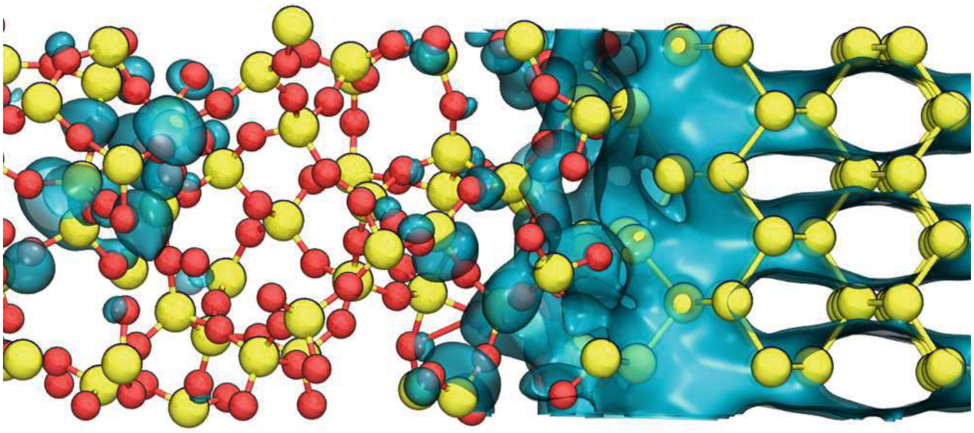
Dissertation

Multiscale Modeling of Defect-Related Reliability Phenomena in Nanoelectronics



zur Erlangung des akademischen Grades
Doktor der technischen Wissenschaften

eingereicht and der
Technische Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

von
Dipl.-Ing. **Dominic Waldhör**
Matr. Nr. 01226100
geboren am 11. Mai 1992 in Linz., Österreich

unter Betreuung von
Univ.Prof. Dipl.-Ing. Dr.techn. **Tibor Grasser**

Wien, Oktober 2024

Abstract

The ongoing rapid miniaturization of integrated circuits has led to numerous breakthroughs, including global communication networks, industrial automation, and artificial intelligence (AI). In particular AI is expected to drive further device down-scaling in the coming years to meet its increasing computational demands. However, as transistors — the fundamental building blocks of digital circuits — approach atomic scales, new challenges emerge. A key issue is the unavoidable presence of point defects within the gate insulator or at the channel/insulator interface. These defects can capture or emit charges during device operation, causing device characteristics to drift, a phenomenon known as bias temperature instability (BTI). Additionally, the same charge trapping processes contribute to random telegraph noise (RTN), which persists even at cryogenic temperatures and complicates the realization of robust solid-state qubits. Since even a single defect can render a small-scale device inoperable, a detailed understanding of defects and the charge trapping physics is essential for accurately modeling and predicting non-ideal device behavior.

This work employs a multiscale modeling approach to investigate both the microscopic charge trapping mechanisms and their macroscopic impact on device performance. At the atomic level, the electronic structure of defects is explored using density functional theory (DFT), providing insights into the charge trapping process and enabling the prediction of key defect parameters, such as thermodynamic charge trap levels and relaxation energies. These parameters can then be integrated into macroscopic device models to, for instance, predict threshold voltage drifts over a device's lifetime under specific operating conditions. Conversely, experimental data from accelerated stress tests can be used to infer underlying defect parameters and identify responsible defects through comparisons with DFT predictions.

Several new models and methods have been developed and implemented in the compact physics device simulator *Comphy*, which is based on nonradiative multiphonon (NMP) theory to describe charge trapping at the device level. By combining DFT results with *Comphy*, new insights in device reliability physics have been achieved. For example, it is demonstrated that self-trapped electrons (polarons) in the amorphous gate oxide contribute to gate-leakage currents via trap-assisted tunneling. Furthermore, persistent noise in CMOS devices at cryogenic temperatures has been linked to Si dangling bonds (P_b centers) at the Si/SiO_2 interface. To facilitate the extraction of defect parameters from degradation experiments, a novel method called the *Effective Single Defect Decomposition* (ESiD) is introduced. This method enables fast and robust parameter extractions from averaged defect responses in large-area devices, offering an alternative to more time-consuming single-defect characterizations.

In addition to studying pre-existing oxide defects, this work confirms earlier experimental observations of Si−H bond dissociation in SiO_2 through hole trapping, which generates electrically active Si dangling bonds in the oxide. Beyond these more applied advances, this thesis also makes a contribution to NMP theory by incorporating nonlocal electron-phonon couplings into the widely used 1-dimensional treatment of nonradiative transitions.

Kurzfassung

Die fortlaufende rasante Miniaturisierung integrierter Schaltkreise hat zahlreiche Durchbrüche ermöglicht, darunter globale Kommunikationsnetzwerke, industrielle Automatisierung und Künstliche Intelligenz (KI). Insbesondere wird erwartet, dass KI in den kommenden Jahren die weitere Verkleinerung von Bauelementen vorantreiben wird, um den steigenden Rechenanforderungen gerecht zu werden. Allerdings treten hier neue Herausforderungen auf, da Transistoren – die grundlegenden Bausteine digitaler Schaltungen – atomare Größenordnungen erreichen. Ein zentrales Problem ist die unvermeidbare Präsenz von Punktdefekten innerhalb des Gate-Isolators oder an der Kanal-Isolator-Grenzfläche. Diese Defekte können während des Betriebs Ladungen einfangen oder abgeben, was zu einem Drift der Bauelementeigenschaften führt, ein Phänomen, das als Bias-Temperature-Instability (BTI) bekannt ist. Darüber hinaus führt der Ladungseinfang an Defekten zu Random Telegraph Noise (RTN), das selbst bei kryogenen Temperaturen bestehen bleibt und beispielsweise die Realisierung robuster Qubits in Festkörpern erschwert. Da bereits ein einzelner Defekt ein nanoelektronisches Bauelement unbrauchbar machen kann, ist ein detailliertes Verständnis von Defekten und der Physik des Ladungseinfangs entscheidend für die genaue Modellierung und Vorhersage des nicht-idealen Bauelementverhaltens.

Diese Arbeit verfolgt einen Mehrskalenmodellierungsansatz, um sowohl die mikroskopischen Mechanismen des Ladungseinfangs als auch deren makroskopischen Auswirkungen auf die Charakteristiken des Bauelements zu untersuchen. Auf atomarer Ebene wird die elektronische Struktur von Defekten mithilfe der Dichtefunktionaltheorie (DFT) erforscht, um Einblicke in den Einfangprozess zu gewinnen und wichtige Defektparameter wie das thermodynamische Defektlevel und die Relaxationsenergien vorherzusagen. Diese Parameter können dann in makroskopische Bauelementmodelle integriert werden, um beispielsweise die Verschiebung der Schwellenspannung über die Lebensdauer eines Bauelements unter spezifischen Betriebsbedingungen vorherzusagen. Umgekehrt können experimentelle Daten aus beschleunigten Stresstests genutzt werden, um zugrunde liegende Defektparameter abzuleiten und die verantwortlichen Defekte durch Vergleiche mit DFT-Vorhersagen zu identifizieren.

Mehrere neue Modelle und Methoden wurden entwickelt und im kompakten physikbasierten Bauelementsimulator *Comphy* implementiert, der auf der Nonradiative Multiphonon (NMP) Theorie basiert, um den Ladungseinfang im Oxid auf Bauelementebene zu beschreiben. Durch die Kombination von DFT-Erkenntnissen mit *Comphy* wurden neue Erkenntnisse in der Zuverlässigkeitsphysik von Bauelementen erreicht. So wird beispielsweise gezeigt, dass Polaronen im amorphen Gate-Oxid durch Trap-Assisted Tunneling zu Gate-Leckströmen beitragen. Darüber hinaus wurde das anhaltende Rauschen in CMOS-Bauelementen bei kryogenen Temperaturen mit Si-Dangling-Bonds (P_b-Zentren) an der Si/SiO$_2$-Grenzfläche in Verbindung gebracht. Um die Extraktion von Defektparametern aus Degradationsexperimenten zu erleichtern, wird eine neuartige Methode namens *Effective Single Defect Decomposition* eingeführt. Diese Methode ermöglicht eine schnelle und robuste Parameterextraktion aus einfachen BTI Experimenten und bietet daher eine Alternative zu zeitaufwendigeren Einzeldefekt-Charakterisierungen.

Neben der Untersuchung bereits bestehender Oxiddefekte bestätigt diese Arbeit frühere experimentelle Beobachtungen über die Dissoziation von Si$-$H Bindungen durch Löchereinfang, wodurch elektrisch aktive Si-Dangling-Bonds im Oxid entstehen. Über diese anwendungsbezogenen Fortschritte hinaus trägt diese Dissertation auch grundlegend zur NMP-Theorie bei, indem nicht-lokale Elektron-Phonon-Kopplungen in die weit verbreitete eindimensionale Behandlung nichtstrahlender Übergänge einbezogen werden.

Contents

Contents vii

List of Figures xi

List of Tables xv

List of Abbreviations xvii

1 Introduction 1
- 1.1 MOS Devices . . . 1
- 1.2 Reliability Issues . . . 2
  - 1.2.1 Bias Temperature Instability . . . 3
  - 1.2.2 Random Telegraph Noise . . . 4
  - 1.2.3 Trap-Assisted Tunneling . . . 4
  - 1.2.4 Hot-Carrier Degradation . . . 5
- 1.3 4-State Model . . . 5
- 1.4 Thesis Outline . . . 6

I First Principles Modeling 9

2 Computational Methods 11
- 2.1 Density Functional Theory . . . 11
  - 2.1.1 Born-Oppenheimer Approximation . . . 12
  - 2.1.2 Hohenberg-Kohn Theorems . . . 14
  - 2.1.3 Kohn-Sham Theory . . . 15
  - 2.1.4 Exchange-Correlation Functionals . . . 18
  - 2.1.5 Pseudopotentials . . . 19
- 2.2 CP2K Software Package . . . 21
  - 2.2.1 Gaussian and Plane Wave Method . . . 22
  - 2.2.2 Auxiliary Density Matrix Method . . . 24

2.3 Lattice Dynamics . . . 25
2.3.1 Phonons . . . 25
2.3.2 Thermal Transitions . . . 27
2.3.3 Molecular Dynamics . . . 30
2.3.4 Beyond Classical Dynamics . . . 33
2.4 Defect Calculations . . . 33
2.4.1 Defect Formation Energies . . . 34
2.4.2 Finite-Size Corrections . . . 34
2.4.3 Charge Transition Levels . . . 35
2.5 Computational Settings . . . 36

3 Nonradiative Multiphonon Transitions 39
3.1 Basic Theory . . . 39
3.1.1 Nonradiative Multiphonon Emission . . . 39
3.1.2 Electron-Phonon Coupling in CP2K . . . 45
3.2 Extension of the Effective Mode Model . . . 48
3.2.1 Constrained Phonon Modes . . . 48
3.2.2 GaP:$Zn_{Ga}O_P$ as a Test Case . . . 49
3.2.3 Environmental Coupling . . . 52
3.3 Classical Limit . . . 56
3.3.1 Classical Lineshape Function . . . 56
3.3.2 Minimum Energy Crossing Point . . . 58

4 Defects in Gate Dielectrics 61
4.1 Si−H Bond Dissociation in SiO_2 . . . 61
4.1.1 Si−H Bonds in MOS Structures . . . 62
4.1.2 Earlier Theoretical Works . . . 63
4.1.3 Si−H Model . . . 64
4.1.4 Results . . . 65
4.2 Modeling Amorphous Oxides . . . 71
4.2.1 Melt-and-Quench Technique . . . 72
4.2.2 Structural Verification . . . 73
4.3 Defects in a-SiO_2 . . . 75
4.3.1 Oxygen Vacancies . . . 76
4.3.2 Hydrogen-Induced Defects . . . 77
4.3.3 Self-Trapped Charges . . . 78
4.3.4 Trap Level Distributions . . . 80
4.3.5 Effective Mode Barriers . . . 82
4.4 Calcium Fluoride as a Gate Dielectric . . . 84
4.4.1 Material Properties . . . 84
4.4.2 Si-Impurity Defects . . . 85

II Device Modeling

II Device Modeling 91

5 Compact Modeling Framework – *Comphy* 93
- 5.1 Introduction 93
- 5.2 Models and Features 94
  - 5.2.1 Parameter Extraction Methods 95
  - 5.2.2 Trap-Assisted Tunneling 95
  - 5.2.3 Charge Trapping at Cryogenic Temperatures 95
  - 5.2.4 Coupling to TCAD 96
- 5.3 Device Model 97
- 5.4 Defect Model 99
  - 5.4.1 Defect States 100
  - 5.4.2 Limitations of the 2-State Defect Model 102
  - 5.4.3 Transition Rates 104
  - 5.4.4 Defect Parameters 105
- 5.5 Gate-leakage Currents 108
  - 5.5.1 Tsu-Esaki Model 108
  - 5.5.2 Efficient Trap-Assisted Tunneling Current Computation 109
- 5.6 Cryogenic Charge Trapping 111
  - 5.6.1 WKB-based approximation 112
  - 5.6.2 Closed Form for Linear Coupling 114

6 Defect Identification 115
- 6.1 Experimental Characterization Techniques 115
  - 6.1.1 Extended Measure-Stress-Measure Scheme 116
  - 6.1.2 Random Telegraph Noise Analysis 117
  - 6.1.3 Time-Dependent Defect Spectroscopy 117
- 6.2 Parameter Extraction Methods 120
  - 6.2.1 Fitting Problem 120
  - 6.2.2 Gaussian Defect Bands 121
  - 6.2.3 Effective Single Defect Decomposition (ESiD) 122
- 6.3 Applications 125
  - 6.3.1 Trap Extractions with ESiD 125
  - 6.3.2 RTN in Cryo-CMOS 131
  - 6.3.3 Gate-Leakage Currents through Polarons 133

7 Conclusions & Outlook 137

References 141

List of Publications 173

List of Figures

1.1 NBTI degradation observed on a pMOS with SiON dielectric. 3
1.2 RTN signals measured on a small-area pMOS 4
1.3 4-State model for oxide defects. 6

2.1 PESs of a diatomic molecule in two different electronic states i and j . . . 14
2.2 Illustration of pseudo-potentials 21
2.3 Minimum Energy Path for Thermal Transitions 27
2.4 Formation Energy Diagram for a Hydrogen Bridge Defect in a-SiO_2 36

3.1 Schematic Illustration of Shockley-Read-Hall recombination 40
3.2 Parameters of the harmonic effective mode model 45
3.3 Comparison of Electron-Phonon Couplings between CP2K and VASP . . 47
3.4 Electron Trapping at the $Zn_{Ga}O_P$ Defect in GaP 50
3.5 Constrained mode approach applied to the $Zn_{Ga}O_P$ defect in GaP 51
3.6 Coupling distribution among different constrained modes 52
3.7 Effect of coupling to bath modes on the configuration coordinate diagram and the capture coefficient 55
3.8 Similar to Fig. 3.7 but with a modified displacement ΔQ 55
3.9 Classical limit for NMP charge transitions 57
3.10 Schematic transition between PESs in multiple dimensions 59

4.1 Band diagram of a Si/SiO_2 interface with the Kohn-Sham energies of the Si−H bond and the Si dangling bond in various configurations 63
4.2 Model structure of an isolated Si dangling bond. 65
4.3 Charge transition levels of the Si−H bond and a Si dangling bond (Si-DB) within the SiO_2 bandgap 66
4.4 Configurations of the Si−H bond before and after capturing a hole 67
4.5 PECs for the hole capture process of a Si−H bond 68
4.6 Hole capture cross sections for the Si−H bond for harmonic and anharmonic PECs 69

4.7 Migration steps of H^+ after hole capture and Si−H bond dissociation . . . 70
4.8 Potential energy along the H^+ migration path obtained from NEB calculations 71
4.9 Schematic depiction of the Melt-and-Quench procedure for creating amorphous oxide structures . . . 72
4.10 Comparison of experimental and theoretical structure factors for a-SiO_2 . . 74
4.11 Comparison of experimental and theoretical structure factors for a-SiO_2 . . 75
4.12 Different configurations of the oxygen vacancy in a-SiO_2 . . . 76
4.13 Different configurations of the hydrogen bridge defect in a-SiO_2 . . . 77
4.14 Different configurations of the hydroxyl-E' center in a-SiO_2 . . . 78
4.15 Inverse participation ratio of electronic states in amorphous SiO_2 . . . 79
4.16 Self-trapped electrons in amorphous SiO_2 . . . 80
4.17 Schematic trap level distributions from experiments and DFT predictions for various defect candidates in a-SiO_2 . . . 81
4.18 Resulting PECs of the effective mode approach and the minimum energy path (MEP) for representative defects in a-SiO_2 . . . 82
4.19 Correlation plots showing the agreement for capture and emission barriers for the PECs of the effective mode and the minimum energy path for various bias conditions . . . 83
4.20 Si(111)/CaF_2/MoS_2 gate stack for novel 2D electronics . . . 85
4.21 PDOS and atomic structure of the Si_i and Si_{Ca} defect in CaF_2 . . . 86
4.22 Formation energy diagram for the Si_i and Si_{Ca} defect in CaF_2 . . . 87
4.23 Trap levels of the Si_i and Si_{Ca} defects within the bandgap of CaF_2 together with band edges of common 2D semiconductors. . . . 89

5.1 Capacitance-voltage (CV) measurements of a planar SiON pMOS device at $T = 100°C$ and $f = 100\,\mathrm{kHz}$. . . . 99
5.2 Full state diagram of the amphoteric hydroxyl E′ defect in SiO_2 . . . 101
5.3 Capture and emission times for a switching trap. . . . 103
5.4 DFT calculations of an atomistic Si/SiO_2 interface. . . . 104
5.5 Classical (high temperature) limit of the NMP model. . . . 106
5.6 Correlation between the parameters E_R and R. . . . 107
5.7 Different tunneling mechanisms typically appearing in a MOS gate stack . 109
5.8 Monte-Carlo sampling for a TAT simulation . . . 110
5.9 Classical and quantum mechanical lineshape functions for different displacements ΔQ . . . 112
5.10 WKB approximation for quantum mechanical transition rates. . . . 113

6.1 Schematic illustration of an extended measure-stress-measure sequence. . . 117
6.2 RTN analysis on a small-area nMOS device. . . . 118
6.3 Typical results of a TDDS experiment. . . . 119
6.4 Decomposition of an eMSM sequence into single defect contributions using ESiD . . . 120

6.5 Extracted τ_c and E_R distributions for electron traps in SiO_2 using a Gaussian defect band . . . 122
6.6 Comparison between the non-self-consistent and self-consistent defect response of the extracted hole trap band in SiON devices . . . 123
6.7 Impact of the regularization parameter γ on the fitting error. . . . 124
6.8 Resulting hole trap distribution extracted with ESiD for varying interlayer thicknesses and capture cross sections. . . . 126
6.9 Samples drawn from the extracted defect distributions within the band diagram of the device. . . . 127
6.10 Experimental and simulated eMSM sequences derived from the extracted defect bands for varying gate bias and temperature. . . . 128
6.11 Extracted defect bands and a comparison to theoretical defect candidates for hole traps in SiO_2. . . . 129
6.12 Heat map of the (E_T, E_R) distribution for the extracted IL trap bands using the ESiD method together with parameters from single defects obtained with TDDS. . . . 130
6.13 Experimental time constants obtained from RTN signals at cryogenic temperatures together with fits of the classical and full quantum mechanical model. . . . 131
6.14 Defect parameters extracted from the RTN signals plottet on an (E_T, E_R) map. . . . 132
6.15 TAT simulations compared to experimental leakage currents of a SiC MOSCAP 133
6.16 Trap parameters extracted from the TAT model compared to values for polarons predicted by DFT . . . 135

List of Tables

3.1 Transition parameters for electron trapping at the $Zn_{Ga}O_{P}$ center in GaP . 49

4.1 Transition parameters for hole trapping at an isolated Si–H bond in α-quartz 67
4.2 Comparison of CaF_2 bulk properties calculated by DFT against experimental data. 86
4.3 Theoretical defect transitions and parameters for the Si_i and Si_{Ca} defect in CaF_2 88

6.1 Parameters used to define the different search regions in the oxide and the interlayer. 126
6.2 Trap parameters extracted from RTN signals at cryogenic temperatures . . 132
6.3 Trap parameters extracted from the TAT model compared to values for polarons predicted by DFT 134

List of Abbreviations

ADMM	auxiliary density matrix method
BOA	Born-Oppenheimer approximation
BTI	bias temperature instability
CB	conduction band
CBM	conduction band minimum
CC	configuration coordinate
CCS	capture cross section
CI-NEB	climbing-image nudged elastic band
CMOS	complementary MOS
CP	crossing point
CSA	charge sheet approximation
CTL	charge transition level
CV	capacitance-voltage
DB	dangling bond
DFT	density functional theory
DOS	density of states
DT	direct tunneling
DTCO	design technology co-optimization
ECM	embedded cluster method
eMSM	extended MSM

EOM equations of motion

EPR electron paramagnetic resonance

ERI electron-repulsion integral

ESiD effective single-defect decomposition

ESR electron spin resonance

EXX exact exchange

FC-LSF Franck-Condon lineshape function

FET field-effect transistor

FF force-field

FGR Fermi's golden rule

FN Fowler-Nordheim

FNV Freysoldt-Neugebauer-van de Walle

FQM full quantum mechanical

GGA generalized-gradient approximation

GPR Gaussian process regression

GTH Goedecker-Teter-Hutter

GTO Gaussian-type orbital

HB hydrogen bridge

HE hydroxyl-E'

HF Hartree-Fock

HK Hohenberg-Kohn

HO harmonic oscillator

HOMO highest occupied molecular orbital

HSE Heyd-Scuseria-Ernzerhof

HT-LSF Herzberg-Teller lineshape function

IPR inverse participation ratio

KS Kohn-Sham

LDA local density approximation

LSF lineshape function

LUMO lowest unoccupied molecular orbital

MD molecular dynamics

MECP minimum energy crossing point

MEP minimum energy path

MLFF machine-learning force-field

MOS metal-oxide-semiconductor

MOSCAP MOS capacitor

MOSFET MOS field-effect-transistor

MP Makov-Payne

MSM measure stress measure

NBTI negative bias temperature instability

NEB nudged elastic band

NMP non-radiative multi-phonon

NNLS non-negative linear least-squares

NQE nuclear quantum effects

OV oxygen vacancy

PBE Perdew-Burke-Ernzerhof

PEC potential energy curve

PES potential energy surface

positive BTI positive bias temperature instability

PP pseudo-potential

RTN random telegraph noise

SCA static coupling approximation

SCAN strongly constrained and appropriately normed

SCF self-consistent field

SIE self-interaction error

SiON silicon oxynitride

SRH Shockley–Read–Hall

STE self-trapped electron

STH self-trapped hole

TAT trap-assisted tunneling

TCAD technology computer-aided design

TDDS time-dependent defect spectroscopy

TS transition state

TST transition state theory

VB valence band

VBM conduction band maximum

VDOS vibrational density of states

vdW van der Waals

WKB Wentzel–Kramers–Brillouin

XC exchange-correlation

1

Introduction

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

It is fair to say that the invention of transistors, particularly metal-oxide-semiconductor field-effect transistors (MOSFETs), revolutionized the world. The ability to build integrated circuits containing billions of transistors — such as the Nvidia Blackwell GPU architecture introduced in 2024, which features 208 billion transistors on a single chip — has led to an extraordinary increase in computational power over the past 60 years. However, this progress has necessitated continuous downscaling of transistor size, pushing towards the atomic scale. This ongoing miniaturization presents scientists and engineers with increasingly challenging, yet fascinating, problems to solve. At such scales, even individual point defects can render a transistor nonfunctional, making precise process control crucial to ensure the reliability of nanoelectronic devices. One significant issue in these devices is charge trapping at atomic defects, which can alter device characteristics.

This thesis presents a multiscale modeling approach that incorporates both first-principles calculations on the atomic scale and device-level simulations, with the aim to provide a deeper understanding of the mechanisms involved and to assist device manufacturers in mitigating the impact of inevitable defects on device performance.

This introductory chapter begins with a brief history of the MOSFET as the central component in nanoelectronics, followed by a discussion of the most relevant reliability issues faced by device manufacturers and current efforts to understand and model them. The chapter concludes with an outline of the thesis structure.

1.1 MOS Devices

While early concepts of a MOS device date back to 1925, when Julius Edgar Lilienfeld proposed the idea, followed by Oskar Heil's insulated-gate FET patent in 1935, the practical realization of these concepts was achieved only later. In the early 1960s, Kahng and Atalla proposed and fabricated the first MOSFET using silicon (Si) with a thermally grown silicon dioxide (SiO_2) layer as a gate insulator [1]. This invention became the fundamental building block of digital circuits due to its scalability, low power consumption, and relative ease of manufacturing compared to bipolar junction transistors. However, even at this early stage, measures had to be taken to reduce the high concentration of defects at the Si/SiO_2 interface to produce a functional device [2].

Devices based on the Si/SiO_2 material system were used almost exclusively for digital logic circuits until the early 1990s. By then, manufacturers began introducing nitrogen into the gate oxide to form silicon oxynitride (SiO_xN_y), or SiON, to suppress gate-leakage currents [3] caused by continued downscaling and the corresponding reduction in oxide layer thickness. However, by the 2000s, further downscaling led to intolerable gate-leakage currents. To address this issue, high-κ materials such as hafnium dioxide (HfO_2) and zirconium dioxide (ZrO_2) were introduced as gate oxides. These materials have significantly higher permittivities compared to SiO_2, allowing for thicker oxides without compromising electrostatic control over the device channel. However, the introduction of high-κ materials resulted in a higher threshold voltage, which had to be compensated by replacing the traditional poly-silicon gate with a metal gate to adjust the gate work function [4]. This is the so-called high-κ/metal gate (HKMG) technology.

The next step in the evolution of nanoelectronic devices was the development of the FinFET, in which the channel is raised like a fin, allowing the gate metal to encompass the channel material from multiple sides [5]. Current approaches further develop this concept, resulting in gate-all-around (GAA) and nanosheet transistors [6].

All the approaches discussed thus far have relied on transistors utilizing silicon as the channel material. However, for power electronics, wide-bandgap semiconductors such as silicon carbide (SiC) and gallium nitride (GaN) offer significant advantages over silicon [7]. Additionally, to address the scaling challenges of transistors, two-dimensional (2D) materials such as molybdenum disulfide (MoS_2) are being investigated, which introduces entirely new challenges [8, 9].

1.2 Reliability Issues

This section is based on work published in [DWJ1].

Reliability challenges in MOSFETs primarily stem from the degradation of materials and interfaces, driven by electrical, thermal, and environmental stresses, such as irradiation, during operation. A detailed understanding of these degradation mechanisms is essential for designing robust devices that meet the stringent requirements of modern technology. Although other components of the device, such as interconnects, can also fail, this thesis focuses primarily on reliability issues caused by point defects in the gate insulator or the semiconductor/insulator interface. These defects can trap or emit charges, thereby altering the device electrostatics, or induce adverse reactions such as bond breakage. The main reliability issues resulting from such defects will be discussed in the following.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

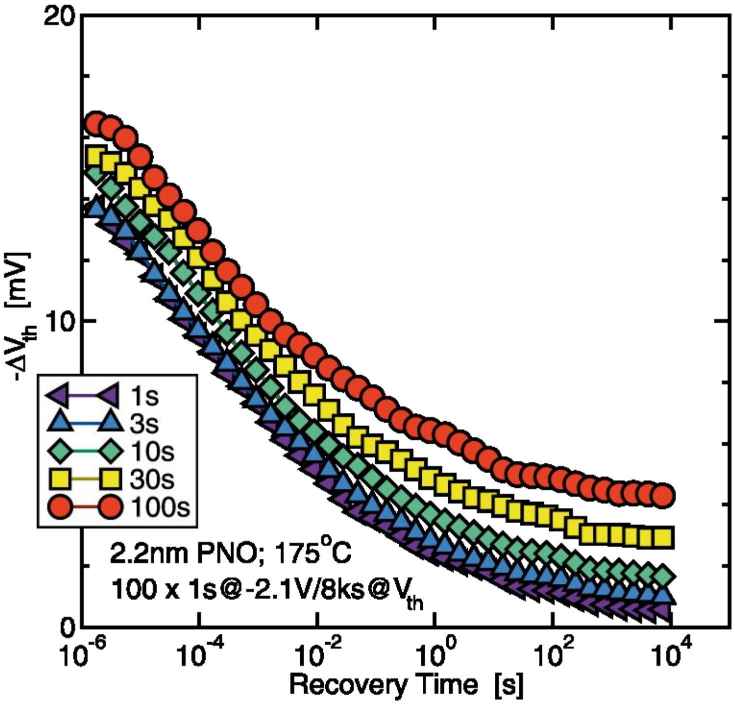


Figure 1.1. NBTI degradation observed in a pMOS device with SiON dielectric for stress times $t_s = 1 \ldots 100\,\mathrm{s}$. The threshold voltage shift $\Delta V_{\mathrm{th}}(t)$ partially recovers over time after the stress conditions are removed. (taken from [13])

1.2.1 Bias Temperature Instability

Bias Temperature Instability (BTI) refers to a shift in the threshold voltage, ΔV_{th}, following combined electrical and thermal stress applied to a device. This phenomenon was first reported in 1966 [10], but it only became a major reliability concern with the increase in electric fields within the gate oxide and the introduction of nitrided oxides as downscaling continued. Negative BTI (NBTI), which refers to a ΔV_{th} shift during negative gate biases, has been particularly problematic in devices using SiO_2 as the gate dielectric [11]. Since pMOS devices typically operate at negative gate voltages (V_{G}), they are primarily affected by NBTI. However, with the introduction of high-κ gate dielectrics, positive BTI (PBTI) in nMOS devices has also emerged as a significant issue [12].

The ΔV_{th} degradation is most pronounced immediately after stress. Following the removal of stress, the device partially recovers, with a permanent degradation persisting even after long recovery times [13]. It is now well established that the recoverable BTI component is due to charge trapping at defects within the gate oxide [14]. However, the exact physical mechanism underlying the permanent component remains a topic of debate. Current evidence suggests that the permanent component may be caused by the release of hydrogen from the gate side, leading to the creation of new defects within the oxide [15].

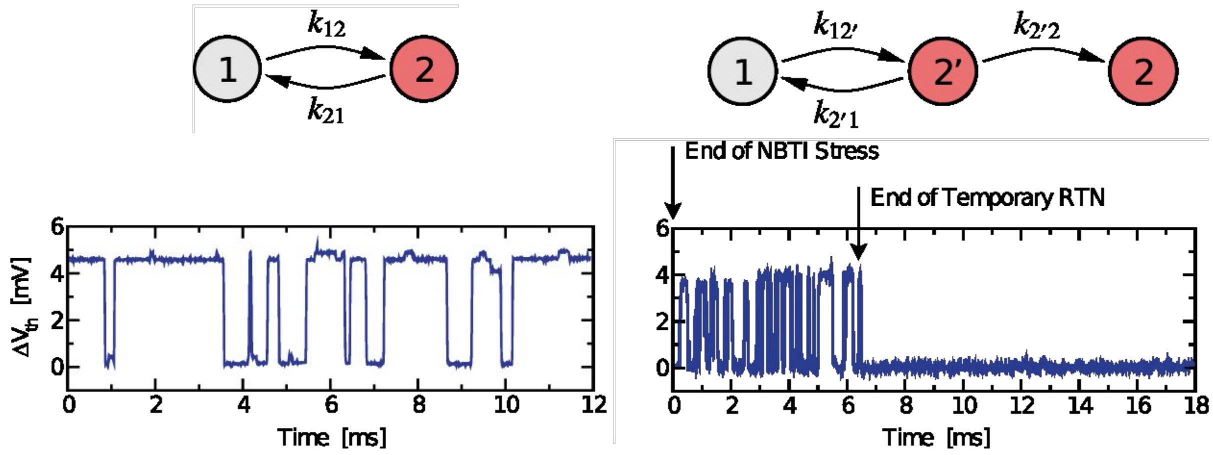


Figure 1.2. RTN signals measured on a small-area pMOS. **Left:** A continous RTN signal, which is produced by an oxide defect with two different states at a constant gate voltage V_G. **Right:** After BTI stress, often a so-called temporary RTN signal can be observed. In this case, the noise signal disappears after some time, indicating an additional electrically inactive defect state. (adapted from [DWJ1, 17])

1.2.2 Random Telegraph Noise

In contrast, charge trapping at the same oxide defects [16] in small-area devices leads to discrete steps in the threshold voltage, as illustrated in Fig. 1.2 (left). In such cases, a defect continuously captures and emits a charge, switching between the two charge states 1 and 2. This stationary random telegraph noise (RTN) signal is observed under a constant gate bias.

However, fundamentally different RTN signals, as depicted in Fig. 1.2 (right), are frequently observed following BTI stress. In this scenario, a temporary RTN signal appears after stress and subsequently vanishes after some time. A defect producing such an RTN signal must have at least three different states. Following stress, the defect switches between charge states 1 and $2'$, producing the stationary RTN signal, until an internal relaxation $2' \Rightarrow 2$ leads the defect into another state, 2, where no further charge exchange occurs. The frequent occurrence of such signals in measurements led to the development of the widely successful 4-state nonradiative multiphonon (NMP) model for describing RTN and BTI [17].

1.2.3 Trap-Assisted Tunneling

While BTI and RTN are caused by charges getting trapped in oxide defects, the defects can also enable charges to transition from the channel to the gate, causing a gate-leakage current, which is contrary to ordinary tunneling strongly temperature-dependent [18]. This so-called trap-assisted tunneling (TAT) emerges from the same NMP charge transfer mechanism as RTN and BTI [19], however, in this case, the charge does not stay at the defect site, but either hops to another defect (multi-TAT) or is released into the gate contact (single-TAT).

1.2.4 Hot-Carrier Degradation

While BTI is typically measured at a small drain-source voltage, V_{DS}, additional degradation is observed for larger V_{DS}. At high drain-source voltages, the carriers in the device channel are significantly accelerated. These hot carriers can interact with Si−H bonds at the Si/SiO$_2$ interface, causing bond breakage due to vibrational heating [DWJ2]. The resulting silicon dangling bonds can act as both electron and hole traps, making them amphoteric [20]. Additionally, their close proximity to the channel can reduce carrier mobility through enhanced scattering and degrade the sub-threshold swing. Hot-carrier degradation (HCD) particularly affects short-channel MOSFETs due to the larger electric fields between the drain and source contacts [21].

1.3 4-State Model

This section is based on work published in [DWJ1].

To account for the complex features of BTI and RTN, such as the aforementioned temporary RTN signals, the existence of switching traps [22], and the decorrelation between capture and emission times, a 4-state model, as depicted in Fig. 1.3 (left), was developed [17]. In this model, the defect can occupy two neutral states (1, $1'$) and two charged states (2, $2'$). The unprimed states 1 and 2 represent stable states, while $1'$ and $2'$ are metastable states. As illustrated in Fig. 1.3 (right), the transition rates between these states are governed by the movement along the defect's potential energy curves (PECs). The charge transfer reactions $1 \Leftrightarrow 2'$ and $1' \Leftrightarrow 2$ are governed by NMP theory, where, in the classical limit, the transition occurs at the crossing point of both PECs. These transitions are electrically detectable since the defect's charge state changes, affecting the device electrostatics. In contrast, the thermal transitions $1 \Leftrightarrow 1'$ and $2 \Leftrightarrow 2'$ do not change the defect's net charge, with their rates being determined by the classical barrier over the transition state.

Using atomistic modeling with density functional theory (DFT), it was confirmed that defects in SiO$_2$ can indeed exhibit the four states required by this model [23]. Here, the thermal transitions are realized by so-called puckering transitions [24]. Moreover, the theoretically predicted defect parameters for hydrogen-related defects in SiO$_2$ align well with experimental data [25], providing strong evidence for the validity of the 4-state model. However, this model is complex and the identification of suitable defect candidates has heavily relied on intricate single-defect characterizations. Therefore, recent efforts have focused on modeling BTI using a simplified 2-state model [26], as shown in Fig. 1.2 (left). Part of this thesis aims to demonstrate that defect identification can also be achieved with this simplified model in conjunction with more straightforward BTI measurements on large-area devices.

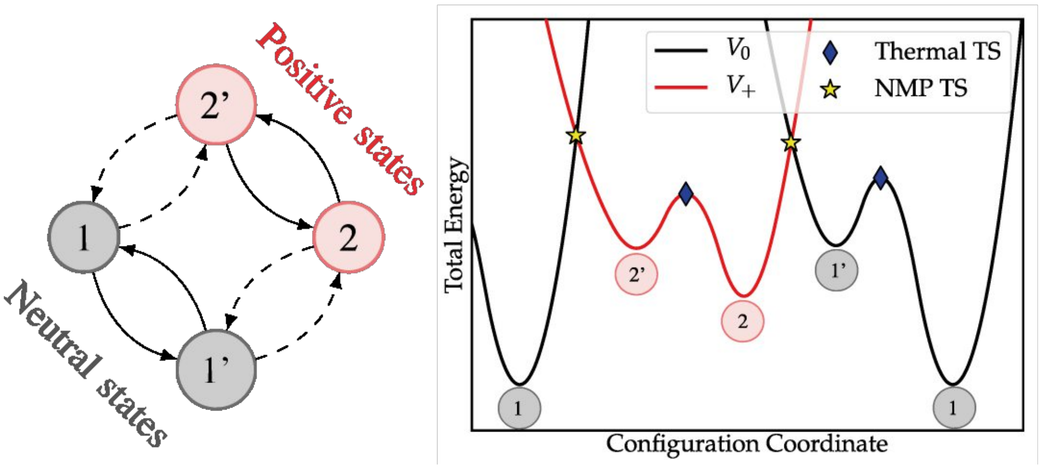


Figure 1.3. 4-State model for oxide defects. **Left:** The defect is assumed to exist in a stable and metastable configuration both in the neutral and charged state. Here the metastable states are indicated with a prime. Thermal transitions are represented by solid lines between states with the same charge, whereas NMP charge transitions are drawn as dashed lines between states with different charges. **Right:** Schematic potential energy curves (PECs) between the different states of the 4-state model. In this picture the states are given by the minima of the energy profiles. The transition rate for thermal transitions is determined by the point with highest energy along the reaction path. In the classical limit of NMP theory charge transitions occur at the intersection points of the surfaces. (taken from [DWJ1])

1.4 Thesis Outline

Chapter 2 introduces the fundamental concepts of atomistic simulations and the computational methods used in this work. It begins with the basics of density functional theory (DFT), the work horse of modern computational material science, and discusses the DFT software package CP2K and its advantages. Additionally, the chapter covers the movement of nuclei on the potential energy surface (PES) and the concept of phonons, relevant for the charge transfer models discussed later.

Chapter 3 reviews the theory of nonradiative multiphonon (NMP) transitions, which underpins the modeling of charge capture and emission at defect sites in the gate oxide. It first presents the core principles of NMP transitions and their quantitative evaluation using DFT. A novel formalism for incorporating additional phonon modes is introduced, enabling efficient multimode treatment while retaining a 1-dimensional representation. The chapter concludes with the classical limit of NMP theory, used in device-level simulations due to its computational efficiency.

Chapter 4 applies the computational methods introduced earlier to study defects in gate dielectrics, focusing on silicon dioxide (SiO_2). It explores the dissociation of Si−H bonds due to hole trapping, the modeling of amorphous SiO_2, and common point defects, including self-trapped charges. The chapter also investigates impurity defects in calcium fluoride (CaF_2) as a potential insulator for future 2D nanoelectronics.

Chapter 5 integrates both classical and quantum mechanical charge trapping models into the *Comphy* device simulator, bridging atomistic processes with their macroscopic impact on devices. It provides an overview of the framework and new features developed within the scope of this thesis, including a novel trap-assisted tunneling model and a WKB-based approximation for studying charge trapping at cryogenic temperatures.

Chapter 6 integrates atomistic and device modeling with experimental defect characterization to link microscopic processes to device degradation, enabling the identification of defects causing reliability issues like BTI, RTN, and gate-leakage currents. It introduces the Effective Single Defect Decomposition (ESiD) algorithm for automatic defect parameter extraction and demonstrates its effectiveness compared to more complex techniques. The chapter concludes by showcasing the multiscale modeling approach to identify defects responsible for detrimental effects in different applications.

Chapter 7 briefly summarizes the most significant results of this thesis and provides a few ideas for future research.

Part I

First Principles Modeling

2

Computational Methods

This chapter is partially based on works previously published in [DWJ1, DWT1].

This chapter introduces the fundamental concepts of atomistic simulations and the computational methods employed in this work. First, the basics of density functional theory (DFT) are presented, as it is the foundation for most electronic structure calculations today. DFT is particularly indispensable for studying point defects due to the large simulation cells required. Following this, the DFT software package CP2K, which is used throughout this work, will be discussed, highlighting its computational advantages.

DFT not only describes the electronic structure but also provides the total energy of a system along with the atomic forces. The movement of nuclei on the resulting potential energy surface (PES) is also covered in this chapter, given its importance for the charge transfer model introduced in Ch. 3, and for creating atomistic structures of amorphous oxides using molecular dynamics, as discussed in Ch. 4.

2.1 Density Functional Theory

On the atomic scale, the behavior of nuclei and electrons is universally governed by the Many-Body-Schrödinger equation of the studied system. However, in order to numerically solve this equation for extended systems like semiconductors, gate dielectrics or point defects within these materials, a number of approximations have to be applied. By far the most widely used approach for quantum mechanical simulations of solids is density functional theory (DFT) due to its favorable scaling with system size compared to other quantum chemistry methods. Here, the necessary theory of this framework will be developed as far as required for subsequent chapters.

2.1.1 Born-Oppenheimer Approximation

In principle, all properties of an arbitrary collection of M nuclei and N electrons are determined by the system's quantum mechanical many-body wavefunction $|\Psi(\mathbf{r}, \mathbf{R})\rangle$, where

$$\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_M\} \in \mathbb{R}^{3M} \quad \text{and} \quad \mathbf{r} = \{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N\} \in \mathbb{R}^{3N} \tag{2.1}$$

are the collective positions of all nuclei and electrons in the system respectively. This wavefunction is the solution of the system's many-body Schrödinger equation

$$\hat{H}\,|\Psi(\mathbf{r}, \mathbf{R})\rangle = E\,|\Psi(\mathbf{r}, \mathbf{R})\rangle \tag{2.2}$$

with E being the total energy of the system and $\hat{H}$ being the Hamiltonian, which consists of multiple different energy contributions:

$$\hat{H} = \hat{T}_\mathrm{e} + \hat{T}_\mathrm{n} + \hat{V}_\mathrm{n} + \hat{V}_\mathrm{en} + \hat{U} \tag{2.3}$$

In Hartree atomic units [27] these contributions are given as follows

(i) The kinetic energy operators $\hat{T}_\mathrm{n}$ and $\hat{T}_\mathrm{e}$ for the nuclei and electrons respectively are given by

$$\hat{T}_\mathrm{n} = \frac{1}{2}\sum_{i=1}^{M} \nabla^2_{\mathbf{R}_i} \quad \text{and} \quad \hat{T}_\mathrm{e} = \frac{1}{2}\sum_{i=1}^{N} \nabla^2_{\mathbf{r}_i}\,. \tag{2.4a}$$

(ii) The potential energy of the nuclei due to their mutual Coulomb repulsion can be written as

$$\hat{V}_\mathrm{n} = \frac{1}{2}\sum_{i \neq j}^{M} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|}\,, \tag{2.4b}$$

with Z_i, Z_j being the nuclear charge or atomic numbers of the nuclei. The factor $1/2$ appears here to avoid double counting when summing over all pairs (i, j) of distinct nuclei.

(iii) Likewise, the potential energy of the electrons due to their mutual Coulomb repulsion is given by

$$\hat{U} = \frac{1}{2}\sum_{i \neq j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}\,. \tag{2.4c}$$

(iv) Finally, the potential energy due to the mutual attraction of nuclei and electrons is simply

$$\hat{V}_\mathrm{en} = -\sum_{i}^{M}\sum_{j}^{N} \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}_j|}\,. \tag{2.4d}$$

By using (2.2), both electrons and nuclei are described fully quantum mechanically and their dynamics are inherently coupled to each other. However, since even the lightest nuclei are much heavier than electrons, the electronic subsystem can almost instantly adjust to a movement of the nuclei. Introducing the dimensionless parameter

$$\xi = \left(\frac{m_\mathrm{e}}{M}\right)^{\frac{1}{4}} \sim 0.1 \tag{2.5}$$

for the electron mass m_e and a typical nucleus mass M_n, it can be shown by perturbation theory, that the energy contributions from the nuclei in the form of vibrations are a factor of $\xi^2 \sim 0.01$ smaller than the electronic energy of the system [28, 29]. For many problems in computational material science it is thus practical and justifiable to separate the dynamics of the nuclei and the electrons, i.e. the vibronic wavefunction is assumed to be separable:

$$|\Psi(\mathbf{r},\mathbf{R})\rangle = |\Phi_i(\mathbf{r};\mathbf{R})\rangle \times |\chi_{im}(\mathbf{R})\rangle \tag{2.6}$$

In this so-called Born-Oppenheimer approximation (BOA) [28, 30], the Schrödinger equation for the electronic subsystem

$$\hat{H}_\mathrm{e}\,|\Phi_i(\mathbf{r};\mathbf{R})\rangle = \left(\hat{T}_\mathrm{e} + \hat{V}_n + \hat{V}_\mathrm{en} + \hat{U}\right)|\Phi_i(\mathbf{r};\mathbf{R})\rangle = V_i(\mathbf{R})\,|\Phi_i(\mathbf{r};\mathbf{R})\rangle \tag{2.7}$$

is first solved for a fixed nuclei configuration $\mathbf{R}$ leading to the adiabatic electronic wavefunction $|\Phi_i(\mathbf{r};\mathbf{R})\rangle$ and its corresponding eigenenergy $V_i(\mathbf{R})$. Here, the electronic state i can denote either the ground state or an excited state of the system. Since $\mathbf{R}$ only enters the electronic Schrödinger equation (2.7) as a parameter, every electronic state i is associated with its own potential energy surface (PES) $V_i(\mathbf{R})$ [29]. This PES acts as an effective potential in which the nuclei are moving and is of profound importance for reaction kinetics of the system as will be further explained in Sec. 2.3. The vibrational wavefunctions $|\chi_{im}\rangle$ for the nuclei are then determined by the Schrödinger equation

$$\hat{H}_\mathrm{vib}\,|\chi_{im}(\mathbf{R})\rangle = \left(\hat{T}_\mathrm{n} + V_i(\mathbf{R})\right)|\chi_{im}(\mathbf{R})\rangle = E_{im}\,|\chi_{im}(\mathbf{R})\rangle\ . \tag{2.8}$$

These vibrational wavefunctions are schematically shown in Fig. 2.1 for two different electronic states i and j of a diatomic molecule. Since a diatomic molecule has only one independent degree of freedom, the otherwise $3M$-dimensional PES reduces to a 1-dimensional potential energy curve (PEC). This simplification to a single degree of freedom is also often applied to describe complex chemical reactions or charge transfer within a system and will be used throughout this work.

While the BOA provides the basis for almost all quantum chemistry methods, it should be noted that there are cases in which this approximation breaks down. The most common example is near crossing points or conical intersections of two different PESs, where the electronic and vibrational movement cannot be separated. However, in such cases the interaction between both electronic states can be reintroduced approximately by adding nonadiabatic coupling terms [31] to the vibrational Hamiltonian (2.8).

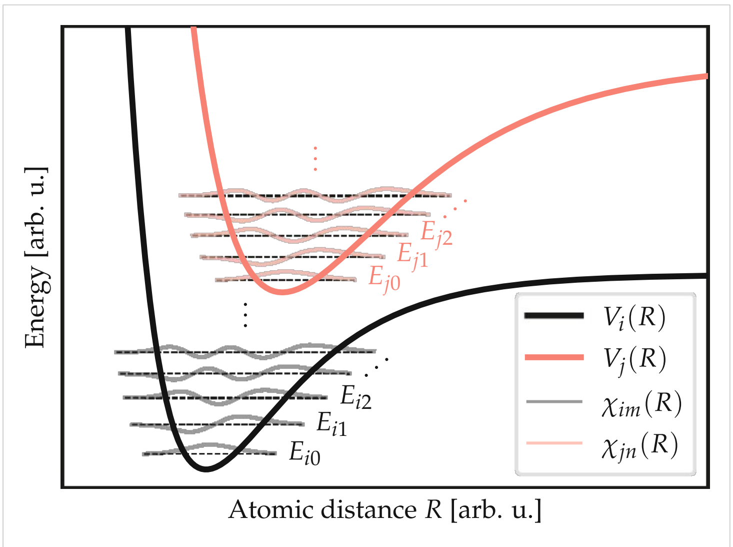


Figure 2.1. Schematic PESs of a diatomic molecule in two different electronic states i and j. In this case the general nuclei coordinates $\mathbf{R}$ reduce to the scalar distance R between the two atoms when accounting for translational and rotational invariance. Each electronic state is associated with its own PES and vibrational eigenstates with energies E_{im} and E_{jn}. (adapted from [DWT1])

2.1.2 Hohenberg-Kohn Theorems

While the BOA provides a significant simplification of the many-body Schrödinger equation (2.2) for a system of interacting nuclei and electrons, solving (2.7) for the electronic subsystem alone is extremely challenging. The main problem here is that for a system with N electrons, the electronic wavefunction $|\Phi\rangle$ depends on all electronic degrees of freedom and is thus a highly complex $3N$-dimensional object. For example, if one would naively try to solve (2.7) for all electrons in a single silicon atom ($N = 14$) by discretizing the spatial domain, even storing the wavefunction evaluated at only 10 points along each electronic coordinate, would require 10^{42} data points. This conundrum is typically referred to as the "curse of dimensionality" and makes directly solving (2.7) impossible for all but the smallest systems. While there are wavefunction based quantum-chemistry methods [32] like Coupled Cluster (CC) or Configuration Interaction (CI) which can solve (2.7) approximately for larger systems of up to 100 electrons, searching for the many-body electronic wavefunction $|\Phi\rangle$ in extended systems like defects in a gate dielectric clearly needs a different approach.

For many problems, including the scope of this work, the electronic ground state is sufficient for describing the system. For the ground state it was shown by Hohenberg and Kohn [33] that the wavefunction and hence every other property of the ground state is *uniquely* determined by the ground state electron density $n_0(\mathbf{r})$ alone. Since the ground state density is a 3-dimensional object regardless of the number of electrons in the system, searching for the density is vastly simpler than searching for the wavefunction itself. This ingenious insight is at the heart of density functional theory (DFT), which revolutionized computational material science and allows to simulate large systems containing 1000s of atoms on modern supercomputer architectures [34]. The special properties of the ground state allowing for this simplification are summarized by the Hohenberg-Kohn (HK) theorems, which can be stated as [DWJ1, 35]

(i) The ground state wavefunction is a unique functional $\Phi_0 = \Phi_0[n_0]$ of the ground state electron density. As a consequence, every ground state observable $\hat{O}$ can be defined as a density functional as well:

$$O[n] = \langle \Phi[n] | \hat{O} | \Phi_0[n] \rangle \tag{2.9}$$

(ii) The energy functional $E[n]$ obtains its global minimum for the true ground state electron density n_0.

While it is trivial that the electron density is uniquely determined by the many-body wavefunction through

$$n(\mathbf{r}) = N \int \mathrm{d}^3\mathbf{r}_2 \cdots \int \mathrm{d}^3\mathbf{r}_N |\Phi(\mathbf{r}, \mathbf{r}_2, \ldots \mathbf{r}_N)|^2 \,, \tag{2.10}$$

the first HK theorem shows that the reverse also holds for the ground state. Hence, there exists a one-on-one mapping between the ground state wavefunction $|\Phi_0\rangle$ and the ground state electron density $n_0(\mathbf{r})$. Using the second HK theorem, finding $n_0(\mathbf{r})$ for a system with N electrons becomes a constrained variational problem for minimizing the energy functional $E[n]$ in the form

$$\frac{\delta E}{\delta n} = 0 \quad \text{subject to} \quad \int n(\mathbf{r}) \mathrm{d}^3\mathbf{r} = N \,, \tag{2.11}$$

where $\delta E / \delta n$ denotes the functional derivative of E with respect to n.

2.1.3 Kohn-Sham Theory

The HK theorems prove the existence of an energy functional

$$E[n] = T[n] + V_{\text{ext}}[n] + U[n], \tag{2.12}$$

where $T[n]$ denotes the electronic kinetic energy functional, $U[n]$ accounts for the interaction between electrons and $V_{\text{ext}}[n]$ incorporates the energy from all other potentials $v_{\text{ext}}(\mathbf{r})$ acting on the electrons, e.g. the Coulomb potential of the nuclei or externally applied electric fields. While minimizing the energy functional according to (2.11) is mathematically equivalent to solving the many-body Schrödinger equation (2.7), only the functional for the external potential V_{ext} is known and can be classically expressed as

$$V_{\text{ext}}[n] = \int v_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) \mathrm{d}^3\mathbf{r} \,. \tag{2.13}$$

On the other hand, the functionals $T[n]$ and $U[n]$ are unknown and would be so intricately complex that they would not be of any practical use. To overcome this, Kohn and Sham proposed to use a fictitious system of non-interacting electrons to construct the unknown energy functionals [36]. In this case, the electronic wavefunction would be separable and hence can be represented as a product of single-particle wavefunctions $\phi_i(\mathbf{r})$:

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \phi_1(\mathbf{r}_1) \times \phi_2(\mathbf{r}_2) \times \cdots \times \phi_N(\mathbf{r}_N) \,. \tag{2.14}$$

However, while such an ansatz was used early on in quantum chemistry and is known as the Hartree method [37], this wavefunction is fundamentally wrong, since it is not antisymmetric with respect to exchanging two particles, violating the Pauli exclusion principle and thus cannot correctly describe a system of Fermions like electrons. This can be fixed by performing a complete anti-symmetrization of (2.14), leading to a so-called Slater determinant [38, 39]:

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \ldots & \phi_N(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2) & \ldots & \phi_N(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{r}_N) & \phi_2(\mathbf{r}_N) & \ldots & \phi_N(\mathbf{r}_N) \end{vmatrix} \tag{2.15}$$

For such a system of non-interaction electrons, the electron density is simply given by

$$n(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2 \,, \tag{2.16}$$

while the kinetic energy functional can be expressed as

$$T_{\mathrm{s}}[n] = -\frac{1}{2} \sum_{i=1}^{N} \langle \phi_i | \nabla^2 | \phi_i \rangle \;. \tag{2.17}$$

Note that $T_{\mathrm{s}}[n]$ does *not* represent the complete functional $T[n]$ for a system of interaction electrons, however, $T_{\mathrm{s}}[n]$ provides a major contribution to $T[n]$. Similarly, the interaction energy $U[n]$ can be approximated by the classical Hartree energy

$$U_{\mathrm{H}}[n] = \frac{1}{2} \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \mathrm{d}^3\mathbf{r} \mathrm{d}^3\mathbf{r}' \,, \tag{2.18}$$

which accounts for the purely electrostatic Coulomb repulsion. Kohn and Sham now expressed the true energy functional as

$$E[n] = T_{\mathrm{s}}[n] + V_{\mathrm{ext}}[n] + U_{\mathrm{H}}[n] + E_{\mathrm{XC}}[n] \,, \tag{2.19}$$

where all the remaining contributions from many-body quantum effects are collected in the so-called exchange-correlation (XC) functional E_{XC}, for which various decent approximations exist as discussed in the next section.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

Applying the Euler-Lagrange formalism to the variational problem (2.11) then leads to a system of single-particle Schrödinger equations, the Kohn-Sham (KS) equations:

$$\hat{H}_{\mathrm{KS}}\,\phi_i(\mathbf{r}) = \left(-\frac{1}{2}\nabla^2 + v_{\mathrm{eff}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r})\,. \tag{2.20}$$

Here, the electrons move independently in an effective potential given by

$$v_{\mathrm{eff}}(\mathbf{r}) = \frac{\delta}{\delta n}\left(U_{\mathrm{H}}[n] + V_{\mathrm{ext}}[n] + E_{\mathrm{XC}}[n]\right) = \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}^3\mathbf{r}' + v_{\mathrm{ext}}(\mathbf{r}) + \frac{\delta E_{\mathrm{XC}}}{\delta n}\,. \tag{2.21}$$

However, since $v_{\mathrm{eff}}(\mathbf{r})$ depends on the electron density, (2.20) becomes a system of nonlinearly coupled differential equations, which have to be solved iteratively in a self-consistent cycle. Hereby, an initial set of trial wavefunctions $\{\phi_i^0(\mathbf{r})\}$ is used to construct an initial electron density $n^0(\mathbf{r})$ via (2.16), from which an effective potential $v_{\mathrm{eff}}^0(\mathbf{r})$ follows. From this potential a new set of wavefunctions $\{\phi_i^1(\mathbf{r})\}$ can be derived by solving (2.20). This so-called self-consistent field (SCF) cycle is repeated until the wavefunctions are self-consistent with the electron density and hence do not change anymore from one iteration to the next. The converged wavefunctions are usually referred to as the auxiliary KS states with eigenenergies ϵ_i. The total energy of the system is then given by [36]

$$E_{\mathrm{tot}} = \sum_i^N \epsilon_i - U_{\mathrm{H}}[n] + E_{\mathrm{XC}}[n] - \int \frac{\delta E_{\mathrm{XC}}}{\delta n} n(\mathbf{r})\mathrm{d}^3\mathbf{r}\,. \tag{2.22}$$

It has to be noted that the KS states are only an auxiliary system of non-interacting electrons, which happens to reproduce the correct ground state electron density $n_0(\mathbf{r})$ (as well as the ionization energy of the system via Janak's theorem [40, 41]) but beyond that does not have a rigorous physical meaning. With that being said, it is a successful practice to interpret the KS states as good approximation for true single-electron states in a solid, i.e. to derive properties like the bandstructure [DWJ3] or energy levels of defects [DWJ4–DWJ6].

The main advantage of the KS approach is that it allows for a good approximation of $T[n]$ by separating out the kinetic energy $T_{\mathrm{s}}[n]$ of the fictitious non-interacting electrons. However, in order to obtain the KS wavefunctions from (2.20), an eigenvalue decomposition of the KS-Hamiltonian $\hat{H}_{\mathrm{KS}}$ has to be performed, which scales as $\mathcal{O}(N^3)$ and thus becomes the limiting step for large systems. While most DFT investigations, including the work presented in this thesis, rely on the auxiliary KS system, for the sake of completeness it should be mentioned that there are other approaches, collectively termed orbital-free DFT, which attempt to directly approximate the unknown energy functional $E[n]$. Although they allow the quantum mechanical calculation of millions of atom [42–44] due to their favorable quasi-linear $\mathcal{O}(N\log N)$ scaling, they usually are less accurate due to a poorer approximation of $T[n]$ [45].

2.1.4 Exchange-Correlation Functionals

In order to enable practical electronic structure calculations, suitable approximations to the unknown XC functional have been developed. Most of these XC functionals are constructed to reproduce some well-known properties of the exact functional. In the following a brief overview about different XC functionals is given, for a more in-depth discussion, the reader is referred to the literature [46, 47].

Already in their seminal work which laid the foundation for DFT [33], Hohenberg and Kohn proposed the first practically useful XC functional, the local density approximation (LDA). Here, it is assumed that the XC functional depends only on the local electron density, i.e. it has the form

$$E_{\mathrm{XC}}^{\mathrm{LDA}}[n] = \int \varepsilon_{\mathrm{XC}}^{\mathrm{LDA}}(n(\mathbf{r}))\, n(\mathbf{r}) \mathrm{d}^3\mathbf{r}\,. \tag{2.23}$$

The function $\varepsilon_{\mathrm{XC}}(n)$ is then determined by assuming a general electron gas behaves locally as an homogeneous electron gas (HEG) with the same density, for which the exchange part can be derived from the Thomas-Fermi model [48, 49] as [35]

$$\varepsilon_{\mathrm{X}}^{\mathrm{LDA}}(n) = -\frac{3}{4\pi}(3\pi^2 n)^{1/3}\,. \tag{2.24}$$

While no exact closed form for the remaining correlation term $\varepsilon_{\mathrm{C}}^{\mathrm{LDA}}(n)$ exists, it has been calculated numerically by quantum Monte Carlo simulations [50]. Despite its simplicity, the LDA is capable of qualitatively describing chemical trends and bonding correctly [40]. However, it tends to overestimate the binding energy [51] and is not applicable for rapidly varying electron densities, like localized charges in defects or polarons [52] for instance.

In order to mitigate some of the shortcomings of the LDA, the functional form can be extended by also including a dependency on the local density gradient:

$$E_{\mathrm{XC}}^{\mathrm{GGA}}[n] = \int \varepsilon_{\mathrm{XC}}^{\mathrm{GGA}}(n, \nabla n)\, n(\mathbf{r}) \mathrm{d}^3\mathbf{r} \tag{2.25}$$

Such semi-local or generalized-gradient approximation (GGA) functionals provide a substantial improvement over the LDA [53], however, there is no unique way to expand upon the LDA, leading to a plethora of different GGA functionals [54]. By far the most widely used parameter-free GGA functional for solid state systems was developed by Perdew, Burke and Ernzerhof (PBE) [55]. GGA functionals can be further extended by also including a dependency on second-order derivatives of the density, leading to so-called meta-GGA functionals. Probably the most widely adopted representative of this class is the rather recently proposed strongly constrained and appropriately normed (SCAN) functional [56].

