Resistive Switching of Amorphous Non-Stoichiometric Gallium Oxide Thin Films

Von der Fakultät für Mathematik und Naturwissenschaften
der RWTH Aachen University zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Philipp Andreas Hein, M.Sc.

aus

Koblenz

Berichter: Univ.-Prof. Dr. rer. nat. Manfred Martin
Univ.-Prof. Dr. rer. nat. Arne Lüchow

Tag der mündlichen Prüfung: 06.10.2025

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.

Abstract

Resistive Switching of Amorphous Non-Stoichiometric Gallium Oxide Thin Films

Amorphous, non-stoichiometric gallium oxide (a-GaO_x, $x < 1.5$) is a promising material for resistive switching memory and logic, especially for neuromorphic computing. However, despite considerable research efforts in recent years, its resistive switching mechanism is still far from understood. A wide range of different switching patterns has been reported, without consensus in the interpretations. This work aims at a deeper understanding of the resistive switching processes in a-GaO_x, by focusing on the electron transport and the ways it can be influenced by ion migration.

The resistive switching measurements not only confirm counter-8, figure-8 and abnormal bipolar resistive switching, but also unveil new links between these switching patterns. The surface oxidation of a-GaO_x, which is quantified for the first time, and the voltage amplitude seem to play a decisive role. The highlight of this work is the discovery of variable-range hopping (VRH) as the dominating electron transport mechanism in a-GaO_x at room temperature. This has never been reported before, as all previous works explicitly or implicitly assumed band conduction. The measured conductivity closely follows the exponential $T^{-1/4}$-dependency of Mott's law for VRH. The presence of localized electronic states near the Fermi level is confirmed by photoemission spectroscopy and density functional theory. The electron hopping between the localized states is simulated with the kinetic Monte Carlo (KMC) method, in what are likely the first VRH simulations with an ab-initio density of states of an amorphous material. It is demonstrated that the KMC-VRH model can reproduce the conductivity measurements quantitatively. Experiment and simulation also agree that the conductivity decreases non-linearly with increasing O/Ga ratio. At high electric field strength, the simulations yield an elevated electron temperature and an exponential increase of the conductivity, in accordance with VRH theory and with notable resemblance to the resistive switching curves. Finally, the consequences of VRH on the device scale, particularly for metal|a-GaO_x Schottky barriers, are investigated by finite element modeling. The device simulations reveal low-conductive regions near both interfaces, which typically dominate the total resistance. The electric field dependence of the VRH conductivity in these regions leads to symmetric, S-shaped current-voltage curves that are comparable to the experiments. Resistive switching thus occurs most likely via changes of the O/Ga ratio in these regions. Based on this conclusion, a new unified switching mechanism is devised that would explain the various experimental observations.

Resistives Schalten von Amorphen Nicht-Stöchiometrischen Galliumoxid-Dünnschichten

Amorphes, nicht-stöchiometrisches Galliumoxid (a-GaO_x, $x < 1.5$) ist ein vielversprechendes Material für resistiv schaltende Speicher- und Logikelemente, insbesondere im Bereich neuromorpher Computerschaltkreise. Trotz erheblicher Forschungsbemühungen in den letzten Jahren ist der Mechanismus des resistiven Schaltens von a-GaO_x bisher nicht geklärt. In der zugehörigen Literatur ist ein breites Spektrum unterschiedlichen Schaltverhaltens dokumentiert, zusammen mit vielen verschiedenen Interpretationen. Das Ziel dieser Arbeit ist ein tieferes Verständnis der Schaltprozesse in a-GaO_x, indem primär die Elektronenleitung und ihre potentielle Beeinflussung durch Ionenmigration untersucht wird.

Die durchgeführten Messungen zum resistiven Schalten zeigen nicht nur alle drei bekannten bipolaren Schaltmuster ("counter-8", "figure-8" und "abnormal"), sondern auch deren Beziehung untereinander. Die erstmals quantifizierte Oberflächenoxidation von a-GaO_x und die maximal angelegte Spannung scheinen eine entscheidende Rolle zu spielen. Das Highlight dieser Arbeit ist die Entdeckung, dass "Variable-Range Hopping" (VRH) der vorherrschende Elektronenleitungsmechanismus in a-GaO_x bei Raumtemperatur ist. Zuvor wurde ubiquitär angenommen, dass wie üblich Bandleitung dominieren würde. Die gemessenen Leitfähigkeiten folgen genau der exponentiellen $T^{-1/4}$-Abhängigkeit, die nach dem Mott'schen Gesetz für VRH erwartet wird. Die lokalisierten Elektronenzustände nahe des Fermi-Niveaus, zwischen denen die Übergänge bei VRH stattfinden, werden durch Photoemissionsspektroskopie und Dichtefunktionaltheorie-Rechnungen nachgewiesen. Letztere bilden die Basis für Kinetik-Monte-Carlo-Simulationen, welche die Leitfähigkeitsmessungen quantitativ reproduzieren. Dies sind sehr wahrscheinlich die ersten VRH-Simulationen mit einer ab-initio Zustandsdichte eines amorphen Materials. Sie ergeben in Übereinstimmung mit den Experimenten eine nicht-lineare Verringerung der Leitfähigkeit bei steigendem O/Ga-Verhältnis x. Hohe elektrische Feldstärken führen zu einer stark erhöhten Elektronentemperatur und zu einem exponentiellen Anstieg der Leitfähigkeit, welcher zur Krümmung der gemessenen Schaltkurven passt. Die Konsequenzen des VRH-Transports auf der Ebene der Schaltelemente, insbesondere für die Schottky-Barrieren an den Grenzflächen, werden schließlich durch Finite-Elemente-Modellierung untersucht. Es zeigt sich, dass der Gesamtwiderstand typischerweise von schmalen Bereichen nahe der Metall|a-GaO_x-Grenzflächen dominiert wird. Die Veränderung der VRH-Leitfähigkeit in diesen Bereichen in Abhängigkeit von der lokalen elektrischen Feldstärke führt zu symmetrischen, S-förmigen Strom-Spannungs-Kurven, die denen im Experiment ähneln. Das resistive Schalten erfolgt daher wahrscheinlich durch Veränderungen des O/Ga-Verhältnisses in diesen Regionen. Ausgehend von dieser Schlussfolgerung wird ein neuer Mechanismus für das resistive Schalten von a-GaO_x vorgeschlagen, der die vielfältigen experimentellen Beobachtungen erklären würde.

Acknowledgement

First and foremost, I want to express my deep gratitude to my supervisor, Prof. Manfred Martin, for his extraordinary support, the inspiring discussions and his guidance in the world of science. I am especially grateful for the personal development he promoted, for example, through the participation in large projects and international conferences, through a research stay in South Korea and by entrusting me with lectures, exams and presentations on his behalf. I owe him a lot, more than I could possibly list here, so I just want to emphasize: Thank you very much, Prof. Martin!

A special thanks goes to Prof. Arne Lüchow, for his engagement as the second reviewer of this thesis.

Naturally, this thesis was supported by the work of many others. First of all, I want to thank "my" students, Steffen Bochenek, Sabine Schneider, Lars Pritzlaff, Tobias Romstadt, Esther Jaekel and Dieter Hoffzimmer, who carried out *I*-*V*, *C*-*V*, van-der-Pauw, UV-Vis and SIMS measurements under my supervision. Each of them has made great efforts in the face of the experimental challenges with gallium oxide. In this context, I also like to thank Alexandra von der Heiden, Andreas Falkenstein and Christian Schwab for helping us with the van-der-Pauw and SIMS measurements.

Next, a big thank you goes to the external cooperation partners, who provided measurements that would not have been possible otherwise: Silvia Richter from the Central Facility for Electron Microscopy (GFE, RWTH Aachen University), for the EPMA measurements; Bo Thöner from the Institute of Physics IA (RWTH Aachen University), for the PPMS measurements; Slavomir Nemsak from the Peter Grünberg Institute (PGI-6, FZ Jülich), for the HAXPES measurements; Kootak Hong and his colleagues from the Korea Basic Science Institute (KBSI, Daejeon, Republic of Korea), for the UPS measurements; Jinseok Ryu and Sanghyo Lee from the Electron Microscopy & Spectroscopy Lab (EMSL) of Seoul National University (SNU, Republic of Korea), for the STEM measurements. I also want to thank the respective group leaders, Prof. Joachim Mayer, Prof. Matthias Wuttig, Prof. Claus Schneider, Prof. Ho Won Jang and Prof. Miyoung Kim, for enabling these measurements. I especially thank Prof. Ho Won Jang and Kootak Hong from the Department of Materials Science and Engineering of Seoul National University (SNU, Republic of Korea), for welcoming me to South Korea, for the opportunity to use their SKPM instrument and for the great time together in Seoul.

My final thanks in terms of explicit contributions belongs to Fabian Draber, whose elaborate AIMD structures of amorphous gallium oxide were the basis for essentially all simulations in this work. As luck would have it, he calculated them as a side project, long before their value for my work became apparent. I am very grateful for his generous permission to use them in my simulations.

Furthermore, I would like to thank for the financial support from the Deutsche Forschungsgemeinschaft (DFG) within the Sonderforschungsbereich "Nanoswitches" (SFB 917), from the Deutsche Akademische Austauschdienst (DAAD) within the German-Korean Partnership Programme GEnKO and from RWTH Aachen University within the RWTH Scholarship for Doctoral Students. The provision of experimental facilities by DESY, a member of the Helmholtz Association HGF, and of computing resources on the supercomputer CLAIX by RWTH Aachen University (project "rwth0945") is gratefully acknowledged as well.

I am thankful to all members of the workgroup, especially Christian Schwab, Alexandra von der Heiden, Andreas Falkenstein and Benjamin Grope, as well as my office roommates Julius Koettgen and Fabian Draber, for the productive and positive working atmosphere, the great team spirit and the many joyful experiences. I also want to thank Alexandra von der Heiden specifically, for the good teamwork in "our" SFB sub-project about gallium oxide, and Benjamin Grope, for bringing me into this workgroup and for the valuable introduction to the Monte Carlo method. Special thanks also goes to Prof. Roger De Souza and Prof. Steffen Neitzel-Grieshammer at the Institute of Physical Chemistry, as well as Prof. Peter Schmidt from TU Darmstadt, for many instructive and fruitful discussions.

Finally, I want to wholeheartedly thank all my family, including my dearest and my four-legged friends, for the immeasurable amounts of support, advice and encouragement during this thesis!

Declarations

Eidesstattliche Erklärung

Ich, Philipp Andreas Hein, erkläre hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung in meiner Zeit am Lehrstuhl I des Instituts für Physikalische Chemie an der RWTH Aachen University, unter der Anleitung von Professor Dr. Manfred Martin generiert wurden.

Wann immer andere eigene oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt. Wenn aus anderen eigenen oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate.

Alle wesentlichen Quellen von Unterstützung wurden benannt und wann immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde dies von mir klar gekennzeichnet. Teile dieser Arbeit wurden wie auf der nächsten Seite aufgeführt zuvor veröffentlicht.

Aachen, den 04.06.2025

Parts of this work have been published in the following article:

Philipp Hein, Tobias Romstadt, Fabian Draber, Jinseok Ryu, Thorben Böger, Andreas Falkenstein, Miyoung Kim and Manfred Martin, "Variable-Range Hopping Conduction in Amorphous, Non-Stoichiometric Gallium Oxide", *Advanced Electronic Materials*, **2025**, 11, 2400407.

and in the developed software:

Philipp Hein, "VHoppAS: Variable-range Hopping in Amorphous Solids", *Zenodo*, **2024**, DOI: 10.5281/zenodo.10974845.

and in oral presentations and posters:

Alexandra von der Heiden, Philipp Hein and Manfred Martin, "Oxygen diffusion and resistive switching in amorphous gallium oxide", poster at the 114th General Assembly of the German Bunsen Society for Physical Chemistry, Bochum, Germany, May 15, **2015**.

Alexandra von der Heiden, Philipp Hein and Manfred Martin, "Influence of structure and stoichiometry on oxygen diffusion and resistive switching in gallium oxide", poster at the E-MRS Spring Meeting, Lille, France, May 5, **2016**.

Philipp Hein and Manfred Martin, "Stoichiometry-dependence of resistive switching in gallium oxide", presentation at the 21st International Conference on Solid State Ionics (SSI-17), Padua, Italy, June 20, **2017**.

Philipp Hein and Manfred Martin, "Stoichiometry-dependence of resistive switching in amorphous gallium oxide", poster at the 17th Non-Volatile Memory Technology Symposium (NVMTS), Aachen, Germany, August 30, **2017**.

Ionic conductivity simulations, which have been developed as a parallel project but which are not discussed in this work, have been published in the following article:

Philipp Hein, Benjamin O.H. Grope, Julius Koettgen, Steffen Grieshammer and Manfred Martin, "Kinetic Monte Carlo simulations of ionic conductivity in oxygen ion conductors", *Materials Chemistry and Physics*, **2021**, 257, 123767.

The following students finished their work under my supervision and their work is (partially) used in this thesis:

Steffen Bochenek, "Synthese und resistives Schalten von amorphen Galliumoxid- und Hafniumoxid-Dünnschichten", Research project, RWTH Aachen University, Aachen, **2015**.

Fabian Michael Draber, "Sauerstoffionen-Leitfähigkeit in dotiertem $SrTiO_3$ und $CaTiO_3$ mittels ab initio und Kinetic Monte Carlo Berechnungen", Master thesis, RWTH Aachen University, Aachen, **2015**.

Sabine Schneider, "I-V characteristics of different amorphous non-stoichiometric oxides", Research project, RWTH Aachen University, Aachen, **2016**.

Lars Pritzlaff, "Impedance spectroscopy on amorphous gallium oxide resistive switching devices", Bachelor thesis, RWTH Aachen University, Aachen, **2017**.

Sebastian Eisele, "A Kinetic Monte Carlo code for multiple mobile species in arbitrary solid solutions", Master thesis, RWTH Aachen University, Aachen, **2017**.

Tobias Romstadt, "Stochiometry-dependence of optical and electrical properties of amorphous gallium oxide thin films", Master thesis, RWTH Aachen University, Aachen, **2018**.

Esther Elisabeth Jaekel, "Effect of the bottom electrode material on resistive switching devices based on non-stochiometric GaO_x thin films", Research project, RWTH Aachen University, Aachen, **2019**.

Dieter Hoffzimmer, "Resistives Schalten von lithografisch strukturiertem a-$GaO_{1.0}$", Research project, RWTH Aachen University, Aachen, **2020**.

Contents

1 Introduction **1**

2 Fundamentals **5**

2.1 Irreversible Thermodynamics of Electron Conduction 5

2.1.1 Effective Charge Carrier Density 10

2.2 Electron Transport in Amorphous Solids 16

2.2.1 Extended and Localized States 16

2.2.2 Mobility Edge Conduction 19

2.2.3 Variable-Range Hopping Conduction 21

2.2.4 Schottky Barriers 28

2.3 Gallium Oxide 32

2.3.1 Density of States 35

2.3.2 Electronic Conductivity 38

2.3.3 Metal|a-GaO_x Interfaces 41

2.3.4 Gallium and Oxygen Ion Transport 43

2.4 Resistive Switching Mechanisms 46

2.4.1 Filamentary Switching 47

2.4.2 Interface Switching 51

2.4.3 Bulk Switching 55

3 Experiments **59**

3.1 Thin Film Deposition 59

3.2 Thin Film Characterization 61

3.2.1 Electron Microscopy (EPMA, SAED and EELS) 62

3.2.2 Photoelectron Spectroscopy (HAXPES and UPS) 64

3.2.3 Scanning Kelvin Probe Microscopy (SKPM) 66

3.3 Current-Voltage Measurements 68

3.4 Conductivity Measurements (van-der-Pauw) 73

4 Simulations **79**

4.1 Density Functional Theory (DFT) 79

4.1.1 Amorphous Model Structures 80

4.1.2 Density of States (DOS) 81

4.1.3 Localization of States (LOS) 82

4.2 Kinetic Monte Carlo (KMC) 84

4.2.1 Concept of the KMC Simulations 85

4.2.2 Physical and Technical Parameters . . . 86
4.2.3 Setup of the KMC Cell . . . 89
4.2.4 KMC Algorithm . . . 92
4.2.5 Calculation of the Results . . . 93
4.2.6 Implementation Details . . . 96
4.3 Finite Element Modeling (FEM) . . . 98
4.3.1 Zero Voltage Model . . . 99
4.3.2 Current-Voltage Model . . . 101
4.3.3 FEM Settings . . . 102

5 Results and Discussion 105
5.1 Material Characterization . . . 105
5.1.1 Amorphicity . . . 106
5.1.2 Non-Stoichiometry . . . 107
5.1.3 Surface Oxidation . . . 110
5.1.4 Work Function . . . 114
5.2 Resistive Switching . . . 116
5.2.1 Counter-8 Bipolar Switching . . . 117
5.2.2 Figure-8 Bipolar Switching . . . 124
5.2.3 Abnormal Bipolar Switching . . . 130
5.2.4 Interpretation of the Resistive Switching Behavior . . . 134
5.3 Experimental Bulk Conductivity . . . 140
5.3.1 Thermal Relaxation . . . 140
5.3.2 Variable-Range Hopping . . . 142
5.3.3 Influence of the Stoichiometry . . . 146
5.3.4 Magnetoresistance and Hall Effect . . . 149
5.4 Localized Electronic States . . . 150
5.4.1 Defect States from Oxygen-Deficiency . . . 151
5.4.2 Localized States from Amorphicity . . . 154
5.5 Simulated Variable-Range Hopping . . . 157
5.5.1 Validation of the KMC Simulations . . . 157
5.5.2 Quantitative Agreement with Experiments . . . 167
5.5.3 Influence of the Stoichiometry . . . 171
5.5.4 Hopping at High Electric Field Strength . . . 174
5.5.5 Variation of the Chemical Potential . . . 178
5.5.6 Electron Diffusion . . . 181
5.6 Modeling of Pt|a-$GaO_{1.0}$|ITO Devices . . . 185
5.6.1 Electron Equilibrium at Zero Voltage . . . 186
5.6.2 Steady State for Non-Zero Voltages . . . 188

6 Conclusion **197**

A Calculations **203**

A.1 Random Distances within a Sphere 203

A.2 Effective Charge Carrier Density in Constant or Linear DOS 204

A.3 Weighted Mean, Inner Error and Outer Error 204

A.4 Tracer Statistics and Random Walk Theory for Electrons 206

A.5 Variational Formulation of the FEM Models 213

B Figures **217**

C Tables **235**

References **237**

Abbreviations **249**

Symbols **253**

1

Introduction

The digital evolution of humankind in the information age is tightly coupled to the improvement of data storage and processing. The current advent of artificial intelligence (AI), based on enormous amounts of training data, is just one example. Symbolic for the rapid development of computation hardware are Kryder's law, describing the exponential growth of storage capacity by about 40% each year,[1] and Moore's law, which states that the transistor density in integrated circuits approximately doubles every two years.[2] In recent years however, the technological progress struggled to meet these predictions, as the down-scaling of field-effect transistors is approaching its physical limits.[3, 4] Another challenge is the increasing disparity between computation and memory speed. Data transfer bandwidth ("von Neumann bottleneck") and read/write access times ("memory wall") are becoming a limiting factor, due to the separation of data storage and processing that is inherent to the von Neumann architecture.[5] Several transistor alternatives are under development in order to overcome these limitations, and resistive random-access memory (RRAM) is one of the promising candidates.

RRAM is based on resistive switching (RS), i.e., the non-volatile, reversible modification of a material's resistance through an applied voltage. The individual memory cells are two-terminal, capacitor-like devices with an RS-capable material sandwiched between two electrode layers. Ion migration, driven by the electric field between the electrodes, causes local compositional changes in the material and thereby alters the overall device resistance ("write" operation). Depending on where these internal processes occur, a distinction is typically made between

- Filamentary switching: A more conductive phase in form of a filament grows in the RS material, decreasing the resistance when it short-circuits the electrodes and increasing the resistance when it ruptures.
- Interface switching: A Schottky barrier at one of the interfaces between RS material and electrodes dominates the device resistance, which varies when the local composition at the interface changes.
- Bulk switching: The device resistance depends on the bulk conductivity in the entire cross-section of the RS material, no just in a filamentary region, and varies when the local stoichiometry changes throughout the material.

The resulting resistance variations can be abrupt, such that the device switches between two distinct resistance states (digital RS), or gradual, allowing the device to switch to any resistance in a certain range (analog RS). Without an applied voltage, the reached state of the device is conserved. At sufficiently low voltages, the resistance state can be probed without altering it ("read" operation). Digital RS devices that switch between a high resistive state (HRS) and a low resistive state (LRS) are ideal for conventional computer logic and memory. Their rather simple, two-terminal structure allows for a higher device density than the common three-terminal transistors, e.g., in the form of stacked crossbar arrays.[6] High switching speeds in the range of nanoseconds and low power consumption in the range of picojoules per read/write operation make them suitable both for data storage and for data processing, even in combination to avoid the von Neumann bottleneck by in-memory computation.[5, 6] Analog RS devices that switch gradually can serve as artificial synapses in neuromorphic computing, as their ability to adjust on a continuous scale, based on an incoming signal, mimics the synaptic plasticity of biological neurons.[7] Together with the other aforementioned advantages, this in-memory learning ability makes analog RRAM suitable for hardware-based AI that is more efficient than the current AI software running on digital systems. The high potential of RRAM and the successive accumulation of promising results have led to tremendous research efforts in the field of resistive switching in recent years.[5–8] In this time, digital RRAM has reached a quite mature stage, with first commercial prototypes produced by the semiconductor industry.[6] Still, several issues must be solved in order to compete for example with state-of-the-art flash memory. Apart from architectural challenges in scaling and stacking of the crossbar arrays, especially the cell-to-cell and cycle-to-cycle variability of the memory cells is problematic.[6, 8] This is even more so for analog RRAM that relies on the reproducibility of a whole range of resistance states. The non-linearity and asymmetry of the switching response, too short retention times and complex control circuits are further exemplary reasons, why RRAM-based AI is still in its infancy.[7] Even at this stage, remarkable pattern-recognition results have already been obtained with rather small analog RRAM chips.[7] Given that in most cases the underlying atomistic processes during resistive switching are still insufficiently understood, it is the investigation of RS mechanisms – like in this work – that paves the route for RRAM optimization.

The inspiration for this work has been the observation of resistive switching in amorphous, non-stoichiometric gallium oxide thin films (a-GaO_x, $x < 1.5$).[9, 10] This material raised special attention because of the severe oxygen-deficiency of about one third ($x \approx 1.1$), because the proposed bulk switching mechanism differed from the established filament- or interface-based RS,[10] and because crystalline β-Ga_2O_3 was already a prominent research topic, e.g., for power electronic devices and solar-blind photodetectors, due to its ultra-wide band gap of about 4.9 eV.[11–16] The use of β-Ga_2O_3 in high-temperature oxygen and hydrogen sensors underlined the sensitive relation between stoichiometry and conductivity in gallium oxide.[11] In the following years, promising results for a-GaO_x-based memory cells [17–20] and neuromorphic computing devices [21–23] have been reported, but the underlying switching mechanism(s) remained controversial. A large variety of switching patterns has been observed, e.g., both digital RS [18, 24, 25] and analog RS [9, 10, 26], and

all major RS mechanisms have been attributed to a-GaO_x by now: filamentary switching [18, 24, 25], interface switching [9, 26–28] and bulk switching.[10, 29, 30] It stands out that in these explanatory approaches a-GaO_x is mostly treated like a crystalline n-type semiconductor doped with oxygen vacancies, even though amorphous solids typically have unique properties that differ from their crystalline counterparts. For instance, point defects are not well-defined in an amorphous solid, which makes the picture of self-doping by oxygen vacancies merely an illustrative description for local non-stoichiometry. It is also well-known that amorphicity leads to localized electronic states in the band gap that may affect the electron transport.[31] The common analogy between a-GaO_x and the well-understood n-type semiconductors originates from the assumption that electrons in a-GaO_x move by mobility edge conduction (MEC), i.e., band transport via extended electronic states above the so-called conduction band mobility edge, which separates localized from extended states. The localized electronic states below this mobility edge have been essentially ignored in the models. So far, no one considered variable-range hopping (VRH), i.e., electron conduction via phonon-assisted tunneling between localized states, in amorphous gallium oxide. Clarification of the electron transport mechanism is however mandatory in order to understand how local changes in stoichiometry may affect it and lead to resistive switching.

It is the aim of this work to investigate the RS mechanism of a-GaO_x. The sample setup thereby follows essentially the original work that proposed bulk switching: a-GaO_x prepared by pulsed laser deposition (PLD) in Ar atmosphere at room temperature, sandwiched between an indium tin oxide (ITO) bottom electrode (BE) and a platinum top electrode (TE).[10] In order to gather mechanistic clues, the switching behavior of these Pt|a-GaO_x|ITO devices – and modified variants thereof – is characterized by I-V measurements. They lead to a central question of this work: How is the electric field distributed inside the devices? Since the local electric field is the driving force for ion migration, it is its spatial and temporal development that controls the RS process. The question is approached through the investigation of the electron conduction in a-GaO_x. In a given resistance state, the electrons are responsible for the majority of the measured current (otherwise the RS could not be non-volatile), such that according to the current continuity equation the local variations of the electron conductivity throughout the device essentially determine the electric field. Consequently, the bulk conductivity of a-GaO_x, in relation to temperature and stoichiometry, is measured with the van-der-Pauw technique. In combination with further assessment of the localized electronic states, mainly through density functional theory (DFT) and hard X-ray photoelectron spectroscopy (HAXPES), it becomes evident that the dominant conduction mechanism is most probably variable-range hopping. This is confirmed further by kinetic Monte-Carlo (KMC) simulations of the VRH mechanism. The resulting KMC data is then used as a basis for simulating the internal electric field and the I-V curve of as-prepared Pt|a-$GaO_{1.0}$|ITO devices with the finite element method (FEM). Alongside this electron-focused approach, also the bulk and surface oxygen-deficiency of the a-GaO_x samples is investigated through electron probe microanalysis (EPMA), electron energy loss spectroscopy (EELS) and HAXPES.

2

Fundamentals

2.1 Irreversible Thermodynamics of Electron Conduction

Electron transport can be described in several different ways. Probably the most coarse depiction of conductive behavior is Ohm's law

$$I = \frac{V}{R} \tag{2.1}$$

that relates the current I to the voltage V applied to a two-terminal device. The resistance R – namesake for the resistive switching effect – characterizes the device like a black box where all intricacies of charge transport are hidden.[1] Ohm's law is even indifferent to the nature of the charge carriers, but it is known that in a-GaO_x the electrons are the majority charge carriers, while migration of oxygen or gallium ions may influence the electron transport but is negligible for a measured current.[10, 32, 33] The high level of abstraction provided by Ohm's law and likewise relations is for example useful for the analysis of extended electronic circuits. The other extreme is quantum mechanics. Therein, individual electrons are described as wave packets that occupy distinct quantum states with discrete energy levels. The temporal evolution of the wave functions and the transition probabilities between the electronic states govern the motion of the electrons on the atomic scale. Section 2.1.1 will introduce this quantum theory of electrons in more detail, also as a basis for the conduction mechanisms discussed later in sections 2.2.2 and 2.2.3.

