAM performance [379]), let us consider the voltage transfer characteristics of PR-NR (Figure 9.10c) that follows different points over the I_D - V_D characteristics of both NMOS and PMOS of an inverter (Figure 9.10d). Since N_{IT} self-compensation reduces $\Delta I_{D,lin}$ more effectively than $\Delta I_{D,sat}$, regions 1-3 of Figure 9.10c-d will be less effected by N_{IT} in a self-compensated transistor – leading to reduced amount of SNM_1 degradation.

However, SNM is not the only performance criteria for the operation of an SRAM cell. Read delay of an SRAM cell (studied using $\text{BL}_B = \text{BL} = ‘1’$) depends on the strength of NMOS access transistors and thus it is unaffected by NBTI [381], but should be affected by PBTI. On the other hand, NBTI (PBTI) weakens the PMOS (NMOS) compared to NMOS (PMOS) and thus should cause a decrease (increase) in write delay. Similar to the improvement in delay degradation for logic circuits (see section 9.8.1), self-compensation is again expected to reduce the read and write delay variations in SRAM, as well.

³⁶ Actual signal probabilities of V_L and V_R in practical memory operation are far more complicated and should be subjected to a statistical evaluation [379].

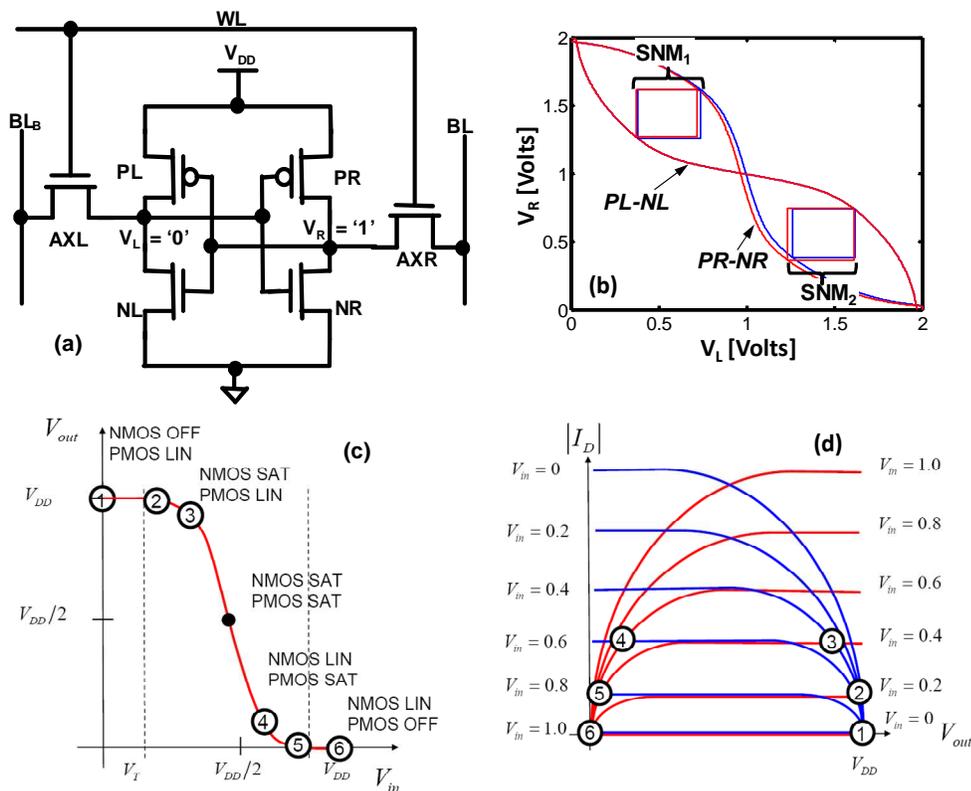


Figure 9.10: (a) Schematic of a 6T-SRAM Cell. PL and PR are PMOS transistors within the storage framework; whereas NL and NR are corresponding NMOS transistors. AXL and AXR are NMOS transistors used for accessing (reading or writing) stored data. (b) Static Noise Margin (SNM) calculation using voltage-transfer characteristics of back-to-back inverters (blue lines: before N_{IT} generation; red lines: after N_{IT} generation). (c) Schematic of voltage transfer characteristics for PR-NR inverter. (d) Schematic of I_D - V_D characteristics for the NMOS/PMOS transistors of PR-NR inverter. Similar modes of operation for NMOS/PMOS transistors are also marked in Fig. c,d. Fig. a is re-drawn from [379] and Fig. c-d is taken from the Purdue's ECE 612 lecture (available at [383]).

9.8.3. Large-scale SPICE analysis: Issues with Unphysical ΔE_{eff}

In the last two sections, we have analyzed the effect of self-compensation on circuit performance using the measured I-V characteristics of a transistor before and after N_{IT} generation and then calculating the V_{IN} - V_{OUT} relationship through equation (9.12). However, for large scale circuit, a detailed SPICE simulation would be necessary. In

current SPICE-based circuit analysis for N_{IT} , $\Delta\mu_{eff}$ is calculated either empirically by adding an extra fitting parameter in ΔV_T estimation [346] or analytically by using the following expression [360] –

$$\mu_{eff} = \frac{\mu_0}{(1 + \alpha_{eff} N_{IT}) f(V_G, V_T)}, \quad (9.13)$$

where α_{eff} is a fitting parameter obtained from mobility degradation experiments performed at constant E_{eff} [384], and based on the mobility model commonly used in BSIM/SPICE analysis,

$$f(V_G, V_T) = 1 + \theta_1 \left(\frac{V_{gsteff} + 2V_T}{T_{ox}} \right) + \theta_2 \left(\frac{V_{gsteff} + 2V_T}{T_{ox}} \right)^2, \quad (9.14)$$

where V_{gsteff} is the gate overdrive (which is $\sim V_G - V_T$ in strong inversion or when $V_G \gg V_T$), and θ_1, θ_2 are fitting parameters. In fact, $(V_{gsteff} + 2V_T)/T_{ox}$ in the right hand side of equation (9.14) effectively serves the purpose of E_{eff} in the BSIM/SPICE model, where V_{gsteff} and $2V_T$ terms reflect the effect of variation in Q_{inv} and Q_{dep} , respectively.

So according to equations (9.13)-(9.14), an increase in N_{IT} not only reduces μ_{eff} through the $\mu_0/(1 + \alpha_{eff} N_{IT})$ factor, but also reduces μ_{eff} through an increase in the $(V_{gsteff} + 2V_T)/T_{ox}$ factor of equation (9.14). Therefore, $\mu_{eff}@V_G$ in the existing SPICE analysis always decreases with N_{IT} , irrespective of the $\mu_{eff}-E_{eff}$ steepness, which is inconsistent with our experimental observations in section 8.4. This discrepancy suggests that the variation of E_{eff} due to N_{IT} through the $(V_{gsteff} + 2V_T)/T_{ox}$ factor of equation (9.14) is completely unphysical in the existing SPICE model.³⁷ And, as shown in the following section, consideration of physical ΔE_{eff} has significant consequences for self-compensation, especially for strained technology with steep $\mu_{eff}-E_{eff}$ characteristics.

³⁷ In fact, $(V_{gsteff} + 2V_T)/T_{ox}$ term of equation (9.14) should fail to reflect the change in E_{eff} due to almost all sources of threshold variation (except RDF and LER), where only Q_{inv} or V_{gsteff} changes with ΔV_T .

9.8.4. Circuit Analysis: Using Physics-based ΔE_{eff}

Our experimentally calibrated physical $\mu_{eff} - E_{eff}$ model, *i.e.* equations (9.2)-(9.5), enables us to simulate I-V characteristics before and after N_{IT} generation, considering proper change in transistor parameters like V_T , μ_{eff} , E_{eff} , *etc.* To perform these I-V simulations as a function of V_G and V_D , we use the equations similar to the one used in PETE [385, 386]. The transistor parameters used in the I-V simulation are: $EOT = 1.2\text{nm}$, $N_{Dop} = 10^{17} \text{ cm}^{-3}$, $T = 300 \text{ }^0\text{K}$, $L = 45\text{nm}$, $W = 1\mu\text{m}$, $|V_{T0}| = 0.4\text{V}$ (defined at $I_{D,lin} \sim 10^{-6} \text{ A}/\mu\text{m}$), $I_{OFF} = 1\text{nA}/\mu\text{m}$, $\lambda = 0.1$, $SS = 106\text{mV}/\text{dec}$, $DIBL = 107\text{mV}/\text{V}$. In addition, we also use the classical SPICE-mobility model, *i.e.* equations (9.13)-(9.14), to perform another set of I-V calculations using PETE. Figure 9.11a,b clearly show the difference between the classical SPICE model (based on equations (9.13)-(9.14)) and the physical model (based on equations (9.2)-(9.5)) in predicting the change of I_D - V_G (at $V_{DS} = -0.1\text{V}$) due to N_{IT} -induced ΔV_T . Mobility model in classical SPICE simulation [360] estimates an unphysical increase in E_{eff} due to N_{IT} generation and hence degrades $\mu_{eff}@V_G$ or $I_D@V_G$ significantly, especially for 2.5% strained transistor (Figure 9.11b) that has more μ_{eff} - E_{eff} steepness. On the other hand, the physical mobility model considers the effect of partial self-compensation due to N_{IT} -induced E_{eff} reduction, and hence predicts significantly less I-V degradation for 2.5% strained transistor, compared to the one predicted from SPICE-like simulation.

Next, we study the effect of N_{IT} -induced ΔV_T at $V_{DD} = 1\text{V}$ for (i) NAND gate driving an INV gate and (ii) 5-stage ring oscillator RO. These circuit configurations are further elaborated in [385]. Our calculation of delay degradation due to N_{IT} -induced ΔV_T (Figure 9.11c,d) demonstrates the importance of using correct sign for ΔE_{eff} in the circuit analysis, especially for the strained transistor (see Figure 9.11d). While unphysical variation of E_{eff} due to N_{IT} in classical SPICE-mobility analysis predicts significant delay degradation with increase in strain, a physical mobility analysis predicts much less delay degradation for the strained transistor. More importantly, although the $I_{D,lin}$ - V_G characteristics of 2.5% strained transistor has no flatness up to 1V (Figure 9.11b), presence of partial self-compensation in strained transistor (through positive $\Delta\mu_{eff}$ in

equation (9.1)) reduces delay degradation by $\sim 15\%$ at $\Delta V_T = 30\text{mV}$ (e.g., changes NAND delay degradation from 4.8% for the unstrained to 4.07% for the strained transistor).

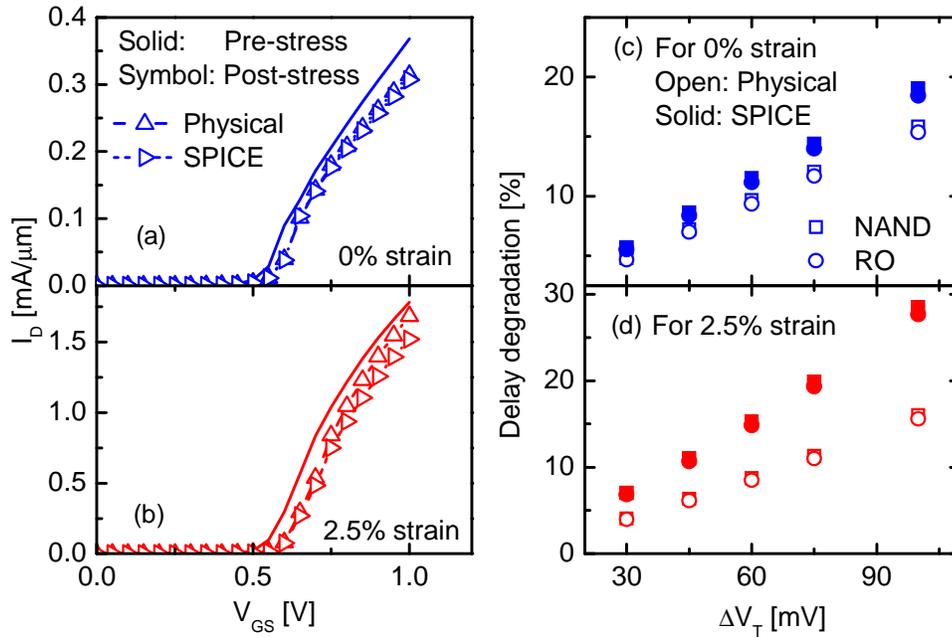


Figure 9.11: Simulated I_D - V_G characteristics at $|V_{DS}| = 0.1\text{V}$ for (a) unstrained and (b) 2.5% uniaxial strained transistor. Effect of N_{IT} -induced ΔV_T in NAND gate and 5-stage RO circuits for (c) unstrained and (d) 2.5% uniaxial strained transistor. The difference between classical SPICE simulation based on equations (9.13)-(9.14) (solid symbols) and circuit simulation based on equations (9.2)-(9.5) (open symbols) is clearly evident in strained CMOS transistor.

Thus, our analysis with uniaxial strained transistor suggests the importance of physical mobility modeling in transistor/circuit level variability analysis. Although total self-compensation (zero delay degradation) is not achieved for $V_{DD} = 1\text{V}$ at the practical limits of uniaxial strain (Figure 9.11), the reduction of delay degradation with strain (albeit from 4.8% to 4.07%) suggests the importance of μ_{eff} - E_{eff} steepness in reducing the effect of CMOS variability. The recent reports of I_D - V_G characteristics in III-V transistors

[358] indicate the presence of large $\mu_{eff}-E_{eff}$ steepness. Therefore, these advanced CMOS technologies may reduce the delay-degradation further, and hence result in considerable increase in IC lifetime (Figure 9.9c), only if ΔE_{eff} due to ΔV_T is correctly taken into account.

9.9. Summary

We have demonstrated how the concept of variation resilience in existing CMOS architecture is related to the negative steepness of the $\mu_{eff} - E_{eff}$ relationship, or equivalently to the presence of smaller $g_{m,ON}$. *Here, one needs to note that by reducing $g_{m,ON}$, we will not lose the current drive capability of a transistor, which mainly depends on $g_{m,max}$.* Such variation resilience or smaller $g_{m,ON}$ is expected to make transistors less sensitive to various sources of V_T variations, including those related to fluctuations in interfacial/oxide defects, oxide thickness, metal work-function, *etc.* Broadly speaking, the increase in $\mu_{eff} - E_{eff}$ steepness can compensate any V_T fluctuations that arise due to parametric variations within the oxide and gate regions, as well as in the oxide/substrate interface. In addition, we have studied the extent of variation resilience in uniaxial strained CMOS transistors, and have shown presence of partial self-compensation in these transistors. This study has also identified the importance of physical mobility modeling in circuit analysis for advanced CMOS technology, by considering appropriate changes in transistor parameters (especially effective electric field) due to threshold voltage variation. Therefore, it is predicted that the effect of self-compensation will be observable in circuit/system level by further increase in the $\mu_{eff} - E_{eff}$ steepness (for example, using III-V materials as substrate), provided the circuit analysis reflects physical changes in transistor parameters. This might substantially reduce the burden on guard-band voltage, expensive design algorithms and/or extra circuitry – that are currently being used for mitigating the effect from V_T variations.

MOS technology has so far evolved to maximize transistor's performance by introducing new concepts/materials. In this chapter, we speculate that the new substrate materials (*e.g.*, strained-Si, III-V) currently being studied for performance improvement,

can also provide better resilience against the effects of CMOS variability. However, only a detailed optimization study by varying the steepness of $\mu_{eff} - E_{eff}$ relationship over a wide range can answer “how much steepness we will need to design an ultimate variation-resilient transistor, totally prone to different sources of V_T fluctuation”.

10. THESIS SUMMARY AND FUTURE WORK

10.1. Thesis Summary

This thesis has been focused in studying one form of temporal fluctuation in MOS transistor, namely Negative Bias Temperature Instability, which is one of the major MOS reliability concerns now-a-days. The detailed modeling and characterization effort, presented here, clarified the physical basis of a phenomenological modeling (*i.e.*, Reaction-Diffusion theory) and also identified many subtle issues for proper characterization of defect-related temporal variability in MOS transistors. We have analyzed the impact of defect on transistor's parametric degradation and thus, proposed the design for a degradation-free transistor. And finally by utilizing the analogy between temporal and spatial variability in MOS transistors, we have shown that a degradation-free transistor can also lead to the design of a variation-resilient transistor, robust against many sources of spatial variability of current MOS technology. The following contains a brief description of the outcome of this thesis, along with the reference to associated publications.

10.1.1. Reaction-Diffusion Modeling of Interface Defect

In chapter 3 and 4 of this thesis, we have developed a physical basis for interface defect generation and its statistics using Reaction-Diffusion theory. We have proposed a self-consistent Si-H bond dissociation model for explaining the field dependence of interface defect generation [119]. Subsequently, we have developed a complete analytical solution for Reaction-Diffusion framework to predict the temporal evolution of N_{IT} over several decades. Moreover, our statistical analysis of R-D system has identified the

existence of log-normal distribution, which (due to practical and theoretical limits) eventually evolves towards a skew-normal or Gaussian distribution.

10.1.2. NBTI for Oxynitride Dielectric Transistors

In chapters 3, 5, and 6, we have identified two competing mechanisms responsible for NBTI in transistors having oxynitride dielectric – (i) defect generation at the oxide-dielectric interface or N_{IT} and (ii) hole trapping into pre-existing oxide defects N_{HT} . For transistors having *plasma oxynitride dielectric*, N_{IT} dominates over N_{HT} . This information encouraged us to propose a scheme (in chapter 7) for co-optimization of both gate leakage and NBTI [128, 184] – two of the most important issues in oxynitride CMOS technology. The gate leakage study resolved the long-standing controversy related to the nitrogen dependence of leakage parameters [184]. Similarly, for transistors having *thermal oxynitride dielectric*, we model N_{HT} within the Shockley-Read-Hall framework to suggest that N_{HT} in nanoscale transistors should saturate quickly within \sim msec [119], a prediction that was later verified by detailed experimental study [203, 387].

10.1.3. NBTI for Strained Transistors

In chapter 3, our proposed Si-H bond dissociation model is applied to clarify whether NBTI in strained transistors is different from that of unstrained transistors [133]. Here, a careful analysis of the strain-dependent NBTI and gate leakage experiments identified hole capture into N_{IT} as the main strain-dependent component. This allowed us to explain the NBTI experiments over a wide range of uniaxial/biaxial, compressive/tensile channel strain within a self-consistent framework.

10.1.4. Interpretation of NBTI Measurements

While working on NBTI physics and comparing it with experiments, we have identified the necessity of explaining “what is measured”. In recent years, various NBTI

characterization results have been mapped back to transistor degradation parameters by using simple approximations to prove/disprove existing NBTI models. As shown in chapter 6, such simplification are often unjustified due to the importance of – (a) first measurement point [161], (b) sub-threshold slope variation [165], (c) mobility variation [203], and (d) valence band electron trapping [165]. This study has reconfirmed the importance of considering N_{HT} in explaining NBTI for oxynitride transistors. Therefore, it resolved the recent controversies between modern ultra-fast NBTI relaxation measurements and Reaction-Diffusion theory [387].

10.1.5. Concept of Variation-Resilient Transistor

Defect formation within a transistor is known to degrade mobility μ and channel carrier concentration Q_C . However, through a careful set of experiments (in chapters 8 and 9), we have shown a fundamentally counter-intuitive result that, depending on the measurement conditions, type of defects and shape of the universal mobility-field characteristics, μ can actually improve with defect generation [165]. Using this principle, we have proposed a novel concept of self-compensated transistor, where scarcity of Q_C due to generated defect is exactly compensated by the increase in μ , thereby keeping the transistor's ON current ($\sim \mu Q_C$) less sensitive to defect generation [347, 376]. The concept has already been generalized [388] to handle different sources of process variation in CMOS transistors. For many years, transistor variabilities have always been handled by using appropriate guard band voltage such that the drivability of a transistor remains above some optimum value (despite maximum parametric variation). Such guard band voltage is significant part of transistor's supply voltage, which is clamped to $\sim 1.0V$ for the last four CMOS generations. The concept of variation-resilience may relax the guard band voltage requirement, hence reduce the supply voltage and power consumption.

Therefore, by using NBTI as a stepping stone, we have provided the novel concept of self-compensation for handling CMOS variabilities. Although circuit designers often adopt different self-correction schemes to handle variations in circuit-level (*e.g.*,

obtaining defect tolerance in HP's Teramac computer [389], also in University of Minnesota's FPGA-based Nanobox [390]), the feasibility of self-compensation within existing CMOS architecture without additional circuitry or extra process steps should prove beneficial and thus ease the requirement of circuit-level correction.

10.2. Future Work

Despite extensive effort has been given in this thesis to understand the dynamics of defect formation and its effect transistor parameters, we believe that the following aspects of defect formation and transistor variability require further work:

- As discussed in section 3.6.5, reverse-annealing of generated interfacial defects might have some concerns in terms of non-zero N_{IT} recovery at long relaxation time. Though the robustness of experimental observations (like Figure 3.14) is yet to be justified, our statistical R-D analysis of chapter 4 might lead a way for potential resolution of this issue (see section 10.2.1).
- The concept of variation-resilient transistor is proposed in chapter 9. Such concept might lead a way of relaxing the requirement of circuit and system level optimization, used in current CMOS architecture. Though we have experimentally verified the concept for temporal V_T fluctuations on transistors, the concept still requires experimental verification for resilience to process variations. Moreover, the concept is yet to be analyzed on large circuits and systems, as well as on III-V transistor technology. In section 10.2.2, we discuss how one could address the problem in a systematic manner.
- The work on trapping into oxide defects (chapter 5) sets up a platform for studying the reliability of transistors having high- κ gate dielectric (see section 10.2.3). One can also study the degradation of α -Si solar cell, by utilizing its analogy [66] with N_{IT} formation during NBTI stress (presented in chapter 3). Moreover, the variation-resilient concept may have implications in the fields of chemical sensors, cryogenic electronics, and in designing radiation-prone transistors.

In the following sections, we show how the formalism and understanding (discussed in this thesis) might be used to address each of the above problems.

10.2.1. Statistical Distribution Si-H Bonds

From the discussion in section 2.2, we know that the density of Si-H bonds at the Si-SiO₂ interface is $\sim 10^{12}$ cm⁻², which is statistically distributed among the $\sim 10^{14}$ cm⁻² Si-lattice sites, as schematically shown in Figure 10.1a. So when hydrogen diffuse away from the Si-SiO₂ interface, they cannot initially maintain a 1D diffusion-front, as assumed in the discussion so far (*e.g.*, see Figure 3.1). Some of the hydrogen, generated from the dissociation of Si-H bonds completely isolated from each other, can even proceed with a 3D diffusion at the initial stage of NBTI stress (see Figure 10.1a). After some period of NBTI stress, sufficient amount of Si-H bonds will be dissociated and hydrogen diffusion front from each of the sources (or Si-H sites) will mix up. So, at long t_{STS} , hydrogen front can equivalently be represented using a 1D profile, as used in chapter 3. As a result, during initial phase of NBTI stress, we expect to observe significantly different time dynamics due to higher dimensionality of hydrogen diffusion (see our preliminary simulation result in Figure 10.2), compared to the one-dimensional hydrogen diffusion used in chapter 3.

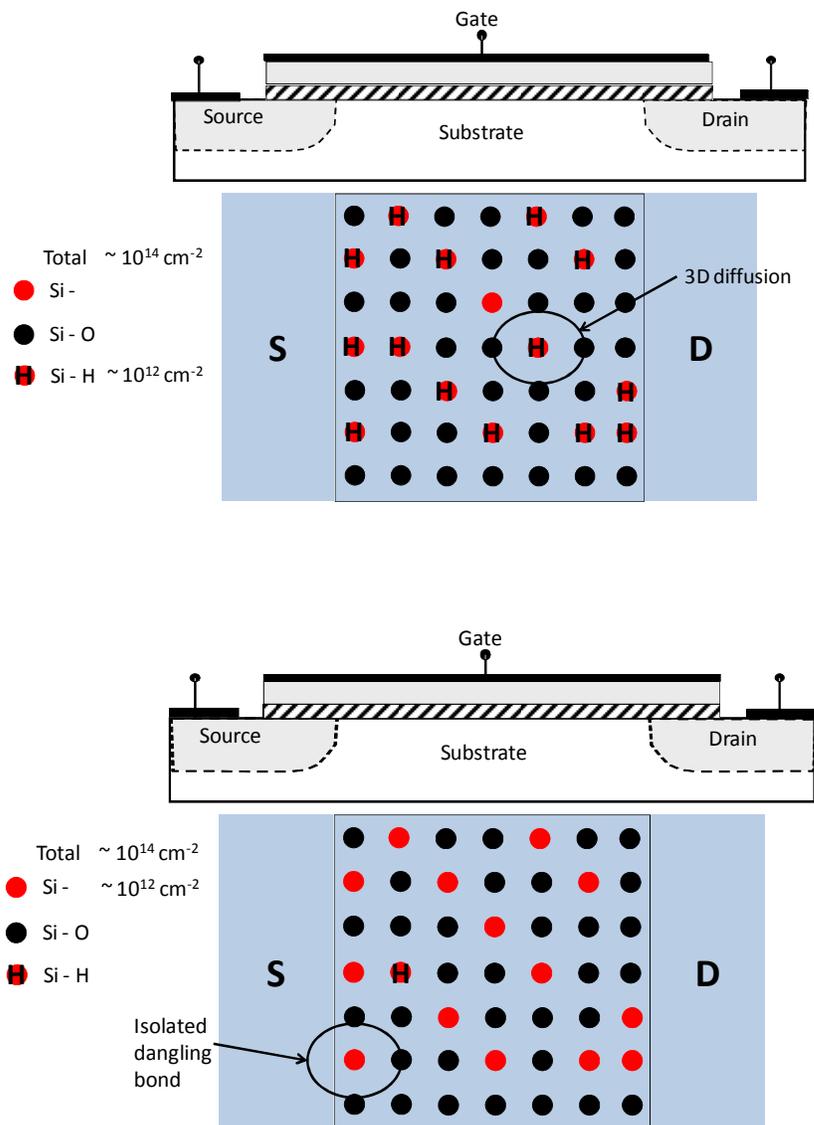


Figure 10.1: Schematic representation of the Si-SiO₂ interface, seen from the top of a MOS transistor. Most of the dangling Si- bonds (total density $\sim 10^{14} \text{ cm}^{-2}$) are terminated by O atoms during oxidation of Si substrate. (a) Remaining Si- bonds (density $\sim 10^{12} \text{ cm}^{-2}$) are annealed using hydrogen. These Si-H bonds are spatially distributed at the interface. (b) After most of the Si-H bonds are dissociated due to the application of NBTI stress, the resultant dangling Si- bonds can be sparsely distributed at the Si/SiO₂ interface.

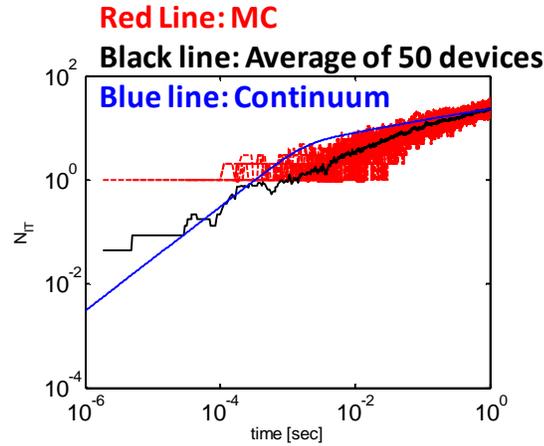


Figure 10.2: Statistical R-D simulation considers distribution of interface defects at the Si/SiO₂ interface. This increases the dimensionality of hydrogen diffusion at smaller time. As such, time exponent of N_{IT} generation at smaller time will be considerably higher than the one obtained from continuum simulation having one-dimensional hydrogen diffusion (Figure 3.1).

Now assuming that most of the Si-H bonds have dissociated at the end of stress phase, we will have a sea of hydrogen species on top of the Si-SiO₂ interface at the start of NBTI relaxation phase. Moreover, as the Si-H bonds are only 1% of total Si-O + Si-H bonds, the generated dangling Si- bonds are indeed sparsely distributed at the interface (see Figure 10.1b for a schematic of Si-SiO₂ interface). So, during NBTI relaxation, hydrogen's back-diffusion into the isolated dangling Si- bonds can be difficult. As such, these isolated dangling Si- bonds may remain unpassivated even after long relaxation time. The situation is similar to the Hot-Carrier-Injection (HCI) phenomenon, studied in [391]. During HCI stress, Si-H bond dissociation happens near the drain end by energetic carriers, located in that region. Hence, the generated hydrogen species has a 2D diffusion front, as shown in Figure 10.3a. During relaxation, hydrogen species has less probability to anneal the broken Si- bonds near the drain end, which results in negligible recovery. Such trend is evidently present in 2D R-D simulation (Figure 10.3b), as well as in HCI experiment (Figure 10.3c). Again, we perform some preliminary analysis of the impact of sparse distribution of dangling Si- bonds on N_{IT} relaxation, by extending our statistical R-D

formalism (presented in chapter 4). Our analysis suggests an increase in the dimensionality of N_{IT} relaxation, over its value of unity for equation (3.37). Therefore, sparse nature of dangling Si- bond distribution indeed can delay the amount of N_{IT} relaxation. However, consideration of higher dimensionality can not provide absolute lock-in (no relaxation for some part of N_{IT} , generated during stress), as reported in some experiments (Figure 3.13a).

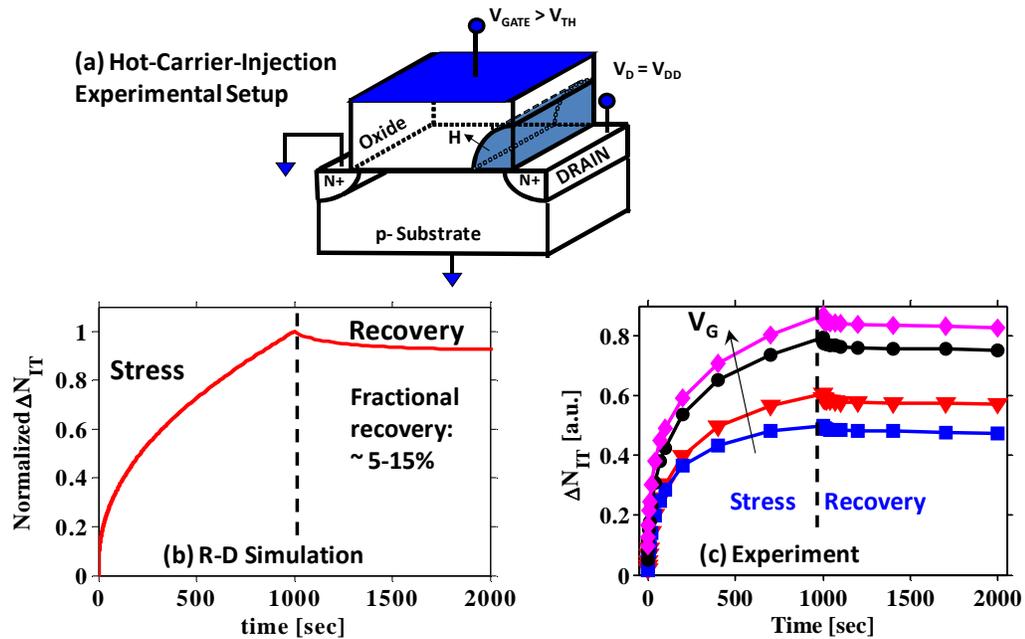


Figure 10.3: (a) During Hot-Carrier-Injection (HCI) stress, Si-H bonds are broken only near the drain end by energetic carriers. Resultant hydrogen species have a 2D diffusion front. (b) R-D simulation and (c) experiment both indicates negligible amount of HCI relaxation or annealing of Si- bonds due to difficulty in the back-diffusion of hydrogen species. Taken from [391].