While the development of (meta-)GGA functionals was undoubtedly a major improvement, these functionals still struggle to correctly predict certain material properties. For instance, also the SCAN functional can not reliably predict the properties of localized charges [57]. Fundamentally, the inability to describe localized charges is not a problem

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

of the used functional, but rather stems from the self-interaction error (SIE) inherent to KS-DFT [58]. This SIE becomes particularly large for charge transfer or dissociation reactions [59]. Another problem is that these functionals pathologically underestimate the electronic bandgap in solids [60, 61]. Both of these drawbacks can be mitigated by the use of hybrid functionals, in which the exchange part of a GGA functional is mixed with the exact exchange (EXX)

$$E_{\mathrm{X}}^{\mathrm{HF}}[n] = -\frac{1}{2}\sum_{i\neq j}^{N}\iint \phi_i^*(\mathbf{r})\phi_j^*(\mathbf{r}')\frac{1}{|\mathbf{r}-\mathbf{r}'|}\phi_i(\mathbf{r}')\phi_j(\mathbf{r})\mathrm{d}^3\mathbf{r}\mathrm{d}^3\mathbf{r}' \tag{2.26}$$

from Hartree-Fock theory [39]. For example, in the PBE0 [62] hybrid functional, the exchange functional is calculated as

$$E_{\mathrm{X}}^{\mathrm{PBE0}} = \alpha E_{\mathrm{X}}^{\mathrm{HF}} + (1-\alpha)E_{\mathrm{X}}^{\mathrm{PBE}} \tag{2.27}$$

with the mixing parameter α. Contrary to the other functionals discussed so far, $E_{\mathrm{X}}^{\mathrm{HF}}[n]$ is fundamentally nonlocal due to the 4-center electron-repulsion integrals (ERIs) required for its evaluation. Initially, $\alpha = 0.25$ was determined on theoretical grounds [63]. However, it was shown that the mixing parameter for a given material is inversely related to the dielectric constant of the material [64, 65]. Hence, the mixing parameter α is material dependent and is frequently modified in order to match the experimental bandgap. The most common form of hybrid functionals in use today are range-separated or screened, meaning the EXX is short-ranged and acts only on a characteristic length scale. Popular screened hybrid functionals include the Heyd-Scuseria-Ernzerhof (HSE) [66] and the PBE0_TC_LRC functionals [67]. The latter uses a truncated Coulomb (TC) operator, which allows to screen ERIs prior to computation and thus is highly efficient.

Despite their much higher computational costs, HSE and PBE0_TC_LRC are used throughout this work, since a correct description of localized charges and electronic bandgaps is of utmost importance when it comes to predicting the behavior of defects and their charge trapping dynamics [64, 68].

2.1.5 Pseudopotentials

While in principle all electrons of a system can be calculated by self-consistently solving the KS equations (2.20), typically only the valence electrons are of importance for chemical bonding. The core electrons on the other hand are barely effected by the chemical environment but are difficult to treat computationally, since they have very high kinetic energies, resulting in their wavefunctions rapidly oscillating in space. Since the main effect of the core electrons is to partially screen the charge of the nucleus, it is thus desirable to incorporate their precalculated wavefunctions in the core potential for each chemical element (frozen core). However, the valence wavefunctions still have to be orthogonal to the core wavefunctions and thus also rapidly oscillate around the nucleus. In order to solve this, one typically introduces *pseudo-wavefunctions* for describing the valence states.

Following the explanations given in [69], the real valence wavefunction $|\psi\rangle$ is expressed by a superposition of the pseudowavefunction $|\psi_{\mathrm{PS}}\rangle$ and the precomputed core-states $|\phi_n\rangle$ from an all-electron calculation:

$$|\psi\rangle = |\psi\rangle_{\mathrm{PS}} + \sum_n^{\mathrm{core}} a_n |\phi_n\rangle \tag{2.28}$$

Due to the required orthogonality, the coefficients a_n follow as

$$\langle\phi_m|\psi\rangle = 0 = \langle\phi_m|\psi_{\mathrm{PS}}\rangle + \sum_n^{\mathrm{core}} a_n \underbrace{\langle\phi_m|\phi_n\rangle}_{\delta_{mn}} \quad \Rightarrow \quad a_m = -\langle\phi_m|\psi_{\mathrm{PS}}\rangle \tag{2.29}$$

and the full wavefunction $|\psi\rangle$ can thus be constructed from $|\psi_{\mathrm{PS}}\rangle$ by projecting out the core states:

$$|\psi\rangle = |\psi\rangle_{\mathrm{PS}} - \sum_n^{\mathrm{core}} |\phi_n\rangle\langle\phi_n|\psi_{\mathrm{PS}}\rangle = \left(\mathbb{1} - \sum_n^{\mathrm{core}} |\phi_n\rangle\langle\phi_n|\right)|\psi_{\mathrm{PS}}\rangle \tag{2.30}$$

The Schrödinger equation $\hat{H}|\psi\rangle = E|\psi\rangle$ then yields

$$\hat{H}|\psi_{\mathrm{PS}}\rangle - \sum_n^{\mathrm{core}} E_n |\phi_n\rangle\langle\phi_n|\psi_{\mathrm{PS}}\rangle = E|\psi_{\mathrm{PS}}\rangle - E\sum_n^{\mathrm{core}} |\phi_n\rangle\langle\phi_n|\psi_{\mathrm{PS}}\rangle \tag{2.31}$$

or rather

$$\Bigg(\hat{H} + \underbrace{\sum_n^{\mathrm{core}}(E - E_n)|\phi_n\rangle\langle\phi_n|}_{\hat{V}_{\mathrm{NL}}}\Bigg)|\psi_{\mathrm{PS}}\rangle = E|\psi_{\mathrm{PS}}\rangle\ . \tag{2.32}$$

Hence, by working with the pseudowavefunction $|\psi_{\mathrm{PS}}\rangle$ an energy dependent nonlocal potential term $\hat{V}_{\mathrm{NL}}$, the so-called pseudo-potential (PP), is added to the Hamiltonian. Since the core levels E_n are much deeper than the valence levels E of interest, the energy dependence of the PP can be dropped by using $(E - E_n) \approx (E_{\mathrm{VB}} - E_n)$, where E_{VB} denotes the valence electron level of the isolated atom. Since the PP is nonlocal, it acts differently on valence states with different quantum numbers. However, when a particular valence state $|\psi\rangle$ is picked, the action of the PP can be visualized as shown in Fig. 2.2. For this particular state, the atom without its valence electrons is then described by the PP

$$V_{\mathrm{PS}}(r) = -\frac{Z_{\mathrm{eff}}}{r} + V_{\mathrm{NL}}(r) \tag{2.33}$$

with the effective core charge $Z_{\mathrm{eff}} = Z - N_{\mathrm{valence}}$. As can be seen, V_{PS} is much shallower than the corresponding Coulomb potential of the (screened) nucleus. However, beyond a certain cutoff radius r_{C} the resulting pseudowavefunction from this potential is identical to the real valence wavefunction and has the same eigenenergy E. Furthermore, the pseudowavefunction is nodeless in the core region, which makes it much easier to deal with numerically.

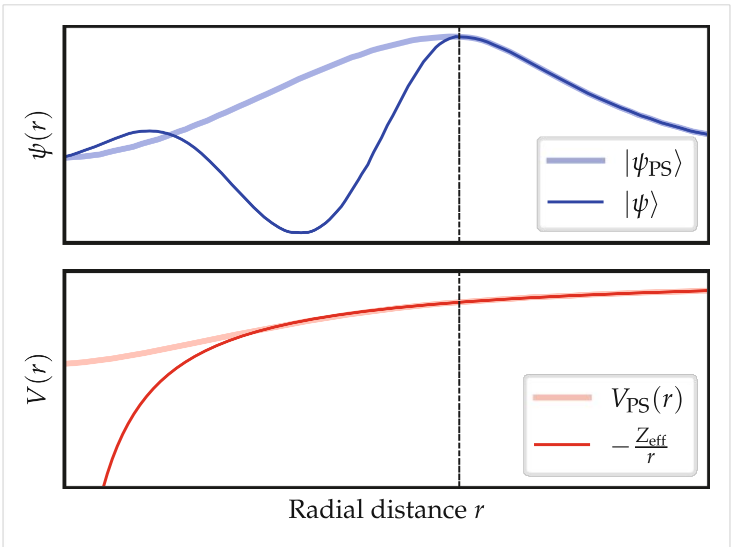


Figure 2.2. **Top:** The valence wavefunction $|\psi\rangle$ strongly fluctuates around the core as a result of the required orthogonality to the core states. To the contrary, the constructed pseudowavefunction $|\psi_{\text{PS}}\rangle$ is nodeless in the core region. **Bottom:** The corresponding pseudopotential V_{PS} is much shallower than the screened core potential $-Z_{\text{eff}}/r$, but produces the same wavefunction beyond a certain cutoff radius r_{C} (dashed line) with the same eigenenergy E. (adapted from [69])

Due to their vast numerical advantages, most electronic structure calculations rely on pseudopotentials to describe the atoms in the system [70]. While the explanation of PPs given above is based on one of the earliest and simplest PP approaches [71], the general concept is the same for more advanced PPs. Among the most popular PPs for electronic structure calculations are ultrasoft [72] and projector augmented-wave (PAW) [73, 74] PPs. In this work, the Goedecker-Teter-Hutter (GTH) [75] PPs are used, due to their efficient implementation in CP2K using Gaussian basis sets.

2.2 CP2K Software Package

The previous section gave an introduction to the basic theory of electronic structure calculations, in particular to DFT. Since the emergence of computational material science as its own discipline, many different codes have been developed to perform practical DFT calculations on real systems, see e.g. [76] for a list of widely used DFT codes. While codes like VASP [74, 77–79] or QUANTUMESPRESSO [80, 81] have been extensively used in the past to study defects in various materials [64, 68, 82, 83], they require a large amount of computational resources, due to their use of a planewave basis set, in particular for large defective supercells and when using hybrid functionals [84]. Since one of the aims in this work is to simulate defects in amorphous gate oxides with fluctuating chemical environments, a large number of defects has to be calculated in order to draw statistically meaningful conclusions. For this reason, the CP2K package [85] was used, since it offers a couple of key performance advantages, which will be discussed in the following.

2.2.1 Gaussian and Plane Wave Method

In order to solve the KS equations (2.20) numerically, the KS wavefunctions first have to be expanded into a *finite* set of basis functions $\{\varphi_k(\mathbf{r})\}$, i.e.

$$\psi_i(\mathbf{r}) = \sum_k C_{ik}\, \varphi_k(\mathbf{r})\,. \tag{2.34}$$

Like other quantum chemistry codes [86, 87], CP2K employs a Gaussian-type orbital (GTO) basis set [88], where every basis function $\varphi(\mathbf{r})$ mimics a particular atomic orbital belonging to an atom at position $\mathbf{R}_0$. Hereby, each basis function is constructed as a superposition

$$\varphi(\mathbf{r}) = \sum_n d_n g_n(\mathbf{r} - \mathbf{R}_0)\,, \tag{2.35}$$

where $g_n(\mathbf{r})$ denotes a primitive Gaussian function of the general form

$$g_n(\mathbf{r}) = r^l e^{-\zeta_n r^2} Y_{lm}(\mathbf{r})\,. \tag{2.36}$$

Each primitive Gaussian is characterized by an exponent ζ_n and its angular momentum quantum numbers (l, m), which define the shape and orientation of the spherical harmonic $Y_{lm}(\mathbf{r})$. Typically, a *contracted* GTO basis set [89] is used, where the coefficients d_n are not subject to the variational minimization in (2.11). The parameters d_n and ζ_n are therefore only optimized once for each basis function of a given chemical element and then kept fixed during electronic structure calculations.

Compared to a plane wave basis, which would seem as a natural choice in solid-state physics, a properly constructed GTO basis set is much smaller, since its basis functions already resemble chemically meaningful orbitals, leading to a computational advantage of Gaussian basis sets over plane waves. However, contrary to plane waves, GTO basis sets are non-orthogonal, i.e. in general the overlap integrals of the basis functions are given by the *overlap matrix* $\underline{\mathbf{S}}$ with coefficients

$$\langle \varphi_k | \varphi_l \rangle = S_{kl} \neq \delta_{kl}\,. \tag{2.37}$$

This makes implementations difficult compared to plane wave codes. For instance, the KS equations (2.20) become the generalized eigenvalue problem [90]

$$\underline{\mathbf{H}}\, \mathbf{c}_i = \epsilon_i\, \underline{\mathbf{S}}\, \mathbf{c}_i \tag{2.38}$$

when expanded in a GTO basis set. Here, $\underline{\mathbf{H}}$ denotes the Kohn-Sham matrix, which represents the KS Hamiltonian $\hat{H}_{\mathrm{KS}}$ in the given basis set, hence the coefficients of this matrix are given by

$$H_{ij} = \langle \varphi_i | \hat{H}_{\mathrm{KS}} | \varphi_j \rangle\,. \tag{2.39}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek
Your knowledge hub
WIEN

The column vector $\mathbf{c}_i$ represents the KS wavefunction $|\psi_i\rangle$ and contains the expansion coefficients C_{ik} from (2.34). In CP2K, this eigenvalue problem is solved iteratively using an efficient orbital transformation method [91]. Furthermore, when calculating matrix elements involving derivatives with respect to the ionic coordinates, e.g. for evaluating atomic forces using the Hellman-Feynman theorem [92], the position dependence of the GTO basis has to be considered, leading to additional Pulay forces, see [90, 93] for further details. This position dependence will also become relevant in Sec. 3.1.2 for evaluating electron-phonon coupling matrix elements.

While the non-orthogonality and position dependence of the GTO basis set are cumbersome to work with analytically, they do not significantly impede on the numerical performance. However, the bottleneck for GTO based calculations is typically the evaluation of the Hartree energy U_H (2.18). To understand why, we first construct the electron density $n(\mathbf{r})$ according to (2.16) using the GTO basis:

$$n(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 = \sum_{jk} \underbrace{\sum_i^N C_{ij} C_{ik}}_{=P_{jk}} \varphi_j(\mathbf{r}) \varphi_k(\mathbf{r}) = \sum_{jk} P_{jk} \varphi_j(\mathbf{r}) \varphi_k(\mathbf{r}) \,. \tag{2.40}$$

In this basis the density is therefore represented by the density matrix

$$\underline{\mathbf{P}} = \underline{\mathbf{C}}\, \underline{\mathbf{C}}^\mathrm{T} \tag{2.41}$$

with $\underline{\mathbf{C}} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n]$ being the complete coefficient matrix of all occupied KS orbitals. Evaluating (2.18) then yields

$$U_\mathrm{H}[n] = \frac{1}{2} \sum_{ijkl} P_{ij} P_{kl} \iint \varphi_i(\mathbf{r}) \varphi_j(\mathbf{r}) \frac{1}{|\mathbf{r}-\mathbf{r}'|} \varphi_k(\mathbf{r}') \varphi_l(\mathbf{r}') \mathrm{d}^3\mathbf{r} \mathrm{d}^3\mathbf{r}' \,. \tag{2.42}$$

for the Hartree energy. However, similarly to the Hartree-Fock exchange in (2.26), the number of involved 4-center ERIs scales with $\mathcal{O}(N^4)$ and thus evaluating the Hartree energy quickly becomes the limiting computational step. The Gaussian and Plane Wave (GPW) method [90, 94] in CP2K and other similar approaches [95, 96] get around this problem by introducing an auxiliary plane wave basis to represent the electron density as

$$n(\mathbf{r}) = \frac{1}{\Omega} \sum_{\mathbf{G}} \widetilde{n}(\mathbf{G}) e^{\mathrm{i}\mathbf{G}\cdot\mathbf{r}} \tag{2.43}$$

with $\mathbf{G}$ being a reciprocal lattice vector and Ω being the volume of the simulation cell. Here, the Fourier transform of the density $\widetilde{n}(\mathbf{G})$ can be efficiently calculated using the fast fourier transform (FFT) algorithm [97], allowing for a dual-space representation of $n(\mathbf{r})$. Using the convolution theorem of the Fourier transform, the Hartree energy can then

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

simply be calculated in Fourier space as

$$U_\mathrm{H}[n] = 2\pi\Omega \sum_{\mathbf{G}} \frac{|\widetilde{n}(\mathbf{G})|^2}{|\mathbf{G}|^2}\,. \tag{2.44}$$

Throughout this work, the double-ζ Goedecker-Teter-Hutter basis sets [98] with additional polarization functions (DZVP) were used in combination with the corresponding Gaussian pseudo-potentials [75].

2.2.2 Auxiliary Density Matrix Method

As already discussed, the number of 4-center ERIs occuring in the evaluation of the HF-exchange and the Hartree energy severely restrict the possible system size due to their unfavorable $\mathcal{O}(N^4)$ scaling. While this problem can be eliminated for the Hartree energy by performing the calculation in Fourier space, the same is unfortunately not possible for the HF-exchange integrals due to the changed order of the indices compared to (2.42):

$$E_\mathrm{X}^\mathrm{HF}[\underline{\mathbf{P}}] = -\frac{1}{2}\sum_{ijkl} P_{ij}P_{kl} \iint \varphi_i(\mathbf{r})\varphi_k(\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}'|}\varphi_j(\mathbf{r}')\varphi_l(\mathbf{r}')\mathrm{d}^3\mathbf{r}\mathrm{d}^3\mathbf{r}' \tag{2.45}$$

As can be seen, the number of necessary ERI computations depends on the size and the sparsity of the density matrix $\underline{\mathbf{P}}$. The Auxiliary Density Matrix Method (ADMM) [99] implemented in CP2K mitigates the computational effort for Hartree-Fock exchange by approximately representing the KS wavefunctions in a smaller auxiliary basis set $\{\hat{\varphi}_k(\mathbf{r})\}$:

$$\hat{\psi}_i(\mathbf{r}) = \sum_k \hat{C}_{ik}\hat{\varphi}_k(\mathbf{r}) \approx \psi_i(\mathbf{r}) \tag{2.46}$$

The ideal coefficients $\hat{C}_{ik}$ are then determined by minimizing the square difference between the KS wavefunctions and their approximations, i.e.

$$\underline{\hat{\mathbf{C}}} = \arg\min_{\hat{C}} \sum_i^N \int \left|\psi_i(\mathbf{r}) - \hat{\psi}_i(\mathbf{r})\right|^2 \mathrm{d}^3\mathbf{r}\,. \tag{2.47}$$

This procedure leads to the linear mapping

$$\underline{\hat{\mathbf{C}}} = \underline{\hat{\mathbf{S}}}^{-1}\underline{\mathbf{Q}}\,\underline{\mathbf{C}} \quad \text{with} \quad \hat{S}_{ik} = \langle\hat{\varphi}_i|\hat{\varphi}_k\rangle \quad \text{and} \quad Q_{ik} = \langle\hat{\varphi}_i|\varphi_k\rangle \tag{2.48}$$

between the original and the auxiliary coefficient matrices $\underline{\mathbf{C}}$ and $\underline{\hat{\mathbf{C}}}$ respectively. It is then further assumed that the *difference* in exchange energy between these two basis sets is roughly the same for the semi-local GGA and the hybrid functional, resulting in an approximation for the HF exchange with the full basis set:

$$E_\mathrm{X}^\mathrm{HF}[\underline{\mathbf{P}}] \approx E_\mathrm{X}^\mathrm{HF}[\underline{\hat{\mathbf{P}}}] + E_\mathrm{X}^\mathrm{GGA}[\underline{\mathbf{P}}] - E_\mathrm{X}^\mathrm{GGA}[\underline{\hat{\mathbf{P}}}] \tag{2.49}$$

This approximation has been demonstrated to be very effective and can accelerate the calculation of the HF exchange by a factor of ∼20 without any significant loss in accuracy [99].

2.3 Lattice Dynamics

As discussed in Sec. 2.1.1, the BOA allows for a separation of the nuclei and electron dynamics. Here, the electrons give rise to a PES $V(\mathbf{R})$, which can be calculated with the electronic structure methods outlined above. While the electrons have to be always treated quantum mechanically, describing the movement of the nuclei (semi-)classically is often justified, due to their large mass. Depending on the magnitude of movement different descriptions can be applied for the nuclei. For instance, if the nuclei are moving around their respective equilibrium positions at finite temperature, the PES often can be approximated by an harmonic oscillator, leading to the concept of phonons. However, studying e.g. the melting process of a material at higher temperatures or chemical reactions, requires a more detailed (and computationally demanding) sampling of the PES far away from the equilibrium positions. The various computational methods used in this work for describing the lattice dynamics in different regimes are discussed in the following.

2.3.1 Phonons

For small thermal vibrations around an equilibrium position $\mathbf{R}_0$, the PES can be expanded to second order, yielding

$$V(\mathbf{u}) = V(\mathbf{R}_0) + \frac{1}{2}\sum_{ij}^{3N} \Phi_{ij} u_i u_j \tag{2.50}$$

where $\mathbf{u} = \mathbf{R} - \mathbf{R}_0$ denotes the displacement from the equilibrium position and $\underline{\mathbf{\Phi}}$ is the force matrix defined as

$$\Phi_{ij} = \left.\frac{\partial^2 V}{\partial u_i \partial u_j}\right|_{\mathbf{u}=\mathbf{0}} = -\left.\frac{\partial F_j}{\partial u_i}\right|_{\mathbf{u}=\mathbf{0}} . \tag{2.51}$$

While the first derivatives of energy with respect to the displacement, i.e. the (negative) atomic forces $\mathbf{F} = -\nabla V$, can be calculated from a single energy calculation using the Feynman-Hellman theorem [92], the same does *not* hold true for the force matrix, which is defined by the second derivatives of the energy. This is due to the well-known “(2n+1) theorem” [100, 101], which states that in order to calculate the $(2n + 1)$th derivative of the energy, the nth derivative of the electron density or of the wavefunctions needs to be known. Within the frozen-phonon approach [102], constructing the full harmonic approximation of the PES requires at least[1] $3N$ additional self-consistent energy calculations to evaluate the derivative of the forces using finite displacements.

---

[1] The number could be reduced due to symmetries of the lattice.

The classical equations of motion (EOM) for the nuclei are then given as

$$\sum_j \Phi_{ij} u_j = -M_i \ddot{u}_i = -F_i\,, \tag{2.52}$$

where M_i is the mass of the nucleus associated with the ith coordinate. Transforming to mass-weighted coordinates $q_i = \sqrt{M_i} u_i$ and assuming a time-harmonic motion, i.e. $q_i(t) = \hat{q}_i \exp(\mathrm{i}\omega t)$, results in the eigenvalue problem

$$\underline{\mathbf{D}}\hat{\mathbf{q}} = \omega^2 \hat{\mathbf{q}} \tag{2.53}$$

for the dynamical matrix $\underline{\mathbf{D}}$ with coefficients

$$D_{ij} = \frac{1}{\sqrt{M_i M_j}} \Phi_{ij}\,. \tag{2.54}$$

Diagonalization of $\underline{\mathbf{D}}$ then yields the eigenvectors $\hat{\mathbf{q}}_n$, called *normal modes* or *phonons*. The square root of the corresponding eigenvalue is the phonon frequency ω_n. Even though this is a classical treatment of phonons, the same coordinate transformation also diagonalizes the vibrational Hamiltonian (2.8), resulting in $3N$ decoupled 1-dimensional quantum mechanical harmonic oscillators:

$$\hat{H}_{\mathrm{vib}} = \sum_i^{3N} \hbar\omega_i \left(\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2} \right)\,, \tag{2.55}$$

where $\hat{a}_i^\dagger$ and $\hat{a}_i$ are the usual creation and annihilation operators respectively for increasing or decreasing the number of phonons in mode i. This also separates the corresponding vibrational wavefunction into a product state of the form

$$\chi(\mathbf{q}) = \chi_1(q_1) \times \chi_2(q_2) \cdots \times \chi_{3N}(q_{3N})\,, \tag{2.56}$$

which greatly simplifies the mathematics. The concept of phonons and the harmonic oscillator are fundamental to solid state physics and is frequently used to predict e.g. thermodynamic material quantities like the heat capacity, the temperature-dependent bandgap renormalization due to thermal atomic fluctuations [103] or phonon-limited carrier mobilities in a semiconductor [104]. In this work, HOs are used to describe defects in their different charge states. In particular for a multi-dimensional treatment of the charge transfer, the HO model is essential to derive an analytical model for the transition rates as will be discussed in Sec. 3.2.3.

It should be noted that the frozen-phonon derivation as given above, in particular (2.53), only describes phonons at the Γ-point, i.e. with a vanishing wave vector. However, this is sufficient for the purposes of this work, since defects have to be simulated in a large supercell of the material, where many wavevectors in the primitive unit cell are folded onto Γ. When calculating phonons with arbitrary wavevector, e.g. to obtain the full phonon

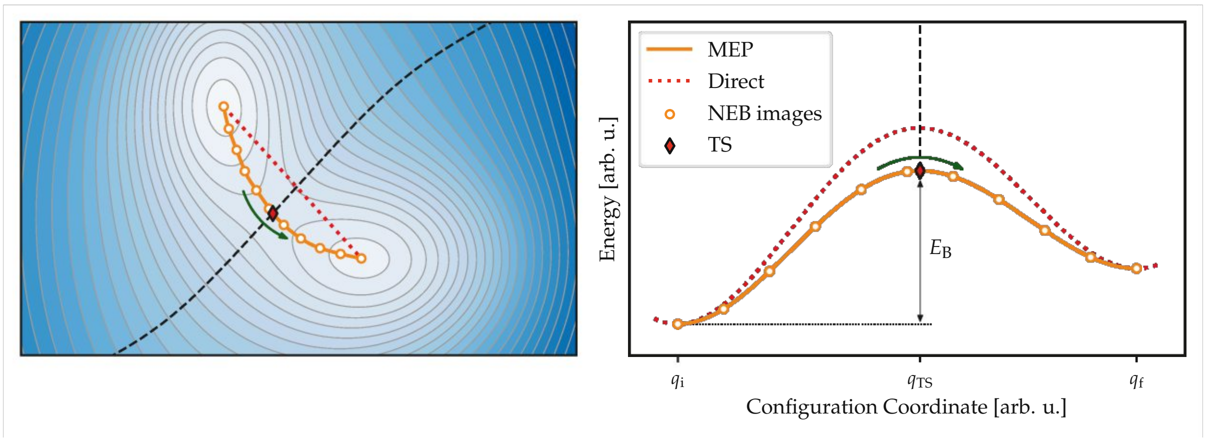


Figure 2.3. Left: Transition between two minima on an analytical 2-dimensional PES. The two minima are connected via a first-order saddle-point, the so-called transition state (TS). The two regions in configuration space are separated by the surface $\mathcal{S}$ (dashed line). **Right:** Potential energy curves for the minimum energy path (MEP) and the direct path. The MEP leads over the TS and is the path with the lowest possible energy barrier E_B. Numerically, the MEP is approximated by a chain of configurations ("images") in the nudged elastic band (NEB) method. (adapted from [DWT1])

dispersion, it is usually more efficient to use the density functional perturbation theory (DFPT) approach [105], which does not rely on calculations in a large supercell like the frozen-phonon approach.

2.3.2 Thermal Transitions

While phonons are a valuable concept to describe a solid in the vicinity of an equilibrium position, they cannot be used directly to describe chemical reactions. During a chemical reaction, the system moves from one equilibrium position on the PES to another. As shown in Fig. 2.3, this motion necessarily leads over a saddlepoint of the PES, the so-called transition state, and thus has to be naturally anharmonic.

For such reactions, the transition rate is governed by classical transition state theory (TST), which was first derived by Eyring, Evans and Polanyi [106, 107] in 1935. In TST the transition rate is given by the total flux in configuration space which crosses the surface $\mathcal{S}$ separating the basins of the initial and final state (dashed line in Fig. 2.3). Further assuming a thermal equilibrium in the initial state, the transition rate can be expressed as [108, 109]

$$k_\mathrm{TST} = \frac{\mathcal{Z}_\mathcal{S}}{\mathcal{Z}_\mathcal{I}} \sqrt{\frac{k_\mathrm{B} T}{2\pi}}\,, \tag{2.57}$$

where $\mathcal{Z}_\mathcal{S}$ denotes the classical partition function when integrating over the separating surface $\mathcal{S}$:

$$\mathcal{Z}_\mathcal{S} = \int_\mathcal{S} e^{-V(\mathbf{q})/k_\mathrm{B}T} \mathrm{d}\mathbf{q} \tag{2.58}$$

Similarly, the partition function of the initial state basin $\mathcal{I}$ is given by

$$\mathcal{Z}_{\mathcal{I}} = \int_{\mathcal{I}} e^{-V(\mathbf{q})/k_{\mathrm{B}}T} \mathrm{d}\mathbf{q}\,. \tag{2.59}$$

While (2.57) in principle is valid for any PES $V(\mathbf{q})$, it is impractical to compute from first principles, since evaluating the partition functions would require an extensive sampling of the PES. However, due to the Boltzmann factor both partition functions are dominated by their lowest energy regions which allows to use harmonic approximations for the PES in those regions. For the initial basin $\mathcal{I}$ the expansion at the minimum configuration can be written as

$$V(\mathbf{q}) = V(\mathbf{q}_{\mathrm{i}}) + \frac{1}{2}\sum_{n=1}^{3N} \omega_{\mathrm{i},n}^2 (q_n - q_{\mathrm{i},n})^2\,. \tag{2.60}$$

The separating surface on the other hand obtains its lowest energy at the transition state (TS) and hence can be expanded as

$$V(\mathbf{q}) = V(\mathbf{q}_{\mathrm{TS}}) - \frac{1}{2}\omega_{\perp}^2 (q_{\perp} - q_{\mathrm{TS},\perp})^2 + \frac{1}{2}\sum_{n=1}^{3N-1} \omega_{\mathrm{TS},n}^2 (q_n - q_{\mathrm{TS},n})^2\,. \tag{2.61}$$

Note that the normal mode $q_{\perp}$, which facilitates the transition and is perpendicular to $\mathcal{S}$, has an imaginary frequency due the TS being a first order saddlepoint [109]. Within this harmonic TST (HTST), the partion functions can be integrated analytically, yielding

$$k_{\mathrm{TST}} = \frac{1}{2\pi}\frac{\prod_n^{3N} \omega_{\mathrm{i},n}}{\prod_n^{3N-1} \omega_{\mathrm{TS},n}} e^{-E_{\mathrm{B}}/k_{\mathrm{B}}T} \tag{2.62}$$

for the transition rate. As can be seen, the transition rate follows the well known Arrhenius law [110] and hence depends exponentially on the energy barrier E_{B} along the reaction path. Although other paths can also contribute to the overall transition rate, the minimum energy path (MEP) leading over the TS and therefore having the lowest possible barrier E_{B} dominates the transition rate due this exponential energy dependence. Although (2.62) already makes calculating the transition rate feasible compared to the general form (2.57), it still requires to expand the PES to second order at the initial minimum and the TS. However, (2.62) can be further simplified by assuming the reaction only occurs along the 1-dimensional MEP, hence the reaction can be described by a so-called configuration coordinate q and the corresponding PEC $V(q)$ as shown in Fig. 2.3 (right). In this case, the transition rate reduces to

$$k_{\mathrm{TST}} = \frac{\omega_0}{2\pi} e^{-E_{\mathrm{B}}/k_{\mathrm{B}}T} \tag{2.63}$$

with ω_0 being the angular frequency associated with the initial minimum in the local direction of the MEP. The prefactor $\nu = \omega_0/2\pi$ is often referred to as the *attempt frequency*, which can be interpreted as the number of attempts per unit time to overcome the barrier. Within this approach, the main task is to locate the TS between the two minima and to construct the MEP as will be explained in the following.

Nudged Elastic Band Method

The standard approach to determine the MEP is the nudged elastic band (NEB) method [111], in which the reaction path is represented by a finite number of different configurations, so-called "images", as indicated by the circles in Fig. 2.3. Each image $\mathbf{R}_i$ is connected to its neighbors by virtual springs enacting a force

$$\mathbf{F}_i^{\text{spring}} = k(\mathbf{R}_{i+1} - \mathbf{R}_i) - k(\mathbf{R}_i - \mathbf{R}_{i-1}) \tag{2.64}$$

on the images, where k denotes the spring constant. These spring forces ensure equally spaced images along the reaction path. The images are further subject to the physical forces $\mathbf{F}_i = -\nabla V|_{\mathbf{R}_i}$ arising from the PES. Note that the images at the previously determined minimum configurations are not influenced by the spring forces in order to keep them at their respective minimum positions. The main idea is now to relax all the images in the band with respect to the total force $\mathbf{F}_i^{\text{total}} = \mathbf{F}_i^{\text{spring}} + \mathbf{F}_i$. However, such a scheme, known as plain-elastic band (PEB) method [112], would not be guaranteed to converge towards the MEP, due to its tendency to cut corners when using larger spring constants. Instead, the NEB method, introduces a small modification to the forces acting on the images in such a way that the physical forces only act perpendicular to the local reaction path, whereas the spring forces only act tangentially to it. Introducing an approximation

$$\mathbf{t}_i = \frac{\mathbf{R}_{i+1} - \mathbf{R}_{i-1}}{|\mathbf{R}_{i+1} - \mathbf{R}_{i-1}|} \tag{2.65}$$

for the local tangent of the reaction path, the NEB forces for each image are then given by

$$\mathbf{F}_i^{\text{NEB}} = \mathbf{F}_i - (\mathbf{F}_i \cdot \mathbf{t}_i)\mathbf{t}_i + (\mathbf{F}_i^{\text{spring}} \cdot \mathbf{t}_i)\mathbf{t}_i\,. \tag{2.66}$$

Although this algorithm uses a very simple approximation for the local tangent, it already provides a reliable method to determine the MEP independent of the chosen spring constant k [111]. However, improved estimates for the local tangent leading to a better convergence can also be applied [113]. Since the TS and the resulting barrier E_B are the most important parameters within the classical transition state theory, the converged NEB path should contain the TS as the image with highest energy. However, as shown in Fig. 2.3, while the NEB images lie on the MEP, they could miss the TS due to the enforced spacing between images. To overcome this, the widely used climbing-image NEB (CI-NEB) [114] releases the highest energy image from the spring forces and lets it move upwards towards to the TS, hence providing a better estimate for the barrier E_B.

Due to its importance for determining activation energies and chemical reaction pathways, the NEB method has been adapted and improved many times over the years. For instance, the original version of NEB used a simple linear interpolation, i.e. the direct path, between the minimum configurations as an initial guess for the MEP. However, much more efficient initial guesses closer to the MEP can be constructed by means of pre-optimizing the inital straight band e.g. by using an auxiliary potential is based on simple pairwise distances [115] or by searching for a geodesic on a manifold spanned by

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

non-redundant coordinates [116]. Further improvements have been achieved by observing that images close to the TS move the most during the optimization, while the other images do not move much at all after a few optimization cycles. This is exploited in the AutoNEB method [117], which only reevaluates an image using DFT, if the forces acting on it are above a certain threshold. Other approaches [118, 119] construct a surrogate PES model using Gaussian Process Regression (GPR) [120] on which the images are optimized. As long as the images do not move much, the surrogate PES is accurate enough to converge to the full NEB band with significant performance improvements but no loss of accuracy.

It should be noted that next to NEB, other methods exist, which directly optimize towards a TS without the need of simulating a whole band of images. The prototypical example for such approaches is the dimer method [121], which rotates a local pair of images ("dimer") to align it with the direction of the smallest curvature on the PES while also moving it uphill out of an equilibrium position. Building on the dimer method, other similar approaches have been developed, which search for a TS starting from a given minimum ("activation") and from then on explore the PES further by searching for new minima ("relaxation"). Contrary to the NEB method, those so-called activation-relaxation techniques (ART) [122–124] do not require prior knowledge of the final state of a transition, since they directly search for a local saddlepoint on the PES. Such methods are hence particularly useful for mapping out possible reaction pathways involved in more complicated chemical reactions.

In this work, the CI-NEB method is utilized to study the migration of hydrogen in SiO_2 after hole-induced Si-H bond dissociation, see Sec. 4.1.4 for further details.

2.3.3 Molecular Dynamics

Molecular dynamics (MD) is a widely used method for studying complex phenomena such as thin film deposition [125], microscopic heat transport [126, 127], ion-implantation in semiconductors [128], and Li-ion migration in batteries [129], among many other applications. These processes are often too intricate to be broken down into discrete steps that can be analyzed by methods like the aforementioned NEB. Instead, it is often more effective to observe the time evolution of the entire system rather than focusing on individual reaction steps. In MD simulations, atoms are typically modeled as classical particles moving on the PES $V(\mathbf{R})$, hence their dynamics are described by the classical equations of motion

$$\dot{\mathbf{R}}_i = \mathbf{v}_i \quad \text{and} \quad \dot{\mathbf{v}}_i = \mathbf{a}_i = \frac{1}{M_i}\mathbf{F}_i = -\frac{1}{M_i}\nabla_{\mathbf{R}_i} V(\mathbf{R}) \tag{2.67}$$

where $\mathbf{R}_i$, $\mathbf{v}_i$ and $\mathbf{a}_i$ represent the position, velocity, and acceleration of atom i, respectively.

This set of coupled equations is usually integrated over time using the Verlet algorithm [130, 131], in which the dynamical quantities are updated from time t_n to t_{n+1} according to

$$\mathbf{R}_i(t_{n+1}) = \mathbf{R}_i(t_n) + \mathbf{v}_i(t_n)\Delta t + \frac{1}{2}\mathbf{a}_i(t_n)\Delta t^2 \tag{2.68a}$$

$$\mathbf{a}_i(t_{n+1}) = \frac{1}{M_i}\mathbf{F}_i\big(\mathbf{R}(t_{n+1})\big) \tag{2.68b}$$

$$\mathbf{v}_i(t_{n+1}) = \mathbf{v}_i(t_n) + \frac{1}{2}\big(\mathbf{a}_i(t_n) + \mathbf{a}_i(t_{n+1})\big)\Delta t\,. \tag{2.68c}$$

The time step $\Delta t = t_{n+1} - t_n$ has to be small enough to capture the dynamics of the physical process accurately and to keep the simulation stable while also being large enough to computationally efficient. Typically, the time step is chosen to be on the order of $\sim$1 fs [132].

Within the scope of this work, MD simulations are mainly used to create atomistic models of amorphous oxide which serve as host materials for studying defects. Such structures are typically created with the melt-quench technique [133] from a crystalline configuration by melting and subsequent cooling in a MD simulation, cf. Sec. 4.2.1.

Statistical Ensembles

The classical equations of motion (2.67) naturally preserve the total energy of the system, since all forces stem from the potential $V(\mathbf{R})$ and are thus conservative. In the context of statistical mechanics, this would correspond to a microcanonical NVE ensemble, in which the system is closed meaning the number of particles (N), the volume (V) and the total energy (E) are conserved. However, it is often required to study the system while being coupled to its environment with which it can exchange energy, particles etc. For instance, when simulating a system at a constant temperature T, i.e. a canonical NVT ensemble, the system must be in thermodynamic equilibrium with a reservoir, necessitating an exchange of energy. In such cases, the EOMs have to be adapted in order to account for the coupling to the environment. In the following, algorithms for keeping the temperature constant, so-called thermostats, are discussed.

The equipartition theorem states that at thermal equilibrium each particle in the system carries on average a kinetic energy of $3k_\mathrm{B}T/2$. This can be used to define the instantanous temperature of a system as [134]

$$T = \frac{2}{3}\frac{\langle K\rangle}{k_\mathrm{B}} \quad \text{with} \quad \langle K\rangle = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}M_i|\mathbf{v}_i|^2\,. \tag{2.69}$$

A simple way to enforce the temperature on the system proposed by Berendsen *et al.* [135] would be to rescale the velocities $\mathbf{v}_i$ at each time step so that the instantaneous temperature evolves according to

$$\frac{\mathrm{d}T}{\mathrm{d}t} = -\frac{1}{\tau_T}(T - T_0) \tag{2.70}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

with T_0 being the temperature of the reservoir. The time constant τ_T is a measure for the coupling strength between the system and the reservoir, and determines the rate at which the temperature in the system can fluctuate. However, the Berendsen thermostat and other methods based on velocity scaling produce a non-canonical ensemble, which can result in simulation artifacts such as the "flying ice cube effect" [136]. Here, kinetic energy is transferred to low frequency modes, in particular to rigid translations, over time. To overcome this, a stochastic scaling algorithm was proposed [137], in which the kinetic energy of the system is separately propagated in a stochastic process and then enforced on the system by velocity rescaling. It can be shown that such an approach results in a correct canonical ensemble, allowing to determine thermodynamical quantities from the resulting MD trajectories.

Another widely used algorithm for a NVT ensemble is the *Nosé-Hoover* thermostat [138, 139]. Here, the EOMs are augmented with an additional degree of freedom ξ representing the thermal bath [140]:

$$\dot{\mathbf{R}}_i = \mathbf{v}_i\,, \quad \dot{\mathbf{v}}_i = \frac{1}{M_i}\mathbf{F}_i - \xi\mathbf{v}_i \quad \text{and} \quad \dot{\xi} = \alpha\sum_{i=1}^{N}\left(M_i|\mathbf{v}_i|^2 - 3k_\mathrm{B}T\right) \tag{2.71}$$

The parameter $1/\alpha$ is an effective mass associated with the bath degree ξ. It was demonstrated that in very simple systems like a single harmonic oscillator, the Nosé-Hoover thermostat behaves in a non-ergodic way and thus is also inconsistent with the usual description of a canonical ensemble [140]. However, this problem can be practically solved by using Nosé-Hoover chains [141] with multiple bath variables ξ_i coupled to each other.

The aforementioned methods are based on modifying the EOMs in a deterministic way, hence the entire trajectory of the system is fully determined by its initial conditions. This can be avoided by using a thermostat based on *Langevin dynamics*, where the EOMs are given as [142]

$$M_i\ddot{\mathbf{R}}_i = -\nabla_{\mathbf{R}_i}V(\mathbf{R}) - \gamma_i M_i\dot{\mathbf{R}}_i + \mathbf{X}_i\,. \tag{2.72}$$

Here, the thermal fluctuations are described by a Gaussian stochastic force component $\mathbf{X}_i$ with vanishing mean

$$\langle\mathbf{X}_i(t)\rangle = \mathbf{0} \tag{2.73}$$

and autocorrelation function

$$\langle\mathbf{X}_i(t)\cdot\mathbf{X}_j(t')\rangle = 2\gamma_i M_i k_\mathrm{B}T\delta(t-t')\delta_{ij}. \tag{2.74}$$

The friction coefficient γ_i not only dampens the motion of particles, but also needs to be related to the variance of the random force component in order to obey the fluctuation-dissipation theorem [143].

For a more in-depth discussion of different thermostat algorithms, the reader is referred to the literature [134]. Besides the discussed canonical NVT ensemble, the Gibbs or NpT ensemble is also of considerable practical importance, since most chemical reactions occur under constant pressure conditions. In this case, not only the temperature but also the pressure needs to be controlled during the simulation. Similarly to the discussed thermostat algorithms, this can be achieved by the use of barostats, see for instance [144]. In particular instances it can also be necessary to simulate a grand-canonical ensemble for systems exchanging particles with their environment. Examples for such MD simulations include the oxidation of Si in an O_2 atmosphere, which needs to be replenished regularly [DWJ7] or crystal growth from a solution at constant chemical potential [145].

2.3.4 Beyond Classical Dynamics

While the description of lattice dynamics in a harmonic potential via phonons can be done quantum mechanically by solving (2.55), the treatment within TST and MD as described above is purely classical. However, considering nuclear quantum effects (NQE) [146] such as tunneling, zero-point motion or the isotope effect can be important for understanding reaction kinetics, in particular when light elements like hydrogen are involved [147–150] or when studying reactions at cryogenic temperatures [DWC1, DWC2, 151]. The main difficulty for including quantum effects in dynamical simulations is that the (anharmonic) PES can only be calculated in a pointwise fashion using DFT, hence it is infeasible to calculate the vibrational wavefunctions for a system with multiple degrees of freedom. Methods enabling the inclusion of NQEs therefore follow different approaches which are usually derived from the path-integral formulation of quantum mechanics. For instance, ring-polymer molecular dynamics (RPMD) [152] considers a closed loop of images connected via virtual springs comparable to the NEB method, and propagates the whole loop classically in an extended phase space. It can then be shown that the partition function of the loop approximates the true quantum mechanical partition function, allowing for the extraction of thermodynamic quantities and reaction rates with the inclusion of quantum effects. Recently a link between RPMD and instanton theory has been established [153], in which transitions are described by a semiclassical tunneling pathway analogous to the MEP for classical transitions. While these approaches neglect realtime quantum coherence [152], their semi-classical nature allows to scale them to large system sizes.

Within the scope of this work, the classical NEB and MD approaches are deemed sufficient to describe the relevant physics. However, NQEs are essential for modeling charge transfer kinetics, especially at cryogenic temperatures, and thus are included in the 1-dimensional charge trapping model as will be discussed in Ch. 5.

2.4 Defect Calculations

This section is based on works published in [DWJ1].

In order to study the electronic structure of point defects in DFT, a suitable atomistic model for the host material is needed. There are basically two different approaches of modeling the local environment of a defect. Coming from a chemical perspective, a small cluster of the host material can be used in a simulation cell with closed boundary conditions to emulate the defect environment. In order to increase the accuracy of this rather crude model, the isolated cluster can be embedded in a potential, which mimics the impact of the surrounding host material. This is known as the embedded cluster method (ECM) [154, 155]. Such cluster models were primarily used in earlier theoretical works due to their comparatively low computational costs. Nowadays however, defects are typically simulated in a large periodic supercell of the material containing a few hundred atoms [156]. All results presented in this work are obtained with this supercell approach. While creating a suitable supercell for defects in crystalline semiconductors is straightforward, studying defects in a typically amorphous gate dielectric is more complex. In this case, a physical atomistic model of the amorphous host structure must first be prepared, for example, by performing a melt-quench molecular dynamics simulation (see Sec. 4.2.1).

2.4.1 Defect Formation Energies

Simulating charge trapping at a defect site requires to perform DFT calculations in different charge states of the defect. However, when comparing total energies from differently charged supercells, additional alignments and corrections have to be accounted for. The standard method to address this is to calculate defect formation energies. Within the established supercell approach, the formation energy of a defect configuration X_q in charge state q is given by [156–158]

$$E_\mathrm{f}\left[X_q\right] = E_\mathrm{tot}\left[X_q\right] - E_\mathrm{tot}\left[\mathrm{bulk}\right] - \sum_i n_i \mu_i + q(\epsilon_\mathrm{VBM} + E_\mathrm{F}) + \Delta_q \,, \tag{2.75}$$

where $E_\mathrm{tot}\left[X_q\right]$ and $E_\mathrm{tot}\left[\mathrm{bulk}\right]$ denote the total energies of the defective and pristine bulk supercells, respectively. The formation energy of the charged defect also depends on the Fermi level E_F in the interacting carrier reservoir. Usually, E_F is given with respect to the valence band maximum (VBM) of the host material, which is approximated by the energy ϵ_VBM of the highest occupied KS state in the bulk. The chemical potential μ_i accounts for the energy required to add ($n_i > 0$) or remove ($n_i < 0$) atoms of kind i to construct the defect of interest. The chemical potentials depend on the process conditions during growth and are therefore variable. However, firm bounds for their values can be established by calculating the energy of elemental or limiting phases within the considered process [157, 158]. Contributions from the chemical potential are relevant when total formation energies are required, e.g. to estimate defect concentrations.

2.4.2 Finite-Size Corrections

When simulating charged defects in a periodic cell, the electrostatic potential would become infinite due to the superposition of long-range Coulomb contributions from infinitely many cells. To avoid this, the net charge in the periodic cell is compensated by a fictitious background charge of opposite sign, ensuring neutrality and a finite potential. While this approach results in a bounded electrostatic potential, it introduces an unwanted interaction between the defect charge and the neutralizing background charge. Additionally, there is

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

a spurious electrostatic interaction between the charged defect and its periodic images, which, especially in small simulation cells, must be corrected by a charge correction term Δ_q.

Various correction schemes have been developed, the simplest of which is the Makov-Payne (MP) scheme. For a cubic supercell with lattice constant L, the first-order correction term in the MP scheme is given by [159]

$$\Delta_q = \frac{\alpha q^2}{2\varepsilon L}\,, \tag{2.76}$$

where $\alpha = 2.8373$ is the Madelung constant for a simple cubic lattice. For a general lattice, the Madelung constant can be calculated using the Ewald summation method [160, 161]. The electrostatic potential is screened by the macroscopic permittivity ε of the host material. In this simple correction scheme, the defect charge q is assumed to be a point charge, neglecting its spatial distribution. While the MP scheme can be extended to include higher moments of the charge distribution [159], a correction scheme that incorporates the actual charge distribution from first-principles calculations would be preferable. This is the main idea behind the Freysoldt-Neugebauer-van de Walle (FNV) correction scheme [162]. For a more detailed discussion about different finite-size corrections the reader is referred to the literature [160].

In this work, the FNV correction, as implemented in the SXDEFECTALIGN package [163], is used to compensate for finite-size effects in crystalline materials. However, this method requires manual alignment of the electrostatic potential. Therefore, the simpler MP scheme in the form (2.76) is used for defects in amorphous materials due to the large number of sampled defects in these materials.

2.4.3 Charge Transition Levels

Formation energies determine the relative thermodynamic stability of different defect states as a function of E_F. As shown in the formation energy diagram Fig. 2.4, the thermodynamic charge transition level (CTL), $E_\mathrm{T}^{q/q'}$, between two charge states q and q', is defined as the Fermi energy at which both states are equally stable. Therefore, $E_\mathrm{T}^{q/q'}$ marks the boundary between different charge stability regions and is given by [164]

$$E_\mathrm{T}^{q/q'} = \frac{(E_\mathrm{tot}[X_q] + \Delta_q) - (E_\mathrm{tot}[X_{q'}] + \Delta_{q'})}{q' - q} - \epsilon_\mathrm{VBM}\,. \tag{2.77}$$

CTLs are a crucial defect parameter, as they determine whether a defect can change its charge state during device operation. This occurs when the Fermi level in the device channel crosses the CTL. Consequently, CTLs near the band edges of the semiconductor channel are especially important for device stability. It is also important to note that CTLs are independent of the chemical potentials μ_i, as they only depend on differences in formation energies rather than on absolute formation energies.

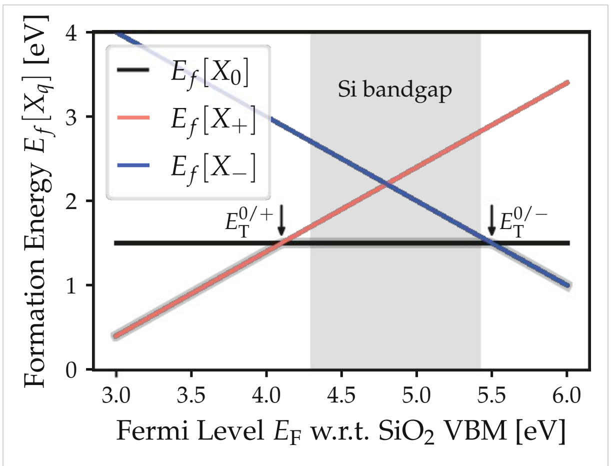


Figure 2.4. Formation energies of a hydrogen bridge defect in a-SiO_2 for different charge states, see also Sec. 4.3.2. The Fermi level E_F is given with respect to the SiO_2 valence band maximum. The lowest energy state at a given E_F is thermodynamically stable and highlighted by the thick gray line. The shaded region indicates the bandgap of the Si channel. The charge transition levels (CTLs) are given by the Fermi levels at which two formation energies cross. If they are close to the Si band edges, the defect can change its charge state during device operation, leading to instabilities. (taken from [DWJ1])

Moreover, CTLs can be directly measured in random telegraph noise (RTN) experiments, since at these levels both charge states have equal occupancy, leading to maximum noise power. The CTL is therefore easily accessible in experiments by varying the gate bias voltage [17] and allows to identify defects by comparison to theoretical CTLs for various defect candidates [23, DWJ8, 165] as will be discussed in Sec. 4.3.4 and Ch. 6.

2.5 Computational Settings

All DFT calculations in this work were performed with the CP2K [85] software package. The evaluation of the PES $V(\mathbf{R})$ and the atomic forces $\mathbf{F} = -\nabla V$ for a given set of nuclei positions $\mathbf{R}$ is the fundamental building block for all first-principles calculations in this thesis. These so-called single-point calculations were performed with the QUICKSTEP routine [90], which self-consistently solves the KS equations (2.20) within the GPW approach described in Sec. 2.2.1. Unless stated otherwise, the following settings were used throughout this work to perform DFT calculations:

Pseudopotentials and Basis Sets

The GTH-PBE pseudopotential [75] is used for all calculations. The molecular orbitals and the electron density are represented in the contracted double-ζ Gaussian basis set DZVP-MOLOPT-PBE-GTH [98], which was optimized for the aforementioned pseudopotential. Here, double-ζ indicates that two 2 primitive Gaussians are used per atomic orbital in (2.35). The basis set also contains additional polarization functions to account for asymmetric charge distributions around the atoms.

Hybrid Functionals

The hybrid functionals HSE [66] and PBE0_TC_LRC [67] are used due to their ability of providing an accurate electronic structure even when localized charges are present, which is highly important for describing defects. In order to mitigate the high computational costs compared to semi-local functionals like PBE, the auxiliary density matrix method (ADMM) [99] in conjunction with the admm-dzp auxiliary basis set is employed.

SCF Convergence

The electronic structure is converged to a threshold of 2.7×10^{-6} eV for the change in total energy during a SCF cycle. A kinetic energy cutoff of 800 Ry for the auxiliary planewave basis in the GPW method [90, 94] is used.

Geometry Optimizations

Based on the forces **F** from subsequent single-point calculations, the minimum configurations on the PES are found by the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [166–169]. The convergence threshold for the atomic forces and the total pressure within the simulation cell where set to $5\,\mathrm{meV\,Å^{-1}}$ and 0.01 GPa respectively. The spring constants in NEB calculations were set to $2.5\,\mathrm{eV\,Å^{-1}}$.

Phonon Calculations

Since calculating phonons requires a very well converged minimum configuration, the force convergence threshold was set to $1\,\mathrm{meV\,Å^{-1}}$ for these calculations. Additionally, evaluating the electron-phonon coupling strength with a finite-differences scheme as explained in Sec. 3.1.2 is done with a tighter SCF convergence threshold of 2.7×10^{-8} eV and a higher planewave cutoff of 1200 Ry in order to reduce numerical noise.

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

3

Nonradiative Multiphonon Transitions

In this chapter, the theory of nonradiative multiphonon (NMP) transitions is reviewed. This theoretical framework forms the foundation for modeling charge capture and emission processes at defect sites in the gate oxide, as will be discussed in Ch. 5. Before addressing the necessary modifications for oxide defects, we first focus on the core principles of NMP transitions in semiconductors and how the resulting charge transition rates can be estimated for a given defect candidate using DFT.

For practical calculations, it is typically assumed that the transition occurs along a single effective phonon mode. In the second part of this chapter, a novel formalism is introduced that incorporates the effects of additional phonon modes, while still being representable in a 1-dimensional configuration coordinate diagram. This approach enables an efficient multimode treatment of charge transfer. The chapter concludes with a discussion of the classical limit of quantum mechanical NMP theory, which will be employed for device-level simulations in Ch. 5 and 6 due to its significant computational advantages.

3.1 Basic Theory

Point defects in semiconductors have been studied for decades due to their detrimental effects on device performance. In particular, defects can act as charge trapping or recombination centers, leading to shorter carrier lifetimes [170], limited carrier mobility due to charged impurity scattering [171], or lower quantum efficiency in optoelectronic devices like solar cells [172] or light-emitting diodes [173]. These trapping and recombination processes typically happen nonradiatively, i.e. without the emission or absorption of photons. The basic theory of such nonradiative transitions and their treatment within first-principles calculations are discussed in the following.

3.1.1 Nonradiative Multiphonon Emission

A prime example for the aforementioned charge trapping and recombination processes is the well-known Shockley–Read–Hall (SRH) recombination [175, 176], in which carriers from the semiconductor band edges recombine via a localized defect state in the bandgap

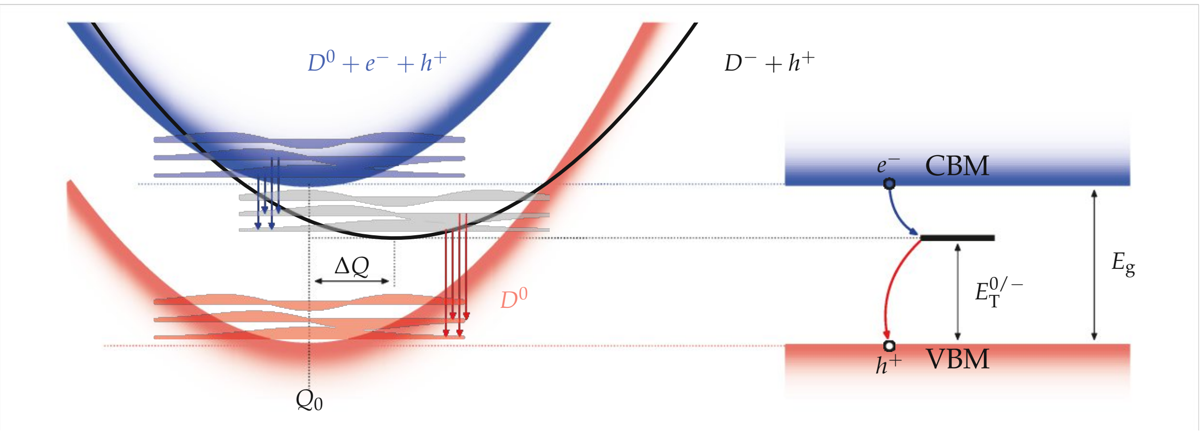


Figure 3.1. Schematic illustration of Shockley-Read-Hall (SRH) recombination at a deep trap in a semiconductor. In this case, the defect is acceptor-like, meaning it can be in a neutral state D^0 and a negatively charged state D^-. Initially, the defect is assumed to be neutral and surrounded by an electron-hole pair (state $D^0 + e^- + h^+$). The electron can be trapped from the CBM of the semiconductor into the localized defect state at the trap level $E_\mathrm{T}^{0/-}$, leading to the defect being negatively charged (state $D^- + h^+$). Each of these two states is represented by its own individual PES shown on the left side. Since the defect oscillates around its respective equilibrium position, both PES can be reasonably approximated by harmonic oscillators. Due to the large ionic displacement ΔQ, the two states can couple via many different vibrational states, which do not need to be related by the well-known selection rules for transitions within a harmonic oscillator. The now negatively charged defect can further emit its electron into the VB (trap a hole from the VB), where it recombines with a hole. Note that in the state before $(D^0 + e^- + h^+)$ and after (D^0) recombination the defect is neutral. Both of these states thus share the same PES but with an energetic offset equal to the bandgap E_g of the semiconductor. (adapted from [174])

as shown in Fig. 3.1. For an acceptor (donor-like) trap, the defect first traps an electron (hole) from the semiconductor CBM (VBM) followed by emission of the carrier into the VBM (CBM). During such transitions the excess energy is dissipated via phonon emission. In particular for deep traps the energy differences between initial and final states are larger than typical phonon frequencies, hence the energy has to be dissipated by emitting multiple phonons. This is enabled by the ionic relaxation ΔQ associated with a change in the electronic state, which lifts the selection rule $\Delta n = \pm 1$ for phonons in a harmonic oscillator [174, 177]. Hence, for a given initial electronic state $|\psi_i\rangle$ and a final state $|\psi_f\rangle$ many different vibronic transitions $|\psi_i \otimes \chi_{im}\rangle \rightarrow |\psi_f \otimes \chi_{fn}\rangle$ need to be considered for calculating the total transition rate.

In order for an initially delocalized carrier in the CB or VB to get trapped at a defect site, a perturbation is required, which couples the electronic wavefunction of the isolated defect in the bandgap to the carrier wavefunction, leading to a nonvanishing transition matrix element. For nonradiative transitions, electron-phonon coupling is usually the dominant coupling mechanism. Since the timescales for nonradiative transitions are usually much larger than phonon lifetimes or time periods of individial phonon modes, the perturbation can be best formulated within the *static coupling approximation* (SCA) [174, 178]. Here, the electronic wavefunctions are evaluated at a given ionic configuration $\mathbf{Q}_0$, leading to a vibronic wavefunction of the form

$$|\psi_i \otimes \chi_{im}\rangle = \psi_i(\mathbf{r}; \mathbf{Q}_0)\chi_{im}(\mathbf{Q}) , \tag{3.1}$$

where $\mathbf{r}$ denotes the electronic coordinates. The configuration $\mathbf{Q}_0$ is usually chosen to be the equilibrium configuration of either the initial or final state of the transition. This approximation has been shown to yield capture coefficients consistent with available experimental data, unlike other schemes such as the adiabatic approximation [179]. Within the SCA, the electron-phonon coupling operator is given by [174, 179]

$$\Delta\hat{H}_{\mathrm{el-ph}}(\mathbf{Q}) = \hat{H}_{\mathrm{el}}(\mathbf{Q}) - \hat{H}_{\mathrm{el}}(\mathbf{Q}_0) . \tag{3.2}$$

The overall temperature-dependent capture coefficient then follows by applying Fermi's golden rule and summing over all vibronic states:

$$C(T) = \frac{2\pi}{\hbar} fgV \sum_{mn} P_m(T) \left| \langle\psi_i \otimes \chi_{im}|\hat{H}_{\mathrm{el-ph}}(\mathbf{Q})|\psi_f \otimes \chi_{fn}\rangle \right|^2 \delta(E_{im} - E_{fn}) \tag{3.3}$$

Here, f denotes the so-called Sommerfeld parameter, which accounts for the electrostatic attraction or repulsion of carriers to an already charged defect. This modifies the carrier wavefunction in the vicinity of the defect and thus can lead to enhancement or suppression of the capture coefficient. While this effect can be treated by ab-initio calculations in principle, it is common to account for it using analytical approximations; see [164, 174, 180] for further details. For a neutral defect, the Sommerfeld parameter is simply $f = 1$. The degeneracy factor g accounts for the possibility of multiple equivalent configurations for the final state. The volume V of the defective supercells normalizes the capture coefficient to become independent of the cell size. Assuming that the nonradiative transition rate is much smaller than the time required to reach thermal equilibrium in the initial state, the probability of finding the system in a certain initial vibrational state m is given by [179]

$$P_m(T) = \frac{1}{Z} e^{-E_{im}/k_\mathrm{B}T} \quad \text{with} \quad Z = \sum_m e^{-E_{im}/k_\mathrm{B}T} \,. \tag{3.4}$$

The Dirac delta in (3.3) ensures the conservation of energy during a vibronic transition from a state with eigenenergy E_{im} to a state with eigenenergy E_{fn}.