For the treatment of macroscopic electron fluxes in a material, however, the framework of non-equilibrium thermodynamics is ideal, as it is situated in between the two aforementioned variants.[34] It describes the electrons as an ensemble with average properties that govern the behavior. In order to focus on the electrons, primarily the case of a homogeneous material with spatially and timely constant atomic composition shall be discussed.[2] Consider a point-like region within the material as a thermodynamic system. The local temperature T is a measure for the average kinetic

[1] The treatment here shall be limited to slowly varying voltages for the sake of brevity. Otherwise, the time-dependent I-V response would need to be described by the impedance.

[2] Homogeneity with regard to the ions is not a requirement for the theory presented here, but it is helpful for a concise notation. It is imposed already at this point, since in this work the equations will only be used under such conditions and since it will be a prerequisite for the later definition of the effective charge carrier density anyway.

energy of the electrons in such system. When T differs nearby, i.e., when there is a temperature gradient ∇T, the electrons tend to diffuse either towards higher or lower T, depending on the material,[3] which is called thermoelectric effect. Another average property is the local chemical potential μ of the electrons, which describes how the Gibbs free energy G of such system would change with the number of electrons N when there would be no electric field:[34]

$$\mathrm{d}G_{\mathrm{int}} = V\,\mathrm{d}p - S\,\mathrm{d}T + \mu\,\mathrm{d}N \quad \Rightarrow \quad \mu = \left(\frac{\partial G_{\mathrm{int}}}{\partial N}\right)_{T,p} \tag{2.2}$$

Here, G_{int} is the Gibbs free energy without contributions from a macroscopic electric field, V is the volume, p is the pressure, S is the entropy and T is the temperature. It is thermodynamically favorable when electrons redistribute from regions with high μ to those with low μ, since this reduces G_{int} of the material as a whole. The result is a diffusion flux in the direction of $-\nabla\mu$. In an electric field $\vec{F} = -\nabla\varphi$, where φ is the electric potential, the electrons with charge $q = -e$ also possess an electric potential energy $q\varphi$ relative to a reference (e.g., electrical ground). The so-called drift flux of electrons in the direction of $\nabla\varphi$ is beneficial as it reduces $q\varphi$. The sum of chemical potential and electric potential energy is the electrochemical potential $\tilde{\mu} = \mu - e\varphi$ of the electrons. It describes how G of the system changes with N when the electric field is taken into account:

$$\mathrm{d}G = \mathrm{d}G_{\mathrm{int}} - e\varphi\,\mathrm{d}N = V\,\mathrm{d}p - S\,\mathrm{d}T + \tilde{\mu}\,\mathrm{d}N \quad \Rightarrow \quad \tilde{\mu} = \left(\frac{\partial G}{\partial N}\right)_{T,p} \tag{2.3}$$

In combination, it is the redistribution of electrons in the direction of $-\nabla\tilde{\mu}$ that is thermodynamically favorable. From the 1st and 2nd law of thermodynamics essentially follows for the total macroscopic electron flux $\vec{J}_{\mathrm{e}}$ [34, 37, 38]

$$\begin{aligned}\vec{J}_{\mathrm{e}} &= -L_{\mathrm{ee}}\left[T\nabla\left(\frac{\mu}{T}\right) - e\nabla\varphi\right] - L_{\mathrm{eQ}}\frac{1}{T}\nabla T = -L_{\mathrm{ee}}\left[\nabla\mu - \frac{\mu}{T}\nabla T - e\nabla\varphi\right] - L_{\mathrm{eQ}}\frac{1}{T}\nabla T \\ &= -L_{\mathrm{ee}}\nabla\tilde{\mu} + L_{\mathrm{ee}}\frac{\mu}{T}\nabla T - L_{\mathrm{eQ}}\frac{1}{T}\nabla T = -L_{\mathrm{ee}}\nabla\tilde{\mu} - L_{\mathrm{ee}}\left[\frac{L_{\mathrm{eQ}}}{L_{\mathrm{ee}}} - \mu\right]\frac{\nabla T}{T}\end{aligned} \tag{2.4}$$

where L_{ee} and L_{eQ} are the Onsager coefficients.[4] L_{ee} essentially corresponds to the conductivity σ of the electrons according to [37]

$$L_{\mathrm{ee}} = \frac{\sigma}{e^2} \tag{2.5}$$

while the thermoelectric term can be expressed in simpler form with the electrons' Seebeck coefficient [35, 39]

$$S_{\mathrm{e}} = -\frac{1}{eT}\left[\frac{L_{\mathrm{eQ}}}{L_{\mathrm{ee}}} - \mu\right] \tag{2.6}$$

[3] According to the Sommerfeld-Bethe relation, the direction of the thermoelectric effect depends on which of the electronic states around the Fermi level contribute most to the conductivity.[35, 36]

[4] Further terms with Onsager cross-coefficients that link the electron flux to the flux of ions are omitted here, because the discussion was limited to a material without ion migration. In a-GaO_x these terms should not be relevant anyway, since the ions can be expected to migrate on a significantly slower timescale than the electrons.[10, 33]

The result is the fundamental macroscopic flux equation for electrons as the sum of the diffusion driven by $\nabla\mu$, the drift due to $\nabla\varphi$ and the thermoelectric effect through ∇T:

$$\vec{J}_\mathrm{e} = -\frac{\sigma}{e^2}\nabla\tilde{\mu} + \frac{\sigma}{e}S_\mathrm{e}\nabla T = -\frac{\sigma}{e^2}\nabla\mu + \frac{\sigma}{e}\nabla\varphi + \frac{\sigma}{e}S_\mathrm{e}\nabla T \tag{2.7}$$

Note that the flux of ions with charge q_i in a solid is governed by the analog equation

$$\vec{J}_i = -\frac{\sigma_i}{q_i^2}\nabla\tilde{\mu}_i - \frac{\sigma_i}{q_i}S_i\nabla T = -\frac{\sigma_i}{q_i^2}\nabla\mu_i - \frac{\sigma_i}{q_i}\nabla\varphi - \frac{\sigma_i}{q_i}S_i\nabla T \tag{2.8}$$

where the ionic conductivity σ_i, the electrochemical potential $\tilde{\mu}_i = \mu_i + q_i\varphi$ and the Seebeck coefficient S_i refer to the respective ionic species. In comparison to Ohm's law (eq. 2.1), the definition of $\vec{J}_\mathrm{e}$ in eq. 2.7 offers more insight into the driving forces of electron flux, yet it is still general enough to serve as the basis for the later KMC and FEM simulations. In a-GaO_x, the thermoelectric term most likely can be neglected, since the Seebeck coefficient S_e is in the order of µV/K while $\nabla\varphi$ may even reach MV/cm in thin film devices ([40]: $S_\mathrm{e} = -(1.9 \pm 0.4)$ µV/K for a-$GaO_{0.9}$ at room temperature). Large ∇T are not expected either, because – as will be discussed later – a-GaO_x is typically well conductive, which suggests a high thermal conductivity and a low impact of Joule heating. Henceforth, just isothermal electron transport shall be considered ($\nabla T = 0$).

While electric field and temperature are intuitive quantities, the chemical potential is a rather abstract driving force that shall be examined further. Essentially, μ is a measure for the local electron density n. More precisely, it is the non-electric potential energy inherent to the presence of a certain amount of electrons. In thermodynamics, μ and n are linked by the activity coefficient γ according to [37]

$$\mu = \mu_0 + k_\mathrm{B}T\ln\left(\gamma\frac{n}{n_0}\right) \tag{2.9}$$

where μ_0 is the reference chemical potential belonging to the reference electron density n_0. The density of core electrons is constant in the absence of ion transport, such that n shall refer to the valence electrons only. In general, electrons are neither ideal ($\gamma \neq 1$), nor is γ a constant. In the here regarded homogeneous and isothermal material, γ just varies with n or μ. This allows to transform the $\nabla\mu$-term of eq. 2.7 into Fick's first law of diffusion with ∇n as driving force:

$$\nabla\mu = k_\mathrm{B}T\nabla\ln\left(\gamma\frac{n}{n_0}\right) = \frac{k_\mathrm{B}T}{\gamma n}\nabla(\gamma n) = \frac{k_\mathrm{B}T}{\gamma n}\left[\gamma\nabla n + n\frac{\partial\gamma}{\partial n}\nabla n\right] = \frac{k_\mathrm{B}T}{n}\left[1 + \frac{n}{\gamma}\frac{\partial\gamma}{\partial n}\right]\nabla n \tag{2.10}$$

$$\Rightarrow \quad \nabla\mu = \frac{k_\mathrm{B}T}{n}\Phi\nabla n \quad \text{with} \quad \Phi = 1 + \frac{n}{\gamma}\frac{\partial\gamma}{\partial n} \tag{2.11}$$

$$\Rightarrow \quad \vec{J}_\mathrm{e} = -\frac{\sigma}{e^2}\nabla\mu + \frac{\sigma}{e}\nabla\varphi = -D\nabla n + \frac{\sigma}{e}\nabla\varphi \quad \text{with} \quad D = \frac{\sigma}{e^2}\frac{k_\mathrm{B}T}{n}\Phi \tag{2.12}$$

Here, Φ is the thermodynamic factor and D is the Fickian diffusion coefficient of the electrons. In summary, the electrons flow towards regions with lower electron density and against the direction

of the electric field ($\sigma > 0$ and $D > 0$ always). Again, analog expressions describe ion transport, with the consequence that both gallium and oxygen ions in a-GaO_x would diffuse towards where their concentration is low. Simultaneously, the positively charged gallium ions would flow in the direction of $\vec{F} = -\nabla\varphi$, while the negatively charged oxygen ions would move antiparallel to $\vec{F}$ like the electrons. The extent of these gallium and oxygen fluxes of course depends on the respective ionic conductivities (cf. sec. 2.3.4). Regarding the interplay of electron and ion transport, it was already mentioned that ionic fluxes would mainly affect the electrons by altering their local stoichiometry-dependent conductivity σ. For resistive switching, however, it is also important how the electrons affect the flux of ions through their mutual driving force, the electric field.

An electric field is always present when there are uncompensated positive or negative charges, as described by the Poisson equation [41]

$$\nabla^2\varphi = -\nabla\vec{F} = -\frac{\rho}{\epsilon_0\epsilon_\mathrm{r}} \tag{2.13}$$

Here, ϵ_0 is the vacuum permittivity, ϵ_r is the material's relative permittivity and ρ is the charge density, i.e., the sum of positive charges (e.g., gallium ions) and negative charges (e.g., electrons and oxygen ions). According to the Poisson equation, the variation of the electric field $\vec{F}$ throughout the material depends on how electrons and ions are distributed. Consider for example the homogeneous, isothermal material regarded so far. When the valence electrons would be evenly distributed as well, the material would be charge neutral everywhere ($\rho = 0$), resulting in a constant electric field throughout the material ($\nabla\vec{F} = 0$ according to eq. 2.13). The actual strength $F = |\vec{F}|$ of this electric field depends on the applied voltage V. In general, V corresponds to the difference of the electrons' electrochemical potential $\tilde{\mu}$ between the contacts.[5] However, when not only the material composition but also n is equal at the contacts, then μ is equal there as well (cf. eq. 2.9). In this case, V is just the difference of φ between the contacts and F is the voltage divided by the distance between the contacts. This idealized situation is not persistent of course, due to the ensuing movement of electrons and ions. On the one hand, the electric field acts upon the electrons and ions. On the other hand, electrons might diffuse from the regarded material into the electrodes, because μ is typically lower in metals (cf. sec. 2.2.4 about Schottky barriers). The drift fluxes due to the electric field cause gradients of the electron density and the ion concentrations, leading to diffusion fluxes that work against the drift fluxes (eqs. 2.7 and 2.8). In turn, the diffusion fluxes due to chemical potential gradients cause charge imbalances, i.e., counter-acting variations of the electric field (eq. 2.13). Since the electrons typically move much faster than the ions, a balance between electron drift and electron diffusion is reached on timescales on which the ions are practically static. For each momentary, quasi-static ion distribution, the redistribution of electrons adapts both μ and ρ – and thereby the electric field according to the Poisson equation – until a spatially constant steady-state

[5] Primarily, an applied voltage V is the difference of φ between the outlets of the voltage source. Since these are made from the same metal, i.e., μ is equal, V is also the difference of $\tilde{\mu}$ between the outlets. When the cabling to the sample is isothermal and has negligible resistance, then $\tilde{\mu}$ virtually does not change from each outlet to the respective contact point/electrode on the sample (cf. eq. 2.7). Therefore, V corresponds to the difference of $\tilde{\mu}$ between the contacts.

electron flux is established according to the continuity equation [41]

$$\frac{\partial n}{\partial t} = -\nabla \vec{J}_\mathrm{e} = 0 \quad \Rightarrow \quad \nabla \left[\frac{\sigma}{e^2} \nabla \tilde{\mu} \right] = \nabla \left[\frac{\sigma}{e^2} (\nabla \mu - e \nabla \varphi) \right] = 0 \tag{2.14}$$

The difference of $\tilde{\mu}$ across the material thereby equals the applied voltage V. The electrons redistribute under this constraint until $\sigma \nabla \tilde{\mu}$ is constant throughout the material. The resulting electric field is still controlled by the voltage, although not in the trivial way mentioned before. As soon as $\rho \neq 0$ somewhere within the material, the electric field is no longer constant (eq. 2.13) and F does not equal V divided by the distance of the contacts anymore. Instead, V influences $\vec{F} = -\nabla \varphi$ indirectly, since it defines $\tilde{\mu} = \mu - e\varphi$ at the contacts. Together with any variation of μ that may results from the redistribution of electrons (or ions), the applied voltage therefore enforces certain φ at the contacts. In this way, V affects the varying electric potential φ (and hence also the electric field) that establishes within the material according to the Poisson equation (eq. 2.13).

According to eq. 2.14, the steady-state electron distribution depends on the local conductivity σ of the electrons throughout the material. The electrochemical potential gradient of the electrons becomes large everywhere where σ is low. As will be described in sec. 2.3 in more detail, in a-GaO_x the conductivity depends on the local stoichiometry x, i.e., on eventual fluxes of oxygen or gallium ions. Typically, σ is higher the lower x, but the actual relation between σ and x depends on the electron conduction mechanism on the atomic scale (cf. sec. 2.2). Additionally, the conductivity σ generally varies also with n (and μ), in a way that is determined by the conduction mechanism as well. In both ways, the electron conduction mechanism impacts the steady-state electron distribution and the steady-state flux $\vec{J}_\mathrm{e}$. Even in the model case with homogeneous material composition, σ may not be constant as soon as n (or μ) varies spatially. This for example happens when Schottky barriers form in the vicinity of a-GaO_x|metal interfaces (sec. 2.2.4). In this work, the electron conduction mechanism will be investigated, followed by KMC simulations that elucidate the relation of σ to μ and x, and FEM calculations that yield the steady-state electron distributions by coupling the Poisson equation (eq. 2.13) with the continuity equation (eq. 2.14).

From the perspective of the ions, the electrons instantaneously adapt to any changes of the ions' distribution and the electric field that results from the steady-state electron distribution is the driving force for the further movement of the ions. Thereby, the local electron conductivity σ influences where within the material the electric field becomes strong enough to cause significant ion redistribution, which in turn changes the local σ. This feedback loop is the general principle behind resistive switching and the three major RS mechanisms – filamentary, interface or bulk switching – basically differ in where it takes place, as indicated by their names. They will be described in more detail in sec. 2.4. From outside, the interplay of electrons and ions is observed as a timely varying measured current I, which essentially corresponds to the momentary steady-state electron flux when the electrons are the majority charge carriers. More precisely, I equals the current density $\vec{j} \approx -e\vec{J}_\mathrm{e}$ integrated over the contact area. Therefore, the resistive switching effect can be observed as a hysteresis in the I-V response of the material.

2.1.1 Effective Charge Carrier Density

So far, the treatment of electron transport followed the conventional description of particle fluxes in non-equilibrium thermodynamics.[34] The analog equations for ion migration are part of every textbook about solid-state diffusion.[37, 38] Even on the macroscopic scale, there is however a difference between electron and ion conduction.[6] It is common to consider all ions of a mobile species (e.g., all ions from a certain element or all ions in a certain sub-lattice) as charge carriers for the respective ionic flux.[37] This is not the case for electrons, where according to quantum mechanics only a certain fraction of n, which is called effective charge carrier density n_{eff}, contributes to the macroscopic flux.[42, 43] In the following, the reasons for this will be explained alongside a condensed introduction to the quantum mechanical description of electrons in solids. The subsequent combination with the irreversible thermodynamics discussed before will lead to the definition of n_{eff}.

As mentioned before, on the quantum scale the electrons occupy electronic states that are described by complex wave functions ψ. These wave functions are solutions to the Schrödinger equation of the regarded quantum system, or more precisely, they are part of the many-body wave function that describes how the system of atoms and electrons evolves with time as a whole.[7] Therein, each single-electron wave function ψ represents a quantum state that an individual electron may adopt.[8] In other words, the electron "is" a particle wave ψ that spreads over a certain region of space. Consequently, ψ contains all information about the electron. For example, the probability density of observing the electron at a precise point in space (and time) is given by the squared modulus $|\psi|^2$. The energy of the electron is $E = \langle\psi|\hat{h}|\psi\rangle$ (bra-ket notation), where $\hat{h}$ is the single-electron Hamiltonian that includes the electron's interactions with other electrons, the cores of the atoms and external fields. A wave function may describe either a stationary state, which means that $|\psi|^2$ does not vary with time, or a non-stationary state. The wave functions of stationary states are the eigenfunctions of $\hat{h}$, with eigenvalue E according to the time-independent form of the Schrödinger equation $\hat{h}|\psi\rangle = E|\psi\rangle$.[44] Since $|\psi|^2$ does not evolve with time, they do not describe macroscopic movement of electrons per se, but electrons may transition between stationary states for example via excitation, relaxation or tunneling (cf. sec. 2.2.3). The wave functions of non-stationary states depict dynamic processes, where the region where the electron is likely to be observed changes with time. An example is the acceleration of an electron in an external electric field $\vec{F}$. As long

[6] The here discussed difference between electron and ion conduction is merely a difference in how they are commonly treated than in the actual physics. For example, ions also have an energetic distribution and require a vacant space nearby in order to migrate. The fraction of ions that actually migrates at a time is limited like for electrons, but since the underlying probability functions are mostly unknown except for idealized cases, this typically remains hidden inside the activity coefficient or the thermodynamic factor. Note that eq. 2.9 is a modification of such an idealized case, namely Boltzmann-distributed particle energies.

[7] An in-depth or more formal introduction to quantum mechanics in general and solid-state electron wave functions in particular would go beyond the scope of this work and can be found in respective textbooks.[35, 44–47] This includes the decomposition of the many-body wave function, e.g., using the Born-Oppenheimer approximation, Slater determinants and Bloch waves. Concepts that are exclusive to crystalline compounds are avoided here as well.

[8] Throughout this work, the term "electronic state" refers to a single-electron state, which means that the spin function is included within ψ. Thus, for each state there is another one with equal $|\psi|^2$ and opposite spin.

as $\hat{h}$ is not explicitly time-dependent (e.g., $\vec{F}$ approximately constant in time), it is possible to express any non-stationary state as a superposition of stationary eigenstates ψ_j (times the respective phase factor $\exp(-iE_jt/\hbar)$).[44] As a result, ψ of a non-stationary state has the form of a more or less extended wave packet that corresponds to the macroscopically moving electron as a particle. In such a superposed state, the electron simultaneously occupies each of the partaking eigenstates with a certain probability. This also means that there is a quantum uncertainty for the energy of the electron and $E = \langle\psi|\hat{h}|\psi\rangle$ is just its expectation value but no longer a sharp eigenvalue of $\hat{h}$. The formation of such a wave packet essentially corresponds to a partial excitation of an electron to other eigenstates, which is only possible if such states are available. The resulting probability distribution among the eigenstates governs the magnitude and the direction of the group velocity, i.e., the velocity of the position expectation value of such wave packets. In absence of an external electric field, the wave packets have time-invariant group velocities and, on average, each wave packet moving in one direction is compensated by another traveling in the opposite direction, which leads to zero net flux. An external electric field breaks this balance. The eigenstates adapt to the additional potential gradient and exhibit a superimposed variation of $|\psi_j|^2$ in the direction of $\vec{F}$.[9] The energy of the eigenstates changes and becomes lower the higher $|\psi_j|^2$ is in regions with low $q\varphi$ ($\langle\psi_j|q\varphi|\psi_j\rangle$ term in the expression for E_j). The superposition of these modified stationary states results in wave packets that accelerate at the expense of electric potential energy, in principle causing a non-zero net flux of electrons within the material.[44] However, a wave packet cannot traverse the whole material undisturbed. The accelerated motion only continues until a scattering event, e.g., with static or dynamic disturbances in the atomic structure, forces a transition to another non-stationary state that propagates in a different direction (cf. sec. 2.2.2). If no other state is available, the electron bounces back and oscillates in its confined non-stationary state. The same applies to other transition options like the excitation/relaxation to states with higher/lower energy. In summary, any macroscopic electron flux requires transitions between quantum states. The Pauli exclusion principle however forbids that two electrons occupy the same state simultaneously.[10] This means that only those electrons can contribute to $\vec{J}_e$ that can reach unoccupied states. In order to quantify this restriction, the energetic distributions of the eigenstates and the electrons have to be regarded.

In a solid, the stationary electronic states are so numerous that the energy differences between them are minuscule. Therefore, the technically quantized energy eigenvalues are typically considered to form a continuous energy scale E. The density of states (DOS) $g(E)$ describes how the electronic states are distributed on E.[11] Put in more precise terms, $g(E)\mathrm{d}E$ is the number density of electronic

[9] For example, the stationary wave functions of a free electron in a constant electric field (i.e., in a triangular potential well) are Airy functions instead of plane waves.[45, 48]

[10] More specifically, the Pauli exclusion principle states that the total many-body wave function must be antisymmetric with regard to electrons, i.e., change sign upon the exchange of two electrons. For different non-stationary single-electron states, each spread among the eigenstates, this can be interpreted as the requirement to differ "sufficiently" in their probability distribution among the stationary states. The total probability to find an electron in a certain eigenstate must not exceed unity.[35] This also means that an electron cannot move in the form of a wave packet when all spatially and energetically close stationary states are occupied.

[11] The basic layout of the DOS for metals and semiconductors should be familiar to the reader, incl. the respective terminology (e.g., valence band, band gap, conduction band edge, etc.). Otherwise see for example [41, 44, 46, 47].

states with an energy in the infinitesimal interval dE around E. At $T = 0\,\mathrm{K}$, all electrons occupy the lowest possible states. In this case, all states are filled with electrons up to the Fermi level $E_\mathrm{F}(0\,\mathrm{K})$, which is therefore related to the electron density n via

$$n = \int_{-\infty}^{E_\mathrm{F}(0\,\mathrm{K})} g(E)\,\mathrm{d}E \tag{2.15}$$

When $T > 0\,\mathrm{K}$, then a part of the electrons is thermally excited to higher states, leaving behind unoccupied states below the Fermi level, as described by the Fermi-Dirac distribution f_FD [44]

$$f_\mathrm{FD} = \left[\exp\left(\frac{E - E_\mathrm{F}}{k_\mathrm{B}T}\right) + 1\right]^{-1} \tag{2.16}$$

that yields the occupation probability for a state with energy E, provided that the electrons are in mutual thermal equilibrium. The Fermi level E_F is then the energy where the occupation probability is 50% ($f_\mathrm{FD}(E_\mathrm{F}) = 0.5$) and the relation to the electron density n becomes

$$n = \int_{-\infty}^{\infty} g(E) f_\mathrm{FD}(E)\,\mathrm{d}E \tag{2.17}$$

Note that $1 - f_\mathrm{FD}(E)$ corresponds to the probability that a state with energy E is unoccupied. It already becomes visible that the DOS and the Fermi-Dirac distribution together determine where occupied states are energetically close to unoccupied states, i.e., which electrons can reach the empty states required for macroscopic electron flux. However, before the derivation of n_eff, the connection to thermodynamics and several common approximations regarding the influence of temperature and electric field need to be clarified.

In the case of $T = 0\,\mathrm{K}$, it is obvious that an additional electron would occupy a state at the Fermi level. This would increase the system's Gibbs free energy G by E_F. The same also applies to $T > 0\,\mathrm{K}$ when the added electron is in thermal equilibrium. Therefore, the Fermi level equals the electrochemical potential of the electrons (cf. eq. 2.3): $E_\mathrm{F} = \tilde{\mu} = \mu - e\varphi$. When $\varphi = 0$, then E_F is also the chemical potential of the electrons.

When the temperature T of the electrons changes, this mainly affects their energetic distribution according to eq. 2.16. However, any variation in the state occupations in principle also influences the other electronic states in the system through the electron-electron interaction, thereby changing the DOS. It is common to consider this effect negligible and treat $g(E)$ as T-independent. Even with this approximation, E_F generally varies slightly with T as a result of eq. 2.17 when n stays constant. The change of $E_\mathrm{F}(T)$ is more pronounced the more asymmetric $g(E)$ is in the vicinity of the Fermi level. When the DOS is symmetric around E_F, then follows from eq. 2.17 that E_F does not vary with T. Note that in this way the quantum mechanics also yield the T-dependence of the thermodynamic quantities $\tilde{\mu}$ and μ (since φ does not depend on T).

In general, the influence of an external electric field on the DOS is less subtle, since $q\varphi$ becomes part of the Hamiltonian and, as explained above, affects both energy and wave function of the eigenstates directly. It is however common to approximate the effect of $\vec{F}$ as a shift of the whole DOS by $q\varphi$ along the energy axis, without any change of the actual density values.[49, 50] This assumption is widely used for example for the description of band bending in space charge zones of Schottky barriers or pn-junctions. It is justified as long as the external electric field is a weak perturbation in comparison to the other forces on the atomic scale (which is typically the case) and as long as it is approximately constant over the extent of the wave functions (each state shifts approx. by $\langle\psi|q\varphi|\psi\rangle \approx q\varphi$). Accordingly, the DOS may as well be defined on an F-independent energy axis $\varepsilon = E - q\varphi$ that denotes the state energy without electric potential energy. This not only means that $g(\varepsilon)$ has the same value as the respective $g(E)$, but also $E - E_\mathrm{F} = \varepsilon - \mu$. From eq. 2.17 then the common thermodynamic notion is recovered that the local electron density n just depends on the material (i.e., $g(\varepsilon)$), the chemical potential μ and the temperature T, but not directly on φ (cf. eq. 2.9):

$$n = \int_{-\infty}^{\infty} g(\varepsilon) f_\mathrm{FD}(\varepsilon, \mu)\,\mathrm{d}\varepsilon = \int_{-\infty}^{\infty} g(\varepsilon)\left[\exp\left(\frac{\varepsilon-\mu}{k_\mathrm{B}T}\right)+1\right]^{-1}\mathrm{d}\varepsilon \tag{2.18}$$

The chemical potential therefore denotes the field-independent position of the Fermi level relative to the electronic states. An increase of μ equals a shift of E_F upwards in the DOS, corresponding to higher electron density and a negative charging of an otherwise charge-neutral material. In turn, decreasing μ describes positive charging. Note that any such change of n in principle affects all electronic states through electron-electron interaction. As discussed for T above, this effect is typically ignored.