10.2.2. Research on Variation Resilience

In chapter 9, we have proposed the possibility of designing variation-resilient transistor for reducing power consumption in CMOS technology. Though variation-

resilient transistors ensure negligible drive current fluctuation at a certain voltage range, one needs to quantify the advantage of using variation-resilience transistor in CMOS circuits and systems, where the transistors operate in a random manner. To achieve this, we suggest the use of HSPICE and MEDICI simulation to identify appropriate mobility-field relationship (achieved using some advanced CMOS substrates, like strain or III-V or Ge) and operating voltage (V_{DD}) for designing transistors, suitable for variation-resilient circuit operation. We summarize the proposed approach using a flowchart in Figure 10.4.

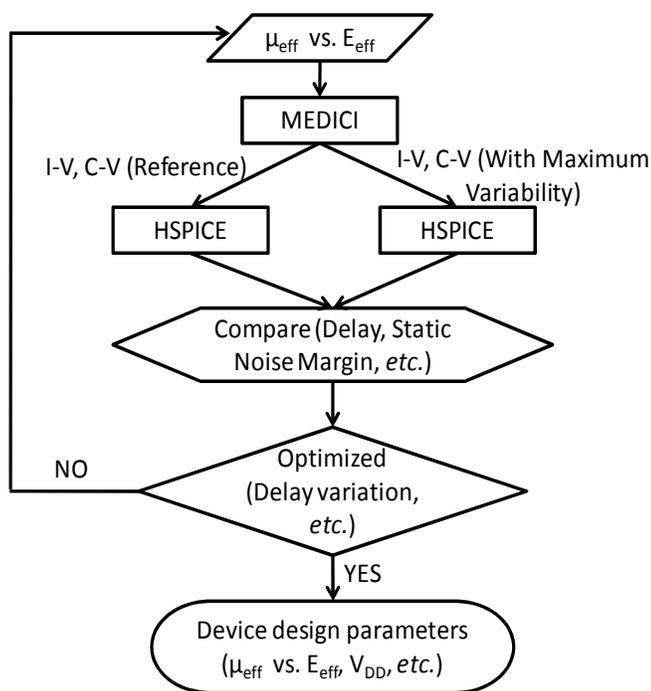


Figure 10.4: Flowchart for designing transistors for variation-resilient circuit operation.

In addition, we expect our variation-resilient concept to be viable for designing: (1) radiation-tolerant transistors (*i.e.*, making the I_D of these transistors insensitive to any defects created by radiation damage), (2) variation-resilient cryogenic electronics, (3) degradation-free Ion-Sensitive FET (ISFET) or chemical FET (CHEMFET) devices or other biosensors (making them sensitive only to existing chemical substances, rather than oxide traps), *etc.*

10.2.3. Extension/Implication of Hole Trapping Analysis

The hole trapping analysis performed in chapter 5 can be extended in analyzing PBTI and NBTI studies for transistors with high- κ gate dielectric. Here, the temperature dependency of electron trapping observed in long-term high- κ PBTI measurements [47, 242, 392] are inconsistent with our trapping/detrapping analysis (Figure 5.8b). This indicates that in addition to trapping into pre-existing oxide defects, there must be an additional contribution from some other mechanism, which is a part of our future study. As high- κ dielectrics are extensively used in carbon nanotube based transistors [393-397], trapping/detrapping study in these transistors during bias temperature instability (BTI) and hot carrier injection (HCI) experiments should also be an interesting topic for future study.

Moreover, ferroelectric materials are recently proposed for use within CMOS architecture [398] for improving the subthreshold properties of transistors by reducing the sub-threshold slope. These ferroelectric materials have a higher coordination number (*i.e.*, number of nearest neighbors for a particular atom within the material). As described in [108, 110], higher coordination number in high- κ materials makes it difficult to satisfy all the chemical bonding within the amorphous network, thus transistors with high- κ dielectric are prone to high density of pre-existing defects. As a result, these transistors generally suffer more from BTI effects compared to their oxynitride counterparts. Since ferroelectric materials have a coordination number higher than the high- κ materials currently used in literature, it is expected to suffer more from different reliability issues. This is evident from the reliability study in ferroelectric memories, which indicates presence of reliability problems like: retention failures [399], fatigue (defined as a resistance to polarization reversal that develops after repeated cycling of the memory capacitor) [399-402], imprint failures (defined as the inability to read a signal after a long-time holding during multiple same data write operations) [403-406], *etc.* Thus, reliability of ferroelectric transistors will be an interesting issue for future study.

10.3. Epilogue

This PhD research was initiated to understand the implication of Reaction-Diffusion theory in interface defect formation of nanoscale transistors. The formalism, developed during this period, eventually allowed the proposal of a set of algorithms for interpreting a wide range of experiments. Moreover, study of the effect of defect on transistor performance enabled the concept of ‘self-compensation’, whose utilization should reduce the impact of defect formation on transistor performance. Finally, utilizing the analogy between process variation and defect formation in nanoscale transistor, the concept of ‘self-compensation’ has been generalized to handle many sources of transistor variabilities – which should overcome one of the significant challenge in transistor scaling following Moore’s law.

In recent days, nano-research involves many emerging concepts, including the development of flexible electronics, improving the efficiency of renewable energy sources, and utilizing floating body conduction through materials placed on insulators. Amorphous materials and their interfaces play significant role in these researches. So, the familiarity with defects in amorphous materials and (the widely used) Reaction-Diffusion theory should enable one to study many emerging issues of modern technology.

LIST OF REFERENCES

LIST OF REFERENCES

- [1] J. Bardeen and W. H. Brattain, "Three Electrode Circuit Element Utilizing Semiconductive Materials," US, 1948.
- [2] W. Shockley, "Circuit Element Utilizing Semiconductive Materials," US, 1948.
- [3] G. E. Moore, "Cramming more components onto integrated circuits (Reprinted from Electronics, pg 114-117, April 19, 1965)," *Proceedings of the Ieee*, vol. 86, pp. 82-85, Jan 1998.
- [4] S. Thompson, P. Packan, and M. Bohr, "MOS Scaling: Transistor Challenges for the 21st Century," *Intel Technology Journal*, vol. 2, 1998.
- [5] K. Bernstein, D. J. Frank, A. E. Gattiker, W. Haensch, B. L. Ji, S. R. Nassif, E. J. Nowak, D. J. Pearson, and N. J. Rohrer, "High-performance CMOS variability in the 65-nm regime and beyond," *IBM Journal of Research and Development*, vol. 50, pp. 433-449, Jul-Sep 2006.
- [6] D. K. Schroder and J. A. Babcock, "Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing," *Journal of Applied Physics*, vol. 94, pp. 1-18, JUL 1 2003.
- [7] "International Technology Roadmap for Semiconductors," 2009 ed, 2009.
- [8] A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub-0.1 μm MOSFET's: A 3-D "atomistic" simulation study," *Ieee Transactions on Electron Devices*, vol. 45, pp. 2505-2513, Dec 1998.
- [9] A. Asenov, A. R. Brown, J. H. Davies, S. Kaya, and G. Slavcheva, "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs," *Ieee Transactions on Electron Devices*, vol. 50, pp. 1837-1852, Sep 2003.
- [10] A. Asenov, S. Kaya, and A. R. Brown, "Intrinsic parameter fluctuations in decananometer MOSFETs introduced by gate line edge roughness," *Ieee Transactions on Electron Devices*, vol. 50, pp. 1254-1260, May 2003.
- [11] A. Asenov, S. Kaya, and J. H. Davies, "Intrinsic threshold voltage fluctuations in decanano MOSFETs due to local oxide thickness variations," *Ieee Transactions on Electron Devices*, vol. 49, pp. 112-119, Jan 2002.
- [12] K. Huang, *Statistical Mechanics*, 2nd ed.: John Wiley & Sons, 1987.

- [13] T. Mizuno, J. Okamura, and A. Toriumi, "Experimental-Study of Threshold Voltage Fluctuation Due to Statistical Variation of Channel Dopant Number in Mosfets," *Ieee Transactions on Electron Devices*, vol. 41, pp. 2216-2221, Nov 1994.
- [14] H. S. Momose, M. Ono, T. Yoshitomi, T. Ohguro, S. Nakamura, M. Saito, and H. Iwai, "1.5 nm direct-tunneling gate oxide Si MOSFET's," *Ieee Transactions on Electron Devices*, vol. 43, pp. 1233-1242, Aug 1996.
- [15] H. Dadgour, K. Endo, V. De, and K. Banerjee, "Modeling and Analysis of Grain-Orientation Effects in Emerging Metal-Gate Devices and Implications for SRAM Reliability," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 705-708, 2008.
- [16] N. D. Lang and W. Kohn, "Theory of Metal Surfaces - Charge Density and Surface Energy," *Physical Review B*, vol. 1, pp. 4555-&, 1970.
- [17] G. F. Cardinale, C. C. Henderson, J. E. M. Goldsmith, P. J. S. Mangat, J. Cobb, and S. D. Hector, "Demonstration of pattern transfer into sub-100 nm polysilicon line/space features patterned with extreme ultraviolet lithography," *Journal of Vacuum Science & Technology B*, vol. 17, pp. 2970-2974, Nov-Dec 1999.
- [18] Christen.S, Lundstro.I, and C. Svensson, "Low Frequency Noise in Mos Transistors .I. Theory," *Solid-State Electronics*, vol. 11, pp. 797-&, 1968.
- [19] D. M. Fleetwood, H. D. Xiong, Z. Y. Lu, C. J. Nicklaw, J. A. Felix, R. D. Schrimpf, and S. T. Pantelides, "Unified model of hole trapping, 1/f noise, and thermally stimulated current in MOS devices," *Ieee Transactions on Nuclear Science*, vol. 49, pp. 2674-2683, Dec 2002.
- [20] S. T. Hsu, Fitzgera.Dj, and A. S. Grove, "Surface-State Related L/F Noise in P-N Junctions and Mos Transistors," *Applied Physics Letters*, vol. 12, pp. 287-&, 1968.
- [21] K. K. Hung, P. K. Ko, C. M. Hu, and Y. C. Cheng, "A Unified Model for the Flicker Noise in Metal Oxide-Semiconductor Field-Effect Transistors," *Ieee Transactions on Electron Devices*, vol. 37, pp. 654-665, Mar 1990.
- [22] A. L. Mcwhorter, "Surface Traps and 1-F Noise in Germanium," *Physical Review*, vol. 98, pp. 1191-1192, 1955.
- [23] C. T. Sah and F. Hielscher, "Evidence of Surface Origin of 1/F Noise," *Physical Review Letters*, vol. 17, pp. 956-&, 1966.
- [24] C. Kittel, *Introduction to Solid State Physics*, 7th ed.: John Wiley & Sons, Inc., 1996.
- [25] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*: Thompson Learning Inc., 1976.
- [26] R. Zallen, *The Physics of Amorphous Solids*, 1st ed.: John Wiley & Sons, 1983.

- [27] J. E. Lilienfield, "Method and Apparatus for Controlling Electrical Current," US, 1926.
- [28] J. E. Lilienfield, "Device for Controlling Electric Current," US, 1928.
- [29] B. E. Deal, M. Sklar, A. S. Grove, and E. H. Snow, "Characteristics of Surface-State Charge (Q_{ss}) of Thermally Oxidized Silicon," *Journal of the Electrochemical Society*, vol. 114, pp. 266-&, 1967.
- [30] M. A. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectronics Reliability*, vol. 45, pp. 71-81, JAN 2005.
- [31] S. Chakravarthi, A. Krishnan, V. Reddy, C. F. Machala, and S. Krishnan, "A comprehensive framework for predictive modeling of negative bias temperature instability," in *IEEE International Reliability Physics Symposium*, 2004, pp. 273 - 282.
- [32] V. Huard, M. Denais, F. Perrier, N. Revil, C. Parthasarathy, A. Bravaix, and E. Vincent, "A thorough investigation of MOSFETs NBTI degradation," *Microelectronics Reliability*, vol. 45, pp. 83-98, JAN 2005.
- [33] B. Kaczer, V. Arkhipov, R. Degraeve, N. Collaert, G. Groeseneken, and M. Goodwin, "Disorder-controlled-kinetics model for negative bias temperature instability and its experimental verification," in *IEEE International Reliability Physics Symposium*, 2005, pp. 381 - 387.
- [34] M. A. Alam, H. Kufluoglu, D. Varghese, and S. Mahapatra, "A Comprehensive Model for PMOS NBTI Degradation: Recent Progress," in *in press, Micro. Reliability*, 2006.
- [35] "International Technology Roadmap for Semiconductors," 2005 ed, 2005.
- [36] S. S. Tan, T. P. Chen, J. M. Soon, K. P. Loh, C. H. Ang, and L. Chan, "Nitrogen-enhanced negative bias temperature instability: An insight by experiment and first-principle calculations," *Applied Physics Letters*, vol. 82, pp. 1881-1883, MAR 24 2003.
- [37] Y. Mitani, "Influence of nitrogen in ultra-thin SiON on negative bias temperature instability under AC stress," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2004, pp. 117 - 120.
- [38] M. L. Green, E. P. Gusev, R. Degraeve, and E. L. Garfunkel, "Ultrathin (< 4 nm) SiO₂ and Si-O-N gate dielectric layers for silicon microelectronics: Understanding the processing, structure, and physical and electrical limits," *Journal of Applied Physics*, vol. 90, pp. 2057-2121, Sep 1 2001.
- [39] G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High-kappa gate dielectrics: Current status and materials properties considerations," *Journal of Applied Physics*, vol. 89, pp. 5243-5275, MAY 15 2001.

- [40] C. C. Hobbs, L. R. C. Fonseca, A. Knizhnik, V. Dhandapani, S. B. Samavedam, W. J. Taylor, J. M. Grant, L. G. Dip, D. H. Triyoso, R. I. Hegde, D. C. Gilmer, R. Garcia, D. Roan, M. L. Lovejoy, R. S. Rai, E. A. Hebert, H. H. Tseng, S. G. H. Anderson, B. E. White, and P. J. Tobin, "Fermi-level pinning at the polysilicon/metal oxide interface - Part I," *Ieee Transactions on Electron Devices*, vol. 51, pp. 971-977, Jun 2004.
- [41] C. C. Hobbs, L. R. C. Fonseca, A. Knizhnik, V. Dhandapani, S. B. Samavedam, W. J. Taylor, J. M. Grant, L. G. Dip, D. H. Triyoso, R. I. Hegde, D. C. Gilmer, R. Garcia, D. Roan, M. L. Lovejoy, R. S. Rai, E. A. Hebert, H. H. Tseng, S. G. H. Anderson, B. E. White, and P. J. Tobin, "Fermi-level pinning at the polysilicon metal oxide interface - Part II," *Ieee Transactions on Electron Devices*, vol. 51, pp. 978-984, Jun 2004.
- [42] S. B. Samavedam, L. B. La, P. J. Tobin, B. White, C. Hobbs, L. R. C. Fonseca, A. Demkov, J. Schaeffer, E. Luckowski, A. Martinez, M. Raymond, D. Triyoso, D. Roan, V. Dhandapani, R. Garcia, S. G. H. Anderson, K. Moore, H. H. Tseng, C. Capasso, O. Adetutu, D. C. Gilmer, W. J. Taylor, R. Hegde, and J. Grant, "Fermi level pinning with sub-monolayer MeOx and metal gates," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 307-310, 2003.
- [43] G. Ribes, J. Mitard, M. Denais, S. Bruyere, F. Monsieur, C. R. Parthasarathy, E. Vincent, and G. Ghibaudo, "Review on high-k dielectrics reliability issues," *Ieee Transactions on Device and Materials Reliability*, vol. 5, pp. 5-19, MAR 2005.
- [44] G. Bersuker, J. H. Sim, C. S. Park, C. D. Young, S. V. Nadkarni, R. Choi, and B. H. Lee, "Mechanism of electron trapping and characteristics of traps in HfO₂ gate stacks," *Ieee Transactions on Device and Materials Reliability*, vol. 7, pp. 138-145, MAR 2007.
- [45] C. M. Hu, S. C. Tam, F. C. Hsu, P. K. Ko, T. Y. Chan, and K. W. Terrill, "Hot-Electron-Induced Mosfet Degradation - Model, Monitor, and Improvement," *Ieee Transactions on Electron Devices*, vol. 32, pp. 375-385, 1985.
- [46] H. Kufluoglu and M. A. Alam, "Theory of interface-trap-induced NBTI degradation for reduced cross section MOSFETs," *Ieee Transactions on Electron Devices*, vol. 53, pp. 1120-1130, MAY 2006.
- [47] S. Pae, M. Agostinelli, M. Brazie, R. Chau, G. Dewey, T. Ghani, M. Hattendorf, J. Hicks, J. Kavalieros, K. Kuhn, M. Kuhn, J. Maiz, M. Metz, K. Mistry, C. Prasad, S. Ramey, A. Roskowski, J. Sandford, C. Thomas, J. Thomas, C. Wiegand, and J. Wiedemer, "BTI reliability of 45 nm high-k plus metal-gate process technology," *International Reliability Physics Symposium Proceedings*, pp. 352-357, 2008.
- [48] S. Pae, J. Maiz, C. Prasad, and B. Woolery, "Effect of BTI Degradation on Transistor Variability in Advanced Semiconductor Technologies," *Ieee Transactions on Device and Materials Reliability*, vol. 8, pp. 519-525, Sep 2008.

- [49] S. E. Rauch, "The Statistics of NBTI-Induced V_T and β Mismatch Shifts in pMOSFETs," *Ieee Transactions on Device and Materials Reliability*, vol. 2, pp. 89-93, 2002.
- [50] K. Kang, S. P. Park, K. Roy, and M. A. Alam, "Estimation of Statistical Variation in Temporal NBTI Degradation and its Impact on Lifetime Circuit Performance," in *IEEE/ACM International Conference on Computer-Aided Design*, 2007, pp. 730-734.
- [51] N. Klein, "Mechanism of Self-Healing Electrical Breakdown in Mos Structures," *Ieee Transactions on Electron Devices*, vol. Ed13, p. 788, 1966.
- [52] P. Solomon, "Breakdown in Silicon-Oxide," *Journal of Vacuum Science & Technology*, vol. 14, pp. 1122-1130, 1977.
- [53] K. F. Schuegraf and C. M. Hu, "Hole Injection Sio₂ Breakdown Model for Very-Low Voltage Lifetime Extrapolation," *Ieee Transactions on Electron Devices*, vol. 41, pp. 761-767, May 1994.
- [54] R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, and H. E. Maes, "A consistent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 863-866, 1995.
- [55] D. J. DiMaria and J. H. Stathis, "Explanation for the oxide thickness dependence of breakdown characteristics of metal-oxide-semiconductor structures," *Applied Physics Letters*, vol. 70, pp. 2708-2710, May 19 1997.
- [56] J. H. Stathis, "Percolation models for gate oxide breakdown," *Journal of Applied Physics*, vol. 86, pp. 5757-5766, Nov 15 1999.
- [57] M. Alam, B. Weir, and P. Silverman, "A future of function or failure?," *Ieee Circuits & Devices*, vol. 18, pp. 42-48, Mar 2002.
- [58] M. Alam, "Reliability- and process-variation aware design of integrated circuits," *Microelectronics Reliability*, vol. 48, pp. 1114-1122, Aug-Sep 2008.
- [59] Y. H. Lee, W. McMahon, N. Mielke, Y. L. R. Lu, and S. Walstra, "Managing bias-temperature instability for product reliability," *Proceedings of International Symposium on VLSI Technology*, pp. 52-53, 2007.
- [60] S. M. Martin, K. Flautner, T. Mudge, and D. Blaauw, "Combined dynamic voltage scaling and adaptive body biasing for lower power microprocessors under dynamic workloads," *IEEE/ACM International Conference on CAD*, pp. 721-725, 2002.
- [61] J. W. Tschanz, J. T. Kao, S. G. Narendra, R. Nair, D. A. Antoniadis, A. P. Chandrakasan, and V. De, "Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and leakage," *IEEE Journal of Solid-State Circuits*, vol. 37, pp. 1396-1402, Nov 2002.

- [62] K. Kang, K. Kim, and K. Roy, "Variation resilient low-power circuit design methodology using on-chip phase locked loop," *ACM/IEEE Design Automation Conference*, pp. 934-939, 2007.
- [63] M. Agarwal, B. C. Paul, M. Zhang, and S. Mitra, "Circuit failure prediction and its application to transistor aging," *Proceedings of IEEE VLSI Test Symposium*, pp. 277-284, 2007.
- [64] B. C. Paul, K. H. Kang, H. Kuflluoglu, M. A. Alam, and K. Roy, "Negative bias temperature instability: Estimation and design for improved reliability of nanoscale circuits," *Ieee Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 26, pp. 743-751, APR 2007.
- [65] T. L. Anderson, *Fracture Mechanics: Fundamentals and Applications*, 2nd ed.: CRC Press, 1994.
- [66] D. R. Allee, L. T. Clark, R. Shringarpure, S. M. Venugopal, Z. P. Li, and E. J. Bawolek, "Degradation effects in a-Si : H thin film transistors and their impact on circuit performance," *IEEE International Reliability Physics Symposium*, pp. 158-167, 2008.
- [67] W. Gopel, "Ultimate limits in the miniaturization of chemical sensors," *Sensors and Actuators a-Physical*, vol. 56, pp. 83-102, Aug 1996.
- [68] P. J. Caplan, E. H. Poindexter, B. E. Deal, and R. R. Razouk, "ESR Centers, Interface States, and Oxide Fixed Charge in Thermally Oxidized Silicon Wafers," *Journal of Applied Physics*, vol. 50, pp. 5847-5854, 1979.
- [69] C. R. Helms and E. H. Poindexter, "The Silicon Silicon-Dioxide System - Its Microstructure and Imperfections," *Reports on Progress in Physics*, vol. 57, pp. 791-852, Aug 1994.
- [70] M. L. Reed and J. D. Plummer, "Chemistry of Si-SiO₂ Interface Trap Annealing," *Journal of Applied Physics*, vol. 63, pp. 5776-5793, JUN 15 1988.
- [71] J. A. Appelbaum and D. R. Hamann, "Theory of Reconstruction Induced Subsurface Strain - Application to Si(100)," *Surface Science*, vol. 74, pp. 21-33, 1978.
- [72] D. E. Eastman, "Geometrical and Electronic-Structure of Si(001) and Si(111) Surfaces - a Status-Report," *Journal of Vacuum Science & Technology*, vol. 17, pp. 492-500, 1980.
- [73] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed.: Wiley, 1981.
- [74] L. D. Thanh and P. Balk, "Elimination and Generation of Si-SiO₂ Interface Traps by Low-Temperature Hydrogen Annealing," *Journal of the Electrochemical Society*, vol. 135, pp. 1797-1801, Jul 1988.
- [75] A. H. Edwards, "Theory of P_b Center at <111> Si/SiO₂ Interface," *Physical Review B*, vol. 36, pp. 9638-9648, 1987.

- [76] Y. Nishi, "Study of Silicon-Silicon Dioxide Structure by Electron Spin Resonance .1.," *Japanese Journal of Applied Physics*, vol. 10, p. 52, 1971.
- [77] Y. Y. Kim and P. M. Lenahan, "Electron-Spin-Resonance Study of Radiation-Induced Paramagnetic Defects in Oxides Grown on (100) Silicon Substrates," *Journal of Applied Physics*, vol. 64, pp. 3551-3557, Oct 1 1988.
- [78] E. H. Poindexter, P. J. Caplan, B. E. Deal, and R. R. Razouk, "Interface States and Electron-Spin Resonance Centers in Thermally Oxidized (111) and (100) Silicon-Wafers," *Journal of Applied Physics*, vol. 52, pp. 879-884, 1981.
- [79] A. Stesmans and V. V. Afanas'ev, "Electron spin resonance features of interface defects in thermal (100)Si/SiO₂," *Journal of Applied Physics*, vol. 83, pp. 2449-2457, Mar 1 1998.
- [80] K. L. Brower, "Structural Features at the Si-SiO₂ Interface," *Zeitschrift Fur Physikalische Chemie Neue Folge*, vol. 151, pp. 177-189, 1987.
- [81] P. M. Lenahan, "Dominating Defects in the MOS System: P_b and E' Centers," in *Defects in Microelectronic Materials and Devices*, S. P. D. Fleetwood, R. D. Schrimpf, Ed.: CRC Press, 2008, pp. 163-214.
- [82] P. V. Gray and D. M. Brown, "Density of SiO₂-Si Interface States - (MOS Devices Si Oxidation 100 Degrees CO₂ + 80 Ppm H₂O Fermi Level E)," *Applied Physics Letters*, vol. 8, p. 31, 1966.
- [83] P. M. Lenahan and P. V. Dressendorfer, "An Electron-Spin Resonance Study of Radiation-Induced Electrically Active Paramagnetic Centers at the Si/SiO₂ Interface," *Journal of Applied Physics*, vol. 54, pp. 1457-1460, 1983.
- [84] P. M. Lenahan and P. V. Dressendorfer, "Hole Traps and Trivalent Silicon Centers in Metal-Oxide Silicon Devices," *Journal of Applied Physics*, vol. 55, pp. 3495-3499, 1984.
- [85] G. J. Gerardi, E. H. Poindexter, P. J. Caplan, and N. M. Johnson, "Interface Traps and Pb Centers in Oxidized (100) Silicon-Wafers," *Applied Physics Letters*, vol. 49, pp. 348-350, Aug 11 1986.
- [86] P. M. Lenahan and J. F. Conley, "What can electron paramagnetic resonance tell us about the Si/SiO₂ system?," *Journal of Vacuum Science & Technology B*, vol. 16, pp. 2134-2153, Jul-Aug 1998.
- [87] J. P. Campbell and P. M. Lenahan, "Density of states of P-b1 Si/SiO₂ interface trap centers," *Applied Physics Letters*, vol. 80, pp. 1945-1947, Mar 18 2002.
- [88] J. P. Campbell, P. M. Lenahan, C. J. Cochrane, A. T. Krishnan, and S. Krishnan, "Atomic-scale defects involved in the negative-bias temperature instability," *Ieee Transactions on Device and Materials Reliability*, vol. 7, pp. 540-557, Dec 2007.
- [89] A. T. Krishnan, V. Reddy, S. Chakravarthi, J. Rodriguez, S. Natarajan, S. John, and S. Krishnan, "NBTI impact on Transistor and Circuit: Models, Mechanisms and Scaling Effects," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2003, pp. 349-352.

- [90] V. Reddy, A. T. Krishnan, A. Marshall, J. Rodriguez, S. Natarajan, T. Rost, and S. Krishnan, "Impact of negative bias temperature instability on digital circuit reliability," in *IEEE International Reliability Physics Symposium*, 2002, pp. 248-254.
- [91] Y. Baryam, D. Adler, and J. D. Joannopoulos, "Structure and Electronic States in Disordered-Systems," *Physical Review Letters*, vol. 57, pp. 467-470, Jul 28 1986.
- [92] H. M. Branz, "Dangling Bonds in Doped Amorphous-Silicon - Equilibrium, Relaxation, and Transition Energies," *Physical Review B*, vol. 39, pp. 5107-5115, Mar 15 1989.
- [93] R. A. Street, "Localized States in Doped Amorphous-Silicon," *Journal of Non-Crystalline Solids*, vol. 77-8, pp. 1-16, Dec 1985.
- [94] D. Jousse, "Direct Electron-Spin-Resonance Determination of Dangling Bond-Energies in Rf-Sputtered Hydrogenated Amorphous-Silicon," *Applied Physics Letters*, vol. 49, pp. 1438-1440, Nov 24 1986.
- [95] A. Stesmans, D. Pierreux, R. J. Jaccodine, M. T. Lin, and T. J. Delph, "Influence of in situ applied stress during thermal oxidation of (111)Si on P-b interface defects," *Applied Physics Letters*, vol. 82, pp. 3038-3040, May 5 2003.
- [96] P. Somers, A. Stesmans, V. V. Afanas'ev, C. Claeys, and E. Simoen, "Paramagnetic point defects at interfacial layers in biaxial tensile strained (100)Si/SiO₂," *Journal of Applied Physics*, vol. 103, pp. -, Feb 1 2008.
- [97] J. P. Campbell, P. M. Lenahan, A. T. Krishnan, and S. Krishnan, "Location, structure, and density of states of NBTI-induced defects in plasma nitrated pMOSFETs," *IEEE International Reliability Physics Symposium*, pp. 503-510, 2007.
- [98] P. M. Lenahan and S. E. Curry, "1st Observation of the Si-29 Hyperfine Spectra of Silicon Dangling Bond Centers in Silicon-Nitride," *Applied Physics Letters*, vol. 56, pp. 157-159, Jan 8 1990.
- [99] A. Y. Kang, P. M. Lenahan, J. F. Conley, and R. Solanki, "Electron spin resonance study of interface defects in atomic layer deposited hafnium oxide on Si," *Applied Physics Letters*, vol. 81, pp. 1128-1130, Aug 5 2002.
- [100] P. D. Kirsch, M. A. Quevedo-Lopez, H. J. Li, Y. Senzaki, J. J. Peterson, S. C. Song, S. A. Krishnan, N. Moumen, J. Barnett, G. Bersuker, P. Y. Hung, B. H. Lee, T. Lafford, Q. Wang, D. Gay, and J. G. Ekerdt, "Nucleation and growth study of atomic layer deposited HfO₂ gate dielectrics resulting in improved scaling and electron mobility," *Journal of Applied Physics*, vol. 99, pp. -, Jan 15 2006.
- [101] B. B. Triplett, P. T. Chen, Y. Nishi, P. H. Kasai, J. J. Chambers, and L. Colombo, "Electron spin resonance study of as-deposited and annealed (HfO₂)_x(SiO₂)_(1-x) high-kappa dielectrics on Si," *Journal of Applied Physics*, vol. 101, pp. -, Jan 1 2007.