While all the essential physics of charge capture is already described within (3.3), further approximations are necessary to make calculating the capture coefficient C computationally tractable. First, the electron-phonon coupling operator (3.2) is expanded to first order around the configuration $\mathbf{Q}_0$ resulting in [174, 179]

$$\Delta\hat{H}_\mathrm{el-ph}(\mathbf{Q}) \approx \nabla_\mathbf{Q}\hat{H}_\mathrm{el}\Big|_{\mathbf{Q}_0} \cdot (\hat{\mathbf{Q}} - \mathbf{Q}_0) = \sum_k (\hat{Q}_k - Q_{k,0}) \left.\frac{\partial \hat{H}_\mathrm{el}}{\partial Q_k}\right|_{Q_k = Q_{k,0}} \,, \tag{3.5}$$

where k runs over *all* normal modes of the system and Q_k is the projection of $\mathbf{Q}$ onto the kth mode. Using this approximation for the coupling operator then yields

$$C(T) = \frac{2\pi}{\hbar} f g V \sum_k |W_{if}^k|^2 \sum_{mn} P_m(T) \left|\langle \chi_{im}|\hat{Q}_k - Q_{k,0}|\chi_{fn}\rangle\right|^2 \delta(E_{im} - E_{fn}) \,, \tag{3.6}$$

where the electron-phonon coupling strength for mode k is given by

$$W_{if}^k = \left\langle \psi_i \middle| \frac{\partial \hat{H}_\mathrm{el}}{\partial Q_k} \middle| \psi_f \right\rangle\Big|_{Q_k = Q_{k,0}} \,. \tag{3.7}$$

Although this formulation has been used for defects in first-principle studies together with a harmonic expansion of the PES [179, 181, 182], the computational costs are still very high due to necessary second-order expansion of the PES, cf. Sec. 2.3.1. Furthermore, evaluating the combinatorial sum in (3.6) is nontrivial and requires special techniques like Monte Carlo sampling [182] or an approach based on a generating function [164, 179, 183, 184].

Effective Mode Model

For these reasons, Alkauskas *et al.* [174] simplified the first-principles treatment of the problem significantly by only considering a single *effective mode* Q along a straight line between the initial and final configuration in mass weighted coordinates as described in Sec. 2.3.1. This approximation is based on the idea that the effective mode introduces a much larger electron-phonon matrix element W_{if}^eff compared to the real phonon modes. As will be shown in Sec. 3.2.3, the effective mode indeed has a dominant matrix element, however, other orthogonal modes can also contribute significantly to the overall capture coefficient. Anyway, in this 1-dimensional approximation the capture coefficient is simply

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

given by

$$C(T) = \frac{2\pi}{\hbar} fgV|W_{if}^{\text{eff}}|^2 G_{if}(0) \tag{3.8}$$

with the *effective electron-phonon matrix element*

$$W_{if}^{\text{eff}} = \left\langle \psi_i \middle| \frac{\partial \hat{H}_{\text{el}}}{\partial Q} \middle| \psi_f \right\rangle \Big|_{Q=Q_0} \tag{3.9}$$

and the vibrational *Herzberg-Teller (HT) lineshape function*

$$G_{if}^{\text{HT}}(\omega) = \sum_{mn} P_m(T) \left| \langle \chi_{im} | \hat{Q} - Q_0 | \chi_{fn} \rangle \right|^2 \delta(E_{im} - E_{fn} - \hbar\omega) \tag{3.10}$$

known from spectroscopy [183, 184]. Note that the lineshape function $G_{if}^{\text{HT}}(\omega)$ is a more general object, which also describes the intensity of absorbing or emitting photons with frequency ω. However, in the nonradiative case, only $G_{if}^{\text{HT}}(0)$ is relevant. As can be seen, the capture coefficient C completely separates into an vibrational part G_{if}^{HT} and an electronic part W_{if}^{eff}. This 1-dimensional formulation allows to represent the charge transfer process in a CC diagram as shown in Fig. 3.1. Furthermore, calculating the vibrational wavefunctions $\chi_{\{im,fn\}}(Q)$ associated with the 1-dimensional PECs $V_{\{i,f\}}(Q)$ can be easily calculated numerically, even for anharmonic potentials, see for instance Fig. 4.5.

In practical computations, the lineshape function $G_{if}^{\text{HT}}(\omega)$ in (3.10) has to be broadened by replacing the Dirac delta with a smooth function of finite width, typically a Gaussian with variance σ_E^2:

$$\delta(E_{im} - E_{fn}) \rightarrow \frac{1}{\sqrt{2\pi\sigma_E^2}} \exp\left(-\frac{1}{2} \frac{(E_{im} - E_{fn})^2}{2\sigma_E^2} \right) \tag{3.11}$$

This broadening not only accounts for the finite phonon lifetime [109] but is also necessary to compensate the artificially sparse density of states in the 1-dimensional representation. Typically, a broadening on the order of the phonon frequency associated with the effective mode is chosen, i.e. $\sigma_E \approx \hbar\Omega$. While this is an empirical parameter, σ_E can be chosen within a rather wide range ($0.5 - 3.0\,\hbar\Omega$) without significantly affecting the final result [164, DWJ9].

Finally, it should be noted that in experimental works the charge trapping efficiency is often quantified using an empirical *capture cross section* (CCS) σ [185–187]. The CCS is related to the more fundamental capture coefficient C through [23, 174, 188]

$$C(T) = v_{\text{th}} \sigma(T)\,, \tag{3.12}$$

where $v_{\text{th}} = \sqrt{k_{\text{B}}T/m^*}$ denotes the thermal velocity of the carrier with effective mass m^*. It should be emphasized however, that the CCS is *not* a physical quantity [189] and thus should only be used to retrieve the experimental capture coefficient C from earlier works.

Harmonic Approximation

When further employing the harmonic approximation, the charge transition is fully characterized by the electron-phonon coupling element W_{if}^{eff}, the energy difference ΔE between the minima of the two parabolas, the ionic displacement ΔQ and the effective phonon frequencies Ω_i and Ω_f of the initial and final state. In this case, the PECs of the two states are simply given by

$$V_i(Q) = c_i Q^2 = \frac{1}{2}\Omega_i^2 Q^2 \tag{3.13a}$$

$$V_f(Q) = c_f (Q - \Delta Q)^2 - \Delta E = \frac{1}{2}\Omega_f^2 (Q - \Delta Q)^2 - \Delta E\,. \tag{3.13b}$$

In the context of device simulations, often the curvatures $c_{\{i,f\}}$ are used to characterize the PECs instead of the phonon frequencies $\Omega_{\{i,f\}}$. A useful dimensionless quantity to assess the validity of the 1-dimensional approximation is the *Huang-Rhys factor* defined by [174, 190]

$$S_{\{i,f\}} = \frac{1}{2\hbar}(\Delta Q)^2 \Omega_{\{i,f\}}\,. \tag{3.14}$$

These factors are a measure for the number of phonons emitted during a transition. A large Huang-Rhys factor $S \gg 1$ indicates strong lattice relaxations, for which the direct path likely dominates the overall transition rate [191]. Another related quantity is the relaxation energy

$$E_{\mathrm{R,\{i,f\}}} = S_{\{i,f\}} \hbar \Omega_{\{i,f\}} = \frac{1}{2}(\Delta Q)^2 \Omega_{\{i,f\}}^2 = |V_{\{i,f\}}(\Delta Q) - V_{\{i,f\}}(0)|\,, \tag{3.15}$$

which describes the amount of energy that is dissipated after a vertical (optical) transition as shown in Fig. 3.2. Typically, the relaxation energies $E_{\mathrm{R,\{i,f\}}}$ are determined from the four points shown as circles in Fig. 3.2. The frequencies $\Omega_{\{i,f\}}$ then follow from (3.15). This has the advantage of probing the PECs on a more global scale instead of only locally via an harmonic expansion around the minimum configuration [DWJ5, DWC3, 192]. For device simulations, also the curvature ratio

$$R = \sqrt{\frac{c_f}{c_i}} = \sqrt{\frac{E_{\mathrm{R}_f}}{E_{\mathrm{R},i}}} \tag{3.16}$$

is often used to characterize the PECs [23, 26, DWJ8, DWJ10], cf. Sec. 5.4.4. The harmonic effective mode approach will be used for a simplified yet physical model for describing charge transitions at the device level as will be discussed in Ch. 5.

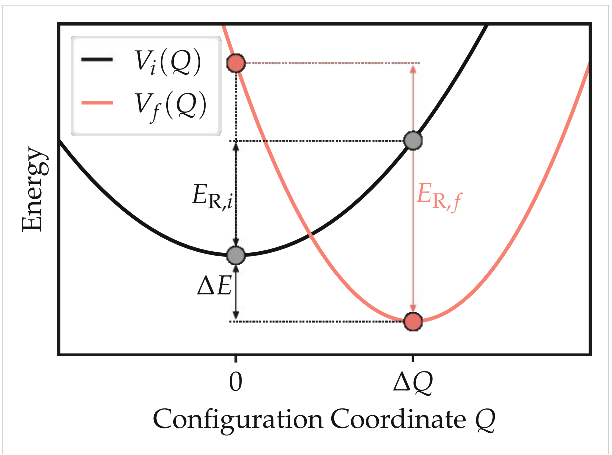


Figure 3.2. Within the harmonic effective mode model, the PECs, and hence also the transition rates, are completely determined by the parameters ΔE, ΔQ and the two relaxation energies $E_{\mathrm{R},\{\mathrm{i,f}\}}$.

3.1.2 Electron-Phonon Coupling in CP2K

The vibrational lineshape function (3.10) can be determined based on the PES of the system and thus only requires the total energy as a function of the ionic coordinates, which can be obtained by every DFT code, including CP2K. However, the electron-phonon coupling elements (3.7) or (3.9) are not provided as output in the CP2K code and hence had to be implemented within the scope of this work.

Using first-order perturbation theory, the electron-phonon matrix element W_{if} can be expressed as [105, 174]

$$W_{if} = \left\langle \psi_i \middle| \frac{\partial \hat{H}_{\mathrm{el}}}{\partial Q} \middle| \psi_f \right\rangle = (\epsilon_f - \epsilon_i) \left\langle \psi_i \middle| \frac{\partial \psi_f}{\partial Q} \right\rangle , \tag{3.17}$$

where $\epsilon_{\{i,f\}}$ denotes the eigenenergies of the initial and final electronic state. Note that for practical computations within DFT, the many-body Hamiltonian $\hat{H}_{\mathrm{el}}$ and corresponding many-body wavefunctions are replaced by their single-particle counterparts in the auxiliary Kohn-Sham system [174], see also Sec. 2.1.3. Contrary to the original definitions in (3.7) and (3.9), this formulation only requires the wavefunctions and their derivatives with respect to the coordinate Q but not the derivative of the electronic Hamiltonian $\hat{H}_{\mathrm{el}}$. Hence, the electron-phonon coupling can be evaluated by using the realspace wavefunctions $\psi_{\{i,f\}}(\mathbf{r}; \mathbf{Q})$ together with a finite-differences scheme, yielding

$$\begin{aligned} W_{if} &\approx \frac{(\epsilon_f - \epsilon_i)}{\delta Q} \Big\langle \psi_i(\mathbf{r}; \mathbf{Q}_0) \Big| \psi_f(\mathbf{r}; \mathbf{Q}_0 + \delta Q\, \mathbf{e}_Q) - \psi_f(\mathbf{r}; \mathbf{Q}_0) \Big\rangle \\ &= \frac{(\epsilon_f - \epsilon_i)}{\delta Q} \langle \psi_i(\mathbf{r}; \mathbf{Q}_0) | \psi_f(\mathbf{r}; \mathbf{Q}_0 + \delta Q\, \mathbf{e}_Q) \rangle , \end{aligned} \tag{3.18}$$

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

with δQ being a small finite displacement and $\mathbf{e}_Q$ is the unit vector in the direction of the phonon mode. Note that the orthogonality relation

$$\langle \psi_m(\mathbf{r}, \mathbf{Q}_0) | \psi_n(\mathbf{r}, \mathbf{Q}_0) \rangle = \delta_{mn} \tag{3.19}$$

was used here. This scheme is convenient and has been frequently used [164, 179], particularly in the 1-dimensional framework, as it only requires the wavefunctions at two slightly displaced ionic configurations and can therefore be applied with any DFT code, regardless of the functionals or basis sets used.

However, in the multidimensional treatment, the normal modes Q_k are not known a priori and must first be evaluated using a finite-difference scheme over the primitive ionic coordinates $\mathbf{R}$, as explained in Sec. 2.3.1. Hence it is beneficial to also calculate the electron-phonon matrix elements in this basis, i.e.

$$\widetilde{W}_{if}^{n} = \left\langle \psi_i \middle| \frac{\partial \hat{H}_\mathrm{el}}{\partial R_n} \middle| \psi_f \right\rangle \Big|_{R_n = R_{n,0}} \tag{3.20}$$

to avoid running two separate finite-difference calculations for the normal modes and the electron-phonon coupling. The actual matrix elements associated with the normal modes then simply follow from a linear transformation by applying the chain rule:

$$W_{if}^{k} = \sum_n \left\langle \psi_i \middle| \frac{\partial \hat{H}_\mathrm{el}}{\partial R_n} \middle| \psi_f \right\rangle \frac{\partial Q_k}{\partial R_n} = \sum_n \widetilde{W}_{if}^{n} \frac{\partial R_n}{\partial Q_k} \tag{3.21}$$

Since the wavefunctions are only determined up to an arbitrary phase factor, only the magnitude of $\widetilde{W}_{if}^{n}$ can be obtained when using (3.18). Normally this is not a problem since the capture coefficient (3.6) only depends on the square modulus of W_{if}^{k}. However, the relative phases between the couplings $\widetilde{W}_{if}^{n}$ in the linear transformation (3.21) are important for obtaining the correct W_{if}^{k}. Hence, the original form (3.7) or rather (3.20) is used, since the Hamiltonian is observable and thus there is no ambiguity in the relative phases[1]. When applying a finite-difference scheme, the coupling elements can be approximated by

$$\begin{aligned}\widetilde{W}_{if}^{n} &\approx \frac{1}{\delta R} \left\langle \psi_i(\mathbf{r}; \mathbf{R}_0) \middle| \hat{H}_\mathrm{el}(\mathbf{R}_0 + \delta R\, \mathbf{e}_R) - \hat{H}_\mathrm{el}(\mathbf{R}_0) \middle| \psi_f(\mathbf{r}; \mathbf{R}_0) \right\rangle \\ &= \frac{1}{\delta R} \left\langle \psi_i(\mathbf{r}; \mathbf{R}_0) \middle| \hat{H}_\mathrm{el}(\mathbf{R}_0 + \delta R\, \mathbf{e}_R) \middle| \psi_f(\mathbf{r}; \mathbf{R}_0) \right\rangle ,\end{aligned} \tag{3.22}$$

when using the orthogonality relation

$$\langle \psi_m(\mathbf{r}; \mathbf{R}_0) | \hat{H}_\mathrm{el}(\mathbf{R}_0) | \psi_n(\mathbf{r}; \mathbf{R}_0) \rangle = \epsilon_m \delta_{mn} \, . \tag{3.23}$$

---

[1] There is still an ambiguity in the global phase. However this is physically irrelevant.

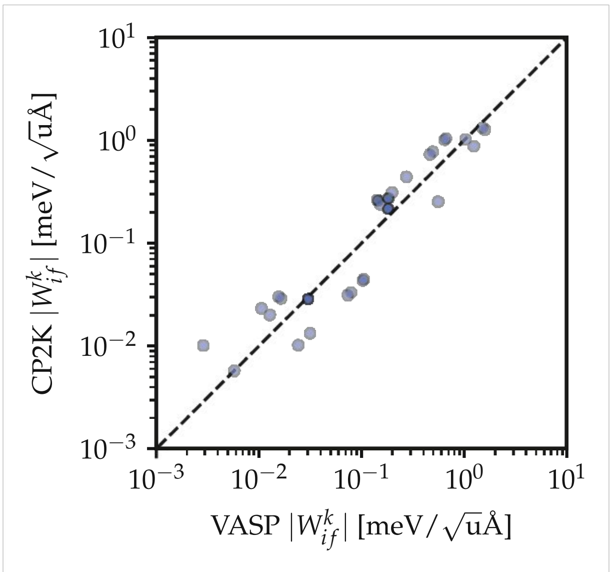


Figure 3.3. Comparison of electron-phonon couplings along different phonon modes from the implementation in CP2K used in this work against an independent reference implementation in VASP. The calculations where done for the electron capture process at a GaP:$\mathrm{Zn_{Ga}O_P}$ donor-acceptor pair, see also Sec. 3.2.2. As can be seen, the two results of the two codes are generally in good agreement over several orders of magnitude considering that both codes use different basis sets and pseudopotentials.

The electron-phonon coupling element can then be further expressed in the eigenbasis of the Hamiltonian:

$$\widetilde{W}^n_{if} = \frac{1}{\delta R} \sum_n \langle \psi_i(\mathbf{r};\mathbf{R}_0) | \psi_n(\mathbf{r};\mathbf{R}_0 + \delta R\, \mathbf{e}_R) \rangle\, \epsilon_n\, \langle \psi_n(\mathbf{R}_0 + \delta R\, \mathbf{e}_R) | \psi_f(\mathbf{r};\mathbf{R}_0) \rangle \ . \tag{3.24}$$

Within the Gaussian basis set used in CP2K, this expression becomes

$$\widetilde{W}^n_{if} = \frac{1}{\delta R} \mathbf{c}_i^{\mathrm{T}} \underline{\mathbf{T}}\,\underline{\mathbf{C}}^{\mathrm{T}} \underline{\mathbf{E}}\,\underline{\mathbf{C}}\,\underline{\mathbf{T}}^{\mathrm{T}} \mathbf{c}_f , \tag{3.25}$$

where $\mathbf{c}_{\{i,f\}}$ are the coefficients of the initial and final state wavefunctions at $\mathbf{R}_0$. The diagonal matrix $\underline{\mathbf{E}} = \mathrm{diag}(\epsilon_k)$ contains all eigenenergies and the matrix $\underline{\mathbf{C}}$ collects the coefficients of the corresponding eigenstates at position $\mathbf{R}_0 + \delta R\, \mathbf{e}_n$. The **transfer matrix $\underline{\mathbf{T}}$** with coefficients

$$T_{mn} = \langle \phi_m(\mathbf{r};\mathbf{R}_0) | \phi_n(\mathbf{r};\mathbf{R}_0 + \delta R\, \mathbf{e}_n) \rangle \tag{3.26}$$

accounts for the shift of the atom-centered basis functions $\{|\phi_m\rangle\}$ between the two displaced configurations.

To verify the implementation and the applicability of the Gaussian basis sets in CP2K for calculating electron-phonon couplings, the results from this implementation were compared against a reference implementation [164] in the planewave code VASP. The electron capture process of the $\mathrm{Zn_{Ga}O_P}$ donor-acceptor pair defect in GaP [186] (cf. Sec. 3.2.2) was used as test system. For this comparison, the electron phonon coupling for a few randomly selected phonon modes was evaluated with both codes. As can be seen in

Fig. 3.3, there is good agreement between the two codes over several orders of magnitude for the electron-phonon coupling strength, considering that both codes use different basis sets and pseudopotentials.

3.2 Extension of the Effective Mode Model

This section is based on work currently prepared for publication [DWJ11].

Due to its simplicity, the 1-dimensional effective model as discussed in Sec. 3.1.1 is frequently used to assess the charge trapping efficiency of defects in different materials [193–196]. Here, an extension to this model is derived which takes into account electron-phonon couplings arising from modes orthogonal to the effective mode. It will be shown that such additional modes not considered in the 1-dimensional model can significantly enhance the charge capture rates especially at low temperatures.

3.2.1 Constrained Phonon Modes

The basic assumption in the 1-dimensional model is that the effective mode, i.e. the direct path between two different charge configurations, provides much stronger electron-phonon coupling than any other individual phonon mode. While other multidimensional approaches to nonradiative transitions abandon the concept of an effective mode [179, 183, 184] in favor of treating all phonon modes equally, the presented approach retains the effective mode because it offers a distinct direction in the problem and indeed introduces a significant electron-phonon coupling. However, retaining the effective mode requires orthogonalizing the real system modes to the effective mode to avoid artificial couplings.

A similar problem arises in other system-bath models, where the interaction of a mode of interest with a set of environmental modes is studied. For instance, the finite lifetimes of stretching and bending modes in a Si−H bond due to couplings with its environment have been investigated in [DWJ2, 197]. Following the derivation given in [DWJ2], a set of *constrained phonon modes* orthogonal to the system mode is constructed by first defining $3N - 1$ random unit vectors $\mathbf{u}_i \in \mathbb{R}^{3N}$. These vectors are then orthogonalized to the unit vector $\mathbf{e}_Q$ along the effective mode using the Gram-Schmidt procedure [198]. These orthogonalized vectors can be collected in the projection matrix

$$\underline{\mathbf{U}} = [\mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_{3N}]^T \in \mathbb{R}^{(3N-1)\times 3N}\,, \tag{3.27}$$

which projects into the subspace orthogonal to $\mathbf{e}_Q$. Using this matrix, the vibrational Hamiltonian can also be projected into this subspace according to

$$\underline{\mathbf{H}}' = \underline{\mathbf{U}}\,\underline{\mathbf{H}}\,\underline{\mathbf{U}}^{\mathrm{T}}\,. \tag{3.28}$$

ΔE	ΔQ	$\hbar\Omega_i$	$\hbar\Omega_f$	W_{if}^{eff}
$0.26\,\text{eV}$	$4.27\sqrt{\text{u}}\text{Å}$	$9.23\,\text{meV}$	$9.27\,\text{meV}$	$2.87\sqrt{\text{u}}\text{Å}$

Table 3.1. Transition parameters for electron trapping at the $\text{Zn}_{\text{Ga}}\text{O}_{\text{P}}$ center in GaP. The corresponding configuration coordinate diagram is shown in Fig. 3.4 (right). The obtained parameters are in good agreement with earlier investigations [201, 202].

Diagonalizating the reduced Hamiltonian $\underline{\mathbf{H}}'$ then yields a set of $3N-1$ eigenvectors $\mathbf{e}_i'$ which can be transformed back to the original space yielding the constrained phonon modes

$$\mathbf{e}_{q_i} = \underline{\mathbf{U}}^{\text{T}}\mathbf{e}_i' \,. \tag{3.29}$$

The new modes $\mathbf{e}_{q_i}$ are orthogonal to each other but are also orthogonal to the effective mode $\mathbf{e}_{\text{Q}}$.

Moving forward, this set of constrained modes with frequencies ω_i will be used for describing the environmental bath modes neglected in the 1-dimensional picture. Contrary to the description with physical modes, the entire displacement ΔQ between both defect states happens along the effective mode. Hence, for all other environmental modes i the displacement is always $\Delta q_i = 0$. This considerably simplifies the evaluation of the vibrational sum in (3.6) and will be utilized in Sec. 3.2.3.

3.2.2 GaP:$\text{Zn}_{\text{Ga}}\text{O}_{\text{P}}$ as a Test Case

In order for this model to be viable, the constrained modes need to be similar to the real phonon modes. This will be demonstrated in the following using the well-studied donor-acceptor pair defect $\text{Zn}_{\text{Ga}}\text{O}_{\text{P}}$ in gallium phosphide (GaP)[179, 186]. In this study, the PBE0_TC_LRC hybrid functional [67] is used with the default mixing parameter of $\alpha = 0.25$, which results in a bandgap of $E_{\text{g}} = 2.31\,\text{eV}$ in good agreement with the experimental value of 2.26 eV at room temperature [199]. The defect was simulated in a $3 \times 3 \times 3$ cubic supercell containing 216 atoms.

The $\text{Zn}_{\text{Ga}}\text{O}_{\text{P}}$ defect is a substitutional defect in which one Ga atom is replaced by Zn acting as a shallow acceptor. However, this is compensated by an O atom substituting a neighboring P, which acts as a deep donor in GaP with an ionization energy of approximately 900 meV [200]. The resulting defect complex has an empty state close to the GaP CBM and thus acts as an acceptor-like trap, i.e. an electron trap. The neutral and singly negative charge state of the defect together with the corresponding configuration coordinate diagram is depicted in Fig. 3.4. The obtained transition parameters within the harmonic approximation are given in Tab. 3.1 and are in good agreement with the findings of earlier works [201, 202]. In a classical picture, the transition is almost barrierless, leading to an efficient electron trapping process consistent with experiments [186].

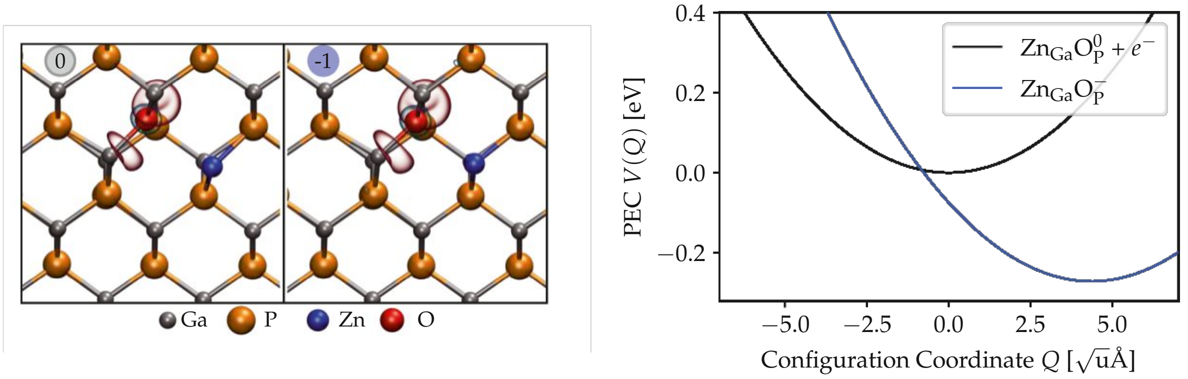


Figure 3.4. Electron trapping at the $\mathrm{Zn_{Ga}O_P}$ donor-acceptor pair defect in GaP. **Left:** This defect introduces a localized state in the bandgap centered at the O site. Upon electron trapping from the CBM, the neighboring Zn moves closer towards the O atom. **Right:** Corresponding configuration coordinate diagram for this transition along the effective mode. In the initial state $\mathrm{Zn_{Ga}O_P}+e^-$, the defect is neutral while a free electron is located at the bottom of the conduction band. The very small classical barrier is the cause of this defect being a very efficient electron trap.

In order to assess the accuracy of the effective mode approximation in this example, the coordinate transformation described in the previous section is applied to construct a set of constrained phonon modes orthogonal to the effective mode. As shown in Fig. 3.5 (left), the constrained modes produce a virtually identical phonon spectrum compared to the real phonon modes. Both phonon spectra lie within the envelope given by the bulk phonon modes, which were calculated in a pristine supercell using the PHONOPY software package [203, 204]. Furthermore, the projections onto the real modes shown in Fig. 3.5 (right) reveal that almost all constrained modes can be unambiguously associated with a real mode, indicating that only small additions have to be made to keep the modes orthogonal to the effective mode. These findings show that the constrained mode approach is well suited as an alternative to the normal mode system for describing nonradiative transitions.

Since the bath modes $\mathbf{e}_{q_i}$ are now decoupled from the effective mode, the electron-phonon coupling induced by these modes can now be studied separately. The bath modes are further distinguished in localized and delocalized modes based on their inverse participation ratio (IPR) [205, 206]

$$\mathrm{IPR} = \frac{\sum_{k=1}^{N} |\mathbf{u}_{ik}|^4}{\left(\sum_{k=1}^{N} |\mathbf{u}_{ik}|^2\right)^2} . \tag{3.30}$$

Here, $\mathbf{u}_{ik}$ denotes the vector associated with movement of atom k when mode i is excited. While the IPR is more commonly used to quantify the localization of electronic states, see for instance Fig. 4.15, it can also be applied to distinguish between a phonon mode localized around a defect site ($\mathrm{IPR} \gg 1/N$) and a delocalized mode which is not associated with the defect ($\mathrm{IPR} \approx 1/N$). The coupling strength of each bath mode as a function

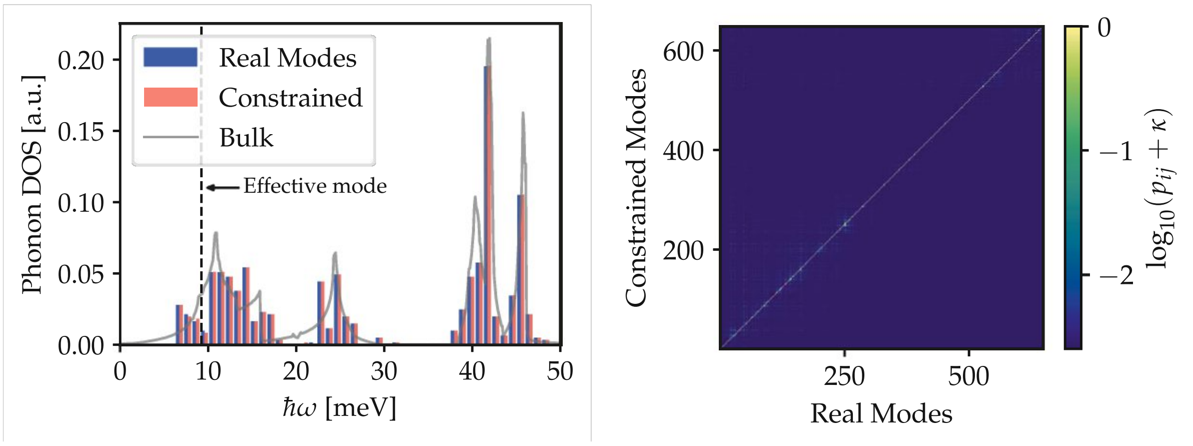


Figure 3.5. Constrained mode approach applied to the $\mathrm{Zn_{Ga}O_P}$ defect in GaP. **Left:** The phonon density of states (DOS) within the constrained mode system is almost indistinguishable from the real phonon DOS. The bulk phonon DOS is also shown for comparison. As can be seen, the DOS at the Γ point of the defective supercell already approaches the bulk DOS, indicating that the chosen supercell is large enough for approximating the phonon bath in the bulk material. **Right:** Projecting the constrained modes back onto the real modes of the system shows that most constrained modes are very close to a physical mode, with only small additions required to maintain orthogonality with the effective mode. Note that instead of the projections p_{ij}, the quantity $\log_{10}(p_{ij}+\kappa)$ with $\kappa = 2.5\times10^{-3}$ is shown here in order to improve the visibility of the relatively small off-diagonal projections.

of the IPR is illustrated in Fig. 3.6 (left). As can be seen, the effective mode is easily distinguishable as the only localized mode with strong coupling. Furthermore, based on this analysis, the bath modes can be divided in 3 categories:

(a) Most bath modes are delocalized and introduce little to no coupling between the initial and final defect states. Hence they are not of great interest for determining the transition rates.

(b) Besides the effective mode there also exist other localized defect modes. However, these modes have to be orthogonal to the effective mode and therefore cannot introduce an efficient coupling.

(c) Surprisingly, there are also optical delocalized bulk modes that can strongly couple to the defect. It is suspected that these modes generate dipole or quadrupole moments, which can create a macroscopic electric field. This field could then couple to the defect, despite the delocalized nature of the phonon mode causing it. While such long-range electron-phonon interactions are known to cause scattering events of free carriers, leading to a mobility reduction [207, 208], to the best of the authors' knowledge, they have not yet been considered in the context of charge trapping.

Although the origin of the nonlocal coupling of modes in category (c) requires further research beyond the scope of this work, it is clear that their influence should be included in the formalism for nonradiative transitions.

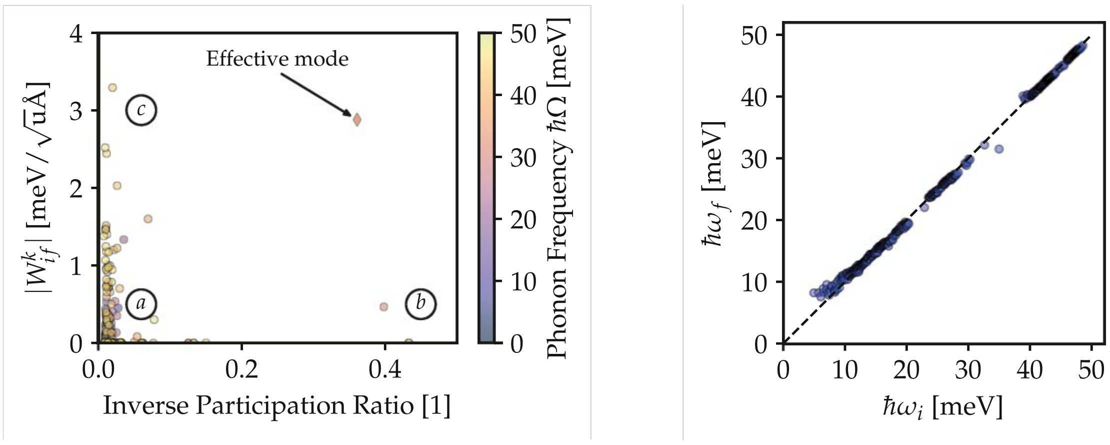


Figure 3.6. Left: Coupling strength of the various constrained bath modes as a function of their inverse participation ratio (IPR) given in (3.30). The IPR is a measure for the localization of the mode at the defect site. The localized effective mode indeed has a strong coupling strength, however, there are also delocalized optical modes (c) with similarly strong coupling. **Right:** Correlation between the frequencies of the bath modes in the initial and final defect state. The frequencies of the bath modes are mostly unaffected by the transition.

Another important observation can be made in Fig. 3.6 (right), where the correlation between the phonon frequencies in the initial and final state is shown. The frequencies of most bath modes are approximately the same in both charge states. This result will be used in the next section to derive a simple analytical expression

3.2.3 Environmental Coupling

With the characteristics of the bath modes established, an analytical expression is now derived to account for their influence on the total capture coefficient. For simplicity, only the effective mode and a single bath mode are considered. However, the derivation below can be easily generalized to include an arbitrary number of bath modes. In what follows, the coordinate along the effective mode and its corresponding vibrational wavefunctions are denoted as Q and $\chi_{\{im,fn\}}(Q)$, respectively, while the coordinate of the bath mode with frequency ω and its wavefunctions are represented as q and $\eta_{\{k,l\}}(q)$. Note that for the bath mode $\Delta q = 0$ and $\omega_i = \omega_f = \omega$ is assumed, therefore the vibrational wavefunctions in the initial and final state are equal, requiring only to account for different occupation numbers k, l.

Although the constrained mode system introduced in Sec. 3.2.1 does not exactly diagonalize the vibrational Hamiltonian, the results presented in Fig. 3.5 show that this auxiliary set of phonon modes is sufficiently close to the true set of phonon modes. Hence, the initial and final vibrational wavefunctions are assumed to be separable in the form

$$\chi_{ikm}(Q,q) = \chi_{im}(Q)\eta_k(q) \quad \text{and} \quad \chi_{fln}(Q,q) = \chi_{fn}(Q)\eta_l(q)\,. \tag{3.31}$$

With these preparations, the capture coefficient can be written as

$$C = \frac{2\pi}{\hbar} f g V \sum_{mnkl} P_m p_k |\left\langle \chi_{im} \otimes \eta_k \middle| W_{if}^{\text{eff}} \hat{Q} + w_{if} \hat{q} \middle| \eta_l \otimes \chi_{fn} \right\rangle|^2 \\ \times \delta\big(E_{im} - E_{fn} + \hbar\omega(k-l)\big)\,. \tag{3.32}$$

Here, P_m and p_k denote the thermal occupancies (see (3.4)) of the effective mode and the bath mode respectively. Evaluating the matrix element within the sum leads to

$$C = \frac{2\pi}{\hbar} f g V \sum_{mnkl} P_m p_k |W_{if}^{\text{eff}} \langle \chi_{im} | \hat{Q} | \chi_{fn} \rangle \langle \eta_k | \eta_l \rangle + w_{if} \langle \chi_{im} | \chi_{fn} \rangle \langle \eta_k | \hat{q} | \eta_l \rangle|^2 \\ \times \delta\big(E_{im} - E_{fn} + \hbar\omega(k-l)\big)\,. \tag{3.33}$$

Since the harmonic oscillator of the bath mode is undisplaced and has the same frequency in both states, its matrix elements are simply given by the well-known expressions [209]

$$\langle \eta_k | \eta_l \rangle = \delta_{kl} \quad \text{and} \quad \langle \eta_k | \hat{q} | \eta_l \rangle = \sqrt{\frac{\hbar}{2\omega}} \Big(\sqrt{k}\, \delta_{k,l+1} + \sqrt{k+1}\, \delta_{k,l-1} \Big)\,. \tag{3.34}$$

Since all Kronecker deltas here are mutually exclusive, plugging these matrix elements into (3.33) and evaluating the square yields

$$\begin{aligned} C = {} & \frac{2\pi}{\hbar} f g V |W_{if}^{\text{eff}}|^2 \sum_{mn} P_m | \langle \chi_{im} | \hat{Q} | \chi_{fn} \rangle |^2 \delta\big(E_{im} - E_{fn}\big) \sum_{k=0}^{\infty} p_k \\ & + \frac{2\pi}{\hbar} f g V |w_{if}|^2 | \sum_{mn} P_m | \langle \chi_{im} | \chi_{fn} \rangle |^2 \delta\big(E_{im} - E_{fn} + \hbar\omega\big) \frac{\hbar}{2\omega} \sum_{k=0}^{\infty} k\, p_k \\ & + \frac{2\pi}{\hbar} f g V |w_{if}|^2 | \sum_{mn} P_m | \langle \chi_{im} | \chi_{fn} \rangle |^2 \delta\big(E_{im} - E_{fn} - \hbar\omega\big) \frac{\hbar}{2\omega} \sum_{k=0}^{\infty} (k+1)\, p_k\,. \end{aligned} \tag{3.35}$$

With the expected number of phonons in the harmonic bath mode given by

$$\langle n \rangle = \frac{1}{e^{\hbar\omega/k_\text{B}T} - 1}\,, \tag{3.36}$$

the occurring sums over k in (3.35) can be expressed as

$$\sum_{k=0}^{\infty} p_k = 1, \quad \sum_{k=0}^{\infty} k\, p_k = \langle n \rangle \quad \text{and} \quad \sum_{k=0}^{\infty} (k+1)\, p_k = 1 + \langle n \rangle = \langle n \rangle\, e^{\hbar\omega/k_\text{B}T}\,. \tag{3.37}$$

Hence, the first term in (3.35) reduces to the previously established expression (3.10) for the capture coefficient. The remaining two terms can be expressed with the *Franck-Condon* lineshape function [210, 211]

$$G_{if}^{\text{FC}}(\omega) = \sum_{mn} w_m(T)\, |\langle \chi_{im} | \chi_{fn} \rangle|^2\, \delta(E_{im} - E_{fn} - \hbar\omega) \tag{3.38}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

resulting in

$$C = \frac{2\pi}{\hbar} fgV \left(|W_{if}^{\mathrm{eff}}|^2 G_{if}^{\mathrm{HT}}(0) + |w_{if}|^2 \frac{\hbar \langle n \rangle}{2\omega} \left(G_{if}^{\mathrm{FC}}(+\hbar\omega) + e^{\hbar\omega/k_{\mathrm{B}}T} G_{if}^{\mathrm{FC}}(-\hbar\omega) \right) \right). \tag{3.39}$$

For multiple bath modes, this expression can be easily generalized to

$$\begin{aligned} C = {} & \frac{2\pi}{\hbar} fgV |W_{if}^{\mathrm{eff}}|^2 G_{if}^{\mathrm{HT}}(0) \\ & + \pi fgV \Big(\sum_k |w_{if}^k|^2 \frac{1}{\omega_k} \big(\langle n_k \rangle G_{if}^{\mathrm{FC}}(-\hbar\omega_k) + (\langle n_k \rangle + 1) G_{if}^{\mathrm{FC}}(\hbar\omega_k) \Big). \end{aligned} \tag{3.40}$$

While the derivation of (3.39) or (3.40) is somewhat tedious, its physical interpretation is quite intriguing. This result shows that modes orthogonal to the effective mode can enhance charge trapping rates as long as they have a nonzero coupling strength w_{if}^k, even though they are not associated with any displacement Δq_k. The two additional terms in (3.40) correspond to the exchange (emission or absorption) of a single phonon with energy $\hbar\omega_k$ between the bath and the effective mode system. Notably, only one phonon can be transferred due to the selection rules of the harmonic oscillator, as given in (3.34).

In the convenient 1-dimensional configuration coordinate diagram, this leads to a broadening of the defect PECs on the order of the phonon frequencies in the bath, as illustrated in Fig. 3.7 (left). Each phonon emission into or absorption from the bath shifts the 1-dimensional PEC of the initial state[2] by $-\hbar\omega_k$ or $+\hbar\omega_k$, respectively, resulting in a family of initial state PECs, each contributing to the total capture coefficient. As shown in Fig. 3.7 (right), these additional contributions can dominate over the effective mode. However, this is not always the case, as indicated in Fig. 3.8. Here, the same parameters for the 1-dimensional effective mode (see Tab. 3.1) and the phonon bath were used as in Fig. 3.7, except for a reduced displacement of $\Delta Q = 2.5\,\sqrt{\mathrm{u}}\text{Å}$. In this case, the configuration coordinate diagram reveals a significant classical barrier, causing PECs with lower barriers to dominate at lower temperatures. At higher temperatures, however, the HT contribution from the effective mode becomes dominant.

Both the FC as well as the HT lineshape function have been used frequently to describe charge transfer reactions. In chemistry and spectroscopy related fields for instance, charge transfer is commonly described with the FC lineshape function, resulting in a capture rate of the form [179, 212, 213]

$$k_{if} = \frac{2\pi}{\hbar} |V_c|^2 G_{if}^{\mathrm{FC}}(0) = \frac{2\pi}{\hbar} |V_c|^2 \sum_{mn} P_m(T) \left| \langle \chi_{im} | \chi_{fn} \rangle \right|^2 \delta(E_{im} - E_{fn}). \tag{3.41}$$

It is assumed here that the coupling between the two electronic states is independent of the phonon coordinates $\mathbf{Q}$ and is therefore expressed through a constant coupling constant V_c. Such coupling naturally arises in radiative transitions, where the interaction is primarily due to the dipole moment μ of the system interacting with the electric field associated with the

---

[2] Since only energy differences are physically significant, the energy shift could also be attributed to the final state with a changed sign.

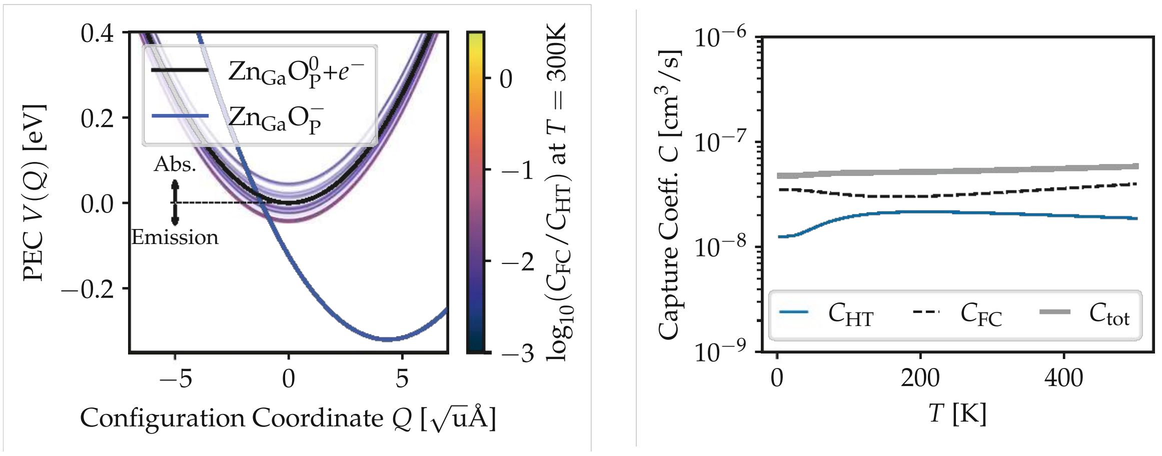


Figure 3.7. Effect of coupling to bath modes on the configuration coordinate diagram and the capture coefficient for electron capture at the $\mathrm{Zn_{Ga}O_P}$ defect in GaP. **Left:** Interaction with the environmental bath allows for phonon absorption or emission, shifting the potential energy curve (PEC) of the inital state by $\pm\hbar\omega_k$. This results in a family of different PECs that must be added together to obtain the total capture coefficient. The color of the inital PECs represents the relative Franck-Condon (FC) contribution to the overall capture coefficient, compared to the Herzberg-Teller (HT) contribution of the effective mode. **Right:** The total capture coefficient, as calculated from (3.40), shows that the sum of all FC contributions can dominate over the HT contribution. However, this is not generally true as demonstrated in Fig. 3.8.

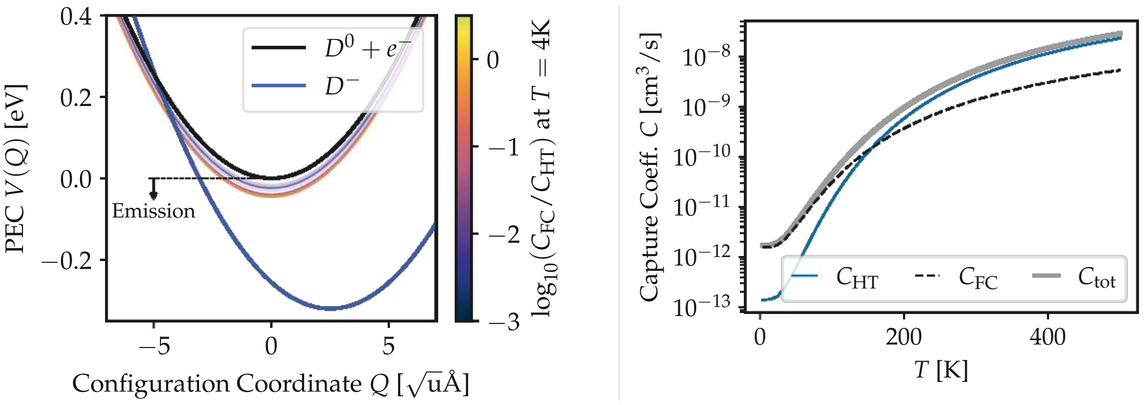


Figure 3.8. Similar to Fig. 3.7, but with a modified displacement of $\Delta Q = 2.5\,\sqrt{\mathrm{u}}$Å. Due to the smaller displacement compared to Fig. 3.7, a classical barrier forms between the two configurations. As a result, transitions with lower barriers dominate at lower temperatures. However, the effective mode contribution can still dominate at higher temperatures. Additionally, the configuration coordinate (CC) diagram shows that at lower temperatures only phonon emission processes, which lower the initial state PEC, contribute to the capture coefficient.

radiation. Since μ is generally nonzero at $\mathbf{Q} = 0$, the leading term in an expansion similar to (3.5) is a constant, resulting in the FC lineshape function as the primary contribution, with the HT lineshape function providing only minor corrections [214]. In nonradiative transitions, this type of coupling can arise because the diabatic (localized) electronic states are not exact eigenstates of the system. As will be discussed in the next section, another reason for assuming a constant coupling constant is that, when a classical barrier exists in the configuration coordinate (CC) diagram, the lineshape function contributes significantly only near the classical crossing point. Therefore, the electron-phonon coupling can be assumed constant around this crossing point.

To the contrary, in the formulation by Alkauskas et al. [174], there is no coupling at $\mathbf{Q} = \mathbf{0}$, resulting in a linear dependence on $\mathbf{Q}$ to first order. This leads to the HT lineshape function providing the dominant contribution. However, as shown here, both lineshape functions can have similar contributions even in nonradiative transitions when considering coupling to the bath modes. Hence, for the sake of simplicity, the FC lineshape function will be employed for the device level modeling in Ch. 5.

3.3 Classical Limit

This section is partially based on works previously published in [DWT1, DWC3]

Although the theory of nonradiative charge transfer is fully quantum mechanical, implicitly accounting for nuclear quantum effects such as tunneling, it is also valuable to explore the classical limit of NMP theory. In particular, for practical implementation in device simulators — where the charge states of thousands of defects must be calculated dynamically — the classical description of charge transfer is highly advantageous due to its significantly lower computational cost compared to the full quantum mechanical framework discussed thus far.

3.3.1 Classical Lineshape Function

A first step towards a classical approximation for the nonradiative transition rate is to replace the discrete energy levels E_{im} and E_{fn} with a continuum. This is justified, when the classical barrier ε_{if} as shown in Fig. 3.9 is much larger than the energy spacing ΔE between subsequent vibrational states. In this case, the *overlap density* $g_{if}^{\mathrm{FC}}(E)$ can be introduced as a smooth function that represents the overlaps of the lineshape function. It is defined at the discrete eigenenergies E_{im} as

$$g_{if}^{\mathrm{FC}}(E_{im}) = \sum_{n} |\langle\chi_{im}|\chi_{fn}\rangle|^2 \delta(E_{im} - E_{fn}) \tag{3.42}$$

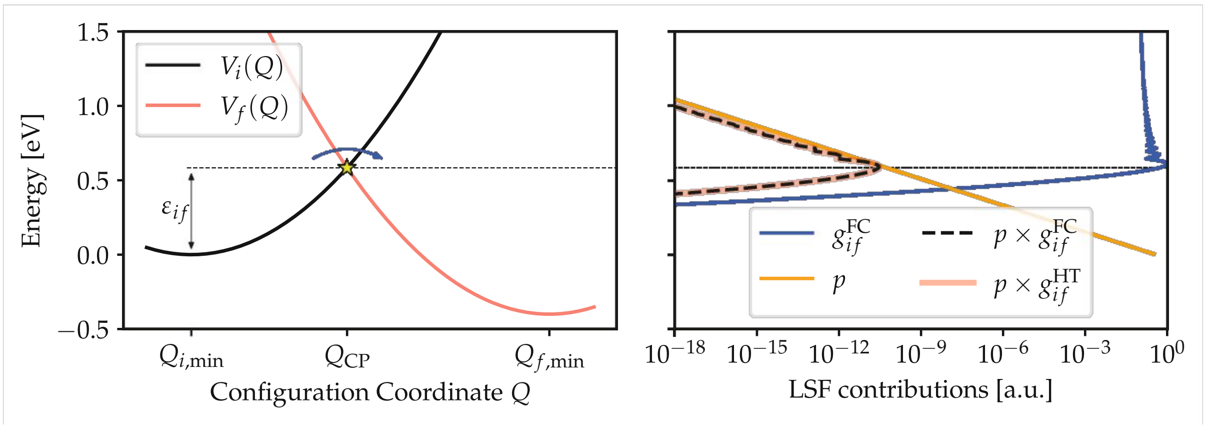


Figure 3.9. Classical limit for NMP charge transitions. **Left:** Schematic configuration coordinate diagram with a classical barrier ε_{if}. **Right:** The corresponding overlap density $g_{if}(E)$ rapidly decays below the crossing point due to the exponential decay of the vibrational wavefunctions in the classically forbidden regions. Above the crossing point, however, the overlap density fluctuates while remaining orders of magnitude larger than in the former case. Combined with the exponentially decaying Boltzmann factor p at higher energies, the product $g_{if} \times p$ sharply peaks at the crossing point energy, justifying the classical approximation of the lineshape function as a δ distribution at the classical crossing point. This observation holds for both the FC and HT lineshape functions. (adapted from [DWT1])

and is interpolated for other values of E. Note that the Dirac δ distribution is again replaced by a Gaussian, as in (3.11). With this overlap density, the FC-LSF (3.38) can be approximated by an integral over the configuration coordinate [DWT1, 215]:

$$G_{if}^{\mathrm{FC}}(0) \approx \frac{1}{\mathcal{Z}} \int_{-\infty}^{\infty} g_{if}^{\mathrm{FC}}(V_i(Q)) e^{-V_i(Q)/k_{\mathrm{B}}T} \mathrm{d}Q \quad \text{with} \quad \mathcal{Z} = \int e^{-V_i(Q)/k_{\mathrm{B}}T} \mathrm{d}Q\,. \tag{3.43}$$

As shown in Fig. 3.9 (right), the overlap density rapidly decays below the energy of the crossing point (CP) at which $V_i(Q_{\mathrm{CP}} = V_f(Q_{\mathrm{CP}}))$ holds, since the vibrational wavefunctions decay exponentially into the classically forbidden regions. Above the CP however, the overlap density fluctuates but stays roughly on the same order of magnitude. On the other hand, the Boltzmann factor $p = \exp(-E/k_{\mathrm{B}}T)/\mathcal{Z}$ falls off exponentially with higher energies. As a result, the integral peaks sharply in the vicinity of the CP. This holds equally true for the FC and HT lineshape functions. It is thus justified to assume that all contributions to the lineshape function stem entirely from the crossing point. Mathematically, this is equivalent to replacing the overlap density with a Dirac δ at the CP, yielding [215]

$$\begin{aligned} G_{if}^{\mathrm{FC}}(0) \approx& \frac{1}{\mathcal{Z}} \int_{-\infty}^{\infty} \delta(V_i(Q) - V_f(Q)) e^{-V_i(Q)/k_{\mathrm{B}}T} \mathrm{d}Q \\ =& \frac{1}{\mathcal{Z}} \frac{1}{|V_i'(Q_{\mathrm{CP}}) - V_f'(Q_{\mathrm{CP}})|} e^{-V_i(Q_{\mathrm{CP}})/k_{\mathrm{B}}T} \\ =& \frac{1}{\mathcal{Z}} \frac{1}{|V_i'(Q_{\mathrm{CP}}) - V_f'(Q_{\mathrm{CP}})|} e^{-\varepsilon_{if}/k_{\mathrm{B}}T}\,. \end{aligned} \tag{3.44}$$

The magnitude of the LSF therefore depends exponentially on the classical barrier ε_{if}. Assuming parabolic PECs, i.e.

$$V_i(Q) = \frac{1}{2}\Omega_i{}^2 Q^2 \quad \text{and} \quad V_i(Q) = \frac{1}{2}\Omega_f^2 (Q - \Delta Q)^2 \tag{3.45}$$

the partition function can be analytically evaluated to be

$$\mathcal{Z} = \int_{-\infty}^{\infty} e^{-V_i(Q)/k_\mathrm{B}T} \mathrm{d}Q = \int_{-\infty}^{\infty} \exp\left(\frac{1}{2k_\mathrm{B}T}\Omega_i^2 Q^2\right) \mathrm{d}Q = \sqrt{\frac{2\pi k_\mathrm{B}T}{\Omega_i^2}}\,. \tag{3.46}$$

Further assuming the curvature ratio to be $R = 1$, i.e. $\Omega_i = \Omega_f = \Omega$, then results in

$$G_{if}^\mathrm{FC}(0) \approx \sqrt{\frac{1}{2\pi k_\mathrm{B}T\Omega^2\Delta Q^2}} e^{-\varepsilon_{if}/k_\mathrm{B}T} = \sqrt{\frac{1}{4\pi k_\mathrm{B}TE_\mathrm{R}}} e^{-\varepsilon_{if}/k_\mathrm{B}T}\,. \tag{3.47}$$

Using this approximation for the lineshape function in (3.41) finally yields the classical rate equation

$$k_{if} = \frac{2\pi}{\hbar}|V_c|^2 \sqrt{\frac{1}{4\pi k_\mathrm{B}TE_\mathrm{R}}} e^{-\varepsilon_{if}/k_\mathrm{B}T}\,, \tag{3.48}$$

which is identical to the transition rate within the famous *Marcus theory* [216, 217] describing classical charge transfer in chemical reactions. Due to its simplicity, such a classical description will be used in Ch. 5 to describe the charging and discharging of thousands of individual defects within a transistor, allowing for modeling the charge trapping related instabilities on the device level.

3.3.2 Minimum Energy Crossing Point

As evident from the derivation above, the charge transition rate depends exponentially on the classical barrier defined by the crossing point Q_CP. In the 1-dimensional effective mode model, only a discrete crossing point exists[3], whereas in the multidimensional case, the potential energy surfaces (PESs) intersect along a continuous seam, as illustrated in Fig. 3.10. Similar to the adiabatic transitions discussed in Sec. 2.3.2, the minimum energy path (MEP) with the lowest possible barrier between the two configurations will dominate the total transition rate due to the exponential dependence. However, the corresponding minimum energy crossing point (MECP) generally does not lie along the effective mode. Furthermore, as demonstrated in Sec. 3.2.3, significant electron-phonon coupling can also occur off the effective mode. Therefore, it can be assumed that also some coupling occurs at the MECP, making it the dominant factor in the transition rate.

---

[3] Depending on the configuration coordinate (CC) diagram, also zero or two crossing points may occur. The former case cannot be treated within the classical limit, while the latter results in an additional term in (3.44); further details are provided in [23, 215].

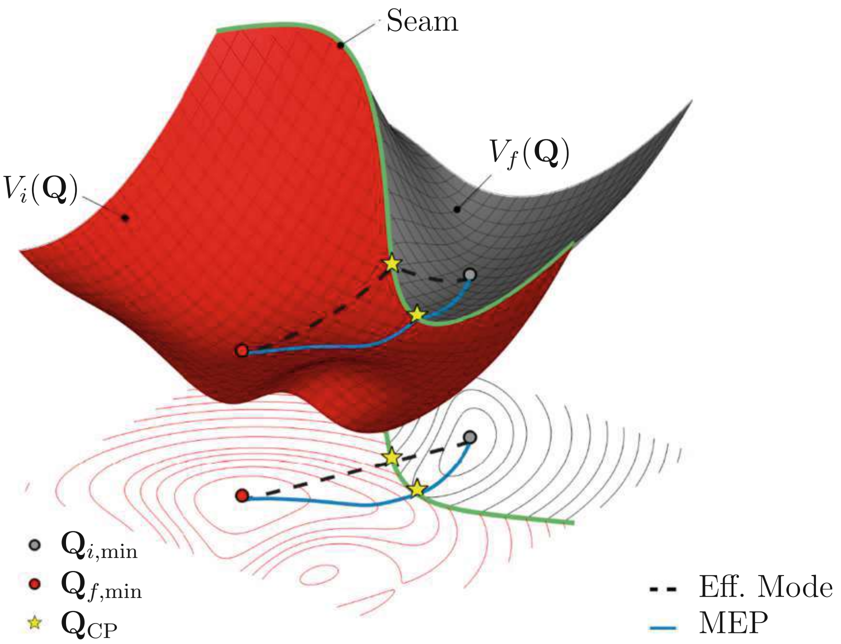


Figure 3.10. Schematic transition between PESs in multiple dimensions with two possible reaction paths marked. The minimum energy path (MEP) (blue) connects the two PESs across the minimum energy crossing point (MECP). Every other path, including the direct path along the effective mode (black), overestimates the classical transition barriers. (adapted from [DWC3])

Contrary to the transition state in adiabatic transitions, the MECP is not a saddle point on a single PES and thus cannot be found with the widely used NEB methods explained in Sec. 2.3.2. However, the MEP across both PESs as well as the MECP can be determined by a two-step process [DWT1]:

(i) The MECP solves the following constrained optimization problem:

$$\mathbf{Q}_{\text{MECP}} = \underset{\mathbf{Q}}{\arg\min}\, V_i(\mathbf{Q}) + V_f(\mathbf{Q}) \tag{3.49a}$$

$$\text{subject to } V_i(\mathbf{Q}) - V_f(\mathbf{Q}) = 0 \tag{3.49b}$$

This problem can be solved using various gradient-based constraint optimization algorithms. Results in this work are obtained by using the sequential quadratic programming [218] scheme.

(ii) After locating the MECP, the MEP can be constructed by connecting the MECP to both minimum configuration. Both segments can then be optimized towards the MEP by a NEB calculation. Note that a climbing-image NEB is not applicable in this case, since one of the endpoints (the MECP) is not a local minimum.

Similar methods have previously been used to study non-adiabatic transitions in molecules [219] and the photoluminescence spectra of doped solids [220]. However, in the context of device simulations and charge transfer to the gate dielectric, this approach is impractical due to its computational complexity. In this work, this scheme is employed to validate the 1-dimensional approximation for typical defects in amorphous SiO_2. As will be shown in Sec. 4.3.5, the effective mode approach, combined with the harmonic approximation discussed in Sec. 3.1.1, provides a good approximation to both the MEP and MECP in the multidimensional treatment, making it a reliable method for calculating charge transition rates. For further technical details on this method, the reader is referred to [DWT1].