The ingredients for the derivation of n_eff are now complete. According to quantum mechanics, only a certain part of the valence electrons may contribute to $\vec{J}_\mathrm{e}$, namely those electrons that are close to the unoccupied states around the Fermi level. The thermodynamic analogon is the thermodynamic factor Φ that determines which fraction of n is "active". This becomes clearer, when examining eq. 2.12 a bit further. While the conductivity σ is a measure for all charge carriers together, the Fickian diffusion coefficient D is a per-particle property. In the relation between the two (eq. 2.12), σ is divided by n/Φ in order to get the contribution of a single charge carrier ($\Phi \geq 1$). This means that

$$n_\mathrm{eff} = \frac{n}{\Phi} \tag{2.19}$$

is the density of electrons that actively participate in charge transport. The same term also connects the two driving forces $\nabla\mu$ and ∇n according to eq. 2.11. This allows to calculate n_eff from the just introduced quantum mechanical relation between μ and n (eq. 2.18): [42]

$$\nabla n = \frac{\partial n}{\partial \mu}\nabla\mu \quad \text{and (eq. 2.11)} \quad \nabla n = \frac{n_\mathrm{eff}}{k_\mathrm{B}T}\nabla\mu \quad \Rightarrow \quad n_\mathrm{eff} = k_\mathrm{B}T\frac{\partial n}{\partial \mu} \tag{2.20}$$

$$\frac{\partial n}{\partial \mu} = \int_{-\infty}^{\infty} g(\varepsilon) \frac{\partial f_{\mathrm{FD}}(\varepsilon, \mu)}{\partial \mu} \, \mathrm{d}\varepsilon \tag{2.21}$$

$$\frac{\partial f_{\mathrm{FD}}(\varepsilon, \mu)}{\partial \mu} = \frac{1}{k_{\mathrm{B}} T} \frac{\exp\left(\frac{\varepsilon-\mu}{k_{\mathrm{B}} T}\right)}{\left[\exp\left(\frac{\varepsilon-\mu}{k_{\mathrm{B}} T}\right)+1\right]^2} = \frac{1}{k_{\mathrm{B}} T} f_{\mathrm{FD}}(\varepsilon, \mu)\left[1 - f_{\mathrm{FD}}(\varepsilon, \mu)\right] \tag{2.22}$$

$$n_{\mathrm{eff}} = \int_{-\infty}^{\infty} g(\varepsilon) f_{\mathrm{FD}}(\varepsilon, \mu)\left[1 - f_{\mathrm{FD}}(\varepsilon, \mu)\right] \, \mathrm{d}\varepsilon \tag{2.23}$$

Note that this derivation is based on a spatially constant DOS, i.e., a homogeneous material, and isothermal conditions, similar to the definition of Φ via eq. 2.10. However, the additional terms that would arise from the chain rule in inhomogeneous or non-isothermal materials would not concern the conductivity but the eventual Onsager cross-coefficients that couple electron, ion and heat flux (cf. eq. 2.4). In the final eq. 2.23 for the effective charge carrier density n_{eff} the term $f_{\mathrm{FD}}[1 - f_{\mathrm{FD}}]$ selects from the DOS the symmetric region around the Fermi level where charge transport happens due to the simultaneous presence of occupied and unoccupied states. Any investigation of electron transport therefore has to focus on this part of the DOS.[12] This is also reflected by the relation between σ and the spectral conductivity $\sigma(\varepsilon)$ from Kubo-Luttinger linear response theory [35, 36, 51]

$$\sigma = \int_{-\infty}^{\infty} \sigma(\varepsilon) \left(-\frac{\partial f_{\mathrm{FD}}(\varepsilon, \mu)}{\partial \varepsilon}\right) \mathrm{d}\varepsilon = \frac{1}{k_{\mathrm{B}} T} \int_{-\infty}^{\infty} \sigma(\varepsilon) f_{\mathrm{FD}}(\varepsilon, \mu)\left[1 - f_{\mathrm{FD}}(\varepsilon, \mu)\right] \, \mathrm{d}\varepsilon \tag{2.24}$$

where $\sigma(\varepsilon)$ is the theoretical conductivity due to all electronic states with energy ε. This equation is popular for linking conduction mechanisms that yield $\sigma(\varepsilon)$ to the macroscopic conductivity. For band transport via wave packets for example, $\sigma(\varepsilon)$ is related to the sum of the group velocities of all states with energy ε.[35] In turn, eq. 2.23 for n_{eff} is useful to express the conductivity in a more macroscopic way as the combination of a per-charge-carrier property and the effective number of charge carriers. The respective relation for the diffusion flux was already explained above and follows from eq. 2.12 as

$$\frac{\sigma}{e^2} = \frac{n_{\mathrm{eff}} D}{k_{\mathrm{B}} T} \tag{2.25}$$

The analog per-charge-carrier quantity for the drift flux driven by $\nabla\varphi$ is the mobility u, defined as the average charge carrier drift velocity $\langle v_{\mathrm{d}} \rangle = |\langle \vec{v}_{\mathrm{d}} \rangle|$ per applied electric field strength $F = |\vec{F}|$:

$$\langle \vec{v}_{\mathrm{d}} \rangle = -u\vec{F} \quad \Rightarrow \quad u = \frac{|\langle \vec{v}_{\mathrm{d}} \rangle|}{|\vec{F}|} = \frac{\langle v_{\mathrm{d}} \rangle}{F} \tag{2.26}$$

According to the outdated Drude model, which considers all valence electrons as charge carriers,

[12] Electrons are indistinguishable and may swap states over time, so always the electrons around E_{F} take part in the flux but not always the same electrons are flowing.

it relates to σ via [44]

$$\sigma = enu_\mathrm{N} \quad \text{with} \quad u_\mathrm{N} = \frac{\langle v_\mathrm{d}\rangle_\mathrm{N}}{F} \quad \text{and} \quad \langle v_\mathrm{d}\rangle_\mathrm{N} = \frac{1}{N}\sum_k^N v_{\mathrm{d}k} \tag{2.27}$$

where $N = nV$ is the number of valence electrons in the regarded volume V and $v_{\mathrm{d}k}$ is each charge carriers' displacement per time in the direction of $\nabla\varphi$. As discussed already, the picture that this model conveys is not appropriate for electrons (in contrast to ions). It is mentioned here, because it can be easily confused with the classical semiconductor theory found in textbooks (sometimes even referred to as "Drude model"). The reason is that in the Drude model u_N is typically written as "u", leading to "$\sigma = enu$". At first sight, the same expression is typically used for n-type non-degenerate semiconductors, albeit with a different meaning of the quantities. There, "n" denotes just those valence electrons that are excited to the conduction band (CB) and "u" is also only an average over the CB electrons. Instead of corresponding to the Drude model, the classical equations in semiconductor theory are special cases of the effective charge carrier model. For n-type non-degenerate semiconductors n_eff equals the conduction band electron density, because $1 - f_\mathrm{FD} \approx 1$ in the CB, $1 - f_\mathrm{FD} \approx 0$ in the valence band (VB) and E_F closer to the CB edge than to the VB edge (cf. eq. 2.23). This example illustrates that the electron mobility in general should refer to the effective charge carriers instead of all valence electrons: [42, 43, 50]

$$\begin{aligned}\sigma =& enu_\mathrm{N} = enu_\mathrm{N}\frac{\Phi}{\Phi} = en_\mathrm{eff}u \quad \text{with} \quad u = u_\mathrm{N}\Phi = \frac{\langle v_\mathrm{d}\rangle_\mathrm{N}\Phi}{F} = \frac{\langle v_\mathrm{d}\rangle}{F} \\ &\text{and} \quad \langle v_\mathrm{d}\rangle = \langle v_\mathrm{d}\rangle_\mathrm{N}\Phi = \frac{\Phi}{N}\sum_k^N v_{\mathrm{d}k} = \frac{1}{N_\mathrm{eff}}\sum_k^N v_{\mathrm{d}k} \quad \text{with} \quad N_\mathrm{eff} = n_\mathrm{eff}V\end{aligned} \tag{2.28}$$

Similar to D, the per-charge-carrier quantity u is multiplied with n_eff to yield the conductivity σ. This has also consequences for the Nernst-Einstein relation that results from eqs. 2.25 and 2.28:

$$D = \frac{uk_\mathrm{B}T}{e} \tag{2.29}$$

Technically, this is the so-called "generalized" Nernst-Einstein relation, since it takes the activity of the flowing particles into account, whereas the "classical" Nernst-Einstein relation connects the mobility and diffusivity of ideal particles (or tracers).[13] Originally, this distinction has been introduced because D naturally refers to the "active" particles (i.e., to $n_\mathrm{eff} = n/\Phi$, cf. eq. 2.12), while the mobility often has been defined with regard to all particles or sub-groups thereof (e.g., w.r.t. n in the Drude model, cf. eq. 2.27). To generalize the Nernst-Einstein relation then of course has required an additional factor to compensate for this deviation (e.g., Φ as a result of combining eqs. 2.12 and 2.27).[46, 52] This is not necessary in the n_eff-formalism where u and D both refer to the "active" particles. As a consequence, eq. 2.29 retains the same form as the original Nernst-Einstein relation. A distinction between "classical" and "generalized" Nernst-Einstein relation is then obsolete.

[13] A more in-depth discussion of Fickian diffusion (where $\nabla n \neq 0$) vs. tracer diffusion (where $\nabla n = 0$) with regard to electrons can be found in the supplement of [40].

Finally it shall be emphasized that the n_{eff}-formalism is not coupled to a specific electron transport mechanism. It mainly requires that the electrons adhere to the Fermi-Dirac distribution and that the electron flux involves empty quantum states. In contrast to the idealized semiconductor theory found in most textbooks, it applies to crystalline and amorphous electron conductors alike, regardless of whether they are metals, degenerate semiconductors or non-degenerate semiconductors. The actual values of u and D, however, depend on the specific electron transport mechanism. In order to complete the picture, the following section is therefore dedicated to the quantum mechanical behavior of the electrons in the vicinity of E_{F} in amorphous solids.

2.2 Electron Transport in Amorphous Solids

Amorphous solids have no long-range order in contrast to crystalline solids. They contain mainly three types of structural disorder: varying bond lengths, varying bond angles and under- and over-coordinated sites.[53] These are tightly connected to a second type of disorder, the energetic disorder, which means that the electron-ion interaction (in $\hat{h}$) differs substantially from periodicity. This section is concerned with the consequences of the disorder for electron conduction. The concept of localized and extended electronic states, the bulk electron transport via mobility edge conduction (MEC) or variable-range hopping (VRH), and the conflicts with the classical Schottky barrier model for metal-semiconductor interfaces will be explained.

2.2.1 Extended and Localized States

In a crystal, the stationary wave functions[14] are Bloch waves $\psi_k(\vec{r}) = u_k(\vec{r})\exp(i\vec{k}\vec{r})$ where u_k is a function with same periodicity as the lattice and $\vec{k}$ is the wave vector of the plane wave backbone of ψ_k, serving also as a quantum number to identify the state.[47] The components of $\vec{k}$ in reciprocal space are the wave numbers k. The wave vector describes the phase velocity $v_{\mathrm{p}} = E/(\hbar k)$ of a Bloch wave with energy E, which is why $\hbar\vec{k}$ is called the quasi-momentum (or crystal momentum) of a Bloch wave. The wave packets in a crystal are very compact in reciprocal space (otherwise they would disperse quickly), which means that they consist of Bloch waves in a small range $\Delta\vec{k}$ around a wave vector $\vec{k}$ ($\Delta k/k \ll 1$ in every direction). Correspondingly, according to the Heisenberg uncertainty principle[15] they extend over a broad region $\Delta\vec{r}$ in real space, spanning several atomic spacings.[44] These Bloch wave packets move with a group velocity $\vec{v}_{\mathrm{g}} = \nabla_k E(\vec{k})/\hbar$ that depends

[14] Note that in presence of an external electric field (or other disturbances of the periodicity like impurities/defects) the Bloch waves are no longer the single-electron eigenfunctions of a crystal. Perturbation theory however allows to express the actual eigenstates through Bloch waves as basis functions. This is especially elegant when the perturbation is weak, leading to the here roughly presented semi-classical Boltzmann transport theory of acceleration and scattering of Bloch wave packets in crystals.[35, 54]

[15] For a wave packet in vacuum (free electron) is $\langle\vec{p}\rangle = \hbar\vec{k}$ the real momentum expectation value. The Heisenberg position-momentum uncertainty principle $\Delta x\Delta p_x \geq \hbar/2$ then becomes $\Delta x\Delta k_x \geq 1/2$ (in x-direction as example).[46, 47] It is not that simple for Bloch wave packets due to u_k, but the tendency of increasing $\Delta\vec{r}$ with decreasing $\Delta\vec{k}$ (and vice versa) typically remains.

on the dispersion relation between E and $\vec{k}$ (∇_k denotes the derivative along the axes in reciprocal space). On their way trough the crystal, the wave packets get scattered when they collide for example with lattice defects or with phonons, which are the quasi-particles of lattice vibration. Scattering alters the trajectory of wave packets in a random fashion, leading to a diffusive motion characterized by the mean free path ℓ_m and the mean free time τ_m between scattering events.[16]

Both ℓ_m and τ_m decrease with rising disorder in a crystal, which gives a rough idea of what happens when a material becomes amorphous. Since scattering is generally not uniform for all wave packets, also the effects of disorder may differ. As long as the mean free path ℓ_m of a wave packet remains rather large, its movement through the material is just slowed down by the additional scattering. However, when ℓ_m approaches interatomic distances ($k\ell_\mathrm{m} \approx 1$) the wave packet becomes confined to that space, i.e., less delocalized, and cannot move without scattering to another state.[17] Any movement is then accompanied by ever-changing $\vec{k}$ and $\vec{v}_\mathrm{g}$ of the wave packet.[47] It only seems logical that when a wave packet would be affected by disorder even stronger it would become even more localized and could no longer scatter freely from state to state. It would become a localized stationary state, comparable to impurity trap states but much more abundant. However, this regime cannot be described within the semi-classical model of scattering Bloch wave packets, since this model already reaches its limits when $k\ell_\mathrm{m} \approx 1$ (Ioffe-Regel criterion).[47, 54] Restricting the extent of Bloch wave packets that much in real space ($|\Delta\vec{r}| \leq \ell_\mathrm{m}$) corresponds to an increase of $\Delta\vec{k}$ to the same order as $\vec{k}$ due to the uncertainty principle.[47] The uncertainty in $\vec{k}$ is then so large that the description of the state through superposed Bloch waves is meaningless.

In amorphous solids, the stationary states cannot be described by Bloch waves ψ_k due to an even more fundamental reason: The total lack of long-range periodicity in the Hamiltonian makes it impossible to solve the Schrödinger equation using the product of a periodic function u_k and a plane wave with wave vector $\vec{k}$.[53] There is no meaningful wave vector associated with each electronic state as a quantum number and the concept of a band structure $E(\vec{k})$ loses its validity. Even though short-range order is typically preserved, the amorphous state is in general too disordered to treat it like a perturbed crystal.[18] The mathematics of the Bloch wave theory are therefore not appropriate, but the concepts developed for crystals are still valuable for amorphous compounds. The orbital overlap combined with the high density of atoms still creates dispersed states with minuscule separation on the energy scale like in crystalline solids. Accordingly, there is a quasi-continuous DOS with valence and conduction band in amorphous solids as well. The wave functions can still be approximated by a linear combination of atomic orbitals (tight binding model), but the coefficients no longer vary periodically but depend on the variation of the potential term in the Hamiltonian throughout the material.[31, 53] In principle, an amorphous structure therefore does not prevent

[16] More generally, ℓ_m and τ_m refer to quasi-momentum relaxation. Since this may take more than one scattering event, they are defined as the average distance and time in which the phase of a wave packet becomes completely random.

[17] In the semi-classical model that uses Bloch wave packets with constant composition in the $\vec{k}$-space, any change of the wave packet's dynamics is a state transition. Closer to reality is that the eigenfunctions adapt to the disorder, leading to non-Bloch wave packets that have varying momentum despite constant distribution among eigenstates.

[18] An exception to this is for example Ziman's theory of liquid metals.[54]

wide-spread wave functions per se, but the form of ψ is not as regular as in crystals. The effects of disorder-induced potential variations on the wave functions can be illustrated by the following Anderson model.[31, 47, 55]

The basic principle already becomes visible in a quantum system of two adjacent potential wells, say with potential energies V_a and V_b.[19] The overlap of the respective "atomic" single-well wave functions ϕ_a and ϕ_b leads to two electronic states with approximate wave functions $\psi_1 = A\phi_\mathrm{a} + B\phi_\mathrm{b}$ and $\psi_2 = B\phi_\mathrm{a} - A\phi_\mathrm{b}$, where the symmetric state ψ_1 has lower energy than ψ_2 ($E_1 < E_2$). When V_a and V_b are quite similar ($|V_\mathrm{a} - V_\mathrm{b}| \ll I$ with the transfer integral $I = |\langle\phi_\mathrm{a}|\hat{h}|\phi_\mathrm{b}\rangle|$), then is $A \approx B$ and $|E_1 - E_2| \approx 2I$.[31] In this case, both states have rather equal $|\psi|^2$ in both wells. Electrons that occupy these states would be almost evenly delocalized between both wells. However, when for example V_a is decreased and V_b is increased until both differ considerably ($|V_\mathrm{a} - V_\mathrm{b}| \gg I$ and $V_\mathrm{a} < V_\mathrm{b}$), then is $A \gg B$ with $A/B \approx |V_\mathrm{a} - V_\mathrm{b}|/(2I)$.[31] In consequence, an electron in state ψ_1 is now much more likely to be found in the deeper V_a-well (where $|\psi_1|^2$ is high) than in the V_b-well. Since this is energetically beneficial, E_1 is now lower than in the balanced case before. The opposite effects result for the antisymmetric ψ_2: An electron in this state would be rather localized to the V_b-well (where $|\psi_2|^2$ is high), which is energetically less favorable such that E_2 is higher than before. In this way, imbalances in the potential landscape can cause localized states with a broader dispersion on the energy scale.

Imagine now a disordered solid as a three-dimensional collection of potential wells with fluctuating depths. The especially deep wells give rise to localized states with especially low energy, while the particularly shallow wells produce localized states with high energy. Those wells with rather average potential create non-localized states with intermediate energy.[47] This has also been shown mathematically by Anderson for a one-dimensional chain of wells with equal separation but random potential modifications.[55] The formation of localized states in amorphous solids is therefore called Anderson localization. Considering instead spatial disorder of equally deep potential wells leads to the same effects. Wells that are especially close together form energetically stable states with rather equal $|\psi|^2$ among them but decaying $|\psi|^2$ all around, while wells that are particularly far from all others give rise to localized states with higher energy.[31, 53] In an amorphous material with structural and energetic disorder it is therefore expected that each band exhibits band tails with localized states, often with an exponential shape of the DOS,[54, 56] while the states in the middle of the band remain extended.[31] The latter allow the formation and movement of wave packets, which differ from Bloch wave packets but lead to comparable band transport.[53, 54] Characteristic for extended states is therefore a non-vanishing spectral conductivity $\sigma(E)$ (cf. eq. 2.24) for $T \to 0\,\mathrm{K}$, while for localized states $\sigma(E) \to 0$ for $T \to 0\,\mathrm{K}$.[47, 53] There cannot be both extended and localized states at the same energy, because any localized states would become integrated into the band transport via the extended states.[31] Based on the above $\sigma(E)$-criterion, it is hence possible to determine the mobility edges, i.e., the energies that separate the regions of extended and localized

[19] A one-dimensional model system could be for example two Dirac delta potential wells, i.e., a potential $V(x) = V_\mathrm{a}\delta(x) + V_\mathrm{b}\delta(x - d)$ with V_a, $V_\mathrm{b} < 0$ and distance d between the wells.

states in the DOS. The conduction band mobility edge E_C separates the CB extended states from the localized CB tail that reaches into the band gap. Likewise, the valence band mobility edge E_V marks the border between VB extended states and the upper localized VB tail. The region between E_C and E_V is called the mobility gap, in analogy to the band gap between the conduction band edge and the valence band edge in crystalline semiconductors.

The localized states are not necessarily centered at specific atoms, since their shape depends on the nature and the number of the partaking atomic orbitals. In analogy to the well-known behavior of wave functions that penetrate the boundaries of finite potential wells, the localization is typically modeled by an exponentially decaying envelope in the outer regions of their stationary wave function: [31, 47, 54]

$$\psi_i(\vec{r}) \propto \exp\left(-\frac{|\vec{r}-\vec{r}_i|}{\alpha}\right) \tag{2.30}$$

Here, $\vec{r}_i$ is the center of the state and α is the localization radius that determines how fast ψ_i decays with distance. A similar exponential decay is also known for impurity states in doped semiconductors, which have the form of inflated spherical s-orbitals.[54] Close to the mobility edges (here for E_C), the localization radius approximately varies according to

$$\alpha(E) \approx \alpha_0 \left(\frac{E_C - E}{E_C}\right)^{2/3} \tag{2.31}$$

where α_0 is its value far from the mobility edge.[31] For $E \to E_C$ with $E < E_C$ it follows $\alpha(E) \to 0$, i.e., the eigenstates become extended when $E = E_C$. Near the mobility edges, the localized states therefore strongly overlap and electrons can transition between them easily as long as $T > 0\,\mathrm{K}$ (since localized states with equal energy must not overlap).[31] Far from the mobility edges, the localized states are much more confined and dilute, limiting either the conduction via extended states in the sense of traps or allowing transport via direct tunneling (hopping) between them.[54] These mechanisms will be discussed in the two following sections.

2.2.2 Mobility Edge Conduction

In amorphous semiconductors, the electron transport via extended states beyond the mobility edges is called mobility edge conduction (MEC). Here, just the n-type case shall be regarded,[20] where the conduction above E_C significantly dominates in comparison to the region below E_V. This typically involves that the Fermi level is closer to E_C than to E_V ($E_V \ll E_F < E_C$). Above E_C, the electrons move in the form of wave packets that consist of superposed extended eigenstates. Although they are not Bloch wave packets, their behavior is comparable to the previously described band transport in successively more disordered crystalline materials (cf. sec. 2.2.1). The wave packets with $E \gg E_C$

[20] The p-type conduction via VB extended states can be treated analogously.

are rather weakly affected by the amorphous structure and have an appreciable mean free path ℓ_m in between scattering (e.g., by phonons) or trapping in localized states. According to the Ehrenfest theorem,[21] they are accelerated by an electric field in a similar way as the Bloch wave packets in crystals. The gain of momentum within the time τ_m is balanced by the loss through scattering/-trapping.[44] For $E \gg E_C$, the Boltzmann transport theory (e.g., in [35, 50]) is therefore often a valid approximation and the mobility can be described by the same expression that is well-known for band transport in crystalline semiconductors: [31, 53, 54]

$$u = \frac{e\tau_m}{m^*} \tag{2.32}$$

where m^* is the effective mass of the electrons.[22] The wave packets close to the mobility edge ($E \gtrapprox E_C$) are much more affected by the disorder and can no longer be treated in a simple scattering model (ℓ_m reaches interatomic distances, cf. sec. 2.2.1). The eigenstates are becoming successively more distorted, up to the limit that the states at E_C have a strongly varying phase from atom to atom.[31] Since they still extend through the material, $\sigma(E_C) \neq 0$ for $T \to 0\,\mathrm{K}$ (as shown with the Kubo-Greenwood method in [31]), but wave packets that consist of such states have a drastically varying momentum in reaction to the disorder instead of a plain acceleration in an external electric field. The analytical treatment of this behavior quickly becomes complex (e.g., [54]), but still a rather general statement can be made. Under the assumption that $\sigma(E)$ does not vary strongly with E (in comparison to $f_{FD}(E)$) in the direct vicinity to E_C, the total MEC conductivity is mainly governed by how many electrons are excited to $E \geqslant E_C$. According to Mott,[31, 53] from eq. 2.24 then follows that the conductivity σ can be described by the MEC law

$$\sigma = -\int_{E_C}^{\infty} \sigma(E)\frac{\partial f_{FD}(E)}{\partial E}\,\mathrm{d}E \approx -\sigma_0 \int_{E_C}^{\infty} \frac{\partial f_{FD}(E)}{\partial E}\,\mathrm{d}E = \sigma_0 f_{FD}(E_C) \approx \sigma_0 \exp\left[-\frac{\Delta E}{k_B T}\right] \tag{2.33}$$

where $\sigma_0 \approx \sigma(E_C)$ and $\Delta E = E_C - E_F > 3k_B T$ in order to use the Boltzmann approximation. The MEC law also follows from $\sigma = e n_{eff} u$ and eq. 2.23 for n_{eff} under the same assumptions. Note that it is similar to the common expression for n-type conduction in crystalline non-degenerate semiconductors,[41, 49] albeit with the conduction band edge replaced by the conduction band mobility edge E_C. The presence of the localized states in the lower region of the conduction band restricts the band transport to $E \geqslant E_C$.

Since σ_0 generally has a much weaker dependence on T and E_F than the exponential term in eq. 2.33,[53] the MEC law captures quite accurately how σ varies with the temperature and the chemical potential of the electrons for this transport mechanism. Accordingly, a linear relation be-

[21] The Ehrenfest theorem says that $\partial\langle\vec{p}\rangle/\partial t = -\langle\psi|\nabla V|\psi\rangle$ for the momentum expectation value $\langle\vec{p}\rangle$ of a wave packet $\psi(\vec{r})$ – not necessarily of the Bloch-type – in a potential gradient $\nabla V(\vec{r})$.[48]

[22] The concept of effective mass, originally developed to treat electrons in crystals analog to free electrons, is also applicable to amorphous materials, as shown for example in [53]. Eq. 2.32 is only valid in this simple form if τ_m and m^* are approximately equal for all involved electronic states and if the external electric field is weak (σ independent of $\vec{F}$).

tween $\ln(\sigma)$ and T^{-1} is a typical experimental observation for amorphous semiconductors, just like for their crystalline counterparts.[31] The dependency on the material composition also approximately follows the same principle as in crystalline semiconductors. Doping or non-stoichiometry leads to defect states, typically in the mobility gap between E_C and E_V, together with an increase or decrease of the number of valence electrons that could potentially be excited to $E \geqslant E_C$. This results in a shift of E_F (and μ), which influences σ exponentially according to the MEC law.[10]

2.2.3 Variable-Range Hopping Conduction

Electron transport in the range of localized states occurs by hopping, i.e., tunneling, of the electrons between the localized states. This process is called variable-range hopping (VRH), because it is not restricted to nearest-neighbors.[23] The foundation for the theoretical treatment of VRH was developed by Miller and Abrahams, who derived a rate equation for electron hopping between localized impurity states in crystalline semiconductors.[57] Although they used concepts that are strictly valid for crystals only, an approximate version of their rate equation was later adopted by Mott and others to describe hopping in amorphous systems.[31, 58–60] This approach was exceptionally successful in explaining the deviations from the MEC law that were measured in amorphous materials, especially at low temperature. The essentials will be explained in the following.