- [102] A. Stesmans and V. V. Afanas'ev, "Si dangling-bond-type defects at the interface of (100)Si with ultrathin HfO₂," *Applied Physics Letters*, vol. 82, pp. 4074-4076, Jun 9 2003.
- [103] R. A. Weeks, "The Many Varieties of E' Centers - a Review," *Journal of Non-Crystalline Solids*, vol. 179, pp. 1-9, Nov 1994.
- [104] R. A. Weeks, "Paramagnetic Resonance of Lattice Defects in Irradiated Quartz," *Journal of Applied Physics*, vol. 27, pp. 1376-1381, 1956.
- [105] T. Takahashi, B. B. Triplett, K. Yokogawa, and T. Sugano, "Electron-Spin-Resonance Observation of the Creation, Annihilation, and Charge State of the 74-Gauss Doublet in Device Oxides Damaged by Soft X-Rays," *Applied Physics Letters*, vol. 51, pp. 1334-1336, Oct 26 1987.
- [106] B. B. Triplett, T. Takahashi, and T. Sugano, "Electron-Spin-Resonance Observation of Defects in Device Oxides Damaged by Soft X-Rays," *Applied Physics Letters*, vol. 50, pp. 1663-1665, Jun 8 1987.
- [107] J. F. Conley and P. M. Lenahan, "Room-Temperature Reactions Involving Silicon Dangling Bond Centers and Molecular-Hydrogen in Amorphous SiO₂ Thin-Films on Silicon," *Ieee Transactions on Nuclear Science*, vol. 39, pp. 2186-2191, Dec 1992.
- [108] J. Robertson, "High dielectric constant gate oxides for metal oxide Si transistors," *Reports on Progress in Physics*, vol. 69, pp. 327-396, FEB 2006.
- [109] G. Lucovsky, "Transition from thermally grown gate dielectrics to deposited gate dielectrics for advanced silicon devices: A classification scheme based on bond ionicity," *Journal of Vacuum Science & Technology A*, vol. 19, pp. 1553-1561, Jul-Aug 2001.
- [110] J. Robertson, "Interfaces and defects of high-K oxides on silicon," *Solid-State Electronics*, vol. 49, pp. 283-293, Mar 2005.
- [111] J. T. Ryan, P. M. Lenahan, G. Bersuker, and P. Lysaght, "Electron spin resonance observations of oxygen deficient silicon atoms in the interfacial layer of hafnium oxide based metal-oxide-silicon structures," *Applied Physics Letters*, vol. 90, pp. - , Apr 23 2007.
- [112] D. Varghese, S. Mahapatra, and M. A. Alam, "Hole energy dependent interface trap generation in MOSFET Si/SiO₂ interface," *Ieee Electron Device Letters*, vol. 26, pp. 572-574, AUG 2005.
- [113] S. Mahapatra, D. Saha, D. Varghese, and P. B. Kumar, "On the generation and recovery of interface traps in MOSFETs subjected to NBTI, FN, and HCI stress," *Ieee Transactions on Electron Devices*, vol. 53, pp. 1583-1592, Jul 2006.
- [114] D. Varghese, H. Kufluoglu, V. Reddy, H. Shichijo, D. Mosher, S. Krishnan, and M. A. Alam, "OFF-State degradation in drain-extended NMOS transistors: Interface damage and correlation to dielectric breakdown," *Ieee Transactions on Electron Devices*, vol. 54, pp. 2669-2678, Oct 2007.

- [115] T. Grasser, B. Kaczer, and W. Goes, "An energy-level perspective of bias temperature instability," in *IEEE International Reliability Physics Symposium*, 2008, pp. 28-38.
- [116] A. Haggag, W. McMahon, K. Hess, K. Cheng, J. Lee, and J. Lyding, "High-performance chip reliability from short-time-tests - Statistical models for optical interconnect and HCI/TDDDB/NBTI deep-submicron transistor failures," *IEEE International Reliability Physics Symposium*, pp. 271-279, 2001.
- [117] K. O. Jeppson and C. M. Svensson, "Negative Bias Stress of Mos Devices at High Electric-Fields and Degradation of Mnos Devices," *Journal of Applied Physics*, vol. 48, pp. 2004-2014, 1977.
- [118] S. Mahapatra and M. A. Alam, "Defect generation in p-MOSFETs under negative-bias stress: An experimental perspective," *Ieee Transactions on Device and Materials Reliability*, vol. 8, pp. 35-46, Mar 2008.
- [119] A. E. Islam, H. Kufluoglu, D. Varghese, S. Mahapatra, and M. A. Alam, "Recent Issues in Negative Bias Temperature Instability: Initial Degradation, Field-Dependence of Interface Trap Generation, Hole Trapping Effects, and Relaxation," *Ieee Transactions on Electron Devices*, vol. 54, pp. 2143-2154, 2007.
- [120] A. S. Grove, *Physics and Technology of Semiconductor Devices*: John Wiley and Sons, 1971.
- [121] K. L. Brower, "Kinetics of H₂ Passivation of P_b Centers at the (111) Si-SiO₂ Interface," *Physical Review B*, vol. 38, pp. 9657-9666, Nov 15 1988.
- [122] E. Cartier, J. H. Stathis, and D. A. Buchanan, "Passivation and Depassivation of Silicon Dangling Bonds at the Si/SiO₂ Interface by Atomic-Hydrogen," *Applied Physics Letters*, vol. 63, pp. 1510-1512, Sep 13 1993.
- [123] S. Ogawa and N. Shiono, "Generalized Diffusion-Reaction Model for the Low-Field Charge-Buildup Instability at the Si-SiO₂ Interface," *Physical Review B*, vol. 51, pp. 4218-4230, FEB 15 1995.
- [124] D. K. Schroder, "Negative bias temperature instability: What do we understand?," *Microelectronics Reliability*, vol. 47, pp. 841-852, JUN 2007.
- [125] M. A. Alam, H. Kufluoglu, D. Varghese, and S. Mahapatra, "A comprehensive model for PMOS NBTI degradation: Recent progress," *Microelectronics Reliability*, vol. 47, pp. 853-862, JUN 2007.
- [126] S. Kumar, C. H. Kim, and S. S. Sapatnekar, "An analytical model for Negative Bias Temperature Instability," in *International Conference on Computer-Aided Design*, 2006, p. 6D.1.
- [127] G. Gupta, S. Mahapatra, L. Madhav, D. Varghese, K. Z. Ahmed, and F. Nouri, "Interface-trap driven NBTI for ultrathin (EOT~12Å) plasma and thermal nitrided oxynitrides," in *IEEE International Reliability Physics Symposium*, 2006, pp. 731-732.

- [128] A. E. Islam, G. Gupta, S. Mahapatra, A. Krishnan, K. Z. Ahmed, F. Nouri, A. Oates, and M. A. Alam, "Gate leakage vs. NBTI in Plasma Nitrided Oxides: Characterization, Physical Principles, and Optimization," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2006, pp. 329-332.
- [129] S. Mahapatra, K. Ahmed, D. Varghese, A. E. Islam, G. Gupta, L. Madhav, D. Saha, and M. A. Alam, "On the Physical Mechanism of NBTI in Silicon Oxynitride p-MOSFETs: Can Differences in Insulator Processing Conditions Resolve the Interface Trap Generation versus Hole Trapping Controversy?," in *IEEE International Reliability Physics Symposium*, 2007, pp. 1-9.
- [130] D. Varghese, D. Saha, S. Mahapatra, K. Ahmed, F. Nouri, and M. A. Alam, "On the Dispersive versus Arrhenius Temperature Activation of NBTI time evolution in Plasma Nitrided Gate Oxides: Measurements, Theory, and Implications," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2005, pp. 684 - 687.
- [131] S. Chakravarthi, A. T. Krishnan, V. Reddy, and S. Krishnan, "Probing negative bias temperature instability using a continuum numerical framework: Physics to real world operation," *Microelectronics Reliability*, vol. 47, pp. 863-872, Jun 2007.
- [132] H. Kufluoglu and M. A. Alam, "A Generalized Reaction-Diffusion Model with Explicit H-H₂ Dynamics for Negative Bias Temperature Instability (NBTI) Degradation," *Ieee Transactions on Electron Devices*, vol. 54, pp. 1101-1107, MAY 2007.
- [133] A. E. Islam, J. H. Lee, W. H. Wu, A. Oates, and M. A. Alam, "Universality of Interface Trap Generation and Its Impact on I_D Degradation in Strained/Unstrained PMOS Devices During NBTI Stress," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2008, pp. 107-110.
- [134] L. Tsetseris, X. J. Zhou, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, "Physical mechanisms of negative-bias temperature instability," *Applied Physics Letters*, vol. 86, p. 142103, APR 4 2005.
- [135] L. Tsetseris, X. J. Zhou, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, "Hydrogen-Related Instabilities in MOS Devices Under Bias Temperature Stress," *Ieee Transactions on Device and Materials Reliability*, vol. 7, pp. 502-508, 2007.
- [136] V. Huard, M. Denais, and C. Parthasarathy, "NBTI degradation: From physical mechanisms to modelling," *Microelectronics Reliability*, vol. 46, pp. 1-23, Jan 2006.
- [137] M. Denais, A. Bravaix, V. Huard, C. Parthasarathy, G. Ribes, F. Perrier, Y. Rey-Tauriac, and N. Revil, "On-the-fly Characterization of NBTI in Ultra-thin Gate Oxide PMOSFET's," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2004, pp. 109-112.

- [138] V. D. Maheta, C. Olsen, K. Ahmed, and S. Mahapatra, "The impact of nitrogen engineering in silicon oxynitride gate dielectric on negative-bias temperature instability of p-MOSFETs: A study by ultrafast on-the-fly I-DLIN technique," *Ieee Transactions on Electron Devices*, vol. 55, pp. 1630-1638, Jul 2008.
- [139] S. Rangan, N. Mielke, and E. C. C. Yeh, "Universal recovery behavior of negative bias temperature instability [PMOSFETs]," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2003, pp. 341-344.
- [140] H. Reisinger, O. Blank, W. Heinrigs, A. Muhlhoff, W. Gustin, and C. Schlunder, "Analysis of NBTI degradation- and recovery- behavior based on ultra fast V_T Measurements," in *IEEE International Reliability Physics Symposium*, 2006, pp. 448 - 453.
- [141] C. Shen, M.-F. Li, C. E. Foo, T. Yang, D. M. Huang, A. Yap, G. S. Samudra, and Y.-C. Yeo, "Characterization and Physical Origin of Fast V_{th} Transient in NBTI of pMOSFETs with SiON Dielectric," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2006, pp. 333-336.
- [142] A. T. Krishnan, S. Chakravarthi, P. Nicollian, V. Reddy, and S. Krishnan, "Negative bias temperature instability mechanism: The role of molecular hydrogen," *Applied Physics Letters*, vol. 88, 153518, APR 10 2006.
- [143] A. T. Krishnan, C. Chancellor, S. Chakravarthi, P. E. Nicollian, V. Reddy, A. Varghese, R. B. Khamankar, S. Krishnan, and L. Levitov, "Material dependence of hydrogen diffusion: implications for NBTI degradation," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2005, pp. 688-691.
- [144] S. Mahapatra and M. A. Alam, "A predictive reliability model for PMOS bias temperature degradation," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 505-508, 2002.
- [145] M. Ershov, S. Saxena, H. Karbasi, S. Winters, S. Minehane, J. Babcock, R. Lindley, P. Clifton, M. Redford, and A. Shibkov, "Dynamic recovery of negative bias temperature instability in p-type metal-oxide-semiconductor field-effect transistors," *Applied Physics Letters*, vol. 83, pp. 1647-1649, AUG 25 2003.
- [146] S. N. Rashkeev, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, "Defect generation by hydrogen at the Si-SiO₂ interface," *Physical Review Letters*, vol. 8716, pp. art. no.-165506, Oct 15 2001.
- [147] C. G. Van de Walle and B. R. Tuttle, "Microscopic theory of hydrogen in silicon devices," *Ieee Transactions on Electron Devices*, vol. 47, pp. 1779-1786, Oct 2000.
- [148] S. Zafar, "Statistical mechanics based model for negative bias temperature instability induced degradation," *Journal of Applied Physics*, vol. 97, pp. -, May 15 2005.

- [149] J. Kakalios, R. A. Street, and W. B. Jackson, "Stretched-Exponential Relaxation Arising from Dispersive Diffusion of Hydrogen in Amorphous-Silicon," *Physical Review Letters*, vol. 59, pp. 1037-1040, AUG 31 1987.
- [150] I. A. Shkrob and A. D. Trifunac, "Time-resolved EPR of spin-polarized mobile H atoms in amorphous silica: The involvement of small polarons," *Physical Review B*, vol. 54, pp. 15073-15078, DEC 1 1996.
- [151] D. Monroe and M. A. Kastner, "Generalizations of Multiple Trapping," *Philosophical Magazine B-Physics of Condensed Matter Statistical Mechanics Electronic Optical and Magnetic Properties*, vol. 47, pp. 605-620, 1983.
- [152] T. Grasser, W. Gos, and B. Kaczer, "Modeling of Dispersive Transport in the Context of Negative Bias Temperature Instability," in *IEEE International Integrated Reliability Workshop*, 2006, pp. 5-10.
- [153] T. Grasser, W. Gos, and B. Kaczer, "Dispersive transport and negative bias temperature instability: Boundary conditions, initial conditions, and transport models," *Ieee Transactions on Device and Materials Reliability*, vol. 8, pp. 79-97, Mar 2008.
- [154] E. N. Kumar, V. D. Maheta, S. Purawat, A. E. Islam, C. Olsen, K. Z. Ahmed, M. A. Alam, and S. Mahapatra, "Material Dependence of NBTI Physical Mechanism in Silicon Oxynitride (SiON) p-MOSFETs: A Comprehensive Study by Ultra-Fast On-The-Fly (UF-OTF) I_{DLIN} Technique," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2007, pp. 809-812.
- [155] T. Yang, M. F. Li, C. Shen, C. H. Ang, C. Zhu, Y. C. Yeo, G. Samudra, S. C. Rustagi, and M. B. Yu, "Fast and slow dynamic NBTI components in p-MOSFET with SiON dielectric and their impact on device life-time and circuit application," in *Symposium on VLSI Technology*, 2005, pp. 92 - 93.
- [156] S. Mahapatra, P. B. Kumar, and M. A. Alam, "Investigation and modeling of interface and bulk trap generation during negative bias temperature instability of p-MOSFETs," *Ieee Transactions on Electron Devices*, vol. 51, pp. 1371-1379, SEP 2004.
- [157] B. J. Fishbein, J. T. Watt, and J. D. Plummer, "Time Resolved Annealing of Interface Traps in Polysilicon Gate Metal-Oxide-Silicon Capacitors," *Journal of the Electrochemical Society*, vol. 134, pp. 674-681, MAR 1987.
- [158] N. H. Nickel, W. B. Jackson, and J. Walker, "Hydrogen migration in polycrystalline silicon," *Physical Review B*, vol. 53, pp. 7750 LP - 7761, 1996/03/15/ 1996.
- [159] C. G. V. d. Walle, "Theory of Hydrogen-related levels in semiconductors and oxides," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2005, pp. 400-403.

- [160] H. Kufluoglu and M. A. Alam, "A Generalized Reaction-Diffusion Model with Explicit H-H₂ Dynamics for Negative Bias Temperature Instability (NBTI) Degradation," in *accepted for IEEE Trans. on ELec. Dev.*, 2007.
- [161] A. E. Islam, H. Kufluoglu, D. Varghese, and M. A. Alam, "Critical analysis of short-term negative bias temperature instability measurements: Explaining the effect of time-zero delay for on-the-fly measurements," *Applied Physics Letters*, vol. 90, 083505, FEB 19 2007.
- [162] J. B. Yang, T. P. Chen, S. S. Tan, and L. Chan, "Analytical reaction-diffusion model and the modeling of nitrogen-enhanced negative bias temperature instability," *Applied Physics Letters*, vol. 88, 172109, APR 24 2006.
- [163] J. Crank, *The Mathematics of Diffusion*, 2nd ed.: Oxford University Press, 1980.
- [164] P. E. Nicollian, A. T. Krishnan, C. Bowen, S. Chakravarthi, C. A. Chancellor, and R. B. Khamankar, "The roles of hydrogen and holes in trap generation and breakdown in ultra-thin SiON dielectrics," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 403-406, 2005.
- [165] A. E. Islam, E. N. Kumar, H. Das, S. Purawat, V. D. Maheta, H. Aono, E. Murakami, S. Mahapatra, and M. A. Alam, "Theory and Practice of On-the-fly and Ultra-fast V_T Measurements for NBTI Degradation: Challenges and Opportunities," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2007, pp. 805-808.
- [166] S. Mahapatra, V. D. Maheta, A. E. Islam, and M. A. Alam, "Isolation of NBTI Stress Generated Interface Trap and Hole-Trapping Components in PNO p-MOSFETs," in *Ieee Transactions on Electron Devices*, 2009.
- [167] M. A. Alam, "A Critical Examination of the Mechanics of Dynamic NBTI for PMOSFETs," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2003, pp. 345 - 348.
- [168] T. Grasser, W. Gos, V. Sverdlov, and B. Kaczer, "The Universality of NBTI Relaxation and its Implications for Modeling and Characterization," in *IEEE International Reliability Physics Symposium*, 2007, pp. 268-280.
- [169] C. R. Parthasarathy, M. Denais, V. Huard, G. Ribes, E. Vincent, and A. Bravaix, "New insights into recovery characteristics post NBTI stress," in *IEEE International Reliability Physics Symposium*, 2006, pp. 471-477.
- [170] H. Reisinger, O. Blank, W. Heinrigs, W. Gustin, and C. Schlunder, "A comparison of very fast to very slow components in degradation and recovery due to NBTI and bulk hole trapping to existing physical models," *Ieee Transactions on Device and Materials Reliability*, vol. 7, pp. 119-129, MAR 2007.
- [171] Y. Mitani, H. Satake, and A. Toriumi, "Influence of nitrogen on negative bias temperature instability in ultrathin SiON," *Ieee Transactions on Device and Materials Reliability*, vol. 8, pp. 6-13, Mar 2008.

- [172] N. Kimizuka, T. Yamamoto, T. Mogami, K. Yamaguchi, K. Imai, and T. Horiuchi, "The impact of bias temperature instability for direct-tunneling ultra-thin gate oxide on MOSFET scaling," in *Symposium on VLSI Technology*, 1999, pp. 73 - 74.
- [173] S. Mahapatra, P. B. Kumar, and M. A. Alam, "A new observation of enhanced bias temperature instability in thin gate oxide p-MOSFET," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2003, pp. 337 - 340.
- [174] S. Tsujikawa, T. Mine, K. Watanabe, Y. Shimamoto, R. Tsuchiya, K. Ohnishi, T. Onai, J. Yugami, and S. Kimura, "Negative bias temperature instability of pMOSFETs with ultra-thin SiON gate dielectrics," *IEEE International Reliability Physics Symposium*, pp. 183-188, 2003.
- [175] C. L. Chen, Y. M. Lin, C. J. Wang, and K. Wu, "A new finding on NBTI lifetime model and an investigation on NBTI degradation characteristic for 1.2nm ultra thin oxide," in *IEEE International Reliability Physics Symposium*, 2005, pp. 704 - 705.
- [176] H. Aono, E. Murakami, K. Okuyama, A. Nishida, M. Minami, Y. Ooji, and K. Kubota, "Modeling of NBTI degradation and its impact on electric field dependence of the lifetime," in *IEEE International Reliability Physics Symposium*, 2004, pp. 23 - 27.
- [177] J. W. McPherson and C. H. Mogul, "Underlying physics of the thermochemical E model in describing low-field time-dependent dielectric breakdown in SiO₂ thin films," *Journal of Applied Physics*, vol. 84, pp. 1513-1523, AUG 1 1998.
- [178] A. Ghetti, A. Hamad, P. J. Silverman, H. Vaidya, and N. Zhao, "Self-consistent simulation of quantization effects and tunneling current in ultra-thin gate oxide MOS devices," in *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 1999, pp. 239 - 242.
- [179] H. M. Nayfeh, J. L. Hoyt, and D. A. Antoniadis, "A physically based analytical model for the threshold voltage of strained-Si n-MOSFETs," *Ieee Transactions on Electron Devices*, vol. 51, pp. 2069-2072, Dec 2004.
- [180] L. Pauling, *The Nature of The Chemical Bond*, 3rd ed.: Cornell University Press, 1960.
- [181] H. Z. Massoud, "Charge-Transfer Dipole-Moments at the Si-SiO₂ Interface," *Journal of Applied Physics*, vol. 63, pp. 2000-2005, MAR 15 1988.
- [182] D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt, and G. Timp, "The electronic structure at the atomic scale of ultrathin gate oxides," *Nature*, vol. 399, pp. 758-761, JUN 24 1999.
- [183] L. F. Register, E. Rosenbaum, and K. Yang, "Analytic model for direct tunneling current in polycrystalline silicon-gate metal-oxide-semiconductor devices," *Applied Physics Letters*, vol. 74, pp. 457-459, JAN 18 1999.

- [184] A. E. Islam, G. Gupta, K. Z. Ahmed, S. Mahapatra, and M. A. Alam, "Optimization of gate leakage and NBTI for plasma-nitrided gate oxides by numerical and analytical models," *Ieee Transactions on Electron Devices*, vol. 55, pp. 1143-1152, May 2008.
- [185] T. Krishnamohan, C. Jungermann, D. Kim, E. Ungersboeck, S. Selberherr, P. Wong, Y. Nishi, and K. Saraswat, "Theoretical investigation of performance in uniaxially- and biaxially-strained Si, SiGe and Ge double-gate p-MOSFETs," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 681-684, 2006.
- [186] K. Uchida, T. Krishnamohan, K. C. Saraswat, and Y. Nishi, "Physical mechanisms of electron mobility enhancement in uniaxial stressed MOSFETs and impact of uniaxial stress engineering in ballistic regime," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 135-138, 2005.
- [187] S. E. Thompson, G. Sun, K. Wu, J. Lim, and T. Nishida, "Key differences for process-induced uniaxial vs. substrate-induced biaxial stressed Si and Ge channel MOSFETs," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 221-224, 2004.
- [188] J. L. Hoyt, H. M. Nayfeh, S. Eguchi, I. Aberg, G. Xia, T. Drake, E. A. Fitzgerald, and D. A. Antoniadis, "Strained silicon MOSFET technology," *International Electron Devices 2002 Meeting, Technical Digest*, pp. 23-26, 2002.
- [189] Y. F. Zhang and J. Singh, "Channel effective mass and interfacial effects in Si and SiGe metal-oxide-semiconductor field effect transistor: A charge control model study," *Journal of Applied Physics*, vol. 83, pp. 4264-4271, Apr 15 1998.
- [190] A. E. Islam, H. Kufluoglu, D. Varghese, S. Mahapatra, and M. A. Alam, "Recent Issues in Negative Bias Temperature Instability: Initial Degradation, Field-Dependence of Interface Trap Generation, Hole Trapping Effects, and Relaxation," *Ieee Transactions on Electron Devices*, vol. 54, pp. 2142-2153, 2007.
- [191] A. E. Islam, G. Gupta, S. Mahapatra, A. Krishnan, K. Ahmed, F. Nouri, A. Oates, and M. A. Alam, "Gate leakage vs. NBTI in Plasma Nitrided Oxides: Characterization, Physical Principles, and Optimization," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2006, pp. 329-332.
- [192] S. S. Tan, T. P. Chen, C. H. Ang, and L. Chan, "Mechanism of nitrogen-enhanced negative bias temperature instability in pMOSFET," *Microelectronics Reliability*, vol. 45, pp. 19-30, JAN 2005.
- [193] S. E. Thompson, G. Y. Sun, Y. S. Choi, and T. Nishida, "Uniaxial-process-induced strained-Si: Extending the CMOS roadmap," *Ieee Transactions on Electron Devices*, vol. 53, pp. 1010-1020, May 2006.
- [194] J. R. Shih and K. Wu, "Reliability considerations of strained silicon on relaxed silicon-germanium (SiGe) substrate," in *IEEE International Reliability Physics Symposium*, 2005, pp. 403 - 408.

- [195] A. Shickova, B. Kaczer, P. Verheyen, G. Eneman, E. S. Andres, M. Jurczak, P. Absil, H. Maes, and G. Groeseneken, "Negligible effect of process-induced strain on intrinsic NBTI behavior," *Ieee Electron Device Letters*, vol. 28, pp. 242-244, Mar 2007.
- [196] T. Irisawa, T. Numata, E. Toyoda, N. Hirashita, T. Tezuka, N. Sugiyama, and S. Takagi, "Physical understanding of strain effects on gate oxide reliability of MOSFETs," *Symposium on VLSI Technology*, pp. 36-37, 2007.
- [197] H. S. Rhee, H. Lee, T. Ueno, D. S. Shin, S. H. Lee, Y. Kim, A. Samoilov, P.-O. Hansson, M. Kim, H. S. Kim, and N.-I. Lee, "Negative Bias Temperature Instability of Carrier-Transport Enhanced pMOSFET with Performance Boosters," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2005, pp. 692-695.
- [198] G. Giusi, F. Crupi, E. Simoen, G. Eneman, and M. Jurczak, "Performance and reliability of strained-silicon nMOSFETs with SiN cap layer," *Ieee Transactions on Electron Devices*, vol. 54, pp. 78-82, Jan 2007.
- [199] C. Y. Lu, H. C. Lin, Y. F. Chang, and T. Y. Huang, "Device characteristics and aggravated negative bias temperature instability in p-channel metal-oxide-semiconductor field-effect transistors with uniaxial compressive strain," *Japanese Journal of Applied Physics Part 1-Regular Papers Brief Communications & Review Papers*, vol. 45, pp. 3064-3069, Apr 2006.
- [200] C. S. Lu, L. Horng-Chih, and Y. J. Lee, "Improvement of negative-bias-temperature instability in SiN-capped p-channel metal-oxide-semiconductor field-effect transistors using ultrathin HfO₂ buffer layer," *Journal of the Electrochemical Society*, vol. 154, pp. H1036-H1040, 2007.
- [201] C. H. Liu and T. M. Pan, "Hot carrier and negative-bias temperature instability reliabilities of strained-si MOSFETs," *Ieee Transactions on Electron Devices*, vol. 54, pp. 1799-1803, Jul 2007.
- [202] D. S. Ang, S. Wang, G. A. Du, and Y. Z. Hu, "A consistent deep-level hole trapping model for negative bias temperature instability," *Ieee Transactions on Device and Materials Reliability*, vol. 8, pp. 22-34, Mar 2008.
- [203] A. E. Islam, V. D. Maheta, H. Das, S. Mahapatra, and M. A. Alam, "Mobility Degradation Due to Interface Traps in Plasma Oxinitride PMOS Devices," in *IEEE International Reliability Physics Symposium*, 2008, pp. 87-96.
- [204] Y. Sun, S. E. Thompson, and T. Nishida, "Physics of strain effects in semiconductors and metal-oxide-semiconductor field-effect transistors," *Journal of Applied Physics*, vol. 101, pp. -, May 15 2007.
- [205] M. Agostinelli, S. Pae, W. Yang, C. Prasad, D. Kenckel, S. Ramey, E. Snyder, S. Kashyap, and M. Jones, "Random charge effects for PMOS NBTI in ultra-small gate area devices," *IEEE International Reliability Physics Symposium*, pp. 529-532, 2005.

- [206] G. La Rosa, W. L. Ng, S. Rauch, R. Wong, and J. Sudijono, "Impact of NBTI induced statistical variation to SRAM cell stability," *IEEE International Reliability Physics Symposium*, pp. 274-282, 2006.
- [207] S. E. Rauch, "Review and reexamination of reliability effects related to NBTI-induced statistical variations," *Ieee Transactions on Device and Materials Reliability*, vol. 7, pp. 524-530, Dec 2007.
- [208] V. Huard, C. Parthasarathy, C. Guerin, T. Valentin, E. Pion, M. Mammasse, N. Planes, and L. Camus, "NBTI degradation: From transistor to SRAM arrays," *IEEE International Reliability Physics Symposium*, pp. 289-300, 2008.
- [209] B. Vaidyanathan, A. S. Oates, Y. Xie, and Y. Wang, "NBTI-Aware Statistical Circuit Delay Assessment," *ISQED: International Symposium on Quality Electronic Design*, pp. 13-18, 2009.
- [210] T. Fischer, E. Amirante, P. Huber, K. Hofmann, M. Ostermayr, and D. Schmitt-Landsiedel, "A 65 nm test structure for SRAM device variability and NBTI statistics," *Solid-State Electronics*, vol. 53, pp. 773-778, Jul 2009.
- [211] S. Das, H. Vikalo, and A. Hassibi, "On scaling laws of biosensors: A stochastic approach," *Journal of Applied Physics*, vol. 105, p. 102021, May 15 2009.
- [212] H. S. Wilf, *Generatingfunctionology*, 2nd ed. Boston: Academic Press, 1994.
- [213] K. W. Kehr, K. Mussawisade, and T. Wichmann, "Mean first-passage times and hopping mobility of particles under bias in nonsymmetric potentials," *Anomalous Diffusion: From Basics to Applications*, vol. 519, pp. 35-44, 1999.
- [214] K. P. N. Murthy and K. W. Kehr, "Mean 1st-Passage Time of Random-Walks on a Random Lattice," *Physical Review A*, vol. 40, pp. 2082-2087, Aug 15 1989.
- [215] B. J. Zwolinski, H. Eyring, and C. E. Reese, "Diffusion and Membrane Permeability .1.," *Journal of Physical and Colloid Chemistry*, vol. 53, pp. 1426-1453, 1949.
- [216] A. Azzalini, "The skew-normal distribution and related multivariate families," *Scandinavian Journal of Statistics*, vol. 32, pp. 159-188, Jun 2005.
- [217] M. F. Bukhori, S. Roy, and A. Asenov, "Statistical aspects of reliability in bulk MOSFETs with multiple defect states and random discrete dopants," *Microelectronics Reliability*, vol. 48, pp. 1549-1552, Aug-Sep 2008.
- [218] M. F. Bukhori, A. R. Brown, S. Roy, and A. Asenov, "Simulation of Statistical Aspects of Reliability in Nano CMOS Transistors," in *IEEE International Integrated Reliability Workshop*, 2009, pp. 82-85.
- [219] L. Lundkvist, I. Lundstrom, and C. Svensson, "Discharge of MNOS Structures," *Solid-State Electronics*, vol. 16, pp. 811-&, 1973.
- [220] R. A. Williams and M. M. E. Beguwala, "Effect of Electrical-Conduction of Si₃N₄ on Discharge of MNOS Memory Transistors," *Ieee Transactions on Electron Devices*, vol. 25, pp. 1019-1023, 1978.