4

Defects in Gate Dielectrics

After introducing the necessary computational methods and tools in Ch. 2 and 3, this chapter focuses on their application to defects in gate dielectrics, with particular emphasis on silicon dioxide (SiO_2) as the native insulator of silicon. First, the developed methods are applied to study the dissociation process of Si−H bonds in SiO_2 due to hole trapping, which leads to the creation of a Si dangling bond — a prototypical electrically active defect in the Si/SiO_2 system — during irradiation or hole injection into the oxide. While this analysis is performed in crystalline SiO_2, the atomistic modeling of real amorphous gate dielectrics is subsequently discussed. Following this, common point defects in amorphous SiO_2 (a-SiO_2), including self-trapped charges, will be examined. The role of these defects in device reliability will be explored in Ch. 6, through comparisons of experimental data with device models.

Although SiO_2 is the most commonly used gate insulator, novel 2D architectures require new materials to achieve competitive device reliability comparable to standard Si devices. The final section of this chapter is thus dedicated to investigating impurity defects in calcium fluoride (CaF_2), a promising insulator candidate for future nanoelectronics.

4.1 Si−H Bond Dissociation in SiO_2

This section is based on work currently prepared for publication [DWJ12].

Silicon dioxide (SiO_2) is one of the most technologically important materials due to its role as a native gate dielectric in silicon-based metal-oxide-semiconductor (MOS) transistors. In these devices, hydrogen is intentionally introduced during forming gas annealing to passivate electrically active defects that would otherwise degrade device performance. While the Si-H bonds formed during this passivation process are highly stable and largely chemically inert, first-principles calculations presented here demonstrate that these bonds can be easily broken upon hole capture. Although hole capture from the silicon valence band is unlikely due to the significant energetic misalignment with the

Si-H bond's charge transition level, the calculations reveal that hole capture from the SiO_2 valence band — such as through irradiation — leads to an efficient Si-H bond dissociation mechanism.

Compared to previous theoretical works, this study quantifies the capture coefficient for this process using the NMP framework discussed in Ch. 3, predicting a large capture cross section of $\sigma = 6 \times 10^{-15}\,\mathrm{cm}^2$. The resulting H^+ can diffuse away, leaving behind a paramagnetic E' center (i.e., an $O_3{\equiv}Si^{\bullet}$ dangling bond), confirming earlier experimental reports of H^+ release and E' center formation upon hole injection into the SiO_2 valence band [185].

4.1.1 Si–H Bonds in MOS Structures

Metal-oxide-semiconductor (MOS) transistors based on silicon are the fundamental building blocks of most electronic devices. Although many other semiconductors have superior electrical characteristics like higher carrier mobilities [221] or larger bandgaps [222], silicon is still the most widely used semiconductor due to its ability to form high-quality interfaces with its native oxide SiO_2. This allows for a simple and robust process to grow a gate dielectric on top of the device channel. While SiO_2 exhibits a lower defect concentration compared to other common gate dielectrics like HfO_2 [223], a passivation of remaining Si dangling bonds (DBs) and other defects at the interface and within the oxide with hydrogen [224] during a forming gas anneal (FGA) still has to be performed for optimal device performance. In this process step Si DBs, which would otherwise act as charge trapping [225] or recombination centers, are converted to highly stable and electrically inactive Si–H bonds.

Since the Si–H bond has a large binding energy of 2.5 eV [226] and 4.2 eV [227] relative to an interstitial H^0 in Si and SiO_2 respectively, purely thermal bond dissociation is highly unlikely under normal conditions. However, it has been demonstrated that hot carriers from the Si channel can facilitate Si–H bond dissociation in Si and at the Si/SiO_2 interface, leading to the formation of $Si_3{\equiv}Si^{\bullet}$ DBs, so-called P_b centers which are thought to be responsible for Hot Carrier Degradation (HCD) [228, 229]. Here, continous interaction with many hot carriers leads to a bond dissociation via vibrational heating [DWJ2, 230, 231].

While the electronic states of Si-H bonds in Si or at the Si/SiO_2 interface are in resonance with the Si valence band (VB), Si-H bonds in the SiO_2 dielectric possess a doubly occupied state slightly above the SiO_2 valence band maximum (VBM) as shown in Fig. 4.1. Since this state is in the bandgap of the oxide, the Si–H bond in principle can capture a hole from the SiO_2 VBM. Compared to a resonance state with a lifetime on the order of femtoseconds [232] however, a captured hole at this isolated Si–H bond would remain there for far longer. Hence, even capturing a single carrier has a reasonable chance to cause a dissociation of the Si-H bond. This process is thus relevant for understanding irradiation damage within the gate dielectric, where holes in the SiO_2 VB can be created by ionizing radiation. Indeed, it was shown experimentally that paramagnetic E' centers are created upon irradiation of SiO_2 samples [233]. Furthermore, Afanas'ev and Stesmans

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

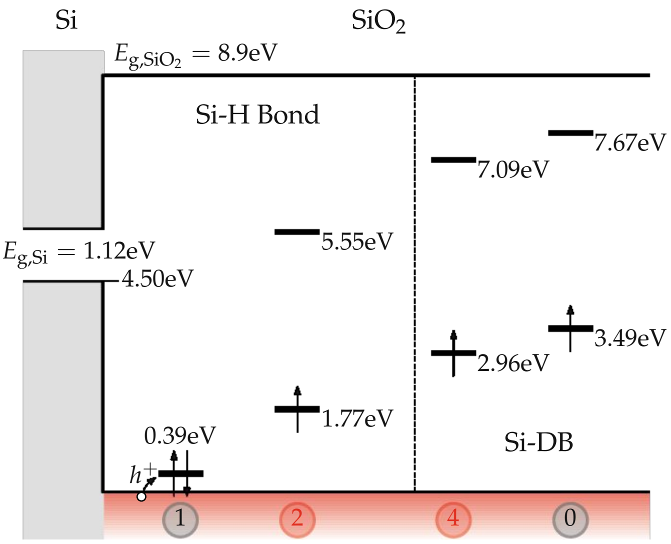


Figure 4.1. Band diagram of a Si/SiO_2 interface with the single-particle Kohn-Sham energies of the Si–H bond and the Si dangling bond in their various configurations (1, 2, 4, 0) shown in Fig. 4.4 and 4.7 as obtained from our calculations. An isolated Si–H bond (1) introduces a bonding state approximately 0.39 eV above the VBM of SiO_2. Due to the large energetic distance to the Si VBM, this state can only trap holes from the SiO_2 VB. The band offsets [241] and the SiO_2 bandgap [242] are taken from the experimental literature.

demonstrated that hole injection into the SiO_2 VB leads to the creation of E' centers, which is also correlated with the release of H^+ [185]. These findings strongly suggest that Si–H bonds dissociate upon hole trapping from the SiO_2 VB. In their experimental work [185], Afanas'ev and Stesmans determined this H^+ release mechanism to have a large effective hole capture cross section (CCS) of $\sigma = 3 \times 10^{-14}\,\text{cm}^2$.

4.1.2 Earlier Theoretical Works

Due to its technological relevance, the Si–H dissociation mechanism in SiO_2 also has been extensively studied theoretically using density functional theory (DFT). Early works done for crystalline α-quartz using the Local Density Approximation (LDA) showed that Si–H bonds within a H-decorated oxygen vacancy, a so-called hydrogen bridge [234], can dissociate with a thermal barrier of $1.15 - 1.3$ eV after capturing a hole [235, 236]. The resulting H^+ then attaches itself to a nearby bridging O atom of the SiO_2 network from which it can further diffuse away via a ring-crossing mechanism with a predicted barrier of 0.5 eV [237, 238] in good agreement with the experimental H^+ diffusion barrier of 0.38 eV [239]. Later refined studies used more accurate hybrid-functionals and also investigated the dissociation of isolated Si–H bonds in amorphous SiO_2 (a-SiO_2) [238]. It was shown that the dissociation process upon hole capture can happen spontaneously but can also be associated with a small thermal barrier in the range of $0.1 - 0.5$ eV depending on the local chemical environment. Similar results were also obtained with hybrid functional calculations for Si–H bonds at explicit atomistic Si/SiO_2 interfaces [240].

All these theoretical studies support the initial hypothesis of Afanas'ev and Stesmans that the presence of holes in the SiO_2 VB leads to a dissociation of Si−H bonds and hence causes the observed H^+ release as well as the formation of E' centers in the oxide. However, these works only studied the behavior of the Si−H bond and its dissociation kinetics *after* a hole was already trapped. As suggested by Fig. 4.1 in a ground state DFT calculation an injected hole always ends up being trapped at the previously doubly occupied Si−H state within the SiO_2 bandgap. However, in this treatment the actual hole trapping process from the SiO_2 VB is neglected under the assumption that the subsequent dissociation and diffusion of H^+ is the rate limiting step. To the best of the author's knowledge this critical assumption has not been thoroughly investigated in the literature yet.

Here, this gap is closed by using a hybrid functional DFT approach combined with the theoretical framework of nonradiative multiphonon (NMP) theory as discussed in Ch. 3 to study the hole capture process directly. For a prototypical isolated Si−H bond in α-quartz it will be shown that the hole trapping indeed happens very fast with a predicted effective hole capture cross section of $\sigma = 6 \times 10^{-15}\,\mathrm{cm}^2$ in good agreement with the experimental values obtained by Afanas'ev and Stesmans [185]. However, these calculations also indicate that the efficient trapping mechanism is caused by the pronounced anharmonicity of the Si−H bond dissociation and thus is *not* self-evident simply from the energetics presented in Fig. 4.1. Furthermore, it will be confirmed that the resulting H^+ only needs to overcome a barrier of 0.72 eV to leave the vicinity of the formed Si DB and thus can be considered to be mobile.

4.1.3 Si−H Model

As a host material for studying the Si−H bond, α-quartz is used. It is a common SiO_2 polymorph, which crystallizes in a trigonal lattice with space group $P3_221$. After performing geometry and cell relaxations using DFT, the primitive unit cell of α-quartz has the lattice parameters $a = b = 4.937\,\text{Å}$ and $c = 5.431\,\text{Å}$ in good agreement with established experimental values ($a = b = 4.912\,\text{Å}$, $c = 5.404\,\text{Å}$) [243]. A Hartree-Fock exchange mixing factor of $\alpha = 0.35$ for the HSE [66] functional was used to match the experimental bandgap of $E_{\mathrm{g,SiO_2}} = 8.9\,\mathrm{eV}$ [242]. In all calculations the primitive unit cell was transformed to an orthorhombic supercell containing 324 atoms with lattice dimensions $a = 14.817\,\text{Å}$, $b = 16.298\,\text{Å}$ and $c = 17.100\,\text{Å}$ using Γ-point sampling for the Brillouin zone. This supercell is deemed large enough to be used as a host structure for the isolated Si−H bond based on similar defect investigations [DWJ6, 227] in a-SiO_2.

Following a recent work studying DBs in bulk semiconductors [244], a single Si−O unit in the middle of the supercell is removed, resulting in one Si- and three O-DBs, each of which is passivated by hydrogen. Afterwards, the Si−H bond was manually puckered [24] to avoid interactions with the OH-groups and to simulate the Si−H bond in a more bulk-like environment. However, further calculations show that both puckered and unpuckered configurations dissociate similarly upon hole trapping. A subsequent geometry optimization is performed, leading to the structure depicted in Fig. 4.2. As can be seen, the 3 OH-groups arrange themselves in a ring due to the formation of weak electrostatic hydrogen bonds

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

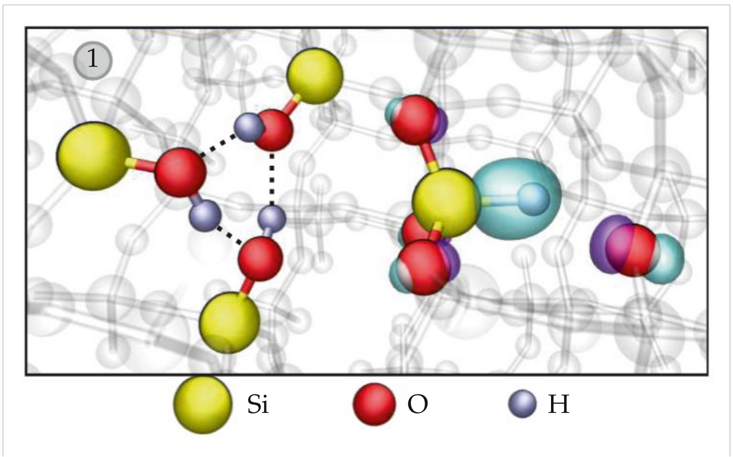


Figure 4.2. Resulting atomic structure after removal of a Si−O unit followed by passivation of all dangling bonds with H. The OH groups form a ring due to mutual H bonding, whereas the Si−H bond is isolated and only weakly interacts with an O nearby as indicated by the plotted isosurfaces for the wavefunction of the Si−H bonding state. All isosurfaces in this work are shown for an isovalue of $\pm 0.1\,{a_0}^{-3/2}$.

(O−H$\cdots$O). This procedure results in an isolated Si−H bond, which does not interact with the OH-groups on the opposite site as indicated by the isosurfaces of the corresponding molecular orbital. Contrary to the Si−H bond, the O−H bonds do not introduce any states in the bandgap since they are below the SiO$_2$ VBM.

Since the focus here is on the dissociation of a pre-existing Si−H bond rather than its formation, the neutral Si-H bond, as shown in Fig. 4.2, is used as the energy reference for the formation energies. For the purposes of this work, the chemical potential of H is assumed to be $\mu_\mathrm{H} = \frac{1}{2} E_\mathrm{tot}[\mathrm{H}_2]$.

4.1.4 Results

In this section the key findings of this work are presented, starting with the basic electronic structure of the Si-H bond, followed by a discussion of the hole-induced dissociation mechanism and the subsequent migration of H$^+$ away from the defect site.

Electronic Structure of the Si−H Bond

As described in Sec. 4.1.3, an isolated Si-H bond in an α-quartz supercell was prepared. The Si-H bond introduces a doubly-occupied bonding state 0.39 eV above the SiO$_2$ VBM. Hence it can in principle capture up to two holes from the top of the SiO$_2$ VB. However, just considering the occupation of the frozen KS states of the defect neglects any electronic or ionic relaxation associated with the charge transition. Looking instead at the charge transition levels of the Si−H bond (which include all these relaxations) in Fig. 4.3, we see that the (+/0) transition occurs at $E_\mathrm{T}^{+/0} = 1.58\,\mathrm{eV}$ above the SiO$_2$ VBM. For a hole at the VBM the Fermi level $E_\mathrm{F} = 0\,\mathrm{eV}$ is below $E_\mathrm{T}^{0/-}$, meaning the positive charge state (trapped hole) of the Si-H bond is indeed thermodynamically favorable. To the contrary, the (+2/+) CTL is 0.54 eV below the VBM. Hence it would only be favorable to capture

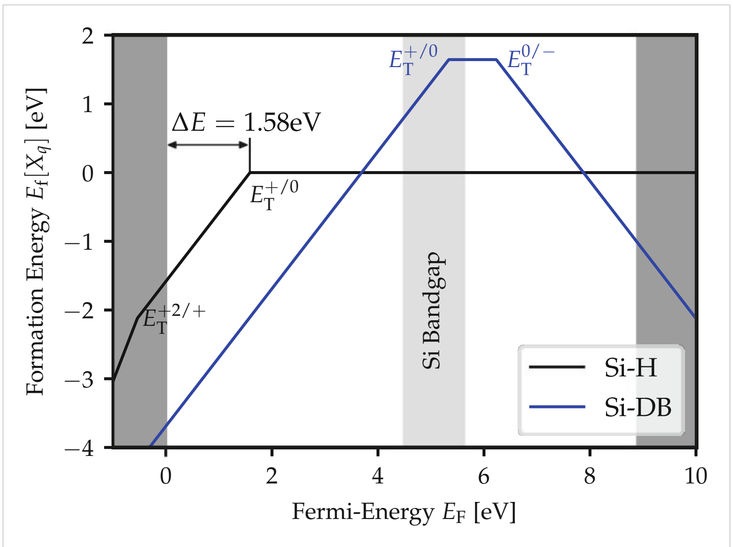


Figure 4.3. Charge transition levels of the Si−H bond and a Si dangling bond (Si-DB) within the SiO_2 bandgap. The Si−H bond has a hole trap level at $E_T^{+/0} = 1.58\,\text{eV}$ above the SiO_2 VBM and thus is able to capture a hole from the VBM. The Si-DB has both hole and electron trap levels in the vicinity of the band edges of the Si channel and hence can act as an amphoteric charge trap, potentially causing device instability.

a second hole from deep within the valence band. However, even if such highly energetic holes are excited via irradiation or carrier injection, they would thermalize and relax to the VBM within a few hundred fs [245], providing only very little time for interaction with the Si−H bond. As will be discussed in the next section, capturing a single hole is already sufficient to cause a spontaneous dissociation of the Si−H bond.

After injecting a hole, the H moves towards the closest nearby bridging O atom as shown in state 2 of Fig. 4.4. During this process the Si-H bond length increases from 1.55 Å to 1.97 Å, while the charge center of the remaining electron is shifted closer to the Si atom, reducing its distance from 1.36 Å to 0.95 Å. This indicates that the missing electron stems from the H atom becoming H^+, while the remaining unpaired electron forms a $O_3{\equiv}Si^\bullet$ dangling bond. The H^+ then attaches itself to the partially negatively charged bridging oxygen atoms as was previously reported in theoretical studies conducted in amorphous SiO_2 [237, 238]. Note that the DB defect (state 2) is not yet an isolated Si-DB (see state 0 in Fig. 4.7) and its KS levels are significantly lower due to the presence of H^+ as can be seen in Fig. 4.1. However, as will be discussed in Sec. 4.1.4, the H^+ can easily diffuse away, leaving behind an isolated neutral Si-DB. Within the simulation, the isolated DB can be created by simply removing the H^+ from the simulation cell followed by a geometry relaxation. Contrary to its Si−H precursor, the isolated dangling bond has a hole trap level $E_T^{+/0} = 5.33\,\text{eV}$ and an electron trap level $E_T^{0/-} = 6.23\,\text{eV}$ in the upper half of the SiO_2 bandgap and hence close to the band edges of the Si channel. The Si-DB is thus amphoteric and susceptible to charge trapping from the channel, and therefore can play a role in device reliability similar to other dangling bond type defects in SiO_2 like the hydrogen bridge or the hydroxyl-E' center [DWJ8, 227, 246], which will be discussed in Sec. 4.3.

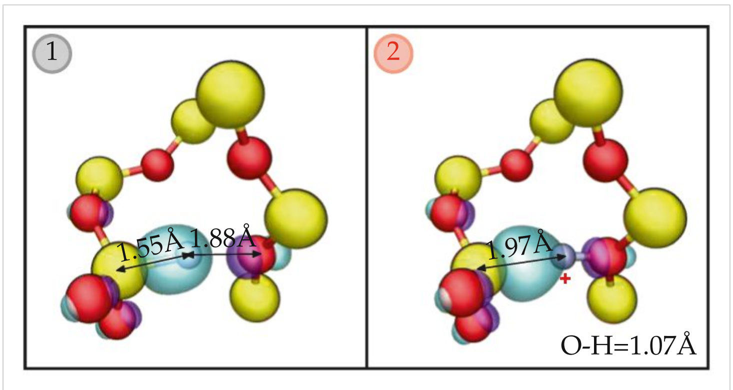


Figure 4.4. Configurations of the neutral Si–H bond (1) and after capturing a hole (2), whereby the H becomes $\mathrm{H^+}$ and attaches to the nearest bridging O atom facing the newly formed dangling bond.

ΔE	ΔQ	$\hbar\Omega_i$	$\hbar\Omega_f$	W_{if}^{eff}
1.58 eV	$4.36\,\sqrt{\mathrm{u}}$Å	46 meV	62 meV	$26\,\mathrm{meV}/\sqrt{\mathrm{u}}$Å

Table 4.1. Transition parameters for hole trapping at an isolated Si–H bond in α-quartz. The corresponding configuration coordinate diagram is shown in Fig. 4.6.

Hole Trapping and Bond Dissocation

When injecting a hole into the simulation cell, an electron is removed from the doubly occupied Si–H bond state, since it represents the highest occupied molecular orbital (HOMO) of the system. In the CC diagram presented in Fig. 4.5 this hole injection would correspond to a vertical transition from the minimum of the neutral Si–H bond PEC (black) at $Q = 0$ to the PEC of the positively charged complex. After this transition, the system quickly relaxes into its new ground state while emitting multiple phonons. As can be seen in Fig. 4.5 (bottom), this process happens spontaneously in α-quartz. In a-$\mathrm{SiO_2}$, however, both spontaneous dissociation and dissociations with small thermal barriers of up to 0.5 eV have been reported due to the stochastic chemical environment [238, 240].

While the forceful injection of holes results in the dissociation of the Si–H bond, understanding the actual hole capture process and its dynamics requires the use of the NMP framework outlined in Sec. 3.1.1. As can be seen from Fig. 4.5 (bottom), the PEC of the positive complex is strongly anharmonic in contrast to the PEC of the neutral Si–H bond. On the one hand, this anharmonicity is attributed to the attraction between the bridging-O and the $\mathrm{H^+}$ being dominated by Coulomb forces, which naturally violate Hooke's law, instead of an actual chemical bonding. On the other hand, the singly occupied Si dangling bond state crosses the $\mathrm{SiO_2}$ VBM close to the minimum configuration of the positive PEC, see Fig. 4.5 (top), leading to strong hybridization with the valence band states. Note that the injected hole still stays localized at the defect as long as the unoccupied dangling bond state remains within the $\mathrm{SiO_2}$ bandgap. Hence the potential energy of the correct electronic structure can still be obtained from ground state DFT calculations. When

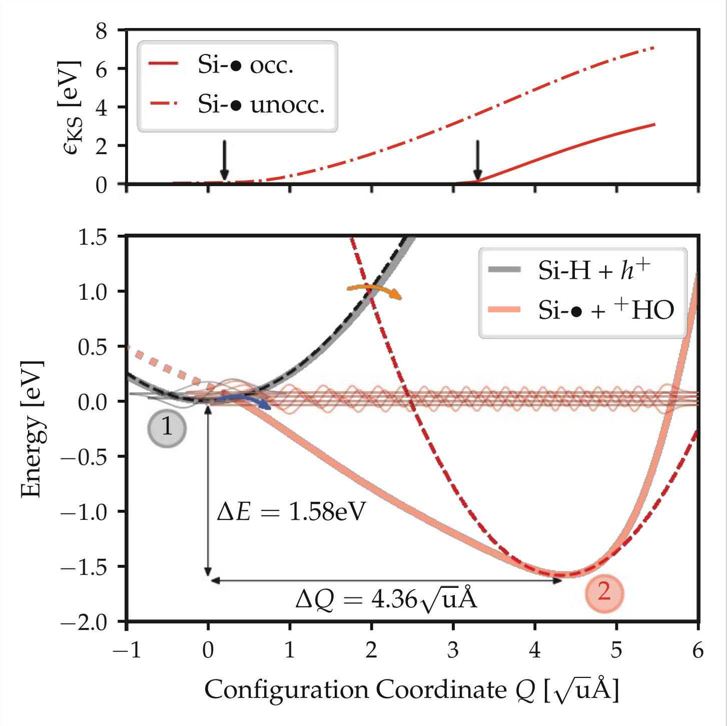


Figure 4.5. Top: Evolution of the Si dangling bond KS states during the relaxation process after hole capture and their crossing into the SiO_2 VB (arrows). The crossing of the occupied state into the SiO_2 VB is suspected to contribute to the pronounced anharmonicity in the positive charge state (2). A crossing of the unoccupied state causes the hole to delocalize near the minimum of the neutral Si−H bond. This is indicated by the dotted PEC in the lower panel. **Bottom:** PECs for the hole capture process from state 1 to 2 in Fig. 4.4 together with the vibrational states facilitating the transition. The blue and orange arrows depict the classical transition points for the anharmonic PECs and their harmonic approximation respectively, corresponding to the capture cross sections shown in Fig. 4.6.

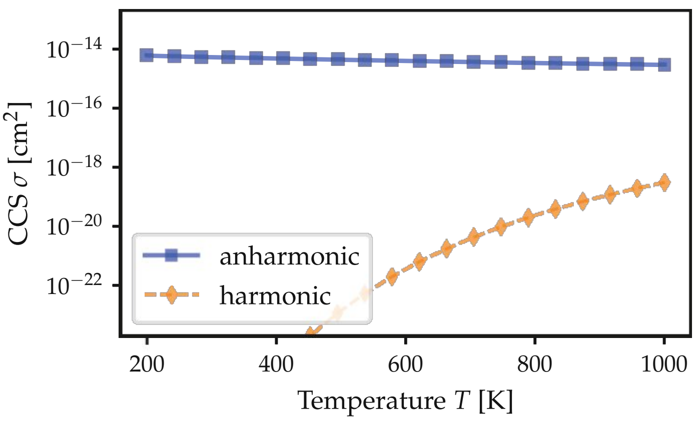


Figure 4.6. Hole capture cross sections for the Si−H bond using the (an)harmonic potential energy curves. In the anharmonic case, the capture cross section is much larger and almost temperature-independent due to the very small classical transition barrier.

the unoccupied KS state crosses into the VB however, the hole delocalizes at the VBM leading to a change of the electronic structure. An excited state with the Si dangling bond remaining unoccupied even below the VBM would be required to correctly describe the electronic structure in this region. Constructing such states would require the use of different techniques like constrained DFT (CDFT) [DWJ2, 232, 247] or the maximum overlap method (MOM) [248]. For the purpose of giving an estimate for the hole capture coefficient, however, the ground state DFT treatment is considered sufficient here, since the invalid portion of the PEC (dotted line in Fig. 4.5) is already beyond the PEC crossing point, which determines [23] the capture rate in Marcus theory [216], the classical limit for NMP transitions (see also Sec. 3.3).

In order to evaluate the hole capture coefficient given in Eq. (3.8), the vibrational modes χ_{im} and χ_{fn} were calculated by solving the vibrational Schrödinger equation numerically up to $m = 40$ and $n = 80$. The electron-phonon coupling element was determined to be $W_{if}^{\mathrm{eff}} = 26\,\mathrm{meV}/\mathrm{Å}\sqrt{\mathrm{u}}$ using the finite-difference scheme described in Sec. 3.1.2. The capture coefficients have been calculated for both the anharmonic PECs, as obtained from evaluating the DFT energies along the configuration coordinate Q, as well as the harmonic approximation based on a second order expansion of the PECs (dashed lines in Fig. 4.5). For the harmonic oscillators the vibrational frequencies $\hbar\Omega_i = 46\,\mathrm{meV}$ and $\hbar\Omega_f = 62\,\mathrm{meV}$ were obtained. For a comparison to experiments the capture coefficients were converted to a capture cross section σ according to (3.12) The effective mass of the hole was assumed to be $m_{\mathrm{h}}^* = 7.5m_0$ [249]. The resulting CCSs are depicted in Fig. 4.6. The CCS from the anharmonic PECs is roughly temperature-independent in a wide temperature range with a value of $\sigma = 6 \times 10^{-15}\,\mathrm{cm}^2$ at $T = 300\,\mathrm{K}$ in good agreement with the experimentally established value [185] of $\sigma = 3 \times 10^{-14}\,\mathrm{cm}^2$. However, the CCS from the harmonic approximation is orders of magnitude smaller and strongly temperature dependent. This stark difference can be easily understood in the classical limit of NMP transitions, in which the overall transition rate and its thermal activation are governed by the classical barrier given by the crossing point of the involved PECs, cf. Sec. 3.3.

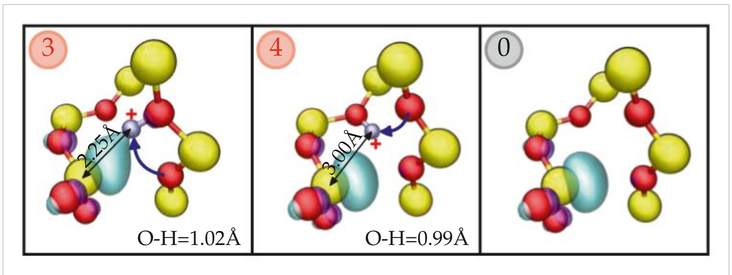


Figure 4.7. Migration steps of H^+ after hole capture and Si–H bond dissociation. Starting from state 2, the migration path leads over a nearest-neighbor hoping (3) followed by a cross-ring hoping (4). After the cross-ring hoping the H^+ distance to the Si dangling bond is large enough, so that its structure and electron distribution is almost indistinguishable from the isolated E' center (0) with the H^+ diffused away entirely.

The classical barrier for capturing a hole is 0.99 eV and only 23 meV for the harmonic and anharmonic PECs respectively. In the latter case, the barrier is so low that the largest vibrational overlap already stems from the ground state and the first excited state of the neutral Si–H bond, leading to such a weak temperature dependence. The comparison between the harmonic and anharmonic model shows that the large hole capture cross section observed in experiments can only be explained when taking the pronounced anharmonicity of the Si–H bond dissociation into account.

H^+ Migration

After the hole capture event and the subsequent dissociation, the defect is in state 2 as shown in Fig. 4.4. While in this state the Si dangling bond has already formed, its electronic structure is still quite different from an isolated DB due to the interaction with the nearby proton as can be seen in Fig. 4.1. Hence we further study the migration path of H^+ away from the dangling bond site using the NEB method.

Since the migration of H^+ in bulk a-SiO_2 has been studied in detail previously [237, 238], this work only focuses on a specific migration path leading away from the newly formed Si dangling bond, showing that the migration is not significantly impeded by the presence of the Si dangling bond. The considered migration path is shown in Fig. 4.7, the corresponding barriers are presented in Fig. 4.8. After the initial hole trapping process, the H^+ hops to a nearest neighbor bridging-O (state 3) with a barrier of 0.61 eV, which is significantly lower than the reported average 0.99 eV barrier for the same hoping process in a-SiO_2 [237] without the presence of the Si DB. In this state, the interaction between the DB and H^+ is still considerable as is indicated by the elongated O–H^+ bond length of 1.02 Å compared to the corresponding value in bulk α-quartz (0.99 Å), as well as the distorted charge distribution of the Si DB. The lowest energy transition leading further away from the DB is found to be a cross-ring hopping (state 4). This transition has a barrier of 0.52 eV with respect to state 3, which is in good agreement with the previously reported average value of 0.55 eV for cross-ring hoping in a-SiO_2 [237]. In state 4, the distance of H^+ to the Si carrying the DB is already 3.00 Å, indicating that the interaction with the DB is weak. This is further supported by the almost undistorted charge distribution compared

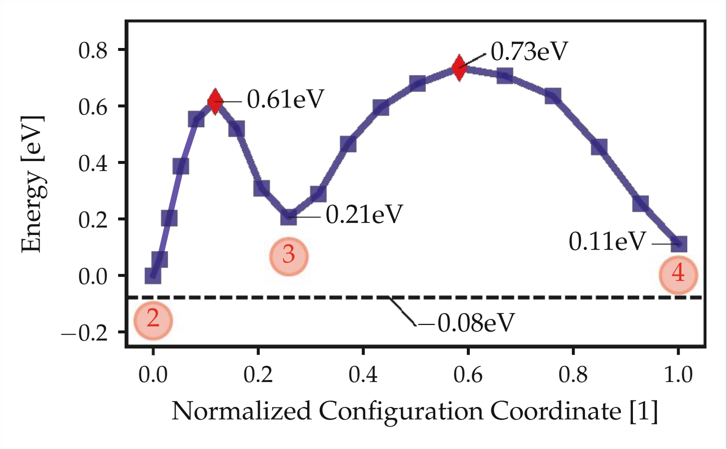


Figure 4.8. Potential energies along the H^+ migration path obtained from NEB calculations. The dashed line indicates the energy of an isolated E' center and an H^+ being attached to a bridging-O. All energies are given with respect to state 2.

to an isolated dangling bond and the $O-H^+$ bond length already reaching its bulk value of 0.99 Å at this distance. Based on these observations it is concluded that the further migration of H^+ will not be significantly different from the diffusion in bulk.

However, as was previously reported for a-SiO_2 [238], even at this distance the presence of the proton still leads to a lowering of the DB KS states by $0.5 - 0.6\,\text{eV}$ due to Coulomb attraction compared to an isolated DB as is also indicated in Fig. 4.1. Despite this, it was found that the formation energy of the DB-H^+ defect complex in its states 2–4 is consistently larger than the energy of the isolated Si-DB and H^+ constituents (dashed line in Fig. 4.8), meaning that the H^+ is not bound to the DB site. Since the involved barriers can be easily overcome at room temperature, hole capture at the $Si-H$ bond thus leads to the release of H^+ which then can diffuse freely throughout the oxide.

4.2 Modeling Amorphous Oxides

This section is based on work previously published in [DWJ1].

Theoretical modeling of defects in crystalline materials is a rich field which provides valuable insights in many phenomena like doping, diffusion processes or the performance of electronic devices. In fact, the first theoretical studies on oxide defects were conducted in crystals like α-quartz [250]. However, in the context of device reliability the modeling of oxide defects requires knowledge about the distribution of defect parameters within the amorphous gate oxide. Furthermore it has been demonstrated that certain defects occuring in a-SiO_2 like the hydroxyl-E' center are not thermodynamically stable in alpha-quartz [246, DWC4] and therefore cannot be properly studied in crystalline models. To get a more complete picture, it is thus required to also study defects in amorphous model structures.

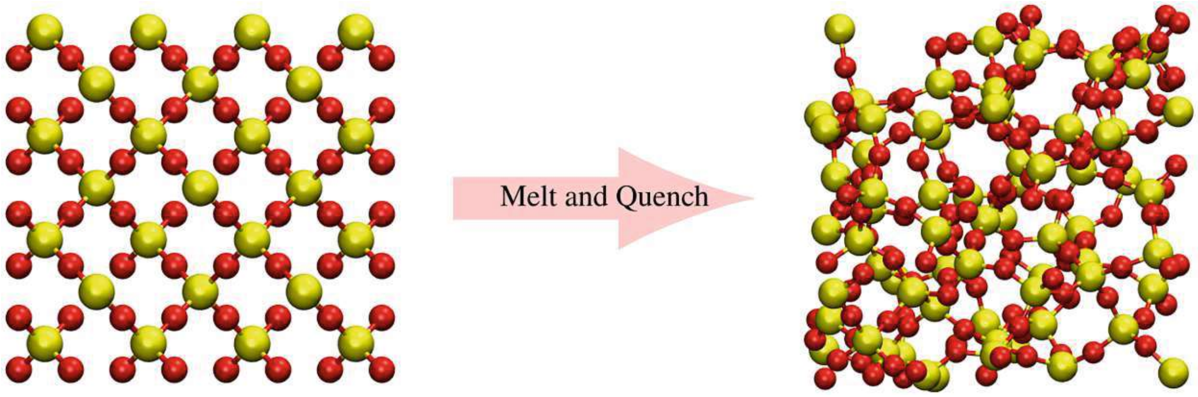


Figure 4.9. Initial crystalline SiO_2 and final a-SiO_2 structure of a melt-quench run. The yellow and red spheres indicate Si and O atoms, respectively. The initial configuration is a $2 \times 2 \times 2$ supercell of β-cristobalite, a high-temperature crystalline polymorph of SiO_2, containing 216 atoms. This crystal is an ideal starting point, since it has a cubic lattice and its density of $2.33\,\mathrm{g\,cm^{-3}}$ is close to the experimental value of $2.2\,\mathrm{g\,cm^{-3}}$ for a-SiO_2. The crystal was melted at 5000 K for 1 ns and afterwards cooled with a rate of $6\,\mathrm{K\,ps^{-1}}$ using the ReaxFF force field [251]. Note that the volume of the simulation cell is allowed to change during this process. (taken from [DWJ1])

4.2.1 Melt-and-Quench Technique

The most popular approach to create amorphous models simulates the melting of a crystal within a periodic cell, followed by an equilibration of the melt at high temperatures and a rapid cooling afterwards. This physically motivated technique is called melt-and-quench method and has been employed frequently to study amorphous materials [DWJ5, 238, 252–254]. The method is illustrated in Fig. 4.9 for the creation of amorphous SiO_2. Hereby the melting and cooling of the material is simulated with classical molecular dynamics (MD) [34]. In MD simulations the temporal evolution of a system is studied within the framework of classical mechanics, where particles move in a force field according to Newton's laws of motion. However, the study of the system behavior at varying temperatures or pressures within MD requires the ability to fix these thermodynamic quantities. This is usually achieved by extending the equations of motion with terms which enforce e.g. the temperature (thermostat) or the pressure (barostat), cf. Sec. 2.3.3.

The simulation of melt and quench processes typically spans several nanoseconds to allow the structure to properly equilibrate. Due to the significant computational demand for energy and atomic force evaluations, density functional theory (DFT) is rarely used during these steps. Instead, molecular dynamics (MD) simulations are performed using classical interatomic force fields. The geometry and cell parameters of the resulting amorphous structures are then further optimized with DFT to reduce internal stress and obtain an accurate electronic structure. The force field must be carefully parameterized to accurately describe the chemical reactions of the involved elements, ensuring physically meaningful structures. For example, high-quality amorphous SiO_2 (a-SiO_2) structures can be generated using the ReaxFF force field [251] or even the simpler potential developed by Van Beest, Kramer, and Van Santen (BKS) [255].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

In addition to classical force fields, machine-learned force fields (MLFF) have gained popularity in recent years due to their flexibility and automatic training for specific material systems. A widely used representative of this class is the Gaussian Approximation Potential (GAP), which infers local atomic properties using Gaussian process regression (GPR) together with a set of local atomic descriptors that map Cartesian atomic coordinates to invariant representations [256, 257]. Applications of such potentials for electronic materials can be found in [DWJ13, DWJ14, 258, 259]. Alternatively to GPR, other MLFFs derive material properties from a multi-layered neural network [260, 261].

It should be mentioned at this point that often highly unphysical parameters are needed for a melt-quench run. For example, the creation of amorphous SiO_2 structures requires temperatures of up to 5000 K. The high temperature can be ascribed to the use of empirical force-fields, the absence of nucleation centers as well as the confinement of the system inside the simulation cell. Also an extremely high cooling rate of up to several $\mathrm{K\,ps^{-1}}$ is typically used to reduce the number of simulation timesteps. It is therefore necessary to verify that important macroscopic and microscropic properties of the material are converged with respect to the unphysical cooling rate. For instance, it has been shown that for a-SiO_2 a cooling rate of $6\,\mathrm{K\,ps^{-1}}$ is sufficiently low to create a credible model structure, whereas other materials like silicon nitride (Si_3N_4) can require substantially lower cooling rates on the order of $0.1\,\mathrm{K\,ps^{-1}}$ [DWJ5, 262].

4.2.2 Structural Verification

The internal structure of amorphous materials strongly depend on the involved manufacturing processes as well as preparation conditions like temperature and pressure. For example, the structure can be partially crystallized due to annealing. This ambiguity makes modeling of amorphous materials for a specific application particularly hard. For that reason, the obtained structures from the melt-quench technique should be compared to available experimental data of the material system in question. Basic structural tests can include comparisons to the experimental density and electronic structure. Furthermore, the spatial behavior of the band edges across the structures should be checked with an analysis of the local density of states. This is necessary to detect spurious electric fields which lead to artificial band bending. More details on structural verification can be found in [263]. In the following, two powerful structural model tests from experimental data are discussed.

Structure Factor

Particularly useful information about the internal structure of amorphous solids can be gained from the structure factor $S(Q)$ obtained from X-ray or neutron diffraction measurements. For isotropic materials the structure factor can be written as [265]

$$S(Q) = \frac{1}{N} \sum_{ik} f_i f_k \frac{\sin(Q|\mathbf{R}_i - \mathbf{R}_k|)}{Q|\mathbf{R}_i - \mathbf{R}_k|} \tag{4.1}$$

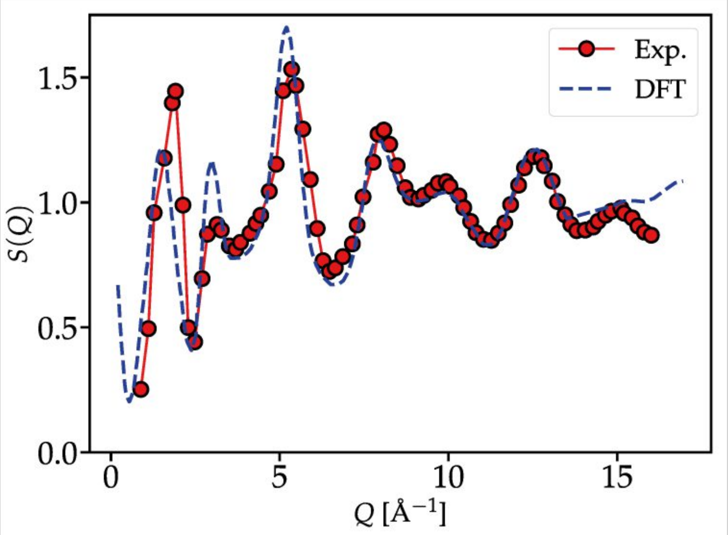


Figure 4.10. Comparison of structure factors between experiments [264] and DFT optimized a-SiO_2 models obtained from the melt-and-quench procedure as described in Fig. 4.9 using the ReaxFF force field [251]. The model structures show good agreement with the experimental data. The match of signal peak positions across a large range of Q values indicate a proper short-range and long-range order of the used models. (taken from [DWJ1], originally published in [263])

Here, f_i and f_k are the atomic form factors, describing the scattering amplitudes of the involved atoms. Q is the amplitude of the scattering vector and $\mathbf{R}_i, \mathbf{R}_k$ are the atomic positions. The structure factor can be obtained from experiment but can also be calculated for the amorphous model structures using software packages like the Atomic Simulation Environment (ASE) [266]. A comparison of $S(Q)$ between experiment and DFT optimized structures for a-SiO_2 is given in Fig. 4.10. As can be seen, the model structures are in good agreement with experimental data, indicating the validity of the melt-quench method and the used force fields.

Vibrational Density of States

Another method to test model structures is through infrared absorption and Raman scattering experiments. In these measurements, incoming photons from an IR laser couple inelastically with the vibrational states of the target material, generating spectra that are linked to the vibrational density of states (VDOS) of the sample. Since the VDOS of a structure can be predicted using DFT calculations, these measurements offer a way to compare model structures with real systems. Within the harmonic approximation, the phonon modes of the structure can be derived from the dynamical matrix, as discussed in Sec. 2.3.1. Fig. 4.11 presents the theoretical phonon density of states for the a-SiO_2 models prepared according to Fig. 4.9 alongside experimental Raman scattering intensities [267]. There is good agreement in the peak positions between the experimental data and the DFT results, providing further evidence for the validity of the structural models. However, it should be noted that within this simplified classical model, the relative intensities are not accurately represented, so only the peak positions are correct. Simulating complete

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

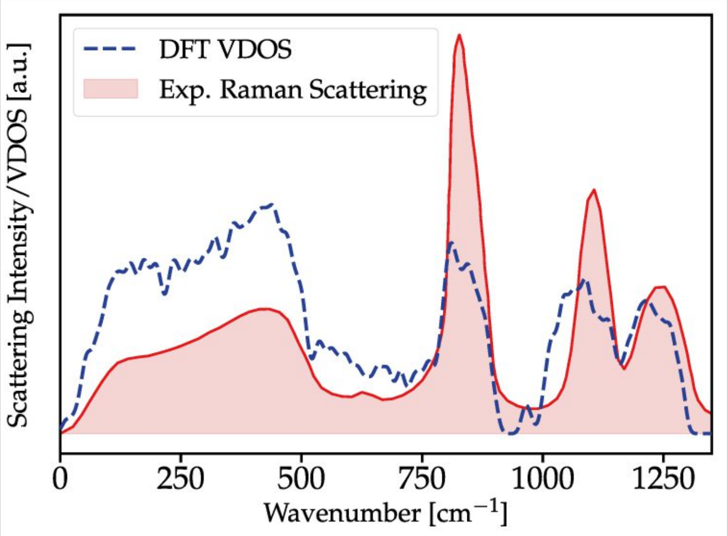


Figure 4.11. Comparison between experimental Raman scattering intensities [267] and the vibrational density of states (VDOS) in an a-SiO$_2$ model obtained from DFT calculations. As can be seen the positions of the signal peaks match well between DFT and experiment. (taken from [DWJ1])

Raman and IR absorption spectra using ab-initio methods is more complex and requires accounting for quantum mechanical selection rules. For more details, the interested reader is referred to the literature [268].

4.3 Defects in a-SiO$_2$

This section is based on work previously published in [DWJ1, DWC3].

SiO$_2$ is the native oxide of the Si substrate and has therefore been used as gate dielectric since the early days of microelectronics. Although SiO$_2$ is increasingly replaced by high$-\kappa$ materials like HfO$_2$ in order to improve the electrostatic channel control, the study of defects in a-SiO$_2$ is still highly relevant for reliability concerns. Firstly, the Si/SiO$_2$ material system is well understood both experimentally and theoretically, therefore a plethora of data is available as a reference for theoretical defect studies using DFT. Furthermore, defect densities can be kept low in the Si oxidation process, which allows an experimental characterization of individual defects in small-area devices using techniques like TDDS or RTN as will be explained in Ch. 6. Defects in SiO$_2$ are also still relevant in modern high$-\kappa$ gate stack structures since even there a thin SiO$_2$ layer at the Si interface is retained to guarantee a good interface quality. Defects in this layer can potentially trap charges rather easily due to the small distance to the channel and therefore contribute considerably to BTI and RTN in high$-\kappa$ devices. Modern devices also use plasma-nitrided SiO$_2$ (SiON) in order to decrease the gate leakage current [3]. It was shown experimentally [26] that defects in SiON behave similarly to the ones in pure SiO$_2$, therefore it is sufficient to model defects

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

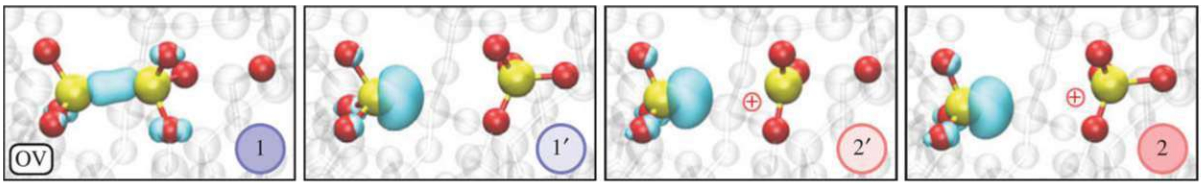


Figure 4.12. Oxygen vacancy in a-SiO_2 with its different configurations mapped to the states in the 4-state model discussed in Sec. 1.3. The turquoise bubbles indicate the HOMO electron distribution. (taken from [DWJ1], originally published in [192])

in the much simpler SiO_2 system. In the following, the key findings of recent ab-initio studies on defects in silica are summerized. All results presented here were obtained from periodic a-SiO_2 random network models containing 216 atoms, as described in Fig. 4.9.

4.3.1 Oxygen Vacancies

Presumably the most studied defect in silica is the oxygen vacancy [269–271]. It is formed when a two-coordinated O atom is missing in the a-SiO_2 network. This defect forms naturally during Si oxidation and the resulting defect concentration depends on processing parameters like temperature and oxygen partial pressure [272]. Depending on the local defect environment, the oxygen vacancy can exist in various stable and metastable states with different charges [254], it thus supports the aforementioned 4-state defect model, see Fig. 4.12. In the primary neutral configuration, the two Si atoms forming the vacancy bind together, which leads to long ranging distortions in the flexible a-SiO_2 network surrounding the defect. Since all electrons in this configuration are paired, state 1 of the OV does not carry a net spin and thus cannot be detected by spin-resonance measurements. However, when the vacancy traps a hole (state 2′), e.g. one of the electrons in the Si-Si bond is removed, the bond can break and the Si atoms move apart from each other, resulting in a Si dangling bond. In other cases, hole trapping only weakens the Si–Si bond instead of breaking it, the resulting spin density is then shared between the two Si atoms, which is known as the dimer configuration [254, 273]. In both cases, there is an unpaired electron which is detectable in electron spin resonance (ESR) measurements. It is assumed that these two configurations correspond to the experimentally observed E′ center defect in a-SiO_2 [274]. In addition, the Si atom without dangling bond can move through the plane spanned by its O neighbors. Depending on the local environment, the backprojected Si atom can be stabilized by a nearby O atom, see states 1′ and 2 in Fig. 4.12. This process is termed puckering and is associated with a thermally activated transition without charge transfer, e.g. the transitions $1 \Leftrightarrow 1'$ and $2 \Leftrightarrow 2'$ in the 4-state model. The positive state 2 is believed to be the E'_γ center known from ESR measurements [274].

Although the 4-state model was originally inspired by the OV [275], multiple theoretical studies [24, 25, 276] clearly show that the hole trap level of this defect is far below the silicon valence band edge ($E_\mathrm{T}^{0/+} - E_\mathrm{VBM,Si} \approx -3.5\,\mathrm{eV}$), see also Fig. 4.17. In a pMOS device with a thin insulator the OV thus remains neutral under all relevant bias conditions and cannot contribute to NBTI or RTN due to hole trapping.

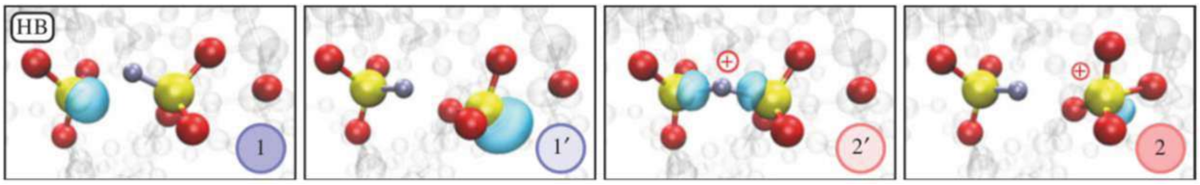


Figure 4.13. Hydrogen bridge defect in a-SiO$_2$ with its different configurations mapped to the states in the 4-state model. (taken from [DWJ1], originally published in [192])

4.3.2 Hydrogen-Induced Defects

Although some reports have suggested oxygen vacancies as the cause of electrically active oxide traps [277], the focus has recently shifted towards hydrogen-related oxide traps [278, 279]. During the forming gas anneal process, hydrogen is used to passivate Si dangling bonds at the Si/SiO$_2$ interface. This crucial step is necessary to reduce the otherwise occuring high density of interface traps, which would significantly degrade the device characteristics. However, experimental [280] as well as theoretical [237] studies indicate that hydrogen easily diffuses in SiO$_2$ with activation energies as low as 0.2 eV. Furthermore it was shown that there is a connection between hydrogen concentration during anneal and the resulting density of electrically active oxide traps [278, 279]. DFT studies indicate that H can form promising candidates for active defects like the hydrogen bridge or the hydroxyl-E' center, which will be discussed in the following.

Hydrogen Bridge

Although the oxygen vacancy itself stays neutral under typical bias conditions, it acts as a precursor site for another related defect, the hydrogen bridge (HB). The HB is formed by the reaction of atomic hydrogen with an OV. DFT studies in α-quartz conducted by Bloechl *et al.* suggested that the HB defect might be responsible for stress-induced leakage current (SILC) in gate oxides [234, 250]. The formation of a HB defect from an interstitial H atom and an OV is nearly barrier-less, whereas the reverse reaction of H release from the HB has a thermal barrier of 2.8 eV on average [227].

The most stable neutral configuration of the HB defect is shown as state 1 in Fig. 4.13. Here the atomic H binds to only one of the Si atoms forming the vacancy, leaving the other one undercoordinated. The resulting ESR from this dangling bond was first identified by [281]. The Si-H bond length is rather constant at 1.47 Å, whereas the weak interaction with the second Si atom is strongly dependent on the local environment of the defect, which results in a large spread of the corresponding Si–H bonds [24, 227]. The HB can capture a hole resulting in a configuration depicted as state 2′ in Fig. 4.13. Here the remaining 2 electrons are shared between the two Si and the H atom, resulting in a nearly symmetrical configuration. Similar to the aforementioned OV in a-SiO$_2$, the HB can undergo a puckering transition in both charge states, where one Si atom moves through the plane of its O neighbors.

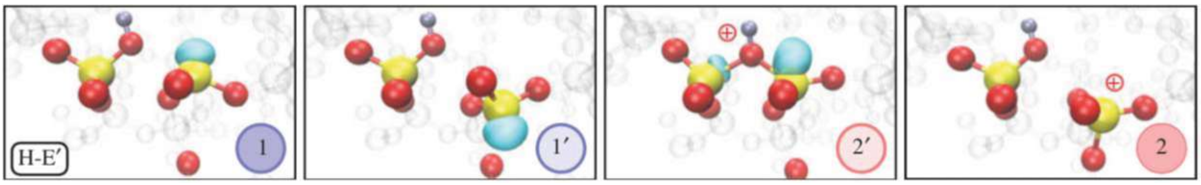


Figure 4.14. Hydroxyl-E' center in a-SiO_2 with its different configurations mapped to the states in the 4-state model. (taken from [DWJ1], originally published in [192])

Hydroxyl-E' Center

In a-SiO_2 the Si-O bond lengths and angles do not have well defined values but are distributed. It was shown that sites with distorted, but otherwise intact, Si-O bonds can interact with atomic hydrogen resulting in new defects, which do not exist in crystalline SiO_2 [227]. One such defect, the so-called hydroxyl-E' center (HE), arises from the interaction of H with strained Si-O bonds which are significantly longer than the equilibrium bond length of 1.61 Å in α-quartz [246]. In this case, H is able to bind to the oxygen atom and break one of the two Si-O bonds, resulting in a hydroxyl group facing a Si dangling bond as depicted in Fig. 4.14 (state 1). Such a configuration is more stable than the interstitial H by 0.8 eV on average. Due to the significant amount of strained bonds in a-SiO_2 and its thermodynamical stability, the HE center is expected to be formed in significant quantities. By trapping a hole, the broken Si-O bond is restored and the positive H ion is bonded to the partially negative O atom by Coulomb interaction. The HE center also supports a 4-state defect model via a similar puckering transition as those of the OV and HB. Due to its thermodynamical trap levels (see Sec. 3.4) near the Si valence band, this defect is a promising candidate for causing NBTI and RTN in pMOS devices [24].

Beside the HE center, another hydrogen related defect in a-SiO_2 was discovered. H can donate its electron to a Si atom with a shortened bond (< 1.61 Å) and binds electrostatically to the neighboring O atom. This results in a defect termed $[SiO_4/H]^0$ center where the Si-O-Si angle is stretched and an additional electron is localized at one of the Si atoms. Naturally, such a defect is formed more easily, when H encounters a site with an already stretched Si-O-Si angle. It was shown that this defect is only slightly more stable (by $0.1 - 0.2$ eV) than an interstitial H atom in a-SiO_2. Due to its lower thermodynamic stability and lower barriers for the H atom to become interstitial again compared to the HE center, it was concluded that the interaction of H with the a-SiO_2 network primarily leads to HE centers instead of $[SiO_4/H]^0$ centers [227].

4.3.3 Self-Trapped Charges

Landau was the first to theorize about the possibility of charge carriers becoming spontaneously localized while moving through a pristine crystal lattice [282]. This can occur if small perturbations in the uniform charge distribution, caused by thermal fluctuations, lead to a stabilizing polarization of the material, further localizing the charge in regions of higher density. If the electron-phonon coupling mediating this interaction is weak, the

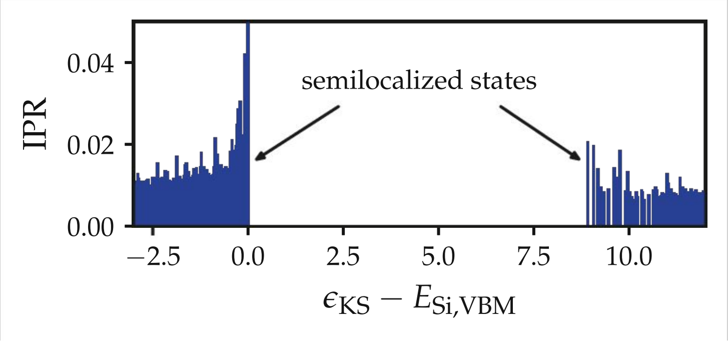


Figure 4.15. Inverse participation ratio (IPR) of electronic states in amorphous SiO$_2$. The high IPR indicates that the electronic states at the band edges are not fully delocalized but rather concentrated in certain regions of the simulation cell, see also Fig. 4.16 (left).

carrier becomes only weakly bound and remains spread out over many lattice sites, forming a so-called *large polaron*. In contrast, with strong electron-phonon interaction, the carrier can ultimately localize at a single site, resulting in a *small polaron* [283].

Polarons, or self-trapped charges, are known to play crucial roles in various phenomena, such as charge transport in organic semiconductors [284] and light emission in perovskites [285]. In amorphous materials, conduction and valence band edges can already be partially localized at certain sites, as is evident by the electronic inverse participation ratio (IPR) of the Kohn-Sham states shown in Fig. 4.15. Therefore, charge self-trapping is expected to be prevalent in amorphous materials. Indeed, in many commonly used gate dielectrics, self-trapping is suspected to play a significant role. For example, the dielectric breakdown of HfO$_2$ is thought to be related to the formation of electron bipolarons [286], and self-trapped charges may be responsible for charge storage in Si$_3$N$_4$, which is used as a trapping layer in charge-trap flash memory devices [DWJ4, 287].

In the case of a-SiO$_2$, the existence of self-trapped holes (STH) has been known for decades. However, experimentally these STHs are only stable below 180 K [288]. Furthermore, spectroscopic studies [289] as well as theoretical works [290] indicate that the trap level of the STH is far below the Si valence band edge and thus irrelevant for electronic devices. However, rather recently also the existence of self-trapped electrons (STE) in a-SiO$_2$ has been predicted [291]. Here, it was discovered that the conduction band edge of the used a-SiO$_2$ models is partially localized at particularly long Si-O bonds, as depicted in Fig. 4.16 (left). Upon electron capture, the electronic state becomes occupied, causing a pronounced structural distortion that localizes the additional electron at a specific Si atom, as shown in Fig. 4.16 (right). This structural relaxation is accompanied by a significant energy gain, resulting in the formation of a deep electron trap. It has been further demonstrated that the additional electron spontaneously localizes at sites where the O-Si-O bond angle exceeds 132°. Based on molecular dynamics models, the concentration of such sites is estimated to be around $4.0 \times 10^{19}\,\mathrm{cm}^{-3}$, which is quite substantial compared to other defects such as oxygen vacancies or hydrogen bridges. It was also speculated that STEs are responsible for a previously unidentified electron trap band located 2.8 eV below the a-SiO$_2$ conduction band edge, with a measured trap concentration of $5 \times 10^{19}\,\mathrm{cm}^{-3}$ [292, 293].

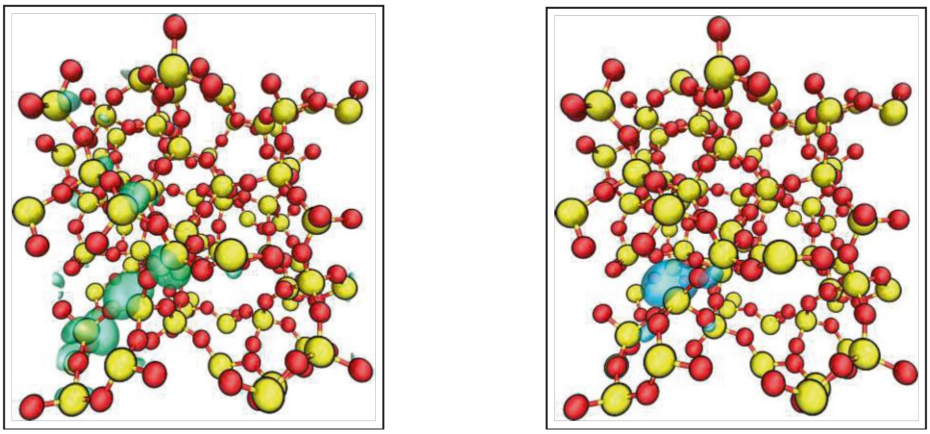


Figure 4.16. Self-trapped electrons in amorphous SiO_2. **Left:** In a-SiO_2, the conduction band edge is semilocalized around strained sites. **Right:** Upon electron trapping, structural relaxation leads to a strong localization of the trapped electron together (adapted from [DWJ15])

While El-Sayed *et al.* [291] established the existence of STEs in a-SiO_2, their work did not focus on the charge trapping dynamics of STEs but rather their structural characterization. Hence, important parameters like the relaxation energy E_R or the charge transition level $E_\mathrm{T}^{0/-}$ were not calculated. Within the scope of this thesis, the distribution of these parameters was determined based on the methodologies described in Sec. 2.4.3 and Sec. 3.1.1 using the a-SiO_2 models described in Fig. 4.9 and the PBE0_TC_LRC [67] hybrid functional. The calculations utilized 485 initial structures from MD simulations which were pre-selected based on the O-Si-O bond angle criterion introduced in [291]. The subsequent DFT calculations yielded 50 structures which show spontaneous electron localization. The obtained parameter distribution will be discussed in Sec. 6.3.3. Using a compact device model, it will be shown that STEs are likely responsible for gate-leakage currents via trap-assisted tunneling (TAT).