Miller and Abrahams regarded the tunneling from an occupied localized state ψ_i to an unoccupied localized state ψ_j, each centered at an impurity in a crystal and defined according to eq. 2.30. The localization radius α is approximated as equal in the regarded energy range (i.e., far from the mobility edges, cf. α_0 in eq. 2.31). Following the definitions in sec. 2.1.1, the state energy difference E_{ij} in presence of a constant electric field $\vec{F}$ is

$$E_{ij} = E_j - E_i = \varepsilon_{ij} + e\vec{r}_{ij}\vec{F} \tag{2.34}$$

where $\vec{r}_{ij} = \vec{r}_j - \vec{r}_i$ is the vector from the center of ψ_i to the center of ψ_j and $\varepsilon_{ij} = \varepsilon_j - \varepsilon_i$ is the state energy difference without the contribution from the electric field. E_{ij} always has to be non-zero, because localized states cannot have partial overlap and equal energy (otherwise they would combine into two larger stationary states, see two-well example in sec. 2.2.1). Variable-range hopping therefore requires that the energy $|E_{ij}|$ is supplied ($E_{ij} > 0$) or received ($E_{ij} < 0$) by the vibrational motion of the atoms.[47] In the common picture of vibrational modes traveling through the material in the form of quasi-particles called phonons, the energy difference between the states is compensated by simultaneous absorption ($E_{ij} > 0$) or emission ($E_{ij} < 0$) of phonons. VRH is thus a phonon-assisted tunneling process. This also illustrates why for hopping conduction $\sigma(E) \to 0$ for $T \to 0\,\mathrm{K}$. In crystals, each phonon is characterized by a well-defined wave vector $\vec{q}$ as a quantum number. Although this is not the case in amorphous compounds, since the non-periodicity

[23] Nearest-neighbor hopping occurs only when the thermal energy $k_B T$ is larger than the energetic spread of all partaking states, such that the tunneling probability only depends on the distance between states (cf. eq. 2.46).[31, 54]

of the structure also prevents periodic vibrational modes,[53] for now the original route of Miller and Abrahams shall be followed. The adaptation of this model to amorphous materials, also with regard to phonons, will be discussed further below.

When ψ_i and ψ_j are close enough to allow tunneling, which means they partially overlap, they also "interact": Each state is influenced by the potential landscape around the other that is apparently also favorable to host an electron. Specifically, e.g., an electron in state i would also be partially delocalized to the potential well of the impurity j. The actual initial and final states of the tunneling process are therefore not the isolated localized wave functions ψ_i and ψ_j (defined by eq. 2.30) but perturbed states ψ_i' and ψ_j'. An important measure for this perturbation is the transfer integral $I_{ij} = \langle \psi_j | \hat{h} | \psi_i \rangle$. As long as the states have rather weak overlap ($|\vec{r}_{ij}| \gg \alpha$), it is approximately[24]

$$I_{ij} \propto \exp\left(-\frac{|\vec{r}_{ij}|}{\alpha}\right) \tag{2.35}$$

Additionally, for a quasi-random distribution of the localized states within the material, it is a fair assumption that for most pairs of rather close states $|E_{ij}| \gg k_\mathrm{B}T$. For weak overlap, i.e., when the concentration of localized states is rather low and E_i, E_j are far from the mobility edge, almost every pair has $|E_{ij}| \gg I_{ij}$. With this, the variational method yields the perturbed states [57, 61]

$$\psi_i' = \psi_i + \frac{I_{ij}}{E_{ij}}\psi_j \quad \text{and} \quad \psi_j' = \psi_j - \frac{I_{ij}}{E_{ij}}\psi_i \tag{2.36}$$

The states ψ_i' and ψ_j' are still well localized around $\vec{r}_i$ and $\vec{r}_j$, respectively, but have small contributions from the other state. Their energy separation still approximately equals E_{ij}. These are the actual states the tunneling happens between. The transition rate with simultaneous absorption or emission of a phonon with energy $|E_{ij}|$ is given by Fermi's golden rule [57, 61]

$$\nu_{ij} \propto \int |M_q|^2 \delta(\hbar c_\mathrm{s}|\vec{q}| - |E_{ij}|)\, \mathrm{d}\vec{q} \tag{2.37}$$

where $\vec{q}$ is the phonon wave vector, c_s is the speed of sound and M_q is the electron-phonon matrix element with (e.g., see [61] for details)

$$\begin{aligned} |M_q|^2 &\propto \langle \psi_j' | e^{i\vec{q}\vec{r}} | \psi_i' \rangle^2 f_\mathrm{BE}(|E_{ij}|) && \text{for absorption } (E_j > E_i) \\ |M_q|^2 &\propto \langle \psi_j' | e^{i\vec{q}\vec{r}} | \psi_i' \rangle^2 [f_\mathrm{BE}(|E_{ij}|) + 1] && \text{for emission } (E_j < E_i) \end{aligned} \tag{2.38}$$

Here, f_BE is the phonon distribution function, i.e., the Bose-Einstein distribution, since phonons

[24] Miller and Abrahams solved the transfer integral analytically for a Hamiltonian with two impurity potential wells and an effective potential term due to the surrounding. Assuming that the non-envelope part of ψ_i and ψ_j oscillates quickly w.r.t. $\vec{r}$, the integration just had to be done for the envelope functions and they arrived exactly at the here given relation for I_{ij}.[57, 61] The same relation is also well-known from the quantum mechanical description of the H_2^+ molecule. For more complex or more general systems, e.g., when $\vec{r}_i$ and $\vec{r}_j$ do not coincide with the position of simple potential wells, an analytical treatment is difficult. Still, all typical terms in an Hamiltonian have weaker $\vec{r}$-dependence than the exponentially decaying envelope functions. For weak overlap, any term in I_{ij} thus should be approximately proportional to $\exp(-|\vec{r}_{ij}|/\alpha)$.

are bosons:[25]

$$f_{\mathrm{BE}}(|E_{ij}|) = \left[\exp\left(\frac{|E_{ij}|}{k_{\mathrm{B}}T}\right) - 1\right]^{-1} \tag{2.39}$$

Eq. 2.37 integrates over all possible phonon states $\vec{q}$, with the Dirac delta selecting those that have suitable energy $\hbar c_{\mathrm{s}}|\vec{q}|$. The matrix element $|M_q|^2$ describes the transition probability w.r.t. to the respective phonon state. For $|E_{ij}| \gg k_{\mathrm{B}}T$, the Boltzmann approximation is applicable to f_{BE}:

$$f_{\mathrm{BE}}(|E_{ij}|) \approx \exp\left(-\frac{E_{ij}}{k_{\mathrm{B}}T}\right) \quad \text{for} \quad E_{ij} > 0 \quad \text{and} \quad f_{\mathrm{BE}}(|E_{ij}|) + 1 \approx 1 \quad \text{for} \quad E_{ij} < 0 \tag{2.40}$$

Inserting eq. 2.36 for ψ_i' and ψ_j' into the electron-phonon interaction term yields

$$\langle\psi_j'|e^{i\vec{q}\vec{r}}|\psi_i'\rangle = \langle\psi_j|e^{i\vec{q}\vec{r}}|\psi_i\rangle + \frac{I_{ij}}{E_{ij}}\langle\psi_j|e^{i\vec{q}\vec{r}}|\psi_j\rangle - \frac{I_{ij}}{E_{ij}}\langle\psi_i|e^{i\vec{q}\vec{r}}|\psi_i\rangle - \frac{I_{ij}^2}{E_{ij}^2}\langle\psi_i|e^{i\vec{q}\vec{r}}|\psi_j\rangle \tag{2.41}$$

Since ψ_i and ψ_j have weak overlap, the mixed integrals can be neglected.[61] With $\psi_j(\vec{r}-\vec{r}_{ij}) \approx \psi_i(\vec{r})$ follows

$$\langle\psi_j'|e^{i\vec{q}\vec{r}}|\psi_i'\rangle \approx \frac{I_{ij}}{E_{ij}}\left[\langle\psi_j|e^{i\vec{q}\vec{r}}|\psi_j\rangle - \langle\psi_i|e^{i\vec{q}\vec{r}}|\psi_i\rangle\right] \approx \frac{I_{ij}}{E_{ij}}\left(e^{i\vec{q}\vec{r}_{ij}} - 1\right)\langle\psi_i|e^{i\vec{q}\vec{r}}|\psi_i\rangle \tag{2.42}$$

The remaining integral $\langle\psi_i|e^{i\vec{q}\vec{r}}|\psi_i\rangle$ can be calculated analytically,[26] at least as long as the phonons have a well-defined wave vector (using $|\vec{q}| = |E_{ij}|/(\hbar c_{\mathrm{s}})$): [61]

$$\langle\psi_i|e^{i\vec{q}\vec{r}}|\psi_i\rangle = \left[1 + \left(\frac{\alpha E_{ij}}{2\hbar c_{\mathrm{s}}}\right)^2\right]^{-2} \tag{2.43}$$

The last term to examine is

$$\left(e^{i\vec{q}\vec{r}_{ij}} - 1\right)\left(e^{-i\vec{q}\vec{r}_{ij}} - 1\right) = 2\left(1 - \cos(\vec{q}\vec{r}_{ij})\right) \tag{2.44}$$

which results when inserting eq. 2.42 into the expression for $|M_q|^2$. This is the only term containing $\vec{q}$ that remains in eq. 2.37. When integrating this term over all possible $\vec{q}$ with given length $|\vec{q}| = |E_{ij}|(/\hbar c_{\mathrm{s}})$, the $\cos(\vec{q}\vec{r}_{ij})$ function oscillates quickly between +1 and −1 such that the areas below and above zero approximately cancel each other: [57, 61]

$$\int 2\left(1 - \cos(\vec{q}\vec{r}_{ij})\right)\delta(\hbar c_{\mathrm{s}}|\vec{q}| - |E_{ij}|)\,\mathrm{d}\vec{q} \approx 2\int \delta(\hbar c_{\mathrm{s}}|\vec{q}| - |E_{ij}|)\,\mathrm{d}\vec{q} \tag{2.45}$$

The integral over the δ-term equals the number density of phonons with energy $|E_{ij}|$. Inserting all discussed terms into Fermi's golden rule (eq. 2.37) shows that the hopping rate ν_{ij} is governed by

[25] The different terms f_{BE} vs. $f_{\mathrm{BE}} + 1$ in eq. 2.38 originate from the boson annihilation and creation operators (e.g., [48]).
[26] With the usual assumption that the non-envelope part of ψ_i oscillates quickly w.r.t. $\vec{r}$, just the integral of the squared envelope function times $\exp(i\vec{q}\vec{r})$ needs to be solved (e.g., [62]).

the exponential terms from I_{ij}^2 (eq. 2.35) and f_{BE} (eq. 2.40). Collecting all other weakly varying terms and here omitted constants in a comparably invariant attempt frequency ν_0 yields finally the Miller-Abrahams rate equation: [57]

$$\nu_{ij} = \begin{cases} \nu_0 \exp\left(-\frac{2|\vec{r}_{ij}|}{\alpha} - \frac{\varepsilon_{ij}+e\vec{r}_{ij}\vec{F}}{k_{\mathrm{B}}T}\right) & \text{for } \varepsilon_{ij} + e\vec{r}_{ij}\vec{F} > 0 \\ \nu_0 \exp\left(-\frac{2|\vec{r}_{ij}|}{\alpha}\right) & \text{for } \varepsilon_{ij} + e\vec{r}_{ij}\vec{F} \leq 0 \end{cases} \tag{2.46}$$

The transition rate for phonon-assisted tunneling thus depends on the spatial and energetic separation of the states in relation to the localization radius and the thermal energy, respectively. Note that the temperature T in the Miller-Abrahams rate equation refers to the vibrational motion of the atoms, since it stems from the phonon distribution f_{BE}. This phonon temperature is what is typically meant by the temperature of a material. Normally, the electron-phonon interaction keeps the electron temperature, which describes the average thermal energy of electrons, identical to the phonon temperature. It is however a known effect that for example high electric field strength or intense laser irradiation can raise the electron temperature significantly above the phonon temperature.[63–67] This is noteworthy, since in the previous thermodynamic description of electron transport (sec. 2.1) the temperature T always referred to the electron temperature (e.g., in f_{FD}). In cases where phonon temperature and electron temperature may differ, it is thus advised to denote the latter by an effective temperature T_{eff} of the electrons instead of T.

The above derivation for VRH between localized states in an otherwise crystalline structure has the major advantage that the whole model can be treated mathematically.[57, 61] In amorphous materials, this is not immediately possible because the electronic states and the phonon states are not as well-defined. For the electronic states, the non-envelope parts of ψ_i and ψ_j are unknown (cf. eq. 2.30), whereas they are Bloch wave packets in the Miller-Abrahams model. Still, it is plausible that also in amorphous solids these factors oscillate much faster w.r.t. $\vec{r}$ than the decay of the envelope functions with $|\vec{r} - \vec{r}_i|$. This justifies to just integrate over the well-defined envelope functions also in the amorphous case (eqs. 2.35, 2.36 and 2.43).[31] The uncertainty in the electronic wave functions therefore should not affect the applicability of the rate equation to amorphous solids. Concerning phonons, the effects of amorphicity are quite comparable to the situation for electrons. Due to the non-periodicity of the structure they can no longer be described by periodic waves with a defined wave vector $\vec{q}$, just like $\vec{k}$ for electrons.[53] Furthermore, three different types of vibrational modes occur in amorphous materials: While some remain phonon-like with rather ballistic movement through the material (called propagons), others are strongly disturbed by the disorder in their path through the material (called diffusons) and still others become localized, i.e., spatially confined, vibrations (called locons).[31, 68–71] In this work for simplicity the term "phonon" will still be used for all these more or less collective vibrations of atoms, but of course their definition (in real space; e.g., [53]) is bound to differ from the one used by Miller and Abrahams. The electron-phonon interaction operator $e^{i\vec{q}\vec{r}}$ in the above derivation thus must be different in amorphous solids (except approx. for propagons). The actual affected term is $\langle\psi_j|e^{i\vec{q}\vec{r}}|\psi_j\rangle - \langle\psi_i|e^{i\vec{q}\vec{r}}|\psi_i\rangle$

in eq. 2.42, which describes how the interaction of the same phonon with either ψ_i or ψ_j differs. As long as this difference is still governed by the phase shift of the phonon wave function from $\vec{r}_i$ to $\vec{r}_j$, i.e., as long as the amplitude of the phonon wave function does not vary exponentially with regard to energy or distance (at least in the typical range of $|E_{ij}|$ and $|\vec{r}_{ij}|$), the argumentation for eq. 2.45 should still be valid in principle. It is then unlikely that this term has a dominant influence on ν_{ij}. Since the phonons in amorphous solids still have a defined energy, in Fermi's golden rule (eq. 2.37) the selective meaning of the δ-term remains the same and the integral over $\vec{q}$ just becomes a sum over all phonons. Both I_{ij} and f_{BE} (eqs. 2.35 and 2.40), which are the origin of the exponentials in the rate equation, are unaffected by how the phonons are defined. As a result of these considerations, the Miller-Abrahams rate equation – in the here presented form with ν_0 not further specified – is generally deemed applicable to amorphous solids.[31, 54]

The individual tunneling rates ν_{ij} now allow to analyze the VRH conductivity of a whole network of localized states. Since VRH requires that the destination state is empty, it predominantly occurs via the states around the Fermi level (cf. eq. 2.23 for n_{eff}). When E_{F} is in a real band gap, e.g., below the CB tail, then hopping can only take place by the few electrons that are excited to the CB tail. The other extreme occurs when E_{F} is so far within the CB tail that it is close to the mobility edge. Then the localized states around the Fermi level are much more wide-spread (cf. sec. 2.2.1), hopping must be described differently (e.g., [31]) and VRH is hardly relevant due to excitation to the extended states. Instead of these two limiting cases, here the common situation of substantial density of localized states around the Fermi level (e.g., E_{F} in the CB tail) shall be discussed. In his according model for 3-dimensional VRH, Mott has used a DOS with $g(\varepsilon)$ = const. around the Fermi level. The current is thereby limited by the hops upwards in energy ($\varepsilon_{ij} + e\vec{r}_{ij}\vec{F} > 0$). Mott's reasoning was then that the flux is dominated by hops with an optimal "range", i.e., with an optimal combination of state distance and state energy difference.[58] In order to find these, he considered a sphere with radius r around a certain occupied state i. Within that sphere, the number of states per energy interval is $g(\mu)$ times the sphere volume. The inverse is thus the average energy separation $\langle\varepsilon_{ij}\rangle$ of state i to the next higher state j within that sphere (which is the most likely destination with $\varepsilon_j > \varepsilon_i$ since lower ε_{ij} is beneficial):[27]

$$\langle\varepsilon_{ij}\rangle = \frac{3}{4\pi r^3 g(\mu)} \tag{2.47}$$

The state j has a random position within the sphere, such that the average distance $\langle|\vec{r}_{ij}|\rangle$ and the average separation $\langle d_{ij}\rangle$ in direction of the electric field are (see sec. A.1)[28]

$$\langle|\vec{r}_{ij}|\rangle = \frac{3}{4}r \quad \text{and} \quad \langle d_{ij}\rangle = \frac{\langle\vec{r}_{ij}\vec{F}\rangle}{F} = \frac{3}{8}r \quad \text{for} \quad \vec{r}_{ij}\vec{F} > 0 \tag{2.48}$$

[27] Following Mott,[31] the optimal "range" for the situation of occupied i and unoccupied j is derived and thereafter multiplied with the number of electrons for which this condition is fulfilled (n_{eff}, cf. eq. 2.54).

[28] In the original treatment by Mott,[31] $\langle|\vec{r}_{ij}|\rangle F$ was used for the field term, which is corrected here by the average in direction of $\vec{F}$. This only influences σ_0 in the final expression.

In half of the cases an electron encounters the situation with $\langle \vec{r}_{ij}\vec{F}\rangle = +\langle d_{ij}\rangle F$, i.e., j lies in the direction of $\vec{F}$, and in the other half it is $\langle \vec{r}_{ij}\vec{F}\rangle = -\langle d_{ij}\rangle F$. From the Miller-Abrahams rate equation (eq. 2.46) therefore follows for the average net velocity $\langle v_\mathrm{d}\rangle$ of an electron in the direction of the electric potential gradient $-\vec{F} = \nabla\varphi$:

$$\begin{aligned}\langle v_\mathrm{d}\rangle &= -\frac{1}{2}\langle d_{ij}\rangle \nu_0 \exp\left(-\frac{2\langle|\vec{r}_{ij}|\rangle}{\alpha} - \frac{\langle\varepsilon_{ij}\rangle + e\langle d_{ij}\rangle F}{k_\mathrm{B}T}\right) \\ &\quad + \frac{1}{2}\langle d_{ij}\rangle \nu_0 \exp\left(-\frac{2\langle|\vec{r}_{ij}|\rangle}{\alpha} - \frac{\langle\varepsilon_{ij}\rangle - e\langle d_{ij}\rangle F}{k_\mathrm{B}T}\right) \\ &= \langle d_{ij}\rangle \nu_0 \exp\left(-\frac{2\langle|\vec{r}_{ij}|\rangle}{\alpha} - \frac{\langle\varepsilon_{ij}\rangle}{k_\mathrm{B}T}\right) \sinh\left(\frac{e\langle d_{ij}\rangle F}{k_\mathrm{B}T}\right)\end{aligned} \tag{2.49}$$

For weak electric fields where $e\langle d_{ij}\rangle F \ll k_\mathrm{B}T$, it is possible to approximate the sinh-term by its argument:

$$\langle v_\mathrm{d}\rangle = \nu_0 \frac{e\langle d_{ij}\rangle^2 F}{k_\mathrm{B}T} \exp\left(-\frac{2\langle|\vec{r}_{ij}|\rangle}{\alpha} - \frac{\langle\varepsilon_{ij}\rangle}{k_\mathrm{B}T}\right) \tag{2.50}$$

According to eqs. 2.47 and 2.48, $\langle\varepsilon_{ij}\rangle$ decreases with increasing r, while $\langle|\vec{r}_{ij}|\rangle$ behaves opposite. Therefore, $\langle v_\mathrm{d}\rangle$ must have a maximum at an optimal sphere radius r_opt. Since the exponential term dominates the value of $\langle v_\mathrm{d}\rangle$, the optimal radius can be found by

$$\begin{aligned}&\frac{\partial}{\partial r}\left(-\frac{2\langle|\vec{r}_{ij}|\rangle}{\alpha} - \frac{\langle\varepsilon_{ij}\rangle}{k_\mathrm{B}T}\right) = 0 \quad \Rightarrow \quad -\frac{3}{2\alpha} + \frac{9}{4\pi r_\mathrm{opt}^4 g(\mu) k_\mathrm{B}T} = 0 \\ &\Rightarrow \quad r_\mathrm{opt} = \left(\frac{3\alpha}{2\pi g(\mu) k_\mathrm{B}T}\right)^{1/4}\end{aligned} \tag{2.51}$$

Combining this result with eqs. 2.47, 2.48 and 2.50 yields

$$\begin{aligned}\langle v_\mathrm{d}\rangle &= \nu_0 \frac{eF}{k_\mathrm{B}T}\frac{9}{64}\left(\frac{3\alpha}{2\pi g(\mu)k_\mathrm{B}T}\right)^{1/2} \\ &\quad \exp\left[-\frac{3}{2\alpha}\left(\frac{3\alpha}{2\pi g(\mu)k_\mathrm{B}T}\right)^{1/4} - \frac{3}{4\pi g(\mu)k_\mathrm{B}T}\left(\frac{3\alpha}{2\pi g(\mu)k_\mathrm{B}T}\right)^{-3/4}\right] \\ &= \nu_0 \frac{eF}{k_\mathrm{B}T}\frac{9}{64}\left(\frac{3\alpha}{2\pi g(\mu)k_\mathrm{B}T}\right)^{1/2} \exp\left[-2\left(\frac{3}{2\pi\alpha^3 g(\mu)k_\mathrm{B}T}\right)^{1/4}\right]\end{aligned} \tag{2.52}$$

According to eq. 2.28, the mobility is given by $u = \langle v_\mathrm{d}\rangle/F$. The charge carrier density in the case of constant DOS follows from eq. 2.23 (see sec. A.2): [72]

$$n_\mathrm{eff} = g(\mu)\int_{-\infty}^{\infty} f_\mathrm{FD}(\varepsilon)\left[1 - f_\mathrm{FD}(\varepsilon)\right]\mathrm{d}\varepsilon = k_\mathrm{B}T g(\mu) \tag{2.53}$$

The classical relation $\sigma = e n_{\text{eff}} u$ (eq. 2.28) then finally leads to Mott's VRH law: [31, 58]

$$\sigma = \sigma_0 \exp\left[-\left(\frac{T_0}{T}\right)^{1/4}\right] \quad \text{with} \quad T_0 = \frac{24}{\pi k_\text{B} \alpha^3 g(\mu)}$$
$$\text{and} \quad \sigma_0 = \nu_0 e^2 \frac{9}{64}\left(\frac{3\alpha g(\mu)}{2\pi k_\text{B} T}\right)^{1/2} \tag{2.54}$$

Mott's approach was later criticized because the notion of all relevant hops occurring with optimal "range" is too simple. However, a more robust treatment with percolation theory then gave the same principal form of eq. 2.54, with T_0 just differing by a unitless factor.[59, 73] This is not surprising, since $2|\vec{r}_{ij}|/\alpha + E_{ij}/(k_\text{B}T) \propto (\alpha^3 g(\mu) k_\text{B} T)^{-1/4}$ for the critical subnetwork in the percolation treatment, just like for the optimal "range" of Mott. Hence, it is established that $T_0 \propto 1/(\alpha^3 g(\mu))$, but estimating the density of states near the Fermi level from eq. 2.54 was shown to yield largely incorrect values.[74]

Analog to the above VRH law for a constant DOS, other relations for σ have been derived for different DOS models. For a power-law DOS of the form $g(\varepsilon) \propto (\varepsilon - \mu)^n$, the resulting expression is $\sigma = \sigma_0 \exp[-(T_0/T)^{-(1+n)/(4+n)}]$.[73, 75] This yields $\sigma = \sigma_0 \exp[-(T_0/T)^{-1/2}]$ (Efros-Shklovskii hopping) for a parabolic DOS ($n = 2$), which is a common consequence of electron-electron repulsion at very low temperatures (Coulomb gap).[61, 76] For an exponential DOS of the type $g(\varepsilon) \propto \exp[(\varepsilon - \mu)/\varepsilon_0]$, which is often assumed for the CB tail,[77] the expected relation is $\sigma = \sigma_0 \exp[-\Delta\varepsilon/(k_\text{B}T)]$, like Mott's MEC law (eq. 2.33), with a weakly (logarithmically) T-dependent activation energy $\Delta\varepsilon$.[54] In general, Mott's VRH law for constant DOS becomes a good approximation at low temperatures – except for cases where a Coulomb gap emerges – since the relevant region around E_F shrinks with decreasing temperature (cf. eq. 2.23).

In most amorphous solids, both mobility edge conduction and variable-range hopping are possible to some extent. Comparing the MEC law (eq. 2.33) and the VRH law (eq. 2.54) shows that in an $\ln(\sigma)$ vs. T^{-1} plot the MEC conductivity would follow a straight line, while the VRH conductivity would be a convex function with decreasing slope. It is therefore typical that VRH dominates at low temperatures, where excitation to the mobility edge is negligible.[31] MEC prevails when T is high or when $E_\text{C} - E_\text{F}$ is small, since the electron mobility is commonly much higher in the extended states than for the hopping process between the localized states. The differing T-dependence of σ allows to identify experimentally which conduction mechanism dominates for which temperatures. Note also that the T-dependencies of MEC and VRH have a totally different origin: For MEC it comes from the Fermi-Dirac distribution (occupation probability), while for VRH it stems from the Miller-Abrahams hopping rate (tunneling probability). This also means that T in the MEC law refers to the electron temperature, while it corresponds to the phonon temperature in the VRH law. The relation to the material composition also differs fundamentally. Adding impurities (doping) or removing native atoms (non-stoichiometry) in principle changes the number of defect states and valence electrons in the system. As a result, the DOS – especially in the mobility gap – is modified

and the position of the Fermi level may vary (depending on eq. 2.18). For MEC the change of the DOS is rather irrelevant but any shift of the Fermi level affects the amount of charge carriers exponentially. For VRH the effect depends mostly on how much $g(\mu)$ differs due to the modified number of defect states near the Fermi level and due to the shift of μ (cf. eq. 2.54).

2.2.4 Schottky Barriers

So far, the discussion focused on MEC and VRH in the bulk of homogeneous amorphous materials, i.e., in regions with spatially constant chemical potential μ of the electrons. Now, the interfaces to metallic electrodes shall be regarded. In comparison to semiconductors, metals typically have a significantly higher work function W, which is the difference between the energy of a free electron at rest just outside the material (vacuum level) and the Fermi level at the surface of the material.[29] According to the definition of μ (eq. 2.2), this means that the chemical potential of the electrons is commonly lower in metals than in semiconductors ($\mu_\mathrm{m} < \mu_\mathrm{s}$), with the difference given by

$$e\varphi_\mathrm{bi} = \mu_\mathrm{s} - \mu_\mathrm{m} \approx -(W_\mathrm{s} - W_\mathrm{m}) = W_\mathrm{m} - W_\mathrm{s} \tag{2.55}$$

when surface effects can be approximately neglected. The quantity φ_bi (> 0) is called the built-in potential. Due to this chemical potential gradient, electrons diffuse from the semiconductor into the metal when the interface is formed. The resulting charge imbalance in turn gives rise to an electric field and a counter-acting drift flux, until a steady state with constant $\sigma\nabla\tilde{\mu}$ is reached (cf. eq. 2.14). In the metal, the excess electrons pile up at the interface because there is virtually no change of the chemical potential (since metals have a large DOS at the Fermi level). On the semiconductor side of the junction, the loss of electrons creates a positively charged depletion zone where the chemical potential is reduced in comparison to the bulk value μ_s. The size of this zone can be obtained by numerically solving the Poisson equation and the continuity equation together (eqs. 2.13 and 2.14), as it will be done in the later FEM simulations (sec. 4.3), but the essentials can already be illustrated by the following one-dimensional full depletion model. Therein, it is assumed that all mobile electrons up to a depletion width x_d flow into the metal, leading to a charge density

$$\rho(x) = \begin{cases} en_\mathrm{d} & \text{for } 0 \le x \le x_\mathrm{d} \\ 0 & \text{for } x > x_\mathrm{d} \end{cases} \tag{2.56}$$

where x denotes the distance from the interface and n_d is the density of depleteable electrons (e.g., the conduction band electron density).[30] Integrating the Poisson equation twice then yields (with

[29] The vacuum level refers to an electron with zero kinetic energy (i.e., at rest) which is far enough outside the material such that its potential energy just corresponds to the macroscopic $q\varphi$ near the material, but is no longer influenced by the material itself (i.e., a free electron). In materials with DOS at the Fermi level, the work function is the minimum energy (ionization energy) required to eject an electron from the surface Fermi level into vacuum (e.g., by photoemission).