- [221] G. L. Heyns and H. E. Maes, "A New Model for the Discharge Behavior of Metal Nitride Oxide Silicon (MNOS) Nonvolatile Memory Devices," *Applied Surface Science*, vol. 30, pp. 153-159, Oct 1987.
- [222] F. P. Heiman and G. Warfield, "Effects of Oxide Traps on Mos Capacitance," *Ieee Transactions on Electron Devices*, vol. Ed12, pp. 167-&, 1965.
- [223] P. Rossel, H. Martinot, and D. Esteve, "Temporary Instability of Mos Transistors .2. Dissymmetrical Behaviour," *Solid-State Electronics*, vol. 13, pp. 425-&, 1970.
- [224] I. Lundstrom, S. Christensson, and C. Svensson, "Carrier trapping hysteresis in MOS transistor," *Phys. Status Solid (A)*, vol. 1, pp. 395-407, 1970.
- [225] V. Huard and M. Denais, "Hole trapping effect on methodology for DC and AC negative bias temperature instability measurements in pMOS transistors," *IEEE International Reliability Physics Symposium*, pp. 40-45, 2004.
- [226] T. Yang, C. Shen, M. F. Li, C. H. Ang, C. X. Zhu, Y. C. Yeo, G. Samudra, S. C. Rustagi, M. B. Yu, and D. L. Kwong, "Fast DNBTI components in p-MOSFET with SiON dielectric," *Ieee Electron Device Letters*, vol. 26, pp. 826-828, NOV 2005.
- [227] M. F. Li, G. Chen, C. Shen, X. P. Wang, H. Y. Yu, Y. C. Yeo, and D. L. Kwong, "Dynamic bias-temperature instability in ultrathin SiO₂ and HfO₂ metal-oxide-semiconductor field effect transistors and its impact on device lifetime," *Japanese Journal of Applied Physics Part 1-Regular Papers Short Notes & Review Papers*, vol. 43, pp. 7807-7814, Nov 2004.
- [228] C. Shen, M. F. Li, X. P. Wang, H. Y. Yu, Y. P. Feng, A. T. L. Lim, Y. C. Yeo, D. S. H. Chan, and D. L. Kwong, "Negative U traps in HfO₂ gate dielectrics and frequency dependence of dynamic BTI in MOSFETs," *Ieee International Electron Devices Meeting 2004, Technical Digest*, pp. 733-736, 2004.
- [229] W. Goes, M. Karner, V. Sverdlov, and T. Grasser, "A rigorous model for trapping and detrapping in thin gate dielectrics," *Proceedings of International Symposium on the Physical & Failure Analysis of Integrated Circuits (IPFA)*, pp. 249-254, 2008.
- [230] T. L. Tewksbury and H. S. Lee, "Characterization, Modeling, and Minimization of Transient Threshold Voltage Shifts in Mosfets," *Ieee Journal of Solid-State Circuits*, vol. 29, pp. 239-252, Mar 1994.
- [231] A. Shanware, M. R. Visokay, J. J. Chambers, A. L. P. Rotondaro, H. Bu, M. J. Bevan, R. Khamankar, S. Aur, P. E. Nicollian, J. McPherson, and L. Colombo, "Evaluation of the positive biased temperature stress stability in HfSiON gate dielectrics," *IEEE International Reliability Physics Symposium*, pp. 208-213, 2003.
- [232] S. Zafar, A. Callegari, E. Gusev, and M. V. Fischetti, "Charge trapping in high K gate dielectric stacks," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 517-520, 2002.

- [233] S. Zafar, A. Kumar, E. Gusev, and E. Cartier, "Threshold voltage instabilities in high-k gate dielectric stacks," *Ieee Transactions on Device and Materials Reliability*, vol. 5, pp. 45-64, MAR 2005.
- [234] X. Garros, J. Mitard, C. Leroux, G. Reimbold, and F. Boulanger, "In-depth Analysis of V_T Instabilities in HfO_2 Technologies by Charge Pumping Measurements and Electrical Modeling," in *IEEE International Reliability Physics Symposium*, 2007, pp. 61-66.
- [235] R. N. Hall, "Electron-Hole Recombination in Germanium," *Physical Review*, vol. 87, pp. 387-387, 1952.
- [236] W. Shockley and W. T. Read, "Statistics of the Recombinations of Holes and Electrons," *Physical Review*, vol. 87, pp. 835-842, 1952.
- [237] T. H. Ning, "Capture Cross-Section and Trap Concentration of Holes in Silicon Dioxide," *Journal of Applied Physics*, vol. 47, pp. 1079-1081, 1976.
- [238] S. Rauf, S. Lim, and P. L. G. Ventzek, "Model for nitridation of nanoscale SiO_2 thin films in pulsed inductively coupled N-2 plasma," *Journal of Applied Physics*, vol. 98, pp. -, JUL 15 2005.
- [239] J. R. Shallenberger, D. A. Cole, and S. W. Novak, "Characterization of silicon oxynitride thin films by x-ray photoelectron spectroscopy," *Journal of Vacuum Science & Technology a-Vacuum Surfaces and Films*, vol. 17, pp. 1086-1090, JUL-AUG 1999.
- [240] C. Shen, M.-F. Li, C. E. Foo, T. Yang, D. M. Huang, A. Yap, G. S. Samudra, and Y.-C. Yeo, "Characterization and Physical Origin of Fast V_{th} Transient in NBTI of pMOSFETs with $SiON$ Dielectric," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2006, p. 12.5.1.
- [241] A. Neugroschel, G. Bersuker, R. Choi, C. Cochrane, P. Lenahan, D. Heh, C. Young, C. Y. Kang, B. H. Lee, and R. Jammy, "An accurate lifetime analysis methodology incorporating governing NBTI mechanisms in High-k/ SiO_2 gate stacks," *Proceedings of International Electron Devices Meeting*, pp. 53-56, 2006.
- [242] D. Heh, C. D. Young, and G. Bersuker, "Experimental evidence of the fast and slow charge trapping/detrapping processes in high-k dielectrics subjected to PBTI stress," *Ieee Electron Device Letters*, vol. 29, pp. 180-182, Feb 2008.
- [243] D. Kaplan, I. Solomon, and N. F. Mott, "Explanation of Large Spin-Dependent Recombination Effect in Semiconductors," *Journal De Physique Lettres*, vol. 39, pp. L51-L54, 1978.
- [244] D. J. Lepine, "Spin-Dependent Recombination on Silicon Surface," *Physical Review B*, vol. 6, pp. 436-&, 1972.
- [245] A. V. Barabanov, V. A. Lvov, and O. V. Tretyak, "About the mechanisms of spin-dependent recombination in semiconductors," *Physica Status Solidi B-Basic Research*, vol. 207, pp. 419-427, Jun 1998.

- [246] C. Boehme and K. Lips, "Electrical detection of spin coherence in silicon," *Physical Review Letters*, vol. 91, pp. -, Dec 12 2003.
- [247] F. Rong, E. H. Poindexter, M. Harmatz, W. R. Buchwald, and G. J. Gerardi, "Electrically Detected Magnetic-Resonance in P-N-Junction Diodes," *Solid State Communications*, vol. 76, pp. 1083-1086, Nov 1990.
- [248] C. Boehme and K. Lips, "Theory of time-domain measurement of spin-dependent recombination with pulsed electrically detected magnetic resonance," *Physical Review B*, vol. 68, pp. -, Dec 2003.
- [249] D. K. Schroder, *Semiconductor Material and Device Characterization*, 3rd ed.: Wiley-IEEE Press, 2006.
- [250] T. Grasser, P. J. Wagner, P. Hehenberger, W. Goes, and B. Kaczer, "A Rigorous Study of Measurement Techniques for Negative Bias Temperature Instability," *Ieee Transactions on Device and Materials Reliability*, vol. 8, pp. 526-535, Sep 2008.
- [251] A. Kerber, E. Cartier, L. Pantisano, M. Rosmeulen, R. Degraeve, T. Kauerauf, G. Groeseneken, H. E. Maes, and U. Schwalke, "Characterization of the V_T -instability in $\text{SiO}_2/\text{HfO}_2$ gate dielectrics," *IEEE International Reliability Physics Symposium*, pp. 41-45, 2003.
- [252] J. F. Zhang, Z. Ji, M. H. Chang, B. Kaczer, and G. Groeseneken, "Real V_{th} instability of pMOSFETs under practical operation conditions," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2007, pp. 817-820.
- [253] R. L. Vranich, B. Henderson, and M. Pepper, "Spin-Dependent Recombination in Irradiated Si/SiO₂ Device Structures," *Applied Physics Letters*, vol. 52, pp. 1161-1163, Apr 4 1988.
- [254] J. H. Stathis and L. Dori, "Fundamental Chemical Differences among Pb Defects on (111) and (100) Silicon," *Applied Physics Letters*, vol. 58, pp. 1641-1643, Apr 15 1991.
- [255] H. Miki, M. Noguchi, K. Yokogawa, B. W. Kim, K. Asada, and T. Sugano, "Electron and Hole Traps in SiO₂-Films Thermally Grown on Si Substrates in Ultra-Dry Oxygen," *Ieee Transactions on Electron Devices*, vol. 35, pp. 2245-2252, Dec 1988.
- [256] P. M. Lenahan and J. F. Conley, "Magnetic resonance studies of trapping centers in high-k dielectric films on silicon," *Ieee Transactions on Device and Materials Reliability*, vol. 5, pp. 90-102, Mar 2005.
- [257] M. A. Jupina and P. M. Lenahan, "A Spin Dependent Recombination Study of Radiation-Induced Defects at and near the Si/SiO₂ Interface," *Ieee Transactions on Nuclear Science*, vol. 36, pp. 1800-1807, Dec 1989.

- [258] S. Fujieda, Y. Miura, M. Saitoh, E. Hasegawa, S. Koyama, and K. Ando, "Interface defects responsible for negative-bias temperature instability in plasma-nitrided SiON/Si(100) systems," *Applied Physics Letters*, vol. 82, pp. 3677-3679, May 26 2003.
- [259] K. Lips, C. Boehme, and T. Ehara, "The impact of the electron spin on charge carrier recombination - The example of amorphous silicon," *Journal of Optoelectronics and Advanced Materials*, vol. 7, pp. 13-24, Feb 2005.
- [260] C. J. Cochrane, P. M. Lenahan, J. P. Campbell, G. Bersuker, and A. Neugroschel, "Negative bias stressing interface trapping centers in metal gate hafnium oxide field effect transistors using spin dependent recombination," *2006 IEEE International Integrated Reliability Workshop, Final Report*, pp. 11-15, 2006.
- [261] J. W. Gabrys, P. M. Lenahan, and W. Weber, "High-Resolution Spin-Dependent Recombination Study of Hot-Carrier Damage in Short-Channel Mosfets - Si-29 Hyperfine Spectra," *Microelectronic Engineering*, vol. 22, pp. 273-276, Aug 1993.
- [262] T. D. Mishima and P. M. Lenahan, "A spin-dependent recombination study of radiation-induced P b1 centers at the (001) Si/SiO₂ interface," *Ieee Transactions on Nuclear Science*, vol. 47, pp. 2249-2255, Dec 2000.
- [263] A. S. Grove and Fitzgera.Dj, "Surface Effects on P-N Junctions - Characteristics of Surface Space-Charge Regions under Non-Equilibrium Conditions," *Solid-State Electronics*, vol. 9, pp. 783-&, 1966.
- [264] F. C. Rong, W. R. Buchwald, E. H. Poindexter, W. L. Warren, and D. J. Keeble, "Spin-Dependent Shockley-Read Recombination of Electrons and Holes in Indirect-Band-Gap Semiconductor P-N-Junction Diodes," *Solid-State Electronics*, vol. 34, pp. 835-841, Aug 1991.
- [265] K. Y. Fu, "Mobility Degradation Due to the Gate Field in the Inversion Layer of Mosfets," *Electron Device Letters*, vol. 3, pp. 292-293, 1982.
- [266] S. Villa, A. L. Lacaita, L. M. Perron, and R. Bez, "Physically-based model of the effective mobility in heavily-doped n-MOSFET's," *Ieee Transactions on Electron Devices*, vol. 45, pp. 110-115, JAN 1998.
- [267] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, "On the Universality of Inversion Layer Mobility in Si MOSFETs .1. Effects of Substrate Impurity Concentration," *Ieee Transactions on Electron Devices*, vol. 41, pp. 2357-2362, DEC 1994.
- [268] N. Arora, *MOSFET Modeling for VLSI Simulation: Theory and Practice*: World Scientific, 2007.
- [269] J. F. Zhang and M. H. Chang, "An assessment of mobility variation during negative bias temperature stress," in *211th ECS Meeting*, 2007, p. 668.
- [270] B. H. Cheng and J. Woo, "A temperature-dependent MOSFET inversion layer carrier mobility model for device and circuit simulation," *Ieee Transactions on Electron Devices*, vol. 44, pp. 343-345, FEB 1997.

- [271] J. T. Watt, B. J. Fishbein, and J. D. Plummer, "Characterization of Surface Mobility in Mos Structures Containing Interfacial Cesium Ions," *Ieee Transactions on Electron Devices*, vol. 36, pp. 96-100, JAN 1989.
- [272] M. Lundstrom, *Fundamentals of Carrier Transport*: Cambridge University Press, 2000.
- [273] P. M. Zeitzoff, C. D. Young, G. A. Brown, and Y. Kim, "Correcting effective mobility measurements for the presence of significant gate leakage current," *Ieee Electron Device Letters*, vol. 24, pp. 275-277, Apr 2003.
- [274] W. J. Zhu, J. P. Han, and T. P. Ma, "Mobility measurement and degradation mechanisms of MOSFETs made with ultrathin high-k dielectrics," *Ieee Transactions on Electron Devices*, vol. 51, pp. 98-105, Jan 2004.
- [275] O. Tonomura, Y. Shimamoto, H. Miki, S. Saito, K. Torii, M. Hiratani, and J. Yugami, "Accurate evaluation of mobility in high gate-leakage-current MOSFETs by using a transmission-line model," *Ieee Transactions on Electron Devices*, vol. 51, pp. 1653-1658, Oct 2004.
- [276] K. M. Cao, W. C. Lee, W. Liu, X. Jin, P. Su, S. K. H. Fung, J. X. An, B. Yu, and C. Hu, "BSIM4 gate leakage model including source-drain partition," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 815-818, 2000.
- [277] C. Schlünder, M. Hoffmann, R.-P. Vollertsen, G. Schindler, W. Heinrigs, W. Gustin, and H. Reisinger, "A new smart V_{th} -extraction methodology considering recovery and mobility degradation due to NBTI," in *IEEE International Integrated Reliability Workshop*, 2007, p. 1.1.
- [278] B. Kaczer, T. Grasser, P. J. Roussel, J. Martin-Martinez, R. O'Connor, B. J. O'Sullivan, and G. Groeseneken, "Ubiquitous relaxation in BTI stressing - New evaluation and insights," in *IEEE International Reliability Physics Symposium*, 2008, pp. 20-27.
- [279] F. Stern, "Self-Consistent Results for N-Type Si Inversion Layers," *Physical Review B*, vol. 5, pp. 4891-4899, 1972.
- [280] N. Yang, W. K. Henson, J. R. Hauser, and J. J. Wortman, "Modeling study of ultrathin gate oxides using direct tunneling current and capacitance-voltage measurements in MOS devices," *Ieee Transactions on Electron Devices*, vol. 46, pp. 1464-1471, JUL 1999.
- [281] M. J. Vandort, P. H. Woerlee, and A. J. Walker, "A Simple-Model for Quantization Effects in Heavily-Doped Silicon Mosfets at Inversion Conditions," *Solid-State Electronics*, vol. 37, pp. 411-414, Mar 1994.
- [282] Y. Taur and T. Ning, *Fundamentals of Modern VLSI Devices*: Cambridge University Press, 1998.

- [283] M. J. Uren, V. Nayar, K. M. Brunson, C. J. Anthony, J. H. Stathis, and E. Cartier, "Interface state capture cross section measurements on vacuum annealed and radiation damaged Si : SiO₂ surfaces," *Journal of the Electrochemical Society*, vol. 145, pp. 683-689, Feb 1998.
- [284] S. Mahapatra, K. Ahmed, D. Varghese, A. E. Islam, G. Gupta, L. Madhav, D. Saha, and M. A. Alam, "On the Physical Mechanism of NBTI in Silicon Oxynitride p-MOSFETs: Can Differences in Insulator Processing Conditions Resolve the Interface Trap Generation versus Hole Trapping Controversy?," in *to appear in IRPS Tech. Dig.*, 2007.
- [285] D. Brisbin, J. Yang, S. Bahl, and C. Parker, "Enhanced PMOS NBTI degradation due to halo implant channeling," in *IEEE International Reliability Physics Symposium*, 2008, pp. 61-66.
- [286] V. Huard, C. R. Parthasarathy, C. Guerin, and M. Denais, "Physical Modeling of Negative Bias Temperature Instabilities for Predictive Extrapolation," in *IEEE International Reliability Physics Symposium*, 2006, pp. 733-734.
- [287] H. Reisinger, U. Brunner, W. Heinrigs, W. Gustin, and C. Schlunder, "A comparison of fast methods for measuring NBTI degradation," *Ieee Transactions on Device and Materials Reliability*, vol. 7, pp. 531-539, DEC 2007.
- [288] T. Grasser, P.-J. Wagner, P. Hehenberger, W. Goes, and B. Kaczer, "A Rigorous Study of Measurement Techniques for Negative Bias Temperature Instability," in *IEEE International Integrated Reliability Workshop*, 2007, p. 1.2.
- [289] R. Fernández, B. Kaczer, A. Nackaerts, S. Demuynck, R. Rodríguez, M. Nafría, and G. Groeseneken, "AC NBTI studied in the 1 Hz – 2 GHz range on dedicated on-chip CMOS circuits," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2006, pp. 337-340.
- [290] A. E. Islam, E. N. Kumar, H. Das, S. Purawat, V. Maheta, H. Aono, E. Murakami, S. Mahapatra, and M. A. Alam, "Theory and Practice of On-the-fly and Ultra-fast V_T Measurements for NBTI Degradation: Challenges and Opportunities," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2007, pp. 805-808.
- [291] V. Huard, C. Parthasarathy, N. Rallet, C. Guerin, M. Mammase, D. Barge, and C. Ouvrard, "New characterization and modeling approach for NBTI degradation from transistor to product level," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2007, pp. 797-800.
- [292] T. Grasser, B. Kaczer, R. Hehenberger, W. Gos, R. O'Connor, H. Reisinger, W. Gustin, and C. Schlunder, "Simultaneous extraction of recoverable and permanent components contributing to bias-temperature instability," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 801-804, 2007.
- [293] T. Grasser, B. Kaczer, W. Goes, T. Aichinger, P. Hehenberger, and M. Nelhiebel, "A Two-Stage Model for Negative Bias Temperature Instability," *IEEE International Reliability Physics Symposium*, pp. 33-44, 2009.

- [294] V. D. Maheta, E. N. Kumar, S. Purawat, C. Olsen, K. Ahmed, and S. Mahapatra, "Development of an Ultrafast On-the-Fly I-DLIN Technique to Study NBTI in Plasma and Thermal Oxynitride p-MOSFETs," *Ieee Transactions on Electron Devices*, vol. 55, pp. 2614-2622, Oct 2008.
- [295] D. Ielmini, M. Manigrasso, F. Gattel, and M. G. Valentini, "A New NBTI Model Based on Hole Trapping and Structural Relaxation in MOS Dielectrics," *Ieee Transactions on Electron Devices*, vol. 56, pp. 1943-1952, Sep 2009.
- [296] J. H. Lee, W. H. Wu, A. E. Islam, M. A. Alam, and A. S. Oates, "Separation method of hole trapping and interface trap generation and their roles in NBTI reaction-diffusion model," *IEEE International Reliability Physics Symposium*, pp. 745-746, 2008.
- [297] K. Sakuma, D. Matsushita, K. Muraoka, and Y. Mitani, "Investigation of nitrogen-originated NBTI mechanism in sion with high-nitrogen concentration," *IEEE International Reliability Physics Symposium*, pp. 454-460, 2006.
- [298] A. Haggag, G. Anderson, S. Parihar, D. Burnett, G. Abeln, J. Higman, and M. Moosa, "Understanding SRAM high-temperature-operating-life NBTI: Statistics and permanent vs recoverable damage," *IEEE International Reliability Physics Symposium*, pp. 452-456, 2007.
- [299] S. Mahapatra, V. D. Maheta, A. E. Islam, and M. A. Alam, "Isolation of NBTI Stress Generated Interface Trap and Hole-Trapping Components in PNO p-MOSFETs," *Ieee Transactions on Electron Devices*, vol. 56, pp. 236-242, Feb 2009.
- [300] S. Deora, V. D. Maheta, A. E. Islam, M. A. Alam, and S. Mahapatra, "A Common Framework of NBTI Generation and Recovery in Plasma-Nitrided SiON p-MOSFETs," *Ieee Electron Device Letters*, vol. 30, pp. 978-980, Sep 2009.
- [301] K. Hess, A. Haggag, W. McMahon, B. Fischer, K. Cheng, J. Lee, and J. Lyding, "Simulation of Si-SiO₂ Defect Generation in CMOS Chips: From Atomistic Structure to Chip Failure Rates," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2000, pp. 93-96.
- [302] B. Tuttle, "Energetics and diffusion of hydrogen in SiO₂," *Physical Review B*, vol. 61, pp. 4417-4420, FEB 15 2000.
- [303] G. Lucovsky, "Ultrathin nitrided gate dielectrics: Plasma processing, chemical characterization, performance, and reliability," *Ibm Journal of Research and Development*, vol. 43, pp. 301-326, MAY 1999.
- [304] D. Kapila, S. Hattangady, M. Douglas, R. Kraft, and M. Gribelyuk, "Modeling and optimization of oxynitride gate dielectrics formation by remote plasma nitridation of silicon dioxide," *Journal of the Electrochemical Society*, vol. 146, pp. 1111-1116, MAR 1999.

- [305] S. S. Tan, T. P. Chen, C. H. Ang, and L. Chan, "Relationship between interfacial nitrogen concentration and activation energies of fixed-charge trapping and interface state generation under bias-temperature stress condition," *Applied Physics Letters*, vol. 82, pp. 269-271, JAN 13 2003.
- [306] S. Inaba, T. Shimizu, S. Mori, K. Sekine, K. Saki, H. Suto, H. Fukui, M. Nagamine, M. Fujiwara, T. Yamamoto, M. Takayanagi, I. Mizushima, K. Okano, S. Matsuda, H. Oyamatsu, Y. Tsunashima, S. Yamada, Y. Toyoshima, and H. Ishiuchi, "Device performance of sub-50 nm CMOS with ultra-thin plasma nitrided gate dielectrics," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2002, pp. 651-654.
- [307] P. A. Kraus, K. Z. Ahmed, T. C. Chua, M. Ershov, H. Karbasi, C. S. Olsen, F. Nouri, J. Holland, R. Zhao, G. Miner, and A. Lepert, "Low-energy nitrogen plasmas for 65-nm node oxynitride gate dielectrics: a correlation of plasma characteristics and device parameters," in *Symposium on VLSI Technology*, 2003, pp. 143-144.
- [308] P. A. Kraus, K. Z. Ahmed, C. S. Olsen, and F. Nouri, "Model to predict gate tunneling current of plasma oxynitrides," *Ieee Transactions on Electron Devices*, vol. 52, pp. 1141-1147, Jun 2005.
- [309] G. Lucovsky, Y. Wu, H. Niimi, H. Yang, J. Keister, and J. E. Rowe, "Separate and independent reductions in direct tunneling in oxide/nitride stacks with monolayer interface nitridation associated with the (i) interface nitridation and (ii) increased physical thickness," *Journal of Vacuum Science & Technology a-Vacuum Surfaces and Films*, vol. 18, pp. 1163-1168, JUL-AUG 2000.
- [310] X. Guo and T. P. Ma, "Tunneling leakage current in oxynitride: Dependence on oxygen/nitrogen content," *Ieee Electron Device Letters*, vol. 19, pp. 207-209, JUN 1998.
- [311] H. Y. Yu, Y. T. Hou, M. F. Li, and D. L. Kwong, "Investigation of hole-tunneling current through ultrathin oxynitride/oxide stack gate dielectrics in p-MOSFETs," *Ieee Transactions on Electron Devices*, vol. 49, pp. 1158-1164, Jul 2002.
- [312] K. Z. Ahmed, P. A. Kraus, C. Olsen, S. Hung, and F. Nouri, "Observation of nitrogen-enhanced doping deactivation at the polysilicon-oxynitride interface of pMOSFETs with 12-angstrom gate dielectrics," *Ieee Electron Device Letters*, vol. 24, pp. 445-447, JUL 2003.
- [313] "Effect of increase in sub-threshold slope and body-effect parameter (due to degradation) causes negligible contribution to I_D degradation⁶, hence can be neglected in this discussion."
- [314] D. M. Brown, P. V. Gray, F. K. Heumann, H. R. Philipp, and E. A. Taft, "Properties of Sixoynz Films on Si," *Journal of the Electrochemical Society*, vol. 115, pp. 311-&, 1968.

- [315] S. V. Hattangady, H. Niimi, and G. Lucovsky, "Integrated processing of silicon oxynitride films by combined plasma and rapid-thermal processing," *Journal of Vacuum Science & Technology a-Vacuum Surfaces and Films*, vol. 14, pp. 3017-3023, NOV-DEC 1996.
- [316] K. Muraoka, K. Kurihara, N. Yasuda, and H. Satake, "Optimum structure of deposited ultrathin silicon oxynitride film to minimize leakage current," *Journal of Applied Physics*, vol. 94, pp. 2038-2045, AUG 1 2003.
- [317] T. Aoyama, K. Suzuki, H. Tashiro, Y. Tada, and K. Horiuchi, "Nitrogen concentration dependence on boron diffusion in thin silicon oxynitrides used for metal-oxide-semiconductor devices," *Journal of the Electrochemical Society*, vol. 145, pp. 689-693, FEB 1998.
- [318] B. Kaczer, V. Arkhipov, M. Jurczak, and G. Groeseneken, "Negative Bias Temperature Instability (NBTI) in SiO₂ and SiON gate dielectrics understood through disorder-controlled kinetics," *Microelectronic Engineering*, vol. 80, pp. 122-125, JUN 2005.
- [319] G. Chen, M. F. Li, C. H. Ang, J. Z. Zheng, and D. L. Kwong, "Dynamic NBTI of p-MOS transistors and its impact on MOSFET scaling," *Ieee Electron Device Letters*, vol. 23, pp. 734-736, DEC 2002.
- [320] K. Ahmed, E. Ibok, and J. Hauser, "Capacitor Test Structures for C-V Measurements on CMOS Devices with Sub-20 Å Oxides," in *Proceedings of the European Solid-State Device Research Conference*, 2000, pp. 308 - 311.
- [321] C. H. Choi, Y. Wu, J. S. Goo, Z. P. Yu, and R. W. Dutton, "Capacitance reconstruction from measured C-V in high leakage, nitride/oxide MOS," *Ieee Transactions on Electron Devices*, vol. 47, pp. 1843-1850, OCT 2000.
- [322] B. C. Paul, K. Kang, H. Kufluoglu, M. A. Alam, and K. Roy, "Negative Bias Temperature Instability: Estimation and Design for Improved Reliability of Nanoscale Circuits," in *accepted for IEEE Trans. on CAD of ICs*, 2006.
- [323] R. F. Pierret, *Semiconductor Device Fundamentals*, 2nd ed.: Prentice Hall, 1996.
- [324] M. Terai, K. Watanabe, and S. Fujieda, "Effect of nitrogen profile and fluorine incorporation on negative-bias temperature instability of ultrathin plasma-nitrided SiON MOSFETs," *Ieee Transactions on Electron Devices*, vol. 54, pp. 1658-1665, JUL 2007.
- [325] E. Cartier, B. P. Linder, V. Narayanan, and V. K. Paruchuri, "Fundamental Understanding and Optimization of PBTI in nFETs with SiO₂/HfO₂ Gate Stack," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2006, pp. 317-320.
- [326] D. Heh, C. D. Young, G. A. Brown, P. Y. Hung, A. Diebold, E. M. Vogel, J. B. Bernstein, and G. Bersuker, "Spatial distributions of trapping centers in HfO₂/SiO₂ gate stack," *Ieee Transactions on Electron Devices*, vol. 54, pp. 1338-1345, JUN 2007.

- [327] M. B. Zahid, R. Degraeve, L. Pantisano, J. F. Zhang, and G. Groeseneken, "Defects Generation in SiO₂/HfO₂ Studied With Variable T_{CHARGE}-T_{DISCHARGE} Charge Pumping (VT²CP)," in *IEEE International Reliability Physics Symposium*, 2007, pp. 55-60.
- [328] T. Krishnamohan, D. Kim, C. D. Nguyen, C. Jungemann, Y. Nishi, and K. C. Saraswat, "High-mobility low band-to-band-tunneling strained-germanium double-gate heterostructure FETs: Simulations," *Ieee Transactions on Electron Devices*, vol. 53, pp. 1000-1009, May 2006.
- [329] G. Eneman, P. Verheyen, R. Rooyackers, F. Nouri, L. Washington, R. Schreutelkamp, V. Moroz, L. Smith, A. D. Keersgieter, M. Jurczak, and K. D. Meyer, "Scalability of the Si_{1-x}Ge_x source/drain technology for the 45-nm technology node and beyond," *Ieee Transactions on Electron Devices*, vol. 53, pp. 1647-1656, JUL 2006.
- [330] D. Esseni and A. Abramo, "Modeling of electron mobility degradation by remote Coulomb scattering in ultrathin oxide MOSFETs," *Ieee Transactions on Electron Devices*, vol. 50, pp. 1665-1674, JUL 2003.
- [331] S. Takagi and M. Takayanagi, "Experimental evidence of inversion-layer mobility lowering in ultrathin gate oxide metal-oxide-semiconductor field-effect-transistors with direct tunneling current," *Japanese Journal of Applied Physics Part 1- Regular Papers Short Notes & Review Papers*, vol. 41, pp. 2348-2352, APR 2002.
- [332] S. Saito, K. Torii, Y. Shimamoto, O. Tonomura, D. Hisamoto, T. Onai, M. Hiratani, S. Kimura, Y. Manabe, M. Caymax, and J. W. Maes, "Remote-charge-scattering limited mobility in field-effect transistors with SiO₂ and Al₂O₃/SiO₂ gate stacks," *Journal of Applied Physics*, vol. 98, pp. -, DEC 1 2005.
- [333] M. S. Lundstrom, "On the mobility versus drain current relation for a nanoscale MOSFET," *Ieee Electron Device Letters*, vol. 22, pp. 293-295, JUN 2001.
- [334] K. Natori, "Ballistic/quasi-ballistic transport in nanoscale transistor," *Applied Surface Science*, vol. 254, pp. 6194-6198, Jul 30 2008.
- [335] A. Lochtefeld and D. A. Antoniadis, "On experimental determination of carrier velocity in deeply scaled NMOS: How close to the thermal limit?," *Ieee Electron Device Letters*, vol. 22, pp. 95-97, Feb 2001.
- [336] R. Clerc, P. Palestri, and L. Selmi, "On the physical understanding of the kT-layer concept in quasi-ballistic regime of transport in nanoscale devices," *Ieee Transactions on Electron Devices*, vol. 53, pp. 1634-1640, Jul 2006.
- [337] M. Zilli, P. Palestri, D. Esseni, and L. Selmi, "On the experimental determination of channel back-scattering in nanoMOSFETs," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2007, pp. 105-108.