4.3.4 Trap Level Distributions

Various degradation experiments, combined with simulations using the compact physics model *Comphy*, have revealed the positions of hole and electron trap bands in SiO_2 [26]. Based on electrical MSM sequences, defect bands for NBTI and PBTI degradation were extracted for multiple commercial device architectures within this framework. A comparison between DFT calculations and the resulting experimental trap levels for a planar SiON[1] device[DWJ8] is shown in Fig. 4.17. It is important to note that the OV, HB, and HE defects have multiple different states due to the possibility of an internal puckering transition, as discussed earlier. However, to compare the trap levels to the effective 2-state model implemented in *Comphy* (cf. Sec. 5.4.4, in particular Fig. 5.2), only the NMP transition

---

[1] SiO_2 and SiON behave rather similarly in terms of their BTI degradation, allowing to compare trap bands in both oxides [26].

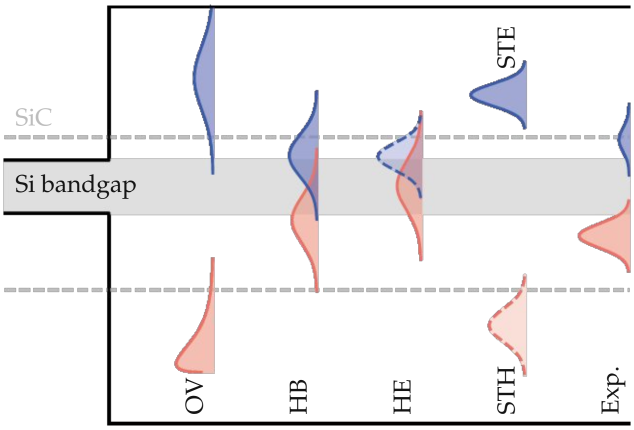


Figure 4.17. Schematic trap level distributions from experiments and DFT predictions for various defect candidates in a-SiO$_2$. The shaded area and the horizontal dashed lines correspond to the position of the Si and SiC bandgap. Red and blue lines indicate the distribution of hole and electron traps respectively. The dashed lines show distributions from insufficient data sets which may need further investigation. The CTLs for OV, HB and HE defects are taken from [DWJ6]. The experimental data stems from planar SiON devices [DWJ8]. STE trap levels are subject of this work and are published in [DWJ15]. Note that in the case of STHs the theoretical absorption spectrum was plotted, since the CTL was not calculated in these studies. (adapted from [DWJ1])

levels between the unpuckered configurations are considered here, e.g. between states 1 and 2′ in Fig. 4.12-4.14 for hole trapping. A comprehensive account of trap levels for both puckered and unpuckered configurations can be found in [DWJ6].

In the case of NBTI degradation, it can be observed that hydrogen-related defects, such as the HB and the HE center, can act as hole traps near the Si valence band, and could explain the available experimental data. Although the oxygen vacancy (OV) defect can capture holes, as discussed in Sec. 4.3.1, the resulting trap level is too deep to account for the observed NBTI degradation. Under typical bias conditions, this defect type remains mostly neutral and does not interfere with device electrostatics. The same holds true for self-trapped holes (STHs).

PBTI degradation due to electron trapping is minimal in Si/SiO$_2$-based devices and has not been as extensively studied as hole trapping. Here, hydrogen-related defect candidates such as the HB and the HE center are known to act as electron traps [DWJ6, 227], with their resulting trap levels aligning closely with the extracted electron trap band. The exact dynamics of electron trapping in these defects, however, are not yet fully understood and require further investigation. For instance, it has been discovered recently that the HE center exhibits additional, deeper electron trap levels, which might render the shown levels near the SiO$_2$ conduction band edge inactive [DWJ6]. Fig. 4.17 also indicates that in wide bandgap semiconductors, such as SiC, STEs and OVs may play a role due to different band alignments.

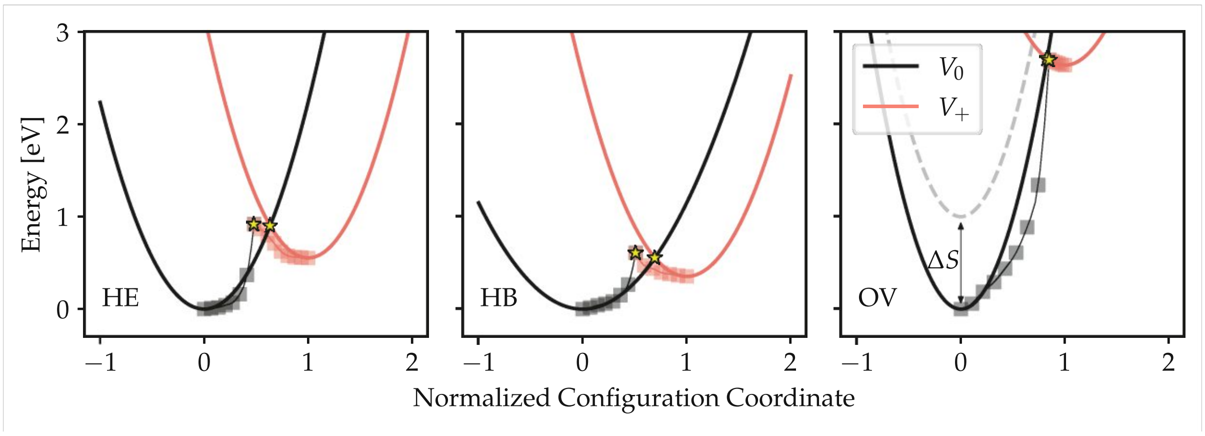


Figure 4.18. Resulting PECs of the effective mode approach (lines) and the minimum energy path (dots) for representative defects in a-SiO$_2$. The crossing point energy is reproduced well by the effective mode approximation. While the shape of the MEP energy curve can differ from the parabolic shape in the effective mode model, this does not affect the classical transition rates. Note that in a device, the PECs can be shifted relative to each other by an amount ΔS due to the electric field inside the oxide. (adapted from [DWC3])

Finally it should be emphasized that calculations of trap levels from DFT depend on the used XC functionals and are subject to uncertainties. The given distributions should therefore be considered as qualitative guidelines for the identification of possible defect candidates rather than a precise quantitative analysis. In order to get more accurate estimates for charge trap levels other theories like the GW approximation could be used to calculate correction terms for the studied defects [294].

4.3.5 Effective Mode Barriers

This section is based on work previously published in [DWC3].

While the effective mode approximation described in Sec. 3.1.1 is frequently used to predict charge transition rates [23, 164], it is not immediately clear that such a simple 1-dimensional approximation yields accurate results. This is particularly important in cases where a classical crossing point exists, as the minimum energy crossing point (MECP) must be accurately represented, since the associated thermal barrier dominates the transition rates, cf. Sec. 3.3. The effective mode method is validated here for the previously discussed OV, HE, and HB defects in a-SiO$_2$. Using model structures containing 27 HE centers, 11 HBs, and 14 OVs from previous studies [23], the minimum energy path (MEP) and corresponding barriers for the hole trapping process $1 \Leftrightarrow 2'$ were computed using the approach outlined in Sec. 3.3.2 and compared to the results obtained from the simpler effective mode method.

Before conducting a statistical evaluation of these various defect samples, it is helpful to examine the potential energy curves (PECs) of a representative defect from each type, as shown in Fig. 4.18. The effective mode model accurately predicts the MECP energy, and consequently the transition barriers for hole capture and emission, for all three defect types.

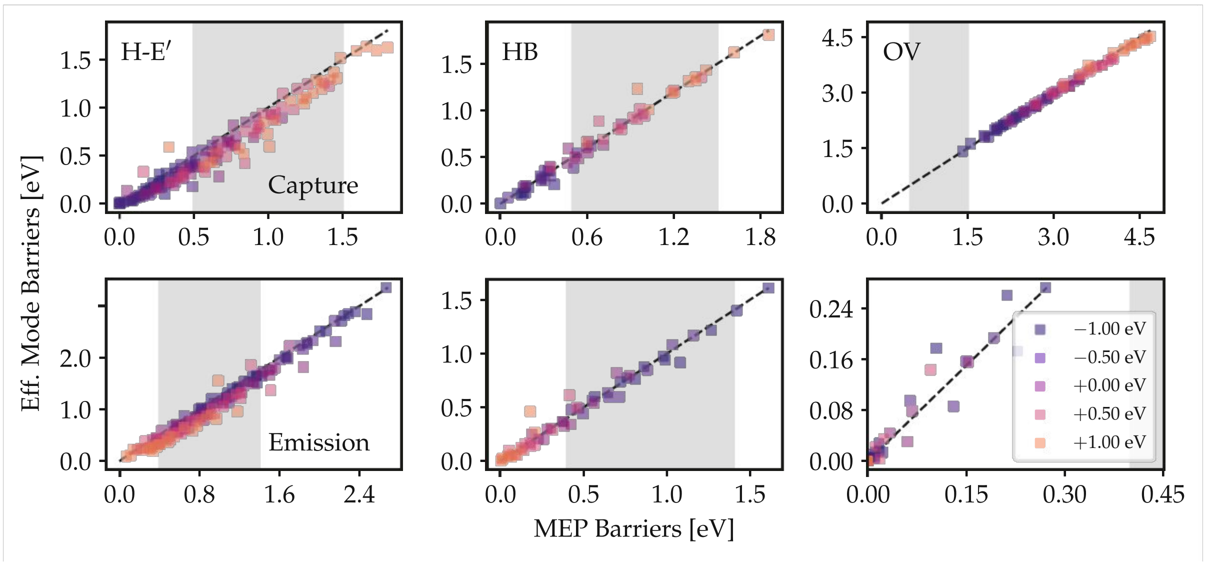


Figure 4.19. Correlation plots showing the agreement for capture ($1 \Rightarrow 2'$, top) and emission ($2' \Rightarrow 1$, bottom) barriers for the PECs of the effective mode and the minimum energy path for various bias conditions $\Delta S = [-1.0, \ldots, 1.0\,\mathrm{eV}]$. As can be seen, the simple 1-dimensional approximation within the effective mode model gives good results for all defects and bias conditions. In these test cases, the error introduced by this approximation is always small compared to the parameter spreading due to the amorphous structure. The shaded areas mark the region [25] which is experimentally accessible. (adapted from [DWC3])

While the shape of the energy curves may deviate from the minimum energy path (MEP), this is immaterial for the classical transition rates and even has only a minor impact in the quantum mechanical case. It is also important to note that Fig. 4.18 presents the PECs without an applied electric field across the oxide. During device operation, however, the electric field alters the electrostatic potential, shifting the PECs relative to one another by an amount ΔS, see also Sec. 5.4.4.

The findings from Fig. 4.18 also hold true when considering an entire ensemble of defects to account for the statistical distribution of defect parameters in the amorphous gate oxide. Fig. 4.19 presents the correlation between barriers obtained using the effective mode approach and the MEP for gate biases causing shifts in the range $\Delta S = [-1.0, \ldots, 1.0\,\mathrm{eV}]$. As seen in the plots, the largest deviations occur for HE center defects. This is expected, as the structure of HE defects is more complex compared to OV and HB defects, resulting in a longer transition path in configuration space. However, in all cases studied, the error from using a 1-dimensional PEC remains small compared to the intrinsic variation in barrier heights introduced by the amorphous oxide.

Fig. 4.18 also demonstrates that a significant portion of hydrogen-related defects lies within the experimentally accessible range of barriers and can contribute to NBTI degradation, whereas OVs remain mostly inactive. This finding aligns with previous theoretical studies [23–25] on a-SiO$_2$ defects. It should be noted that OV defects may still get positively charged in devices with sufficiently thick oxides, where larger values of ΔS are possible under realistic bias conditions. However, in modern devices with thin oxide layers, the achievable shift is too small to activate OV defects based on these results.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

The findings above validate the use of the effective mode approximation for treating oxide defects in a-SiO_2 in reliability models. In the subsequent chapters on device modeling, this approximation is employed in place of the more accurate, yet computationally and conceptually challenging, MEP approach.

4.4 Calcium Fluoride as a Gate Dielectric

This section is based on work published in [DWC5].

Being able to form high quality insulator/semiconductor interfaces is essential for creating operational nanoelectronic devices, metal-oxide-semiconductor field-effect transistors (MOSFETs) in particular. While conventional Si-based devices have settled on using a thin layer of native oxide SiO_2 followed by high-k materials like HfO_2 as gate dielectric, finding suitable dielectric materials for 2D devices is challenging [295]. First of all, chemical bonding with the dielectric layer disrupts the electronic structure of the delicate 2D channel material, leading to a significant reduction of carrier mobility compared to its theoretical maximum [296]. Furthermore, the usually employed amorphous 3D oxides, e.g. SiO_2 and HfO_2, have large defect concentrations at their surfaces [297]. Those defects are in close proximity to the device channel, since the 2D layer is directly exfoliated/transferred on a prepared dielectric surface for frequently manufactured back-gated prototype devices. This leads to a large electrostatic impact and enhanced bias temperature instability (BTI) or $I_\mathrm{D}(V_\mathrm{G})$ hysteresis.

For these reasons, layered insulators which can form quasi van der Waals (vdW) interfaces with the channel have been considered [9]. The most prominent candidate within this category is hexagonal boron-nitride (hBN), which however is not scalable due to its rather small bandgap of $E_\mathrm{G} = 5.95\,\mathrm{eV}$ and a comparatively low permittivity ($\varepsilon \approx 5$), leading to excessive gate leakage currents [298]. It was recently suggested to use the ionic crystal calcium fluoride (CaF_2) as a gate dielectric for 2D materials [299]. Its large bandgap of $E_\mathrm{G} = 12.1\,\mathrm{eV}$ and its fluorine-terminated inert (111) surface make it appear like an almost ideal insulator. Moreover, thin layers of CaF_2 have been deposited on Si(111) [300] but also free standing slabs of CaF_2 are dynamically stable according to theoretical studies [301]. Indeed, initial reports show that employing CaF_2 as a substrate leads to significantly reduced $I_\mathrm{D}(V_\mathrm{G})$ hysteresis and BTI [299]. This improvement is attributed to the clean quasi vdW interface between MoS_2 and CaF_2 as schematically illustrated in Fig. 4.20. All these findings suggest that CaF_2 might be used as a reliable gate insulator in (2D) nanoelectronics in the near future.

4.4.1 Material Properties

Even though, the overall defect density is expected to be smaller in CaF_2 than in amorphous oxides, the defect levels and their prevalence will still determine the electrical stability of devices based on CaF_2 as a gate insulator. What is more, in order to include CaF_2 in TCAD simulation workflows, in particular when focusing on device reliability, not

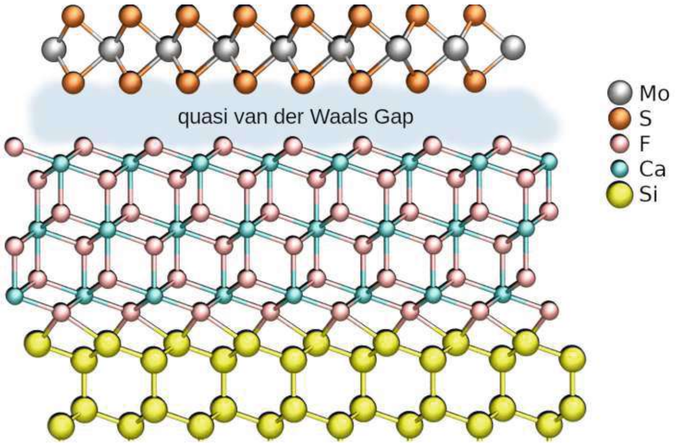


Figure 4.20. Si(111)/CaF_2/MoS_2 gate stack envisioned in [299]. Due to its good lattice match, CaF_2 can be grown on Si(111) in thin layers. It provides a virtually defect-free quasi van der Waals interface to the 2D material and hence considerably improves the reliability of 2D devices. (taken from [DWC5])

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

only do the basic material properties like permittivity, bandgap etc. have to be known, but also information about different defects in the material. Here, DFT is employed to determine important defect parameters like the thermodynamic trap levels, relaxation and formation energies. These are important input parameters for physical charge trapping models based on NMP theory, which allow the prediction of degradation over the course of the device lifespan.

All calculations were performed on a 3×3×3 cubic cell of CaF_2 using the PBE0_TC_LRC hybrid functional with the PBEsol [302] parametrization and a mixing factor $\alpha = 0.39$ for the Hartree-Fock (HF) exchange. This value was determined by calibrating the theoretical single-particle bandgap to the known *indirect* bandgap $E_G = 11.8\,\mathrm{eV}$ [303]. Since the mixing factor is inversely related to the optical permittivity ε_∞ and thus material dependent [304], there is some theoretical justification for tuning the mixing factor to match certain experimental results. However, since this is still an empirical procedure, the simulation setup was verified by calculating various other well-known properties of bulk CaF_2 and comparison to experiments as summarized in Tab. 4.2. As shown, the theoretical predictions are in good agreement with experimental values for lattice constants, bulk modulus, and static and optical permittivities, indicating that the DFT setup is well-suited to describe this material. The permittivities were calculated by evaluating the change in polarization under a finite electric field using the Berry phase formalism [305].

4.4.2 Si-Impurity Defects

The focus of this study lies on Si-impurity defects, which are relevant since many 2D device prototypes are fabricated as back-gated devices with a silicon wafer acting both as substrate for the dielectric and as a back-gate for the 2D layer. Furthermore, the stable (7×7) reconstructed Si(111) surface has four excess Si atoms per unit cell [310], which either enter the gas phase as SiF_4 during CaF_2 deposition or are incorporated in CaF_2. It will be shown that Si-impurity defects like the Si interstitial (Si_i) or the Si substituting

		PBE0 $\alpha = 0.39$	exp.	ref.
Lattice Constant a	[Å]	5.454	5.463	[306]
Bulk Modulus K	[GPa]	77.29	81.67	[307]
Indirect Bandgap E_G	[eV]	11.90	11.8	[303]
Static Permittivity ε_0	[1]	6.43	6.81 - 6.84	[308, 309]
Optical Permittivity ε_∞	[1]	1.85	2.05	[308]

Table 4.2. Comparison of CaF_2 bulk properties calculated by DFT against experimental data. (taken from [DWC5])

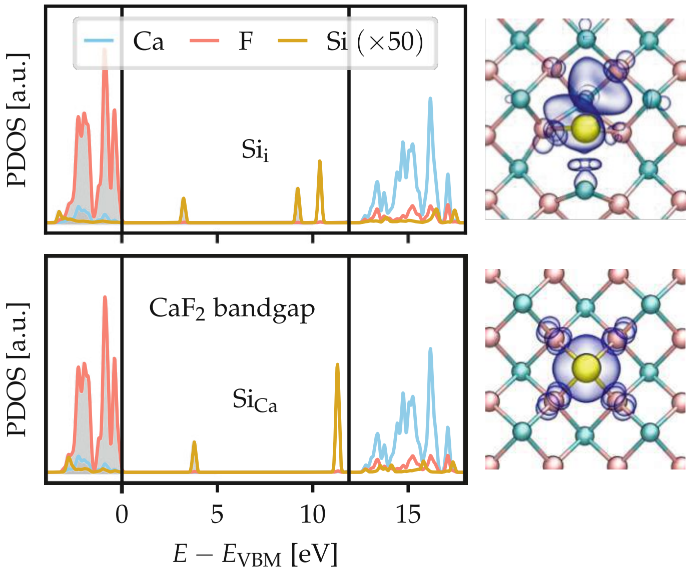


Figure 4.21. Projected density of states (PDOS) and localized highest occupied molecular orbitals of the Si_i **(top)** and Si_{Ca} **(bottom)** defect. Both defects introduce localized states in the bandgap as can be seen from the DOS projected onto the Si impurity. Note that the Si PDOS is scaled by a factor of 50 for illustrative purposes. (taken from [DWC5])

for Ca (Si_{Ca}) can act as amphoteric traps and might contribute to hysteresis and BTI in gate-stacks similar to Fig. 4.20. Given the novelty of this material in 2D nanoelectronics, to the best of the author's knowledge, this is the first report on such impurity defects in CaF_2.

The results for the Si_i and Si_{Ca} defects are presented below, along with a discussion of their potential implications for 2D MOSFETs on CaF_2. Additionally, defect parameters suitable for use in NMP charge trapping models in reliability-aware TCAD simulations are provided to support device simulations for this new material system.

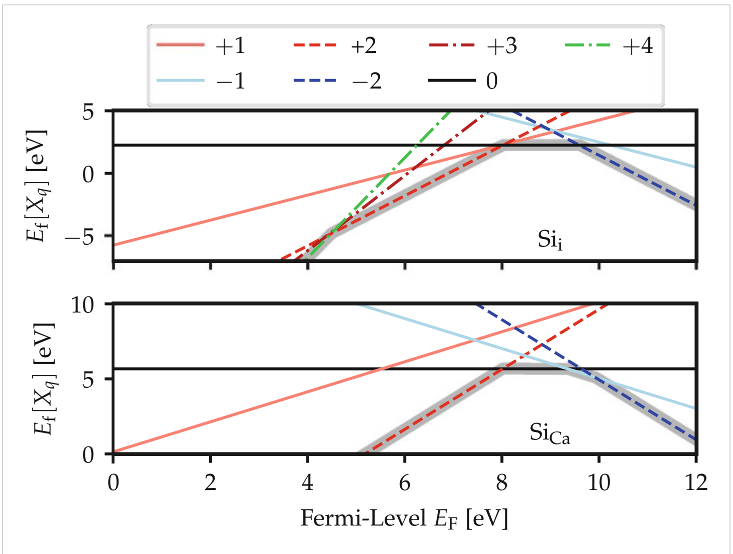


Figure 4.22. Formation energies for the Si_i **(top)** and Si_{Ca} **(bottom)** defect as a function of the Fermi level E_F for different charge states. The lowest energy state at a given E_F is referred to as *stable* and highlighted by the thick gray line. Note that there are *unstable* states, which are never lowest in energy, e.g. the +1 state for both defects. E_F is taken w.r.t. the valence band maximum (VBM) of CaF_2. (taken from [DWC5])

Electronic Structure

As shown in the atomistic models provided in Fig. 4.21 (right), the substitutional Si atom (bottom) is incorporated into the host crystal with small lattice distortions. On the other hand, the Si at an interstitial site (top) binds to the surrounding F atoms and displaces a nearby Ca in the process, leading to quite strong local distortions of the host lattice. As can be seen from the density of states projected onto the Si impurity in Fig. 4.21 (left), both defects introduce localized electronic states within the CaF_2 bandgap. In particular, the Si_i defect exhibits two filled and one empty state, compared to only one filled and one empty state for the Si_{Ca}. Therefore, the Si_{Ca} potentially can carry the net charges $\{-2, -1, 0, +1, +2\}$, whereas Si_i additionally could also have a +3 and +4 state as well. The defects were relaxed in all their possible respective charge states in order to determine whether or not a particular state actually exists, i.e. the excess charge stays localized at the impurity site after relaxation. For the defects studied in this work, this is fulfilled for all the aforementioned charge states, hence they all have to be considered for subsequent studies on charge transitions.

Charge Trapping Parameters

Following the relaxation, the thermodynamic stability of the differently charged states was examined by calculating their respective formation energy $E_f[X_q]$ as a function of the Fermi level according to Eq. 2.75. The resulting stability ranges are presented in Fig. 4.22. As can be seen, the Si_i and Si_{Ca} defect can be thermodynamically stable in their $\{-2, 0, +2, +4\}$ and $\{-2, -1, 0, +2\}$ states respectively, hence both are amphoteric defects which can act as hole *and* electron traps depending on the applied bias.

Defect	Transition	$E_\mathrm{f}[X_0]$ [eV]	E_T [eV]	E_R [eV]	R [1]
$\mathrm{Si_i}$		2.25			
	$0/+1$		7.99	1.16	0.77
	$+1/+2$		8.05	3.03	0.99
	$+2/+3$		4.39	1.94	0.95
	$+3/+4$		4.57	3.70	1.04
	$0/-1$		10.24	3.54	1.41
	$-1/-2$		8.95	0.73	0.50
$\mathrm{Si_{Ca}}$		5.67			
	$0/+1$		5.53	2.23	0.90
	$+1/+2$		10.50	0.92	2.60
	$0/-1$		9.35	2.12	0.92
	$-1/-2$		9.92	1.67	0.91

Table 4.3. Theoretical defect transitions and parameters predicted with DFT. E_T is given w.r.t. the $\mathrm{CaF_2}$ valence band edge. E_R always refers to the state with the more positive charge, the relaxation energy of the other state is then given by E_R/R^2. (taken from [DWC5])

The charge transition levels (CTLs) together with the relaxation energies as defined in Sec. 3.1.1 are summarized in Tab. 4.3 for each individual charge transition. Note that only transitions were considered during which only one electron or hole is transferred. The reason is that transitions including multiple charges, e.g. 0/+2, typically are associated with larger barriers [DWJ6, 192], which is also reflected in the considerably higher relaxation energies compared to single-charge transitions. Furthermore, such many-particle processes naturally have a small capture cross section. Therefore, even though a direct transition would be thermodynamically optimal, this process is generally still more likely to occur via multiple steps, e.g. $0 \Rightarrow +1 \Rightarrow +2$ instead of $0 \Rightarrow +2$.

The impact of charge trapping on the device reliability is mostly governed by the energetic alignment between the trap level of the defect and the carrier reservoirs provided by the band edges of the semiconductor. Fig. 4.23 shows the calculated CTLs for both defects considered in this work together with the band edges of silicon and commonly used 2D semiconductors such as $\mathrm{MoS_2}$. As can be seen, for a $\mathrm{Si/CaF_2/MoS_2}$ stack as depicted in Fig. 4.20, the $0/+1$ level of the $\mathrm{Si_i}$ is aligned with the valence band of the Si back-gate, while for the the $\mathrm{Si_{Ca}}$ the $-1/-2$ and $0/-1$ level are aligned with the conduction band of $\mathrm{MoS_2}$ and Si back-gate respectively. Hence charge trapping at those defects could provide an explanation for the comparatively low, but still existing hysteresis observed in such stacks. Based on the band diagram in Fig. 4.23 we would predict more severe effects of charge trapping for stacks including $\mathrm{WSe_2}$ or $\mathrm{MoSe_2}$ due to the additional alignment between the valence band of these semiconductors and the $\mathrm{Si_i}$ $0/+1$ level. Note that the $\mathrm{Si_i}$ $-1/-2$ level is also close to the conduction band of the back-gate. However, since the $0/-1$ level is higher up, the defect is initially in state 0 and going to state -2 requires two electrons being transferred at once, which is very unlikely. Hence the $-1/-2$ level of this defect is considered inactive.

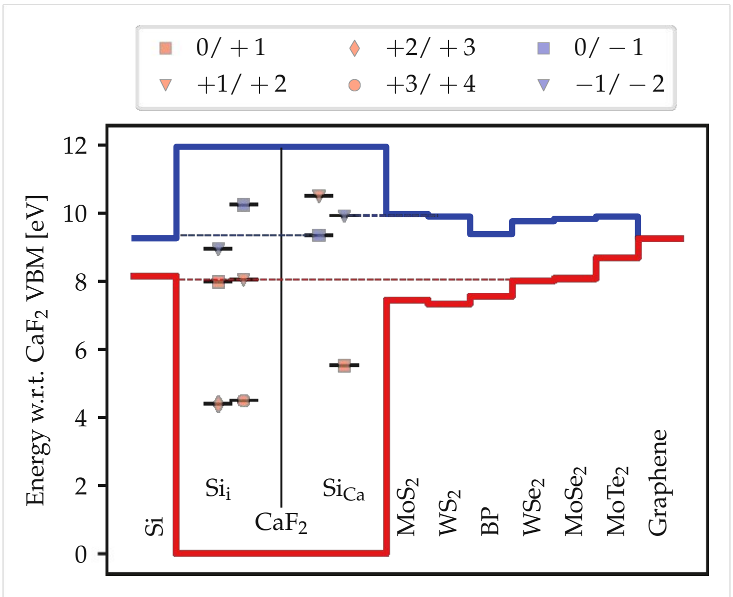


Figure 4.23. Trap levels of the $\mathrm{Si_i}$ and $\mathrm{Si_{Ca}}$ defects within the bandgap of $\mathrm{CaF_2}$ together with the band edges of silicon and commonly used 2D semiconductors. Alignments between band edges and active defect states are indicated with a dashed line. The band offsets were taken from [9]. (taken from [DWC5])

Part II

Device Modeling

5

Compact Modeling Framework – Comphy

This chapter is based on works previously published in [DWJ10].

Building on the fundamentals of charge trapping and nonradiative transitions discussed in Ch. 3, both classical and quantum mechanical charge trapping models are integrated into the compact device simulator *Comphy*. Integrating a microscopic charge trapping model with a device simulator is essential to bridge the gap between atomistic processes and their macroscopic impact on the device, such as threshold voltage shifts. By utilizing a simple yet effective 1-dimensional device model, *Comphy* facilitates the development and testing of novel physical reliability models.

This chapter provides an overview of the framework and the new features developed as part of this thesis, including a novel model for trap-assisted tunneling based on NMP theory, as well as a WKB-based approximation for accelerating quantum mechanical rate calculations, enabling the study of charge trapping at cryogenic temperatures.

5.1 Introduction

Charge trapping at defects in the gate dielectric or at the oxide/semiconductor interface of a MOSFET device causes various reliability challenges in different applications. For instance, charge trapping due to temperature and/or bias stress induces a drift in the device characteristic, commonly referred to as bias temperature instability (BTI) [10, 11, 311], leading to a potential failure in both digital and analog circuits over the course of the device lifetime. In memory applications on the other hand, trap-assisted tunneling (TAT) [312] through oxide defects can limit the data retention times of the stored information [313]. In non-standard technologies, on the other hand, the defect densities are typically higher, making charge trapping already relevant during initial operation and can lead to a clearly visible hysteresis in the transfer characteristics [314, 315].

While the aforementioned phenomena typically exhibit a pronounced temperature activation [316–318], which might lead to the conclusion that charge trapping becomes irrelevant at lower temperatures, it has been demonstrated that charge trapping processes also occur at cryogenic temperatures due to nuclear tunneling. One example is the prominence of positive BTI at 4 K observed in high-κ gate stacks [DWJ16]. Another manifestation of nuclear tunneling is the occurrence of random telegraph noise (RTN) [319] in control circuits for solid-state qubits, which causes a loss of quantum coherence [320, 321].

Besides these detrimental effects, charge trapping can also be leveraged for certain applications, e.g. to store information in charge trap flash devices [DWJ4, 287, 322] or as physical unclonable function by providing a unique device-specific noise signature [323].

These examples emphasize the importance of understanding the microscopic nature of defects as well as the need to accurately describe charge trapping in devices. Modern TCAD simulation packages like Minimos-NT [324, 325] or Sentaurus Device [326] include a multi-state nonradiative multiphonon (NMP) model [17] to incorporate the effects of charge trapping on the device performance and lifetime. However, such models are often too complex to be well calibrated with the available experimental data. Furthermore, an overly elaborate modeling approach can distract from the essential physics of charge trapping, which, in many relevant cases, can be well described within a simplified 2-state NMP model [327], as has been discussed in Sec. 3.1.1. In this spirit, the *Comphy* (compact-physics) framework [26] was recently introduced, a light-weight Python package, which is designed to simulate various charge-trapping related reliability issues based on the 2-state NMP model within a compact 1D device simulator.

While *Comphy* in its original form was intended as a proof-of-concept to demonstrate the feasibility of accurate and efficient reliability modeling of core logic FETs, it is now frequently employed to assess and understand the degradation in experimental devices, ranging from novel memory applications [328] to SiC based power electronics [DWJ15, 329, DWJ17]. Although accurate reliability models are valuable in their own right for device lifetime predictions, the physics-based modeling approach used in *Comphy* also allows to draw conclusions about the atomistic nature of defects responsible for charge trapping by comparison of model parameters to theoretical predictions based on density functional theory (DFT) [DWJ1, 23]. Furthermore, a physical model can provide clues for possible design improvements. For example, based on *Comphy* predictions, a dipole layer introduced at the SiO_2/HfO_2 interface of high-κ gate stacks was demonstrated to improve BTI by suppressing the defect-carrier interaction [330, 331].

Here, recent efforts to further extend the original *Comphy* framework towards a powerful and user-friendly reliability code are summarized, culminating in the publically available *Comphy v3.0* package [DWJ10, 332].

5.2 Models and Features

Compared to *Comphy v1.0*, several newly developed models and features have been added in order to meet the current demands for nanoscale device modeling. In the following a brief of the most relevant new features in *Comphy v3.0* is given, a detailed description is then provided in later sections.

5.2.1 Parameter Extraction Methods

One of the main motivations behind a framework like *Comphy* is to obtain a physically-motivated degradation model based on experimental data points at accelerated stress conditions. A well-calibrated model can then be used to estimate the device lifetime by extrapolating the degradation to typical use conditions. As most gate oxides are amorphous thin films, they show a wide distribution of defect parameters due to the varying local chemical environment of the defects. Furthermore, most degradation experiments are conducted on large-area devices where only the collective response of a whole defect ensemble can be observed. Hence, the underlying defect parameter distributions have to be inferred from these observations. Previous studies mostly assumed Gaussian distributions for the parameters [26, 333] and their mean and sigma values were obtained from a non-linear optimization procedure by fitting experimental data. However, such a procedure can lead to extraction artifacts because it enforces a certain shape of the distribution. Furthermore, the optimization requires a good initial guess, becomes more tedious and requires frequent manual intervention particularly when multiple defect bands are involved, e.g. interacting electron and hole traps as observed in SiC/SiO_2 devices [329, DWJ17, 334]. In order to circumvent these issues, *Comphy v3.0* offers a novel method of parameter extraction named *Effective Single Defect Decomposition* (ESiD) [DWJ8], which allows for a semi-automatic extraction of defect parameters from experimental measure-stress-measure traces without the aforementioned assumptions about their distribution and, in particular, without requiring a good initial guess. This new method will be discussed in Sec. 6.2.3

5.2.2 Trap-Assisted Tunneling

While the most pronounced effects of charge trapping at defects in the oxide are electrostatic shifts manifesting as BTI, RTN and hysteresis, the same mechanism can also facilitate a parasitic gate leakage current by conductance (“hopping”) over defects [312]. Contrary to direct (DT) or Fowler-Nordheim (FN) tunneling [335], the resulting currents exhibit a strong temperature dependence, indicating a charge transfer mediated by a multiphonon process similar to charge trapping. Although several different models have been proposed in literature to describe this phenomenon [19, 336, 337], it is usually treated separately from charge trapping in the context of BTI. *Comphy v3.0* includes a recently developed unified approach [DWJ18] for TAT and BTI, where both are treated on the same footing within the NMP framework, see Sec. 5.5.2. This approach allows to obtain a consistent defect parameter set for the defect-channel interaction which is then transformed to a corresponding parameter set describing the defect-defect interaction.

5.2.3 Charge Trapping at Cryogenic Temperatures

For applications at room temperature and above, the full quantum mechanical NMP model can be reasonably well approximated by its classical analog, the celebrated Marcus theory for charge transfer reactions [216, 217] (see also Sec. 3.3). This model is preferable from a computational point of view, since it only requires calculating the classical transition barrier instead of all vibrational overlaps in NMP theory. However, the classical model predicts a complete freeze-out of charge transfer towards cryogenic temperatures, whereas it is well known that charge trapping and its resulting effects on the device, i.e. BTI and RTN, can be observed even at cryogenic temperatures [DWC1, DWC2, DWJ16, 338]. In this regime, the defect reconfiguration upon charge transfer is no longer temperature activated but rather dominated by *nuclear tunneling*. In order to efficiently model charge trapping under these conditions, e.g. for studying RTN in emerging quantum information applications, a Wentzel–Kramers–Brillouin (WKB) approximation to the full quantum mechanical model has been developed [DWJ9] and incorporated into *Comphy v3.0*. This approach remedies the computational overhead associated with the quantum mechanical description while remaining sufficiently accurate, hence allowing to simulate a whole ensemble of defects in the nuclear tunneling regime.

5.2.4 Coupling to TCAD

While *Comphy* provides a fast way to assess device reliability, an inherent limitation of the employed 1D geometry is related to the charge sheet approximation (CSA) which is used to model the impact of oxide charges on device electrostatics [339, 340]. Even for ideal planar devices, this approximation fails to describe the distribution of threshold voltage shifts caused by the charges of individual traps [DWJ19, DWJ20], furthermore the deviations worsen in the subthreshold regime, where current percolation paths are formed, for example through random discrete dopants (RDD) and random trapped charges [341]. In the case of more complex device geometries such as FinFETs or gate-all-around (GAA) FETs, the modeling of transmission coefficients and the strongly inhomogeneous electric fields and carrier concentrations lead to further inaccuracies in the 1D approximation.

An efficient simulation method to go beyond this 1D approach and to accurately describe the variability-aware reliability of 3D structures is to couple TCAD simulations with *Comphy*. In this approach, the 3D electrostatics are simulated with TCAD for all bias and temperature conditions which are relevant for a given stress scenario to obtain the quantities required by *Comphy* to compute the defect occupancies: The local lattice temperature, the electrostatic potential, the transmission coefficients to all charge reservoirs, as well as the carrier concentrations, energies, and effective masses of the semiconductor and the metal gate. In addition, the perturbative impedance field method (IFM) [342] can be applied on top of the quasi-stationary TCAD solutions to accurately model the impact of the defect charges on the device characteristics instead of using the CSA.

With this approach, the efficient *Comphy* simulations can leverage full 3D TCAD accuracy as recently demonstrated for a reliability and variability aware design technology co-optimization (DTCO) study of FinFET and nanosheet devices, including BTI degradation of ring-oscillators for 10 years of AC operation [343].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

5.3 Device Model

Modern TCAD codes typically solve the drift-diffusion equations [344] on a 2- or 3-dimensional device model to obtain quantities such as the electric field inside the oxide, the carrier concentration within the channel, or macroscopic device characteristics. While this approach is accurate, even for complex device geometries like trench-MOSFETs, FinFETs, or nanosheets, it complicates the development of physical models due to the challenging numerical treatment required.

Comphy, on the other hand, takes a different approach that is more suitable for straightforward implementation of novel reliability models. In *Comphy*, the device geometry is assumed to be 1-dimensional, simulating a planar MOSFET, which significantly simplifies the programming and numerical complexity.

In *Comphy*, all device quantities are derived from the surface potential φ_S at the channel/oxide interface. For uniform channel doping, the total surface charge per unit area in a nMOSFET[1] can be determined by solving the 1-dimensional Poisson equation under the assumption of charge neutrality in the bulk semiconductor, yielding [347]

$$Q_\mathrm{S} = \pm\frac{\sqrt{2}k_\mathrm{B}T}{q_0 L_\mathrm{D,e}}\sqrt{\frac{p_0}{n_0}\left(e^{-q_0\varphi_\mathrm{S}/k_\mathrm{B}T} + q_0\varphi_\mathrm{S}/k_\mathrm{B}T - 1\right) + \left(e^{+q_0\varphi_\mathrm{S}/k_\mathrm{B}T} - q_0\varphi_\mathrm{S}/k_\mathrm{B}T - 1\right)}\,. \tag{5.1}$$

where the positive sign refers to $\varphi_\mathrm{S} > 0$, and the negative sign to $\varphi_\mathrm{S} < 0$. The Debye length for electrons is given by

$$L_\mathrm{D,e} = \sqrt{\frac{k_\mathrm{B}T\varepsilon_0\varepsilon_\mathrm{r,chan}}{n_0 q_0^2}}\,, \tag{5.2}$$

where $\varepsilon_\mathrm{r,chan}$ is the relative permittivity of the channel material and q_0 is the elemental charge. The hole and electron concentrations at thermal equilibriums, p_0 and n_0 respectively, follow from the doping concentrations $N_\mathrm{A}, N_\mathrm{D}$ and the semiconductor bandgap E_g by using basic semiconductor physics [347]. For a given gate bias V_G, it follows that the relation [345]

$$V_\mathrm{G} = \frac{Q_\mathrm{S}(\varphi_\mathrm{S})}{C'_\mathrm{ox}} + \varphi_\mathrm{S} + \Delta E_\mathrm{W,0} + E_\mathrm{F,0} + \Delta V_\mathrm{th}\,. \tag{5.3}$$

must hold. Here, $\Delta E_\mathrm{W,0}$ is the work function difference between the channel semiconductor and the gate metal, $E_\mathrm{F,0}$ denotes the Fermi level in the bulk semiconductor at thermal equilibrium. For a device with channel width W, length L and oxide thickness t_ox, the oxide capacitance per area is given by

$$C'_\mathrm{ox} = \frac{C_\mathrm{ox}}{WL} = \frac{\varepsilon_0\varepsilon_\mathrm{r,ox}}{t_\mathrm{ox}}\,. \tag{5.4}$$

---

[1] Analogous expressions can be derived for pMOSFETs, see e.g. [345, 346].

Charges trapped in oxide defects shift the device electrostatics and are considered by the threshold voltage shift ΔV_{th}. In order to obtain the surface potential φ_{S} as a function of the gate bias V_{G}, (5.3) is solved for φ_{S} numerically. Based on how the oxide charges are treated, there are multiple different operating modes implemented in *Comphy* [345]:

(i) **Oxide charges decoupled from Poisson equation (default)**
The simplest way to consider the impact of oxide charges is to use the charge sheet approximation (CSA) [339, 340]. Within this approximation, it is assumed that a trapped charge in the oxide can be represented by a uniform sheet of charge within the 1-dimensional model. Each trap causes its own threshold voltage shift independently of all the other traps, the total shift is then given by

$$\Delta V_{\mathrm{th}} = -q_0 C_{\mathrm{ox}} \sum_{i}^{\mathrm{traps}} f_i \left(1 - \frac{x_{\mathrm{T},i}}{t_{\mathrm{ox}}}\right) , \tag{5.5}$$

where f_i and $x_{\mathrm{T},i}$ are the occupancy and the position of trap i relative to the channel/insulator interface, respectively. In this operating mode, ΔV_{th} is *not* considered when solving (5.3). Instead, the device electrostatics only needs to be calculated once without oxide charges. The effect of the oxide charges is then calculated in a postprocessing step. This scheme is used per default due to its vast performance benefits. However, it is only valid, if the oxide charges can be treated as a small perturbation. This is usually the case for established technologies, where the concentration of electrically active oxide defects is small.

(ii) **Non-self-consistent Poisson Scheme (NCP)**
In this scheme, the oxide charges at the first time step are calculated within the decoupled scheme described above. However, in each consecutive timestep the previous oxide charge is considered through an effective gate bias

$$V_{\mathrm{G}}^{\mathrm{eff}}[i] = V_{\mathrm{G}}[i] - \Delta V_{\mathrm{th}}[i-1] \, . \tag{5.6}$$

This scheme takes into account the reduction of the oxide field due to the trapped charges. However, it requires a rather dense sampling in the time-domain in order to be numerically stable.

(iii) **Self-consistent Poisson Scheme (SCP)**
The most accurate way to account for the effect of oxide charges on the device electrostatics is to include them directly when solving (5.3). However, since the oxide charges themselves depend on the device electrostatics, this approach requires an additional outer loop to recalculate the updated oxide charges based on the new charge capture and emission rates, making the self-consistent potential (SCP) scheme computationally expensive. As a result, it is typically only used to verify the validity of simpler approximations, see for instance Fig. 6.6.

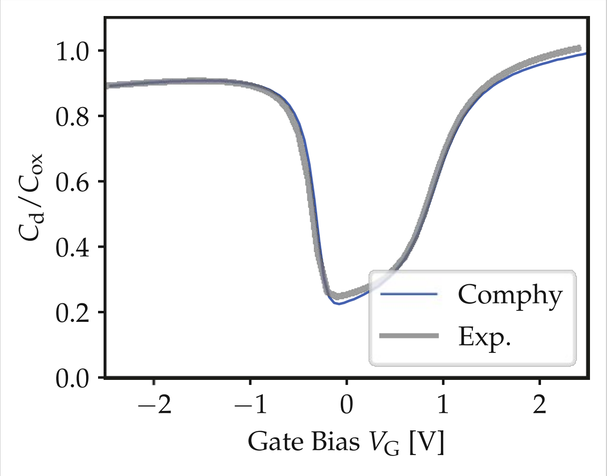


Figure 5.1. Capacitance-voltage (CV) curve recorded for a SiON pMOS device [DWJ8] at a temperature of $T = 100°C$ with frequency $f = 100\,\mathrm{kHz}$. The measured CV curve can be used to calibrate the electrostatics of the device model. As can be seen, there is good agreement between the measurements and the device model, indicating a well-calibrated model usable for accurate defect parameter extractions.

Once the surface potential φ_S has been calculated for a given gate bias V_th and temperature T, other important quantities like the Fermi level E_F or the carrier concentrations n, p at the interface easily follow [345]. These quantities are required to determine the charge transition rates between the device channel and the oxide defects as will be discussed in Sec. 5.4.3.

In order to obtain accurate defect parameters from experimental data using *Comphy*, the device model has to be carefully calibrated to the actual devices. A convenient way to do this are capacitance-voltage (CV) measurements, in which the differential capacitance

$$C_\mathrm{d} = \frac{\mathrm{d}Q_\mathrm{S}}{\mathrm{d}V_\mathrm{G}} \tag{5.7}$$

is measured. The device model can then be calibrated to match the experimental CV curve in order to make sure that essential device parameters like the threshold voltage V_th are accurately respresented in the model. An example for a well-calibrated device model is shown in Fig. 5.1. More details regarding CV measurements can be found in [348].

5.4 Defect Model

Due to the strong impact on electronic devices, charge trapping at oxide defects has been the subject of numerous experimental and theoretical investigations to this day. While RTN is evidently caused by charge trapping [319, 349], a connection between oxide defects and BTI has long been suspected [350, 351] but has only recently been firmly established [14, 15]. Although early theoretical modeling attempts for BTI assumed a reaction-diffusion or reaction-dispersive transport of hydrogen as underlying cause [352–354], elaborate recent measurements show that such a mechanism is incompatible with the experimental evidence [22, 355]. In addition, the parameters used in the reaction-diffusion model, e.g. those related to the properties of hydrogen, are in stark contradiction to both independent experiments and theoretical expectations [311]. Instead, a reaction-limited process involving the phonon-mediated charge transfer to oxide defects is now widely accepted as the dominant cause of BTI [14]. This conclusion is mainly based on the pronounced bias- and temperature dependence observed in nanoscale devices [22] as well as the observation that BTI and RTN are two manifestations of the same microscopic mechanism at shorter time scales [16]. At longer time scales additional interface defects are created by a gate-sided hydrogen release mechanism [15]. Furthermore, recent atomistic studies based on DFT support the involvement of defects in BTI for SiO_2 [23, 25, DWJ8] (see also Sec. 6.3.1) and HfO_2 [356], the most widely used gate dielectrics.

While the essential physics of charge trapping and NMP theory have been discussed already in Ch. 3, here, its application to device models, its limitations for reliability modeling as well as approximations necessary for its implementation in *Comphy* are discussed.

5.4.1 Defect States

In order to model the dynamics of a defect in a device, details about the atomic configurations and the pathways between them are often too complicated to be dealt with explicitly. Following a certain transition, a defect typically relaxes into its new equilibrium configuration in a matter of picoseconds [357]. This implies that the properties of a defect and its possible future pathways only depend on the current defect state. In other words, the defect can be treated as a memory-less system. Using this assumption, the defect states and the transition rates between different states form a Markov chain. All the defect physics, i.e. temperature and bias dependence, is encoded in the transition rates k_{ij} from an initial state i to a final state j. The probabilities P_j of finding a defect in a certain state at a given time evolve according to the master equations

$$\frac{\mathrm{d}P_j}{\mathrm{d}t} = \sum_{i \neq j} k_{ij} P_i(t) - k_{ji} P_j(t) \quad \text{with} \quad \sum_i P_i(t) = 1\,. \tag{5.8}$$

This set of coupled linear differential equations can be solved analytically for an arbitrary number of interacting states using the matrix exponential approach [358].

Time-dependent defect spectroscopy (TDDS) [359] and RTN experiments [17] have unambiguously shown that defects can not only undergo charge transitions, but also relaxations to meta-stable configurations within the same charge state. Prime examples for the existence of such meta-stable states are observations of anomalous RTN [188] and so-called switching traps [360], which show an unusual bias-dependence in their emission time constants, as explained in Sec. 5.4.2. The existence of these states was also established using paramagnetic resonance studies [275, 360] and formed the basis for the development of the 4-state NMP model [17] (see Sec. 1.3) currently implemented in commercial TCAD software like Minimos-NT or Sentaurus Device. While this model can

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

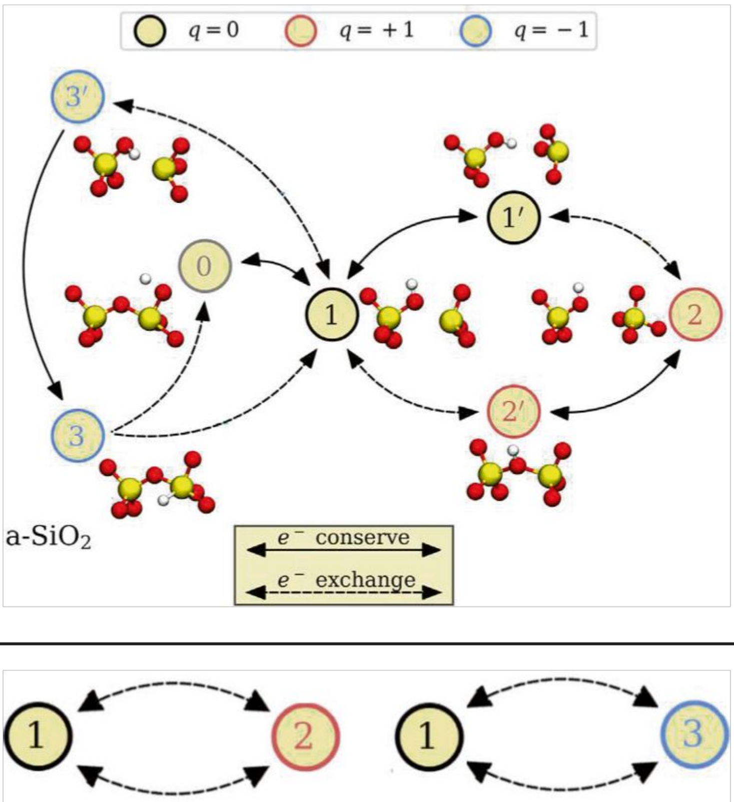


Figure 5.2. Top: Full state diagram of the amphoteric hydroxyl E′ defect in SiO_2. A single amphotheric defect can capture either electrons or holes by going through its various states. **Bottom:** Simplified 2-state models for hole and electron trapping used in *Comphy*. Note that in this approximation electron and hole traps become independent. (taken from [DWJ10], originally published in [DWJ6])

give a comprehensive description of hole trapping in Si/SiO_2 devices during negative BTI (NBTI) stress, treatment of the less pronounced electron trapping during positive BTI (PBTI) was still lacking. As recently has been demonstrated [DWJ6], defects in amorphous SiO_2 are amphoteric, i.e. they can act as both hole and electron traps. The resulting 7-state diagram is depicted in Fig. 5.2 (top) for the case of the hydroxyl-E′ (HE) defect. Compared to the previous 4 states shown in Fig. 4.14, the negatively charged states $3, 3'$ as well as the neutral H interstitial (state 0) are added. Although the defect dynamics in such a model are quite complex, its behavior would still be described by (5.8). However, for a proper parameterization this model would require extensive experimental data based on single-defect characterizations, which is exceedingly hard to obtain.

For this reason, *Comphy* condenses the defect dynamics into effective 2-state models for electron and hole traps separately, as shown in Fig. 5.2 (bottom). Such a simplified model accurately captures the essential physics of charge trapping in most scenarios [327]. Furthermore, this model can be easily parameterized with only a few extended measure-stress-measure (eMSM) [361] sequences using the novel ESiD approach implemented in *Comphy v3.0*, as will be discussed in Sec. 6.2.3 and Sec. 6.3.1. Besides these advantages, (5.8) has the simple analytical solution

$$P_i(t) = P_i(\infty) + (P_i(0) - P_i(\infty)) \cdot \exp\left(-\frac{t}{\tau}\right) \tag{5.9}$$

with the time constant $\tau = 1/(k_{ij} + k_{ji})$ and the equilibrium solution $P_i(\infty) = \tau k_{ji}$. This is exploited to efficiently evaluate the occupancies of defects analytically for arbitrary V_G and T profiles. Note, however, that this approach is only valid for non-interacting defects, an assumption that must be dropped in fully self-consistent calculations or to describe phenomena like multi-trap TAT, see Sec. 5.5.2 for more details.

5.4.2 Limitations of the 2-State Defect Model

While the effective 2-state model is well-suited for reproducing the averaged response of a large ensemble of defects, it is important to consider its limitations. The following subsections summarize the features introduced by additional states that cannot be captured by the 2-state model. The relevance of these effects for a specific application can be evaluated by comparing results to the full 4-state model available in TCAD solutions.

Metastable State 2′

Although it has long been known that charge transitions between states 1 and 2 in Fig. 5.2 (top) occur via a metastable state 2′ [275], the necessity of including this transition into the model was originally deduced from the stronger temperature activation of the emission time compared to the capture time constant [362]. Later studies demonstrated that state 2′ also leads to a frequency dependence of the capture time constant since at higher frequencies back-transitions from 2′ to 1 slow down the overall charge capture [165].

Switching Traps

Based on TDDS measurements, oxide defects can be distinguished based on their behavior during charge emission. On the one hand, fixed charge traps exhibit a rather constant emission time τ_e as a function of V_G, whereas switching traps show a pronounced V_G dependence, in particular they show a steep decrease of τ_e below V_th [17, 22] as shown in Fig. 5.3. This behavior can be explained by the existence of an additional neutral state 1′, see Fig. 5.2 (top), leading to the well-established 4-state model [17, 22] for hole trapping and NBTI. After trapping a hole via the path 1 ⇒ 2′ ⇒ 2, there are two possible pathways to release the hole [108]. For a fixed trap, the transition 2 ⇒ 1′ is unfavorable, hence the hole can only be emitted via the pathway 2 ⇒ 2′ ⇒ 1. However, this process will be limited by the thermal barrier between the two positive states 2 and 2′ and hence will be bias independent. On the other hand, a switching trap can also release its hole via the transition 2 ⇒ 1′, which is strongly bias dependent as it involves another NMP transition.

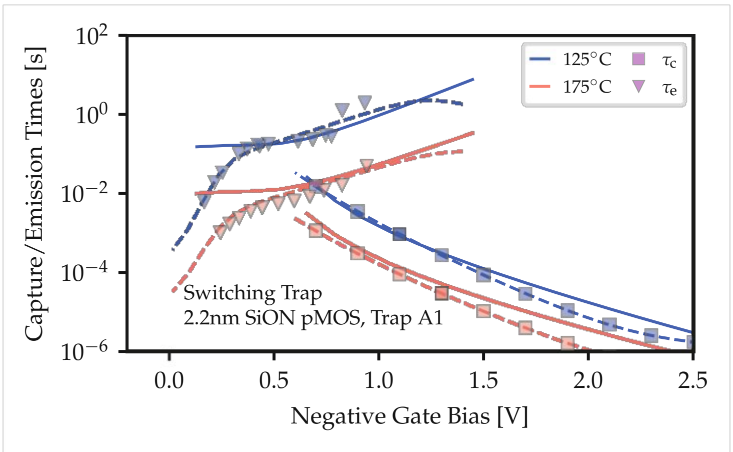


Figure 5.3. 2-state **(solid)** and 4-state **(dashed)** model fits to the capture and emission times of a switching trap obtained from TDDS in [25]. While most features can be described by the simpler 2-state model, the rapidly decreasing emission times of this switching trap below the threshold voltage ($V_G \approx -0.5\,\mathrm{V}$) cannot be explained within the 2-state model. (adapted from [DWJ8])

Precursor Activation

Another interesting property of oxide defects observed in TDDS experiments is their volatility: they can disappear without a trace just to randomly reappear weeks or months later with exactly the same properties [363]. This volatility was interpreted as a transition into the precursor state 0 in Fig. 5.2, e.g. by releasing hydrogen from the defect site. The hydrogen is then free to move around, potentially activating new defects nearby via a transition $0 \Rightarrow 1$, or even depassivating Si-H bonds at the interface. Note that contrary to the assumptions of the reaction-diffusion model, the diffusivity of neutral H is very high [311]. Thus, H moves around very quickly, not limiting any subsequent reactions. An interesting consequence of the H released in the insulators is that it is available to trigger further reactions. In fact, it has long been known that the Si-H bond is too strong to easily release its H and that the reaction $\mathrm{Si{-}H} + \mathrm{H} \Rightarrow \mathrm{Si}^\bullet + \mathrm{H}_2$ is the only feasible reaction [311, 364–366]. Note that the depassivation could also proceed via a proton instead of neutral hydrogen [367]. However, the hole-induced depassivation as discussed in Sec. 4.1 is not possible under normal operating conditions, since for this process the hole has to come from the SiO_2 valence band. Assuming that hydrogen is trapped at the gate side in HE centers, they would mostly be positively charged (state 2). Upon application of a bias, they could be neutralized (transition to state 1) or release their H (transition to state 0). This idea led to the gate-sided hydrogen release model [15] and results in a coupling between charge trapping in the oxide and interface trap creation [368]. While providing a physical explanation, this model in its current stage is computationally demanding and suffers from numerical instabilities, hence in *Comphy* the creation of interface states is modeled via an empirical double-well barrier [26].

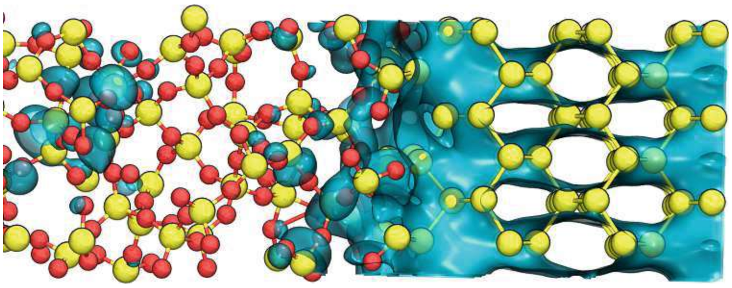


Figure 5.4. DFT calculations of an atomistic Si/SiO$_2$ interface. **Left:** Localized defect state in the SiO$_2$. **Right:** Delocalized wavefunction in the silicon substrate penetrating into the oxide. (taken from [DWJ10])

5.4.3 Transition Rates

While mathematically the defects are modeled as Markov chains, the transition rates k_{ij} containing the physical details within this model still have to be determined. As indicated in Fig. 5.2, there are two fundamentally different types of transitions. First, there are so-called thermal transitions (solid lines) between states with the same charge. In such cases, the transition rates are determined by a reaction barrier and can be evaluated within the classical transition state theory, see Sec. 2.3.2. Since these transitions are not treated explicitly in the simplified *Comphy* model, we will not discuss them further here and refer the interested reader to the literature [109]. The second type of transitions involves a charge transfer event typically accompanied by the emission or absorption of multiple phonons from the environment. These transitions are governed by the NMP formalism [23, 164, 174, 186, 190] as discussed in Sec. 3.1.1.

In *Comphy*, a model similar to (3.41) is used to describe the transition rates quantum mechanically. However, except for cryogenic applications, a classical model in the form of (3.48) is employed to describe charge transfer. Unlike the charge transition theory for defects in the semiconductor discussed in Ch. 3, oxide defects are spatially separated from the semiconductor (or gate contact), as shown in Fig. 5.4. This spatial separation has several consequences, as will be discussed in the following.

First, oxide defects responsible for BTI are energetically located either within the valence or conduction band of the semiconductor, cf. Fig. 4.17. As a result, in a Si/SiO$_2$ interface structure shown in Fig. 5.4, the charge state of such a defect cannot be directly controlled using ground-state DFT, since an added carrier would simply delocalize into the Si bands. This presents a challenge for evaluating the coupling constant V_c. Recent attempts have been made to evaluate V_c from first-principles calculations by inducing an avoided crossing between the defect state and a semiconductor band edge using an electric field [369, 370]. However, this methodology probes the diabatic coupling term rather than the electron-phonon coupling, which likely dominates in this case. In this work, the coupling is instead described by a capture cross section of

$$\sigma_0 = 1 \times 10^{-15}\,\mathrm{cm}^2\,, \tag{5.10}$$

which is consistent with the first-principles findings for hole trapping at the Si−H bond in SiO_2, as presented in Sec. 4.1. This value is also in agreement with recent experimental findings that probed charge trapping at fast interface traps using charge pumping [20].

Another effect caused by the spatial separation of the defect wavefunction and the wavefunctions of the semiconductor bands is an attenuation of the coupling strength with the distance x_T of the defect to the interface. This is due to the carrier wavefunction exponentially decaying into the oxide. In this compact modeling approach, this is taken into account by modifying the capture cross section with a tunneling factor $\vartheta_\mathrm{WKB}(E)$, see also Sec. 5.5.1. Here, E denotes the energy of the carrier in the semiconductor band before getting trapped at the oxide defect.

This leads to a classical transition rate of the form [345]

$$k_{ij}^\mathrm{cl}(E) \propto v_\mathrm{th}\sigma_0\theta_\mathrm{WKB}(E)\exp\left(-\frac{\varepsilon_{ij}(E)}{k_\mathrm{B}T}\right) , \tag{5.11}$$

with $v_\mathrm{th} = \sqrt{k_\mathrm{B}T/m^*}$ being the thermal velocity of the carriers. Note that the carrier energy E not only influences the WKB factor, but also the relative positions of the defect PECs and thus the classical barrier ε_{ij}, see Fig. 5.5. Finally, (5.11) has to be integrated over the whole conduction and valence band of the substrate semiconductor in order to obtain the full transition rates. However, this integration can only be done numerically and is hence associated with large computational costs. To circumvent this, the so-called band-edge approximation [345] is employed, which assumes that all the available carriers are located exactly at the conduction or valence band edge. Under this assumption the classical transition rates can be expressed analytically and, in the case of an electron trap interacting with the semiconductor conduction band edge, are given by

$$k_{ij}^\mathrm{cl} = n v_\mathrm{th,n}\sigma_{0,n}\vartheta_\mathrm{n}\mathrm{e}^{-(\varepsilon_{ij}-E_\mathrm{F}+E_\mathrm{CB})/k_\mathrm{B}T} \tag{5.12a}$$

$$k_{ji}^\mathrm{cl} = n v_\mathrm{th,n}\sigma_{0,n}\vartheta_\mathrm{n}\mathrm{e}^{-\varepsilon_{ji}/k_\mathrm{B}T} . \tag{5.12b}$$

Note however, that for some special cases like the weak electron-phonon coupling regime, the correct value of the classical barrier ε_{ij} can deviate from the definition given in Fig. 5.5. Details about this intricacy can be found in the original *Comphy* paper [26], in particular Tab. 4. For the fully quantum mechanical case see also [DWJ9].