[30] The full depletion model also assumes that the regarded interface dominates the device resistance by far, such that $\nabla\varphi(x) \approx 0$ (and $\nabla\mu(x) \approx 0$) for $x < 0$ and $x > x_\mathrm{d}$.

$\varphi(x \leq 0) = 0$ as arbitrary reference) [50]

$$\nabla^2\varphi = -\frac{en_\mathrm{d}}{\epsilon_0\epsilon_\mathrm{r}} \quad \Rightarrow \quad \varphi(x) = \frac{en_\mathrm{d}}{2\epsilon_0\epsilon_\mathrm{r}}\left[x_\mathrm{d}^2 - (x_\mathrm{d} - x)^2\right] \tag{2.57}$$

for $0 \leq x \leq x_\mathrm{d}$. The electric potential is thus a quadratic function that reaches the bulk value for $x = x_\mathrm{d}$. In the steady state, the difference between the Fermi levels of the semiconductor bulk and the metal is known to be

$$\tilde{\mu}_\mathrm{s} - \tilde{\mu}_\mathrm{m} = \mu_\mathrm{s} - \mu_\mathrm{m} - e(\varphi_\mathrm{s} - \varphi_\mathrm{m}) = eV \quad \Rightarrow \quad \varphi(x \geq x_\mathrm{d}) = \varphi_\mathrm{bi} - V \tag{2.58}$$

where V is an eventual voltage applied to the metal (relative to the semiconductor), $\varphi_\mathrm{m} = \varphi(x \leq 0) = 0$ is the electric potential in the metal and $\varphi_\mathrm{s} = \varphi(x \geq x_\mathrm{d})$ is the electric potential in the bulk of the semiconductor. With this boundary condition, the depletion width x_d can be obtained:

$$\varphi(x_\mathrm{d}) = \frac{en_\mathrm{d}}{2\epsilon_0\epsilon_\mathrm{r}}x_\mathrm{d}^2 = \varphi_\mathrm{bi} - V \quad \Rightarrow \quad x_\mathrm{d} = \left[\frac{2\epsilon_0\epsilon_\mathrm{r}}{en_\mathrm{d}}(\varphi_\mathrm{bi} - V)\right]^{1/2} \tag{2.59}$$

The larger the work function difference between the materials (and the larger ϵ_r of the semiconductor), the broader is the depletion zone. Increasing the voltage applied to the metal decreases x_d, while decreasing V broadens the depletion zone. According to the band bending model introduced in sec. 2.1.1, the varying electric potential modifies the energy of the electronic states in the semiconductor by $-e\varphi(x)$. In a crystalline n-type semiconductor, this means that in the depletion zone the conduction band edge is bent upwards for $x \to 0$, thus imposing an obstacle for electron flow that is called Schottky barrier.[31] Fig. 2.1 illustrates this situation. For $V > 0$ (forward bias), the electrons flow from the semiconductor to the metal, thereby encountering a potential barrier with height $\varphi_\mathrm{bi} - V$. The equivalent for $V < 0$ (reverse bias) is the Schottky barrier height φ_B, which according to the Schottky-Mott rule is given by [78]

$$\varphi_\mathrm{B} = \varphi_\mathrm{bi} + \varphi_\mathrm{n} \tag{2.60}$$

where φ_n is the difference between the semiconductor's Fermi level and CB edge in the bulk ($\varphi_\mathrm{n} > 0$ in the non-degenerate case). In a crystal, there are no (or few) defect states in the band gap, such that the electrons need to cross the Schottky barrier by excitation over the barrier, called thermionic emission, or by tunneling through the barrier.[32] The tunneling can take place either at the lowest possible energy level, then called field emission, or thermally activated at a higher energy where the barrier is thinner, which is called thermionic field emission (cf. fig. 2.1).[41] As mentioned above, thermionic emission benefits from a decreasing barrier $\varphi_\mathrm{bi} - V$ for increasing forward bias, but faces

[31] For extended states, the band bending is just a picture to describe that each wave function with a certain energy below the tip of the Schottky barrier begins to decay exponentially when it encounters the potential barrier $q\varphi(x)$ at a certain distance away from the interface. Quantum mechanically, individual stationary states cannot "bend" as they have a unique, not a spatially varying eigenvalue.

[32] There are further conduction processes through a Schottky barrier like for example electron-hole recombination, which are not discussed here (e.g., [41, 78]).

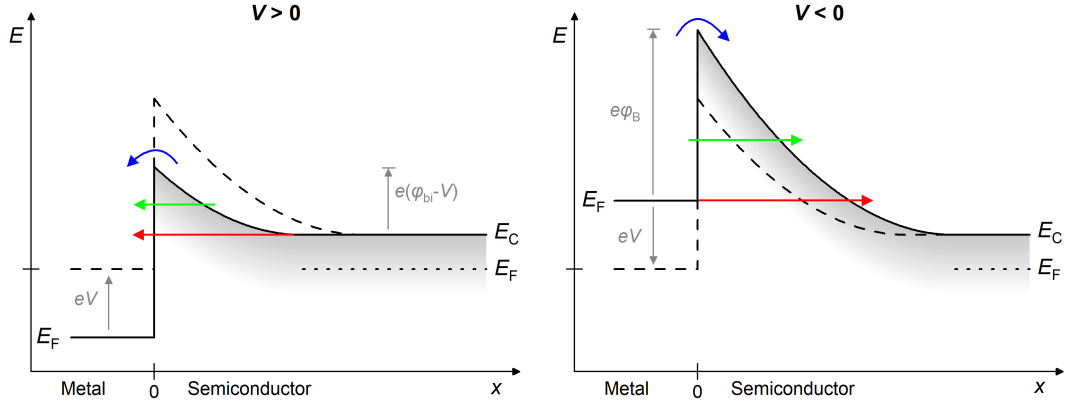


Figure 2.1: Schematic band diagrams for a Schottky barrier at a metal-semiconductor interface, with either positive voltage (left) or negative voltage (right) applied to the metal. For a crystalline n-type semiconductor, the solid line represents the conduction band edge and the shaded area can be ignored. For an amorphous n-type semiconductor, the solid line is the CB mobility edge E_C and the shaded area indicates the localized tail states. The dashed line shows the situation for $V = 0$ and the dotted line is the Fermi level of the semiconductor. The colored arrows symbolize thermionic emission (blue), thermionic field emission (green) and field emission (red).

an (approximately) constant barrier φ_B for reverse bias. This leads to the classical I-V relation of a diode with exponentially increasing current for $V > 0$ and a minor saturation current for $V < 0$: [41, 78]

$$\text{Thermionic emission:} \quad I \propto \exp\left(-\frac{e\varphi_B}{k_B T}\right)\left[\exp\left(\frac{eV}{\eta k_B T}\right) - 1\right] \tag{2.61}$$

The ideality factor η, which equals unity for pure thermionic emission, is typically introduced to account for minor deviations for example due to additional tunneling. Eq. 2.61 shows that a metal-semiconductor interface may represent a directional resistance which allows current for forward bias but almost blocks it in the opposite direction. A reverse-biased Schottky barrier is thus often the dominating resistance in electronic devices. The Schottky barrier height typically has a weak V-dependence due to image force lowering (Schottky effect),[33] such that

$$\varphi_B = \varphi_{B,0} - \left[\frac{e^3 n_d}{8\pi^2 \epsilon_0^3 \epsilon_r^3}(\varphi_{bi} - V)\right]^{1/4} \tag{2.62}$$

where $\varphi_{B,0}$ is the original barrier height from eq. 2.60.[41, 79] This effect is generally negligible for forward bias (where the lowering decreases with increasing V), due to the stronger V-dependence in eq. 2.61. For thermionic emission at $V > 3k_B T/e$ (and $\eta = 1$), $\ln(I)$ therefore increases linearly with V, until at $V > \varphi_{bi}$ there is no barrier anymore and the interface becomes ohmic ($I \propto V$). For reverse bias, the Schottky effect prevents the complete saturation that eq. 2.61 alone would indicate

[33] Image force barrier lowering: Each electron that crosses the Schottky barrier in either direction repels the electrons that are piled up at the metal surface, when it is close to the metal (i.e., near the tip of the Schottky barrier). The electrons in the metal are slightly pushed away from the interface, which lowers the electric field experienced by the nearby electron as if there would be a positive image charge in the metal.[41]

and leads to an exponentially increasing current also for this direction, albeit with a much weaker V-dependence. For $V < -3k_\mathrm{B}T/e$, a linear relation between $\ln(|I|)$ and $(\varphi_\mathrm{bi} - V)^{1/4}$ is therefore expected for thermionic emission.

As indicated in fig. 2.1, the tunneling width through the Schottky barrier decreases both for positive and for negative voltage. Field emission and thermionic field emission therefore have more symmetric I-V relations: [41]

$$\text{Field emission } (V \gg 0): \quad I \propto \exp\left[-\frac{e(\varphi_\mathrm{B} - V)}{E_{00}}\right] \quad \text{with} \quad E_{00} = \frac{e\hbar}{2}\left(\frac{n_\mathrm{d}}{\epsilon_0 \epsilon_\mathrm{r} m^*}\right)^{1/2} \tag{2.63}$$

$$\text{Field emission } (V \ll 0): \quad I \propto \exp\left[-\frac{2e\varphi_\mathrm{B}^{3/2}}{3E_{00}(\varphi_\mathrm{B} - V)^{1/2}}\right] \tag{2.64}$$

$$\text{Thermionic field emission } (V \gg 0): \quad I \propto \exp\left[-\frac{e\varphi_\mathrm{n}}{k_\mathrm{B}T} - \frac{e(\varphi_\mathrm{B} - \varphi_\mathrm{n} - V)}{E'_{00}}\right] \tag{2.65}$$

$$\text{Thermionic field emission } (V \ll 0): \quad I \propto \exp\left[-\frac{e\varphi_\mathrm{B}}{E'_{00}} - \frac{eV}{E''_{00}}\right] \tag{2.66}$$

$$\text{with} \quad E'_{00} = E_{00}\coth\left(\frac{E_{00}}{k_\mathrm{B}T}\right) \quad \text{and} \quad E''_{00} = E_{00}\left[\frac{E_{00}}{k_\mathrm{B}T} - \tanh\left(\frac{E_{00}}{k_\mathrm{B}T}\right)\right]^{-1}$$

The more tunneling occurs, the less diode-like is the resistance of a Schottky barrier. As a rough criterion serves the quantity E_{00}: Thermionic emission dominates for $E_{00} \ll k_\mathrm{B}T$, thermionic field emission for $E_{00} \approx k_\mathrm{B}T$ and the almost T-independent field emission for $E_{00} \gg k_\mathrm{B}T$.[41, 78]

In amorphous n-type semiconductors, the CB mobility edge E_C essentially takes the role of the CB edge in the treatment of Schottky barriers, as it marks the onset of band transport (fig. 2.1). The Schottky barrier height φ_B then refers to the difference between E_C and E_F at the interface (eq. 2.60), while $e\varphi_\mathrm{n}$ is defined as $E_\mathrm{C} - E_\mathrm{F}$ in the bulk of the amorphous semiconductor. In this regard, the picture is analog to the crystalline case discussed before, but now there are localized states below E_C that partially (or even completely) fill the formerly state-free Schottky barrier region (fig. 2.1). The aforementioned model equations for thermionic emission, thermionic field emission and field emission (eqs. 2.61 to 2.66)[34] – originally derived for crystalline n-type semiconductors – therefore have only limited applicability for amorphous solids. Even when MEC dominates in the bulk of an amorphous semiconductor, VRH could play a vital role at Schottky barriers where both conduction mechanisms are affected to a different degree. Hopping transport below the Schottky barrier tip could for example significantly reduce the necessary excitation for thermionic emission. Localized states in the regions of field emission or thermionic field emission present an efficient way to split the tunneling process into multiple shorter hops with higher tunneling probability (cf. fig. 2.1). Furthermore, the full depletion approximation becomes problematic when there is substantial DOS around the Fermi level, since in that case it is not reasonable that ρ has a constant value in the depletion zone. These possibilities are often ignored in literature and the above models are used

[34] Note that the here given I-V relations are popular but not unique. Depending on the utilized approximations and included effects, for each of these emission currents also other model equations have been proposed (e.g., [79, 80]).

for amorphous Schottky barriers with the implicit assumption that the density of localized states is too low to have any impact.[9, 26, 27, 81–83] This is rarely mentioned or explicitly justified. In materials where VRH dominates in the bulk, it is even thinkable that electrons cross the Schottky barrier via localized states close to E_F, i.e., by tunneling (hopping) at an even lower energy level than field emission. The later FEM simulations will show that this is indeed a viable option (sec. 5.6).

2.3 Gallium Oxide

Now that the major theoretical foundations have been introduced, it is time to move on to the material under study. Three varieties of gallium oxide are of concern for this work: crystalline β-Ga_2O_3 as quite well-understood starting material, amorphous, stoichiometric a-$GaO_{1.5}$, which reflects the sole influence of amorphicity, and of course amorphous, non-stoichiometric a-GaO_x ($x < 1.5$) – the investigated material.

β-Ga_2O_3 is the thermodynamically stable phase of Ga_2O_3 on the whole temperature range up to its melting temperature of about 1800 °C.[11] It has a monoclinic structure with 8 gallium atoms and 12 oxygen atoms in the unit cell.[84] The Ga atoms are tetrahedrally and octahedrally coordinated by oxygen, while the O atoms have trigonal and tetrahedral coordination by gallium. In these polyhedra, the Ga-O nearest-neighbor distance varies between 1.83 and 2.07 Å.

The amorphous varieties a-$GaO_{1.5}$ and a-GaO_x are metastable materials that are commonly prepared in thin film form either by radio-frequency (RF) magnetron sputtering [18, 24, 85–90], pulsed laser deposition (PLD) [9, 10, 26, 27, 32, 91] or atomic layer deposition (ALD) [92–94] at low temperature.[35] PLD and sputtering typically use β-Ga_2O_3 targets as source material, from which gallium oxide fragments are ejected and deposited on a substrate through irradiation with a laser beam or accelerated ions, respectively. In contrast, ALD employes molecular gallium and oxygen precursors (e.g., trimethylgallium and O_2 plasma), which sequentially react with the surface of a substrate, thereby forming gallium oxide layer by layer.

For the PLD method, which is used in this work, the decisive process parameters to obtain a-$GaO_{1.5}$ or a-GaO_x are the substrate temperature, the pressure and the gas composition in the deposition chamber. Even though X-ray amorphous gallium oxide films have been reported for PLD at room temperature [10, 33, 96, 97], 300 °C [98], 400 °C [91, 96], 450 °C [9], 500 °C [26, 27, 99, 100] and even up to 600 °C [32], the actual temperature limit for permanently and fully amorphous gallium oxide seems to be at around 300 °C. During post-deposition annealing in inert atmosphere, as-prepared X-ray amorphous films developed the first crystalline features in X-ray diffraction (XRD) and transmission electron microscopy (TEM) already at 330–350 °C, even when they were

[35] According to [95], sputtering, PLD and ALD are the predominant preparation methods for a-GaO_x – in that order – in the literature up to the year 2021. See also [95] for a detailed overview over these and other preparation methods for a-GaO_x.

deposited at much higher temperature.[32, 33, 96] Films that were deposited at room temperature stayed fully XRD- and TEM-amorphous during up to 40 days annealing at 300–320 °C.[33, 97] Whether stoichiometric a-$GaO_{1.5}$ or non-stoichiometric a-GaO_x is obtained in PLD depends mainly on the atmosphere in the deposition chamber. Pure Ar gas always leads to oxygen-deficient a-GaO_x, regardless of the pressure.[10, 32, 96, 97] In air, pure O_2 or Ar/O_2 mixtures, a-GaO_x results just when the pressure is especially low [9, 26, 27, 91], while oxygen partial pressures (p_{O_2}) of 0.04–0.008 mbar lead to a-$GaO_{1.5}$ films.[32, 33, 96, 97] The underlying principles for the non-stoichiometric material transfer from the Ga_2O_3 target to the substrate have not been fully clarified, but it is common belief that scattering between the ablated species and the gas plays a major role.[101] Scattering events have a weaker effect on the trajectories of heavier particles (e.g., Ga), while lighter species (e.g., O) are easily deflected away from the substrate.[28] In presence of O_2, the oxygen loss through scattering is compensated by incorporation of oxygen from the gas phase into the thin film.

A fine-grained control of the non-stoichiometry x of a-GaO_x through PLD conditions has proven difficult so far. In [102], it was reported that x may be varied rather smoothly between $x = 1.5$ and $x = 1.0$ by decreasing the oxygen partial pressure from 10^{-1} to 10^{-7} mbar and increasing the laser fluence, i.e., the pulse energy per illuminated target area. They however found that most of their films consisted of a stoichiometric and an oxygen-deficient phase, probably due to the deposition at 400–500 °C, and that the variation of x was just a changing ratio between the two phases. Likewise, in annealing studies of a-GaO_x at 480 °C the disproportionation into a β-Ga_2O_3 phase and an almost pure Ga phase was observed, with an average $x = 1.2$.[96] This was also deemed responsible for a metal-insulator transition observed in a-$GaO_{1.2}$ at 390 °C.[32, 96, 103] Such a phase separation could also be involved in the apparent trend that PLD in Ar or vacuum at high temperature (400–600 °C) results in $x \approx 1.2$–1.4, while room temperature (RT) leads to stoichiometries around $x \approx 1.0$.[26, 27, 91, 96] These films with $x \approx 1.2$–1.4 might not have been homogeneous. Instead, PLD in low-pressure O_2 seems to give access to these intermediate oxygen deficiencies. With $p_{O_2} \approx 10^{-5}$ mbar, $x \approx 1.4$ has been obtained at 300 °C [104], while $x \approx 1.3$ has resulted at 100–160 °C and $p_{O_2} \approx 10^{-3}$ mbar.[105] Based on the aforementioned annealing experiments in [96], a phase separation is less probable at these temperatures (even though it has not been ruled out). PLD-made films with amorphicity and homogeneity explicitly verified by TEM have been reported for $x \approx 0.8$–1.1 (deposited at RT in Ar, 0.04 mbar) and for $x = 1.5$ (deposited at RT in O_2, 0.04 mbar).[10, 33, 96, 97] These works also hint at another trend: For an equal gas atmosphere (Ar, 0.04 mbar) and temperature (RT), $x \approx 0.8$ has resulted with a laser energy setting of 200 mJ/pulse [97], while 300 mJ/pulse has led to $x \approx 1.0$–1.1.[10, 96] However, this cannot be assessed with certainty because the actual measure for the laser energy delivered to the PLD target, the laser fluence that takes beamline losses and beam diameter into account, has not been reported.

The structures of a-$GaO_{1.5}$ and a-GaO_x deposited at RT have been investigated by extended X-ray absorption fine structure (EXAFS) measurements.[96, 106] According to the radial distribution functions (RDF) extracted from the Ga K-edge signal, these materials still have short-range or-

der, i.e., a rather defined first coordination shell around Ga, but no further ordering beyond that. The corresponding RDF peak at about 1.7–1.9 Å had the same position but significantly increased amplitude in the amorphous films in comparison to crystalline reference films, which indicated a higher average coordination number of Ga in a-$GaO_{1.5}$ and a-GaO_x than in β-Ga_2O_3.[96, 106] It was found that the EXAFS at the Ga K-edge of a-$GaO_{1.5}$ could be modeled with the octahedral and tetrahedral coordination polyeder known from β-Ga_2O_3, albeit with slightly different Ga-O distances and a substantially higher proportion of octahedra (50% octahedra and 50% tetrahedra in β-Ga_2O_3, 57% octahedra and 43% tetrahedra in a-$GaO_{1.5}$).[96] Since it is questionable whether the ideal polyhedra employed in the modeling can really be present in an amorphous compound, considering that even those in β-Ga_2O_3 are distorted, this result has to be interpreted rather qualitatively: The polyhedra present in β-Ga_2O_3 still seem to be a rather appropriate description of the immediate environment around Ga atoms in a-$GaO_{1.5}$, but with a higher proportion of distorted 6-fold coordination enabled by the flexibility of the amorphous structure (less dense packing). This is in line with the commonly lower density of a-$GaO_{1.5}$ (5.1–6.0 g/cm^3 [106–111]) in comparison to β-Ga_2O_3 (6.0–6.3 g/cm^3 [11, 106]). The density is typically even lower for a-GaO_x (4.5–5.8 g/cm^3 [92, 107]), which is plausible for a lack of oxygen.[36] For a-GaO_x with $x \approx 1.0$, the EXAFS-RDF amplitude was lower than for a-$GaO_{1.5}$, corresponding to lower average coordination number of Ga, as anticipated for oxygen-deficiency.[96] The effect was however weaker than expected for 1/3 of oxygen missing in the coordination polyhedra, which led to the conclusion that either the octahedra/tetrahedra ratio or the average polyhedra shape differs significantly from a-$GaO_{1.5}$.[37] A rearrangement of the local structure around Ga is highly probable, considering that at such high oxygen deficit on average 2 of 6 and more than 1 of 4 oxygen atoms would be missing in octahedra and tetrahedra, respectively.

This not clarified question of the short-range order in a-GaO_x immediately translates into difficulties in describing the defect chemistry. In pure β-Ga_2O_3, the intrinsic majority defects in the p_{O_2} range from 10^{-17} to 1 bar are oxygen vacancies $v_O^{\bullet\bullet}$ or $v_O^{\times}$, with relative charge +2 or 0 depending on the p_{O_2}.[112–114] The respective defect equilibria can be expressed in Kröger-Vink notation as

$$O_O^{\times} \rightleftharpoons v_O^{\bullet\bullet} + 2\,e' + \frac{1}{2}\,O_2(g) \quad \text{and} \quad O_O^{\times} \rightleftharpoons v_O^{\times} + \frac{1}{2}\,O_2(g) \tag{2.67}$$

where $v_O^{\times}$ is an oxygen vacancy with two bound electrons (in the Ga dangling bonds), $v_O^{\bullet\bullet}$ is an oxygen vacancy without bound electrons and e' is an excess electron excited to the CB. In an amorphous structure, there are no crystallographically equivalent positions and therefore vacancies are – in principle – not a valid concept. Nevertheless, one could argue that a vacancy picture is still appropriate when there would be rather equally shaped Ga coordination polyhedra which either

[36] Note that most of the here mentioned densities of a-$GaO_{1.5}$ and a-GaO_x refer to films prepared by other methods than PLD (mostly by ALD). Just [106] contains a value for PLD-made a-$GaO_{1.5}$: 6.01 g/cm^3.

[37] Note that an EXAFS analysis for a-$GaO_{1.2}$ (deposited by PLD at 500–600 °C) was reported in [32, 103]. There, an equal octahedra/tetrahedra ratio was obtained as stated for a-$GaO_{1.5}$ in [96]. The RDF of a-$GaO_{1.2}$ however contained a second peak at about 2.8 Å, just like β-Ga_2O_3, which makes an interpretation difficult (see the above discussion of PLD at high temperature).

host an oxygen ion or not. As discussed above, in almost stoichiometric a-GaO_x this might actually be the case, but for high oxygen-deficiency a short-range order with distinguished "empty" positions is unlikely. The Kröger-Vink notation is thus inappropriate for a-GaO_x in general and the oxygen-deficiency defect equilibrium just reads

$$O^{2-}(s) \rightleftharpoons 2\,e^- + \frac{1}{2}\,O_2(g) \tag{2.68}$$

In X-ray photoelectron spectroscopy (XPS) of a-GaO_x, it has been observed that the excess electrons e^- predominantly increase the electron density around Ga^{3+} ions (electron configuration: [Ar] $3d^{10}$), formally leading to Ga^+ valence defects (electron configuration: [Ar] $3d^{10}$ $4s^2$): [10, 92, 115]

$$Ga^{3+}(s) + 2\,e^- \rightleftharpoons Ga^+(s) \tag{2.69}$$

Instead of the oxygen vacancy picture, the effects of oxygen-deficiency in a-GaO_x thus may be roughly described as a reduced average coordination number of the Ga ions and the presence of Ga^+ valence defects.

Naturally, the structural and compositional differences between β-Ga_2O_3, a-$GaO_{1.5}$ and a-GaO_x also have consequences for the material properties, e.g., for the band gap. β-Ga_2O_3 has an indirect band gap of about 4.85 eV and a direct band gap of 4.9 eV, as determined congruently by ab-initio band-structure calculations, UV-Vis spectroscopy and angle-resolved photoemission spectroscopy.[11, 14, 116, 117] It is commonly considered a direct-band gap material due to the ten times higher probability for this transition.[14] This large band gap is an outstanding characteristic of β-Ga_2O_3 and led to its use in high power diodes and transistors, as well as in solar-blind UV photodetectors.[12, 13, 15, 16] For a-$GaO_{1.5}$, an almost equal direct band gap of 4.8–5.0 eV has been measured via spectroscopic ellipsometry and UV-Vis spectroscopy in the vast majority of works.[9, 85, 109, 110, 118–122] Therefore, it is often used for the same applications as β-Ga_2O_3, with the benefit of preparation at low T and on flexible substrates in exchange for lower thermal stability.[86, 87, 89, 95, 123–125] Only few studies report a significantly lower band gap in the range of 4.3–4.5 eV for a-$GaO_{1.5}$, probably as a result of unintentional doping or oxygen-deficiency.[105, 107, 117] Such narrow band gaps have been consistently found for non-stoichiometric a-GaO_x. Direct band gaps of 4.4–4.9 eV for $x \approx 1.4$ [104, 126, 127] and of 4.0–4.4 eV for $x \approx 1.3$ [26, 105, 126] have been measured by UV-Vis spectroscopy. The following discussion of the DOS will show that this clear trend of narrower band gap for lower x originates from defect states that emerge in the band gap.