- [338] M. J. Chen, H. T. Huang, Y. C. Chou, R. T. Chen, Y. T. Tseng, P. N. Chen, and C. H. Diaz, "Separation of channel backscattering coefficients in nanoscale MOSFETs," *Ieee Transactions on Electron Devices*, vol. 51, pp. 1409-1415, Sep 2004.
- [339] A. Khakifirooz and D. A. Antoniadis, "MOSFET performance scaling - Part I: Historical trends," *Ieee Transactions on Electron Devices*, vol. 55, pp. 1391-1400, Jun 2008.
- [340] M. V. Fischetti, T. P. O'Regan, S. Narayanan, C. Sachs, S. Jin, J. Kim, and Y. Zhang, "Theoretical study of some physical aspects of electronic transport in nMOSFETs at the 10-nm gate-length," *Ieee Transactions on Electron Devices*, vol. 54, pp. 2116-2136, Sep 2007.
- [341] A. Rahman and M. S. Lundstrom, "A compact scattering model for the nanoscale double-gate MOSFET," *Ieee Transactions on Electron Devices*, vol. 49, pp. 481-489, Mar 2002.
- [342] R. Kim and M. S. Lundstrom, "Physics of Carrier Backscattering in One- and Two-Dimensional Nanotransistors," *Ieee Transactions on Electron Devices*, vol. 56, pp. 132-139, Jan 2009.
- [343] P. Palestri, D. Esseni, S. Eminent, C. Fiegna, E. Sangiorgi, and L. Selmi, "Understanding quasi-ballistic transport in nano-MOSFETs: Part I - Scattering in the channel, and in the drain," *Ieee Transactions on Electron Devices*, vol. 52, pp. 2727-2735, Dec 2005.
- [344] F. Assad, Z. Ren, D. Vasileska, S. Datta, and M. S. Lundstrom, "On the performance limits for Si MOSFET's: A theoretical study," *Ieee Transactions on Electron Devices*, vol. 47, pp. 232-240, Jan 2000.
- [345] J. C. Ho, R. Yerushalmi, Z. A. Jacobson, Z. Fan, R. L. Alley, and A. Javey, "Controlled nanoscale doping of semiconductors via molecular monolayers," *Nature Materials*, vol. 7, pp. 62-67, Jan 2008.
- [346] K. Kang, K. Kim, A. E. Islam, M. A. Alam, and K. Roy, "Characterization and estimation of circuit reliability degradation under NBTI using on-line I_{DDQ} measurement," *ACM/IEEE Design Automation Conference*, pp. 358-363, 2007.
- [347] A. E. Islam and M. A. Alam, "On the possibility of degradation-free field effect transistors," *Applied Physics Letters*, vol. 92 (173504), Apr 28 2008.
- [348] A. E. Islam and M. A. Alam, "Mobility Enhancement due to Charge Trapping & Defect Generation: Physics of Self-Compensated BTI," in *to appear in IEEE International Reliability Physics Symposium*, 2010.
- [349] O. Bonno, S. Barraud, D. Mariolle, and F. Andrieu, "Effect of strain on the electron effective mobility in biaxially strained silicon inversion layers: An experimental and theoretical analysis via atomic force microscopy measurements and Kubo-Greenwood mobility calculations," *Journal of Applied Physics*, vol. 103, pp. -, Mar 15 2008.

- [350] Y. Zhao, M. Takenaka, and S. Takagi, "On Surface Roughness Scattering-Limited Mobilities of Electrons and Holes in Biaxially Tensile-Strained Si MOSFETs," *Ieee Electron Device Letters*, vol. 30, pp. 987-989, Sep 2009.
- [351] G. Y. Sun, Y. K. Sun, T. K. Nishida, and S. E. Thompson, "Hole mobility in silicon inversion layers: Stress and surface orientation," *Journal of Applied Physics*, vol. 102, pp. -, Oct 15 2007.
- [352] O. Weber and S. Takagi, "Experimental examination and physical understanding of the Coulomb scattering mobility in strained-Si nMOSFETs," *Ieee Transactions on Electron Devices*, vol. 55, pp. 2386-2396, Sep 2008.
- [353] A. Cros, K. Romanjek, D. Fleury, S. Harrison, R. Cerutti, P. Coronel, B. Dumont, A. Pouydebasque, R. Wacquez, B. Duriez, R. Gwoziecki, F. Boeuf, H. Brutt, G. Ghibaudo, and T. Skotnicki, "Unexpected mobility degradation for very short devices: A new challenge for CMOS scaling," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2006, pp. 663-666.
- [354] J. L. Egley and D. Chidambarrao, "Strain Effects on Device Characteristics - Implementation in Drift-Diffusion Simulators," *Solid-State Electronics*, vol. 36, pp. 1653-1664, Dec 1993.
- [355] V. Moroz, N. Strecker, X. P. Xu, L. Smith, and I. Bork, "Modeling the impact of stress on silicon processes and devices," *Materials Science in Semiconductor Processing*, vol. 6, pp. 27-36, Feb-Jun 2003.
- [356] C.-C. Wang, W. Zhao, M. Chen, and Y. Cao, "Compact modeling of stress effects in scaled CMOS," in *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 2009, pp. 131-134.
- [357] Y. Zhao, M. Takenaka, and S. Takagi, "Comprehensive understanding of surface roughness and Coulomb scattering mobility in biaxially-strained Si MOSFETs," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2008, pp. 577-580.
- [358] C. L. Hinkle, A. M. Sonnet, R. A. Chapman, and E. M. Vogel, "Extraction of the Effective Mobility of In_{0.53}Ga_{0.47}As MOSFETs," *Ieee Electron Device Letters*, vol. 30, pp. 316-318, 2009.
- [359] K. Kuhn, C. Kenyon, A. Kornfeld, M. Liu, A. Maheshwari, W.-K. Shih, S. Sivakumar, G. Taylor, P. VanDerVoorn, and K. Zawadzki, "Managing Process Variation in Intel's 45nm CMOS Technology," *Intel Technology Journal*, vol. 12, pp. 93-110, 2008.
- [360] W. P. Wang, S. Q. Yang, S. Bhardwaj, R. Vattikonda, S. Vrudhula, F. Liu, and Y. Cao, "The impact of NBTI on the performance of combinational and sequential circuits," *ACM/IEEE Design Automation Conference*, pp. 364-369, 2007.

- [361] K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau, C.-H. Choi, G. Ding, K. Fischer, T. Ghani, R. Grover, W. Han, D. Hanken, M. Hattendorf, J. He, J. Hicks, R. Huessner, D. Ingerly, P. Jain, R. James, L. Jong, S. Joshi, C. Kenyon, K. Kuhn, K. Lee, H. Liu, J. Maiz, B. McIntyre, P. Moon, J. Neiryneck, S. Pae, C. Parker, D. Parsons, C. Prasad, L. Pipes, M. Prince, P. Ranade, T. Reynolds, J. Sandford, L. Shifren, J. Sebastian, J. Seiple, D. Simon, S. Sivakumar, P. Smith, C. Thomas, T. Troeger, P. Vandervoorn, S. Williams, and K. Zawadzki, "A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2007, pp. 247-250.
- [362] D. Bishop, "Nanotechnology and the end of Moore's Law?," *Bell Labs Technical Journal*, vol. 10, pp. 23-28, Fal 2005.
- [363] S. Borkar, "Design challenges of technology scaling," *Ieee Micro*, vol. 19, pp. 23-29, Jul-Aug 1999.
- [364] L. B. Kish, "End of Moore's law: thermal (noise) death of integration in micro and nano electronics," *Physics Letters A*, vol. 305, pp. 144-149, Dec 2 2002.
- [365] C. C. Mann, "The end of Moore's law?," *Technology Review*, vol. 103, pp. 42-48, May-Jun 2000.
- [366] J. J. Welser, G. I. Bourianoff, V. V. Zhirmov, and R. K. Cavin, "The quest for the next information processing technology," *Journal of Nanoparticle Research*, vol. 10, pp. 1-10, Feb 2008.
- [367] S. Datta and B. Das, "Electronic Analog of the Electrooptic Modulator," *Applied Physics Letters*, vol. 56, pp. 665-667, Feb 12 1990.
- [368] K. C. Hall and M. E. Flatte, "Performance of a spin-based insulated gate field effect transistor," *Applied Physics Letters*, vol. 88, pp. -, Apr 17 2006.
- [369] J. Schliemann, J. C. Egues, and D. Loss, "Nonballistic spin-field-effect transistor," *Physical Review Letters*, vol. 90, pp. -, Apr 11 2003.
- [370] L. Leem and J. S. Harris, "Magnetic Coupled Spin-torque Devices and Magnetic Ring Oscillator," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2008, pp. 159-162.
- [371] D. E. Nikonov, G. I. Bourianoff, and P. A. Gargini, "Simulation of highly idealized, atomic scale magnetic quantum cellular automata logic circuits," *Journal of Nanoelectronics and Optoelectronics*, vol. 3, pp. 3-11, Mar 2008.
- [372] S. Yang, S. Ahmed, B. Arcot, R. Arghavani, P. Bai, S. Chambers, P. Charvat, R. Cotner, R. Gasser, T. Ghani, M. Hussein, C. Jan, C. Kardas, J. Maiz, F. McGregor, B. McIntyre, P. Nguyen, P. Packan, I. Post, S. Sivakumar, J. Steigerwald, M. Taylor, B. Tufts, S. Tyagi, and M. Bohr, "A high performance 180 nm generation logic technology," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 197-200, 1998.

- [373] S. Thompson, M. Alavi, R. Arghavani, A. Brand, and R. Bigwood, "An Enhanced 130 nm Generation Logic Technology Featuring 60 nm Transistors Optimized for High Performance and Low Power at 0.7 - 1.4 V," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2001, pp. 257-260.
- [374] T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson, and M. Bohr, "A 90nm high volume manufacturing logic technology featuring novel 45nm gate length strained silicon CMOS transistors," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 978-980, 2003.
- [375] S. Tyagi, C. Auth, P. Bai, G. Curello, H. Deshpande, S. Gannavaram, O. Golonzka, R. Heussner, R. James, C. Kenyon, S. H. Lee, N. Lindert, M. Liu, R. Nagisetty, S. Natarajan, C. Parker, J. Sebastian, B. Sell, S. Sivakumar, A. St Amour, and K. Tone, "An advanced low power, high performance, strained channel 65nm technology," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 1070-1072, 2005.
- [376] A. E. Islam and M. A. Alam, "Self-Optimizing Defect Generation for Advanced CMOS Substrates," in *IEEE International Integrated Reliability Workshop*, 2009, pp. 97-101.
- [377] T. H. Morshed, S. P. Devireddy, Z. Celik-Butler, A. Shanware, K. Green, J. J. ChamberS, M. R. Visokay, and L. Colombo, "Physics-based 1/f noise model for MOSFETs with nitrided high-kappa gate dielectrics," *Solid-State Electronics*, vol. 52, pp. 711-724, May 2008.
- [378] H. D. Xiong, D. Heh, M. Gurfinkel, Q. Li, Y. Shapira, C. Richter, G. Bersuker, R. Choi, and J. S. Suehle, "Characterization of electrically active defects in high-k gate dielectrics by using low frequency noise and charge pumping measurements," *Microelectronic Engineering*, vol. 84, pp. 2230-2234, SEP-OCT 2007.
- [379] K. Kang, H. Kufluoglu, K. Roy, and M. A. Alam, "Impact of Negative Bias Temperature Instability in Nano-Scale SRAM Array: Modeling and Analysis," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 26, pp. 1770-1781, OCT 2007.
- [380] S. V. Kumar, C. H. Kim, and S. S. Sapatnekar, "Impact of NBTI on SRAM read stability and design for reliability," *ISQED: Proceedings of the International Symposium on Quality Electronic Design*, pp. 210-215, 2006.
- [381] S. Bhardwaj, W. P. Wang, R. Vattikonda, Y. Cao, and S. Vrudhula, "Predictive modeling of the NBTI effect for reliable design," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 189-192, 2006.
- [382] J. C. Lin, A. S. Oates, and C. H. Yu, "Time dependent V_{ccmin} degradation of SRAM fabricated with high-k gate dielectrics," *Proceedings of International Reliability Physics Symposium*, pp. 439-444, 2007.

- [383] M. S. Lundstrom, "ECE 612 Nanoscale Transistors," Aug 8, 2006 ed: www.nanoHUB.org.
- [384] S. C. Sun and J. D. Plummer, "Electron-Mobility in Inversion and Accumulation Layers on Thermally Oxidized Silicon Surfaces," *Ieee Transactions on Electron Devices*, vol. 27, pp. 1497-1508, 1980.
- [385] C. Augustine, A. Raychowdhury, Y. F. Gao, M. Lundstrom, and K. Roy, "PETE: A Device/Circuit Analysis Framework for Evaluation and Comparison Of Charge Based Emerging Devices," *International Symposium on Quality Electronic Design (ISQED)*, pp. 80-85, 2009.
- [386] A. Raychowdhury, C. Augustine, Y. Gao, M. Lundstrom, and K. Roy, "PETE : Purdue Emerging Technology Evaluator," 2007.
- [387] A. E. Islam, S. Mahapatra, S. Deora, V. D. Maheta, and M. A. Alam, "On The Differences Between Ultra-fast NBTI Experiments and Reaction-Diffusion Theory," in *International Electron Devices Meeting (IEDM) Technical Digest*, 2009, pp. 733-736.
- [388] A. E. Islam, C. Augustine, K. Roy, and M. A. Alam, "Extent of Variation Resilience in Strained CMOS: From Transistors to Digital Circuits," *Submitted to IEEE Transaction of Electron Devices*, 2010.
- [389] J. R. Heath, P. J. Kuekes, G. S. Snider, and R. S. Williams, "A defect-tolerant computer architecture: Opportunities for nanotechnology," *Science*, vol. 280, pp. 1716-1721, Jun 12 1998.
- [390] A. J. Kleinsowski and D. J. Lilja, "The NanoBox project: Exploring fabrics of self-correcting logic blocks for high defect rate molecular device technologies," *Proceeding of IEEE Computer Society Annual Symposium on VLSI*, pp. 19-24, 2004.
- [391] H. Kufluoglu, "MOSFET degradation due to negative bias temperature instability (NBTI) and hot carrier injection (HCI), and its implications for reliability-aware VLSI design," in *Electrical and Computer Engineering West Lafayette: Purdue University*, 2007, p. 200.
- [392] J. Mitard, X. Garros, L. P. Nguyen, C. Leroux, G. Ghibaudo, F. Martin, and G. Reimbold, "Large-scale time characterization and analysis of PBTI in HfO₂/metal gate stacks," *International Reliability Physics Symposium Proceedings*, pp. 174-178, 2006.
- [393] Z. Arefinia and A. A. Orouji, "Investigation of the novel attributes of a carbon nanotube FET with high-kappa gate dielectrics," *Physica E-Low-Dimensional Systems & Nanostructures*, vol. 40, pp. 3068-3071, Sep 2008.
- [394] Q. Cao, H. S. Kim, N. Pimparkar, J. P. Kulkarni, C. J. Wang, M. Shim, K. Roy, M. A. Alam, and J. A. Rogers, "Medium-scale carbon nanotube thin-film integrated circuits on flexible plastic substrates," *Nature*, vol. 454, pp. 495-U4, Jul 24 2008.

- [395] A. Javey, J. Guo, D. B. Farmer, Q. Wang, D. W. Wang, R. G. Gordon, M. Lundstrom, and H. J. Dai, "Carbon nanotube field-effect transistors with integrated ohmic contacts and high-k gate dielectrics," *Nano Letters*, vol. 4, pp. 447-450, Mar 2004.
- [396] A. Javey, H. Kim, M. Brink, Q. Wang, A. Ural, J. Guo, P. McIntyre, P. McEuen, M. Lundstrom, and H. J. Dai, "High-kappa dielectrics for advanced carbon-nanotube transistors and logic gates," *Nature Materials*, vol. 1, pp. 241-246, Dec 2002.
- [397] A. A. Orouji and Z. Arefinia, "The impact of high-kappa gate dielectrics on carbon nanotube transistors," *Proceedings of the 2007 International Workshop on the Physics of Semiconductor Devices: Iwpsd-2007*, pp. 831-833, 2007.
- [398] S. Salahuddin and S. Datta, "Use of negative capacitance to provide voltage amplification for low power nanoscale devices," *Nano Letters*, vol. 8, pp. 405-410, Feb 2008.
- [399] N. Setter, D. Damjanovic, L. Eng, G. Fox, S. Gevorgian, S. Hong, A. Kingon, H. Kohlstedt, N. Y. Park, G. B. Stephenson, I. Stolitchnov, A. K. Tagantsev, D. V. Taylor, T. Yamada, and S. Streiffner, "Ferroelectric thin films: Review of materials, properties, and applications," *Journal of Applied Physics*, vol. 100, pp. -, SEP 1 2006.
- [400] A. K. Tagantsev, I. Stolichnov, E. L. Colla, and N. Setter, "Polarization fatigue in ferroelectric films: Basic experimental findings, phenomenological scenarios, and microscopic features," *Journal of Applied Physics*, vol. 90, pp. 1387-1402, AUG 1 2001.
- [401] H. M. Duiker, P. D. Beale, J. F. Scott, C. A. P. Dearaujo, B. M. Melnick, J. D. Cuchiaro, and L. D. Mcmillan, "Fatigue and Switching in Ferroelectric Memories - Theory and Experiment," *Journal of Applied Physics*, vol. 68, pp. 5783-5791, DEC 1 1990.
- [402] M. Dawber and J. F. Scott, "A model for fatigue in ferroelectric perovskite thin films (vol 76, pg 1060, 2000)," *Applied Physics Letters*, vol. 76, pp. 3655-3655, JUN 12 2000.
- [403] W. L. Warren, B. A. Tuttle, D. Dimos, G. E. Pike, H. N. AlShareef, R. Ramesh, and J. T. Evans, "Imprint in ferroelectric capacitors," *Japanese Journal of Applied Physics Part 1-Regular Papers Short Notes & Review Papers*, vol. 35, pp. 1521-1524, FEB 1996.
- [404] M. Grossmann, O. Lohse, D. Bolten, U. Boettger, T. Schneller, and R. Waser, "The interface screening model as origin of imprint in PbZrxTi1-xO3 thin films. I. Dopant, illumination, and bias dependence," *Journal of Applied Physics*, vol. 92, pp. 2680-2687, SEP 1 2002.

- [405] M. Grossmann, O. Lohse, D. Bolten, U. Boettger, and R. Waser, "The interface screening model as origin of imprint in $\text{PbZr}_x\text{Ti}_{1-x}\text{O}_3$ thin films. II. Numerical simulation and verification," *Journal of Applied Physics*, vol. 92, pp. 2688-2696, SEP 1 2002.
- [406] M. Grossmann, O. Lohse, D. Bolten, U. Boettger, R. Waser, W. Hartner, M. Kastner, and G. Schindler, "Lifetime estimation due to imprint failure in ferroelectric $\text{SrBi}_2\text{Ta}_2\text{O}_9$ thin films," *Applied Physics Letters*, vol. 76, pp. 363-365, JAN 17 2000.
- [407] M. Houssa, M. Aoulaiche, S. De Gendt, G. Groeseneken, M. M. Heyns, and A. Stesmans, "Reaction-dispersive proton transport model for negative bias temperature instabilities," *Applied Physics Letters*, vol. 86, pp. -, Feb 28 2005.
- [408] H. Scher and E. W. Montroll, "Anomalous Transit-Time Dispersion in Amorphous Solids," *Physical Review B*, vol. 12, pp. 2455-2477, 1975.
- [409] J. C. Dyre, "Master-Equation Approach to the Glass-Transition," *Physical Review Letters*, vol. 58, pp. 792-795, Feb 23 1987.
- [410] J. Noolandi, "Multiple-Trapping Model of Anomalous Transit-Time Dispersion in a-Se," *Physical Review B*, vol. 16, pp. 4466-4473, 1977.
- [411] V. I. Arkhipov and A. I. Rudenko, "Drift and Diffusion in Materials with Traps .2. Non-Equilibrium Transport Regime," *Philosophical Magazine B-Physics of Condensed Matter Statistical Mechanics Electronic Optical and Magnetic Properties*, vol. 45, pp. 189-207, 1982.
- [412] W. B. Jackson, "Connection between the Meyer-Neldel Relation and Multiple-Trapping Transport," *Physical Review B*, vol. 38, pp. 3595-3598, Aug 15 1988.
- [413] D. L. Scharfetter and H. K. Gummel, "Large-Signal Analysis of a Silicon Read Diode Oscillator," *Ieee Transactions on Electron Devices*, vol. Ed16, p. 64, 1969.
- [414] V. Huard, F. Monsieur, G. Ribes, and S. Bruyere, "Evidence for hydrogen-related defects during NBTI stress in p-MOSFETs," *IEEE International Reliability Physics Symposium*, pp. 178-182, 2003.
- [415] M. Houssa, M. Aoulaiche, J. L. Autran, C. Parthasarathy, N. Revil, and E. Vincent, "Modeling negative bias temperature instabilities in hole channel metal-oxide-semiconductor field effect transistors with ultrathin gate oxide layers," *Journal of Applied Physics*, vol. 95, pp. 2786-2791, Mar 1 2004.
- [416] M. Houssa, J. L. Autran, A. Stesmans, and M. M. Heyns, "Model for interface defect and positive charge generation in ultrathin $\text{SiO}_2/\text{ZrO}_2$ gate dielectric stacks," *Applied Physics Letters*, vol. 81, pp. 709-711, Jul 22 2002.
- [417] M. A. Alam and H. Kufluoglu, "On Quasi-Saturation of Negative Bias Temperature Degradation," in *208th Meeting of The Electrochemical Society (ECS)*, 2005.

- [418] M. F. Li, D. M. Huang, C. Shen, T. Yang, W. J. Liu, and Z. Y. Liu, "Understand NBTI mechanism by developing novel measurement techniques," *Ieee Transactions on Device and Materials Reliability*, vol. 8, pp. 62-71, Mar 2008.
- [419] W. J. Liu, Z. Y. Liu, D. M. Huang, C. C. Liao, L. F. Zhang, Z. H. Gan, W. S. Wong, C. Shen, and M. F. Li, "On-the-fly interface trap measurement and its impact on the understanding of NBTI mechanism for p-MOSFETs with SiON gate dielectric," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 813-816, 2007.
- [420] G. Groeseneken, H. E. Maes, N. Beltran, and R. F. Dekeersmaecker, "A Reliable Approach to Charge-Pumping Measurements in Mos-Transistors," *Ieee Transactions on Electron Devices*, vol. 31, pp. 42-53, 1984.
- [421] J. H. Stathis and S. Zafar, "The negative bias temperature instability in MOS devices: A review," *Microelectronics Reliability*, vol. 46, pp. 270-286, Feb-Apr 2006.
- [422] C. G. Van de Walle, "Stretched-exponential relaxation modeled without invoking statistical distributions," *Physical Review B*, vol. 53, pp. 11292-11295, May 1 1996.
- [423] J. P. Mckelvey, T. P. Brody, and R. L. Longini, "Alternative Approach to Solution of Added Carrier Transport Problems in Semiconductors," *Physical Review*, vol. 1, pp. 51-&, 1961.
- [424] S. Tanaka and M. S. Lundstrom, "A Compact HBT Device Model-Based on a One-Flux Treatment of Carrier Transport," *Solid-State Electronics*, vol. 37, pp. 401-410, Mar 1994.
- [425] H. S. Wilf, *Generatingfunctionology*, 3rd ed. Wellesley, Mass.: A K Peters, 2006.
- [426] G. Timp, J. Bude, K. K. Bourdelle, J. Garno, A. Ghetti, H. Gossmann, M. Green, G. Forsyth, Y. Kim, R. Kleiman, F. Klemens, A. Kornblit, C. Lochstampfor, W. Mansfield, S. Moccio, T. Sorsch, D. M. Tennant, W. Timp, and R. Tung, "The Ballistic Nano-transistor," in *International Electron Devices Meeting (IEDM) Technical Digest*, 1999, pp. 55-58.
- [427] M. Zilli, D. Esseni, P. Palestri, and L. Selmi, "On the apparent mobility in nanometric n-MOSFETs," *Ieee Electron Device Letters*, vol. 28, pp. 1036-1039, Nov 2007.
- [428] M. Xiao, I. Martin, E. Yablonovitch, and H. W. Jiang, "Electrical detection of the spin resonance of a single electron in a silicon field-effect transistor," *Nature*, vol. 430, pp. 435-439, Jul 22 2004.
- [429] M. J. Chen and L. F. Lu, "A parabolic potential barrier-oriented compact model for the $k_B T$ layer's width in nano-MOSFETs," *Ieee Transactions on Electron Devices*, vol. 55, pp. 1265-1268, May 2008.

APPENDICES

A. DEFINITION OF PARAMETERS

A.1. Effective Oxide Thickness (EOT)

Introduction of high- κ materials (having higher dielectric constant, ϵ_{HK}) as gate dielectric in MOS transistor have enabled us to increase the physical thickness of the dielectric layer (T_{PHY}), without sacrificing the dielectric capacitance (C_{di}); since $C_{di} = \epsilon_{HK}/T_{PHY}$. However, comparison of different CMOS technologies require one to define EOT as,

$$EOT = \frac{\epsilon_{SiO_2}}{\epsilon_{HK}} T_{PHY}, \quad (A.1)$$

where ϵ_{SiO_2} is the dielectric constant of SiO_2 (the original insulator material used in MOS technology). Since 1960s, CMOS scaling has progressed in such a way that there have always been a continuous reduction in EOT and hence an increase in $C_{di} = \epsilon_{HK}/T_{PHY} = \epsilon_{SiO_2}/EOT$ [7].

A.2. Lifetime (t_{life})

Lifetime (t_{life}) is a reliability parameter, commonly used in defining the operating period of a device, up to which it can maintain a certain optimum performance. In MOS reliability, t_{life} is defined as the time, when the transistor degradation or ΔV_T reaches certain maximum value ($\Delta V_{T(max)}$). As shown in Figure 3.21, one can determine t_{life} by extrapolating the stress data up to $\Delta V_{T(max)}$ and then reading-out the required time for having $\Delta V_{T(max)}$.

A.3. Voltage Acceleration (Γ_V)

Voltage acceleration (Γ_V) is another MOS reliability parameter, used for defining the change in the degradation magnitude of a transistor, as the operating voltage is increased.

In this thesis, Γ_V is defined as the absolute value for the slope of $\log_{10}(t_{life})$ vs. V_G plot, as shown in Figure 6.21.

A.4. Safe Operating Voltage (V_{safe})

This is defined as the stress voltage (V_{STS}) of a MOS transistor, at which it can operate up to a specified lifetime (t_{life}) without crossing a certain failure criteria ($\Delta V_{T(max)}$). Figure 3.21 indicates the determination of V_{safe} for a particular transistor using $\Delta V_{T(max)} = 60\text{mV}$ and $t_{life} = 5$ years.

A.5. Activation Energy (E_A)

In MOS reliability, E_A is obtained from the absolute magnitude of the slope of $\log_e(\Delta V_T @ t_{STS})$ vs. $1/kT_{STS}$ plot, where T_{STS} and t_{STS} indicate the stress temperature and stress time, respectively. During NBTI stress on type-I transistor, having dominance of interface defect (N_{IT}) generation over hole trapping (N_{HT}), ΔV_T can be expressed as –

$$\Delta V_T \sim \exp(E_{A,IT}/kT_{STS})t_{STS}^n. \quad (\text{A.2})$$

Thus, from a plot of $\log_e(1/t_{STS})$ vs. $1/kT_{STS}$ and taking its absolute slope, we can obtain $E_{A,IT}/n$. Then, using equation (3.6), we have $E_{A,IT}/n \equiv E_{D1}$, which in limiting cases of $E_{A,F} \approx E_{A,R}$ [30, 125] and negligible aE_{ox} can provide an estimation for the activation energy of hydrogen diffusion (E_D in equation (3.6)) using, $E_{A,IT}/n \equiv E_D$. Therefore, $E_{A,IT}$ and E_D follows the following relationship during N_{IT} generation (under NBTI stress) in a type-I transistor:

$$E_{A,IT} \equiv nE_D, \quad (\text{A.3})$$

which is defined as the generalized scaling law for NBTI activation in type-I transistors [125].

B. ALTERNATE MODELS FOR N_{IT} FORMATION

In addition to the R-D model involving Arrhenius drift/diffusion of hydrogen species (as discussed in chapter 3), many other models are available in literature for handling N_{IT} formation. Most of them were focused on capturing signatures obtained from one or two experiments, hence failed to have a broader impact. More interestingly, some of the models (though presented in a different way) can be shown (section B.4.1) to have direct equivalence with different versions of R-D model, discussed in chapter 3.