5.4.4 Defect Parameters

In order to simulate the threshold voltage shift ΔV_th caused by a (partially) charged defect ensemble, the PECs of the individual defects have to be parameterized. As depicted in Fig. 5.5, within the harmonic approximation the PECs of the two defect states can be described by

$$U_i(Q) = c_iQ^2 \quad \text{and} \quad U_j(Q) = c_j(Q-\Delta Q)^2 + \Delta E . \tag{5.13}$$

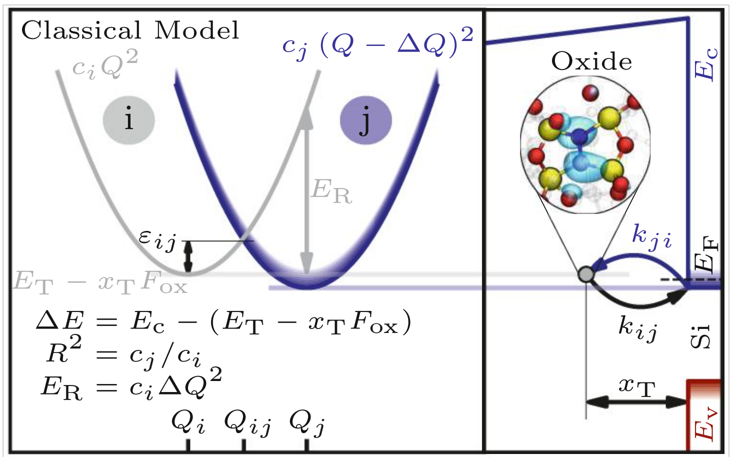

Classical Model
$c_i Q^2$
$c_j\ (Q - \Delta Q)^2$
i
j
E_{R}
ε_{ij}
$E_{\mathrm{T}} - x_{\mathrm{T}} F_{\mathrm{ox}}$
$\Delta E = E_{\mathrm{c}} - (E_{\mathrm{T}} - x_{\mathrm{T}} F_{\mathrm{ox}})$
$R^2 = c_j / c_i$
$E_{\mathrm{R}} = c_i \Delta Q^2$
Q_i Q_{ij} Q_j
Oxide
E_{c}
k_{ji}
E_{F}
k_{ij}
Si
x_{T}
E_{v}

Figure 5.5. Classical (high temperature) limit of the NMP model. Here, the transition rate is only determined by the classical barrier ε_{ij}. The PECs and their relative position are defined by the trap level E_{T}, the relaxation energy E_{R} and the curvature ratio R. Additionally, the position of the trap (x_{T}) together with the applied electric field F_{ox} across the oxide introduces a bias-dependent energy offset ΔE. (taken from [DWJ10])

While the PECs can be directly parameterized by the curvatures $c_{\{i,f\}}$ at their respective minima, it is more common to use the relaxation energy

$$E_{\mathrm{R}} = U_i(\Delta Q) - U_i(0) = c_i \Delta Q^2 \tag{5.14}$$

and the *curvature ratio* $R = \sqrt{c_j/c_i}$ instead, cf. Sec. 3.1.1. The energetic offset ΔE between the two parabolas is determined by the intrinsic trap level of the defect E_{T} as well as the applied electric field F_{ox} in the oxide. For electron trapping from the semiconductor conduction band, the offset is then given by

$$\Delta E = E_{\mathrm{C}} - E_{\mathrm{T}} - x_{\mathrm{T}} F_{\mathrm{ox}} \tag{5.15}$$

for a charge-free dielectric. Note that such a non-self-consistent treatment is only valid if the charges accumulating in the oxide defects cause only a small shift in the device electrostatics. Otherwise a self-consistent approach solving the Poisson equation including the oxide charges is necessary.

For the defect PECs and their relative position with respect to each other to be uniquely defined, one has to specify the parameter tuple

$$\mathbf{p} = (E_{\mathrm{T}}, E_{\mathrm{R}}, R, x_{\mathrm{T}}, \Delta Q)\ . \tag{5.16}$$

Note that in the classical limit the transition rates are only determined by the energetic barrier defined by the PEC crossing point. Hence, the displacement ΔQ along the configuration coordinate is immaterial in this case. However, in a quantum mechanical treatment, ΔQ is an important defect parameter determining the likelihood of nuclear tunneling, see for instance Fig. 5.9.

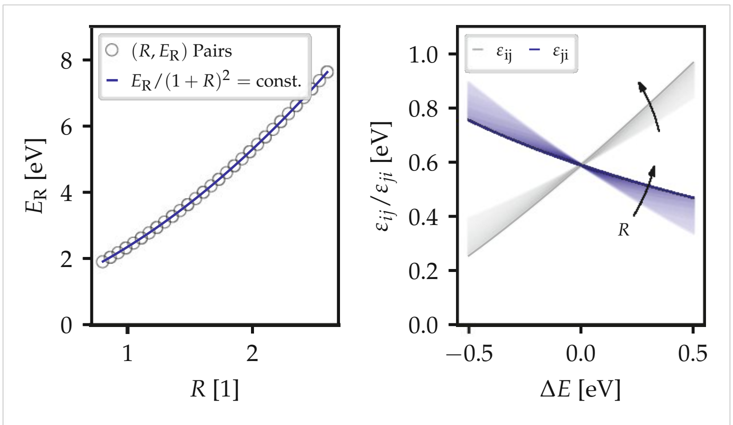


Figure 5.6. As long as (R, E_R) pairs are moved along the curve $E_R/(1+R)^2 = \text{const.}$ **(left)**, the corresponding charge transition barriers for these pairs are very similar **(right)**. Experimentally these barriers are nearly indistinguishable. Therefore, in order to extract defect parameters experimentally, one has to fix either R or E_R. (taken from [DWJ10], originally published in [DWJ8])

Parameter Correlations

While such a tuple $\mathbf{p}$ is sufficient to determine the transition rates of a defect and hence also its capture (τ_c) and emission times (τ_e), for the inverse problem, i.e. extracting this tuple from a set of experimental ΔV_{th} curves, one has to check for cross-correlations between the model parameters. Framed differently, the question is, whether or not a *unique* set of defect parameters can be extracted from the available experimental data obtained with electrical characterization methods.

Considering only the classical limit of the NMP model, a non-linear cross-correlation between the parameters R and E_R has been identified [DWJ8]. Within the harmonic approximation, the classical barrier can be calculated analytically and is given by [17]

$$\varepsilon_{ij}(\Delta E) = \frac{E_R}{(R^2-1)^2}\left(1 - R\sqrt{1 + \frac{\Delta E}{E_R}(R^2-1)}\right)^2 . \tag{5.17}$$

A Taylor expansion around $\Delta E = 0$ then yields [362]

$$\varepsilon_{ij}(\Delta E) = \frac{E_R}{(1+R)^2} + \frac{R}{1+R}\Delta E + \mathcal{O}(\Delta E^2) . \tag{5.18}$$

As illustrated in Fig. 5.6, while keeping the zeroth order term in this expansion fixed, R and E_R can vary over a wide parameter range resulting in very similar barriers. However, since the barrier is the defining feature for the electrically measurable ΔV_{th} response in the classical model, this implies that simultaneously extracting both R and E_R uniquely from experiments is challenging [DWJ8].

In an earlier work [26], a rather high curvature ratio of $R = 2.59$ together with a mean relaxation energy of 7.95 eV have been reported for hole traps in SiON. However, such values seem unphysical and disagree with theoretical DFT studies on defects in amorphous silica, which typically predict values in the range $R = 0.8{-}1.2$ and $E_\mathrm{R} = 2{-}4\,\mathrm{eV}$ [23]. The cause for these unphysically high relaxation energies is precisely the (R, E_R) correlation, which prohibits extracting a unique parameter set from electrical characterizations [DWJ8]. In order to obtain a unique, physical parameter set, it is therefore recommended to restrict the simulations to the so-called *linear coupling regime* defined by $R = 1$ [109, 371]. This choice is supported by extensive theoretical DFT studies on oxide defects in a-SiO$_2$ [23]. Besides this empirical justification, one could also make an argument based on the curvatures of the PECs being related to the phonon frequencies in the material. Since in an amorphous material many different phonon modes will be involved in a charge transition, the phonon frequency changes in different charge states will mostly average out, resulting in an effective curvature ratio of $R = 1$.

While the (R, E_R) correlation has been demonstrated for the classical 2-state NMP model, other cross-correlations might still be present in the quantum mechanical model. For example, it is suspected that a correlation between ΔQ and E_R exists, however, this is subject to further investigations.

5.5 Gate-leakage Currents

Besides charge-trapping which leads to considerable V_th shifts in MOSFETs, as described in the previous sections, time-zero gate-leakage currents pose a severe threat for power dissipation and to gate oxide reliability. Within *Comphy*, an efficient modeling approach is incorporated to enable the computation of these leakage currents that arise from the same inelastic charge transfer reactions driving BTI, but without additional parameters compared to the BTI simulations [DWJ18]. Depending on the gate stack material and defect properties within the insulating layers, either band-to-band tunneling or defects acting as charge transition centers between gate and channel dominate the leakage mechanism. As shown in Fig. 5.7, the trap-assisted tunneling (TAT) component appears as either a single- or multi-step process.

5.5.1 Tsu-Esaki Model

The most commonly used approximation of the current density due to tunneling of charge carriers between two carrier reservoirs separated by an energetic barrier is given by the Tsu-Esaki formalism [335]. For electron tunneling the current density is given as

$$J_\mathrm{TE,e} = \frac{4\pi m_\mathrm{e} q_0}{h^3} \int_{E_\mathrm{CB}}^{\infty} \vartheta_\mathrm{WKB}\left(E\right) N_\mathrm{e}\left(E\right) \mathrm{d}E \qquad (5.19)$$

with the effective electron mass m_e in the semiconductor channel and the supply function

$$N_e\left(E\right) = k_\mathrm{B} T \ln \left(\frac{1 + \exp\left(-\frac{E - E_\mathrm{F1}}{k_\mathrm{B}T}\right)}{1 + \exp\left(-\frac{E - E_\mathrm{F2}}{k_\mathrm{B}T}\right)} \right). \qquad (5.20)$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

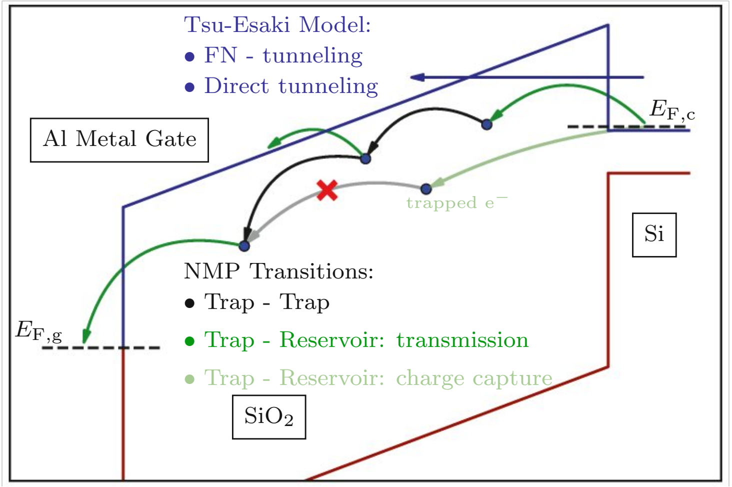


Figure 5.7. The different tunneling mechanisms that typically appear in a MOS gate stack are shown schematically. The band to band tunneling mechanisms (blue) are included in the well-known Tsu-Esaki model as implemented in *Comphy*, while all trap-assisted tunneling reactions (green and black) are described by two-state NMP transitions. (taken from [DWJ10], originally published in [DWJ18])

accounting for the Fermi-Dirac distribution of the carriers in both electrodes. The tunneling probability is calculated using a WKB approximation by

$$\vartheta_{\mathrm{WKB}}\left(E\right) = \exp\left(-\frac{4\pi}{h}\int_{x_1}^{x_2}\sqrt{2m_{\mathrm{e,diel}}\left(W\left(x\right)-E\right)}\,\mathrm{d}x\right), \tag{5.21}$$

with the tunneling effective mass in the dielectric layer $m_{\mathrm{e,diel}}$ and the energetic barrier $W\left(x\right)$. Depending on the shape of $W\left(x\right)$, i.e. triangular or trapezoidal, (5.19) evolves to the well-known Fowler-Nordheim or direct tunneling formulas.

5.5.2 Efficient Trap-Assisted Tunneling Current Computation

Contrary to charge capture at oxide defects in large-area devices, the calculation of charge transfer between individual defects in amorphous materials is not feasible with a simulation on a equidistant grid spanning the whole defect parameter space. Therefore, for the calculation of the TAT current within an amorphous dielectric, the defects are randomly sampled based on a uniform distribution in a three dimensional volume V_i, as schematically shown in Fig. 5.8. The defect density N_{T} and average defect number $\overline{N}$ thereby *define* $V_i = a^2 t_{\mathrm{ox}} = \overline{N}/N_{\mathrm{T}}$. The actual defect number N is drawn from a Poisson distribution within each simulated V_i. From the three dimensional sampling, the distances d_{T} between the defects can be calculated, while electrostatic quantities are computed in a one-dimensional simulation space, as described in [26]. Defects sampled with unphysically low distances to each other (below 1 nm) are removed from the sample and redrawn. In order to capture the stochastic variation of the simulation results, the simulation can be performed within M simulation volumes in parallel. Here, a Gaussian distribution (see Sec. 6.2.2) is used for the sampling of the energetic defect parameters E_{T} and E_{R} which

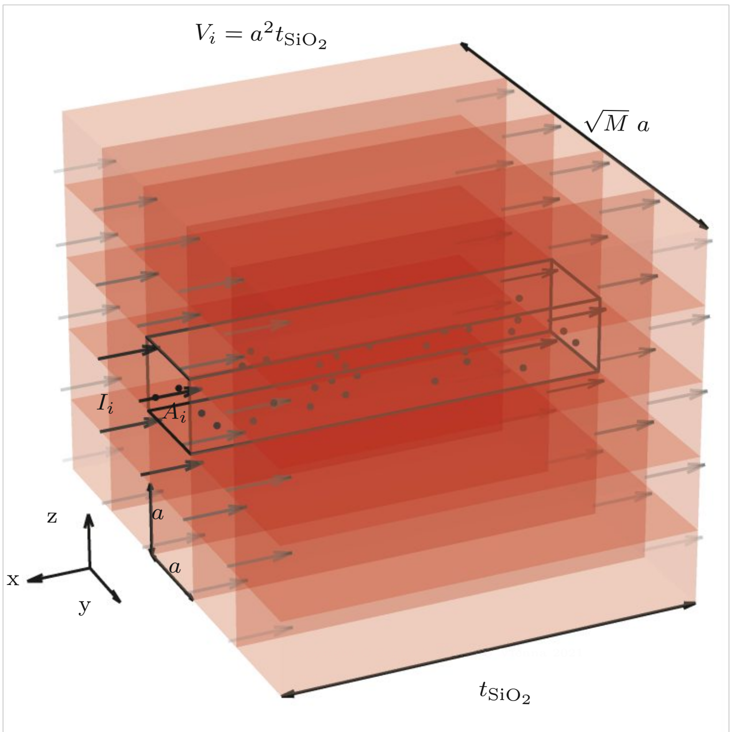


Figure 5.8. The three dimensional spatial simulation domain is divided into M slabs in the dielectric layer. Within each slab a Poisson distributed number of defects N with mean value $\overline{N}$ is sampled throughout the volume $V_i = a^2 t_{\mathrm{SiO_2}}$, with a resulting from the given defect density N_{T}. (taken from [DWJ10], originally published in [DWJ18])

fully determine the classically approximated PECs of the two charge states of a defect within the NMP model. This representation allows for the calculation of the energetic barriers ε_{ij} that determine the classical charge capture and emission rates k_{c} and k_{e}. As described in detail in [DWJ18], considering all possible charge transfer paths, the trap-assisted leakage current can be computed based on a generalized Shockley-Ramo theorem [372] by

$$\begin{aligned} I_{\mathrm{G,TAT}} = \underbrace{C_{\mathrm{ox}} \frac{\mathrm{d}V_{\mathrm{G}}}{\mathrm{d}t}}_{\text{displacement current}} & + \underbrace{q_0 \sum_i^N k_{\mathrm{e},i,\mathrm{gate}} f_i - k_{\mathrm{c},i,\mathrm{gate}} \left(1 - f_i\right)}_{\text{single-TAT current}} \\ & + \underbrace{q_0 \sum_i^N \sum_{j \neq i}^N k_{\mathrm{e},ij} f_i \left(1 - f_j\right) \frac{x_{\mathrm{T},i} - x_{\mathrm{T},j}}{t_{\mathrm{ox}}}}_{\text{multi-TAT current}} \\ & + \underbrace{q_0 \sum_i^N \left[k_{\mathrm{c},i,\mathrm{channel}} \left(1 - f_i\right) - k_{\mathrm{e},i,\mathrm{channel}} f_i\right] \left(1 - \frac{x_{\mathrm{T},i}}{t_{\mathrm{ox}}}\right)}_{\text{charge trapping current (channel)}} . \end{aligned} \tag{5.22}$$

In order to calculate the gate-leakage current, the occupancies f_i of each trap have to be known. They are governed by a set of nonlinear coupled ordinary differential equations [373]

$$\frac{\mathrm{d}f_i}{\mathrm{d}t} = \sum_{j\neq i} \underbrace{k_{ji} f_j (1-f_i)}_{\text{inflow}} - \underbrace{k_{ij} f_i (1-f_j)}_{\text{outflow}} \; . \tag{5.23}$$

Note that this system of equations is fundamentally different to the master equations (5.8), since the hopping rate from trap j to i requires trap j to be occupied and trap i to be unoccupied, resulting in a nonlinear coupling of all occupancies. For this system of equations, no closed solution exists, hence requiring a numerical solution. In *Comphy*, this set of equations is solved using a Newton scheme together with the backward-Euler method for time propagation. Further details can be found in [373]. The Newton scheme used to calculate the occupations in multi-TAT mode leads to a reduced computational efficiency compared to single-TAT mode, whose performance is comparable to a simple ΔV_{th} calculation. However, as shown in [DWJ18] for many applications a multi-step TAT current only becomes relevant at large defect densities N_{T} and for low relaxation energies E_{R} of the conducting defect bands. Due to these prerequisites, single-step TAT already provides efficient and accurate results for the majority of technologies.

It is also important to note that in this context the transition rates between individual traps are calculated rather than the rates between a trap and the channel or gate contact. Therefore, in this case, both traps relax and exchange phonons with the surrounding heat bath simultaneously. As a result, a system of two interacting defects evolves along different PECs compared to interactions involving the channel or gate, which exhibit negligible relaxation. For $R = 1$ and otherwise identical defects, the effective relaxation energy is twice that of a single defect [DWJ18, 373, 374]. This implies that only defects with relatively low relaxation energies can contribute to trap-assisted tunneling (TAT), since the classical transition barrier also scales with E_{R}, as shown in (5.17). This makes polarons a likely candidate for causing TAT leakage currents, as will be demonstrated in Sec. 6.3.3.

5.6 Cryogenic Charge Trapping

While the classical approximation of the 2-state NMP transition rates (5.12) can be used for describing charge transfer kinetics at room temperature or above, it breaks down at cryogenic temperatures. Multiple studies have demonstrated active charge trapping at cryogenic temperatures [DWC1, DWC2, DWJ16, 375], however, the classical limit

$$\lim_{T\to 0\,\mathrm{K}} k_{ij}^{\mathrm{cl}}(T) = 0 \tag{5.24}$$

would predict a complete freeze-out of charge transitions between neutral and charged defect configurations. Therefore, it is necessary to use the full quantum mechanical version of the 2-state NMP model for cryogenic conditions. At these low temperatures, the transition rates are not dominated by the height of the classical barrier ε_{ij} but by the overlap of the vibrational wavefunctions. As temperatures approach the cryogenic regime, the transition

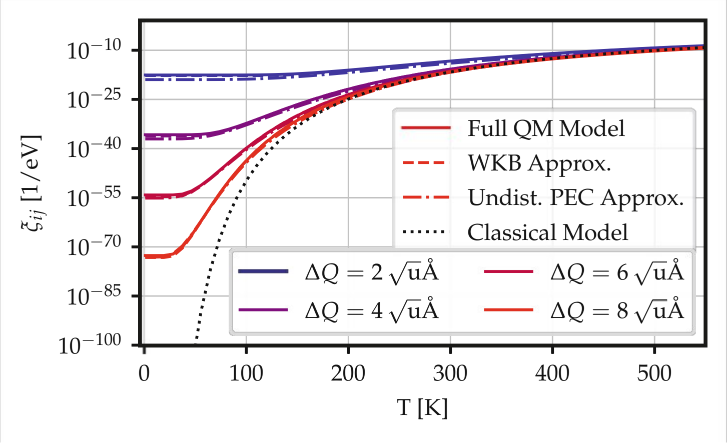


Figure 5.9. Lineshape functions from different approximations for $E_\mathrm{R} = 2.5\,\mathrm{eV}$, $R = 1$ and $\Delta E = 0\,\mathrm{eV}$ but with different displacements ΔQ. The full quantum mechanical model given by (3.41) **(solid)**, the WKB-based model **(dashed)** and the undistorted PEC model **(dash-dotted)** all become temperature independent towards cryogenic temperatures while the classical model (5.12) **(dotted)** freezes out completely. As ΔQ increases, the system exhibits increasingly classical behavior. (taken from [DWJ10], originally published in [DWJ9])

rate is primarily determined by the overlap of the vibrational ground state, since the low thermal energy limits access to excited states. Consequently, the lineshape function, which is proportional to the transition rate, becomes temperature-independent, as illustrated in Fig. 5.9. The transition from one atomic configuration to another at energies below the classical barrier is called nuclear tunneling. Computing the full quantum mechanical transition rate is numerically expensive because it requires calculating the eigenstates of the potential energy surfaces and the corresponding overlap functions, which must then be added up, see e.g. (3.41). This procedure is not feasible for reliability simulations involving thousands of defects. To address this challenge, a numerically efficient model based on the WKB approximation of the vibrational wavefunctions was developed in [DWJ9] and implemented in *Comphy v3.0*.

5.6.1 WKB-based approximation

Within the developed approximation scheme for the 2-state NMP model, the exact vibrational wave functions are replaced by their respective WKB-approximations as shown in Fig. 5.10.

At cryogenic temperatures, the overlaps of the wave functions which are needed for the computation of (3.41) are dominated by the exponentially decaying part of the WKB wave function. This allows to simplify the overlap integrals to a continuous form and approximate them analytically using the stationary phase method [376]. The transition rate k_{ij} then evaluates to [DWJ9]

$$k_{ij}(T) = C(E^*) \exp\left(\frac{-E^*}{k_\mathrm{B}T} + \varphi(E^*)\right) \sqrt{\frac{2\pi}{\varphi''(E^*)}}\,, \tag{5.25}$$

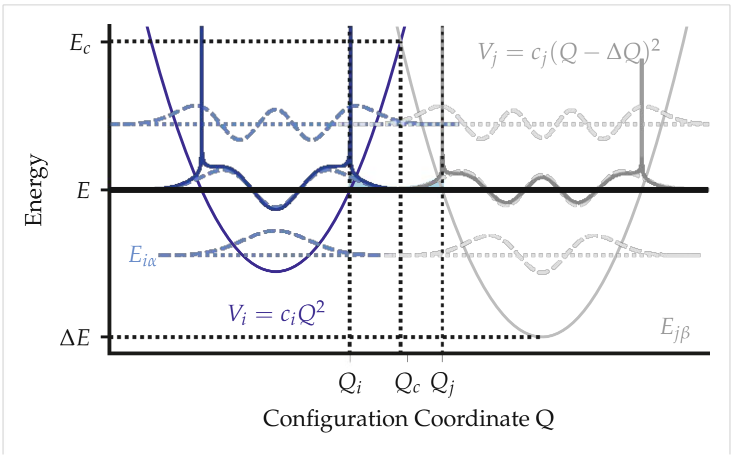


Figure 5.10. The vibrational wave functions **(dashed)** of the potential energy curves representing the neutral and charged states can be approximated using WKB-based wave functions **(solid)**. The overlaps of the wavefunctions are dominated by their exponentially decaying parts which allows the formulation of a simplified expression for the charge transition rates. (taken from [DWJ10], originally published in [DWJ9])

where $\varphi(E)$ is a phase factor given by

$$\varphi(E) = \frac{2}{\hbar}\int_{Q_i}^{Q_c} k_i(Q,E)\mathrm{d}Q + \frac{2}{\hbar}\int_{Q_c}^{Q_j} k_j(Q,E)\mathrm{d}Q \tag{5.26}$$

with the classical momenta

$$k_{\{i,j\}}(Q,E) = \sqrt{2m(U_{\{i,j\}}(Q) - E)} \tag{5.27}$$

for a particle with energy $E < E_c$. The prefactor $C(E)$ contains all non-exponential dependencies on E. While mathematically challenging, the final result shows that the transition happens most likely at an energy $E^* < E_c$, i.e. through tunneling below the classical barrier. This energy E^* can be obtained via the saddle-point method using a numerical optimizer for solving

$$\left.\frac{\mathrm{d}\varphi(E)}{\mathrm{d}E}\right|_{E=E^*} = \frac{1}{k_\mathrm{B}T}. \tag{5.28}$$

This model gives very similar results to the full quantum mechanical (FQM) model while also being computationally much more efficient [DWJ9]. This is shown exemplary in Fig. 5.9, where the lineshape function of the WKB-based approximation and the FQM model are shown for a wide range of temperatures and displacements ΔQ. This example shows the ability of describing the non-freeze-out behavior of the quantum mechanical transition rates compared to the classical model and the experimentally observed temperature independent behavior.

However, due to the required optimization step (5.28) no closed form expression is available which results in a considerably slower computation compared to the classical approximation. For further numerical improvement a model based on the assumption of linear coupling ($R = 1$) is given below.

5.6.2 Closed Form for Linear Coupling

For $R = 1$, the curvatures of the PECs do not change upon charge trapping, implying that the effective phonon frequencies are also the same in both defect charge states ($\Omega_i = \Omega_j = \Omega$). In this case, it is possible to develop a closed form expression for the quantum mechanical transition rates in the strong electron-phonon coupling regime at cryogenic temperatures [377]. Here, the transition rate k_{ij} can be computed as

$$k_{ij}(T) = \omega \frac{\sqrt{2\pi}}{2D\hbar} \Big((1 + \coth\alpha)\, \mathrm{e}^{-E_-/2D^2\hbar^2} - (1 - \coth\alpha)\, \mathrm{e}^{-E_+/2D^2\hbar^2} \Big) \tag{5.29}$$

with $D = \sqrt{\omega E_\mathrm{R}(2n+1)/\hbar}$, $n = 1/(\exp(\hbar\Omega/k_\mathrm{B}T) - 1)$, $\alpha = \hbar\Omega/2k_\mathrm{B}T$ and $E_\pm = \Delta E \pm \hbar\Omega - E_\mathrm{R}$. The phonon frequency Ω is given by

$$\Omega = \frac{\sqrt{2E_\mathrm{R}}}{\Delta Q}. \tag{5.30}$$

As shown in the example in Fig. 5.9, the closed form expression for linear coupling gives the same results as the FQM model over a wide temperature range. Furthermore, computing this analytical expression is comparable to the classical model in terms of computational costs and is thus well suited for efficient reliability simulations with thousands of defects in a device.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek
WIEN Your knowledge hub

6

Defect Identification

After establishing the fundamentals of atomistic and device modeling in the previous chapters, this chapter focuses on integrating these insights with experimental defect characterization to link microscopic processes to observable device degradation. This approach enables the identification of defects responsible for various reliability issues, such as bias temperature instability (BTI), random telegraph noise (RTN), and gate-leakage currents.

The chapter begins with a discussion of the experimental characterization techniques used in this work. This is followed by the introduction of a novel algorithm — the Effective Single Defect Decomposition (ESiD) — which enables the automatic extraction of defect parameters from simple experiments. It will be demonstrated that this new approach performs on par with more complex and time-consuming single-defect characterization techniques.

Finally, the effectiveness of the multiscale modeling approach emphasized throughout this thesis is demonstrated by identifying likely defects causing detrimental effects in different applications, based on experimental data, device modeling, and insights from first-principles calculations.

6.1 Experimental Characterization Techniques

This section is based on works previously published in [DWT1, DWC2].

To identify defects in devices through electrical means, the capture and emission time constants of the defects must be measured. Depending on the device area, different methods are employed. For large-area devices, individual defects cannot be resolved experimentally. Instead, only the collective response of all defects can be measured. In this case, measure-stress-measure (MSM) schemes are typically used, which produce smooth $\Delta V_{\mathrm{th}}(t)$ recovery traces. While these measurements can be conducted relatively quickly, interpreting them is challenging, as they reflect only the aggregated behavior of all defects.

To the contrary, in scaled small-area devices, individual defects can be resolved. Techniques such as random telegraph noise (RTN) analysis and time-dependent defect spectroscopy (TDDS) provide valuable insights by directly measuring the time constants of individual defects, see for instance the discussion in Sec. 5.4.2. However, these studies are typically time-consuming, as the stochastic nature of capture and emission events requires repeated observations to obtain accurate capture and emission times.

The methods used to obtain the experimental data relevant to this work are briefly summarized below. A comprehensive discussion can be found in [348].

6.1.1 Extended Measure-Stress-Measure Scheme

In large-scale devices, individual defects cannot be observed separately due to their minimal impact on the threshold voltage. For such devices, an extended Measure-Stress-Measure (eMSM) technique is typically used [361]. In this approach, the device undergoes repeated stress phases, each followed by a relaxation period during which the transient threshold voltage shift $\Delta V_{\mathrm{th}}(t)$ is recorded. As shown in Fig. 6.1, an eMSM scheme consists of three phases: [346]

(i) **Initial** $I_{\mathrm{D}}(V_{\mathrm{G}})$**:** Before any stress is applied, the initial $I_{\mathrm{D}}(V_{\mathrm{G}})$ characteristic of the device is measured. This initial $I_{\mathrm{D}}(V_{\mathrm{G}})$ is later used to convert a change ΔI_{D} in the drain current to an equivalent shift ΔV_{th} of the threshold voltage.

(ii) **Stress phase:** After measuring the $I_{\mathrm{D}}(V_{\mathrm{G}})$, the device enters the stress phase, in which a constant stress voltage $V_{\mathrm{G,str}}$ is applied at a given temperature T. During this time, traps within the oxide accumulate charges leading to an increase of ΔV_{th}. Note however, that this change is very difficult to observe experimentally during the stress phase, since the device is in strong inversion where ΔV_{th} almost has no impact on the drain current [378].

(iii) **Recovery phase:** Subsequently, the gate voltage is reduced to the recovery level $V_{\mathrm{G,rec}}$, bringing the device into weak inversion. Here, $\Delta V_{\mathrm{th}}(t)$ is recorded via measuring the drain current $I_{\mathrm{D}}(t)$. During this phase, the previously charged defects (partially) release their charge again, leading to the shown recovery trace.

The steps (ii) and (iii) are repeated several times with increasing stress times t_{str} and stress voltages $V_{\mathrm{G,str}}$. From the obtained recovery traces, the parameters of the electrically active defects can be inferred by fitting the theoretical $\Delta V_{\mathrm{th}}(t)$ response of a Gaussian defect band to the experimental data [26], see also Sec. 6.2.2. In Sec. 6.2.3 the novel effective single defect decompositions (ESiD) algorithm is introduced, which can extract the parameter distribution of defects automatically based on multiple recovery traces.

It should be noted that due to a necessary delay t_{d} when switching between stress and recovery phase, very fast charge emission can be missed, as can be seen for instance in Fig. 6.10 (bottom). In order to overcome this issue, so-called on-the-fly methods [379, 380] have been proposed, which attempt to measure ΔV_{th} directly during the stress phase.

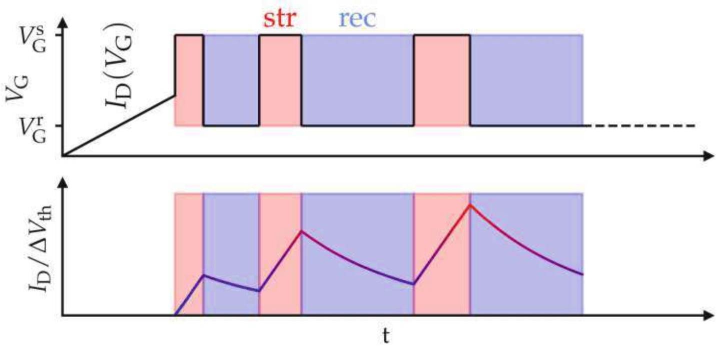


Figure 6.1. Schematic illustration of an extended measure-stress-measure (eMSM) sequence. Initially, an $I_D(V_G)$ characteristic is measured. In the following stress phase, a high gate voltage $V_{G,str}$ is applied, which causes charges to move into the oxide, leading to an increased ΔV_{th} shift. In the subsequent recovery phase, a lower voltage $V_{G,rec}$ is applied and the recovery of ΔV_{th} is determined from measuring I_D and conversion using the initially recorded $I_D(V_G)$. Stress and recovery phases are repeated with progressively increasing stress times t_{str}. (taken from [346])

6.1.2 Random Telegraph Noise Analysis

Analyzing random telegraph noise (RTN) signals [DWC1, DWC2, 381, 382], as illustrated in Fig. 6.2, can be used to measure the characteristics of individual defects in small-area devices. For RTN measurements, the device operates at a constant gate bias (V_G), while the change in threshold voltage, $\Delta V_{th}(t)$, is recorded over a sufficiently long period to capture enough capture and emission events to obtain reliable statistics for the capture and emission times.

After filtering the raw signal, algorithms such as Canny edge detection [383] or Otsu's method [384] can be employed to identify steps (depicted by red lines in Fig. 6.2) and determine when they occur. As shown in the figure, the step heights η follow a Gaussian distribution. On the other hand, the capture and emission times are exponentially distributed for a two-state Markov process [17]. By analyzing RTN signals under varying gate biases and temperatures, the dependence of the time constants on these environmental variables can be determined.

This technique is not limited to simple two-state defects. It can also be applied to characterize multi-state defects with more complex dynamics [385] or to analyze signals composed of multiple defects [386]. In such cases, however, more advanced extraction algorithms, such as the Baum-Welch algorithm [387] for fitting a hidden Markov model to the noise signal, are required.

6.1.3 Time-Dependent Defect Spectroscopy

Similar to RTN measurements, time-dependent defect spectroscopy (TDDS) [388] is a method used to experimentally extract parameters from individual defects. While RTN measurements apply a constant gate voltage V_G, keeping defects in quasi-thermal equilibrium and limiting the measurable range of time constants to $\tau_c \approx \tau_e$, TDDS offers a

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

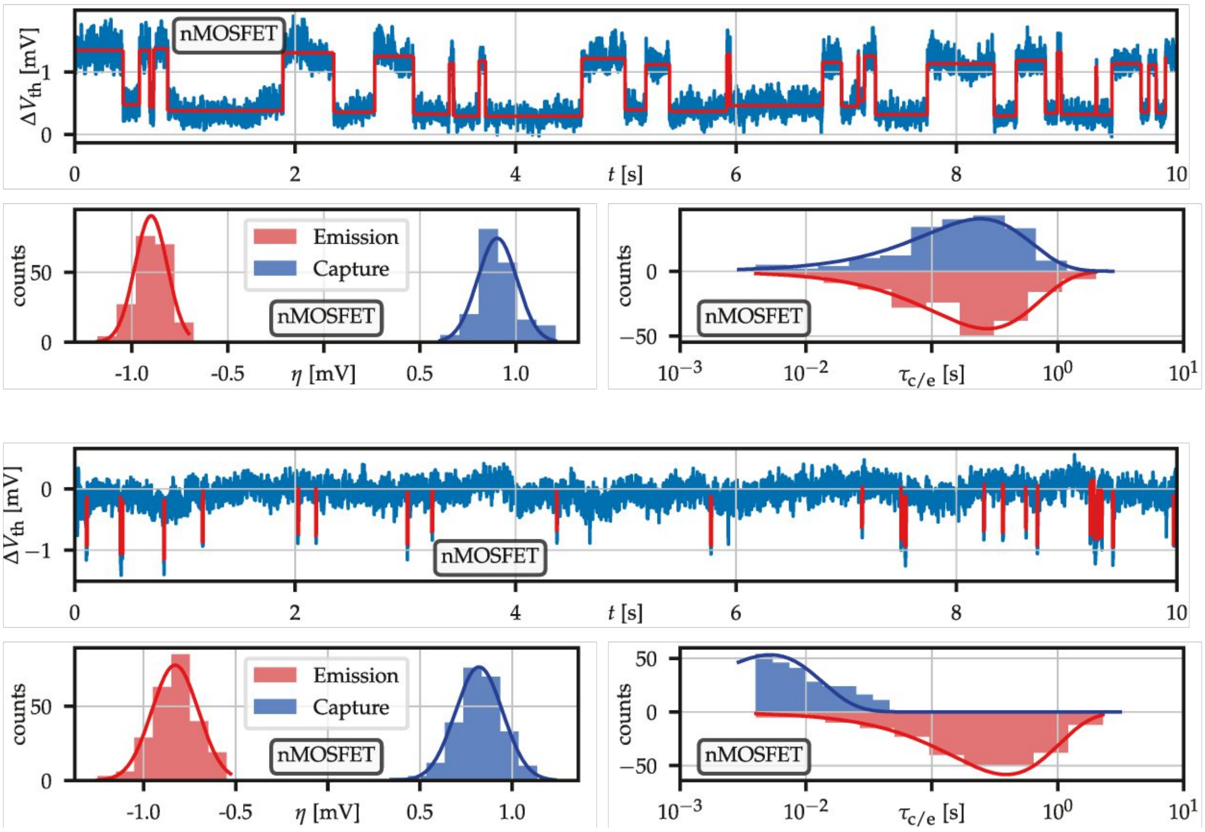


Figure 6.2. RTN analysis for a scaled nMOSFET with SiON dielectric and dimensions $W \times L = 0.22 \times 0.18\,\mu\mathrm{m}^2$ at $T = 9\,\mathrm{K}$. The upper and lower panels show the obtained RTN signal for a gate bias of $V_{\mathrm{G}} = 557.5\,\mathrm{mV}$ and $V_{\mathrm{G}} = 582.5\,\mathrm{mV}$ respectively. The steps (red) were extracted using the Canny edge detection algorithm. The observed step heights η are normally distributed but show two clear peaks with opposite sign for capture and emission events. The time constants $\tau_{\mathrm{c}}, \tau_{\mathrm{e}}$ are exponentially distributed, as can be expected from a 2-state Markov process. While the upper panel shows a situation where $\tau_{\mathrm{c}} \approx \tau_{\mathrm{e}}$, only a slight variation of the gate bias leads to $\tau_{\mathrm{c}} \ll \tau_{\mathrm{e}}$. (taken from [DWC2])

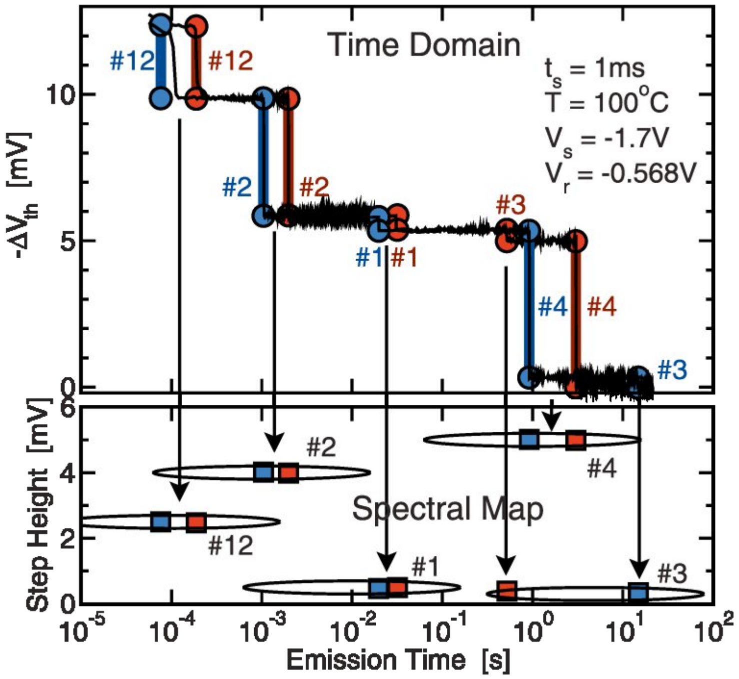


Figure 6.3. Typical results of a TDDS experiment. The device is stressed multiple times and the corresponding emission traces are recorded **(top)**. The resulting emission times τ_e and step heights η are drawn in a spectral map **(bottom)** revealing individual defects as distinct clusters. (taken from [388])

different approach by forcing the device to alternate between two phases. Analogous to MSM techniques, during the stress phase, a high voltage is applied to the MOSFET gate, pushing the device into the strong inversion regime and shifting the defect equilibrium toward the charged state. In the subsequent recovery phase, which occurs in weak inversion, defects discharge with a characteristic emission time τ_e, causing a step η in the threshold voltage shift ΔV_{th}. The step height depends on the defect's position within the oxide as can be easily understood within the charge sheet approximation, see (5.5). It can thus be used to distinguish between different defects.

The experiment is repeated multiple times, and the measured step heights, along with the corresponding emission times, are collected in what is called a spectral map as shown in Fig. 6.3. The capture time τ_c can be determined by varying the stress time t_{str}. When $t_{str} \ll \tau_c$, the defect does not have enough time to charge during the stress phase, so the corresponding cluster in the spectral map only appears for sufficiently long stress times. TDDS allows the characterization of defect dynamics over a much broader time scale than RTN measurements. Additionally, the spectral maps generated can characterize multiple defects simultaneously, whereas analyzing RTN signals becomes increasingly complex when multiple defects are involved.

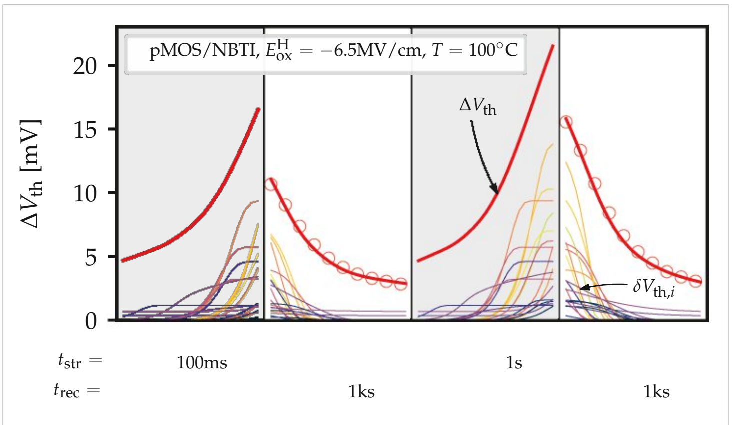


Figure 6.4. Assuming non-interacting defects, the transient ΔV_{th} during an eMSM sequence can be linearly decomposed into small contributions $\delta V_{\mathrm{th,i}}$ from individual defects. The gray and white regions indicate the stress and recovery phases of the eMSM sequence respectively. (taken from [DWJ8])

6.2 Parameter Extraction Methods

This section is based on works previously published in [DWJ8, DWJ10].

One of the main goals in device reliability physics is to accurately predict a device's performance and degradation over its lifetime. Since degradation processes typically occur on the timescale of years under normal operating conditions, accelerated stress conditions are employed to observe degradation within a practical timeframe of hours to days. The main challenge lies in extrapolating these observations back to operating conditions in a physically meaningful manner. In industry, empirical power laws [389] are often used due to their simplicity, however, their predictive capabilities are very limited due to the lack of a physical justification.

In contrast, *Comphy* utilizes a physical description of charge trapping based on the rigorous nonradiative multiphonon (NMP) framework discussed in Ch. 3 and 5. This approach not only provides a more reliable extrapolation scheme but also allows for direct comparison between simulated defect parameters and theoretical DFT studies, aiding in the identification of possible defect candidates responsible for the observed degradation.

6.2.1 Fitting Problem

Due to the amorphous nature of the gate dielectric, the defect parameters are usually distributed, resulting in capture and emission time constants spanning many orders of magnitude. Furthermore, most experiments on ΔV_{th} degradation are carried out on large-area devices, where only the collective response of a whole defect ensemble is observable. Typically, eMSM schemes [361] are used in experiments, resulting in recovery curves as depicted in Fig. 6.4. Using the classical 2-state NMP model, one can deduce from (5.9) or rather (5.12) that these recovery traces have to be composed of multiple exponentially decaying functions. Assuming non-interacting defects, the total ΔV_{th} can be expressed as a linear superposition of individual defect contributions, i.e.

$$\Delta V_{\mathrm{th}}(t, V_{\mathrm{G}}, T) = \int_{\Omega} N(\mathbf{p}) \cdot \delta V_{\mathrm{th}}(t, V_{\mathrm{G}}, T; \mathbf{p}) \mathrm{d}\mathbf{p} \,. \tag{6.1}$$

Here, $\mathbf{p}$ is a set of defect parameters, cf. (5.16), δV_{th} is the contribution of a single defect with parameters $\mathbf{p}$ and $N(\mathbf{p})$ is the distribution function in the parameter space Ω.

Hence, extracting defect parameters from experiments is equivalent to finding a suitable distribution function $N(\mathbf{p})$ in order to fulfill (6.1). However, such an *inverse problem* [390] is mathematically ill-posed and requires to impose further restrictions on $N(\mathbf{p})$ to obtain a physical solution. Another relevant example of such a problem would be the *multiexponential analysis* required in deep level transient spectroscopy (DLTS) [391, 392]. In the following, two different approaches implemented in *Comphy* are discussed to address this parameter extraction problem.

6.2.2 Gaussian Defect Bands

The main challenge in solving (6.1) is that the unknown $N(\mathbf{p})$ is a scalar field and hence has far too many degrees of freedom. A straightforward approach to obtain a solution is to enforce a certain distribution shape. The most common choice found in literature is a normal distribution for the energetic parameters E_{T} and E_{R} resulting in a *Gaussian defect band* [26]. The spatial variable x_{T}, on the other hand, is usually assumed to be uniform, leading to a constant defect density in a certain region of the oxide. Such a Gaussian defect band is therefore characterized by the parameters

$$\mathcal{P} = \left(\langle E_{\mathrm{T}} \rangle, \sigma_{E_{\mathrm{T}}}, \langle E_{\mathrm{R}} \rangle, \sigma_{E_{\mathrm{R}}}, R, x_{\mathrm{min}}, x_{\mathrm{max}}, N_{\mathrm{T}}\right) \tag{6.2}$$

where N_{T} represents the trap concentration in the region from x_{min} to x_{max} within the oxide. Besides this simple model, an exponentially decaying spatial distribution is often chosen to reflect the increasing defect density towards the interface [393].

By fixing the mathematical form of the distribution, only its defining parameters, like the mean and standard deviation in the case of Gaussian defect bands, have to be optimized to fit the experimental ΔV_{th}. This can be achieved by using the simplex-method or other local optimizers.

The concept of Gaussian defect bands is intuitive and has been widely used to reproduce experimental data across different devices. However, artificially constraining the shape of the parameter distribution can introduce artifacts, complicating physical interpretations and making it difficult to link the obtained distributions to theoretical DFT studies. An example of such a misrepresentation is shown in Fig. 6.5, illustrating the case of electron traps in $\mathrm{SiO_2}$, as published in [26]. While the extracted defect band suggests a mean relaxation energy of 5 eV, closer examination reveals that, even under severe stress conditions, most defects in this band remain electrically inactive due to their extremely large time constants. When these defects are filtered out, only the tails of the

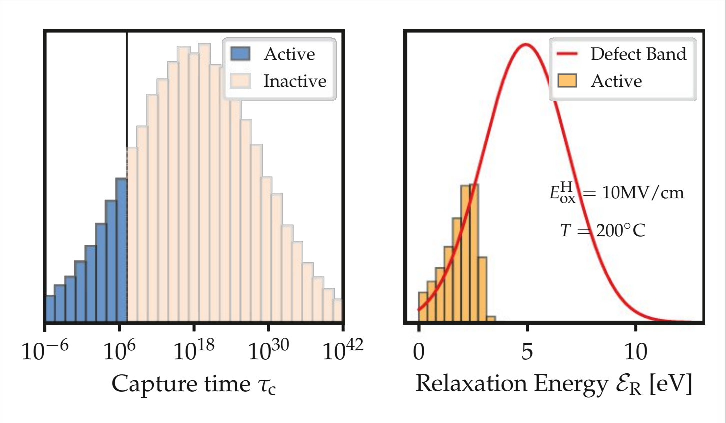


Figure 6.5. Extracted τ_c and E_R distributions for electron traps in SiO_2 using a Gaussian defect band as published in [26]. Even under severe conditions, only a small fraction of the distribution is active (τ_c $<1 \times 10^7$ s), leading to significant deviations from the assumed Gaussian distribution when only considering the active defect set. (taken from [DWJ8])

Gaussian band contribute to the observed charge trapping, resulting in a significantly altered parameter distribution. Similar cases of likely misinterpretations, resulting in unphysically large relaxation energies, can be found in the literature [328, 394, 395].

Another drawback of Gaussian bands is that the optimization procedure often needs manual intervention and a good initial guess. Especially in SiC based devices, where multiple different defect bands are involved, as a result of the increased stoichiometric complexity near the interface [334], the optimization becomes increasingly difficult and time consuming.

6.2.3 Effective Single Defect Decomposition (ESiD)

To overcome the limitations of predefined distribution functions, a novel extraction method called Effective Single Defect Decomposition (ESiD) has been developed [DWJ8]. This method is capable of inferring an unrestricted physical defect distribution from experimental eMSM sequences in a semi-automatic manner.

Algorithm

Returning to the original problem (6.1), the algorithm directly exploits the fact that the macroscopic threshold voltage shift, ΔV_{th}, is a linear superposition of independent defect responses, denoted by δV_{th}. Although this assumption is strictly valid only in the limit of low defect concentrations, it remains a reasonable approximation for typical defect concentrations in the range of $10^{18} - 10^{19}\,\mathrm{cm}^{-3}$. Regardless, the validity of this approximation can always be verified retrospectively by using the resulting defect distributions in a self-consistent Poisson (SCP) calculation and comparing it to the non-self-consistent results initially used to obtain these distributions. Such a comparison is shown in Fig. 6.6 for the extracted hole trap band in SiON devices [DWJ8] with a density of $8.0 \times 10^{18}\,\mathrm{cm}^{-3}$. As can be seen, the non-self-consistent treatment only introduces a minor error.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

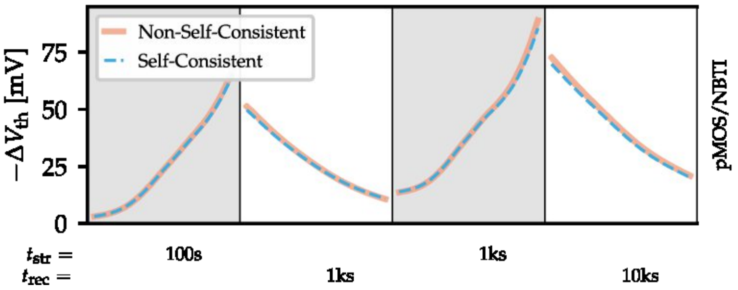


Figure 6.6. Comparison between the non-self-consistent and self-consistent defect response of the extracted hole trap band in SiON devices [DWJ8] with a density of $8.0 \times 10^{18}\,\mathrm{cm}^{-3}$. The assumption of linearly independent defects introduces only a minor error.

Using this assumption, finding an optimal distribution function $N(\mathbf{p})$ can be recast into the non-negative linear least square (NNLS) problem

$$N(\mathbf{p}) = \arg\min_{\hat{N}\geq 0} \left| \Delta V_{\mathrm{th}} - \int_{\Omega} \hat{N}(\mathbf{p}) \delta V_{\mathrm{th}} \mathrm{d}\mathbf{p} \right|^2 . \tag{6.3}$$

It is crucial to include the non-negativity constraint here, as negative values for the defect density would lead to an unphysical solution. However, note that (6.3) is still essentially a reformulation of the original problem and as such remains ill-posed.

Naively trying to solve the optimization problem as stated above would lead to unstable solutions which are highly sensitive to noise in the input data. In order to obtain a physically meaningful solution, the problem has to be regularized. The implementation in *Comphy* uses the *Tikhonov* scheme [396], in which a regularization term depending on the total density

$$\hat{N}_{\mathrm{tot}} = \int_{\Omega} \hat{N}(\mathbf{p}) \mathrm{d}\mathbf{p} \tag{6.4}$$

is introduced. In this scheme, the term to be minimized is given by

$$\left| \Delta V_{\mathrm{th}} - \int_{\Omega} \hat{N}(\mathbf{p}) \delta V_{\mathrm{th}} \mathrm{d}\mathbf{p} \right|^2 + \gamma^2 N_{\mathrm{tot}}^2 \tag{6.5}$$

with γ being an adjustable regularization parameter. The effects of γ on the optimization are twofold. It favors solutions which require only a low defect density to match the experimental data and, although not obvious, also implicitly forces the solutions to be smooth. This can be shown by performing a singular-value decomposition (SVD) on the linear optimization problem and realizing that γ sets a lower bound for the smallest possible singular values, which are responsible for discontinuous solutions. In practice, γ has to be adjusted to the particular problem at hand. However, as illustrated in Fig. 6.7 a reasonable value can be estimated by plotting the total defect density versus the error norm to the experimental data for different values of γ, resulting in a L-shaped curve. As can be seen, for too small values, the approximation error is very low at the expense of a very high defect concentration. Quite to the contrary, if γ is too large, the problem becomes overregularized, meaning that there is a steep increase in approximation error

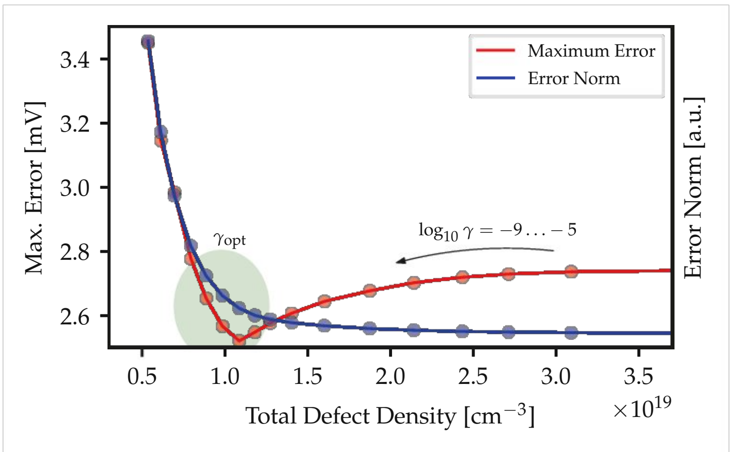


Figure 6.7. Impact of the regularization parameter γ. For small γ, the solution requires an unphysically large defect density, whereas too large values lead to quickly growing approximation errors. The optimal value for γ lies in the green region between these two detrimental regimes. (taken from [DWJ8])

when further increasing γ. According to the *L-criterion* [397], the optimal value for γ lies at the "corner" of this L-shaped curve, providing a good compromise between accurate representation of the experimental data with a reasonably small defect concentration.

Implementation

In order to solve (6.5) efficiently, the regularized estimator can be transformed to the standard NNLS form by introducing [DWJ8]

$$\mathbf{N} = \underset{\hat{\mathbf{N}} \geq \mathbf{0}}{\arg\min} \left\| \mathbf{A} \cdot \hat{\mathbf{N}} - \mathbf{b} \right\|_2^2 \quad \text{with} \tag{6.6}$$

$$\mathbf{A} = \begin{bmatrix} \delta\mathbf{V} \\ \gamma\mathbf{I} \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} \boldsymbol{\Delta}\mathbf{V} \\ \mathbf{0} \end{bmatrix} \tag{6.7}$$

with $\mathbf{I}$ being a unit matrix of proper dimensions. In the *Comphy* implementation, the estimator given by (6.6) is solved iteratively by increasing γ and invoking a fast NNLS algorithm introduced in [398] at each step until the "corner" is detected.

For an efficient implementation in *Comphy*, ESiD is only employed for the energetic parameters E_{T} and E_{R} since they are most relevant for identifying possible defect candidates. As mentioned earlier in Sec. 5.4.4, $R = 1$ is assumed in order to eliminate cross-correlations between the model parameters. The remaining distributions along x_{T} and ΔQ are assumed to be uniform, meaning that the responses along these dimensions can be added up before optimization with ESiD. The NNLS optimization problem can then be easily discretized by defining a search region $[E_{\mathrm{T,min}}, E_{\mathrm{T,max}}] \times [E_{\mathrm{R,min}}, E_{\mathrm{R,max}}]$ on a uniform grid. By doing so, the parameter extraction reduces to determining a 2-dimensional $(E_{\mathrm{T}}, E_{\mathrm{R}})$ distribution function. An example for such a $(E_{\mathrm{T}}, E_{\mathrm{R}})$ map is given in Fig. 6.12.

6.3 Applications

In the following, the modeling tools developed within the scope of this work are applied in combination with insights from first-principles calculations to identify defects causing detrimental effects in different device technologies from experimental data.

6.3.1 Trap Extractions with ESiD

This section is based on works previously published in [DWJ8].

In this first example, hole and electron trap distributions are extracted from large-area MOSFET devices ($W \times L = 10 \times 10\,\mu\text{m}^2$) fabricated using a commercially available 180 nm SiON technology. While hole traps in similar technologies have previously been investigated using TDDS on small-scale devices [25], it will be demonstrated here that the same trap parameters can also be obtained from eMSM measurements in combination with ESiD. Since large-area devices sample many defects simultaneously, the time required to determine the defect distribution is significantly reduced compared to single-defect studies.

Experimental Details

For the extraction of hole and electron trap distributions, both NBTI measurements on pMOS and PBTI measurements on nMOS devices were performed. The temporal recovery of device degradation after stress has been monitored with a custom-built low-noise measurement tool [399] using an eMSM scheme with stress times in the range of $t_\text{str} =$ 100 µs to 10 ks and recovery times up to $t_\text{rec} = 10\,\text{ks}$. The ΔV_th readout has been performed at $I_\text{D} = 100\,\text{nA}{\cdot}W/L$ with a minimal delay after stress of $t_\text{d} = 100\,\mu\text{s}$. In order to accurately account for the bias and temperature dependence of BTI, the experiments were performed at three stress fields $\left|E_\text{ox}^\text{H}\right| = 5$, 6.5 and $8\,\text{MV}\,\text{cm}^{-1}$, and at three temperatures $T = 50$, 100 and 150 °C. The ΔV_th response has been extracted by converting ΔI_D according to the initial $I_\text{D}(V_\text{G})$ characteristics of the unstressed device [361]. For each device, low-frequency CV curves at $f = 100\,\text{kHz}$ were recorded and used to calibrate the electrostatics in our device models. As an example, the calibrated CV curves at 100 °C for the pMOS are given in Fig. 5.1.