2.3.1 Density of States

Fig. 2.2 shows different densities of electronic states (DOS) for β-Ga_2O_3, a-$GaO_{1.5}$ and a-GaO_x, which have been calculated by hybrid DFT.[89, 128] Thereby, the required amorphous structures have been created from crystalline supercells through an artificial melt-quench procedure using

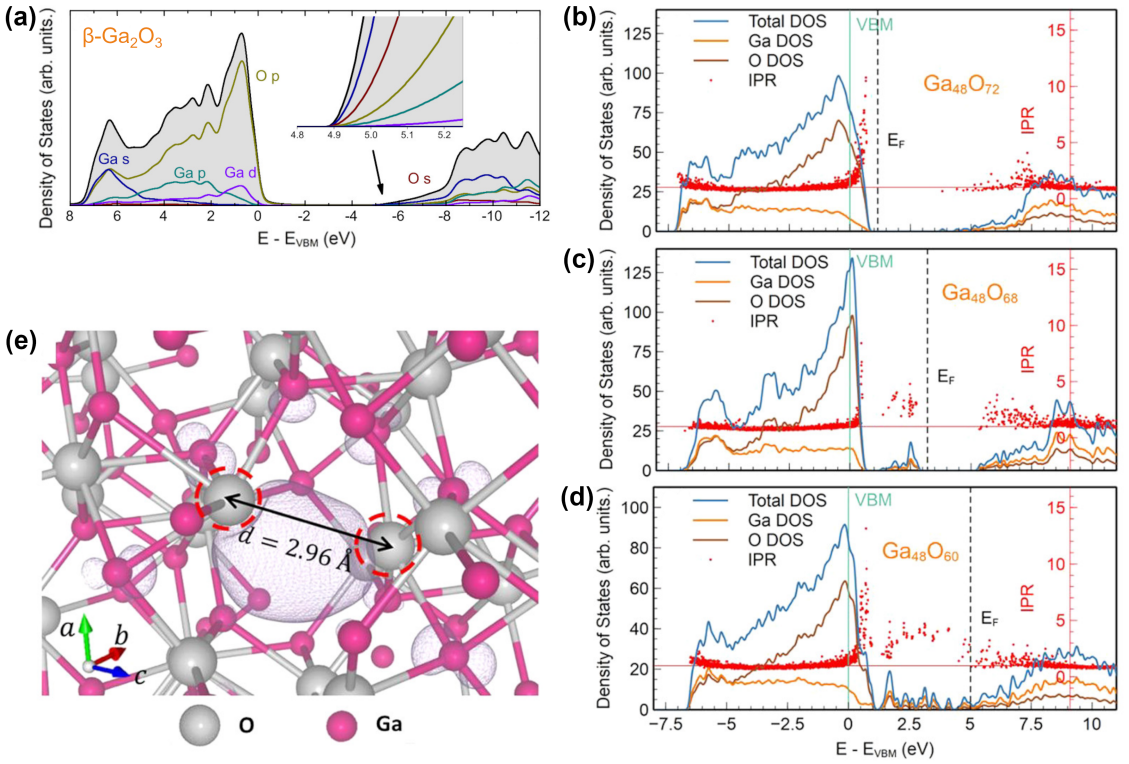


Figure 2.2: Electronic states in crystalline and amorphous gallium oxide. (a) DOS of β-Ga_2O_3, calculated by hybrid DFT (HSE06 functional with 32% Hartree-Fock exchange).[128] (b)–(d) DOS of a-$Ga_{48}O_{72}$ (x = 1.50), a-$Ga_{48}O_{68}$ (x = 1.42) and a-$Ga_{48}O_{60}$ (x = 1.25), calculated by hybrid DFT (HSE06 functional with 25% Hartree-Fock exchange) for amorphous structures generated by AIMD.[89] In (a), VBM marks the VB edge, while in (b)–(d) it refers to the VB mobility edge as inferred from the inverse participation ratio (IPR). (e) Partial charge density of the highest electronic state below E_F in a-$Ga_{48}O_{68}$, shown for an isovalue of −0.008 e/Bohr^3.[89] Reprinted with permission from [128] and [89]. © 2019 American Chemical Society.

ab-initio molecular dynamics (AIMD).[38] The DOS of β-Ga_2O_3 (fig. 2.2a) features a steep VB edge, a more gradually rising conduction band and a broad band gap in between that is in accordance with the values discussed above. Comparison with the DOS of a-$GaO_{1.5}$ (fig. 2.2b) shows that the amorphicity has almost no impact on the shape of the DOS, which is in line with the approximately equal band gaps reported for β-Ga_2O_3 and a-$GaO_{1.5}$ (cf. sec. 2.3). However, the states near the VB and CB edges in a-$GaO_{1.5}$ are localized, as has been concluded from their inverse participation ratio (IPR) that is also shown in fig. 2.2b. The IPR is roughly the inverse of the number of atoms that contribute to a certain electronic state, such that extended states have low IPR while localized states have high IPR.[39] The DOS of a-$GaO_{1.5}$ thus exhibits the localized VB and CB tails that are expected in amorphous compounds (cf. sec. 2.2.1). These are also present in the DOS of non-stoichiometric a-GaO_x (figs. 2.2c,d), but with increasing oxygen deficit also more and more defect states emerge in the band gap. According to their IPR, these defect states are localized as well. Other published

[38] AIMD procedure in [89]: Each structure was subjected to 10 sequential 3 ps-AIMD runs at 10 different temperatures from 2000 K down to 500 K in equal steps (NVT ensemble, time step: 1.5 fs), followed by a full structure optimization until all forces were lower than 0.05 eV/Å.

[39] There exist different definitions of the IPR, e.g., in [69, 129, 130], which differ in the calculation but not in the meaning of the quantity. It is not specified in [89] which definition was used for figs. 2.2b–d.

DFT-DOS for a-$GaO_{1.5}$ [106, 126] and a-GaO_x [126] show the same trends, but lack an analysis of state localization.

For x = 1.42 (fig. 2.2c), the defect states appear as a rather broad defect band in the lower half of the band gap. This not only explains the experimentally determined shrinking of the band gap (cf. sec. 2.3), but also fits to the observation of defect states near the VB edge in photoemission spectra of a-GaO_x (with x = 1.4 in [104]), which have been measured by hard X-ray photoelectron spectroscopy (HAXPES).[104, 123] They are not present in XPS spectra of a-$GaO_{1.5}$.[106] HAXPES and DFT agree further in that the defect states are predominantly occupied by the excess electrons from oxygen-deficiency (cf. eq. 2.68), which results in a shift of the Fermi level to higher energy for a-GaO_x in comparison to a-$GaO_{1.5}$ (figs. 2.2b,c). The defect band in fig. 2.2c is located at significantly lower energy than the electronic defects from oxygen vacancies and gallium interstitials in β-Ga_2O_3, which lie 2.2–4.0 eV and 3.5–4.0 eV above the VB edge, respectively, according to hybrid DFT calculations.[112, 114, 131] This is another sign that the situation in a-GaO_x differs from a mere accumulation of β-Ga_2O_3-like point defects.

The DOS for x = 1.25 (fig. 2.2d) indicates that for decreasing x the band gap is eventually filled by localized states. It is however difficult to assess whether this condition is already reached for x = 1.25, because in these DFT calculations only few defect states in the band gap have actually been calculated (as indicated by the rather limited number of red IPR dots in this region). This is a result of the low number of just 8 k-points (2 × 2 × 2) that have been used due to the high computational demand of hybrid DFT.[89] Estimation of the CB mobility edges E_C from figs. 2.2b–d is difficult for the same reason, but it seems like $E_C - E_F$, i.e., the governing quantity for mobility edge conduction (cf. sec. 2.2.2), decreases substantially with decreasing x. Simultaneously, there are more and more localized states close to E_F, up to the point where E_F is in the localized CB tail in fig. 2.2d. This means that with decreasing x variable-range hopping also becomes a viable option (cf. sec. 2.2.3). Whether MEC or VRH dominates in a-GaO_x is not clear so far, as discussed in the next section.

Fig. 2.2e shows the partial charge density of an exemplary defect state from the DFT-DOS in fig. 2.2c, which confirms that these states are indeed localized as indicated by the IPR. The shown state seems to be localized to a void in the amorphous structure, which illustrates that localized states are not necessarily centered on a specific atom. The void appears to be coordinated by O and Ga atoms, which fits to the rather comparable contributions from O and Ga orbitals to the defect states in the site-projected DOS (fig. 2.2c). This is another argument against the simple picture of oxygen vacancies ("empty" Ga polyhedra) as point defects in a-GaO_x.

The trend indicated by figs. 2.2a–d, i.e., a successive filling of the band gap by defect states through increasing oxygen deficit, is also plausible with regard to the reported colors of gallium oxide thin films. β-Ga_2O_3, a-$GaO_{1.5}$ and even a-GaO_x thin films with $x \geq 1.3$ have been described as colorless and transparent,[11, 33, 92, 95, 105, 110, 111] which means that their band gap must exceed the energy range of 1.6–3.3 eV for visible photons (cf. sec. 2.3). Severely oxygen-deficient films with x = 0.8–1.0 however have been reported as dark and opaque due to strong absorption over the whole

visible spectrum,[96, 97] which requires that their band gap is lower than 1.6 eV.[40] This also means that visible light may influence the electron transport in opaque a-GaO_x films, e.g., by additional photoexcitation of electrons above E_C. While this effect can be used to construct broad-spectrum photodetectors,[132] electrical characterization along the surface of a-GaO_x – e.g., for measuring the conductivity – should generally be done in the dark.

2.3.2 Electronic Conductivity

Pure β-Ga_2O_3 and a-$GaO_{1.5}$ are electrical insulators with $\sigma < 10^{-6}$ S/m at room temperature.[14, 123, 133] This is a direct result of the wide band gap that essentially prevents thermal excitation of electrons from the valence to the conduction band. Consequently, they are used e.g. as inert substrates [11], passivation layers in solar cells [107, 134] or gate dielectrics in field-effect transistors.[135] The low intrinsic conductivity also makes β-Ga_2O_3 and a-$GaO_{1.5}$ highly susceptible to the modification of electron transport through (un)intentional doping or non-stoichiometry. Thereby, another broad range of applications is enabled, e.g., as active material in diodes [136–138], transistors [123, 139, 140] and photodetectors.[86, 87, 89, 90, 141, 142] Furthermore, it is the basis for the resistive switching effect that will be discussed in sec. 2.4.

The most abundant modification of β-Ga_2O_3 is donor doping, e.g., by Sn or Si, which allows to change the conductivity $\sigma = e n_{\text{eff}} u$ over several orders of magnitude.[14] The donor energy levels are typically just about 50 meV below the CB edge (shallow donors), such that most of the respective electrons are excited to the conduction band.[14] Accordingly, Hall measurements have shown that the charge carrier density n_{eff} approximately equals the doping concentration and may thus be varied from 10^{15} to 10^{20} cm^{-3}.[14] The electron mobility u in the conduction band decreases slightly with doping due to additional scattering, but it has been found to stay in the order of 100 $cm^2/(Vs)$ at RT for many different dopants and concentrations.[14, 143] It is tremendously faster than the electron hole mobility in the valence band, which is expected to be in the order of 10^{-6} $cm^2/(Vs)$ at RT.[11] Usually, acceptor doping of β-Ga_2O_3 therefore does not lead to appreciable p-type conduction and may instead enhance its insulating character by compensating eventual donor impurities.[14] The difference of the electron and electron hole mobilities is also the reason why impure β-Ga_2O_3 is commonly an n-type semiconductor. Apart from Ga-substituting impurities like Si, also hydrogen incorporation (e.g., as H_i^+) is discussed as the origin of unintentional n-type conductivity in β-Ga_2O_3.[14, 112, 144–146]

Oxygen-deficiency in β-Ga_2O_3 is comparable to donor doping, since the formation of oxygen vacancies leads to excess electrons that are partially excited to the conduction band (eq. 2.67). As a result, β-Ga_2O_3 exhibits a p_{O_2}-dependence of the conductivity at high temperatures.[41] For ex-

[40] The dark brown color reported for a-$GaO_{1.2}$ and a-$GaO_{1.4}$ in [32] and [96] results from an embedded, more oxygen-deficient phase, as discussed in [96] and in sec. 2.3.

[41] In principle, the conductivity of β-Ga_2O_3 is also p_{O_2}-dependent at lower temperatures, but the adjustment of the oxygen vacancy equilibrium (eq. 2.67) takes an increasingly long time (about 1 s at 1000 °C; > 1 h below 800 °C).[147]

ample, at 1000 °C the oxygen vacancy concentration increases from about 2×10^{10} to 4×10^{12} cm^{-3} and the conductivity increases approx. from 0.013 to 0.25 S/m when the p_{O_2} is lowered from 1 to 10^{-5} bar.[114, 147] Like for extrinsic donor dopants, the electron mobility seems to stay about the same (around 100 $cm^2/(Vs)$ at RT).[148] The effect of the changing oxygen vacancy concentration on n_{eff} (and thus on σ) is much lower than for the aforementioned shallow donors, because oxygen vacancies in β-Ga_2O_3 are deep donors, with defect states rather far from the CB edge (> 0.8 eV, as mentioned in the previous sec. 2.3.1). Only few of the excess electrons are excited to the CB and contribute to conduction, while most occupy the oxygen vacancy defect states ($v_O^\times$, cf. eq. 2.67). Furthermore, even for an extremely low p_{O_2} of 10^{-15} bar (lower limit w.r.t. to phase stability) and rather high T = 1000 °C, just an equilibrium oxygen vacancy concentration of about 10^{18} cm^{-3} is expected,[114] which is in the order of the typical impurity concentrations (mainly Si) in commercial gallium oxide powder.[14] In β-Ga_2O_3, intrinsic oxygen vacancies are thus typically less relevant for the conductivity than extrinsic donors, except when the latter are compensated (e.g., by gallium vacancies, [114]). In amorphous gallium oxide, much more pronounced oxygen-deficiency is reached through non-equilibrium preparation techniques like PLD.[42] For example, x = 1.4 would formally correspond to an oxygen vacancy concentration in the order of 10^{21} cm^{-3} in a-GaO_x (estimate from density values in sec. 2.3). For rather pure a-GaO_x, the effects from non-stoichiometry thus can be expected to outnumber those from impurities by far.

Experimentally, a non-linear increase of the conductivity of a-GaO_x over many orders of magnitude has been observed for a decreasing O/Ga ratio x, from 2.2×10^{-8} S/m for x = 1.5 [126] to 3.3×10^{-6} S/m for x = 1.3 [126], 2.2×10^{1} S/m for x = 1.0 [96] and 2.3×10^{2} S/m for x = 0.8 [97] at room temperature. With increasing oxygen deficit, a-GaO_x converts from an insulator to a good conductor. In analogy to the donor doping by oxygen vacancies in β-Ga_2O_3, this has been attributed to an increase of the number of change carriers taking part in band transport above the CB mobility edge E_C.[32, 126] Mobility edge conduction is considered to be the dominating electron transport mechanism in a-GaO_x, since measurements of σ as a function of T showed an approximately linear relation between $\ln(\sigma)$ and T^{-1} (fig. 2.3a) in accordance with the MEC law (eq. 2.33).[32, 96, 97] Currently, all further models with regard to electrons in a-GaO_x, e.g., for Schottky barriers (sec. 2.3.3) or for resistive switching (sec. 2.4), are explicitly or implicitly built on this conclusion. The difference $\Delta E = E_C - E_F$ obtained from MEC fits decreases from x = 1.2 to 0.8 (fig. 2.3a), which explains the non-linear increase of σ with increasing oxygen deficit (cf. eq. 2.33). The same plausible trend was observed in the DFT-DOS discussed in sec. 2.3.1, where a shift of the Fermi level E_F to higher energy, i.e., towards E_C, can be seen as a result of more and more defect states emerging in the former band gap (fig. 2.2). Whether the actual ΔE values in fig. 2.3a are realistic is difficult to judge from the available DFT-DOS.

The alternative – variable-range hopping – has been reported just for crystalline gallium oxide,

[42] Note that crystallizing a-GaO_x in inert atmosphere leads to phase separation (as discussed in [91, 96] and sec. 2.3) and not to β-Ga_2O_3 with equivalent oxygen deficit. Annealing in vacuum even causes gallium loss through volatilization of Ga_2O.[149]

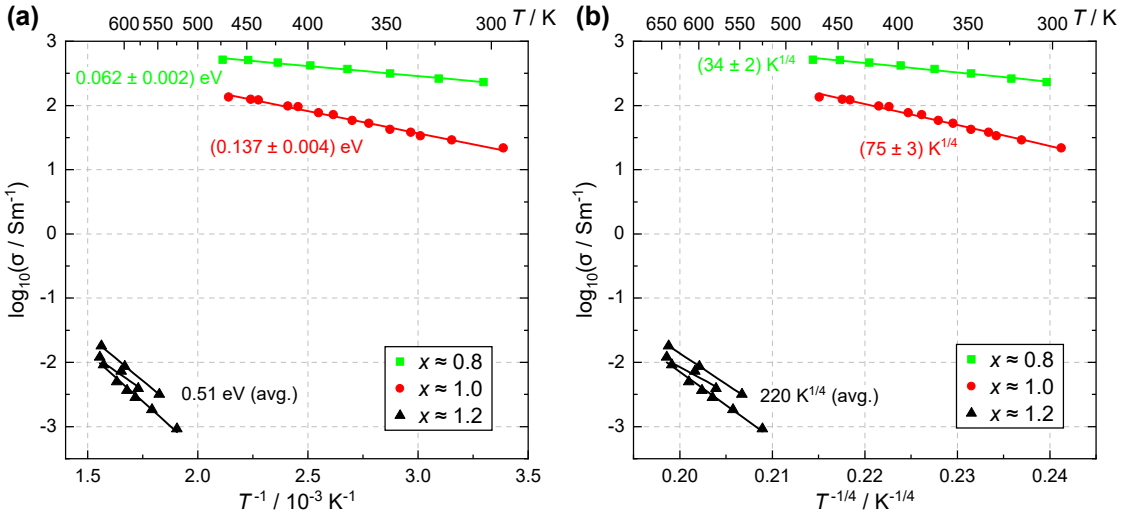


Figure 2.3: Experimental literature data for the conductivity σ of PLD-made a-GaO_x thin films with $x \approx 0.8$ [97], $x \approx 1.0$ [96] and $x \approx 1.2$ [32], shown both as (a) MEC-plot and (b) VRH-plot for comparison. Lines are the respective linear fits according to the MEC law (eq. 2.33) and the VRH law (eq. 2.54). The nearby shown values in (a) and (b) are the fit results ΔE and $T_0^{1/4}$, respectively. Note that the a-GaO_x films with $x \approx 1.2$ have most probably not been phase-pure (cf. sec. 2.3).

where electron hopping between dopant sites or along grain boundaries has been proposed in several cases.[136, 138, 141, 150–152] To the best of the author's knowledge, there are no mentions of VRH in amorphous gallium oxide so far. It probably has never been considered because in amorphous semiconductors there are typically too few localized states near E_F to compete with the much faster band transport beyond the mobility edges. VRH normally just dominates at temperatures far below RT, where the excitation of electrons above E_C is practically suppressed. However, for highly non-stoichiometric a-GaO_x the DFT-DOS (fig. 2.2d) indicate a substantial amount of localized defect states at/near the Fermi level. VRH thus might dominate at higher temperatures than usual. Indeed, the aforementioned $\sigma(T)$-measurements can also be adequately described by Mott's VRH law (eq. 2.54), as demonstrated in fig. 2.3b. The T_0 values obtained from the respective VRH fits decrease from $x = 1.2$ to 0.8, as would be expected from Mott's estimate $T_0 \propto 1/g(E_F)$ (eq. 2.54) in conjunction with the increasing DOS around E_F (fig. 2.2b–d). This trend for T_0 would also be an alternative explanation for the non-linear relation between σ and x (cf. eq. 2.54).

It is not possible to decide whether the MEC model (fig. 2.3a) or the VRH model (fig. 2.3b) is more appropriate for the shown conductivity data. The scatter of the data is too large and the covered T-range is too narrow on the reciprocal scale to observe significant differences in the fit quality. Furthermore, both models yield equally plausible trends that match the experimental results. Hall measurements, where the electron transport in the presence of a magnetic field is studied, could provide additional insight. It is known that for VRH the Hall effect is extremely small, because the electron tunneling events are not affected by the classical Lorentz force,[31, 153, 154] whereas for MEC the charge carrier density and the electron mobility above E_C can be obtained. The fact that charge carrier densities in the order of 10^{14} cm^{-3} and Hall mobilities in the range from 1.5

to 33.6 $cm^2/(Vs)$ have been reported for a-GaO_x at RT is thus a strong argument for MEC.[123, 155, 156] The mobilities are lower than for crystalline gallium oxide (see above), e.g., due to increased scattering in the amorphous structure. Unfortunately, the respective studies lack a determination of the stoichiometry of the a-GaO_x films. Based on the deposition conditions (e.g., PLD with p_{O_2} = 0.05 mbar at RT) and other details (e.g., transparency, band gap ≥ 4.0 eV), it seems like the investigated films had rather low oxygen deficit ($x \approx 1.5$), probably in conjunction with a decent amount of unintentional doping (rather high conductivities of 4.4×10^{-3} to 8.4×10^{-2} S/m). Substantial donor doping by impurities would of course distort the situation in favor of MEC. Still, these results indicate that MEC dominates in a-GaO_x at least for rather high x. This is also the regime where VRH is expected to be weak ($g(E_F) \to 0$). Further measurements are necessary to clarify the main electron conduction mechanism in a-GaO_x, especially for non-transparent films with $x < 1.2$. It is a major aim of this work to address this point (cf. sec. 5.3).

2.3.3 Metal|a-GaO_x Interfaces

The electron transport across metal|a-GaO_x interfaces is commonly treated as if a-GaO_x were a crystalline n-type semiconductor with a large amount of oxygen vacancies. Often, it is not mentioned that this analogy originates from the MEC model, that there is a mobility edge that takes the role of the band edge and that oxygen vacancies are just an illustrative picture for local non-stoichiometry in this case. The classical Schottky barrier models introduced in sec. 2.2.4 are used without accounting for localized states that might be present in the barrier region (except through a phenomenological ideality factor).[105, 156] This is probably a justified approach when the localized regions in the band tails are narrow, i.e., almost no difference between band edge and mobility edge, and when there are no localized states in direct vicinity of the Fermi level. The DFT-DOS (fig. 2.2) show that this is likely not the case for the severely non-stoichiometric gallium oxide that is investigated in this work, but it might be for films that are transparent and conductive.

As explained in sec. 2.2.4, the presence of a Schottky barrier is coupled to a difference of the work functions W of the metal and the regarded semiconductor. However, no work function values seem to be published for a-GaO_x so far. For β-Ga_2O_3 and a-$GaO_{1.5}$, almost equal work functions of W = 4.79 eV and W = 4.68 eV, respectively, have been obtained in Kelvin probe measurements on PLD-made thin films.[157] The other extreme, pure Ga metal, has a significantly lower work function of W = 4.32 eV.[158] Considering that the work functions of the electrode materials employed in this work are 5.5–5.6 eV for Pt [158, 159], 5.3–5.5 eV for Au [158, 160] and 4.3–4.5 eV for ITO [159, 161–163], Schottky barriers can be expected at Pt|a-GaO_x and Au|a-GaO_x interfaces.[43] ITO|a-GaO_x interfaces should either be ohmic or have low Schottky barriers (low w.r.t. φ_{bi}, cf. eq. 2.55).

The extent of the Schottky barrier depletion region depends on the relative permittivity ϵ_r (cf. eq. 2.59). Both β-Ga_2O_3 [164] and a-$GaO_{1.5}$ [121, 122, 135, 165] have $\epsilon_r \approx 10$ (range: 9–14) according

[43] See [15] for a comprehensive review of metal|β-Ga_2O_3 Schottky barriers.

to capacitance measurements (C-V). For non-stoichiometric a-GaO_x, there is apparently only one study, in which $\epsilon_r = 11.9$ is reported for transparent a-$GaO_{1.3}$.[92] Neither amorphicity nor oxygen-deficiency seem to influence ϵ_r, but this may be different for opaque a-GaO_x with a large amount of localized states in the former band gap.[44]

Experimental characterization of metal|a-GaO_x Schottky barriers is rare as well, especially in combination with a determination of the O/Ga ratio. For an Au|a-GaO_x interface, with x approx. 1.3–1.4 according to XPS, $\varphi_{bi} = 1.1\,\text{eV}$ and $n_d \approx 10^{17}\,\text{cm}^{-3}$ have been determined by C-V measurements.[105][45] In combination with the aforementioned work function of Au, this indicates a realistic work function of 4.2–4.4 eV for these transparent a-GaO_x films (cf. eq. 2.55). The barrier height has been extracted from the diode-like I-V curves of the Au|a-GaO_x devices as well. By fitting with eq. 2.61 for thermionic emission ($\log(I)$ vs. V), $\varphi_B \approx 0.7\,\text{eV}$ and $\eta > 2$ have been obtained.[105] The high ideality factor has been attributed to additional trap-assisted tunneling and/or barrier inhomogeneities. Since $\varphi_B < \varphi_{bi}$ would indicate $E_F > E_C$, the authors conclude that φ_B is underestimated as a result of the deviations from ideal thermionic emission.[105] Essentially the same has also been reported for Al|a-GaO_x Schottky barriers with unknown O/Ga ratio.[156]

Pt|a-GaO_x interfaces prepared by PLD with $p_{O_2} = 0.05\,\text{mbar}$ at RT exhibited diode-like I-V curves as well.[123] The stoichiometry of these transparent films has not been measured, but x should be close to 1.5 (cf. sec. 2.3). Taking the DFT-DOS in sec. 2.3.1 and the presence of impurities (high reported σ) into account, it seems plausible that thermionic emission dominates in this case. This seems to change for Pt|a-GaO_x interfaces deposited by PLD in vacuum at 450 °C, where $x \ll 1.5$ can be expected (cf. sec. 2.3).[9] In this case, the non-linear I-V curves are not diode-like but quite symmetric, indicating substantial tunneling or even hopping transport through the barrier region. For $V > 0$ applied to Pt (forward bias), an approximately linear relation between $\ln(I)$ and $V^{1/2}$ has been observed in these measurements, which does not fit to any of the classical Schottky barrier models introduced in sec. 2.2.4.[46]

ITO|a-GaO_x interfaces have been investigated using symmetrical, transparent ITO|a-GaO_x|ITO devices, prepared by RF sputtering at RT in pure Ar atmosphere.[124] As expected, their I-V curves are completely symmetric too. For each voltage polarity, $\ln(|I|)$ vs. V is approximately linear, which would fit to thermionic field emission through the respective reverse-biased ITO|a-GaO_x interface that should dominate the device resistance (cf. eq. 2.66).[47] However, in principle, the non-linearity of the I-V curves could also stem from a V-dependence of the bulk resistance. The rather large distance of about 2 µm between these ITO electrodes and the low work function of ITO makes a

[44] In several works about localized electrons (and VRH) in dielectric ceramics, these were seen as the reason for an observed colossal ϵ_r.[166–168]

[45] In the full depletion approximation, the capacitance C of a Schottky barrier is given by $C = \epsilon_0\epsilon_r/x_d$, with the depletion width x_d from eq. 2.59. This means $C^{-2} \propto \varphi_{bi} - V$. Note that the calculation of n_d from the slope of this relation requires to know ϵ_r, but the used value is not mentioned in [105].

[46] Note that in [9] this forward bias current is erroneously fitted with a reverse bias thermionic emission model that also neglects the V-dependence of the depletion width x_d (cf. eq. 2.59).