B.1. R-D Model: Dispersive Drift/diffusion

The assumption of Arrhenius diffusion for the diffusing hydrogen species, having constant diffusion coefficient, *i.e.*, $D \sim \exp[-E_A/k_B T]$, is only appropriate for isotropic media with spatially and temporally uniform hopping rates. Classical R-D models (described in sections 3.5) consider diffusion to be happening mostly within the poly-Si gate structure [143], thus assumption of uniform hopping rates (Figure B.1a) can consistently explain the experiments, specifically the temperature independence of power-law time exponent, n (Figure 6.26e). However, there are models available in NBTI literature [33, 148, 407], which considers that the diffusion of hydrogen species are restricted within an amorphous media (mostly oxides), where hopping distances and hopping rates are exponentially distributed [149, 150, 408, 409] (Figure B.1b). Therefore, although the diffusion process can be described by a constant diffusion coefficient (D_{00}) at the earliest phases, the diffusive particles soon find themselves trapped in states with exponentially long release times, which in turn slows down or disperse the overall diffusion process. The resultant dispersive diffusion coefficient can be approximated using [30, 149, 150] –

$$D = D_{00} (\omega_0 t)^{-\beta} \quad (\text{B.1})$$

where $\beta = 1 - k_B T/E_0$ and E_0 is the characteristic energy. Such a time-dependent diffusion coefficient, having an inherent slow down of diffusion with time, will obviously reduce n of generated N_{IT} to value less than the one obtained from classical Arrhenius-type

diffusion. Detailed analytical and numerical studies [33, 152, 153] show that N_{IT} generation with dispersive hydrogen diffusion can be expressed as –

$$N_{IT} \sim t^{n_{classical}(1-\beta)} \sim t^{n_{classical}k_B T/E_0} \quad (\text{B.2})$$

where $n_{classical}$ is the power-law time exponent expected from classical drift/diffusion of hydrogen species (Table 3.1). Thus equation (B.2) suggests a temperature dependence of n for N_{IT} generation, modeled within R-D framework, which is a definite signature of the presence of dispersive transport in N_{IT} generation experiments. Such existence of temperature dependent n motivated researchers [33, 407] to predict the presence of dispersive H^+ transport within the R-D framework, which was later shown to be an artifact of measurement delay [130].

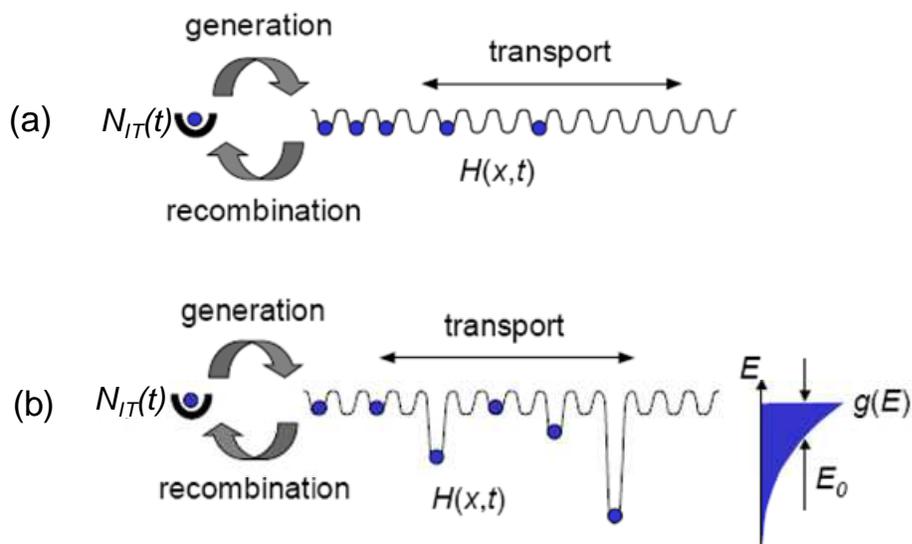


Figure B.1: Schematic representation of the N_{IT} generation model with (a) classical hydrogen transport having single energy level and (b) dispersive hydrogen transport through exponentially distributed multiple energy levels, $g(E) \sim \exp(-E/E_0)$. Adapted from [33].

Incorporation of dispersive hydrogen transport within R-D framework requires one to treat the hydrogen drift/diffusion through multiple trapping [151, 153, 410, 411]. The resultant equations for the entire R-D framework have the following form (shown here for atomic hydrogen):

$$\frac{dN_{IT}}{dt} = k_F (N_0 - N_{IT}) - k_R N_{IT} \left[N_H^{(F)} + \int_{-E_{min}}^0 \frac{\partial \rho(z, E_t, t)}{\partial t} dE_t \right] \quad (B.3)$$

$$\frac{\delta}{2} \frac{\partial N_H^{(F)}(0, t)}{\partial t} = -D_{00} \frac{\partial N_H^{(F)}(0, t)}{\partial z} + \frac{dN_{IT}}{dt} - \frac{\delta}{2} \int_{-E_{min}}^0 \frac{\partial \rho(z, E_t, t)}{\partial t} dE_t \quad (B.4)$$

$$\frac{\partial N_H^{(F)}(z, t)}{\partial t} = D_{00} \nabla^2 N_H^{(F)} - \int_{-E_{min}}^0 \frac{\partial \rho(z, E_t, t)}{\partial t} dE_t \quad (B.5)$$

$$\begin{aligned} \frac{\partial \rho(z, E_t, t)}{\partial t} = & v_0 \frac{g(E_t) - \rho(z, E_t, t)}{N_C - N_H^{(F)}(z, t) + g(E_t) - \rho(z, E_t, t)} N_H^{(F)}(z, t) \\ & - v_0 \exp\left(\frac{E_t}{kT}\right) \rho(z, E_t, t) \frac{N_C - N_H^{(F)}(z, t)}{N_C}; \quad E_t < 0 \end{aligned} \quad (B.6)$$

where N_H^F is the hydrogen concentration in the conduction state (unit: cm^{-3}), N_C is the effective density-of-states (DOS) in the conduction state (unit: cm^{-3}), $\rho(z, E_t, t)$ is the hydrogen concentration at energy level E_t (unit: $\text{cm}^{-3}\text{eV}^{-1}$), $g(E_t) = N_T/E_0 \exp(E_t/E_0)$ [33, 151, 153, 411] is the DOS at E_t (unit: $\text{cm}^{-3}\text{eV}^{-1}$), v_0 is the attempt frequency ($\sim 10^{10}$ Hz [412], $\sim 10^{12}$ Hz [130]), N_T is the DOS for trapped hydrogen (unit: cm^{-3}) and E_{min} is the minimum hydrogen trap levels present in the amorphous media (which is very large for highly dispersive media and finite for amorphous structures [130, 151, 411]).

Our numerical implementation of dispersive R-D model with $v_0 = 10^{12}$ Hz, $E_0 = 47\text{meV}$, $N_T = 5 \times 10^{21} \text{ cm}^{-3}$, and $N_C = 10^{19} \text{ cm}^{-3}$ reveals that as time progresses hydrogen in a dispersive system settles down to deeper energy levels. The transition between these shallow and deep levels (termed as demarcation energy) is expressed as [30, 151, 153, 411] –

$$E_d(t) = -k_B T \ln v_0 t \quad (B.7)$$

which closely follows the average energy (E_{avg}) of trapped hydrogen, when $E_{min} \gg E_d(t)$ [152, 153]. On the other hand, if the DOS of trapped hydrogen is shallow (*i.e.*, E_{min} is finite) [125, 130], then our simulation (Figure B.2) indicates a saturation of E_{avg} at a value, which is independent of temperature. Following equation (B.7), the time for reaching E_{avg} only depends on E_{min} and ν_0 .

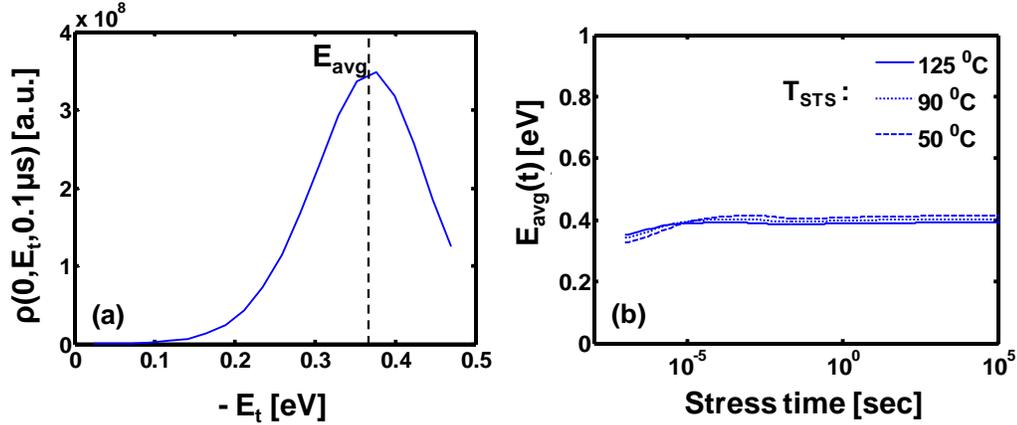


Figure B.2: Simulation results of R-D framework (equations (B.3)-(B.6)) with dispersive hydrogen transport. Simulation parameters: $\nu_0 = 10^{12}$ Hz, $E_0 = 47$ meV, $N_T = 5 \times 10^{21}$ cm $^{-3}$, $N_C = 10^{19}$ cm $^{-3}$, and $E_{min} = 0.47$ eV. (a) Hydrogen distribution near the oxide/substrate interface at $t = 0.1$ us. (b) Evolution of average trap energy level E_{avg} at different temperature.

We also perform dispersive (with $E_{min} = \infty$) H^+ drift/diffusion transport simulation within the R-D framework – a mechanism considered for N_{IT} generation in [33, 407]. Dispersive H^+ drift/diffusion with $E_{min} = \infty$ is implemented using [33, 411] –

$$\frac{\partial N_{H^+}}{\partial \tau} = D_{00} \frac{\partial^2 N_{H^+}}{\partial z^2} + \mu_{00} E_{ox} \frac{\partial N_{H^+}}{\partial z}, \quad (\text{B.8})$$

where $\mu_{00} = D_{00}/(k_B T/q)$ according to Einstein's relation and –

$$\tau(t) = \frac{N_C}{\nu_0} \left[\int_{-\infty}^{-k_B T \ln(\nu_0 t)} g(E_t) dE_t \right]^{-1}. \quad (\text{B.9})$$

Presence of high electric field ($E_{ox} \sim 5\text{-}10$ MV/cm) within the oxide region requires the use of Scharfetter-Gummel discretization scheme [413] for handling proton transport in a numerically stable manner, a problem routinely dealt in any drift-diffusion formalism. The simulation agrees well with the experiment (Figure B.3), reported in [33], albeit using unphysical parameters like: $\nu_0 = 6.33 \times 10^{-4}$ Hz, $E_0 = 105\text{meV}$, and $N_C/N_T = 5 \times 10^2$.

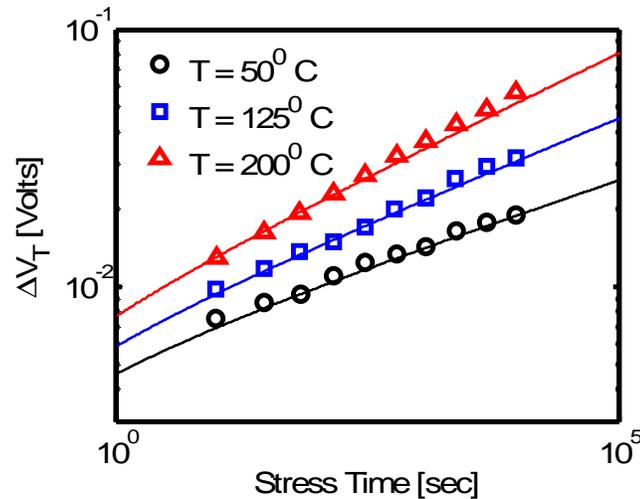


Figure B.3: Dispersive H^+ drift/diffusion transport simulation, implemented using Scharfetter-Gummel discretization, within the R-D framework agrees well with experimental data from [33], albeit using unphysical parameters like: $\nu_0 = 6.33 \times 10^{-4}$ Hz, $E_0 = 105\text{meV}$, and $N_C/N_T = 5 \times 10^2$.

B.2. Bond Dispersion Model

Bond dispersion (B-D) model, first proposed for explaining hot-carrier injection (HCI) experiments [301], became popular in NBTI-induced N_{IT} generation literature [32, 116, 414-416], when it was shown using charge-pumping (CP) measurement that generated N_{IT} re-passivates negligibly under the removal of V_{STS} [32, 414]. Existence of negligible N_{IT} relaxation led the researchers to believe that diffusion of hydrogen species after Si-H bond dissociation is so fast [417] that the Si-H bond dissociation always remains the rate-limiting step; *i.e.*, diffusion terms in R-D model has no significance. As

such, the following equations are used in explaining the N_{IT} dynamics of B-D model [32, 116, 301, 417]:

$$\frac{dN_{IT}}{dt} = \int dE_{A,F} k_F(E_{A,F}) g_F(E_{A,F}), \quad (\text{B.10})$$

where $\int g_F(E_{A,F}) dE_{A,F} = N_0$, $k_F \sim \exp(-E_{A,F}/k_B T)$ and $g_F(E_{A,F})$ indicates the normalized distribution of Si-H bonds having the following expression:

$$g_F(E_{A,F}) = \frac{N_0}{\sigma_F} \frac{\exp\left\{-\left(E_{A,F} - \langle E_{A,F} \rangle\right)/\sigma_F\right\}}{\left[1 + \exp\left\{-\left(E_{A,F} - \langle E_{A,F} \rangle\right)/\sigma_F\right\}\right]^2} \quad (\text{B.11})$$

with the dissociation energy, having a mean of $\langle E_{A,F} \rangle$ and standard deviation σ_F . Solving equations (B.10)-(B.11) one obtains

$$\frac{N_{IT}}{N_0} = 1 - \frac{1}{1 + (t/\tau)^\alpha}, \quad (\text{B.12})$$

where, $\tau \sim \exp(E_{A,F}/k_B T)$, $\alpha = (k_B T/\sigma_F)^p$, and p depends on the range of fit to the experimental data.

However, detailed experiment from several groups [129, 130, 226, 418, 419] later verified that the main assumption of B-D model, *i.e.*, the observation of negligible N_{IT} relaxation in CP experiments during NBTI stress, is a measurement artifact. As conventional CP experiment [420] has inherent recovery during the measurement [129, 130, 226, 418, 419], most of the N_{IT} relaxation happens before the actual measurement is started. So the assumption of negligible N_{IT} recovery, though valid for HCI experiments [114, 301], is not justified for NBTI experiments; which was recently confirmed by showing N_{IT} relaxation in ultra-fast version of CP measurement [419].

B.3. Statistical Mechanics Based Model

Recently, statistical mechanics is applied [148] for estimating the equilibrium concentration of hydrogen species (here, proton or H^+) near the interface, as well as N_{IT} .

Using the well-known concept of partition function from statistical mechanics, interfacial H^+ is estimated as –

$$N_{H^+}^{(0)} = N_i \exp\left(E_{H^+}/k_B T\right) \frac{N_0 - N_{IT}}{N_{IT}}, \quad (\text{B.13})$$

where N_i is the total density of interstitial H^+ sites at the oxide/substrate interface and E_{H^+} is energy level associated with bonded hydrogen at the oxide/substrate interface. Later, H^+ is assumed to drift through oxide (*i.e.*, drift term in equations (3.13) and (3.14) dominates), hence at long-stress time equation (3.14) leads to –

$$\frac{dN_{IT}}{dt} = \mu_{H^+} E_{ox} N_{H^+}^{(0)} = \mu_{H^+} E_{ox} N_i \exp\left(E_{H^+}/k_B T\right) \frac{N_0 - N_{IT}}{N_{IT}}. \quad (\text{B.14})$$

Assuming H^+ transport is dispersive, as in equation (B.1), and using Einstein's relation, *i.e.* $\mu_{H^+} = D_{H^+}/(k_B T/q)$, equation (B.14) can be expressed as –

$$\frac{dN_{IT}}{dt} = \frac{D_{00} (\omega_0 t)^{-\beta}}{k_B T / q} E_{ox} N_i \exp\left(E_{H^+}/k_B T\right) \frac{N_0 - N_{IT}}{N_{IT}}, \quad (\text{B.15})$$

which leads to a stretched exponential solution for N_{IT} generation.

Comparing this statistics based model with the R-D model, we can easily identify that there is a direct equivalence between equations (B.13) and (3.12), if we substitute k_F/k_R in equation (3.12) using $N_i \exp(E_{H^+}/k_B T)$ (such that, $E_{H^+} = E_{A,F} - E_{A,R}$) and assume $N_0 \gg N_{IT}$. So, the role of statistical mechanics [148] is only limited to the derivation of the equilibrium condition, which can alternatively derived by using simple rate equations like (3.2) [421]. So, this statistical mechanics-based model is simply an alternate representation of R-D model with H^+ dispersive drift/diffusion, hence should have the same limitations of dispersive transport, as discussed in section B.1, in explaining N_{IT} generation under NBTI stress condition.

B.4. Energy-level Perspective of N_{IT} Formation

This model for N_{IT} generation during NBTI stress [115] proposes that dissociation of Si-H bond and subsequent transport of hydrogen can be modeled by considering a chain of energy wells, as shown in Figure B.4. As such, by varying the barrier heights between the energy wells and also considering such barrier to have a distribution of values, it becomes possible to fit ΔV_T (or N_{IT} generation) with the use of just two or three energy wells – where, the three energy-well system is providing better fitting to the measurements. As we will show later, such an energy-well representation for N_{IT} generation is basically an alternate representation of the B-D model (alternative to the two-well system, having H^+ generation; see section B.4.3.1) and the R-D model (alternative to the three-well system, having H^+ drift/diffusion; see section B.4.3.2), both having unsubstantiated amount of dispersion in k_F , k_R and drift/diffusion.

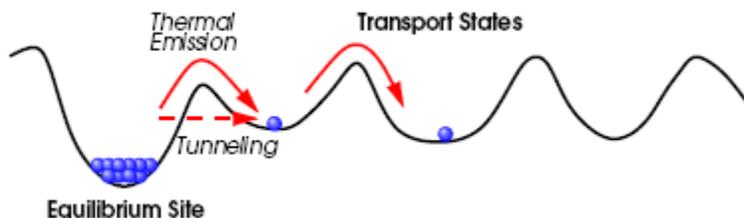


Figure B.4: Energy-well representation of N_{IT} generation (taken from [115]). Here, hydrogen is released from the Si-H bond (equilibrium state) thermally over a barrier, or tunnel through the barrier at $T < 1K$, into transport states. Dissociation of Si-H leaves dangling Si- bonds at the oxide/substrate interface.

B.4.1. Energy-well Representation

An energy-well representation was first used in [422] to explain the stretched exponential relaxation in electronic properties of hydrogenated amorphous-Si (a -Si:H) (see Figure B.5). In such model, H is initially assumed to be trapped in a -Si:H states (T states) and after dissociation H transports through intermediate I states, before getting

capturing in a reservoir R state. While moving through the I states, H can move back towards the T state, thus delaying the a -Si:H dissociation process. Thus, the generated dangling bond (a -Si:) can reduce the electronic properties (e.g. conductivity) of amorphous material in a stretched exponential manner [149], which can be fitted pretty well using this energy-well model [422].

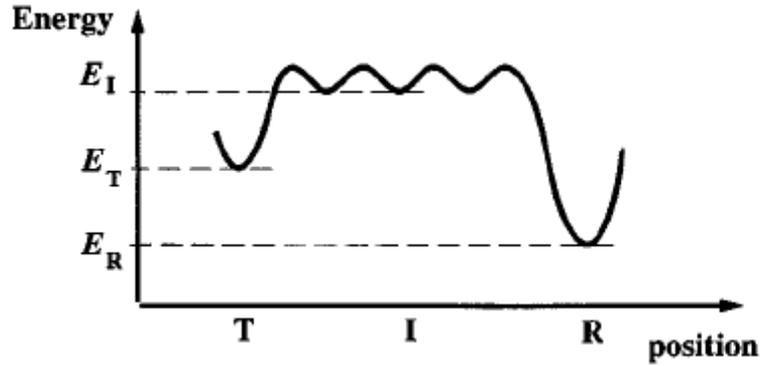


Figure B.5: Energy-well representation of stretched exponential relaxation in a -Si:H samples through dissociation of trapped hydrogen from T states (taken from [422]). Here, R refers to the ground state or reservoir, and I refers to the interstitial hydrogen states for diffusion through the crystal.

In this energy-well representation of stretched exponential relaxation, the reaction describing the trapping at T and R sites can be expressed as [422] –

$$\frac{d[T]}{dt} = -k_T [T] + \sigma_T [I] (N_T - [T]), \quad (\text{B.16})$$

$$\frac{d[R]}{dt} = -k_R [R] + \sigma_R [I] (N_R - [R]), \quad (\text{B.17})$$

where N_T is the concentration of sites available for trapping, $[T]$ is the concentration of trapped hydrogen in trapping sites, N_R is the concentration of sites available at reservoir, $[R]$ is the concentration of trapped hydrogen in reservoir, $[I]$ is the concentration of interstitial trapped hydrogen, k_T and k_R are the rate constant, and σ_T and σ_R are capture parameters.

B.4.2. Energy-well Representation of R-D Model

In this section, we will explain how a R-D system can be represented in terms energy-well parameters of equations (B.16)-(B.17). We first start with equation (B.16) and show how it is equivalent to Si-H bond dissociation reaction of. As shown in Figure B.6, these two equations are analogous, if we use the following replacements (where, δ is the interfacial thickness):

$$N_T \equiv N_0/\delta; \quad [T] \equiv (N_0 - N_{IT})/\delta; \quad [I] \equiv N_H^{(0)}; \quad N_T - [T] \equiv N_{IT}/\delta.$$

As a result, equation (B.16) becomes $\frac{1}{\delta} \frac{d(N_0 - N_{IT})}{dt} = -k_T \frac{(N_0 - N_{IT})}{\delta} + \sigma_T N_H^{(0)} \frac{N_{IT}}{\delta}$;

i.e.,

$$\frac{dN_{IT}}{dt} = k_T (N_0 - N_{IT}) - \sigma_T N_H^{(0)} N_{IT}, \tag{B.18}$$

which is similar to equation (3.2), when $X \equiv H$, $k_T \equiv k_F$, and $\sigma_T \equiv k_R$.

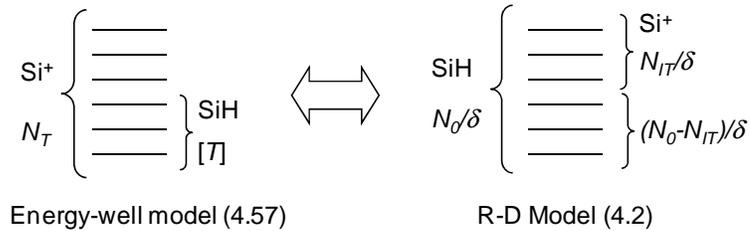


Figure B.6: Analogy between equations (B.16) and (3.2), with $X \equiv H$.

Next, for obtaining the analogy of equation (B.17) in R-D system, we use McKelvey’s flux method [423, 424] to show that forward and backward fluxes in carrier transport can be expressed as (referring to Figure B.7),

$$\frac{df}{dz} = -\alpha f + \alpha' b; \quad \frac{db}{dz} = -\alpha f + \alpha' b; \tag{B.19}$$

where f and b denote incident fluxes ($\text{cm}^{-2}\text{-sec}^{-1}$) from positive (forward flux) and negative sides (backward flux) respectively; α and α' represent characteristic decay length (cm^{-1}) for positive and negative fluxes respectively.

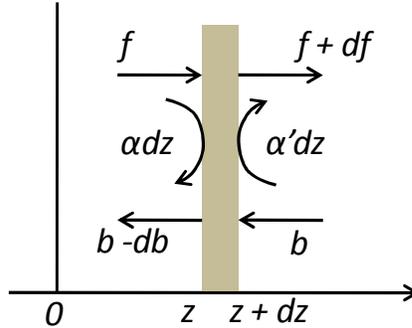


Figure B.7: Schematic for McKelvey's flux method for carrier transport. Here, f and b denote forward and backward fluxes, and α and α' represent characteristic decay length.

B.4.2.1. Energy-well Representation of R-D Model (H Diffusion)

Assuming similar energy levels for R and I in Figure B.5 (*i.e.*, though R acts as a reservoir of infinite capacity, the energy barrier for carrier transport from I to R and R to I is same), in equation (B.19), we can use $\alpha = \alpha'$, $f = [R]^+ v_T = [R] v_T / 2$ and $b = [R]^- v_T = [R] v_T / 2$; where, v_T is the thermal velocity. In other words, we are considering Arrhenius diffusion of hydrogen species, generated from Si-H dissociation (Figure B.8), as considered in R-D model with H-diffusion. Thus, from equation (B.19) we obtain –

$$J_H \equiv J_R = \frac{v_T}{2\alpha} \frac{d[R]}{dz} = D_H \frac{d[R]}{dz}, \quad (\text{B.20})$$

where J represents the net flux in the positive z direction. Now using similar energy barrier for carrier transport from I to R and R to I (*i.e.*, capture cross-section σ_H for hydrogen transport in and out of R is same), left-hand-side of equation (B.17) can be represented as –

$$\begin{aligned} \frac{d[R]}{dt} &\equiv \frac{J_H}{\delta} \equiv \frac{J_R}{\delta} = \sigma_H [R]_- (N_R - [R]_+) - \sigma_H [R]_+ (N_R - [R]_-) \\ &= \sigma_H N_R ([R]_- - [R]_+) = \delta \sigma_H N_R \frac{d[R]}{dz}. \end{aligned} \quad (\text{B.21})$$

Comparing equations (B.20) and (B.21), we have –

$$v_T \equiv \delta \sigma_H N_R, \quad \text{and} \quad D_H = \frac{v_T}{2\alpha} \equiv \delta^2 \sigma_H N_R \sim \exp(-E_{A,H}/k_B T). \quad (\text{B.22})$$

This proves that equation (B.17) will represent H diffusion, when energy levels for R and I are same. As such, the energy-well representation of Figure B.5 will be equivalent to R-D model with H-diffusion (Figure B.8), where the boundary condition between reaction and diffusion term (*i.e.*, equation (3.4) in R-D framework) can be equivalently represented using –

$$\begin{aligned} \frac{d[I]}{dt} &= -\frac{d[T]}{dt} - \frac{d[R]}{dt}, \\ \text{or,} \quad \frac{dN_H^{(0)}}{dt} &= \frac{1}{\delta} \frac{dN_{IT}}{dt} - \frac{1}{\delta} D_H \frac{dN_H^{(0)}}{dz}. \end{aligned} \quad (\text{B.23})$$

Comparing to equation (3.4), we are only missing a factor 2 in equation (B.23).

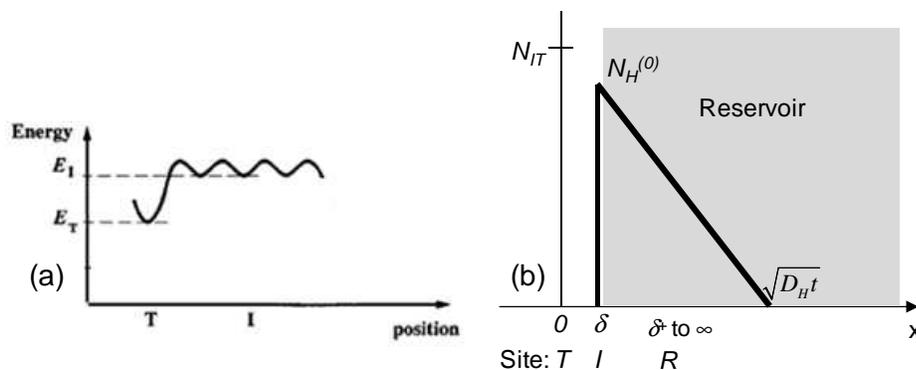


Figure B.8: Analogy between (a) energy-well representation with R and I states having same energy and (b) R-D system, having H diffusion. Analogous parameters are: $N_T \equiv N_0/\delta$; $[T] \equiv (N_0 - N_{IT})/\delta$; $[I] \equiv N_{H(0)}$; $N_T - [T] \equiv N_{IT}/\delta$; $k_T \equiv k_F$; $\sigma_T \equiv k_R$; and $D_H \equiv v_T/2\alpha = \delta^2 \sigma_H N_I$.

B.4.2.2. Energy-well Representation of R-D Model (H^+ Drift)

When energy levels of R and I in Figure B.5 are different, we have $\alpha \neq \alpha'$ in (B.19). Now, suppose the generated hydrogen species from Si-H dissociation is proton (H^+); *i.e.*, after $N_H^{(0)}$ in equation (B.16) is replaced by $N_{H^+}^{(0)}$, it represents the chemical reaction of equation (3.11). As negative bias is applied during N_{IT} generation, H^+ while drifting within the oxide (reservoir) will preferentially move towards the gate (Figure 3.2), *i.e.*, the energy barrier for carrier transport from I to R will be lower than the barrier for transport from R to I (Figure B.9). Thus, $\alpha > \alpha'$ and they can be represented using [424] $\alpha = \alpha_0 + \alpha_1/2$ and $\alpha' = \alpha_0 - \alpha_1/2$. Then equation (B.19) reduces to –

$$-(f - b) = \frac{1}{2\alpha_0} \frac{d(f + b)}{dx} + \frac{\alpha_1}{2\alpha_0} (f + b). \quad (\text{B.24})$$

Again, using $f = [R]^+ v_T = [R]v_T/2$ and $b = [R]^- v_T = [R]v_T/2$, equation (B.24) becomes –

$$J_{H^+} \equiv J_R = \frac{v_T}{2\alpha_0} \frac{d[R]}{dx} + \frac{\alpha_1 v_T}{2\alpha_0} [R] = D_{H^+} \frac{d[R]}{dx} + \mu_{H^+} E_{ox} [R]. \quad (\text{B.25})$$

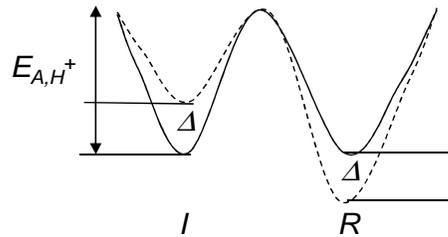


Figure B.9: When the transport species is charged hydrogen (H^+), then application oxide electric field under NBTI stress makes the transport from I to R (towards gate) easier compared to the transport from R to I (towards substrate).