Defect Extraction

The ESiD algorithm is employed for extracting the defect parameter distributions for hole and electron traps in SiON from negative BTI (NBTI) measurements on pMOS devices and positive BTI (PBTI) on nMOS devices. In order to account for the chemically different environments in the oxide and IL, two disjunct and independent search regions for the ESiD extraction were defined. The used parameters for the search regions as well as the extracted total defect densities are summarized in Tab. 6.1. Based on experimental [400] and theoretical [401] studies on the Si/SiO_2 interface, the IL thickness was assumed to be 0.6 nm. Note that the precise value of the assumed IL thickness is only of minor importance

	E_T [eV]	E_R [eV]	x_T [nm]	N_T [cm^{-3}]
Hole traps (IL)	-2.0–0.0	0.1–5.0	0.0–0.6	8.0×10^{18}
Hole traps (oxide)	-2.5–0.0	0.1–5.0	0.6–2.5	2.8×10^{18}
Electron traps (IL)	0.0–2.0	0.1–5.0	0.0–0.6	1.4×10^{18}
Electron traps (oxide)	0.0–2.5	0.1–5.0	0.6–2.5	7.0×10^{16}
Grid spacing	0.05	0.125	0.1	–

Table 6.1. Parameters used to define the different search regions in the oxide and the interlayer (IL). The concentration N_T is the resulting total defect concentration in each defect band as extracted from ESiD. (taken from [DWJ8])

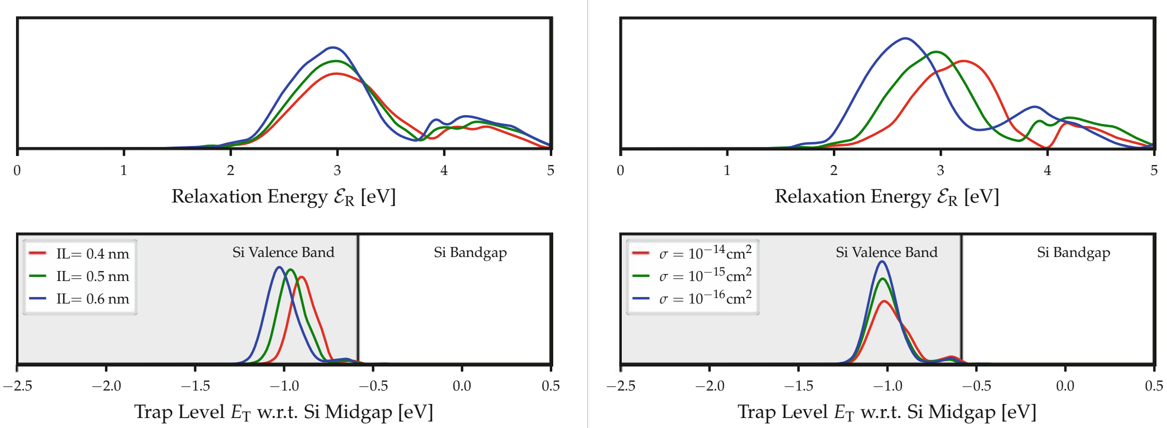


Figure 6.8. Resulting hole trap distribution extracted with ESiD for varying interlayer thicknesses and capture cross sections. A change in the assumed IL thickness only causes a small shift in the trap level distribution. Similarly, varying the capture cross section over 2 orders of magnitude only introduces a small shift in the relaxation energies, while modifying the required total defect density. However, in both cases, the defect distributions qualitatively do not significantly change upon varying those parameters.

for the resulting defect distribution as can be seen from Fig. 6.8 (left). A change in the assumed IL thickness only causes a small shift in the trap level distribution. Similarly, the capture cross section σ enters the transition rates (5.12) only linearly, whereas all other parameters have an exponential impact. As shown in Fig. 6.8 (right), varying the capture cross section over 2 orders of magnitude only introduces a small shift in the relaxation energies, with the distributions staying qualitatively the same.

In accordance with previous findings [402], the ESiD algorithm shows that most of the active defects in both n- and pMOS are located in the IL, as depicted in Fig. 6.9. However, the total active defect density in nMOS is lower compared to pMOS devices by a factor of seven in good agreement with earlier investigations [351, 403]. Although these findings have been well established, note that they naturally emerge from simple eMSM recovery traces using this novel algorithm. Compared to earlier reports utilizing Gaussian parameter distributions [26, 328, 394], the ESiD extraction shows significantly lower defect densities, ranging from 5 to 15 times less depending on the device type. As already explained in

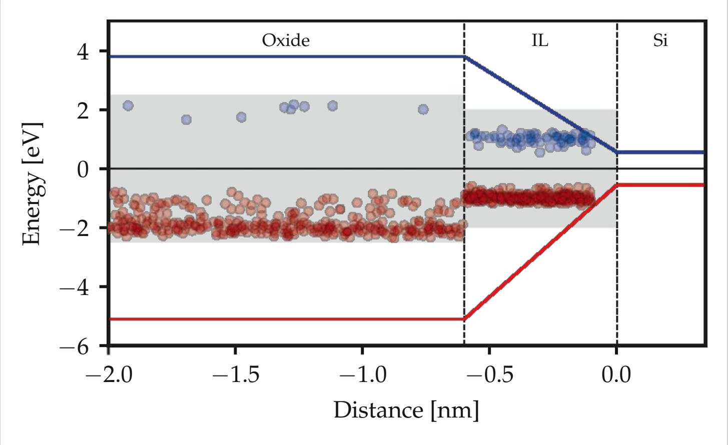


Figure 6.9. Samples drawn from the extracted defect distributions within the band diagram of the device. The shaded areas indicate the search regions used in ESiD. The highest concentration of both electron (blue) and hole (red) traps is located within the interfacial layer (IL). (taken from [DWJ8])

Fig. 6.5, this is an artifact of choosing a specific parameter distribution function, where the observed degradation only stems from the distribution tail, leading to an apparent higher defect density. One advantage of ESiD is that inactive defects within the provided measurement window are automatically discarded since they cannot contribute to the superposition given in (6.3). However, as shown in Fig. 6.10, the extracted parameter set describes all recorded MSM traces very well, despite the much lower defect density. As can be seen, both the dependence on the gate bias as well as the temperature activation of BTI degradation are well captured by this new model.

Most notably, the temperature activation in the PBTI/nMOS case shows a peculiar feature due to fast electron traps. As shown in Fig. 6.10 (right), the measured ΔV_{th} degradation actually decreases with higher temperatures for longer stress times. As pointed out already in [404], simpler BTI characterization schemes like the use of single effective activation energies would predict an unphysical negative activation energy for the charge transfer process in this scenario. The degradation during stress is indeed larger for higher temperatures, however, some traps become too fast and already recover during the delay time between stress and recovery phase. These traps therefore fall out of the measurement window and cause this apparent decrease in degradation.

Defect Distributions

Since we have fixed the parameter $R = 1$ and averaged over x_{T} within a defect band, the defining characteristic of a defect band is its $(E_{\mathrm{T}}, E_{\mathrm{R}})$ distribution. These distributions are presented as stacked histograms in Fig. 6.11 for all four defect bands used during the extraction. As can be seen, most electron and hole traps are located within the IL layer approximately 1.0 eV above and below the Si midgap, respectively. While hole traps are still present in the deeper oxide layer, electron traps are almost entirely absent in this region. This finding explains the more pronounced quasi-permanent BTI component (P) in pMOS devices compared to their nMOS counterparts. Note, however, that P cannot be

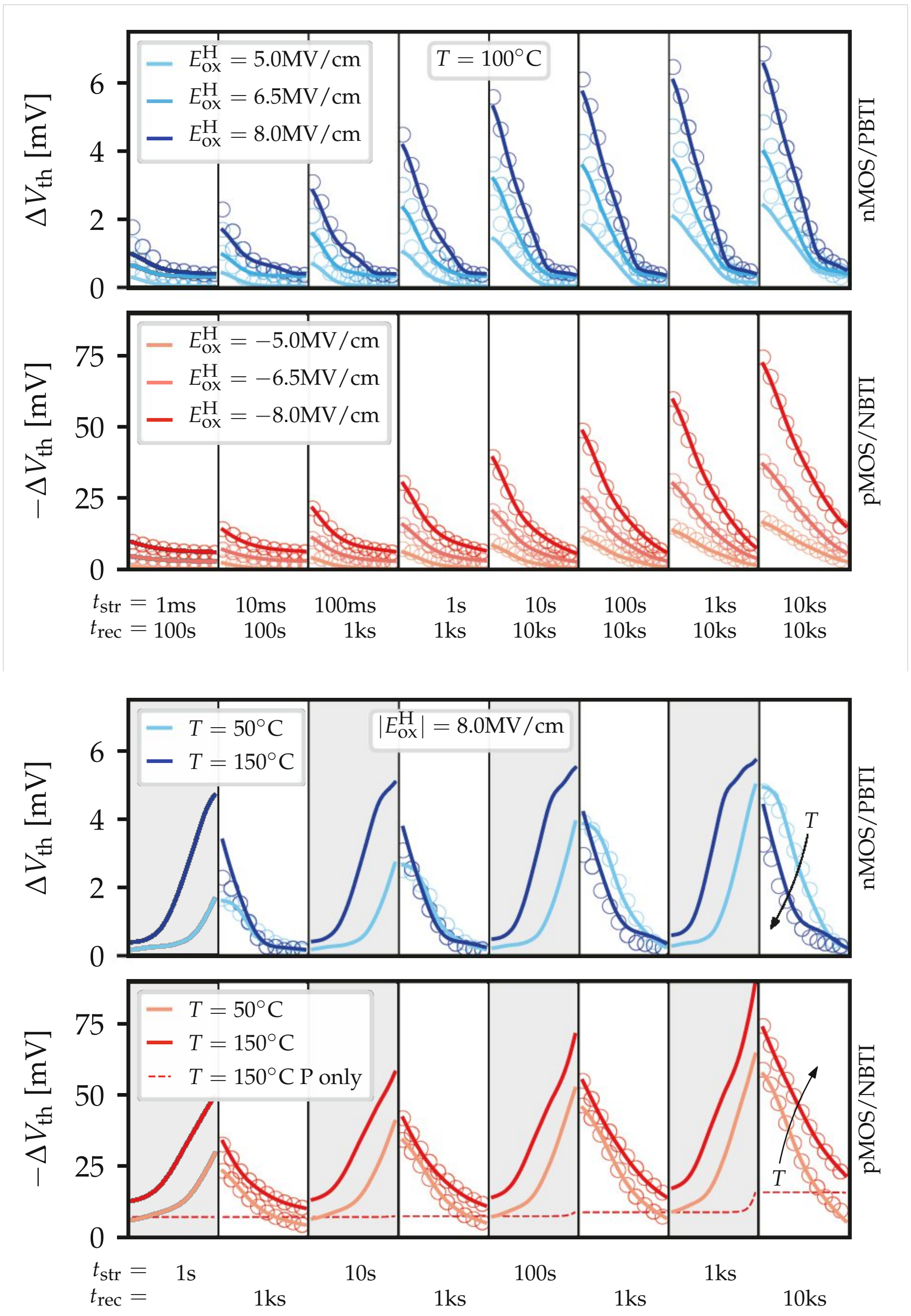


Figure 6.10. Experimental and simulated eMSM sequences derived from the extracted defect bands for varying gate bias (left) and temperature (right). The first data point is measured 100 µs after stress. The extracted (quasi)-permanent component (P) is separately shown as a dashed line for the pMOS in the right plot. As can be seen, there is a very good agreement with the measured ΔV_{th} degradation over the entire measured temperature and bias range. Note the apparent negative temperature activation for PBTI in the lower figure due to fast electron traps. (taken from [DWJ8])

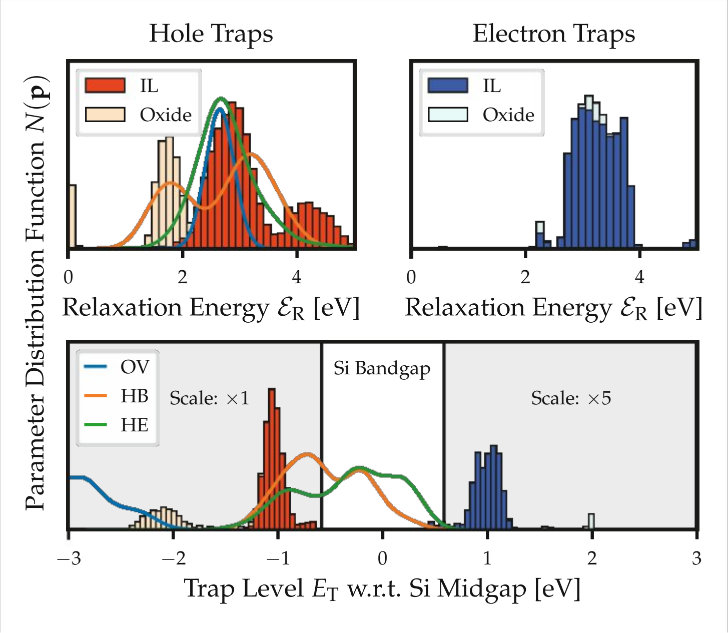


Figure 6.11. Extracted defect bands and a comparison to theoretical defect candidates[23] for hole traps in SiO_2. As can be seen, the hydrogen-related defects are in good agreement with our obtained IL hole traps, whereas the OV defect could act as a deeper trap in the oxide. (taken from [DWJ8])

fully described by charge trapping in pre-existing oxide defects alone since other effects like defect creation and hydrogen kinetics also play an important role [15, 405]. The rigorous treatment of these mechanisms is beyond the scope of this work, instead P is represented by artificial defects within the used NMP framework. In our simulation we consider defects, which can only trap charges but not emit them again within the experimental window ($\tau_e > 5 \times t_{rec}^{max}$) even at the highest temperature, to be responsible for P. The collective ΔV_{th} response of these defects is shown as a dashed line in Fig. 6.10 (right). However, the corresponding defect parameters are unphysical and therefore not shown in Fig. 6.11.

The recoverable BTI component, however, is well described by the NMP model and its extracted defect parameters allow the identification of likely responsible defect types by comparison to DFT predictions. As shown in Fig. 6.11, the (E_T, E_R) distribution of the extracted IL hole traps is in good agreement with theoretical results [23] for the well-known hydrogen bridge (HB) defect and the recently popularized hydroxyl-E′ center (HE) [246], cf. Sec. 4.3.2. Regarding the energetic location of the defect bands, it has to be kept in mind that E_T predictions commonly show errors of up to $\pm 0.5\,\text{eV}$ within standard DFT. Hence, the shown theoretical distributions should be considered as physically reasonable ranges rather than precise distributions of defects. However, the oxygen vacancy (OV), which has been previously suspected [406] to be responsible for BTI, shows a far too deep trap level to account for most of the observed degradation. Since the OV still has been identified by ESR measurements during NBTI stress [407], the author speculates that the small number of deeper oxide defects extracted from the present data could be attributed to OVs.

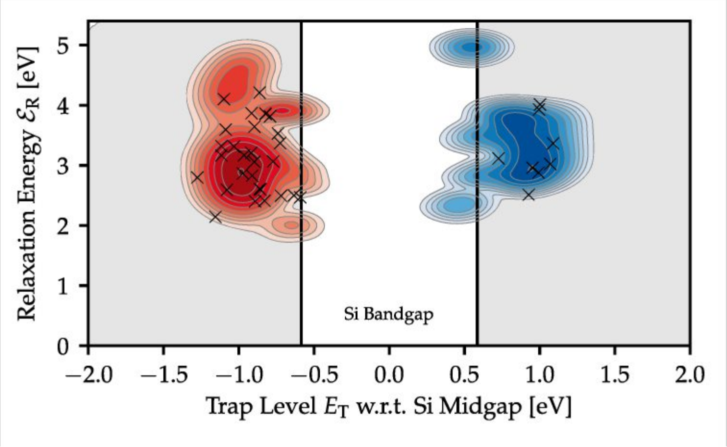


Figure 6.12. Heat map of the $(E_\mathrm{T}, E_\mathrm{R})$ distribution for the extracted IL trap bands using the ESiD method together with parameters from single defects (crosses) obtained with TDDS [16, 408]. As can be seen, both distributions overlap nicely, confirming the physicality of the extracted defect bands. (taken from [DWJ8])

Note that in this study, the identification of electron traps is not attempted due to the presence of a recently discovered very deep negative charge state [DWJ6] (state 3 in Fig. 5.2 (top)). The impact of this state on the trapping characteristics, as well as the feasibility of incorporating its effects into an effective 2-state model with meaningful physical parameters, remains the subject of future investigations.

Comparison to Single Defect Characterizations

In order to validate the defect parameters extracted from eMSM measurements using the ESiD method, they were compared against extracted single defect parameters from earlier data sets [16, 408] on ultra scaled SiON devices otherwise similar to the large-area devices used in this work. Fig. 6.12 shows that the resulting single-defect parameters agree excellently with the extracted parameters for the IL layer from the performed eMSM measurements. Note that TDDS typically requires a smaller experimental window ($\tau_\mathrm{c}, \tau_\mathrm{e} \leq 100\,\mathrm{s}$) compared to eMSM since many repeated recovery traces have to be analyzed in order to obtain sensible mean capture/emission times. Therefore it is expected that TDDS primarily reflects the properties of the faster IL layer instead of the deeper oxide. These findings show that the ESiD algorithm indeed provides physical defect bands which can be used for identification of defects and can act to some degree as a suitable substitute for the much more involved and time-consuming single-defect characterization.

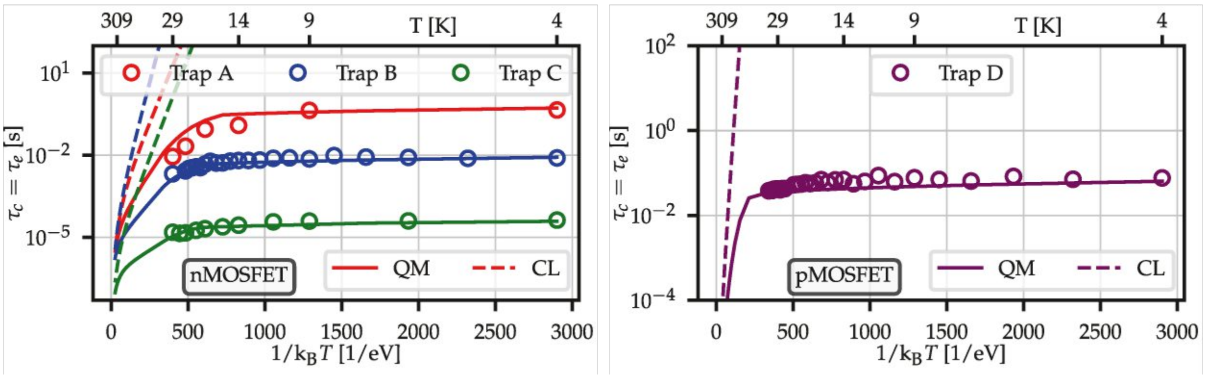


Figure 6.13. Experimental time constants (evaluated at $\tau_\mathrm{c} = \tau_\mathrm{e}$) obtained from RTN signals in nMOS (left) and pMOS (right) devices at cryogenic temperatures together with fits of the classical (CL) and full quantum mechanical (QM) model. As can be seen, both electron and hole traps become temperature independent. This can only be explained by nuclear tunneling within the quantum mechanical charge transition model. (taken from [DWC2])

6.3.2 RTN in Cryo-CMOS

This section is based on works previously published in [DWC2].

The previous example demonstrated the usefulness of ESiD to extract physically meaningful defect parameter from measurements on large-area devices. Here, scaled SiON devices with dimensions ($W \times L = 0.22 \times 0.18\,\mu\mathrm{m}^2$) are studied using RTN analysis at cryogenic temperatures. These devices are based on the same process as the large-area devices studied in Sec. 6.3.1.

In this study, three defects were identified in nMOS devices, compared to a single hole trap in a pMOS device. The corresponding RTN signals for trap A are shown in Fig. 6.2. As can be seen from the Arrhenius plots presented in Fig. 6.13, the time constants become temperature independent towards lower temperatures, indicating that nuclear tunneling is becomes the dominant transition mechanism.

The classical and quantum mechanical models for single defects in *Comphy* were subsequently fitted to this experimental data to extract the corresponding defect parameters. As expected, only the quantum mechanical model correctly describes this temperature independence, while the classical model predicts a complete freeze-out. The defect parameters obtained from the quantum mechanical model are listed in Tab. 6.2 and additionally visualized on an $(E_\mathrm{T}, E_\mathrm{R})$ map in Fig. 6.14. It turns out that all identified defects are located near the interface and exhibit particularly small relaxation energies, inconsistent with the oxide defects extracted in Fig. 6.12. Furthermore, the trap levels for both hole and electron traps are close to the silicon valence and conduction bands, respectively. This indicates that in these devices the RTN at cryogenic temperatures stems entirely from interface defects like the P_b center, which is predicted to have similar charge trapping parameters [20] as the defects extracted in this work.

	Device	E_T [eV]	E_R [meV]	x_T [nm]	ΔQ [$\sqrt{\mathrm{u}}$Å]
Trap A	nMOS	0.44	110	0.2	3.5
Trap B	nMOS	0.48	95	0.3	3.0
Trap C	nMOS	0.46	60	0.05	2.6
Trap D	pMOS	-0.44	100	0.25	3.0

Table 6.2. Trap parameters extracted from RTN signals at cryogenic temperatures. The trap level E_T is given w.r.t. the Si midgap. All defects are close to the interface and have small relaxation energies. (taken from [DWC2])

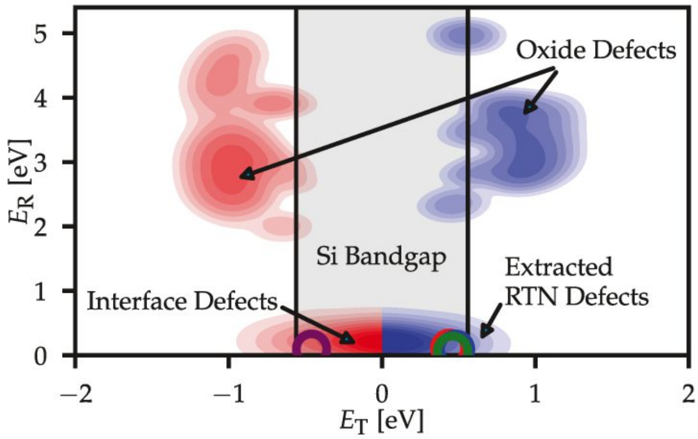


Figure 6.14. Defect parameters extracted from the RTN signals plottet on an $(E_\mathrm{T}, E_\mathrm{R})$ map. The observed electron and hole traps are close to the respective band edges of Si with an exceptionally low relaxation energy compared to the oxide defects depicted in Fig. 6.12. However, DFT predictions for interface defects like the P_b center [20] (elliptic distribution within the Si bandgap) are in good agreement with the extracted RTN defects. (taken from [DWC2])

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

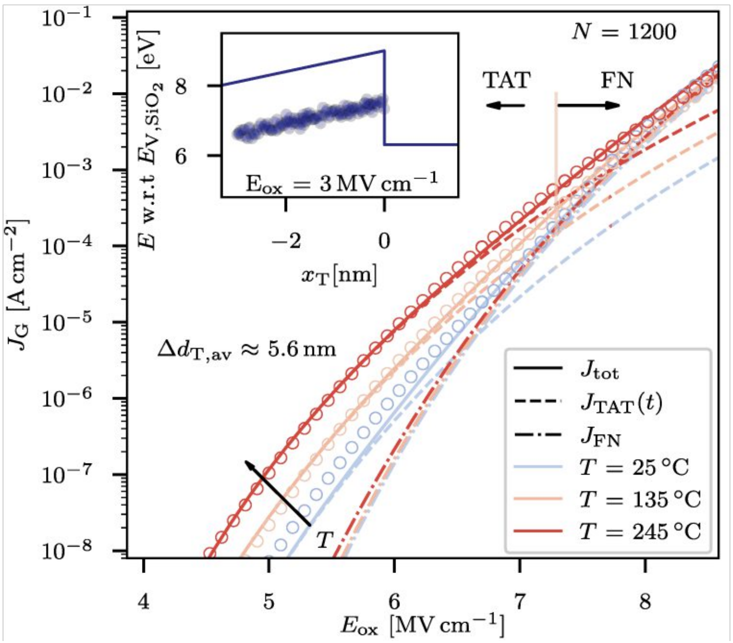


Figure 6.15. TAT simulations (lines) compared to experimental leakage currents [411] (circles) of a SiC MOSCAP. The simulations can accurately reproduce the measurement characteristics both in the TAT and the FN regime. The calculation shows that the thermally activated leakage current can be explained by carriers tunneling from the substrate channel to traps located within 3 nm of the interface and then further to the conduction band of the insulator. (taken from [DWJ15])

6.3.3 Gate-Leakage Currents through Polarons

This section is based on works previously published in [DWJ15].

The previous two examples focused on charge trapping due to oxide or interface defects. However, defects can also facilitate gate-leakage currents across the oxide via trap-assisted tunneling (TAT), as discussed in Sec. 5.5. This phenomenon is particularly relevant for SiC power devices, where comparatively large leakage currents are often reported [409], even for thick gate oxides [410]. Additionally, these leakage currents exhibit strong temperature dependence, a characteristic feature of TAT [18], unlike direct (DT) or Fowler-Nordheim (FN) tunneling.

To better understand the microscopic nature of TAT, this work analyzes leakage currents measured on poly-Si/SiO$_2$/SiC MOSCAPs, as reported in [411], using the TAT model implemented in *Comphy* [DWJ18]. The parameters of a Gaussian electron trap band in SiO$_2$ were optimized to fit the available experimental data. As shown in Fig. 6.15, the experimental gate leakage currents exhibit a pronounced temperature dependence for oxide fields below 8 MV cm^{-1}, indicating a TAT current. For higher fields, Fowler-Nordheim (FN) tunneling, which is largely temperature-independent, begins to dominate. Both tunneling regimes are well captured by the TAT model.

	$\langle E_\mathrm{T}\rangle$ [eV]	σ_{E_T} [eV]	$\langle E_\mathrm{R}\rangle$ [eV]	σ_{E_R} [eV]	N_T [cm^{-3}]
TAT band	2.85	0.1	0.89	0.11	7.6×10^{18}
Polarons (DFT)*	2.53	0.23	1.06	0.23	-

Table 6.3. Trap parameters extracted from the TAT model compared to values for polarons (cf. Sec. 4.3.3) predicted by DFT. *see comment in Fig. 6.16. (taken from [DWJ15])

The required defect parameters are listed in Tab. 6.3 and plotted in Fig. 6.16. The relaxation energy is greater than that of interface defects but still significantly smaller than that of typical oxide defects (cf. Fig. 6.14). However, the DFT calculations conducted for this work [DWJ15] show that polarons or self-trapped electrons (STEs), as discussed in Sec. 4.3.3, occupy this intermediate energy range. It is worth noting that STEs have a curvature ratio of $R = 1.35$ in DFT, resulting in a higher relaxation energy, whereas the TAT model implemented in *Comphy* requires $R = 1$. Nevertheless, there is good agreement between the *Comphy* parameters and DFT calculations when the DFT relaxation energies are rescaled to $R = 1$, using the parameter cross-correlation discussed in Sec. 5.4.4.

This study demonstrates that STEs are a strong candidate to explain the increased gate-leakage currents observed in SiC devices. This finding is consistent with the higher conduction band edge of SiC being closer to the polaron band compared to that of Si (see Fig. 4.17). Additionally, it has been shown that STEs could also account for leakage currents in other gate oxides, such as ZrO_2 [DWJ15]. Based on these simulations, it appears that STEs are generally a prime candidate for enabling charge hopping, due to their relatively low relaxation energy combined with their position in the upper half of the oxide bandgap [DWJ15].

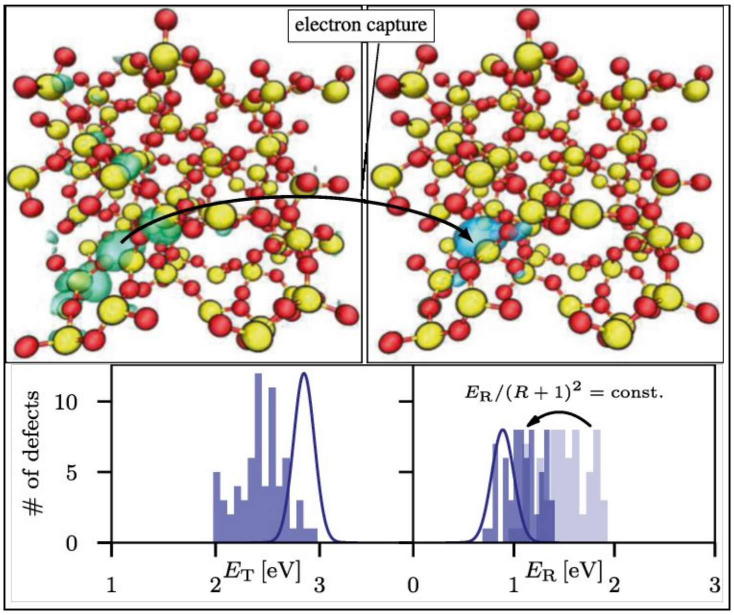


Figure 6.16. Atomistic a-SiO_2 model showing a semilocalized conduction band state (top left). Upon electron injection, the wavefunction localizes, forming a polaron (top right). The bottom panel shows a comparison of the associated charge transition parameters calculated with DFT (histogram) and the model value distributions required to explain the experimental data listed in Tab. 6.3 (lines). *Note that the DFT results initially showed a curvature ratio of $R = 1.35$, whereas the TAT model requires $R = 1$. Hence, the relaxation energy from DFT was scaled according to the (R, E_R) cross-correlation discussed in Sec. 5.4.4. (taken from [DWJ15])

7

Conclusions & Outlook

This thesis has investigated various reliability phenomena linked to charge trapping in oxide and interface defects, such as bias temperature instability (BTI), random telegraph noise (RTN), and gate-leakage currents through trap-assisted tunneling (TAT). A multiscale modeling approach, combining first-principles and device-level simulations, was used to identify potential defects responsible for observable device degradation. The work presented here has led to several methodological advancements and insights into the physics of device reliability, which are summarized as follows:

- Charge trapping is governed by the theory of nonradiative multiphonon (NMP) transitions. Significant progress in this field has been made in recent years, largely due to the development of ab-initio methods that enable the prediction of transition rates. Notably, the works by Alkauskas *et al.* [174] and Goes *et al.* [23] are based on different lineshape functions, the Herzberg-Teller and the Franck-Condon lineshape functions, respectively. However, as demonstrated in Sec. 3.2, both lineshape functions can contribute in comparable amounts. The developed theory further incorporates the effect of coupling to environmental phonon modes within the typically used 1-dimensional effective mode approximation. It was shown that these bath modes can significantly impact the overall transition rate.
- Experimentally, it is well known that hole injection into SiO_2, such as through irradiation, results in the dissociation of $Si-H$ bonds and the formation of electrically active silicon dangling bonds [185]. While this process has been studied theoretically in prior work [238, 240], the study presented in Sec. 4.1 provides a detailed analysis of the hole capture process and predicts the capture cross section in good agreement with experimental data. This not only corroborates the hypothesis of hole-induced $Si-H$ bond dissociation but also serves as a benchmark for the theoretical methodologies employed in this work.
- Inferring crucial defect parameters, such as charge transition levels or relaxation energies, from BTI experiments on large-area MOSFETs requires a well-calibrated device model and an appropriate parameter extraction method. In Sec. 6.2.3, a novel extraction algorithm, the Effective Single Defect Decomposition (ESiD), was

developed and implemented in the *Comphy* device simulator. Due to its capability to automatically extract defect parameters from experimental BTI data, ESiD has already been widely used across various device technologies, including SiC power devices [DWJ17], high-κ CMOS for cryogenic applications [DWJ16], and ferroelectric FETs [412]. In Sec. 6.3.1, ESiD was employed to obtain physical defect parameters for hole trapping in SiON devices, which showed good agreement with theoretical predictions for hydrogen-related defects, in contrast to previously reported defect bands [26].

- As demonstrated in Sec. 6.3.2, the analysis of RTN signals at cryogenic temperatures using a fully quantum mechanical charge trapping model reveals that noise at these temperatures is predominantly caused by interface traps, such as the P_b center, rather than typical oxide defects [DWC2]. This insight could be instrumental in improving the coherence time of solid-state qubits, which is currently compromised by noise.

- A novel TAT model has been developed in which charge transfer is described using the NMP framework, enabling a unified treatment of BTI and TAT within the *Comphy* device simulator [DWJ18]. Applying this framework to TAT in SiC MOSCAPs, in conjunction with DFT calculations as discussed in Sec. 6.3.3, suggests that polaron hopping is the mechanism responsible for TAT in these devices [DWJ15]. Since polarons also occur in other amorphous gate dielectrics, this mechanism could represent a more universal phenomenon [DWJ15].

Outlook

Understanding the microscopic physics of device degradation and developing device models that accurately capture the essential physics is a rich field of research, one that will continue to gain importance as device dimensions approach the atomic scale. Below, several ideas for future research are presented, with the hope of inspiring interest among researchers and fellow students entering this field.

- From a theoretical perspective, charge trapping at oxide defects is typically studied in bulk oxide structures using DFT, while the effects of the semiconductor/insulator interface are generally addressed through device models. However, directly modeling the charge transfer across the interface with DFT would be advantageous for gaining deeper insights into the behavior of defects near the interface, especially under the influence of strain or a non-stoichiometric environment. This approach currently faces challenges, such as the difficulty in controlling the charge state of defects. Although some pioneering steps have been made in this direction [DWJ2, 232, 369, 370], a generally applicable methodology for tackling this problem has yet to be developed.

- Currently, most DFT studies investigating NMP rates for charge capture or emission at a defect employ a simplified 1-dimensional effective mode model. As discussed in this thesis, the multidimensional nature of the physical problem influences the system in two key ways. First, environmental phonon modes can introduce additional electron-phonon coupling. Second, the classical minimum energy crossing point generally does not lie along the effective mode. While the first effect can be treated quantum mechanically, as described in Sec. 3.2, the second is currently addressed only in a classical manner. Developing a unified framework to simultaneously incorporate both effects would be highly desirable.

- This thesis primarily focused on the development of suitable modeling methodologies, which were largely applied to established material systems like Si/SiO$_2$ as a proof of concept. However, the developed methods can also be applied to emerging technologies, where reliability is less well understood. For instance, current transistors based on 2D materials, despite their tremendous theoretical potential, face significant reliability challenges [9]. Applying the multiscale modeling approach outlined in this thesis could help improve device stability, potentially enabling the transition from experimental devices to industrial manufacturing.
- While knowledge of defect physics and charge trapping has already become integral to the electronic device community, driven by the necessity to enable further optimization and downscaling, other emerging fields could also benefit from this understanding. For example, building solid-state qubits currently faces challenges due to noise-limited coherence times [320]. Given the similarities in underlying physical mechanisms, a multiscale modeling approach could be applied to understand, predict, and enhance qubit lifetimes.

References

[1] M. Attila and D. Kahng, "Silicon-silicon dioxide field induced surface devices", in IRE Solid State Device Res. Conf. (1960).

[2] B. Deal, "Standardized terminology for oxide charges associated with thermally oxidized silicon", IEEE Transactions on Electron Devices **27**, 606–608 (1980).

[3] X. Guo and T. Ma, "Tunneling leakage current in oxynitride: Dependence on oxygen/nitrogen content", IEEE Electron Device Letters **19**, 207–209 (1998).

[4] M. M. Frank, "High-k/metal gate innovations enabling continued CMOS scaling", in 2011 Proceedings of the European Solid-State Device Research Conference (ESSDERC) (IEEE, 2011), pp. 25–33.

[5] C. Wu, D. Lin, A. Keshavarzi, C. Huang, C. Chan, C. Tseng, C. Chen, C. Hsieh, K. Wong, M. Cheng, et al., "High performance 22/20nm FinFET CMOS devices with advanced high-K/metal gate scheme", in 2010 International Electron Devices Meeting (IEEE, 2010), pp. 27–1.

[6] S. Mukesh and J. Zhang, "A review of the gate-all-around nanosheet FET process opportunities", Electronics **11**, 3589 (2022).

[7] G. Iannaccone, C. Sbrana, I. Morelli, and S. Strangio, "Power electronics based on wide-bandgap semiconductors: Opportunities and challenges", IEEE Access **9**, 139446–139456 (2021).

[8] T. Schram, S. Sutar, I. Radu, and I. Asselberghs, "Challenges of wafer-scale integration of 2D semiconductors for high-performance transistor circuits", Advanced Materials **34**, 2109796 (2022).

[9] Y. Illarionov et al., "Insulators for 2D Nanoelectronics: The Gap to Bridge", Nature Communications **11**, 3385 (2020).

[10] Y. Miura and Y. Matukura, "Investigation of Silicon-Silicon Dioxide Interface Using MOS Structure", Japanese Journal of Applied Physics **5**, 180 (1966).

[11] D. K. Schroder, "Negative bias temperature instability: What do we understand?", Microelectronics Reliability **47**, 841–852 (2007).

[12] S. Zafar, A. Kumar, E. Gusev, and E. Cartier, “Threshold voltage instabilities in high-κ gate dielectric stacks”, IEEE Transactions on Device and Materials Reliability **5**, 45–64 (2005).

[13] T. Grasser, T. Aichinger, G. Pobegen, H. Reisinger, P.-J. Wagner, J. Franco, M. Nelhiebel, and B. Kaczer, “The ‘permanent’ component of NBTI: Composition and annealing”, in 2011 International Reliability Physics Symposium (IEEE, 2011), 6A–2.

[14] T. Grasser et al., “The Paradigm Shift in Understanding the Bias Temperature Instability: From Reaction–Diffusion to Switching Oxide Traps”, IEEE Transactions on Electron Devices **58**, 3652–3666 (2011).

[15] “Gate-sided hydrogen release as the origin of ”permanent” NBTI degradation: From single defects to lifetimes”, in 2015 IEEE International Electron Devices Meeting (IEDM) (2015), pp. 20.1.1–20.1.4.

[16] T. Grasser, K. Rott, H. Reisinger, M. Waltl, J. Franco, and B. Kaczer, “A unified perspective of RTN and BTI”, in 2014 IEEE International Reliability Physics Symposium (2014), 4A.5.1–4A.5.7.

[17] T. Grasser, “Stochastic charge trapping in oxides: From random telegraph noise to bias temperature instabilities”, Microelectronics Reliability **52**, 39–70 (2012).

[18] M. Houssa, M. Tuominen, M. Naili, V. Afanas’ ev, A. Stesmans, S. Haukka, and M. M. Heyns, “Trap-assisted tunneling in high permittivity gate dielectric stacks”, Journal of Applied Physics **87**, 8615–8620 (2000).

[19] L. Larcher, “Statistical simulation of leakage currents in MOS and flash memory devices with a new multiphonon trap-assisted tunneling model”, IEEE Transactions on Electron Devices **50**, 1246–1253 (2003).

[20] B. Ruch, M. Jech, G. Pobegen, and T. Grasser, “Applicability of Shockley–Read–Hall theory for interface states”, IEEE Transactions on Electron Devices **68**, 2092–2097 (2021).

[21] E. Amat, T. Kauerauf, R. Degraeve, R. Rodríguez, M. Nafría, X. Aymerich, and G. Groeseneken, “Channel hot-carrier degradation in pMOS and nMOS short channel transistors with high-k dielectric stack”, Microelectronic Engineering **87**, 47–50 (2010).

[22] T. Grasser, K. Rott, H. Reisinger, M. Waltl, F. Schanovsky, and B. Kaczer, “NBTI in Nanoscale MOSFETs – The Ultimate Modeling Benchmark”, IEEE Transactions on Electron Devices **61**, 3586–3593 (2014).

[23] W. Goes, Y. Wimmer, A.-M. El-Sayed, G. Rzepa, M. Jech, A. L. Shluger, and T. Grasser, “Identification of oxide defects in semiconductor devices: A systematic approach linking DFT to rate equations and experimental evidence”, Microelectronics Reliability **87**, 286–320 (2018).

[24] Y. Wimmer, A.-M. El-Sayed, W. Gös, T. Grasser, and A. L. Shluger, "Role of hydrogen in volatile behaviour of defects in SiO_2-based electronic devices", Proceedings of the Royal Society A **472**, 20160009 (2016).

[25] T. Grasser et al., "On the microscopic structure of hole traps in pMOSFETs", in 2014 IEEE International Electron Devices Meeting (2014), pp. 21.1.1–21.1.4.

[26] G. Rzepa et al., "Comphy - A Compact-Physics Framework for Unified Modeling of BTI", Microelectronics Reliability **85**, 49–65 (2018).

[27] H. Shull and G. G. Hall, "Atomic Units", Nature **184**, 1559–1560 (1959).

[28] M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln", Annalen der Physik **389**, 457–484 (1927).

[29] J. C. Tully, "Perspective on "Zur Quantentheorie der Molekeln" Born M, Oppenheimer R (1927) Ann Phys 84: 457", Theoretical Chemistry Accounts **103**, 173–176 (2000).

[30] M. Born and K. Huang, *Dynamical theory of crystal lattices* (Oxford university press, 1996).

[31] G. A. Worth and L. S. Cederbaum, "Beyond Born-Oppenheimer: Molecular Dynamics Through a Conical Intersection", Annu. Rev. Phys. Chem. **55**, 127–158 (2004).

[32] A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Books on Chemistry (Dover Publications, 1996).

[33] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas", Physical Review **136**, B864 (1964).

[34] S. Mohr, L. E. Ratcliff, L. Genovese, D. Caliste, P. Boulanger, S. Goedecker, and T. Deutsch, "Accurate and efficient linear scaling DFT calculations with universal applicability", Physical Chemistry Chemical Physics **17**, 31360–31370 (2015).

[35] E. Engel and R. M. Dreizler, *Density Functional Theory* (Springer Berlin Heidelberg, 2011).

[36] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects", Physical Review **140**, A1133 (1965).

[37] D. R. Hartree, "The wave mechanics of an atom with a non-coulomb central field. Part II. Some results and discussion", in Mathematical Proceedings of the Cambridge Philosophical Society, Vol. 24, 1 (Cambridge University Press, 1928), pp. 111–132.

[38] J. C. Slater, "Note on Hartree's method", Physical Review **35**, 210 (1930).

[39] V. Fock, "Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems", Zeitschrift für Physik **61**, 126–148 (1930).

[40] R. O. Jones and O. Gunnarsson, “The density functional formalism, its applications and prospects”, Reviews of Modern Physics **61**, 689 (1989).

[41] J. F. Janak, “Proof that $\partial E/\partial n_i = \varepsilon$ in density-functional theory”, Physical Review B **18**, 7165 (1978).

[42] W. C. Witt, B. G. Del Rio, J. M. Dieterich, and E. A. Carter, “Orbital-free density functional theory for materials research”, Journal of Materials Research **33**, 777–795 (2018).

[43] V. Gavini, K. Bhattacharya, and M. Ortiz, “Quasi-continuum orbital-free density-functional theory: A route to multi-million atom non-periodic DFT calculation”, Journal of the Mechanics and Physics of Solids **55**, 697–718 (2007).

[44] W. Mi and M. Pavanello, “Orbital-free density functional theory correctly models quantum dots when asymptotics, nonlocality, and nonhomogeneity are accounted for”, Physical Review B **100**, 041105 (2019).

[45] V. V. Karasiev and S. B. Trickey, “Issues and challenges in orbital-free density functional calculations”, Computer Physics Communications **183**, 2519–2527 (2012).

[46] G. E. Scuseria and V. N. Staroverov, “Progress in the development of exchange-correlation functionals”, in *Theory and Applications of Computational Chemistry* (Elsevier, 2005), pp. 669–724.

[47] J. Toulouse, “Review of Approximations for the Exchange-Correlation Energy in Density-Functional Theory”, in *Density functional theory: modeling, mathematical analysis, computational methods, and applications*, edited by E. Cances and G. Friesecke (Springer International Publishing, Cham, 2023), pp. 1–90.

[48] L. H. Thomas, “The calculation of atomic fields”, in Mathematical Proceedings of the Cambridge Philosophical Society, Vol. 23, 5 (Cambridge University Press, 1927), pp. 542–548.

[49] E. Fermi, “Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihre Anwendung auf die Theorie des periodischen Systems der Elemente”, Zeitschrift für Physik **48**, 73–79 (1928).

[50] D. M. Ceperley and B. J. Alder, “Ground state of the electron gas by a stochastic method”, Physical Review Letters **45**, 566 (1980).

[51] A. Van de Walle and G. Ceder, “Correcting overbinding in local-density-approximation calculations”, Physical Review B **59**, 14992 (1999).

[52] M. Reticcioli, U. Diebold, and C. Franchini, “Modeling polarons in density functional theory: lessons learned from TiO_2”, Journal of Physics: Condensed Matter **34**, 204006 (2022).

[53] D. C. Langreth and M. J. Mehl, “Beyond the local-density approximation in calculations of ground-state electronic properties”, Physical Review B **28**, 1809 (1983).

[54] P. Haas, F. Tran, P. Blaha, K. Schwarz, and R. Laskowski, “Insight into the performance of GGA functionals for solid-state calculations”, Physical Review B **80**, 195109 (2009).

[55] J. P. Perdew, K. Burke, and M. Ernzerhof, “Generalized gradient approximation made simple”, Physical Review Letters **77**, 3865 (1996).

[56] J. Sun, A. Ruzsinszky, and J. P. Perdew, “Strongly constrained and appropriately normed semilocal density functional”, Physical Review Letters **115**, 036402 (2015).

[57] D. Wickramaratne and J. L. Lyons, “Assessing the SCAN functional for deep defects and small polarons in wide band gap semiconductors and insulators”, Physical Review B **109**, 245201 (2024).

[58] M. Lundberg and P. Siegbahn, “Quantifying the effects of the self-interaction error in DFT: When do the delocalized states appear?”, The Journal of Chemical Physics **122**, 224103 (2005).

[59] Y. Zhang and W. Yang, “A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons”, The Journal of Chemical Physics **109**, 2604–2608 (1998).

[60] P. Borlido, J. Schmidt, A. W. Huran, F. Tran, M. A. Marques, and S. Botti, “Exchange-correlation functionals for band gaps of solids: benchmark, reparametrization and machine learning”, npj Computational Materials **6**, 1–17 (2020).

[61] J. P. Perdew, W. Yang, K. Burke, Z. Yang, E. Gross, M. Scheffler, G. E. Scuseria, T. M. Henderson, I. Y. Zhang, A. Ruzsinszky, et al., “Understanding band gaps of solids in generalized Kohn–Sham theory”, Proceedings of the National Academy of Sciences **114**, 2801–2806 (2017).

[62] C. Adamo and V. Barone, “Toward reliable density functional methods without adjustable parameters: The PBE0 model”, The Journal of Chemical Physics **110**, 6158–6170 (1999).

[63] J. P. Perdew, M. Ernzerhof, and K. Burke, “Rationale for mixing exact exchange with density functional approximations”, The Journal of Chemical Physics **105**, 9982–9985 (1996).

[64] A. Alkauskas, P. Broqvist, and A. Pasquarello, “Defect levels through hybrid density functionals: Insights and applications”, physica status solidi (b) **248**, 775–789 (2011).

[65] M. Marques, J. Vidal, M. Oliveira, L. Reining, and S. Botti, “Density-based mixing parameter for hybrid functionals”, Physical Review B **83**, 035119 (2011).

[66] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, “Influence of the exchange screening parameter on the performance of screened hybrid functionals”, The Journal of Chemical Physics **125**, 224106 (2006).

[67] M. Guidon, J. Hutter, and J. VandeVondele, "Robust periodic Hartree-Fock exchange for large-scale simulations using Gaussian basis sets", Journal of Chemical Theory and Computation **5**, 3010–3021 (2009).

[68] P. Deák, B. Aradi, T. Frauenheim, E. Janzén, and A. Gali, "Accurate defect levels obtained from the HSE06 range-separated hybrid functional", Physical Review B **81**, 153203 (2010).

[69] P. D. Haynes, "Linear-scaling methods in ab initio quantum-mechanical calculations", PhD Thesis, University of Cambridge (1998).

[70] W. E. Pickett, "Pseudopotential methods in condensed matter applications", Computer Physics Reports **9**, 115–197 (1989).

[71] C. Herring, "A new method for calculating wave functions in crystals", Physical Review **57**, 1169 (1940).

[72] K. Laasonen, R. Car, C. Lee, and D. Vanderbilt, "Implementation of ultrasoft pseudopotentials in ab initio molecular dynamics", Physical Review B **43**, 6796 (1991).

[73] P. E. Blöchl, "Projector augmented-wave method", Physical Review B **50**, 17953 (1994).

[74] G. Kresse and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method", Physical Review B **59**, 1758 (1999).

[75] S. Goedecker, M. Teter, and J. Hutter, "Separable dual-space Gaussian pseudopotentials", Physical Review B **54**, 1703 (1996).

[76] P. Makkar and N. N. Ghosh, "A review on the use of DFT for the prediction of the properties of nanomaterials", RSC Advances **11**, 27897–27924 (2021).

[77] G. Kresse and J. Hafner, "Ab initio molecular dynamics for liquid metals", Physical Review B **47**, 558 (1993).

[78] G. Kresse and J. Furthmüller, "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set", Computational Materials Science **6**, 15–50 (1996).

[79] G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set", Physical Review B **54**, 11169 (1996).

[80] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, et al., "Advanced capabilities for materials modelling with Quantum ESPRESSO", Journal of Physics: Condensed Matter **29**, 465901 (2017).

[81] P. Giannozzi, O. Baseggio, P. Bonfà, D. Brunato, R. Car, I. Carnimeo, C. Cavazzoni, S. De Gironcoli, P. Delugas, F. Ferrari Ruffino, et al., "Quantum ESPRESSO toward the exascale", The Journal of Chemical Physics **152**, 154105 (2020).

[82] A. Janotti and C. G. Van de Walle, “Native point defects in ZnO”, Physical Review B **76**, 165202 (2007).

[83] I.-T. Lu, J.-J. Zhou, and M. Bernardi, “Efficient ab initio calculations of electron-defect scattering and defect-limited carrier mobility”, Physical Review Materials **3**, 033804 (2019).

[84] S. Tosoni, C. Tuma, J. Sauer, B. Civalleri, and P. Ugliengo, “A comparison between plane wave and Gaussian-type orbital basis sets for hydrogen bonded systems: Formic acid as a test case”, The Journal of Chemical Physics **127**, 154102 (2007).

[85] T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, et al., “CP2K: an electronic structure and molecular dynamics software package-Quickstep: efficient and accurate electronic structure calculations”, The Journal of Chemical Physics **152**, 194103 (2020).

[86] E. Apra, E. J. Bylaska, W. A. De Jong, N. Govind, K. Kowalski, T. P. Straatsma, M. Valiev, H. van Dam, Y. Alexeev, J. Anchell, et al., “NWChem: Past, present, and future”, The Journal of Chemical Physics **152**, 184102 (2020).

[87] A. Garcia, N. Papior, A. Akhtar, E. Artacho, V. Blum, E. Bosoni, P. Brandimarte, M. Brandbyge, J. I. Cerda, F. Corsetti, et al., “Siesta: Recent developments and applications”, The Journal of Chemical Physics **152**, 204108 (2020).

[88] S. F. Boys, “Electronic wave functions-I. A general method of calculation for the stationary states of any molecular system”, Proceedings of the Royal Society of London. Series A. **200**, 542–554 (1950).

[89] R. C. Raffenetti, “General contraction of Gaussian atomic orbitals: Core, valence, polarization, and diffuse basis sets; Molecular integral evaluation”, The Journal of Chemical Physics **58**, 4452–4458 (1973).

[90] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, “Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach”, Computer Physics Communications **167**, 103–128 (2005).

[91] J. VandeVondele and J. Hutter, “An efficient orbital transformation method for electronic structure calculations”, The Journal of Chemical Physics **118**, 4365–4369 (2003).

[92] R. P. Feynman, “Forces in molecules”, Physical Review **56**, 340 (1939).

[93] P. Pulay, “Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules: I. Theory”, Molecular Physics **17**, 197–204 (1969).

[94] G. Lippert, J. Hutter, and M. Parrinello, “A hybrid Gaussian and plane wave density functional scheme”, Molecular Physics **92**, 477–488 (1997).

[95] X. Chen, J.-M. Langlois, and W. A. Goddard III, “Dual-space approach for density-functional calculations of two-and three-dimensional crystals using Gaussian basis functions”, Physical Review B **52**, 2348 (1995).

[96] P. Ordejón, E. Artacho, and J. M. Soler, “Self-consistent order-N density-functional calculations for very large systems”, Physical Review B **53**, R10441 (1996).

[97] J. W. Cooley, P. A. W. Lewis, and P. D. Welch, “The Fast Fourier Transform and Its Applications”, IEEE Transactions on Education **12**, 27–34 (1969).

[98] J. VandeVondele and J. Hutter, “Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases”, The Journal of Chemical Physics **127**, 114105 (2007).

[99] M. Guidon, J. Hutter, and J. VandeVondele, “Auxiliary density matrix methods for Hartree- Fock exchange calculations”, Journal of Chemical Theory and Computation **6**, 2348–2364 (2010).

[100] X. Gonze, “Perturbation expansion of variational principles at arbitrary order”, Physical Review A **52**, 1086 (1995).

[101] J. G. Ángyán, “Wigner’s (2 n+ 1) rule for nonlinear Schrödinger equations”, Journal of Mathematical Chemistry **46**, 1–14 (2009).

[102] P. K. Lam and M. L. Cohen, “Ab initio calculation of phonon frequencies of Al”, Physical Review B **25**, 6139 (1982).

[103] S. Poncé, G. Antonius, Y. Gillet, P. Boulanger, J. Laflamme Janssen, A. Marini, M. Côté, and X. Gonze, “Temperature dependence of electronic eigenenergies in the adiabatic harmonic approximation”, Physical Review B **90**, 214304 (2014).

[104] S. Poncé, E. R. Margine, and F. Giustino, “Towards predictive many-body calculations of phonon-limited carrier mobilities in semiconductors”, Physical Review B **97**, 121201 (2018).

[105] S. Baroni, S. De Gironcoli, A. Dal Corso, and P. Giannozzi, “Phonons and related crystal properties from density-functional perturbation theory”, Reviews of Modern Physics **73**, 515 (2001).

[106] H. Eyring, “The activated complex in chemical reactions”, The Journal of Chemical Physics **3**, 107–115 (1935).

[107] M. G. Evans and M. Polanyi, “Some applications of the transition state method to the calculation of reaction velocities, especially in solution”, Transactions of the Faraday Society **31**, 875–894 (1935).

[108] M. Jech, “The Physics of Non-Equilibrium Reliability Phenomena”, PhD Thesis, TU Wien (2020).

[109] A. Nitzan, *Chemical Dynamics in Condensed Phases: Relaxation, Transfer, and Reactions in Condensed Molecular Systems* (Oxford University Press, 2006).

[110] S. Arrhenius, “Über die Dissociationswärme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte”, Zeitschrift für physikalische Chemie **4**, 96–116 (1889).

[111] H. Jónsson, G. Mills, and K. W. Jacobsen, “Nudged elastic band method for finding minimum energy paths of transitions”, in *Classical and quantum dynamics in condensed phase simulations* (World Scientific, 1998), pp. 385–404.

[112] R. Elber and M. Karplus, “A method for determining reaction paths in large molecules: Application to myoglobin”, Chemical Physics Letters **139**, 375–380 (1987).

[113] G. Henkelman and H. Jónsson, “Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points”, The Journal of Chemical Physics **113**, 9978–9985 (2000).

[114] G. Henkelman, B. P. Uberuaga, and H. Jónsson, “A climbing image nudged elastic band method for finding saddle points and minimum energy paths”, The Journal of Chemical Physics **113**, 9901–9904 (2000).

[115] S. Smidstrup, A. Pedersen, K. Stokbro, and H. Jónsson, “Improved initial guess for minimum energy path calculations”, The Journal of Chemical Physics **140**, 214106 (2014).

[116] X. Zhu, K. C. Thompson, and T. J. Martínez, “Geodesic interpolation for reaction pathways”, The Journal of Chemical Physics **150**, 164103 (2019).

[117] E. L. Kolsbjerg, M. N. Groves, and B. Hammer, “An automated nudged elastic band method”, The Journal of Chemical Physics **145**, 094107 (2016).

[118] O.-P. Koistinen, F. B. Dagbjartsdóttir, V. Ásgeirsson, A. Vehtari, and H. Jónsson, “Nudged elastic band calculations accelerated with Gaussian process regression”, The Journal of Chemical Physics **147**, 152720 (2017).

[119] J. A. Garrido Torres, P. C. Jennings, M. H. Hansen, J. R. Boes, and T. Bligaard, “Low-scaling algorithm for nudged elastic band calculations using a surrogate machine learning model”, Physical Review Letters **122**, 156001 (2019).

[120] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, Vol. 2, 3 (MIT press Cambridge, MA, 2006).

[121] G. Henkelman and H. Jónsson, “A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives”, The Journal of Chemical Physics **111**, 7010–7022 (1999).

[122] L. K. Béland, P. Brommer, F. El-Mellouhi, J.-F. Joly, and N. Mousseau, “Kinetic activation-relaxation technique”, Physical Review E **84**, 046704 (2011).

[123] A. Jay, C. Huet, N. Salles, M. Gunde, L. Martin-Samos, N. Richard, G. Landa, V. Goiffon, S. De Gironcoli, A. Hémeryck, et al., “Finding reaction pathways and transition states: r-ARTn and d-ARTn as an efficient and versatile alternative to string approaches”, Journal of Chemical Theory and Computation **16**, 6726–6734 (2020).

[124] A. Jay, M. Gunde, N. Salles, M. Poberžnik, L. Martin-Samos, N. Richard, S. de Gironcoli, N. Mousseau, and A. Hémeryck, “Activation–Relaxation Technique: An efficient way to find minima and saddle points of potential energy surfaces”, Computational Materials Science **209**, 111363 (2022).

[125] X. W. Zhou, D. A. Murdick, B. Gillespie, and H. N. G. Wadley, “Atomic assembly during GaN film growth: Molecular dynamics simulations”, Physical Review B **73**, 045337 (2006).

[126] Z. Fan, H. Dong, A. Harju, and T. Ala-Nissila, “Homogeneous nonequilibrium molecular dynamics method for heat transport and spectral decomposition with many-body potentials”, Physical Review B **99**, 064308 (2019).

[127] M. Puligheddu, F. Gygi, and G. Galli, “First-principles simulations of heat transport”, Physical Review Materials **1**, 060802 (2017).

[128] B. Liu, Z. Xu, Y. Wang, X. Gao, and R. Kong, “Effect of ion implantation on material removal mechanism of 6H-SiC in nano-cutting: A molecular dynamics study”, Computational Materials Science **174**, 109476 (2020).

[129] K. Leung and J. L. Budzien, “Ab initio molecular dynamics simulations of the initial stages of solid–electrolyte interphase formation on lithium ion battery graphitic anodes”, Physical Chemistry Chemical Physics **12**, 6583–6586 (2010).

[130] L. Verlet, “Computerexperiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules”, Physical Review **159**, 98 (1967).

[131] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, “A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters”, The Journal of Chemical Physics **76**, 637–649 (1982).

[132] D. Fincham, “Choice of timestep in molecular dynamics simulation”, Computer Physics Communications **40**, 263–269 (1986).

[133] J. Sarnthein, A. Pasquarello, and R. Car, “Model of vitreous SiO_2 generated by an ab initio molecular-dynamics quench from the melt”, Physical Review B **52**, 12690 (1995).

[134] P. H. Hünenberger, “Thermostat Algorithms for Molecular Dynamics Simulations”, in *Advanced Computer Simulation: Approaches for Soft Matter Sciences I*, edited by D. Holm and P. D. Kremer (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005), pp. 105–149.

[135] H. J. C. Berendsen, J. Postma, W. F. Van Gunsteren, A. DiNola, and J. R. Haak, “Molecular dynamics with coupling to an external bath”, The Journal of Chemical Physics **81**, 3684–3690 (1984).

[136] S. C. Harvey, R. K.-Z. Tan, and T. E. Cheatham III, “The flying ice cube: velocity rescaling in molecular dynamics leads to violation of energy equipartition”, Journal of Computational Chemistry **19**, 726–740 (1998).

[137] G. Bussi, D. Donadio, and M. Parrinello, “Canonical sampling through velocity rescaling”, The Journal of Chemical Physics **126**, 014101 (2007).

[138] S. Nosé, “A unified formulation of the constant temperature molecular dynamics methods”, The Journal of Chemical Physics **81**, 511–519 (1984).

[139] W. G. Hoover, “Canonical dynamics: Equilibrium phase-space distributions”, Physical Review A **31**, 1695 (1985).

[140] H. A. Posch, W. G. Hoover, and F. J. Vesely, “Canonical dynamics of the Nosé oscillator: Stability, order, and chaos”, Physical Review A **33**, 4253 (1986).

[141] G. J. Martyna, M. L. Klein, and M. Tuckerman, “Nosé–Hoover chains: The canonical ensemble via continuous dynamics”, The Journal of Chemical Physics **97**, 2635–2643 (1992).

[142] E. Paquet and H. L. Viktor, “Molecular dynamics, monte carlo simulations, and Langevin dynamics: A computational review”, BioMed Research International **2015**, 183918 (2015).

[143] R. Kubo, “The fluctuation-dissipation theorem”, Reports on Progress in Physics **29**, 255 (1966).

[144] G. J. Martyna, D. J. Tobias, and M. L. Klein, “Constant pressure molecular dynamics algorithms”, The Journal of Chemical Physics **101**, 4177–4189 (1994).

[145] C. Perego, M. Salvalaglio, and M. Parrinello, “Molecular dynamics simulations of solutions at constant chemical potential”, The Journal of Chemical Physics **142**, 144113 (2015).

[146] T. E. Markland and M. Ceriotti, “Nuclear quantum effects enter the mainstream”, Nature Reviews Chemistry **2**, 0109 (2018).

[147] H. S. Mei, M. E. Tuckerman, D. E. Sagnella, and M. L. Klein, “Quantum nuclear ab initio molecular dynamics study of water wires”, The Journal of Physical Chemistry B **102**, 10446–10458 (1998).

[148] Y. Litman and M. Rossi, “Multidimensional hydrogen tunneling in supported molecular switches: The role of surface interactions”, Physical Review Letters **125**, 216001 (2020).

[149] A. Cahlik, J. Hellerstedt, J. I. Mendieta-Moreno, M. Svec, V. M. Santhini, S. Pascal, D. Soler-Polo, S. I. Erlingsson, K. Vyborny, P. Mutombo, et al., “Significance of nuclear quantum effects in hydrogen bonded molecular chains”, ACS nano **15**, 10357–10365 (2021).

[150] J. Guo, J.-T. Lü, Y. Feng, J. Chen, J. Peng, Z. Lin, X. Meng, Z. Wang, X.-Z. Li, E.-G. Wang, et al., "Nuclear quantum effects of hydrogen bonds probed by tip-enhanced inelastic electron tunneling", Science **352**, 321–325 (2016).

[151] O. Kirshenboim, A. Frenklah, and S. Kozuch, "Switch chemistry at cryogenic conditions: quantum tunnelling under electric fields", Chemical Science **12**, 3179–3187 (2021).