[47] Note that in [124] the I-V results are fitted with a wrong parallel circuit of two back-to-back thermionic emission equations (instead of a serial circuit). This effectively leads to a forward-bias model for both polarities.

dominance of the a-GaO_x bulk more likely, but it is unknown whether the conductivity of a-GaO_x would actually increase exponentially for the electric field strengths reached within these devices. As a result, the nature of ITO|a-GaO_x interfaces is yet unclear. Still, for asymmetric devices like Pt|a-GaO_x|ITO it is very common to assume a low resistance for the ITO|a-GaO_x interface in comparison to the other interface with higher work function difference.[9, 10, 18, 28, 169] In these cases, ITO is treated like an ohmic contact, even if – strictly speaking – it may not be one.[48]

2.3.4 Gallium and Oxygen Ion Transport

Although this work mostly focuses on electron transport, ion transport is no less important, especially for the resistive switching mechanisms that will be discussed in the next section. However, even for crystalline gallium oxide there are only few studies on gallium or oxygen ion diffusion.

The tracer diffusion of oxygen in β-Ga_2O_3 has been characterized by ^{18}O exchange experiments.[170, 171] Therein, β-Ga_2O_3 single crystals have been annealed in $^{18}O_2$ atmosphere at temperatures from 1200 °C to 1600 °C for about 24 h. The resulting diffusion profiles for the ^{18}O isotope have been measured by secondary-ion mass spectrometry (SIMS). The oxygen tracer diffusion coefficient D_O^* then has been obtained by fitting the profiles with the appropriate diffusion equation. For example, D_O^* is in the order of 10^{-17} cm^2/s for T = 1200 °C. This means that the tracer diffusion of oxygen is quite slow, even at such high temperatures. The main reason for this is most likely the low concentration of point defects (e.g., oxygen vacancies) that are a prerequisite for oxygen movement in the crystal lattice. This is also reflected in the high activation energies of 3.2 eV and 5.0 eV that have been extracted from the T-dependence of D_O^* for two different orientations of the crystals.[170, 171] According to theoretical calculations, these activation energies are dominated by the formation energy of the required defects.

Extrapolating the above results to room temperature yields extremely low D_O^* in the range of 10^{-60} to 10^{-84} cm^2/s. Still, it must be taken into account that this tracer diffusion coefficient just describes the diffusion without a concentration gradient. It may differ substantially from the Fickan diffusion coefficient that describes chemical diffusion driven by an oxygen concentration gradient. This is illustrated by a study, where Ti (50 nm) has been deposited on β-Ga_2O_3.[172] Before annealing, the Ti|β-Ga_2O_3 interface gave diode-like, non-ohmic I-V curves. However, after rapid thermal annealing at 470 °C for 120 s in inert atmosphere, the I-V curves became linear. Subsequent cross-sectional TEM-EELS measurements revealed that this ohmic contact formation is the result of oxygen diffusion from β-Ga_2O_3 into the Ti layer. The whole Ti layer is oxidized rather homogeneously to TiO_x within the short heating period. The EELS analysis also showed that the O/Ga ratio near the interface in β-Ga_2O_3 is drastically reduced to about 1.0 (50% Ga, 50% O). During the annealing process, the steep chemical potential gradient for oxygen causes an oxygen flux

[48] Real ohmic contacts to a-GaO_x, with linear I-V curves of the respective symmetric devices, have been reported for In and Ti electrodes.[86, 88, 90]

into the Ti layer, creating an increasing number of oxygen vacancies in β-Ga_2O_3 that greatly facilitate the further oxygen transport in this region. These results demonstrate that, depending on the conditions, oxygen transport in gallium oxide can be quite fast even at rather low temperatures.

For gallium diffusion, there is no suitable isotope that could be used in tracer diffusion experiments. Instead, Al has been used as a substitute to investigate the diffusion of metal ions in β-Ga_2O_3.[171] Aluminum diffusion profiles have been obtained by SIMS after deposition of Al_2O_3 on β-Ga_2O_3 single crystals and annealing at temperatures from 1100 °C to 1500 °C. The resulting aluminum tracer diffusion coefficients D^*_{Al} are about two orders of magnitude higher than the aforementioned D^*_O for oxygen (e.g., $D^*_{Al} \approx 10^{-15}\,cm^2/s$ for $T = 1200\,°C$). The activation energy for Al tracer diffusion (4.2 eV) is also lower than the respective value for oxygen (5.0 eV for the same crystal orientation), mainly due to a lower formation energy for gallium vacancies (at high p_{O_2} and typical impurity concentrations).[114] Correspondingly, the extrapolated tracer diffusion coefficient at RT is higher for Al ($D^*_{Al} \approx 10^{-72}\,cm^2/s$) than for oxygen ($D^*_O \approx 10^{-84}\,cm^2/s$ for the same crystal orientation), but, of course, still extremely low. These results indicate that gallium tracer diffusion is most likely faster than oxygen tracer diffusion in β-Ga_2O_3. However, as discussed above, this has only limited significance for chemical diffusion in non-homogeneous devices.

Chemical diffusion in β-Ga_2O_3 is addressed by another study, where single crystals have been bombarded with Ni^+ ions to introduce artificial disorder.[173] The simultaneous "healing" of the disorder[49] through diffusion has been monitored by Rutherford backscattering spectrometry (RBS) for different temperatures (25–250 °C). From the T-dependence, an activation energy of 0.8 eV has been determined. Due to the defect creation by ion irradiation, this energy does not contain the formation energy of the defects. Instead, it corresponds to the migration energy for Ga diffusion, since only the disorder in the Ga-sublattice has been measured by RBS. The obtained value is quite low, at least for ion transport, which indicates that appreciable Ga diffusion can be expected when enough vacancies are present in the Ga-sublattice. This is also demonstrated by the short timescale of this experiment. Despite the low temperatures, "healing" periods of only 12–720 s were sufficient to get meaningful results.

The oxygen and gallium ion transport in β-Ga_2O_3 has also been studied theoretically by DFT calculations.[113, 114, 174] They essentially confirm the above experimental findings. Migration enthalpies of 1.7–4.9 eV have been reported for the different migration paths of neutral oxygen vacancies ($v_O^\times$) in the β-Ga_2O_3 unit cell.[114, 174] The migration enthalpies tend to be lower for charged oxygen vacancies ($v_O^{\bullet\bullet}$; 1.1–2.7 eV), but their concentration is much lower under the usual experimental conditions. For gallium vacancies (v_{Ga}'''), migration enthalpies of 0.6–2.2 eV have been obtained for the different possible paths.[114, 174] They fit quite well with the aforementioned experimental value and they are significantly lower than those for oxygen. This indicates that gallium movement through the crystal is fundamentally easier.

[49] It has been verified by RBS that the introduced disorder is well below the amorphization level. Further, it is claimed that the accumulated dose of Ni^+ is too low to have any chemical influence.

The investigation of ion transport in amorphous gallium oxide is more difficult. For theoretical calculations, there are no longer well-defined migration paths or defect positions. For experiments, the crystallization temperature of about 350 °C sets an upper boundary for the observable diffusion lengths. For example, ^{18}O exchange annealing times of about 20 days have been necessary to get meaningful tracer diffusion profiles for a-$GaO_{1.5}$.[33, 97] Even with such long procedures, these experiments have only been possible because the tracer diffusion of oxygen is significantly faster in amorphous than in crystalline Ga_2O_3. The obtained tracer diffusion coefficients of oxygen in a-$GaO_{1.5}$ are much higher (e.g., $D^*_O = 8.1{\times}10^{-20}\,cm^2/s$ for $T = 300\,°C$) and the determined activation energy of 0.83 eV is considerably lower than for β-Ga_2O_3.[33, 97] Extrapolated to RT, a value of $D^*_O \approx 10^{-35}\,cm^2/s$ can be expected. This is many orders of magnitude higher than the aforementioned estimates for β-Ga_2O_3. The same trend of drastically easier diffusion in the amorphous state has also been found for other materials, e.g., for amorphous Si.[175] Still, it cannot compete with electron transport. With the analog of eq. 2.12 for O^{2-} tracers – using the density stated in sec. 2.3 and coarsely assuming ideality for the thermodynamic factor ($\Phi^* = 1$) – an ionic conductivity of $\sigma_O \approx 10^{-27}\,S/m$ can be estimated for oxygen in a-$GaO_{1.5}$ at room temperature. This is many orders of magnitude lower than the electronic conductivities discussed in sec. 2.3.2 (e.g., $2.2{\times}10^{-8}\,S/m$ for rather pure a-$GaO_{1.5}$ [126]).

Apart of these results for a-$GaO_{1.5}$, there seem to be no other experimental or theoretical works on ion transport in amorphous gallium oxide. Specifically, none were found for amorphous, non-stoichiometric gallium oxide. It is even more difficult to investigate due to its high susceptibility to oxidation.[33] The even lower density (cf. sec. 2.3) and the larger amount of electronic defects (i.e., less ideal bonds) in a-GaO_x could probably benefit ion diffusion, but, so far, that is just speculation. However, some general observations can be made:

- The known long-term stability of opaque a-GaO_x films during storage in air indicates that oxygen diffusion is slow at room temperature.
- The macroscopic phase separation observed during annealing at temperatures over 300 °C requires ion transport.
- On the one hand, resistive switching is based on ionic movement driven by an applied electric field. On the other hand, the resistance states would be volatile if ion transport would be too fast at room temperature.

It is unclear if the trends seen for β-Ga_2O_3 and a-$GaO_{1.5}$ are applicable to a-GaO_x. For example, it is unknown whether gallium or oxygen ions are more mobile in a-GaO_x. Either one, or even both, could be responsible for local changes within the material when a large enough electric field is present. Since they have opposite charge, however, the outcome for the local composition x and charge density ρ is qualitatively the same (e.g., $\nabla^2\varphi > 0$ causes decreasing x and increasing ρ, cf. eqs. 2.8, 2.14). Therefore, in order to simplify the further descriptions throughout this work, the picture of changing x (and ρ) through an oxygen ion flux will be used. This is common in the resistive switching literature on a-GaO_x, in analogy to the well-known oxygen vacancy migration

in crystalline oxides, but it must be kept in mind that also an opposite gallium flux could lead to the discussed effects.

2.4 Resistive Switching Mechanisms

The basic terms and mechanisms of resistive switching (RS) were already mentioned in the introduction (sec. 1). Here, the observations and proposed explanations for resistive switching of a-GaO_x will be examined more closely. But first of all, the nomenclature of different switching patterns must be clarified.

Resistive switching behavior can be classified into two major categories, unipolar or bipolar.[50] For unipolar switching, no change of the voltage polarity is necessary to switch back and forth between LRS and HRS (e.g., [17, 98, 100, 176]). Typically, this involves specifying a compliance current (i.e., a limit for I) for the HRS → LRS process. Applying a large enough positive (or negative) voltage with activated current limit triggers the HRS → LRS transition but too high currents are prevented. Afterwards, with deactivated current limit, applying again a sufficient positive (or negative) voltage causes a high current that triggers the reverse LRS → HRS transition (e.g., through Joule heating). In contrast, bipolar switching requires to alternate between positive and negative voltage in order to observe RS behavior. A current limit is typically not needed. Depending on which transition occurs for which voltage polarity, three types of bipolar resistive switching can be distinguished:

- **Figure-8 bipolar RS** (e.g., [9, 20, 23, 127, 177, 178]): An HRS → LRS transition occurs for $V > 0$, while LRS → HRS is observed for $V < 0$.
- **Counter-8 bipolar RS** (e.g., [9, 10, 18, 28, 29, 169]): An LRS → HRS transition occurs for $V > 0$, while HRS → LRS is observed for $V < 0$.
- **Abnormal bipolar RS** (e.g., [26, 27]): It looks like the same transition occurs at $V > 0$ and $V < 0$ (e.g., LRS → HRS), while at $V = 0$ the opposite process is observed (e.g., HRS → LRS).

The "abnormal" type usually involves at least one directional resistance (Schottky barrier). There is no driving force for real switching at $V = 0$, but due to the directional resistance it looks like a transition occurs when the polarity changes. The actual reverse transition is hidden. There are several possibilities how this can come about. In the easiest case, just one directional resistance is relevant, which behaves like this: Whatever internal mechanism increases its resistance with regard to positive voltage also simultaneously decreases its resistance for an opposite electron flow, and vice versa. For example, after an LRS → HRS transition was observed at $V > 0$, the device is – at the same time – in an HRS with regard to $V > 0$ and in an LRS with regard to $V < 0$. This can be probed using small positive and negative "read" voltages. Still, $V < 0$ would have to be applied to reverse these states. This distinguishes abnormal bipolar RS from unipolar RS.

[50] The citations in this paragraph refer to corresponding examples specifically for a-GaO_x.

A proper definition of the voltage polarity is mandatory both for the categorization and the interpretation of figure-8 and counter-8 bipolar behavior. The device notation usually follows the convention for electrochemical cells, i.e., anode|electrolyte|cathode, assigned for $V > 0$. The voltage is applied to the first mentioned electrode (= anode for $V > 0$), relative to the second electrode that equals electrical ground in most cases (= cathode for $V > 0$). For example, in a Pt|a-GaO_x|ITO device, positive voltage means that electrons flow from ITO to Pt, while negative voltage causes an electron flux from Pt to ITO.

The basic operation of resistive switching devices is typically investigated either by voltage sweeps or by voltage pulses. The continuous voltage sweeps yield I-V curves, where the switching property is visible as a hysteresis. It is important to verify the non-volatility of this hysteresis (e.g., by multiple sweeps with the same voltage polarity or by pauses within a sweep), because the RS devices are typically built like capacitors, which exhibit an I-V hysteresis as well – albeit a volatile one due to charging/de-charging. In pulsed measurements, positive or negative "write" pulses trigger resistive switching, while the momentary resistance state is monitored by "read" pulses with low voltage. This allows a more step-like operation of the devices, where switching and state-readout are decoupled.

The following sections give a condensed overview over the a-GaO_x-related resistive switching literature, with focus on the proposed mechanisms. Common to all explanation attempts is the persisting but reversible modification of the devices' electron conduction by ion transport. Typically, they are based on an electric field-driven modification of the local oxygen deficit in different regions of the device. Note that neither HRS → LRS nor LRS → HRS transitions may rely on Fickian diffusion of ions, because otherwise the device state could change even during zero voltage, i.e., when not used and during storage, which means that the state would be volatile. Of course, diffusion is a major factor for long-term state retention, but the switching processes must be driven by the applied voltage. It is also important to note that a-GaO_x devices stay amorphous during resistive switching, which distinguishes them from phase change memory.[10]

2.4.1 Filamentary Switching

The filamentary RS mechanism is based on the formation of a well-conductive, filament-like phase between the two electrodes of an RS device. Along this filamentary path, the conductivity is significantly enhanced in comparison to the surrounding phase. In general, this is due to a differing material composition, resulting from ion transport in an inhomogeneous electric field. Especially in highly resistive materials, the electric field due to an applied voltage is not uniform but guided by already present defects. In consequence, electron and ion transport is concentrated to certain regions, leading to the growth of individual filaments through the material. This process is self-amplified by the further accumulation of defects and the Joule heating along the filaments. When both electrodes get connected by such a continuous filament, they are essentially short-circuited and the device is in the LRS. Then, reversible resistive switching can proceed via rupture (LRS → HRS)

and re-connection (HRS → LRS) of the filament. The resulting abrupt changes of the device resistance are typical for filamentary RS. This actual RS operation of the device usually involves only a small part of the filament, where the connection is lost for example due to ion transport in the opposite direction (reversal of the voltage polarity). Therefore, the initial filament formation often requires much higher applied voltage than the ensuing RS transitions. This preliminary creation of the filament prior to actual RS is called electroforming. It is a soft variant of dielectric breakdown, and it often requires to limit the current in order to prevent complete device failure. Even though electroforming is common for filamentary RS devices, it is not mandatory. Especially in very thin layers, the normal RS voltages may be sufficient to grow a filament. Such devices, which exhibit RS directly from the as-prepared state, are called forming-free.

The resistive switching of a-GaO_x-based RS devices has often been attributed to this mechanism. Bipolar filamentary RS has been reported in the following studies:

- [177]: **TiN|a-GaO_x**(30 nm)**|Pt**, with a-GaO_x deposited by RF sputtering at room temperature, either in O_2/Ar (1:3) or pure Ar atmosphere. The O/Ga ratio for the two variants is $x \approx 1.3$ and $x \approx 1.2$, respectively, according to Ga $3d_{5/2}$ XPS peak analysis. After electroforming at +10 V, both exhibit figure-8 bipolar RS with abrupt HRS → LRS transition and rather gradual LRS → HRS process. For the sample prepared in O_2/Ar, the difference between HRS and LRS is larger, which is attributed to the higher x. For both samples, a linear relation between $\ln(I)$ and $V^{1/2}$ is observed in the HRS.
- [18, 169]: **Al|a-GaO_x**(15–20 nm)**|ITO**, with a-GaO_x deposited by RF sputtering at room temperature in different O_2/Ar gas mixtures (0.007 mbar). For samples prepared with up to 4% O_2, no RS is observed. Those with 20% to 50% O_2 during deposition exhibit stable counter-8 bipolar RS with abrupt HRS → LRS and LRS → HRS transitions. From the I-V curves, ohmic conduction in the LRS (approx. $I \propto V$) and space charge limited conduction in the HRS (approx. $I \propto V^2$) is concluded.[51] Both HRS and LRS current are independent of the device area.
- [178]: **Ta|a-GaO_x**(20 nm)**|Pt**, with a-GaO_x deposited by RF sputtering at room temperature in pure Ar atmosphere (0.01 mbar). Initially, the devices have high resistance. After electroforming at +5 V, figure-8 bipolar switching with abrupt HRS → LRS and rather gradual LRS → HRS transitions is observed. The I-V characteristics are quasi-ohmic both in LRS and HRS (approx. $I \propto V$).
- [20]: **Al|a-GaO_x**(30 nm)**|Pt**, with a-GaO_x deposited by RF sputtering at room temperature in different O_2/Ar gas mixtures with 0%, 10% or 20% O_2 (0.007 mbar). All three variants show figure-8 bipolar switching with abrupt transitions, without explicit electroforming but rather high voltage for HRS → LRS is required in each cycle. Approximately the same switching characteristics are observed with Ti instead of Al, but the device endurance is lower (presumably due

[51] Space charge limited conduction (SCLC) can occur when $W_m < W_s$ at a metal|semiconductor interface. In this case, contrary to a Schottky barrier, a negative space charge is present in the semiconductor, which hinders electron flow (n-type case). Under ideal conditions (e.g., no defects, no disorder), the current then follows the Mott-Gurney law with $I \propto V^2$.[179]

to oxidation of Ti at the interface). No resistive switching is observed for a Pt|a-GaO_x|Pt device.

- [127]: **Ag|a-GaO_x|Pt** and **Pt|a-GaO_x|Pt**, with a-GaO_x deposited by RF sputtering at room temperature in O_2/Ar gas mixtures with 0% to 1.5% O_2 (0.004 mbar). The O/Ga ratio increases slightly with the O_2 percentage, but all samples have $x \approx 1.4$ according to Ga $3d_{5/2}$ XPS peak analysis. Figure-8 bipolar RS with abrupt switching events is observed after electroforming at high positive voltages. Higher O_2 percentage in the deposition atmosphere leads to higher required voltages for forming and switching. The Ag|a-GaO_x|Pt device requires lower voltages for forming and switching than the Pt|a-GaO_x|Pt device, which is attributed to a cooperative filament formation through Ag^+ and O^{2-} ion fluxes. In all devices, the LRS has a more than one order of magnitude higher capacitance than the HRS.
- [23]: **TiN|a-GaO_x(20 nm)|Pt**, with a-GaO_x deposited by RF sputtering at room temperature in O_2(11%)/Ar atmosphere (0.0045 mbar). For the O/Ga ratio, $x \approx 0.9$ is reported based on Ga $2p_{3/2}$ XPS peak analysis, but the XPS signal is more like a single, symmetric peak instead of the two assumed sub-peaks for Ga^{3+} and Ga^+. Most probably, x is substantially higher. Figure-8 bipolar switching is observed, with abrupt HRS → LRS but gradual LRS → HRS transition, apparently without prior electroforming. A switching speed of 20 ns is reported.

In these works, filaments consisting of oxygen vacancies are proposed, even though vacancies are not well-defined in an amorphous structure. Instead, one should speak of filaments with lower O/Ga ratio x in comparison to the surrounding bulk material. Such filaments would have an increased conductivity due to the lower oxygen content, as discussed in sec. 2.3.2. It is speculated that the filament formation and the observed HRS → LRS transitions are the result of an O^{2-} flux towards the positive electrode, which is typically the less inert that may even incorporate oxygen ions (e.g., Al, Ta or ITO). Through this O^{2-} flux, a filament with lower x grows from the negative to the positive electrode and short-circuits the device. When the voltage polarity is reversed, the O^{2-} ions flow into the opposite direction, thereby increasing x at least in some region of the filament. This blocks the path of lower σ through the material and results in the LRS → HRS transition of the observed bipolar switching.

Unipolar filamentary switching has been reported as well, in the following studies:

- [17]: **Pt|a-GaO_x(60 nm)|Ga|ITO**, with a-GaO_x prepared by local anodic oxidation of elemental Ga with conductive atomic force microscopy (C-AFM) at room temperature and humid atmosphere. Unipolar RS is observed, using positive voltages applied to the Pt-coated AFM tip. At +2.5 V, with current limited to about 1 µA, an abrupt HRS → LRS transition occurs. During the following I-V sweep, at +1.7 V without current limitation, the device reaches a current of about 1 mA and abruptly switches back to the HRS.
- [98]: **Pt|a-GaO_x(200 nm)|Pt**, with a-GaO_x prepared by PLD at 300 °C in vacuum (10^{-8} mbar). Unipolar switching with abrupt transitions is observed. The HRS → LRS transition occurs at about +4 V (current limited to 6 mA), followed by the LRS → HRS transition at about +2 V and 14 mA when no current limitation is used. The same is observed for negative voltages, as

expected for the symmetric devices. The LRS current is ohmic ($I \propto V$) and shows a metal-like T-dependence (decreases slightly with increasing T), while the HRS current is interpreted as Poole-Frenkel emission (approx. $\ln(I/V) \propto V^{1/2}$; I increases with T).[52]

- [100]: **Ag|a-GaO$_x$|Nb:SrTiO$_3$**, with a-GaO$_x$ from PLD at 500 °C in vacuum (2×10^{-7} mbar). Initially, the device is in the HRS. Positive voltage of +5 V, in conjunction with a 0.7 mA current limit, triggers an abrupt transition to an LRS. Positive voltage of about +4 V without current limitation then causes an abrupt transition back to the HRS. Using negative voltages yields the same unipolar RS behavior. The LRS current is ohmic and metal-like ($I \propto V$, I decreases with increasing T), while the HRS current is interpreted as Poole-Frenkel emission (approx. $\ln(I/V) \propto V^{1/2}$; I increases with T).
- [176]: **Pt|a-GaO$_x$**(100 nm)**|Pt**, with a-GaO$_x$ deposited by RF sputtering at room temperature in pure Ar gas (0.004 mbar). Unipolar, abrupt switching is observed, with an HRS → LRS transition at +5 V (current limited to 10 mA) and an LRS → HRS transition at +4 V (without current limit). The same results for negative voltages. The I-V curves show ohmic conduction in the LRS ($I \propto V$) and space charge limited conduction in the HRS ($I \propto V^2$). Illumination with ultraviolet light has a much stronger impact on the HRS current than on the LRS current, and it also reduces the voltages necessary for triggering the switching events.

In these works about unipolar filamentary switching of a-GaO$_x$, again, the filaments are assumed to consist of oxygen vacancies (i.e., lower O/Ga ratio x). The proposed mechanism for the HRS → LRS transitions is the same as for the bipolar filamentary switching discussed above. The reverse transition back to the HRS is different, however, since it occurs with the same voltage polarity. In the unipolar case, it is thought that the LRS → HRS transition results from the diffusion of O^{2-} ions into the weakest spots of the filament, which is decisively enhanced through Joule heating by the high local currents.

A special type of filamentary RS is the electrochemical metallization mechanism, where the filament originates from an "active" electrode. When a positive voltage is applied to said metal electrode (= anode), it partially dissolves through the formation of metal ions. These ions migrate through the central layer of the device, towards the cathode, where they are reduced to metal atoms. In this way, a metallic filament grows from the cathode to the anode, consisting of material from the anode. Filament rupture is induced by opposite voltage polarity, which reverses these processes. The respective RS devices are typically called conductive-bridging memory. Electrochemical metallization in a-GaO$_x$-based devices has been proposed in the following studies:

- [19, 24]: **Cu|TiW**(1.5 nm)**|a-GaO$_x$**(20 nm)**|Pt**, with a-GaO$_x$ deposited by RF sputtering at room temperature in an O_2/Ar (1:1) or in a pure Ar atmosphere (0.004 mbar). Abrupt, figure-8 bipolar RS is observed in both cases, after electroforming at high positive voltage. The LRS resistance increases approximately linear with temperature, indicating a metal-like filament, and the slope

[52] Poole-Frenkel emission is an electron transport mechanism common in insulators, where the electrons move between localized states (traps) through intermediate electric field-assisted thermal emission to the conduction band.

of this T-dependence is close to the expectation for Cu metal. It is thus concluded that Cu-based filaments are responsible for the RS behavior. It is assumed that the oxygen-deficiency of the material facilitates the formation of the Cu filament. When the devices are annealed at 200 °C in O_2 atmosphere prior to the RS experiment, they show no reproducible switching behavior.

- [25]: **EGaIn|a-GaO$_x$**(1–3 nm)**|a-SiO$_x$**(2–3 nm)**|p$^+$-Si**, with a-GaO$_x$ prepared by oxidation of Ga-In eutectic (EGaIn) in air (= native oxide). Figure-8 bipolar switching with abrupt transitions is observed, without prior electroforming. The same is obtained with pure Ga or galinstan (alloy of Ga, In and Sn) instead of EGaIn. The LRS resistance is independent of the device area, while the HRS resistance varies with the area. It is proposed that filaments consisting of Ga metal are responsible for the RS behavior. Mainly, this is concluded from a volume expansion of the device that occurs during the HRS → LRS transition (as measured by C-AFM).

Electrochemical metallization may also be involved in some of the studies discussed further above, where oxygen-deficient filaments have been proposed. For example, Ag and Al could act as "active" electrodes, too.

Notably, there is no direct evidence for a filament, or for any ion transport, in all these studies about filamentary switching of a-GaO$_x$.[53] Instead, the main arguments for this mechanism are:

- Abrupt RS transitions, with HRS and LRS resistance differing by orders of magnitude
- Electroforming prior to RS behavior
- LRS resistance independent of the electrode area
- I-V and I-T relations consistent with bulk conduction (especially the metal-like T-dependence sometimes observed)

It is noticeable that the a-GaO$_x$ layers, for which filamentary RS has been proposed, are quite thin and intrinsically possess a relatively high resistance. These conditions might be beneficial for filament formation. For a high bulk resistance (= high O/Ga ratio x), the major voltage drop occurs over the bulk and the influence of defects on the course of electron and ion fluxes is enhanced. In combination with a thin layer, this may increase the probability that the switching process occurs along distinct paths instead of homogeneous across the device area.

2.4.2 Interface Switching

In the interface RS mechanism, the switching process takes place either at one or at both electrode interfaces. It requires that the interface resistances dominate over the bulk resistance, typically because of Schottky barriers at the involved interface(s). This is a common situation when the semiconductor bulk has rather high conductivity. The I-V response of the device is then governed by the electron transport over and through the Schottky barrier(s). The interface RS mechanism is based

[53] Filaments have been detected experimentally in several other materials, which is a major reason for the popularity of this mechanism.[180]

on the modulation of the Schottky barrier(s) through ion transport in the interface region. The height, width and shape of a Schottky barrier depends sensibly on the local material composition, which changes due to the applied voltage.