Now, using same condition of Figure B.9 in the left-hand-side of equation (B.17), we can write –

$$\begin{aligned}
\frac{d[R]}{dt} &\equiv \frac{J_H}{\delta} \equiv \frac{J_R}{\delta} = \sigma_H \exp(\Delta/k_B T) [R]_- (N_R - [R]_+) \\
&\quad - \sigma_H \exp(-\Delta/k_B T) [R]_+ (N_R - [R]_-) \\
&\sim \delta \sigma_H N_R \frac{d[R]}{dz} \exp(\Delta/k_B T),
\end{aligned} \tag{B.26}$$

where Δ represents barrier height reduction due to applied oxide electric field for transport from I to R . For $\Delta < k_B T$, equation (B.26) can be written as

$$\frac{d[R]}{dt} \equiv \frac{J_H}{\delta} \equiv \frac{J_R}{\delta} \equiv \delta \sigma_H N_R \frac{d[R]}{dz} + 2\sigma_H N_R [R] \Delta/k_B T. \tag{B.27}$$

Comparing equations (B.25) and (B.27), we have –

$$v_T \equiv \delta \sigma_H N_R, \quad D_{H^+} = \frac{v_T}{2\alpha} \equiv \delta^2 \sigma_H N_R, \quad \text{and} \quad \mu_{H^+} E_{ox} \equiv \frac{\alpha_1 v_T}{2\alpha_0} \equiv \frac{2D_{H^+} \Delta}{\delta k_B T}. \tag{B.28}$$

Using Einstein's relation in equation (B.28) results –

$$qE_{ox} \delta \equiv 2\Delta. \tag{B.29}$$

Thus equations (B.27) and (B.17) will represent H^+ drift, when energy levels for R and I are different. As such, the energy-well representation of Figure B.5 will be equivalent to R-D model with H^+ drift/diffusion (Figure B.10), where the boundary condition between reaction and diffusion term (*i.e.*, equation (3.14) in R-D framework) can be equivalently represented using –

$$\begin{aligned}
\frac{d[I]}{dt} &= -\frac{d[T]}{dt} - \frac{d[R]}{dt}, \\
\text{or,} \quad \frac{dN_H^{(0)}}{dt} &= \frac{1}{\delta} \frac{dN_{IT}}{dt} - \frac{1}{\delta} D_{H^+} \frac{dN_{H^+}^{(0)}}{dz} - \frac{1}{\delta} \mu_{H^+} E_{ox} N_{H^+}^{(0)}.
\end{aligned} \tag{B.30}$$

Comparing to equation (3.14), we are again missing a factor 2 in equation (B.30).

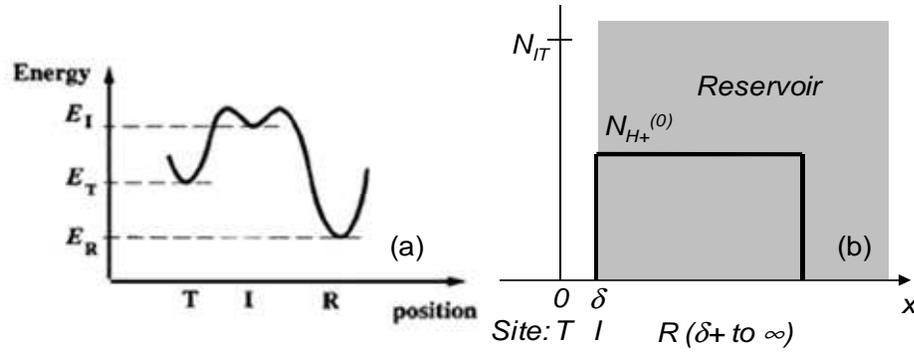


Figure B.10: Analogy between (a) energy-well representation with R and I states having same energy and (b) R-D system, having H^+ drift/diffusion. Analogous parameters are: $N_T \equiv N_0/\delta$; $[T] \equiv (N_0 - N_{IT})/\delta$; $[I] \equiv N_{H(0)}$; $N_T - [T] \equiv N_{IT}/\delta$; $k_T \equiv k_F$; $\sigma_T \equiv k_R$; $D_{H+} \equiv v_T/2\alpha = \delta^2 \sigma_H N_I$; and $\mu_{H+} E_{ox} \equiv 2D_{H+} \Delta/\delta k_B T$.

B.4.3. Energy-level Perspective: Revisited

In section B.4.2, we have shown how a chain of energy levels can be used for representing a R-D system, where McKelvey's flux method is used for explaining the analogy between energy-wells and drift/diffusion³⁸. Following this, we can easily show how the double-well model of [115] is equivalent to a bond-dispersive (B-D) model with H^+ generation (*i.e.*, by using the concepts of section B.2) and the triple well model of [115] is equivalent to a R-D model with H^+ drift/diffusion, along with bond-dispersion (*i.e.*, by adding the concepts discussed in sections B.2 and 3.5.2).

B.4.3.1. Double-well Model

The equations used in double well model (Figure B.11a [115]) are:

$$\frac{df_1}{dt} = -\frac{df_3}{dt} = -f_1 k_{13} + (1 - f_1) k_{31} f_3, \quad (\text{B.31})$$

³⁸ Please refer to Appendix C, where similar analogy is explained using the concept of generating function.

where $k_{13} = \nu \exp[-(E_2 - \Delta)/k_B T]$, and $k_{31} = \nu \exp[-(E_2 - \Delta - E_3 + 2\Delta)/k_B T]$. At equilibrium, well 1 (with potential $E_1 = 0$) is the energetically preferred configuration. After the application of the oxide electric field E_{ox} tilts the wells by an amount Δ (Figure B.11a), thus favoring the second well (having potential E_3 before the tilt and $E_3 - 2\Delta$ after the tilt).

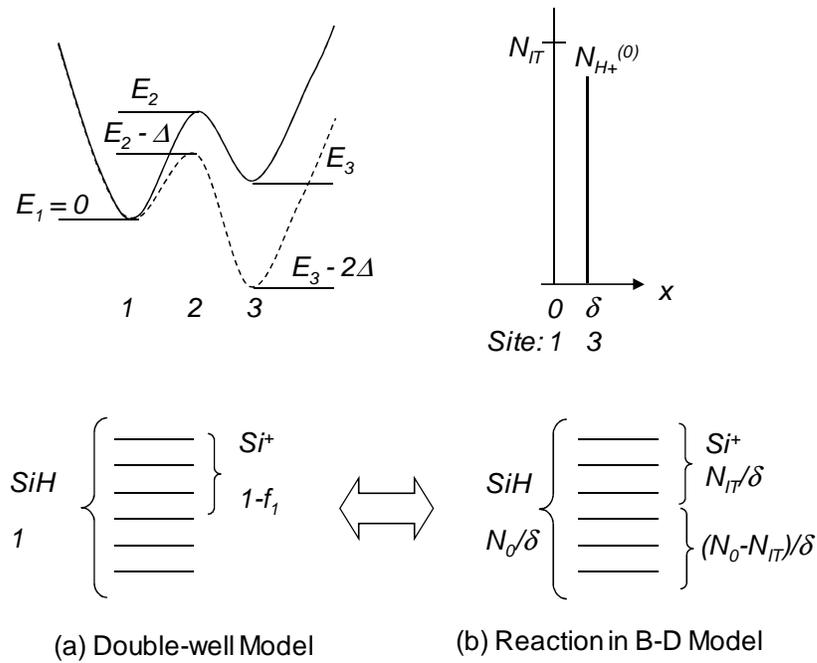


Figure B.11: Analogy between double-well model and B-D model with H^+ generation.

Following the analysis similar to Figure B.6, we can again establish an analogy between the double-well system and the reaction terms in R-D or B-D model (Figure B.11). As shown in Figure B.11, these two models are analogous, if we use the following replacements:

$$I \equiv N_0/\delta; \quad f_1 \equiv (N_0 - N_{IT})/\delta; \quad f_3 \equiv N_{H^+}^{(0)}; \quad I - f_3 \equiv N_{IT}/\delta.$$

As a result, equation (B.31) becomes $\frac{1}{\delta} \frac{d(N_0 - N_{IT})}{dt} = -k_{13} \frac{(N_0 - N_{IT})}{\delta} + k_{31} N_{H^+}^{(0)} \frac{N_{IT}}{\delta}$; *i.e.*,

$$\frac{dN_{IT}}{dt} = k_{13}(N_0 - N_{IT}) - k_{31}N_{H^+}^{(0)}N_{IT}, \quad (\text{B.32})$$

where the use of H^+ will be evident from the following discussion. Some additional analogy of $k_{13} \equiv k_F$ and $k_{31} \equiv k_R$ can also be drawn by comparing equations (3.2) and (B.32). As such, by replacing E_2 with $E_{A,F}$ (activation energy for forward dissociation of Si-H) and $E_2 - E_3$ with $E_{A,R}$ (activation energy for reverse annealing of Si-H), we can represent k_F and k_R using:

$$\begin{aligned} k_F &\sim \exp\left[-(E_{A,F} - \Delta)/k_B T\right], \\ \text{and } k_r &\sim \exp\left[-E_{A,R}/k_B T\right]\exp\left[-\Delta/k_B T\right]. \end{aligned} \quad (\text{B.33})$$

Now, the expression for k_F in equation (B.33) is equivalent to the one we proposed in Figure 3.16 [119, 128, 133], with Δ in equation (B.33) being analogous to aE_{ox} in Figure 3.16. In addition, equation (B.33) uses extra barrier Δ for reverse annealing (k_R) of dangling bonds, which was not used in Figure 3.16. Note that Figure 3.16 uses neutral hydrogen as the by-product of Si-H dissociation, thus the reverse passivation is considered to be independent of E_{ox} . Since equation (B.33) uses extra barrier Δ for k_R , reverse passivation is considered to be obstructed by E_{ox} , which is possible only if we have H^+ as the by-product of Si-H dissociation. Thus use of H^+ as by product in the double-well model, reverse reaction is suppressed during NBTI stress phase and comparatively enhanced during NBTI relaxation phase. This also justifies our use of $N_{H^+}^{(0)}$ in equation (B.32) for explaining the analogy between double-well systems and B-D model.

B.4.3.2. Triple Well Model

A triple-well model [115] incorporates an extra well to the right hand side of a double-well system of Figure B.11a. As discussed in section B.4.2, adding this extra well is equivalent to adding drift/diffusion within the overall framework of energy-wells. Since the barriers between the second and third energy-well in this model is electric field dependent, it is analogous to the representation of Figure B.9 between I and R states.

Thus, the energy wells in the triple-well model is equivalent to T, I, R states of the original energy-well representation [422] (Figure B.12). Thus, the newly proposed triple-well model is exactly same as the R-D model, with dispersion for Si-H forward and reverse reaction, as well as H^+ drift-diffusion.

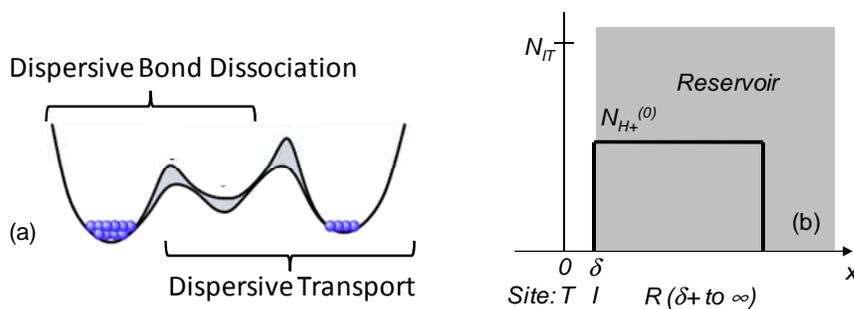


Figure B.12: (a) Triple-well representation for explaining N_{IT} formation has unsubstantiated amount of dispersion in Si-H bond dissociation and hydrogen transport [115]. (b) As discussed in section B.4.2, such a triple-well model is nothing but an alternate representation of R-D model.

C. DRIFT/DIFFUSION SYSTEM USING GENERATING FUNCTION

Following the analysis in [213-215], we can understand the energy-well representation of a drift/diffusion system (presented using McKelvey's flux method in section B.4.1) using the concept of generating functions. Here, generating functions are used to calculate mean first passage time (MFPT) or average arrival time or transit time of particles, hopping randomly in a one-dimensional lattice having $N + 1$ sites; hence it is used for calculate drift/diffusion fluxes for the same particle. As shown in Figure C.1, we denote the hopping sites by $j = 0, 1, 2, \dots, N$ and hopping is modeled by assigning the following transition probabilities for each lattice site, j :

- p_j The probability of hopping from site j to $(j+1)$ – right jump
- q_j The probability of hopping from site j to $(j-1)$ – left jump
- $1-p_j-q_j$ The probability of staying at site j (sojourn probability)

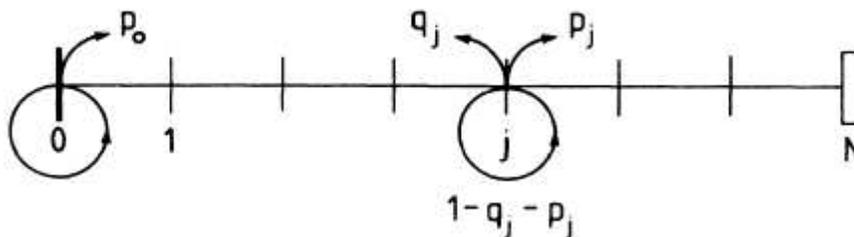


Figure C.1: Random walks in 1D lattice having $N+1$ sites ($j = 0, 1, 2, \dots, N$). Probability of right and left jumps at site j is represented by p_j and q_j respectively. Hence, $(1 - p_j - q_j)$ represents the staying probability at grid j . Taken from [214].

Consider that a particle starts its random walk at $j = 0$ (hence, $q_0 = 0$) towards site N . Hence, this 1D random walk can be modeled using a distribution of $2N-1$ random variables, *i.e.*, $p_0, p_1, \dots, p_{N-1}, q_1, q_2, \dots, q_{N-1}$. We need to find MFPT for the particle to reach at site N , which is average of first-passage time (FPT) over the ensemble of all possible random walks within the 1D lattice. Here FPT is defined as the number of steps required for the particle to reach the site N for the first time. Let us denote FPT and

MFPT by $t_{0,N}$ and $\overline{t_{0,N}}$ respectively. We determine $\overline{t_{0,N}}$ by using the concept of generating function, which is very popular in solving such discrete problems [425].

Suppose, $g_{i,j}(n)$ is the probability for a particle to start at site ‘ i ’ and eventually reach site ‘ j ’ for the first time in ‘ n ’ steps. Then the corresponding generating function, $G_{i,j}(z)$, is defined as –

$$G_{i,j}(z) = \sum_{n=0}^{\infty} z^n g_{i,j}(n) \quad (\text{C.1})$$

and hence, MFPT will be –

$$\overline{t_{0,N}} = \frac{\sum_{n=0}^{\infty} n g_{0,N}(n)}{\sum_{n=0}^{\infty} g_{0,N}(n)} = \frac{G_{0,N}'(z)}{G_{0,N}(z)} \Big|_{z=1} = \frac{d}{dz} \ln G_{0,N}(z) \Big|_{z=1} \quad (\text{C.2})$$

Similarly, formalism of generating function can also be used to redefine different transition probabilities as below:

Table C.1: Transition probabilities from lattice site ‘ i ’ in Figure C.1.

| | Probability for the random walk | z-probability using generating function |
|--|---|---|
| Particles stays at site ‘ i ’ for any ‘ n ’ steps | $\hat{\chi}_i(n) = [1 - (q_i + p_i)]^n$ | $\chi_i(z) = [1 - z\{1 - (q_i + p_i)\}]^{-1}$ |
| Particle to jump from site ‘ i ’ to ‘ $i-1$ ’ | q_i | zq_i |
| Particle to jump from site ‘ i ’ to ‘ $i+1$ ’ | p_i | zp_i |
| Particle making a first passage from site ‘ $i-1$ ’ to ‘ i ’ | $g_{i-1,i}(n)$ | $G_{i-1,i}(z)$ |
| Particle making a first passage from site ‘ i ’ to ‘ $i+1$ ’ | $g_{i,i+1}(n)$ | $G_{i,i+1}(z)$ |

Considering all these, z -probability of a particle (no need for normalization) to go to site ' $i+1$ ' from ' i ' will be –

$$G_{i,i+1}(z) = \frac{\chi_i(z) z p_i}{1 - z q_i G_{i-1,i}(z) \chi_i(z)} \quad (\text{C.3})$$

Equation (C.3) is not valid near boundaries, i.e. for i close to 0 and N . Substituting for $\chi_i(z)$ we obtain –

$$G_{i,i+1}(z) = \frac{z p_i}{1 - z [1 - (q_i + p_i) + q_i G_{i-1,i}(z)]} \quad (\text{C.4})$$

which is equal to unity at $z=1$ (follows from $G_{0,1}(z) = 1$, then all others are 1). Finally, we have $G_{0,N}(z) = \prod_{j=1}^N G_{j-1,j}(z)$, which can be used in equation (C.2) to calculate MFPT as –

$$\overline{t_{0,N}} = \frac{d}{dz} \ln G_{0,N}(z) = \frac{d}{dz} \sum_{j=1}^N \ln G_{j-1,j}(z) = \sum_{j=1}^N G'_{j-1,j}(z) / G_{j-1,j}(z) = \sum_{j=1}^N G'_{j-1,j}(z) \quad (\text{C.5})$$

Differentiating equation (C.4) with respect to z and using $G_{i,i+1}(z=1) = 1$ for any i , we have –

$$G'_{i,i+1}(z=1) = \frac{1}{p_i} + \frac{q_i}{p_i} G'_{i-1,i}(z=1) \quad (\text{C.6})$$

Using $i = 0, 1, 2, \dots, N-1$ in equation (C.6), we can write –

$$\begin{aligned} G'_{0,1}(z=1) &= \frac{1}{p_0} \\ G'_{1,2}(z=1) &= \frac{1}{p_1} + \frac{q_1}{p_1} G'_{0,1}(z=1) = \frac{1}{p_1} + \frac{1}{p_0} \frac{q_1}{p_1} \\ G'_{2,3}(z=1) &= \frac{1}{p_2} + \frac{q_2}{p_2} G'_{1,2}(z=1) = \frac{1}{p_2} + \frac{q_2}{p_2} \left(\frac{1}{p_1} + \frac{q_1}{p_1} \frac{1}{p_0} \right) = \frac{1}{p_2} + \left(\frac{1}{p_1} \frac{q_2}{p_2} + \frac{1}{p_0} \frac{q_1}{p_1} \frac{q_2}{p_2} \right) \\ G'_{3,4}(z=1) &= \frac{1}{p_3} + \frac{q_3}{p_3} G'_{2,3}(z=1) = \frac{1}{p_3} + \left[\frac{1}{p_2} \frac{q_3}{p_3} + \left(\frac{1}{p_1} \frac{q_2}{p_2} \frac{q_3}{p_3} + \frac{1}{p_0} \frac{q_1}{p_1} \frac{q_2}{p_2} \frac{q_3}{p_3} \right) \right], \dots \end{aligned}$$

Putting these values in equation (C.5), MFPT becomes –

$$\overline{t_{0,N}} = \sum_{i=1}^N G'_{i-1,i}(z) = \sum_{k=0}^{N-1} \frac{1}{p_k} + \sum_{k=0}^{N-2} \frac{1}{p_k} \sum_{i=k+1}^{N-1} \prod_{j=k+1}^i \frac{q_j}{p_j} \quad (\text{C.7})$$

Now, to obtain expression for steady state flux (*i.e.*, J_i is constant at any site i), we use (C_i is the concentration of particle at site i and δ is the uniform grid spacing) –

$$J/\delta = p_0 C_0 - q_1 C_1 = p_1 C_1 - q_2 C_2 = p_2 C_2 - q_3 C_3 = \dots = p_{N-1} C_{N-1} - q_N C_N \quad (\text{C.8})$$

A sequential elimination of $C_{N-1}, C_{N-2}, \dots, C_2, C_1$ from equation (C.8) yields –

$$\begin{aligned} C_{N-1} &= \frac{1}{p_{N-1}} \left(\frac{J}{\delta} + q_N C_N \right); & C_{N-2} &= \frac{1}{p_{N-2}} \left[\frac{J}{\delta} \left(1 + \frac{q_{N-1}}{p_{N-1}} \right) + \frac{q_{N-1}}{p_{N-1}} q_N C_N \right] \\ C_{N-3} &= \frac{1}{p_{N-3}} \left[\frac{J}{\delta} \left(1 + \frac{q_{N-2}}{p_{N-2}} \left(1 + \frac{q_{N-1}}{p_{N-1}} \right) \right) + \frac{q_{N-2}}{p_{N-2}} \frac{q_{N-1}}{p_{N-1}} q_N C_N \right] \\ &\dots\dots\dots \\ C_1 &= \frac{1}{p_1} \left[\frac{J}{\delta} \left(1 + \sum_{r=2}^{N-1} \prod_{i=2}^r \frac{q_i}{p_i} \right) + \prod_{i=2}^{N-1} \frac{q_i}{p_i} q_N C_N \right] \end{aligned} \quad (\text{C.9})$$

Putting values from equation (C.9) into (C.8), we obtain –

$$\frac{J}{\delta} = \frac{p_0 C_0 - \prod_{i=1}^{N-1} \frac{q_i}{p_i} q_N C_N}{\left(1 + \sum_{r=1}^{N-1} \prod_{i=1}^r \frac{q_i}{p_i} \right)} \quad (\text{C.10})$$

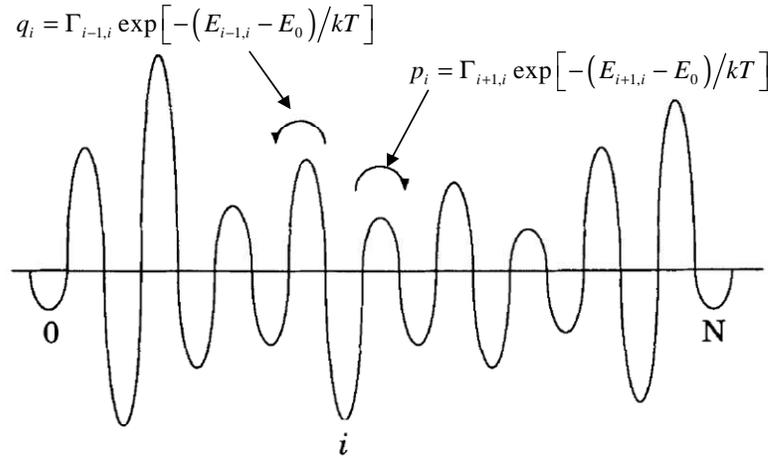
Equation (C.10) can be used for calculating the steady state flux (*e.g.*, drift or diffusion fluxes) of particles moving from site 0 to N in any random 1D network, as discussed next.

C.1. Calculation of Diffusion Flux:

We consider zero external electric field in this situation. However, the barrier between neighboring sites can still be random (like the one shown in Figure C.2). As a result, the random walk has symmetric transition rates, *i.e.*, $p_i = q_{i+1}$, $p_{i-1} = q_i$, etc. Under this condition, equations (C.7) and (C.10) will be–

$$\overline{t_{0,N}} = \sum_{i=0}^{N-1} \frac{(i+1)}{p_i}, \quad (\text{C.11})$$

$$\frac{J}{\delta} = \frac{p_0 C_0 - q_1 (= p_0) C_N}{\left(1 + \sum_{i=1}^{N-1} \frac{q_1 (= p_0)}{p_i}\right)} = \frac{C_0 - C_N}{\sum_{i=0}^{N-1} \frac{1}{p_i}} = \frac{C_0 - C_N}{\sum_{i=0}^{N-1} \frac{1}{\Gamma_{i+1,i} \exp[-(E_{i+1,i} - E_0)/kT]}}. \quad (\text{C.12})$$



$$q_i = p_{i-1} = \Gamma_{i,i-1} \exp[-(E_{i,i-1} - E_0)/kT] \quad \text{and} \quad p_i = q_{i+1} = \Gamma_{i,i+1} \exp[-(E_{i,i+1} - E_0)/kT]$$

Figure C.2: One dimensional random network with transition rates between neighboring sites. E_0 represents the site energy at $i = 0$.

Equation (C.12) represents Fick's first law (note that it contains both drift and diffusion flux), which can be rewritten as (lattice spacing is assumed unity)-

$$\frac{J}{\delta} = \frac{(C_0 - C_N)/L}{\frac{1}{L} \sum_{i=0}^{N-1} \frac{1}{\Gamma_{i+1,i} \exp[-(E_{i+1,i} - E_0)/kT]}} \quad \text{or, } J = D_T (C_0 - C_N)/L = D_T \frac{dC}{dx},$$

where the co-efficient D_T can be expressed as (note that it contains both drift and diffusion flux) -

$$D_T^{-1} = \frac{1}{\delta L} \sum_{i=0}^{N-1} \frac{1}{\Gamma_{i+1,i} \exp[-(E_{i+1,i} - E_0)/kT]}. \quad (\text{C.13})$$

Transit time under such condition is,

$$\tau_{transit} = \frac{L^2}{2D_T} = \frac{L}{2\delta} \sum_{i=0}^{N-1} \frac{1}{\Gamma_{i+1,i} \exp[-(E_{i+1,i} - E_0)/kT]} = \frac{L}{2\delta} \sum_{i=0}^{N-1} \frac{1}{p_i}. \quad (\text{C.14})$$

Now, for uniform diffusion, we have $E_{i+1,i} = E_0$ and equation (C.13) represents diffusion co-efficient (D_H), which can be expressed as –

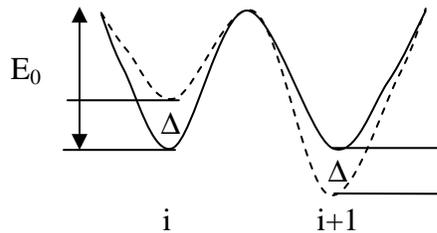
$$D_H^{-1} = \frac{1}{\delta L} \sum_{i=0}^{N-1} \frac{1}{\Gamma_{i+1,i}}. \quad (\text{C.15})$$

Also assuming, $\Gamma_{i+1,i} = \Gamma_0$, equations (C.14) and (C.11) become –

$$\tau_{transit} = \frac{L}{2\delta} \sum_{i=0}^{N-1} \frac{1}{\Gamma_0} = \frac{N^2}{2\Gamma_0}; \quad \overline{t_{0,N}} = \sum_{i=0}^{N-1} \frac{(i+1)}{p_i} = \sum_{i=0}^{N-1} \frac{(i+1)}{\Gamma_0} = \frac{N(N+1)}{2\Gamma_0}. \quad (\text{C.16})$$

Thus, we obtain $\tau_{transit} \sim \overline{t_{0,N}}$; the reason behind this is rooted on the fact that although $\overline{t_{j,j+1}}$ is a reflection of first passage time, it is calculated by considering the injection of infinite number of injections at j ; i.e., $\overline{t_{j,j+1}} = \frac{\sum_{n=1}^{\infty} t_{j,j+1}(n)}{\sum_{n=1}^{\infty} n}$. As such, $\overline{t_{j,j+1}}$ is denoted as MFPT, not just FPT.

C.2. Calculation of Drift/Diffusion Flux:



Suppose, a uniform bias field along $0 \rightarrow N$ is influencing particles to move preferentially in the forward direction (i to $i+1$) compared to its movement in reverse

direction (*i.e.*, $p_i' > q_{i+1}'$; where `s are used to denote probabilities under this modified condition). Under this condition, we can write $E_{i+1,i}' = E_0 - \Delta$ and $E_{i,i+1}' = E_0 + \Delta$. Hence, using $p_i' = q_{i+1}' e^{2\Delta/kT}$ for any 'i' in equation (C.7), we can write –

$$\overline{t_{0,N}} = \frac{1}{e^{\Delta/kT}} \sum_{k=0}^{N-1} \frac{1}{p_k} + \frac{1}{e^{\Delta/kT}} \sum_{k=0}^{N-2} \frac{1}{e^{-2k\Delta/kT}} \sum_{i=k+1}^{N-1} \frac{1}{e^{2i\Delta/kT}} \frac{1}{p_i} \quad (\text{C.17})$$

where p_i indicates the values without any applied bias. Under the influence of very high bias ($\Delta \gg kT$), equation (C.17) becomes

$$\overline{t_{0,N}} \approx \frac{1}{e^{\Delta/kT}} \sum_{k=0}^{N-1} \frac{1}{p_k} \quad (\text{C.18})$$

Hence, drift-diffusion flux can be calculated as,

$$J = \delta \frac{C_0 - C_N}{\overline{t_{0,N}}} = e^{\Delta/kT} \frac{(C_0 - C_N)/L}{\frac{1}{\delta L} \sum_{i=0}^{N-1} \frac{1}{\Gamma_{i+1,i}}} = e^{\Delta/kT} D \frac{dC}{dx} \quad (\text{C.19})$$

which is similar to the one obtained in equation (B.26).

D. MOMENTS OF STATISTICAL DISTRIBUTION: USING GENERATING FUNCTION

In Appendix C, we have discussed the method for calculating mean first passage time for the random 1D walk of Figure C.1 using (*c.f.*, equation (C.2)) –

$$\overline{t}_{0,N} = \frac{\sum_{n=0}^{\infty} n g_{0,N}(n)}{\sum_{n=0}^{\infty} g_{0,N}(n)} = \frac{G_{0,N}'(z)}{G_{0,N}(z)} \Big|_{z=1} = \frac{d}{dz} \ln G_{0,N}(z) \Big|_{z=1}, \quad (\text{D.1})$$

where $g_{i,j}(n)$ is the probability for a particle to start at site ‘ i ’ and eventually reach site ‘ j ’ for the first time in ‘ n ’ steps; and $G_{i,j}(z)$ is the corresponding generating function, which is defined as –

$$G_{i,j}(z) = \sum_{n=0}^{\infty} z^n g_{i,j}(n). \quad (\text{D.2})$$

Actually, such power-series generating are well adapted to determine the other moments of statistical distributions, as well. As shown in [425], generating function can be used to derive the variance of a statistical distribution, as shown below:

$$\begin{aligned} \sigma_{0,N}^2 &= \frac{\sum_{n=0}^{\infty} (n - \overline{t}_{0,N})^2 g_{0,N}(n)}{\sum_{n=0}^{\infty} g_{0,N}(n)} = \frac{\sum_{n=0}^{\infty} n^2 g_{0,N}(n)}{\sum_{n=0}^{\infty} g_{0,N}(n)} - 2\overline{t}_{0,N} \frac{\sum_{n=0}^{\infty} n g_{0,N}(n)}{\sum_{n=0}^{\infty} g_{0,N}(n)} + \overline{t}_{0,N}^2 \\ &= \frac{\sum_{n=0}^{\infty} n^2 g_{0,N}(n)}{\sum_{n=0}^{\infty} g_{0,N}(n)} - \overline{t}_{0,N}^2 = \frac{G_{0,N}''(z) + G_{0,N}'(z)}{G_{0,N}(z)} \Big|_{z=1} - \left[\frac{G_{0,N}'(z)}{G_{0,N}(z)} \Big|_{z=1} \right]^2 \\ &= \left[\frac{d}{dz} \ln G_{0,N}(z) + \frac{d^2}{dz^2} \ln G_{0,N}(z) \right] \Big|_{z=1}. \end{aligned} \quad (\text{D.3})$$

Using, generating function for each step, equation (D.3) can be re-written as:

$$\sigma_{0,N}^2 = \sum_{j=1}^N G_{j-1,j}'(z) \Big|_{z=1} + \sum_{j=1}^N G_{j-1,j}''(z) \Big|_{z=1} - \sum_{j=1}^N [G_{j-1,j}'(z)]^2 \Big|_{z=1}. \quad (\text{D.4})$$

Equation (D.4), along with equation (C.4), is used for deriving the expression of equation (4.14).