[152] S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. Miller III, "Ring-polymer molecular dynamics: Quantum effects in chemical dynamics from classical trajectories in an extended phase space", Annual Review of Physical Chemistry **64**, 387–413 (2013).

[153] J. O. Richardson, "Perspective: Ring-polymer instanton theory", The Journal of Chemical Physics **148**, 200901 (2018).

[154] C. Pisani, F. Corà, R. Nada, and R. Orlando, "Hartree-Fock perturbed-cluster treatment of local defects in crystals: I. The EMBED program: general features", Computer Physics Communications **82**, 139–156 (1994).

[155] P. V. Sushko, S. Mukhopadhyay, A. S. Mysovsky, V. B. Sulimov, A. Taga, and A. L. Shluger, "Structure and properties of defects in amorphous silica: new insights from embedded cluster calculations", Journal of Physics: Condensed Matter **17**, S2115 (2005).

[156] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, "First-principles calculations for point defects in solids", Reviews of Modern Physics **86**, 253–305 (2014).

[157] C. G. Van de Walle and J. Neugebauer, "First-principles calculations for defects and impurities: Applications to III-nitrides", Journal of Applied Physics **95**, 3851–3879 (2004).

[158] S. B. Zhang and J. E. Northrup, "Chemical potential dependence of defect formation energies in GaAs: Application to Ga self-diffusion", Physical Review Letters **67**, 2339 (1991).

[159] G. Makov and M. C. Payne, "Periodic boundary conditions in ab initio calculations", Physical Review B **51**, 4014–4022 (1995).

[160] H.-P. Komsa, T. T. Rantala, and A. Pasquarello, "Finite-size supercell correction schemes for charged defect calculations", Physical Review B **86**, 045112 (2012).

[161] P. P. Ewald, "Die Berechnung optischer und elektrostatischer Gitterpotentiale", Annalen der Physik **369**, 253–287 (1921).

[162] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, "Fully Ab Initio Finite-Size Corrections for Charged-Defect Supercell Calculations", Physical Review Letters **102**, 016402 (2009).

[163] C. Freysoldt, "Manual for SXDEFECTALIGN version 3.0", (2022).

[164] M. E. Turiansky, A. Alkauskas, M. Engel, G. Kresse, D. Wickramaratne, J.-X. Shen, C. E. Dreyer, and C. G. Van de Walle, “Nonrad: Computing nonradiative capture coefficients from first principles”, Computer Physics Communications **267**, 108056 (2021).

[165] T. Grasser, H. Reisinger, K. Rott, M. Toledano-Luque, and B. Kaczer, “On the microscopic origin of the frequency dependence of hole trapping in pMOSFETs”, in 2012 International Electron Devices Meeting (IEDM) (2012), pp. 19.6.1–19.6.4.

[166] C. G. Broyden, “The convergence of a class of double-rank minimization algorithms 1. general considerations”, IMA Journal of Applied Mathematics **6**, 76–90 (1970).

[167] R. Fletcher, “A new approach to variable metric algorithms”, The Computer Journal **13**, 317–322 (1970).

[168] D. Goldfarb, “A family of variable-metric methods derived by variational means”, Mathematics of Computation **24**, 23–26 (1970).

[169] D. F. Shanno, “Conditioning of quasi-Newton methods for function minimization”, Mathematics of Computation **24**, 647–656 (1970).

[170] A. Stoneham, “Non-radiative transitions in semiconductors”, Reports on Progress in Physics **44**, 1251 (1981).

[171] E. D. Siggia and P. C. Kwok, “Properties of electrons in semiconductor inversion layers with many occupied electric subbands. I. Screening and impurity scattering”, Physical Review B **2**, 1024 (1970).

[172] C. Vargas, Y. Zhu, G. Coletti, C. Chan, D. Payne, M. Jensen, and Z. Hameiri, “Recombination parameters of lifetime-limiting carrier-induced defects in multicrystalline silicon for solar cells”, Applied Physics Letters **110**, 092106 (2017).

[173] H. Cheng, Y. Feng, Y. Fu, Y. Zheng, Y. Shao, and Y. Bai, “Understanding and minimizing non-radiative recombination losses in perovskite light-emitting diodes”, Journal of Materials Chemistry C **10**, 13590–13610 (2022).

[174] A. Alkauskas, Q. Yan, and C. G. Van de Walle, “First-principles theory of nonradiative carrier capture via multiphonon emission”, Physical Review B **90**, 075202 (2014).

[175] W. Shockley and W. T. Read Jr, “Statistics of the recombinations of holes and electrons”, Physical Review **87**, 835 (1952).

[176] R. N. Hall, “Electron-hole recombination in germanium”, Physical Review **87**, 387 (1952).

[177] V. N. Abakumov, V. I. Perel, and I. N. Yassievich, *Nonradiative recombination in semiconductors* (Elsevier, 1991).

[178] R. Pässler, “Calculation of nonradiative multiphonon capture coefficients and ionization rates for neutral centres according to the static coupling scheme: I. Theory”, physica status solidi (b) **68**, 69–79 (1975).

[179] L. Shi, K. Xu, and L.-W. Wang, “Comparative study of ab initio nonradiative recombination rate calculations under different formalisms”, Physical Review B **91**, 205315 (2015).

[180] R. Pässler, “Relationships between the nonradiative multiphonon carrier-capture properties of deep charged and neutral centres in semiconductors”, physica status solidi (b) **78**, 625–635 (1976).

[181] L. Shi and L.-W. Wang, “Ab initio Calculations of Deep-Level Carrier Nonradiative Recombination Rates in Bulk Semiconductors”, Physical Review Letters **109**, 245501 (2012).

[182] G. D. Barmparis, Y. S. Puzyrev, X.-G. Zhang, and S. T. Pantelides, “Theory of inelastic multiphonon scattering and carrier capture by defects in semiconductors: Application to capture cross sections”, Physical Review B **92**, 214111 (2015).

[183] A. Baiardi, J. Bloino, and V. Barone, “General time dependent approach to vibronic spectroscopy including Franck–Condon, Herzberg–Teller, and Duschinsky effects”, Journal of Chemical Theory and Computation **9**, 4097–4115 (2013).

[184] R. Borrelli, A. Capobianco, and A. Peluso, “Generating function approach to the calculation of spectral band shapes of free-base chlorin including Duschinsky and Herzberg–Teller effects”, The Journal of Physical Chemistry A **116**, 9934–9940 (2012).

[185] V. Afanas’ev and A. Stesmans, “Hydrogen release related to hole injection into SiO_2 layers on Si”, Materials Science in Semiconductor Processing **4**, 149–151 (2001).

[186] C. H. Henry and D. V. Lang, “Nonradiative capture and recombination by multiphonon emission in GaAs and GaP”, Physical Review B **15**, 989–1016 (1977).

[187] A. M. Stoneham, *Theory of Defects in Solids: Electronic Structure of Defects in Insulators and Semiconductors* (Oxford University Press, 2001).

[188] M. Kirton and M. Uren, “Noise in solid-state microstructures: A new perspective on individual defects, interface states and low-frequency (1/f) noise”, Advances in Physics **38**, 367–468 (1989).

[189] J. T. Ryan, A. Matsuda, J. P. Campbell, and K. P. Cheung, “Interface-state capture cross section—Why does it vary so much?”, Applied Physics Letters **106**, 163503 (2015).

[190] K. Huang, A. Rhys, and N. F. Mott, “Theory of light absorption and nonradiative transitions in F-centres”, Proceedings of the Royal Society of London. Series A **204**, 406–423 (1950).

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

[191] A. Alkauskas, M. D. McCluskey, and C. G. Van de Walle, “Tutorial: Defects in semiconductors—Combining experiment and theory”, Journal of Applied Physics **119**, 181101 (2016).

[192] Y. Wimmer, “Hydrogen related defects in amorphous SiO_2 and the negative bias temperature instability”, PhD Thesis, TU Wien (2017).

[193] Y. Yuan, G. Yan, R. Hong, Z. Liang, and T. Kirchartz, “Quantifying efficiency limitations in all-inorganic halide perovskite solar cells”, Advanced Materials **34**, 2108132 (2022).

[194] X. Zhang, M. E. Turiansky, J.-X. Shen, and C. G. Van de Walle, “Defect tolerance in halide perovskites: A first-principles perspective”, Journal of Applied Physics **131**, 090901 (2022).

[195] X. Wang, S. R. Kavanagh, D. O. Scanlon, and A. Walsh, “Upper efficiency limit of Sb_2Se_3 solar cells”, Joule **8**, 2105–2122 (2024).

[196] S. Li, G. Thiering, P. Udvarhelyi, V. Ivády, and A. Gali, “Carbon defect qubit in two-dimensional WS_2”, Nature Communications **13**, 1210 (2022).

[197] U. Lorenz and P. Saalfrank, “A novel system-bath Hamiltonian for vibration-phonon coupling: Formulation, and application to the relaxation of Si–H and Si–D bending modes of H/D: Si (100)-(2×1)”, Chemical Physics **482**, 69–80 (2017).

[198] W. Cheney and D. R. Kincaid, *Linear Algebra: Theory and Applications* (Jones and Bartlett Learning, Burlington, MA, USA, 2011).

[199] M. R. Lorenz, G. D. Pettit, and R. Taylor, “Band gap of gallium phosphide from 0 to 900 K and light emission from diodes at high temperatures”, Physical Review **171**, 876 (1968).

[200] P. Dean, C. Henry, and C. Frosch, “Infrared donor-acceptor pair spectra involving the deep oxygen donor in gallium phosphide”, Physical Review **168**, 812 (1968).

[201] D. Wickramaratne, J.-X. Shen, A. Alkauskas, and C. G. Van de Walle, “Comment on “Comparative study of ab initio nonradiative recombination rate calculations under different formalisms””, Physical Review B **97**, 077301 (2018).

[202] L. Shi, K. Xu, and L.-W. Wang, “Reply to “Comment on ‘Comparative study of ab initio nonradiative recombination rate calculations under different formalisms’””, Physical Review B **97**, 077302 (2018).

[203] A. Togo, “First-principles Phonon Calculations with Phonopy and Phono3py”, Journal of the Physical Society of Japan **92**, 012001 (2023).

[204] A. Togo, L. Chaput, T. Tadano, and I. Tanaka, “Implementation strategies in phonopy and phono3py”, Journal of Physics: Condensed Matter **35**, 353001 (2023).

[205] R. Atta-Fynn, D. A. Drabold, S. R. Elliott, and P. Biswas, “First-principles simulations of vibrational decay and lifetimes in a-Si:H and a-Si:D”, Physical Review B **95**, 104205 (2017).

[206] R. Thapa and D. A. Drabold, “Ab Initio Simulation of Amorphous Materials”, in *Atomistic Simulations of Glasses* (John Wiley & Sons, Ltd, 2022) Chap. 2, pp. 30–59.

[207] V. A. Jhalani, J.-J. Zhou, J. Park, C. E. Dreyer, and M. Bernardi, “Piezoelectric electron-phonon interaction from ab initio dynamical quadrupoles: Impact on charge transport in wurtzite GaN”, Physical Review Letters **125**, 136602 (2020).

[208] J. Park, J.-J. Zhou, V. A. Jhalani, C. E. Dreyer, and M. Bernardi, “Long-range quadrupole electron-phonon interaction from first principles”, Physical Review B **102**, 125203 (2020).

[209] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, 3rd ed. (Cambridge University Press, 2020).

[210] J. Franck and E. G. Dymond, “Elementary processes of photochemical reactions”, Transactions of the Faraday Society **21**, 536–542 (1926).

[211] E. Condon, “A Theory of Intensity Distribution in Band Systems”, Physical Review **28**, 1182–1201 (1926).

[212] J. Jortner, “Temperature dependent activation energy for electron transfer between biological molecules”, The Journal of Chemical Physics **64**, 4860–4867 (1976).

[213] S. H. Lin, C. H. Chang, K. K. Liang, R. Chang, Y. J. Shiu, J. M. Zhang, T.-S. Yang, M. Hayashi, and F. C. Hsu, “Ultrafast Dynamics and Spectroscopy of Bacterial Photosynthetic Reaction Centers”, in *Advances in Chemical Physics* (John Wiley & Sons, Ltd, 2002) Chap. 1, pp. 1–88.

[214] M. Lax, “The Franck-Condon principle and its application to crystals”, The Journal of Chemical Physics **20**, 1752–1760 (1952).

[215] W. Goes, F. Schanovsky, and T. Grasser, “Advanced Modeling of Oxide Defects”, in *Bias temperature instability for devices and circuits*, edited by T. Grasser (Springer New York, New York, NY, 2014), pp. 409–446.

[216] R. A. Marcus and N. Sutin, “Electron transfers in chemistry and biology”, Biochimica et Biophysica Acta **811**, 265–322 (1985).

[217] R. A. Marcus, “Electron transfer reactions in chemistry. Theory and experiment”, Review of Modern Physics **65**, 599–610 (1993).

[218] J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer New York, 2006).

[219] T. Chachiyo and J. H. Rodriguez, “A direct method for locating minimum-energy crossing points (MECPs) in spin-forbidden transitions and nonadiabatic reactions”, The Journal of Chemical Physics **123**, 094711 (2005).

[220] Y. Jia, S. Poncé, A. Miglio, M. Mikami, and X. Gonze, "Beyond the one-dimensional configuration coordinate model of photoluminescence", Physical Review B **100**, 155109 (2019).

[221] R. Chau, S. Datta, and A. Majumdar, "Opportunities and challenges of III-V nanoelectronics for future high-speed, low-power logic applications", in IEEE Compound Semiconductor Integrated Circuit Symposium, 2005. CSIC'05. (IEEE, 2005), p. 4.

[222] F. Roccaforte, P. Fiorenza, G. Greco, R. L. Nigro, F. Giannazzo, F. Iucolano, and M. Saggio, "Emerging trends in wide band gap semiconductors (SiC and GaN) technology for power devices", Microelectronic Engineering **187**, 66–77 (2018).

[223] J. Robertson, "High dielectric constant gate oxides for metal oxide Si transistors", Reports on Progress in Physics **69**, 327 (2005).

[224] K. Brower, P. Lenahan, and P. Dressendorfer, "Defects and impurities in thermal oxides on silicon", Applied Physics Letters **41**, 251–253 (1982).

[225] P. Lenahan, "Atomic scale defects involved in MOS reliability problems", Microelectronic Engineering **69**, 173–181 (2003).

[226] B. Tuttle and C. G. Van de Walle, "Structure, energetics, and vibrational properties of Si-H bond dissociation in silicon", Physical Review B **59**, 12884 (1999).

[227] A.-M. El-Sayed, Y. Wimmer, W. Goes, T. Grasser, V. V. Afanas' ev, and A. L. Shluger, "Theoretical models of hydrogen-induced defects in amorphous silicon dioxide", Physical Review B **92**, 014107 (2015).

[228] J. Krick, P. Lenahan, and G. Dunn, "Direct observation of interfacial point defects generated by channel hot hole injection in n-channel metal oxide silicon field effect transistors", Applied Physics Letters **59**, 3437–3439 (1991).

[229] P. M. Lenahan and J. Conley Jr, "What can electron paramagnetic resonance tell us about the Si/SiO_2 system?", Journal of Vacuum Science & Technology B **16**, 2134–2153 (1998).

[230] K. Stokbro, C. Thirstrup, M. Sakurai, U. Quaade, B. Y.-K. Hu, F. Perez-Murano, and F. Grey, "STM-induced hydrogen desorption via a hole resonance", Physical Review Letters **80**, 2618 (1998).

[231] W. McMahon and K. Hess, "A multi-carrier model for interface trap generation", Journal of Computational Electronics **1**, 395–398 (2002).

[232] Y.-Y. Liu, Z. Wei, S. Meng, R. Wang, X. Jiang, R. Huang, S.-S. Li, and L.-W. Wang, "Electronically induced defect creation at semiconductor/oxide interface revealed by time-dependent density functional theory", Physical Review B **104**, 115310 (2021).

[233] D. L. Griscom, “Characterization of three E'-center variants in X- and γ-irradiated high purity a-SiO_2”, Nuclear Instruments and Methods in Physics Research Section B **1**, 481–488 (1984).

[234] P. E. Blöchl, “First-principles calculations of defects in oxygen-deficient silica exposed to hydrogen”, Physical Review B **62**, 6158 (2000).

[235] N. Lopez, F. Illas, and G. Pacchioni, “Mechanisms of Proton Formation from Interaction of H_2 with E' and Oxygen Vacancy Centers in SiO_2: Cluster Model Calculations”, The Journal of Physical Chemistry B **104**, 5471–5477 (2000).

[236] R. Van Ginhoven, H. Hjalmarson, A. Edwards, and B. Tuttle, “Hydrogen release in SiO_2: Source sites and release mechanisms”, Nuclear Instruments and Methods in Physics Research Section B **250**, 274–278 (2006).

[237] J. Godet and A. Pasquarello, “Proton diffusion mechanism in amorphous SiO_2”, Physical Review Letters **97**, 155901 (2006).

[238] S. Ling, A.-M. El-Sayed, F. Lopez-Gejo, M. B. Watkins, V. Afanas’ev, and A. L. Shluger, “A computational study of Si–H bonds as precursors for neutral E$'$ centres in amorphous silica and at the Si/SiO_2 interface”, Microelectronic Engineering **109**, 310–313 (2013).

[239] R. Devine and G. Herrera, “Electric-field-induced transport of protons in amorphous SiO_2”, Physical Review B **63**, 233406 (2001).

[240] Z.-C. Hong, P. Yao, Y. Liu, and X. Zuo, “First-principles calculations of the hole-induced depassivation of SiO_2/Si interface defects”, Chinese Physics B **31**, 057101 (2022).

[241] V. Afanas’ev, M. Houssa, A. Stesmans, and M. Heyns, “Electron energy barriers between (100) Si and ultrathin stacks of SiO_2, Al_2O_3, and ZrO_2 insulators”, Applied Physics Letters **78**, 3073–3075 (2001).

[242] T. DiStefano and D. Eastman, “The band edge of amorphous SiO_2 by photoinjection and photoconductivity measurements”, Solid State Communications **9**, 2259–2261 (1971).

[243] G. Will, M. Bellotto, W. Parrish, and M. Hart, “Crystal structures of quartz and magnesium germanate by profile analysis of synchrotron-radiation high-resolution powder data”, Journal of Applied Crystallography **21**, 182–191 (1988).

[244] J. Varley, J. Weber, A. Janotti, and C. Van de Walle, “Dangling bonds, the charge neutrality level, and band alignment in semiconductors”, Journal of Applied Physics **135**, 075703 (2024).

[245] H. Ye, G. Wicks, and P. Fauchet, “Hot electron relaxation time in GaN”, Applied Physics Letters **74**, 711–713 (1999).

[246] A.-M. El-Sayed, M. B. Watkins, T. Grasser, V. V. Afanas' ev, and A. L. Shluger, "Hydrogen-induced rupture of strained Si-O bonds in amorphous silicon dioxide", Physical Review Letters **114**, 115503 (2015).

[247] B. Kaduk, T. Kowalczyk, and T. Van Voorhis, "Constrained density functional theory", Chemical Reviews **112**, 321–370 (2012).

[248] A. T. Gilbert, N. A. Besley, and P. M. Gill, "Self-consistent field calculations of excited states using the maximum overlap method (MOM)", The Journal of Physical Chemistry A **112**, 13164–13171 (2008).

[249] J. R. Chelikowsky and M. Schlüter, "Electron states in α-quartz: A self-consistent pseudopotential calculation", Physical Review B **15**, 4020 (1977).

[250] P. E. Blöchl and J. H. Stathis, "Hydrogen Electrochemistry and Stress-Induced Leakage Current in Silica", Physical Review Letters **83**, 372–375 (1999).

[251] T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, et al., "The ReaxFF reactive force-field: development, applications and future directions", npj Computational Materials **2**, 1–14 (2016).

[252] J. Sarnthein, A. Pasquarello, and R. Car, "Structural and electronic properties of liquid and amorphous SiO_2: An ab initio molecular dynamics study", Physical Review Letters **74**, 4682 (1995).

[253] K. Vollmayr, W. Kob, and K. Binder, "Cooling-rate effects in amorphous silica: A computer-simulation study", Physical Review B **54**, 15808 (1996).

[254] S. Mukhopadhyay, P. V. Sushko, A. M. Stoneham, and A. L. Shluger, "Modeling of the structure and properties of oxygen vacancies in amorphous silica", Physical Review B **70**, 195203 (2004).

[255] B. Van Beest, G. J. Kramer, and R. A. Van Santen, "Force fields for silicas and aluminophosphates based on ab initio calculations", Physical Review Letters **64**, 1955 (1990).

[256] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, "Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons", Physical Review Letters **104**, 136403 (2010).

[257] S. Klawohn, J. P. Darby, J. R. Kermode, G. Csányi, M. A. Caro, and A. P. Bartók, "Gaussian approximation potentials: Theory, software implementation and application examples", The Journal of Chemical Physics **159**, 174108 (2023).

[258] D. Unruh, R. V. Meidanshahi, S. M. Goodnick, G. Csányi, and G. T. Zimányi, "Gaussian approximation potential for amorphous Si: H", Physical Review Materials **6**, 065603 (2022).

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[259] G. Sivaraman, G. Csanyi, A. Vazquez-Mayagoitia, I. T. Foster, S. K. Wilke, R. Weber, and C. J. Benmore, "A combined machine learning and high-energy X-ray diffraction approach to understanding liquid and amorphous metal oxides", Journal of the Physical Society of Japan **91**, 091009 (2022).

[260] J. Behler and M. Parrinello, "Generalized neural-network representation of high-dimensional potential-energy surfaces", Physical Review Letters **98**, 146401 (2007).

[261] R. Lot, F. Pellegrini, Y. Shaidu, and E. Küçükbenli, "PANNA: Properties from Artificial Neural Network Architectures", Computer Physics Communications **256**, 107402 (2020).

[262] R. P. Vedula, N. L. Anderson, and A. Strachan, "Effect of topological disorder on structural, mechanical, and electronic properties of amorphous silicon nitride: An atomistic study", Physical Review B **85**, 205209 (2012).

[263] A.-M. El-Sayed, "Atomistic Modelling of Charge Trapping Defects in Silicon Dioxide", PhD Thesis, University Collge London (2015).

[264] D. L. Price and J. M. Carpenter, "Scattering function of vitreous silica", Journal of Non-Crystalline Solids **92**, 153–174 (1987).

[265] S. R. Elliott, *Physics of amorphous materials* (Longman Scientific & Technical, New York, 1984).

[266] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dulak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, et al., "The atomic simulation environment—a Python library for working with atoms", Journal of Physics: Condensed Matter **29**, 273002 (2017).

[267] P. McMillan, B. Piriou, and R. Couty, "A Raman study of pressure-densified vitreous silica", The Journal of Chemical Physics **81**, 4234–4236 (1984).

[268] D. R. Fredkin, A. Komornicki, S. R. White, and K. R. Wilson, "Ab initio infrared and Raman spectra", The Journal of Chemical Physics **78**, 7077–7092 (1983).

[269] K. L. Yip and W. B. Fowler, "Electronic structure of E_1' centers in SiO_2", Physical Review B **11**, 2327 (1975).

[270] E. P. O'Reilly and J. Robertson, "Theory of defects in vitreous silicon dioxide", Physical Review B **27**, 3780 (1983).

[271] J. K. Rudra and W. B. Fowler, "Oxygen vacancy and the E_1' center in crystalline SiO_2", Physical Review B **35**, 8223 (1987).

[272] M. K. Schurman and M. Tomozawa, "Equilibrium oxygen vacancy concentrations and oxidant diffusion in germania, silica, and germania-silica glasses", Journal of Non-Crystalline Solids **202**, 93–106 (1996).

[273] N. Richard, L. Martin-Samos, G. Roma, Y. Limoge, and J.-P. Crocombette, “First principle study of neutral and charged self-defects in amorphous SiO_2”, Journal of Non-Crystalline Solids **351**, 1825–1829 (2005).

[274] S. T. Pantelides, Z.-Y. Lu, C. Nicklaw, T. Bakos, S. N. Rashkeev, D. M. Fleetwood, and R. D. Schrimpf, “The E' center and oxygen vacancies in SiO_2”, Journal of Non-Crystalline Solids **354**, 217–223 (2008).

[275] E. Poindexter and W. Warren, “Paramagnetic Point Defects in Amorphous Thin Films of SiO_2 and Si_3N_4: Updates and Additions”, Journal of the Electrochemical Society **142**, 2508–2516 (1995).

[276] F. Schanovsky, O. Baumgartner, W. Goes, and T. Grasser, “A detailed evaluation of model defects as candidates for the bias temperature instability”, in 2013 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) (IEEE, 2013), pp. 1–4.

[277] P. M. Lenahan, J. P. Campbell, A. T. Krishnan, and S. Krishnan, “A model for NBTI in nitrided oxide MOSFETs which does not involve hydrogen or diffusion”, IEEE Transactions on Device and Materials Reliability **11**, 219–226 (2010).

[278] 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 **88**, 153518 (2006).

[279] T. Aichinger, S. Puchner, M. Nelhiebel, T. Grasser, and H. Hutter, “Impact of hydrogen on recoverable and permanent damage following negative bias temperature stress”, in 2010 IEEE International Reliability Physics Symposium (IEEE, 2010), pp. 1063–1068.

[280] K. Kajihara, L. Skuja, M. Hirano, and H. Hosono, “In situ observation of the formation, diffusion, and reactions of hydrogenous species in F_2-laser-irradiated SiO_2 glass using a pump-and-probe technique”, Physical Review B **74**, 094202 (2006).

[281] C. Nelson and R. Weeks, “Trapped electrons in irradiated quartz and silica: I, optical absorption”, Journal of the American Ceramic Society **43**, 396–399 (1960).

[282] L. D. Landau, “Electron motion in crystal lattices”, Phys. Z. Sowjet. **3**, 664 (1933).

[283] C. Franchini, M. Reticcioli, M. Setvin, and U. Diebold, “Polarons in materials”, Nature Reviews Materials **6**, 560–586 (2021).

[284] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, and J.-L. Brédas, “Charge transport in organic semiconductors”, Chemical reviews **107**, 926–952 (2007).

[285] D. Cortecchia, J. Yin, A. Bruno, S.-Z. Lo, G. G. Gurzadyan, S. Mhaisalkar, J.-L. Brédas, and C. Soci, "Polaron self-localization in white-light emitting hybrid perovskites", Journal of Materials Chemistry C **5**, 2771–2780 (2017).

[286] J. Strand, P. La Torraca, A. Padovani, L. Larcher, and A. L. Shluger, "Dielectric breakdown in HfO_2 dielectrics: Using multiscale modeling to identify the critical physical processes involved in oxide degradation", Journal of Applied Physics **131**, 234501 (2022).

[287] W. I. Choi, W.-J. Son, R. Dronskowski, Y. Oh, S.-Y. Yang, U. Kwon, and D. S. Kim, "Switchable Chemical-Bond Reorganization for the Stable Charge Trapping in Amorphous Silicon Nitride", Advanced Materials **36**, 2308054 (2024).

[288] D. L. Griscom, "Self-trapped holes in amorphous silicon dioxide", Physical Review B **40**, 4224 (1989).

[289] Y. Sasajima and K. Tanimura, "Optical transitions of self-trapped holes in amorphous SiO_2", Physical Review B **68**, 014204 (2003).

[290] A. V. Kimmel, P. V. Sushko, and A. L. Shluger, "Structure and spectroscopic properties of trapped holes in silica", Journal of Non-Crystalline Solids **353**, 599–604 (2007).

[291] A.-M. El-Sayed, M. B. Watkins, V. V. Afanas' ev, and A. L. Shluger, "Nature of intrinsic and extrinsic electron trapping in SiO_2", Physical Review B **89**, 125201 (2014).

[292] V. V. Afanas'ev and A. Stesmans, "Interfacial Defects in SiO_2 Revealed by Photon Stimulated Tunneling of Electrons", Physical Review Letters **78**, 2437 (1997).

[293] N. S. Saks and A. K. Agarwal, "Hall mobility and free electron density at the SiC/SiO_2 interface in 4H–SiC", Applied Physics Letters **77**, 3281–3283 (2000).

[294] S. Lany and A. Zunger, "Many-body GW calculation of the oxygen vacancy in ZnO", Physical Review B **81**, 113201 (2010).

[295] S. Das, S. Amritanand, E. Pop, C. J. McClellan, A. D. Franklin, T. Grasser, T. Knobloch, Y. Illarionov, et al., "Transistors Based on Two-Dimensional Materials for Future Integrated Circuits", Nature Electronics **4**, 786–799 (2021).

[296] Z. Yu, Z.-Y. Ong, S. Li, J.-B. Xu, G. Zhang, Y.-W. Zhang, Y. Shi, and X. Wang, "Analyzing the Carrier Mobility in Transition-Metal Dichalcogenide MoS_2 Field-Effect Transistors", Advanced Functional Materials **27**, 1604093 (2017).

[297] C. Lee, S. Rathi, M. A. Khan, D. Lim, Y. Kim, S. J. Yun, D.-H. Youn, K. Watanabe, T. Taniguchi, and G.-H. Kim, "Comparison of trapped charges and hysteresis behavior in hBN encapsulated single MoS_2 flake based field effect transistors on SiO_2 and hBN substrates", Nanotechnology **29**, 335202 (2018).

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

[298] T. Knobloch et al., "The Performance Limits of Hexagonal Boron Nitride as an Insulator for Scaled CMOS Devices Based on Two-Dimensional Materials", Nature Electronics **4**, 98–108 (2021).

[299] Y. Illarionov et al., "Ultrathin Calcium Fluoride Insulators for Two-Dimensional Field-Effect Transistors", Nature Electronics **2**, 230–235 (2019).

[300] Y. Illarionov, M. I. Vexler, V. V. Fedorov, S. M. Suturin, and N. S. Sokolov, "Light Emission from the Au/CaF2/p-Si(111) Capacitors: Evidence for an Elastic Electron Tunneling Through a Thin (1-2 nm) Fluoride Layer", Thin Solid Films **545**, 580–583 (2013).

[301] J. Weng and S.-P. Gao, "Layer-dependent band gaps and dielectric constants of ultrathin fluorite crystals", Journal of Physics and Chemistry of Solids **148**, 109738 (2021).

[302] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, "Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces", Physical Review Letters **100**, 136406 (2008).

[303] G. W. Rubloff, "Far-Ultraviolet Reflectance Spectra and the Electronic Structure of Ionic Crystals", Physical Review B **5**, 662–684 (1972).

[304] J. H. Skone, M. Govoni, and G. Galli, "Self-consistent hybrid functional for condensed systems", Physical Review B **89**, 195112 (2014).

[305] R. Resta and D. Vanderbilt, "Theory of Polarization: A Modern Approach", in *Physics of ferroelectrics: a modern perspective* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007), pp. 31–68.

[306] M. Sugiyama and M. Oshima, "MBE growth of fluorides", Microelectronics Journal **27**, LDSD: Epitaxial Deposition Techniques and Materials Systems, 361–382 (1996).

[307] C. Wong and D. Schuele, "The pressure derivatives of the elastic constants of CaF_2", Journal of Physics and Chemistry of Solids **28**, 1225–1231 (1967).

[308] R. P. Lowndes, "Dielectric response of the alkaline earth fluorides", Journal of Physics C **2**, 1595–1605 (1969).

[309] J. Hartnett, A. Fowler, M. Tobar, and J. Krupka, "The microwave characterization of single crystal lithium and calcium fluoride at cryogenic temperatures", IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **51**, 380–386 (2004).

[310] M. A. Olmstead, "Heteroepitaxy of Disparate Materials: From Chemisorption to Epitaxy in CaF_2/Si (111)", Thin Films: Heteroepitaxial Systems **15**, 211 (1999).

[311] J. H. Stathis, S. Mahapatra, and T. Grasser, "Controversial issues in negative bias temperature instability", Microelectronics Reliability **81**, 244–251 (2018).

[312] A. I. Chou, K. Lai, K. Kumar, P. Chowdhury, and J. C. Lee, “Modeling of stress-induced leakage current in ultrathin oxides with the trap-assisted tunneling mechanism”, Applied Physics Letters **70**, 3407–3409 (1997).

[313] K. Lee and H. Shin, “Investigation of Retention Characteristics for Trap-Assisted Tunneling Mechanism in Sub 20-nm NAND Flash Memory”, IEEE Transactions on Device and Materials Reliability **17**, 758–762 (2017).

[314] G. Rescher, G. Pobegen, T. Aichinger, and T. Grasser, “On the subthreshold drain current sweep hysteresis of 4H-SiC nMOSFETs”, in 2016 IEEE International Electron Devices Meeting (IEDM) (2016).

[315] T. Knobloch, B. Uzlu, Y. Y. Illarionov, Z. Wang, M. Otto, L. Filipovic, M. Waltl, D. Neumaier, M. C. Lemme, and T. Grasser, “Improving stability in two-dimensional transistors with amorphous gate oxides by Fermi-level tuning”, Nature Electronics **5**, 356–366 (2022).

[316] T. Aichinger, M. Nelhiebel, and T. Grasser, “On the Temperature Dependence of NBTI Recovery”, Microelectronics Reliability **48**, 1178–1184 (2008).

[317] T. Aichinger, M. Nelhiebel, and T. Grasser, “Unambiguous Identification of the NBTI Recovery Mechanism using Ultra-Fast Temperature Changes”, in 2009 International Reliability Physics Symposium (2009), pp. 2–7.

[318] M. Toledano-Luque, B. Kaczer, E. Simoen, P. J. Roussel, A. Veloso, T. Grasser, and G. Groeseneken, “Temperature and voltage dependences of the capture and emission times of individual traps in high-k dielectrics”, Microelectronic Engineering **88**, 1243–1246 (2011).

[319] K. S. Ralls, W. J. Skocpol, L. D. Jackel, R. E. Howard, L. A. Fetter, R. W. Epworth, and D. M. Tennant, “Discrete Resistance Switching in Submicrometer Silicon Inversion Layers: Individual Interface Traps and Low-Frequency (1/f?) Noise”, Physical Review Letters **52**, 228–231 (1984).

[320] A. Ayachi, W. B. Chouikha, S. Jaziri, and R. Bennaceur, “Telegraph noise effects on two charge qubits in double quantum dots”, Physical Review A **89**, 012330 (2014).

[321] R. E. Throckmorton and S. Das Sarma, “Crosstalk- and charge-noise-induced multiqubit decoherence in exchange-coupled quantum dot spin qubit arrays”, Physical Review B **105**, 245413 (2022).

[322] H.-S. Choi, H. Kim, J.-H. Lee, B.-G. Park, and Y. Kim, “AND Flash Array Based on Charge Trap Flash for Implementation of Convolutional Neural Networks”, IEEE Electron Device Letters **41**, 1653–1656 (2020).

[323] J. Chen, T. Tanamoto, H. Noguchi, and Y. Mitani, “Further investigations on traps stabilities in random telegraph signal noise and the application to a novel concept physical unclonable function (PUF) with robust reliabilities”, in 2015 Symposium on VLSI Technology (VLSI Technology) (2015), T40–T41.

[324] Institut für Mikroelektronik, TU Wien, "MINIMOS-NT User's Guide", (2008).

[325] Global TCAD Solutions, "MINIMOS-NT", (2020).

[326] Synopsys, "Sentaurus Device", (2022).

[327] G. Rzepa et al., "Efficient physical defect model applied to PBTI in high-k stacks", in 2017 IEEE International Reliability Physics Symposium (IRPS) (2017), XT-11.1-XT–11.6.

[328] B. J. O'Sullivan et al., "Defect profiling in FEFET Si:HfO_2 layers", Applied Physics Letters **117**, 203504 (2020).

[329] C. Schleich, J. Berens, G. Rzepa, G. Pobegen, G. Rescher, S. Tyaginov, T. Grasser, and M. Waltl, "Physical Modeling of Bias Temperature Instabilities in SiC MOSFETs", in 2019 IEEE International Electron Devices Meeting (IEDM) (2019), pp. 20.5.1–20.5.4.

[330] J. Franco, B. Kaczer, P. J. Roussel, J. Mitard, S. Sioncke, L. Witters, H. Mertens, T. Grasser, and G. Groeseneken, "Understanding the suppressed charge trapping in relaxed- and strained-Ge/SiO_2/HfO_2 pMOSFETs and implications for the screening of alternative high-mobility substrate/dielectric CMOS gate stacks", in 2013 IEEE International Electron Devices Meeting (2013).

[331] J. Franco et al., "BTI Reliability Improvement Strategies in Low Thermal Budget Gate Stacks for 3D Sequential Integration", in 2018 IEEE International Electron Devices Meeting (IEDM) (2018), pp. 34.2.1–34.2.4.

[332] "Comphy v3.0", TU Wien, Austria (2023).

[333] G. Pobegen and T. Grasser, "On the Distribution of NBTI Time Constants on a Long, Temperature-Accelerated Time Scale", IEEE Transactions on Electron Devices **60**, 2148–2155 (2013).

[334] V. V. Afanasev, M. Bassler, G. Pensl, and M. Schulz, "Intrinsic SiC/SiO_2 Interface States", physica status solidi (a) **162**, 321–337 (1997).

[335] R. Tsu and L. Esaki, "Tunneling in a Finite Superlattice", Applied Physics Letters **22**, 562–564 (1973).

[336] M. Herrmann and A. Schenk, "Field and high-temperature dependence of the long term charge loss in erasable programmable read only memories: Measurements and modeling", Journal of Applied Physics **77**, 4522–4540 (1995).

[337] M. Zhang, Z. Huo, Z. Yu, J. Liu, and M. Liu, "Unification of three multiphonon trap-assisted tunneling mechanisms", Journal of Applied Physics **110**, 114108 (2011).

[338] J. H. Scofield, N. Borland, and D. M. Fleetwood, "Temperature-independent switching rates for a random telegraph signal in a silicon metal–oxide–semiconductor field-effect transistor at low temperatures", Applied Physics Letters **76**, 3248–3250 (2000).

[339] T. Tewksbury, "Relaxation effects in MOS devices due to tunnel exchange with near-interface oxide traps", PhD Thesis, Massachusetts Institute of Technology (1992).

[340] O. Roux dit Buisson, G. Ghibaudo, and J. Brini, "Model for Drain Current RTS Amplitude in Small-Area MOS Transistors", Solid-State Electronics **35**, 1273–1276 (1992).

[341] B. Kaczer, S. M. Amoroso, R. Hussin, A. Asenov, J. Franco, P. Weckx, P. J. Roussel, G. Rzepa, T. Grasser, and N. Horiguchi, "On the distribution of the FET threshold voltage shifts due to individual charged gate oxide defects", in 2016 IEEE International Integrated Reliability Workshop (IIRW) (2016), pp. 18–20.

[342] K. El Sayed, E. Lyumkis, and A. Wettstein, "Modeling Statistical Variability with the Impedance Field Method", in 2012 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) (2012), pp. 205–208.

[343] G. Rzepa et al., "Reliability and Variability-Aware DTCO Flow: Demonstration of Projections to N3 FinFET and Nanosheet Technologies", in 2021 IEEE International Reliability Physics Symposium (IRPS) (2021), pp. 1–6.

[344] W. Van Roosbroeck, "Theory of the flow of electrons and holes in germanium and other semiconductors", The Bell System Technical Journal **29**, 560–607 (1950).

[345] G. Rzepa, "Efficient physical modeling of bias temperature instability", PhD Thesis, TU Wien (2018).

[346] J. Michl, "Charge trapping and variability in CMOS technologies at cryogenic temperatures", PhD Thesis, TU Wien (2022).

[347] S. M. Sze, *Physics of semiconductor devices* (Wiley-Interscience, Hoboken, N.J, 2007), 200ff.

[348] B. Stampfer, "Advanced electrical characterization of charge trapping in MOS transistors", PhD Thesis, TU Wien (2020).

[349] D. Fleetwood, H. Xiong, Z.-Y. Lu, C. Nicklaw, J. Felix, R. Schrimpf, and S. Pantelides, "Unified model of hole trapping, 1/f noise, and thermally stimulated current in MOS devices", IEEE Transactions on Nuclear Science **49**, 2674–2683 (2002).

[350] D. J. Breed, "A new model for the negative voltage instability in MOS devices", Applied Physics Letters **26**, 116–118 (1975).

[351] V. Huard, M. Denais, and C. Parthasarathy, "NBTI Degradation: From Physical Mechanisms to Modelling", Microelectronics Reliability **46**, 1–23 (2006).

[352] 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 **48**, 2004–2014 (1977).

[353] 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 **86**, 093506 (2005).

[354] S. Chakravarthi, A. Krishnan, V. Reddy, C. Machala, and S. Krishnan, “A comprehensive framework for predictive modeling of negative bias temperature instability”, in 2004 IEEE International Reliability Physics Symposium. Proceedings (2004), pp. 273–282.

[355] F. Schanovsky and T. Grasser, “On the Microscopic Limit of the RD Model”, in *Bias temperature instability for devices and circuits*, edited by T. Grasser (Springer New York, New York, NY, 2014), pp. 379–408.

[356] M. N. K. Alam, S. Clima, B. J. O’Sullivan, B. Kaczer, G. Pourtois, M. Heyns, and J. Van Houdt, “First principles investigation of charge transition levels in monoclinic, orthorhombic, tetragonal, and cubic crystallographic phases of HfO_2”, Journal of Applied Physics **129**, 084102 (2021).

[357] F. Schanovsky, “Atomistic Modeling in the Context of the Bias Temperature Instability”, PhD Thesis, TU Wien (2013).

[358] V. I. Arnold, *Ordinary differential equations* (Springer Berlin Heidelberg, 1992).

[359] T. Grasser, H. Reisinger, P.-J. Wagner, and B. Kaczer, “Time-dependent defect spectroscopy for characterization of border traps in metal-oxide-semiconductor transistors”, Physical Review B **82**, 245318 (2010).

[360] A. Lelis and T. Oldham, “Time dependence of switching oxide traps”, IEEE Transactions on Nuclear Science **41**, 1835–1843 (1994).

[361] B. Kaczer, T. Grasser, 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 2008 IEEE International Reliability Physics Symposium (2008), pp. 20–27.

[362] T. Grasser, H. Reisinger, P. .-. Wagner, F. Schanovsky, W. Goes, and B. Kaczer, “The time dependent defect spectroscopy (TDDS) for the characterization of the bias temperature instability”, in 2010 IEEE International Reliability Physics Symposium (2010), pp. 16–25.

[363] T. Grasser, K. Rott, H. Reisinger, M. Waltl, P. Wagner, F. Schanovsky, W. Goes, G. Pobegen, and B. Kaczer, “Hydrogen-Related Volatile Defects as the Possible Cause for the Recoverable Component of NBTI”, in 2013 International Electron Device Meeting (2013).

[364] E. Cartier, J. H. Stathis, and D. A. Buchanan, “Passivation and depassivation of silicon dangling bonds at the Si/SiO_2 interface by atomic hydrogen”, Applied Physics Letters **63**, 1510–1512 (1993).

[365] J. H. Stathis and E. Cartier, “Atomic hydrogen reactions with P_b centers at the (100) Si/SiO$_2$ interface”, Physical Review Letters **72**, 2745–2748 (1994).

[366] M. Houssa, V. Afanas’ev, A. Stesmans, M. Aoulaiche, G. Groeseneken, and M. Heyns, “Insights on the Physical Mechanism behind Negative Bias Temperature Instabilities”, Applied Physics Letters **90**, 043505 (2007).

[367] S. N. Rashkeev, D. M. Fleetwood, R. D. Schrimpf, and S. T. Pantelides, “Defect Generation by Hydrogen at the Si- SiO$_2$ Interface”, Physical Review Letters **87**, 165506 (2001).

[368] T. Grasser and B. Kaczer, “Evidence that Two Tightly Coupled Mechanism are Responsible for Negative Bias Temperature Instability in Oxynitride MOSFETs”, IEEE Transactions on Electron Devices **56**, 1056–1062 (2009).

[369] Y.-Y. Liu, F. Zheng, X. Jiang, J.-W. Luo, S.-S. Li, and L.-W. Wang, “Ab Initio Investigation of Charge Trapping Across the Crystalline-Si–Amorphous-SiO$_2$ Interface”, Physical Review Applied **11**, 044058 (2019).

[370] Y.-Y. Liu, F. Liu, R. Wang, J.-W. Luo, X. Jiang, R. Huang, S.-S. Li, and L.-W. Wang, “Characterizing the charge trapping across crystalline and amorphous Si/SiO$_2$/HfO$_2$ stacks from first-principle calculations”, Physical Review Applied **12**, 064012 (2019).

[371] C. S. Kelley, “Moments of semiclassical and classical absorption and emission band shapes of impurities in solids”, Physical Review B **20**, 5084–5089 (1979).

[372] C. Jungemann and C. Zimmermann, “DC, AC and Noise Simulation of Organic Semiconductor Devices Based on the Master Equation”, in 2014 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) (IEEE, 2014), pp. 137–140.

[373] C. Schleich, “Modeling of Defect Related Reliability Phenomena in SiC Power-MOSFETs”, PhD Thesis, TU Wien (2022).

[374] J. Blumberger and K. P. McKenna, “Constrained density functional theory applied to electron tunnelling between defects in MgO”, Physical Chemistry Chemical Physics **15**, 2184–2196 (2013).

[375] A. Grill, E. Bury, J. Michl, S. E. Tyaginov, D. Linten, T. Grasser, B. Parvais, B. Kaczer, M. Waltl, and I. Radu, “Reliability and Variability of Advanced CMOS Devices at Cryogenic Temperatures”, in Proceedings of the IEEE International Reliability Physics Symposium (IRPS) (2020), pp. 1–6.

[376] T. Markvart, “Determination of potential surfaces from multiphonon transition rates”, Journal of Physics C **14**, L435 (1981).

[377] K. F. Freed and J. Jortner, “Multiphonon Processes in the Nonradiative Decay of Large Molecules”, The Journal of Chemical Physics **52**, 6272–6291 (1970).

[378] 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 **7**, 531–539 (2007).

[379] M. Denais, C. Parthasarathy, G. Ribes, Y. Rey-Tauriac, N. Revil, A. Bravaix, V. Huard, and F. Perrier, “On-the-fly characterization of NBTI in ultra-thin gate oxide PMOSFETs”, in IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004. (2004), pp. 109–112.

[380] 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 **8**, 526–535 (2008).

[381] E. Simoen and C. Claeys, “Random Telegraph Signal: a local probe for single point defect studies in solid-state devices”, Materials Science and Engineering: B **91-92**, 136–143 (2002).

[382] Y. Yuzhelevski, M. Yuzhelevski, and G. Jung, “Random telegraph noise analysis in time domain”, Review of Scientific Instruments **71**, 1681–1688 (2000).

[383] J. Canny, “A computational approach to edge detection”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 679–698 (1986).

[384] N. Otsu et al., “A threshold selection method from gray-level histograms”, IEEE Transactions on Systems, Man, and Cybernetics **9**, 62–66 (1979).

[385] A. Grill, B. Stampfer, K.-S. Im, J.-H. Lee, C. Ostermaier, H. Ceric, M. Waltl, and T. Grasser, “Electrostatic coupling and identification of single-defects in GaN/AlGaN Fin-MIS-HEMTs”, Solid-State Electronics **156**, 41–47 (2019).

[386] H. Miki, N. Tega, M. Yamaoka, D. J. Frank, A. Bansal, M. Kobayashi, K. Cheng, C. D’Emic, Z. Ren, S. Wu, et al., “Statistical measurement of random telegraph noise and its impact in scaled-down high-κ/metal-gate MOSFETs”, in 2012 International Electron Devices Meeting (IEEE, 2012), pp. 19–1.

[387] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition”, Proceedings of the IEEE **77**, 257–286 (1989).

[388] T. Grasser, H. Reisinger, P.-J. Wagner, and B. Kaczer, “Time-dependent defect spectroscopy for characterization of border traps in metal-oxide-semiconductor transistors”, Physical Review B **82**, 245318 (2010).

[389] D. Varghese, D. Saha, S. Mahapatra, K. Ahmed, F. Nouri, and M. Alam, “On the dispersive versus Arrhenius temperature activation of NBTI time evolution in plasma nitrided gate oxides: Measurements, theory, and implications”, in IEEE International Electron Devices Meeting, 2005. IEDM Technical Digest. (IEEE, 2005), pp. 684–687.

[390] A. Tarantola, *Inverse problem theory and methods for model parameter estimation* (SIAM, 2005).

[391] T. R. Hanak, R. K. Ahrenkiel, D. J. Dunlavy, A. M. Bakry, and M. L. Timmons, "A new method to analyze multiexponential transients for deep-level transient spectroscopy", Journal of Applied Physics **67**, 4126–4132 (1990).

[392] A. A. Istratov and O. F. Vyvenko, "Exponential analysis in physical phenomena", Review of Scientific Instruments **70**, 1233–1257 (1999).

[393] J. Franco et al., "Low Temperature Atomic Hydrogen Treatment for Superior NBTI Reliability—Demonstration and Modeling across SiO_2 IL Thicknesses from 1.8 to 0.6 nm for I/O and Core Logic", in 2021 Symposium on VLSI Technology (2021), pp. 1–2.

[394] G. Rzepa, F. Schanovsky, and M. Karner, "From Gate Oxide Characterization to TCAD Predictions: Exploring Impact of Defects Across Technologies", in 2019 Silicon Nanoelectronics Workshop (SNW) (2019), pp. 1–2.

[395] D. Claes, J. Franco, N. Collaert, D. Linten, and M. Heyns, "Positive bias temperature instability of HfO_2-based gate stacks at reduced thermal budget for future CMOS technologies", Journal of Applied Physics **128**, 104101 (2020).

[396] A. N. Tikhonov, A. Goncharsky, V. Stepanov, and A. G. Yagola, *Numerical methods for the solution of ill-posed problems*, Vol. 328 (Springer Science & Business Media, 1995).

[397] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari, "Tikhonov regularization and the L-curve for large discrete ill-posed problems", Journal of Computational and Applied Mathematics **123**, 423–446 (2000).

[398] R. Bro and S. De Jong, "A fast non-negativity-constrained least squares algorithm", Journal of Chemometrics **11**, 393–401 (1997).

[399] M. Waltl, "Ultra-low noise defect probing instrument for defect spectroscopy of MOS transistors", IEEE Transactions on Device and Materials Reliability **20**, 242–250 (2020).

[400] 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 **399**, 758–761 (1999).

[401] Y. Yamashita et al., "Direct observation of site-specific valence electronic structure at the SiO_2/Si interface", Physical Review B **73**, 045336 (2006).

[402] Y. Maneglia and D. Bauza, "Extraction of slow oxide trap concentration profiles in metal-oxide-semiconductor transistors using the charge pumping method", Journal of Applied Physics **79**, 4187–4192 (1996).

[403] J. P. Campbell, P. M. Lenahan, A. T. Krishnan, and S. Krishnan, "Location, Structure, and Density of States of NBTI-Induced Defects in Plasma Nitrided pMOSFETs", in 2007 IEEE International Reliability Physics Symposium Proceedings (2007), pp. 503–510.

[404] M. W. Feil, K. Puschkarsky, W. Gustin, H. Reisinger, and T. Grasser, "On the Physical Meaning of Single-Value Activation Energies for BTI in Si and SiC MOSFET Devices", IEEE Transactions on Electron Devices **68**, 236–243 (2021).

[405] T. Grasser, M. Waltl, K. Puschkarsky, B. Stampfer, G. Rzepa, G. Pobegen, H. Reisinger, H. Arimura, and B. Kaczer, "Implications of gate-sided hydrogen release for post-stress degradation build-up after BTI stress", in 2017 IEEE International Reliability Physics Symposium (IRPS) (2017), 6A-2.1-6A–2.6.

[406] 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 **7**, 540–557 (2007).

[407] J. T. Ryan, P. M. Lenahan, T. Grasser, and H. Enichlmair, "Observations of negative bias temperature instability defect generation via on the fly electron spin resonance", Applied Physics Letters **96**, 223509 (2010).

[408] M. Waltl, B. Stampfer, G. Rzepa, B. Kaczer, and T. Grasser, "Separation of electron and hole trapping components of PBTI in SiON nMOS transistors", Microelectronics Reliability **114**, 113746 (2020).

[409] M. Sometani, D. Okamoto, S. Harada, H. Ishimori, S. Takasu, T. Hatakeyama, M. Takei, Y. Yonezawa, K. Fukuda, and H. Okumura, "Temperature-dependent analysis of conduction mechanism of leakage current in thermally grown oxide on 4H-SiC", Journal of Applied Physics **117**, 024505 (2015).

[410] P. Fiorenza, A. La Magna, M. Vivona, and F. Roccaforte, "Near interface traps in SiO_2/4H-SiC metal-oxide-semiconductor field effect transistors monitored by temperature dependent gate current transient measurements", Applied Physics Letters **109**, 012102 (2016).

[411] P. Moens, J. Franchi, J. Lettens, L. De Schepper, M. Domeij, and F. Allerstam, "A charge-to-breakdown (Q_{BD}) approach to SiC gate oxide lifetime extraction and modeling", in 2020 32nd International Symposium on Power Semiconductor Devices and ICs (ISPSD) (IEEE, 2020), pp. 78–81.

[412] Y. Huang et al., "In-Depth Understanding of Nitridation-Induced Endurance Enhancement in FeFETs: Defect Properties and Dynamics Characterized by Nonradiative Multi-Phonon Model", IEEE Transactions on Electron Devices **71**, 5388–5392 (2024).

List of Publications

Scientific Journals and Book Contributions

[DWJ1] **D. Waldhoer**, A.-M. El-Sayed, Y. Wimmer, M. Waltl, and T. Grasser, "Atomistic modeling of oxide defects", in *Noise in nanoscale semiconductor devices*, edited by T. Grasser (Springer, 2020), pp. 609–648.

[DWJ2] M. Jech, A.-M. El-Sayed, S. Tyaginov, **D. Waldhoer**, F. Bouakline, P. Saalfrank, D. Jabs, C. Jungemann, M. Waltl, and T. Grasser, "Quantum chemistry treatment of silicon-hydrogen bond rupture by nonequilibrium carriers in semiconductor devices", Physical Review Applied **16**, 014026 (2021).

[DWJ3] P. Khakbaz, **D. Waldhoer**, M. Bahrami, T. Knobloch, M. R. Davoudi, M. Pourfath, Y. Zhang, H. Peng, M. Waltl, and T. Grasser, "Two-dimensional Bi_2SeO_2 and Its Native Insulators for Next-Generation Nanoelectronics", ACS nano, (under review) (2024).

[DWJ4] C. Wilhelmer, **D. Waldhoer**, L. Cvitkovich, D. Milardovich, M. Waltl, and T. Grasser, "Polaron formation in the hydrogenated amorphous silicon nitride Si_3N_4:H", Physical Review B **110**, 045201 (2024).

[DWJ5] C. Wilhelmer, **D. Waldhoer**, L. Cvitkovich, D. Milardovich, M. Waltl, and T. Grasser, "Over- and Undercoordinated Atoms as a Source of Electron and Hole Traps in Amorphous Silicon Nitride (a-Si_3N_4)", Nanomaterials **13**, 2286 (2023).

[DWJ6] C. Wilhelmer, **D. Waldhoer**, M. Jech, A.-M. El-Sayed, L. Cvitkovich, M. Waltl, and T. Grasser, "Ab initio investigations in amorphous silicon dioxide: proposing a multi-state defect model for electron and hole capture", Microelectronics Reliability **139**, 114801 (2022).

[DWJ7] L. Cvitkovich, **D. Waldhoer**, A.-M. El-Sayed, M. Jech, C. Wilhelmer, and T. Grasser, "Dynamic modeling of Si(100) thermal oxidation: Oxidation mechanisms and realistic amorphous interface generation", Applied Surface Science **610**, 155378 (2023).

[DWJ8] **D. Waldhoer**, C. Schleich, J. Michl, B. Stampfer, K. Tselios, E. G. Ioannidis, H. Enichlmair, M. Waltl, and T. Grasser, "Toward automated defect extraction from bias temperature instability measurements", IEEE Transactions on Electron Devices **68**, 4057–4063 (2021).

[DWJ9] J. Michl, A. Grill, **D. Waldhoer**, W. Goes, B. Kaczer, D. Linten, B. Parvais, B. Govoreanu, I. Radu, M. Waltl, et al., "Efficient modeling of charge trapping at cryogenic temperatures—Part I: Theory", IEEE Transactions on Electron Devices **68**, 6365–6371 (2021).

[DWJ10] **D. Waldhoer**, C. Schleich, J. Michl, A. Grill, D. Claes, A. Karl, T. Knobloch, G. Rzepa, J. Franco, B. Kaczer, et al., "Comphy v3.0 — A compact-physics framework for modeling charge trapping related reliability phenomena in MOS devices", Microelectronics Reliability **146**, 115004 (2023).

[DWJ11] **D. Waldhoer**, M. E. Turiansky, C. G. Van de Walle, and T. Grasser, "Nonlocal Electron-Phonon Coupling in Nonradiative Transitions", (in preparation) (2024).

[DWJ12] **D. Waldhoer**, M. E. Turiansky, W. Lee, C. G. Van de Walle, and T. Grasser, "First-Principles Analysis of Hole-Induced Si-H Bond Dissociation in Silicon Dioxide", (in preparation) (2024).

[DWJ13] D. Milardovich, **D. Waldhoer**, M. Jech, A.-M. El-Sayed, and T. Grasser, "Building robust machine learning force fields by composite gaussian approximation potentials", Solid-State Electronics **200**, 108529 (2023).

[DWJ14] D. Milardovich, C. Wilhelmer, **D. Waldhoer**, L. Cvitkovich, G. Sivaraman, and T. Grasser, "Machine learning interatomic potential for silicon-nitride (Si_3N_4) by active learning", The Journal of Chemical Physics **158**, 194802 (2023).

[DWJ15] C. Schleich, **D. Waldhoer**, A.-M. El-Sayed, K. Tselios, B. Kaczer, T. Grasser, and M. Waltl, "Single-versus multi-step trap assisted tunneling currents—Part II: The role of polarons", IEEE Transactions on Electron Devices **69**, 4486–4493 (2022).

[DWJ16] J. Michl, A. Grill, **D. Waldhoer**, W. Goes, B. Kaczer, D. Linten, B. Parvais, B. Govoreanu, I. Radu, T. Grasser, et al., "Efficient modeling of charge trapping at cryogenic temperatures—Part II: Experimental", IEEE Transactions on Electron Devices **68**, 6372–6378 (2021).

[DWJ17] C. Schleich, **D. Waldhoer**, K. Waschneck, M. W. Feil, H. Reisinger, T. Grasser, and M. Waltl, "Physical modeling of charge trapping in 4H-SiC DMOSFET technologies", IEEE Transactions on Electron Devices **68**, 4016–4021 (2021).

[DWJ18] C. Schleich, **D. Waldhoer**, T. Knobloch, W. Zhou, B. Stampfer, J. Michl, M. Waltl, and T. Grasser, "Single-versus multi-step trap assisted tunneling currents—Part I: Theory", IEEE Transactions on Electron Devices **69**, 4479–4485 (2022).

[DWJ19] K. Tselios, **D. Waldhoer**, B. Stampfer, J. Michl, E. G. Ioannidis, H. Enichlmair, T. Grasser, and M. Waltl, "On the distribution of single defect threshold voltage shifts in SiON transistors", IEEE Transactions on Device and Materials Reliability **21**, 199–206 (2021).

[DWJ20] K. Tselios, T. Knobloch, **D. Waldhoer**, B. Stampfer, E. Ioannidis, H. Enichlmair, R. Minixhofer, T. Grasser, and M. Waltl, "Revealing the Impact of Gate Area Scaling on Charge Trapping Employing SiO_2 Transistors", IEEE Transactions on Device and Materials Reliability **23**, 355–362 (2023).

Conference Proceedings

[DWC1] T. Knobloch, J. Michl, **D. Waldhoer**, Y. Illarionov, B. Stampfer, A. Grill, R. Zhou, P. Wu, M. Waltl, J. Appenzeller, et al., "Analysis of single electron traps in nano-scaled MoS_2 FETs at cryogenic temperatures", in Proceedings of the Device Research Conference (DRC) (2020), pp. 52–53.

[DWC2] J. Michl, A. Grill, B. Stampfer, **D. Waldhoer**, C. Schleich, T. Knobloch, E. Ioannidis, H. Enichlmair, R. Minixhofer, B. Kaczer, et al., "Evidence of tunneling driven random telegraph noise in cryo-CMOS", in 2021 IEEE International Electron Devices Meeting (IEDM) (IEEE, 2021), pp. 31–3.

[DWC3] **D. Waldhoer**, Y. Wimmer, A.-M. El-Sayed, W. Goes, M. Waltl, and T. Grasser, "Minimum energy paths for non-adiabatic charge transitions in oxide defects", in International Integrated Reliability Workshop (IIRW) (IEEE, 2019), pp. 1–5.

[DWC4] D. Milardovich, M. Jech, **D. Waldhoer**, M. Waltl, and T. Grasser, "Machine learning prediction of defect formation energies in a-SiO_2", in 2020 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) (IEEE, 2020), pp. 339–342.

[DWC5] **D. Waldhoer**, B. Manna, A.-M. El-Sayed, T. Knobloch, Y. Illarionov, and T. Grasser, "Silicon-impurity defects in calcium fluoride: a first principles study", in 52nd European Solid-State Device Research Conference (ESSDERC) (IEEE, 2022), pp. 380–383.

Master Thesis

[DWT1] **Dominic Waldhoer**, "Potential Energy Surface Approximations for Nonradiative Multiphonon Charge Transitions in Oxide Defects", MA Thesis, TU Wien (2018).