E. PARAMETRIC CHARACTERIZATION OF SCATTERING THEORY

Since its proposition in [333], scattering theory in nanoscale MOS transistors has been substantiated both by detailed numerical calculations [334, 336, 337, 426] and by indirect experimental evidences based on transistors from one [338] or multiple [339] CMOS technology nodes. These studies established the importance of low-field mobility ($\mu_{eff} \sim v_{inj}\lambda$ [333]), or equivalently back-scattering (only) near the top-of-the-barrier, for explaining the correlation in current transport in both linear and saturation regimes. In theoretical efforts to verify the scattering theory, one may need to consider additional scattering mechanisms in calculation of μ_{eff} for shorter channel transistors [427], as well as influence of ‘down-stream scattering’ on self-consistent potential near the top-of-the-barrier [340]. On the other hand, experimental study of scattering theory involves modulating μ_{eff} or λ , calculating B_{lin} and B_{sat} , and thereby studying its impact on both $I_{D,lin}$ and $I_{D,sat}$. Modulation of μ_{eff} is obtained in [338] by varying V_G and T on a single transistor; whereas, the same is obtained by variation of strain over different technology nodes in [339] (see Figure E.1a). Unfortunately, such studies always require a calculation of v_{inj} [338, 339]. Moreover, when strain is used for varying μ_{eff} , it also changes band-structure and increases velocity overshoot near virtual source (thus reduces ℓ) [339], which needs to be theoretically estimated before the results can be compared to the prediction of the scattering theory.

In the following sections, we analyze scattering theory by modulating μ_{eff} or λ on a single transistor through the generation of coulomb scatterers or defects at the oxide/substrate interface of a PMOS transistor. As expected from scattering theory, we show that change in μ_{eff} ($\Delta\mu_{eff}$) impacts both $I_{D,lin}$ and $I_{D,sat}$, with the impact on $I_{D,sat}$ being lower compared to that on $I_{D,lin}$ (i.e., $B_{sat} > B_{lin}$). The observed variation of B_{lin} and B_{sat} due to defect generation follows essentially the time-zero (or without defect) B_{lin} vs. B_{sat} relationship, *provided* the effect of V_T shift (from defect generation) is accounted for (section E.2). Furthermore, we also analyze the variation in ℓ with V_G and ΔV_T in strongly inverted channel (section E.4). Thus our work allows an intuitive generalization of the

scattering theory for applications in reliability-aware IC design. We also believe this work to be an interesting example of using defect generation (similar to that of [428]) as a controlled technique to verify the fundamentals of a theory.

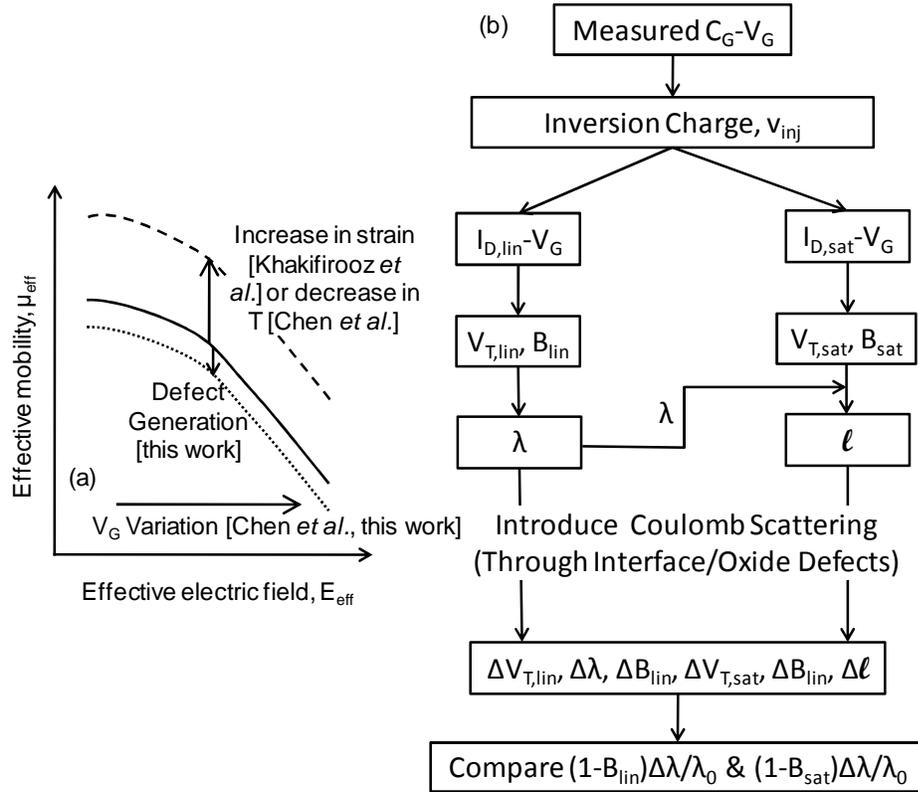


Figure E.1: (a) Experimental analysis of scattering theory can be performed either by varying μ_{eff} using V_G and T or using strain. This work broadens the application of scattering theory, by varying μ_{eff} through defects, as well as V_G . (b) Flowchart for the experimental steps performed in validating scattering theory, where λ is varied by creation of additional coulomb scattering.

E.1. Experimental Details

We vary μ_{eff} or λ systematically by generating controlled amount of defects mostly at the interface of a *single* PMOS transistor (Width/Length: $W/L = 1/0.13 \mu\text{m}$; effective oxide thickness, $EOT \sim 1.5 \text{ nm}$) by using NBTI stress [347]. Whereas, in a NMOS

transistor we change μ_{eff} or λ by having increased amount of electron trapping into oxide defects (pre-existing or generated) under the application of PBTI stress [43]. For estimating changes in different quantities for equations (8.8)-(8.11), we take I_D - V_G sweeps in linear ($V_{DS,lin} = -0.025V$) and saturation ($V_{DS,sat} = -1.1V$) regimes at pre-defined time intervals. We also use C_G - V_G to compute inversion charge (Q_{inv}) and hence calculate v_{inj} using degenerate carrier statistics and single sub-band occupancy [344]. Later, $V_{T,lin}$ and $V_{T,sat}$ are determined from the x-axis intercepts of the slopes on I_D - V_G curves, drawn at maximum transconductance ($g_{m,max}$) point [249]. Estimated Q_{inv} , v_{inj} , and $V_{T,lin}$ are then used in $I_{D,lin}$ - V_G and $I_{D,sat}$ - V_G for calculating B_{lin} - V_G (through equations (8.8) or (8.10)) and B_{sat} - V_G (through equations (8.9) or (8.11)), respectively at different levels of generated defect. B_{lin} - V_G is then used for obtaining λ - V_G , which is hence utilized in B_{sat} - V_G for estimating ℓ - V_G . Here, the magnitude of defect generation is estimated from the shift in $V_{T,lin}$ ($\Delta V_{T,lin}$) [347]. As such, $\Delta V_{T,lin}$ in our experiment plays the analogous role of T in [338] or levels of strain in [339]. Finally, B_{lin} and B_{sat} at different V_G are calculated for different levels of $\Delta V_{T,lin}$ and used for checking the consistency with scattering theory. Figure E.1b summarizes all the above steps we are using for validating scattering theory.

E.2. Variation of Ballistic Efficiencies

We calculate $(1-B_{lin})\Delta\lambda/\lambda_0$ and $(1-B_{sat})\Delta\lambda/\lambda_0$ using equations (8.10)-(8.11), which has Maxwell-Boltzmann approximation for Fermi Integrals, to estimate the impact of $\Delta\lambda$ on $I_{D,lin}$ and $I_{D,sat}$, respectively [333]; where, λ_0 is mean-free path before defect generation and $\Delta\lambda$ is the change in λ due to defect generation. As anticipated in [333], the impact of $\Delta\lambda$ is more prominent on $I_{D,lin}$ compared to $I_{D,sat}$, *i.e.*, $(1-B_{sat})/(1-B_{lin}) < 1$ (see Figure E.2a), for $L = 0.13\mu m$ transistor used in this study. We also obtain L dependency for $(1-B_{sat})/(1-B_{lin})$ using $L = 0.13\sim 1\mu m$ PMOS/NMOS transistors (Figure E.2b) and thus observe this ratio to decrease as L is reduced (*i.e.*, B_{sat} is higher for reduced L , whereas B_{lin} increases negligibly with L), which is again consistent with scattering theory [333]. Here, one needs to note that use of exact Fermi integrals (*i.e.*, equations (8.8)-(8.9)), though will change the magnitudes of B_{lin} and B_{sat} (Figure E.3), but will not change the

broad conclusions of Figure E.2, *i.e.*, $(1-B_{sat})/(1-B_{lin}) < 1$ and this decreases with decrease in L .

Lines in Figure E.3 also indicate negligible variation in the B_{lin} vs. B_{sat} relationship (measured at different V_G) due to interfacial defect generation (maximum $\Delta V_{T,lin} \sim 0.06V$, corresponding to a defect density of $\sim 8 \times 10^{11} \text{ cm}^{-2}$). This is because the change in λ (less than 1nm) and ℓ due to defect generation has negligible impact on B_{lin} (as $\lambda \ll L$) and also causes less than 2% increase in B_{sat} (for same B_{lin}). In spite of this approximate invariance of B_{lin} vs. B_{sat} (due to the defects), defect generation does increase both B_{lin} and B_{sat} , when measured at fixed V_G (symbols in Figure E.3). More interestingly, the increasing trend of B_{lin} and B_{sat} follows approximately the same B_{lin} vs. B_{sat} relationship measured with/without defects (lines in Figure E.3). This suggests that the B_{lin} vs. B_{sat} relationship depends mainly on $|V_G - V_T|$, *i.e.*, increase in $|V_T|$ (due to defect generation) has the same impact as the decrease in $|V_G|$, at least for the defect density of $< 10^{12} \text{ cm}^{-2}$ used in this study. And as discussed earlier, use of exact Fermi integrals though changes the magnitudes of B_{lin} and B_{sat} , but do not affect the broad conclusions.

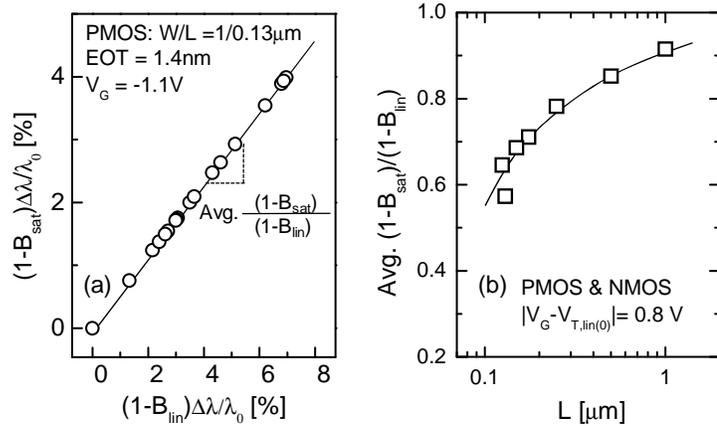


Figure E.2: (a) Creation of interfacial defects (marked using $\Delta V_{T,lin}$) for a PMOS strained transistor increases λ (Figure 8.4d). Resultant change in λ effects both $I_{D,lin}$ and $I_{D,sat}$, which is monitored here using $(1-B_{lin})\Delta\lambda/\lambda_0$ and $(1-B_{sat})\Delta\lambda/\lambda_0$. (b) Estimated average $(1-B_{sat})/(1-B_{lin})$ for PMOS (due to interfacial defects) and NMOS (due to oxide

defects) transistors (having L of 0.13~1 μm) decreases with decrease in L , which indicates an increase in ballisticity with reduced L .

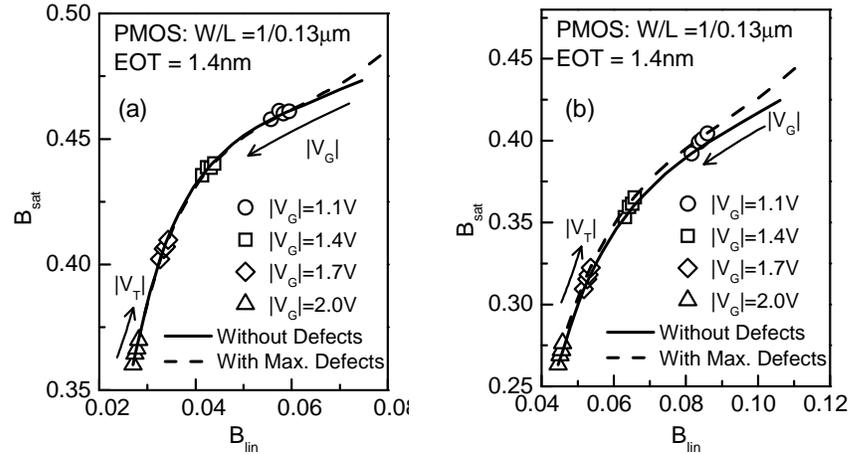


Figure E.3: Interfacial defects in a PMOS transistor increases both B_{lin} and B_{sat} , when measured at fixed V_G , relationship between B_{lin} and B_{sat} remains approximately invariant of defects (a) Using Maxwell-Boltzmann approximation for Fermi integrals, *i.e.*, equations (8.10)-(8.11); (b) Using equations (8.8)-(8.9). Both the figures prove that for a particular transistor B_{lin} and B_{sat} depends mainly on $|V_G - V_T|$.

E.3. Variation of λ

As μ_{eff} , for the PMOS transistor under study, increases due to N_{IT} generation, we always observe an increase of λ , whether we characterize it using equations (8.10)-(8.11) (see Figure E.4a) or using equations (8.8)-(8.9) (Figure E.4b). Figure E.4 also shows that use of Maxwell-Boltzmann approximation underestimates the carrier flux from drain end [341], thus underestimates μ_{eff} , as well as λ . Moreover, Figure E.4b suggests $2k_B T \mu_{eff} / q v_{inj}$ is not a good approximation for λ , as sometimes used in literature [338], and thus suggests $\lambda \sim 2k_B T \mu_{eff} \mathcal{S}_0(\eta_F) / q v_{inj} \mathcal{S}_{-1}(\eta_F)$ [341] for a better estimation.

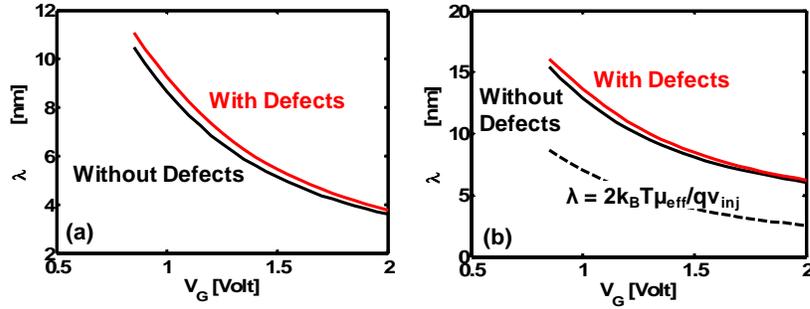


Figure E.4: Variation of λ extracted using (a) Maxwell-Boltzmann approximation for Fermi Integrals, *i.e.*, equations (8.10)-(8.11) or (b) exact expressions of equations (8.8)-(8.9). Figure (b) also shows the underestimation of λ induced by using $\lambda = 2k_B T \mu_{eff} / q v_{inj}$ [341].

E.4. Variation of ℓ

The parameter ℓ is very critical within the framework of scattering theory. Recently, there have been lots of studies to identify the extent of this length along the MOS channel [342, 343]. The extraction of B_{lin-V_G} (hence, $\lambda-V_G$) and B_{sat-V_G} at different levels of V_T enables us to have more insight into variation of ℓ , which was not reported earlier.

Our calculation involving equations (8.10)-(8.11) indicates a decrease of ℓ with $|V_G|$ (Figure E.5a), which is also reported in [338, 429]. Additionally, ℓ is observed to saturate at larger $|V_G|$. Figure E.5a also shows that the generation of defects improves ℓ at lower $|V_G|$, whereas keeps ℓ invariant to defects at higher $|V_G|$. Next, we decrease $|V_{DS,sat}|$ to 0.6V, which not only cause a reduction of ℓ at lower $|V_G|$, but at higher $|V_G|$, ℓ also starts to increase (Figure E.5b).

We also study the limitation of equations (8.10)-(8.11) by calculating variation of ℓ with $|V_G|$ at $|V_{DS,sat}| = 1.1V$ using equations (8.8)-(8.9) (Figure E.5c) and hence compare it with Figure E.5a. We observe that use of exact Fermi integrals causes the turnaround of ℓ variation to appear, even at higher $|V_{DS,sat}|$, which only appeared at lower $|V_{DS,sat}|$ when Maxwell-Boltzmann approximation was used (Figure E.5b). We will perform an

integrated analysis involving transport simulation and further experiment to explain variation in ℓ .

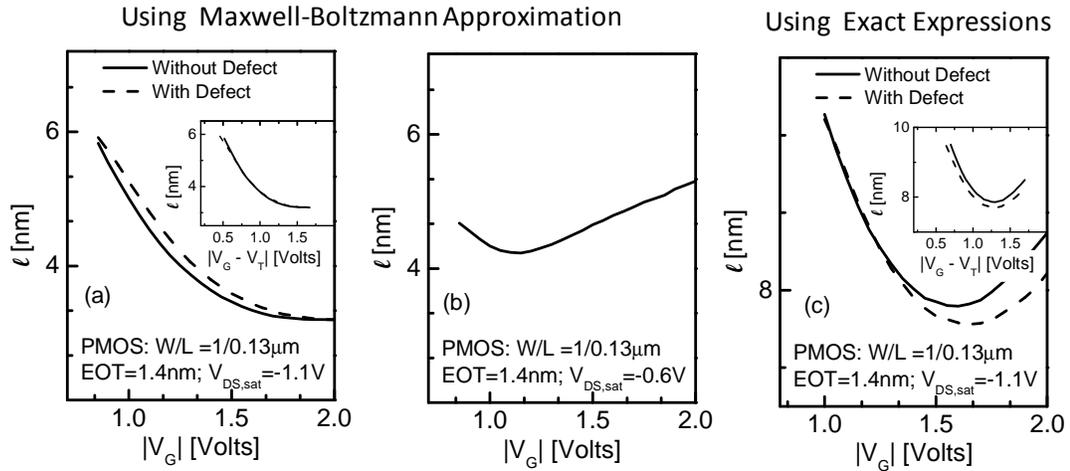


Figure E.5: (a) Extracted ℓ vs. $|V_G|$, using Maxwell-Boltzmann approximation (*i.e.*, equations (8.10)-(8.11)), shows an initial decrease, then saturation, when measured at $|V_{DS,sat}| = 1.1 \text{ V}$. Whereas, (b) variation of ℓ indicates a turnaround when measured at lower $|V_{DS,sat}| = 0.6 \text{ V}$. (c) Turn-around in ℓ variation starts to show up even at $|V_{DS,sat}| = 1.1 \text{ V}$, when exact expressions for drain currents or equations (8.8)-(8.9) are used.

E.5. Scattering Theory Analysis: Summary

In this appendix, we have broadened the application of scattering theory with time-dependent parameter shifts and show that – despite dramatic changes in the device parameters – the relationship between the ballistic efficiencies are essentially invariant with respect to its time-zero value, once the effect of threshold voltage shift is taken into account. Our work also provides a novel scheme for the justifying the parametric variation in the scattering theory and allows an intuitive, yet powerful, generalization of the theory for applications in reliability-aware IC design.

VITA

VITA

Ahmad Ehteshamul Islam was born in Sylhet, Bangladesh in 1979. He received B.S. in electrical and electronic engineering (EEE) from Bangladesh University of Engineering and Technology (BUET) in 2004. Since 2005, he is pursuing the Ph.D. degree in the School of Electrical Engineering and Computer Science, Purdue University, West Lafayette, IN and is expected to receive the degree by May 2010. During 2004–2005, he was a Lecturer with the Department of EEE, BUET. His research mainly focuses on studying the impact of variabilities in semiconductor devices. He is currently working with variation resilience aspects in nanoscale transistors. He has authored and coauthored more than 20 journals and conference papers. Mr. Islam has been a Student Member of the IEEE Electron Devices Society (since 2002) and American Physical Society (since 2008) and also serves as a Reviewer for several IEEE, Elsevier, APS, and Electrochemical Society journals. He is the recipient of Kintar-Ul-Haque Gold Medal (2005) for his undergraduate result and IEEE EDS Ph.D. Fellowship (2008), Intel Foundation PhD Fellowship (2009-2010) for his work on transistor reliability. He was also involved as student team leader (2002-2003) and faculty co-supervisor (2004-2005) of EEE, BUET team in International Future Energy Challenge 2003 and 2005 competitions, respectively.

The following contains a list of (12) journals and (15) conference papers that has been published during Mr. Islam's Ph.D. research at Purdue.

Journals:

- [1] **A. E. Islam**, C. Augustine, K. Roy, and M. A. Alam, "Extent of Variation Resilience in Strained CMOS: From Transistors to Digital Circuits", under review in IEEE Trans. on Electron Devices.

- [2] **A. E. Islam**, and M. A. Alam, "On the Possibility of Degradation-Free Field Effect Transistors", *Applied Physics Letter*, 92, 173504, 2008.
- [3] **A. E. Islam**, G. Gupta, K. Ahmed, S. Mahapatra, M. A. Alam, "Optimization of Gate Leakage and NBTI for Plasma-Nitrided Gate Oxides by Numerical and Analytical Models", *IEEE Trans. on Electron Devices*, 55(5), pp. 1143-1152, 2008.
- [4] **A. E. Islam**, H. Kufluoglu, D. Varghese, S. Mahapatra, and M. A. Alam, "Recent Issues in Negative Bias Temperature Instability: Initial Degradation, Field-Dependence of Interface Trap Generation, and Hole Trapping Effects and Relaxation", (Invited Paper) *IEEE Trans. on Electron Devices*, 54(9), pp. 2143-2154, 2007.
- [5] **A. E. Islam**, H. Kufluoglu, D. Varghese, and M. A. Alam, "A Critical Analysis of Short-term Negative Bias Temperature Instability Measurements: Explaining the effect of time-zero delay for On-the-fly Measurements", *Applied Physics Letter*, 90, 083505, 2007.
- [6] **A. E. Islam**, A. Haque, "Accumulation gate capacitance of MOS devices with ultra-thin high-K gate dielectrics: Modeling and Characterization", *IEEE Trans. on Electron Devices*, 53(6), pp. 1364-1372, 2006. (Best paper award in Engineering & Technology, University Grant Commission-Bangladesh, 2006)
- [7] M. M. Satter, **A. E. Islam**, D. Varghese, M. A. Alam, and A. Haque, "Extraction of Interface Trap States of MOS Devices on Alternative High-Mobility Substrates from the Low Frequency Gate C-V Characteristics," under review in *IEEE Trans. on Electron Devices*.
- [8] M. Masuduzzaman, **A. E. Islam**, R. Degraeve, M. Cho, M. Zahid, and M. A. Alam, "Experimental Identification of Regions A and B in Multi-Frequency Charge Pumping," under review in *IEEE Electron Device Letters*.
- [9] S. Deora, V. D. Maheta, **A. E. Islam**, M. A. Alam, and S. Mahapatra, "A Common Framework of NBTI Generation and Recovery in Plasma Nitrided SiON p-MOSFETs", *IEEE Electron Device Letters*, 30(9), pp. 978-980, 2009.

- [10] S. Mahapatra, V. D. Maheta, S. Deora, E. N. Kumar, S. Purawat, C. Olsen, K. Ahmed, **A. E. Islam** and M. A. Alam, "Material Dependence of Negative Bias Temperature Instability (NBTI) Stress and Recovery in SiON p-MOSFETs", (Invited Paper) Journal of Electrochemical Society, 19(2), pp. 243-263, 2009.
- [11] S. Mahapatra, V. D. Maheta, **A. E. Islam**, and M. A. Alam, "Isolation of NBTI Stress Generated Interface Trap and Hole-Trapping Components in PNO p-MOSFETs", IEEE Trans. on Electron Devices, 56(2), pp. 236-242, 2009.
- [12] M. Masuduzzaman, **A. E. Islam**, and M. A. Alam, "Exploring the Capability of Multi-Frequency Charge Pumping in Resolving Location and Energy Levels of Traps within Dielectric", IEEE Trans. on Electron Devices, 55(12), pp. 3421-3431, 2008.

Conference Proceedings:

- [13] **A. E. Islam**, and M. A. Alam, "Mobility Enhancement Due to Charge Trapping and Defect Generation: Physics of Self-Compensated BTI", to appear in Proc. of Int. Reliability Phys. Symp. (IRPS), 2010.
- [14] **A. E. Islam**, S. Mahapatra, S. Deora, V. D. Maheta, and M. A. Alam, "On The Differences Between Ultra-fast NBTI Experiments and Reaction-Diffusion Theory", International Electron Devices Meeting (IEDM) Tech. Digest, pp. 733-736, 2009.
- [15] **A. E. Islam**, and M. A. Alam, "Self-Compensating Defect Generation in Advanced CMOS Substrates", International Integrated Reliability Workshop (IIRW), session 8.3, 2009.
- [16] **A. E. Islam**, J. H. Lee, W. H. Wu, A. Oates and M. A. Alam, "Universality of Interface Trap Generation and Its Impact on I_D Degradation in Strained/Unstrained PMOS Devices During NBTI Stress", International Electron Devices Meeting (IEDM) Tech. Digest, pp. 107-110, 2008.

- [17] **A. E. Islam**, V. D. Maheta, H. Das, S. Mahapatra, and M. A. Alam, "Mobility Degradation Due to Interface Traps in Plasma Oxinitride PMOS Devices", Proc. of Int. Reliability Phys. Symp. (IRPS), pp. 87-96, 2008.
- [18] **A. E. Islam**, E. N. Kumar, H. Das, S. Purawat, V. D. Maheta, H. Aono, E. Murakami, S. Mahapatra, and M. A. Alam, "Theory and Practice of On-the-fly and Ultra-fast V_T Measurements for NBTI Degradation: Challenges and Opportunities", International Electron Devices Meeting (IEDM) Tech. Digest, pp. 805-808, 2007. (Nominee for Roger A. Haken Best Student Paper Award)
- [19] **A. E. Islam**, G. Gupta, S. Mahapatra, A. T. Krishnan, K. Ahmed, F. Nouri, A. S. Oates, and M. A. Alam, "Gate Leakage vs. NBTI for Plasma Nitrided Oxides: Characterization, Physical Principles and Optimization", International Electron Devices Meeting (IEDM) Tech. Digest, pp. 12.4, 2006.
- [20] A. Zain, **A. E. Islam**, and M. A. Alam, "A Theoretical Study of Negative Bias Temperature Instability in p-type NEMFET," to appear in University Government Industry Micro/Nano Symposium, 2010.
- [21] M. Masduzzaman, **A. E. Islam**, and M. A. Alam, "A multi-probe correlated bulk defect characterization scheme for ultra-thin high- κ dielectric", to appear in Proc. of Int. Reliability Phys. Symp. (IRPS), 2010.
- [22] M. A. Alam and **A. E. Islam**, "On the Reliability of and Self-Compensation in Strained Transistors", (Invited Paper) International Conference on Solid State Devices and Materials (SSDM), session B-9-1, 2009.
- [23] M. Masduzzaman, **A. E. Islam**, and M. A. Alam, "Physics and Mechanics of Dielectric Trap Profiling by Multi-frequency Charge Pumping Method", Proc. of Int. Reliability Phys. Symp. (IRPS), pp. 13-20, 2009.
- [24] J. H. Lee, W. H. Wu, **A. E. Islam**, M. A. Alam and A. S. Oates, "Separation method of hole trapping and interface trap generation and their roles in NBTI Reaction-Diffusion model", Proc. of Int. Reliability Phys. Symp. (IRPS), pp. 745-746, 2008.

- [25] E. N. Kumar, V. D. Maheta, S. Purawat, S. Rani, **A. E. Islam**, K. Ahmed, M. A. Alam and S. Mahapatra, “Material Dependence of NBTI Physical Mechanism in Silicon Oxynitride (SiON) p-MOSFETs: A Comprehensive Study by Ultra-Fast On-The-Fly (UF-OTF) IDLIN Technique”, International Electron Devices Meeting (IEDM) Tech. Digest, pp. 809-812, 2007.
- [26] K. Kang, K. Kim, **A. E. Islam**, M. A. Alam, and K. Roy, “Characterization and Estimation of Circuit Reliability Degradation under NBTI using On-Line IDDQ Measurement”, Proceedings of 44th Design Automation Conference (DAC), p. 20.1, 2007.
- [27] S. Mahapatra, K. Ahmed, D. Varghese, **A. E. Islam**, G. Gupta, L. Madhav, D. Saha and M. A. Alam, “On the Physical Mechanism of NBTI in Silicon Oxynitride p-MOSFETs: Can Differences in Insulator Processing Conditions Resolve the Interface Trap Generation versus Hole Trapping Controversy?”, Proc. of Int. Reliability Phys. Symp. (IRPS), pp. 1-9, 2007.

Non-Refereed Talks and Poster Presentations

- **A. E. Islam**, “Compensating Fluctuation/Variation in Nanoscale Transistors”, Frontiers in Scalable Nanostructured Materials and Interfaces, Purdue University, 2009.
- **A. E. Islam**, “Interfacial Defects: History, Modeling, Effects & Optimization”, Student Seminar, IEEE EDS (Bangladesh Chapter), United International University & East West University, 2009.
- **A. E. Islam**, and M. A. Alam, “Characterization and Modeling of Trap Generation: A Primer on Why and How the Transistors Degrade”, Birck Nanotechnology Center Annual Review, Purdue University, 2008.
- **A. E. Islam**, H. Kuflluoglu, K. Kang, and M. A. Alam, “Integrated Framework for Reliability and Process-Variation Aware Design Methodology for VLSI Circuits”, Semiconductor Research Corporation GRC ICSS Annual Review, Univ. of Massachusetts at Amherst, 2007.

- **A. E. Islam**, H. Kufluoglu, and M. A. Alam, “Modeling of Negative Bias Temperature Instability and Hot Carrier Injection”, Semiconductor Research Corporation GRC ICSS Annual Review, Univ. of Illinois at Urbana-Champaign, 2007.
- H. Kufluoglu, **A. E. Islam**, K. Kang, B. C. Paul, K. Roy, and M. A. Alam, “Computational Modeling of Negative Bias Temperature Instability for Reliability-aware VLSI Design”, Network for Computational Nanotechnology (NCN) Annual Review, Purdue University, 2006.

Outreach

- **A. E. Islam**, H. Kufluoglu and M. A. Alam, “Modeling Interface-Defect Generation (MIG)”, deployed for public use at www.nanohub.org; 271 users around the world (Last update: February 12, 2010).