For example, when a sufficiently high positive voltage is applied to the metal at a metal|a-GaO_x interface (forward bias), it will increase the O/Ga ratio x in the vicinity of the interface, either through an oxygen ion flux towards the metal or a gallium ion flux away from the interface. The DFT-DOS discussed in sec. 2.3.1 indicate that such an increase of x would lead to higher work function W_s of a-GaO_x, lower built-in potential φ_{bi} (cf. eq. 2.55), higher $\varphi_n = E_C - E_F$ and lower density n_d of depleteable electrons. The opposite effects can be expected for negative voltage (i.e., decreasing x). Comparison with the Schottky barrier models in sec. 2.2.4 then gives an idea how a local variation of x would influence the electron transport through the barrier region. However, additional to the big question whether these models are valid for a-GaO_x, it is also far from obvious how Schottky barrier height and width react to changes of x. The above trends have counteractive effects on these two important quantities: Lower φ_{bi} reduces the barrier height φ_B, while higher φ_n increases it (cf. eq. 2.60). Lower n_d increases the depletion width x_d, while lower φ_{bi} decreases it (cf. eq. 2.59). Looking at the DFT-DOS and the trends for the optical band gap (sec. 2.3), it seems like φ_n and n_d might have a stronger influence, leading to an increase of φ_B and x_d for higher x. In this case, a metal|a-GaO_x Schottky barrier would undergo an LRS → HRS transition during forward bias, since both thermionic emission and tunneling would decrease. Accordingly, reverse bias ($V < 0$ applied to the metal) would cause an HRS → LRS transition. This is, of course, still a very simplified view, for example because the O/Ga ratio will not be homogeneous near the interface and ion transport will be influenced also by the intrinsic electric field in the Schottky barrier depletion zone.

Since the underlying ion transport processes affect the aforementioned Schottky barrier parameters in a continuous way, it is typical that interface-type switching is rather gradual than abrupt. While this is beneficial for multi-level switching, it can be detrimental for the switching speed. In contrast to filamentary RS, the interface RS mechanism generally does not require electroforming.

Interface switching of a-GaO_x-based devices, with only one Schottky barrier responsible for the RS behavior, has been reported in the following studies:

- [9]: **Pt|a-GaO_x(250 nm)|ITO**, with a-GaO_x from PLD at 450 °C in vacuum (7×10^{-7} mbar). The O/Ga ratio is unknown, but significant oxygen-deficiency can be assumed based on the deposition conditions and a measured band gap of 3.6 eV. Forming-free, counter-8 bipolar switching with gradual transitions is observed. For $V > 0$ (forward bias), the HRS current approx. follows $\ln(I) \propto V^{1/2}$ and $\ln(I/T^2) \propto T^{-1}$, which is interpreted as thermionic emission over the Pt|a-GaO_x Schottky barrier.[54] The LRS current can be described by $\ln(|I|) \propto |V|$ for positive

[54] This interpretation is based on reverse bias image force barrier lowering (lower barrier height for higher V, eq. 2.62), even though it is a forward bias situation, in conjunction with the wrong assumption of $F \propto V$ at the interface, thereby neglecting the V-dependence of x_d (eq. 2.59).

and negative voltage, which is attributed to thermionic field emission (tunneling) through the Schottky barrier (cf. eqs. 2.65, 2.66). The ITO|a-GaO_x interface is deemed irrelevant for the RS behavior due to the low work function of ITO. Using Ti (low work function) instead of Pt changes the RS behavior to figure-8 bipolar, which is explained with a bulk-type switching mechanism (see sec. 2.4.3).

- [99]: **Nb:$SrTiO_3$|a-GaO_x**(150 nm)**|Au**, with a-GaO_x prepared by PLD at 500 °C in vacuum (2×10^{-7} mbar). Figure-8 bipolar switching with gradual transitions is measured, without prior electroforming. It is claimed that the Au|a-GaO_x interface is ohmic and irrelevant for the RS behavior, because a similarly prepared Au|a-GaO_x|Au device showed linear I-V curves. The proposed mechanism, which is based on electron trapping/detrapping at the Nb:$SrTiO_3$|a-GaO_x interface ($v_O^{\times} \rightleftharpoons v_O^{\bullet\bullet} + 2\,e'$) without any ion transport, has a weak point: HRS and LRS would not be stable without an applied voltage. Since Nb:$SrTiO_3$ is – like ITO – a degenerate semiconductor with a low work function, it is more likely that these results correspond to counter-8 bipolar switching at the Au|a-GaO_x interface.
- [28]: **Pt|a-GaO_x**(60–100 nm)**|ITO**, with a-GaO_x prepared by PLD at room temperature in pure Ar atmosphere (0.005–0.015 mbar). The O/Ga ratio x is approximately 1.1–1.3, with lower x for higher deposition pressure, according to the O 1s to Ga 3d peak intensity ratio from XPS.[55] Counter-8 bipolar RS with gradual transitions is observed, also when measured at elevated temperatures up to 600 K (in vacuum). The higher the temperature, the lower voltages are required to trigger the switching events. The first HRS → LRS transition requires high negative voltage, i.e., electroforming. This is attributed to a surface oxidation layer that may have formed before the Pt deposition. It must be bridged by a short filament of lower x, before the actual interface switching can occur in this region. This is consistent with the observed area-independence of both HRS and LRS currents. The main arguments for non-filamentary RS are the non-abrupt transitions and the decreasing HRS and LRS resistances with increasing temperature (in contrast to metal-like filaments). Note that for both HRS and LRS, the relation of $\ln(\partial I/\partial V)$ vs. T^{-1} (at +0.1 V) is quite non-linear and convex, which could hint at a substantial influence from localized states (or even VRH transport).

In these works, interface switching at a single Pt|a-GaO_x or Au|a-GaO_x Schottky barrier – in the way described above – is a plausible explanation, although the actual evidence for this mechanism is very limited.

Interface-type RS with both interfaces taking part in the switching process has been reported in the following studies:

- [27]: **Pt|a-GaO_x**(200 nm)**|Pt**, with a-GaO_x prepared by PLD at 500 °C in vacuum (10^{-8} mbar). The O/Ga ratio is x = 1.3, as obtained from XPS in a not further specified way. Forming-free, abnormal bipolar switching is observed in the rather symmetrical I-V curves, with gradual

[55] The reported values for x are realistic for the deposition conditions, but it is unclear whether the different photoionization cross-sections of gallium and oxygen where taken into account.

LRS → HRS transitions visible both for positive and negative voltage. Upon change of the voltage polarity at $V = 0$, in both directions, there is an apparent HRS → LRS transition. The HRS current approximately follows $\ln(|I|) \propto |V|^{1/2}$ for both voltage polarities, which is interpreted as reverse bias thermionic emission.[56] The LRS is quasi-ohmic for positive and negative voltage (approx. $I \propto V$), probably due to facile tunneling (i.e., negligible interface resistances).[57]

- [26]: **Au|Ti**(50 nm)**|a-GaO$_x$|Ti**(50 nm)**|Au**, with an opaque a-GaO$_x$ layer deposited by PLD at 500 °C in vacuum (10^{-8} mbar). Like in their previous study [27] (see above), an XPS-derived O/Ga ratio of $x = 1.3$ is reported without details. Abnormal bipolar switching with symmetrical I-V curves is observed without previous electroforming. Gradual LRS → HRS transitions are visible at positive and negative voltage, while the device seems to switch from HRS to LRS at $V = 0$. The LRS has an ohmic $I \propto V$ relation and the HRS approximately follows $\ln(|I|) \propto |V|^{1/2}$ in the high-V regime, as observed also in [9] and [27]. The HRS is influenced both by the maximum voltage V_{max} and by the sweep rate v_{sw} used in the I-V cycles, while the LRS appears inert to these parameters. The higher V_{max}, the more pronounced is the switching to the HRS (i.e., the higher is the HRS resistance). The faster the I-V curve is measured (high v_{sw}), the higher is the voltage where the LRS → HRS transition occurs and the lower is the reached HRS resistance. This shows that the LRS → HRS transition is a gradual, time-dependent process, which can be accelerated by higher voltage.

In these two studies with abnormal bipolar RS, the following mechanism is proposed: In the as-prepared state, due to the high oxygen-deficiency, the Schottky barriers are thin and allow quasi-ohmic tunneling (= LRS). When positive voltage is applied to either electrode, then oxygen ions flow towards the respective interface and increase the barrier width. As a result, forward bias thermionic emission instead of tunneling occurs at this Schottky barrier (= HRS). When the voltage is reduced and close to zero, then the accumulated oxygen ions diffuse away from the interface, which restores the thin barrier width and enables tunneling again (= LRS). Afterwards, for opposite voltage polarity, the same processes occur at the other electrode. This mechanism is not only inconsistent with the reported I-V relations, it also means that the HRS is a volatile state. According to the high-v_{sw} results, the HRS → LRS transition would have to occur within fractions of a second when no voltage is applied. Consecutive negative voltage sweeps and state retention measurements in [27], however, demonstrate that HRS and LRS are non-volatile, at least for up to 10^4 s. This is only possible when there is no actual switching process around $V = 0$. Instead, the Schottky barriers' rectifying property must be responsible for the change of the device resistance around the origin. So far, this has not been respected in the models for abnormal bipolar switching of a-GaO$_x$.[58]

In summary, like for filamentary switching, there is only indirect evidence for interface-type RS of a-GaO$_x$-based devices. The main arguments put forward for this mechanism are:

[56] This interpretation is based on the wrong assumption of $F \propto V$ (with an "effective switching layer width" as inverse proportionality factor), thereby neglecting the V-dependence of the depletion layer width x_d.
[57] The authors in [27] report tunneling with $\ln(|I|) \propto |V|$ for the LRS, based on a very questionable fit.
[58] Note that the drift-diffusion modeling of abnormal bipolar RS in [30], which is based on the bulk switching mechanism, suffers from the same problem of a volatile HRS.

- Gradual, time-dependent RS transitions, with rather low HRS/LRS resistance ratio
- No electroforming
- Influence of the used electrode material on the RS behavior
- I-V and I-T relations attributed to models for electron transport over/through Schottky barriers

The a-GaO_x layers, for which interface switching has been proposed, tend to have a lower O/Ga ratio x and thus a lower bulk resistance than those where a filamentary mechanism has been assumed (cf. sec. 2.4.1). In conjunction with the use of electrode metals with high work function, this makes it plausible that the interfaces dominate the device resistance.

2.4.3 Bulk Switching

The bulk switching mechanism is directly based on the dependency between the bulk conductivity and the material composition. Any redistribution of ions within the central layer of an RS device inevitably changes the local conductivity σ in the affected regions and therefore also alters the bulk resistance of this layer as a whole. When the bulk is rather conductive (no incentive for filament formation) but still dominates over the interface resistances, then these changes of the bulk resistance can be visible as RS behavior with gradual transitions.

The strong, non-linear relation between σ and the O/Ga ratio x in a-GaO_x has already been discussed in sec. 2.3.2. The higher x in a certain region of an a-GaO_x layer, the lower is the local conductivity. In the simplest case with an initially homogeneous a-GaO_x layer, a positive voltage applied to either electrode leads to an increase of x on the respective side of the sample due to oxygen or gallium ion migration. The resulting decrease of σ in this region outweighs the increase of σ further within the a-GaO_x layer, where x decreased due to the ion transport.[59] Since the relation of σ vs. x is a non-linear, concave function (sec. 2.3.2), the result is an increase of the total bulk resistance (LRS → HRS transition). Afterwards, when a negative voltage is applied, the reversed electric field leads to a decrease of the local O/Ga ratio near the respective electrode. This leads to an again more homogenous O/Ga ratio throughout the sample and a lower total resistance (HRS → LRS transition). Thus, in this exemplary case, counter-8 bipolar RS would be observed.

However, when the initial composition of the sample is not homogenous, then also figure-8 bipolar bulk switching is possible. For example, if initially the region near the grounded electrode would have high x (i.e., low local σ), then applying positive voltage to the other electrode could lead to an HRS → LRS transition. The major voltage drop would occur over this region with high x, causing either oxygen ion migration away or gallium ion migration towards it. Thereby, the O/Ga ratio would become more homogeneous and the total bulk resistance would decrease. Afterwards, when

[59] When a device can be treated as 1-dimensional due to symmetry (built like a parallel-plate capacitor), then the total bulk resistance is simply given by the inverse of the path integral over σ^{-1} between the two electrodes (analog to a series circuit of resistances). Regions with low σ therefore have disproportionate influence on the overall resistance.

the voltage polarity is reversed, x would again increase in this region, resulting in a LRS → HRS transition.

Bulk switching of a-GaO_x-based devices has been reported in the following studies:

- [10]: **Pt|a-GaO_x**(50–90 nm)**|ITO**, with a-GaO_x prepared by PLD at room temperature in pure Ar atmosphere (0.04 mbar). The O/Ga ratio is x = 1.1 according to X-ray emission spectroscopy (STEM-EDX). Counter-8 bipolar switching with gradual transitions is observed, without previous electroforming. The HRS and LRS currents are proportional to the electrode area (i.e., current density constant w.r.t. area). The lower the a-GaO_x film thickness and the slower the sweep rate v_{sw}, the more pronounced is the observed RS behavior (higher HRS/LRS resistance ratio). Higher maximum voltage or elevated temperatures (up to 110 °C, in inert atmosphere) during the I-V sweeps also lead to more intense switching. It is mainly the LRS that is qualitatively and quantitatively affected by the variation of these parameters, while the HRS part of the I-V curves stays mostly the same (even w.r.t. film thickness). The more pronounced the switching processes (e.g., at slow sweep rates), the more linear is the I-V relation in the LRS, as expected for bulk electron transport in a well-conductive a-GaO_x layer. In contrast, the HRS has a distinct S-shaped I-V curve, most probably due to significant influence from the Pt|a-GaO_x Schottky barrier. It is detected experimentally, both by XPS and by STEM-EDX, that positive voltage leads to an accumulation of oxygen ions and an increase of x near the Pt|a-GaO_x interface (LRS → HRS), while negative voltage reduces the oxygen ion concentration and x in this region to approximately the same value as in the rest of the film (HRS → LRS). For the XPS experiment, the Pt electrode was removed at different stages of the I-V sweeps and the associated O/Ga ratio at the Pt|a-GaO_x interface was obtained from analysis of the Ga 3d XPS signal (LRS: $x \approx 1.0$, HRS: $x \approx 1.3$). In the STEM-EDX experiment, a cross-sectional Pt|a-GaO_x|ITO lamella was switched in-situ, while measuring the spatial oxygen and gallium ion distributions through EDX mapping. In the HRS, the oxygen ion signal is significantly increased in an about 10 nm wide region at the Pt|a-GaO_x interface (EDX counts almost doubled), while in the LRS the oxygen signal is quite homogeneous throughout the a-GaO_x layer. Even though EDX is known to be inaccurate for light elements like oxygen, this is still strong evidence that oxygen ion migration near the Pt|a-GaO_x interface is responsible for the RS behavior. The gallium signal is homogeneous across the sample and does not change with the applied voltages.
- [29]: **Pt|a-Cr:GaO_x**(120 nm)**|Pt**, with opaque Cr-doped a-GaO_x prepared by RF co-sputtering of gallium oxide and chromium metal at room temperature in pure Ar atmosphere (0.01 mbar). By adjusting the amount of chromium, it is possible to control the oxygen-deficiency of the films. The optimal switching performance is obtained for $Ga_{0.82}Cr_{0.18}O_{1.2}$ (composition measured by EPMA), which – when ignoring the Cr_2O_3 content – is comparable to a-GaO_x with $x \approx 1.2$ (same Ga^{3+}/Ga^{+} ratio according to XPS). Counter-8 bipolar switching is observed without an electroforming step. The slower the sweep rate v_{sw}, the more pronounced is the RS behavior (no switching at $v_{sw} \geq 100$ V/s). The shape of the I-V curves is quite similar to those in [10] (see above), but tend to be more symmetric. The LRS current scales linearly with the electrode area.

In these works, counter-8 bulk switching according to the mechanism discussed above has been proposed. This hypothesis is supported by drift-diffusion simulations with the FEM method in [10] and [30], which reproduce the counter-8 switching qualitatively. The simulations are based on oxygen ion migration and mobility edge conduction of the electrons, but comparable results can be expected when gallium ion transport or variable-range hopping of the electrons would be assumed.

A special kind of bulk switching has been proposed in [9], for a Ti|a-GaO_x(250 nm)|ITO device prepared by PLD at 450 °C in vacuum (7×10^{-7} mbar). In contrast to the Pt|a-GaO_x|ITO devices investigated in the same study (sec. 2.4.2), those with Ti show figure-8 instead of counter-8 bipolar switching. Due to the low work function of Ti, it is assumed that there is no relevant Schottky barrier. Instead, it is proposed that positive voltage applied to the "active" Ti electrode leads to the incorporation of oxygen ions into the Ti layer, such that the O/Ga ratio gets decreased in the vicinity of the Ti|a-GaO_x interface. As a result, the overall bulk resistance of a-GaO_x is reduced (HRS → LRS transition). For negative voltage, oxygen ions flow back from the Ti to the a-GaO_x layer, thereby restoring the as-prepared state (LRS → HRS transition). This mechanism is especially plausible if the O/Ga ratio near this interface would already be elevated in the beginning (e.g., due to surface oxidation). The effect of the reduction would then be all the stronger.

In contrast to the works about filament and interface switching summarized in sec. 2.4.1 and 2.4.2, the bulk switching studies contain first experimental evidence for compositional changes within the a-GaO_x layer. The measurements convincingly show that a reversible oxygen ion accumulation near one of the electrodes is involved in the switching behavior of a-GaO_x, at least for the investigated samples with rather low x. There is no sign of gallium ion migration. Oxygen-deficient filaments are ruled out by the following arguments:

- No filaments found in cross-sectional elemental mappings
- Linear increase of HRS and LRS currents with the electrode area
- Gradual, time-dependent RS transitions, with rather low HRS/LRS resistance ratio
- No electroforming

No conclusive statement can yet be made on the distinction from interface switching. The basic assumption would be that the Schottky barriers are very thin for the low O/Ga ratio x in these samples, such that the bulk resistance dominates despite the use of metals with high work function. This picture is supported by the high number of defect states (high n_d) expected for low x according to the DFT-DOS (sec. 2.3.1). Another major argument for bulk RS instead of interface RS is the observed thickness-dependence of the I-V curves. However, as discussed above, this mainly applies to the LRS, while the HRS could be governed by a Schottky barrier. The RS behavior probably even involves a gradual transition between bulk- and interface-dominated device resistance.

All in all, it is likely that bulk and interfaces together play a role in the non-filamentary resistive switching of a-GaO_x. For the full picture, a more profound understanding of its bulk properties and of the situation at metal|a-GaO_x interfaces is a goal of this work.

3

Experiments

3.1 Thin Film Deposition

The a-GaO_x thin films investigated in this work were prepared by pulsed laser deposition (PLD), following the procedure used in [10], where bulk switching of gallium oxide has been reported. The PLD method, which belongs to the physical vapor deposition techniques, is based on laser ablation from a solid target in a vacuum chamber. The material fragments ejected from the target by the laser pulses form a plasma plume, from which a thin film condensates on a substrate positioned above the target.

The used PLD workstation, built by SURFACE systems+technology, features two PLD chambers (#1 and #2), which are connected to a COMPexPro 205F KrF excimer laser (λ = 248 nm, Coherent). Apart from some explicitly stated exceptions, chamber #2 was used for the samples discussed in this work. Chamber #1 is equipped with a substrate heater, while the newer chamber #2 has a built-in laser energy meter and faster evacuation capabilities. The major difference between them, however, is due to the beamlines from the laser to each chamber. Due to an unequal length and differing optical elements, the energy loss and the focusing on the target differs between the beamlines. This leads to a different areal energy density of the laser beam on the target, called laser fluence, which is the decisive parameter for the ablation of the source material. It can be varied further by exchanging lens systems and adjusting attenuators in the beamlines. For the different used configurations of the beamlines, the fluence was determined by recording the laser spot size on wax paper placed on the target and by measuring the laser pulse energy in the chamber with a portable energy meter. This calibration of the beamlines was repeated from time to time, especially after maintenance.

Turbo pumps (Pfeiffer TMU 262, Pfeiffer HiPace 300) and membrane backing pumps (Vacuubrand) are responsible for the vacuum in the PLD chambers. The supply of Ar gas (99.9992%, max. 2 ppm O_2 and H_2O, Air Products), O_2 gas (99.999%, Westfalen) and H_2/Ar premix (5% H_2) to the chambers is managed through mass flow controllers. Prior to each deposition, the chamber was evacuated to a pressure below 10^{-6} mbar, followed by flushing three times with pure Ar gas. Afterwards, the desired pressure and composition of the deposition atmosphere were established by limited, continuous flow of either Ar, O_2 and/or H_2/Ar gas.

The required β-Ga_2O_3 target was synthesized from 4–5 g Ga_2O_3 powder (99.999%, Alfa Aesar), which was pressed uniaxially (80 kN, 10 min) and isostatically (200 kN, 85 min), followed by sintering at 1673 K for 24 h in air (heating/cooling: 3 K/min). After coarse polishing, a clean target surface was prepared by prolonged laser ablation. A comparable β-Ga_2O_3 target from A. von der Heiden, which is described in [33, 97], was used as well.

Different commercial substrates were utilized, depending on the intended use case. ITO-coated glass slides (120–160 nm ITO layer, SnO_2 content: 10–20 %, Sigma-Aldrich) were used as conductive substrates, e.g., as bottom electrode in RS devices or to prevent charging in electron microscopy. Amorphous quartz glass (SiO_2, Präzisions Glas & Optik) and crystalline sapphire slides (Al_2O_3, MaTecK) were employed as insulating substrates, e.g., for conductivity measurements. The detailed sample setup for the different measurements will be explained in the respective sections further below. The ITO substrates were cleaned by ultrasonification in acetone/isopropanol (1:1), rubbing with acetone-soaked tissue, rinsing with acetone and drying with compressed air. The polished SiO_2 and Al_2O_3 substrates were used without further cleaning. All substrates are squares with an area of $10 \times 10\ mm^2$ and a thickness of 0.5–1 mm. Due to the PLD sample holder with a 1 mm border, the coated area is limited to $8 \times 8\ mm^2$.

The amorphous, non-stoichiometric gallium oxide thin films were deposited with the following PLD parameters:

- Substrate temperature: RT (approx. 300 K)
- Process gas: pure Ar or H_2(5%)/Ar
- Chamber pressure: 0.04 mbar
- Laser fluence on the target: 0.2 to 5.0 J/cm^2 (varied)
- Laser pulse rate: 10 Hz
- Laser pulse count: 10^3–4×10^4 (depending on desired thickness)
- Distance between target and substrate: 40 mm
- Target rotation and toggling: enabled (for homogeneous ablation)

Stoichiometric a-$GaO_{1.5}$ thin films for comparison were prepared in O_2 atmosphere (0.04 mbar), with all other parameters equal to those above.

Pt, Au and ITO electrode layers were deposited by PLD as well, using commercial targets (MaTecK). Both PLD chambers have a target carousel that enables switching between different targets without the need to open the chamber. The basic PLD conditions were the same as for a-GaO_x. Pt and Au metal required high fluence, in the range of 3–5 J/cm^2, for sufficient ablation. They were deposited at room temperature and in the same atmosphere as the respective a-GaO_x film, i.e., either in Ar or H_2/Ar (0.04 mbar). ITO layers were prepared with rather low fluence of 0.8–0.9 J/cm^2, which already gives efficient film growth. Amorphous ITO was deposited at room temperature in O_2,

Ar or H_2/Ar (0.04 mbar), while crystalline ITO was obtained either by a substrate temperature of 300 °C during PLD in O_2 atmosphere (in chamber #1) or by post-annealing of amorphous ITO films at 300 °C in air. These PLD-made ITO layers will be denoted as "a-ITO" and "c-ITO", e.g., in the notation of devices, while the aforementioned commercial substrates are written as "ITO".

Pt electrodes were also deposited by DC magnetron sputtering and by electron-beam physical vapor deposition (EBPVD). In distinction to the PLD-made Pt, these layers will be denoted by "sp-Pt" and "eb-Pt", respectively. The sp-Pt deposition was carried out with an Edwards ScanCoat Six sputter coater, using Ar plasma with a base pressure of 0.1 mbar. The sputtering process lasted 3–6 times 300 s, with at least 2 min cool down in-between, at a deposition rate of about 0.2–0.3 Å/s. The eb-Pt electrodes were deposited with an electron-beam evaporator (Rocky Mountain Vacuum Tech) in a vacuum chamber with a base pressure of 1.3×10^{-6} mbar and a deposition time of 1500 s. The e-beam voltage and current were set to 7.3 kV and 100 mA, respectively, resulting in a deposition rate of about 0.2 Å/s.

Different layer dimensions and patterns, especially for the electrodes, were achieved primarily by using drilled or laser-cut shadow masks (Becktronic). Placed on the sample during the deposition, they limit the film growth to the non-covered regions. The employed masks and the resulting electrode patterns will be explained in sections 3.3 and 3.4. Alternatively, lithographic etching after deposition was used in the case of Pt-TEs (carried out by R. Borowski and M. Grates at PGI-7, FZ Jülich): The PLD-made Pt layer was covered with photo-sensitive resist (AZ 5214 E, MicroChemicals) using a BLE spin coater. After hardening of the resist at 90 °C for 5 min on a hotplate, the sample was exposed with a SÜSS MA6 mask-aligner (350 W, constant power mode, 25 s), developed in AZ MIF 320 solvent (MicroChemicals) and etched with an Oxford IonFab 300 Plus ion beam etcher (125 mA Ar ion beam current, perpendicular, time adjusted to 30 nm Pt, cooling: 15 °C). Residual resist was removed with hot dimethyl sulfoxide (80 °C). Besides a smaller feature size (< 100 µm), the lithographic structuring also allows to avoid air exposure of the Pt|a-GaO_x interface. In the case of lithographic Pt-TEs, the Pt layer was deposited directly on a-GaO_x, without leaving the vacuum conditions. Using shadow masks, in contrast, always requires to open the PLD chamber before the TE deposition. A potential downside of the lithography approach is the moderate heating of the sample.

3.2 Thin Film Characterization

Several analytical methods were used for the characterization of the deposited thin films. They address basic information like spatial dimensions, amorphicity and oxygen-deficiency, but also advanced properties like depth-resolved composition and work function. Together, they provide the basis for the interpretation of the elaborate *I*-*V* and conductivity measurements described further on.

The thickness d of the deposited films was measured by stylus profilometry (Veeco Dektak 150), where the vertical deflection of a diamond needle moving over the sample is recorded (2.5 µm tip radius, contact force: approx. 10 N). Baseline-correction was applied to each resulting profile by fitting the outer substrate regions with a parabola and subtracting it from the raw data. Then, the layer thickness d was calculated as arithmetic average of the aggregated height data from the plateau regions of multiple measurements on the same sample. The associated measurement uncertainty s_d was obtained as sample standard deviation of these collected height data points. Artificial spikes, which originate from the friction between needle and surface, were excluded from the analysis. The growth rate of the PLD films was obtained as the thickness divided by the number of laser pulses.

The surface roughness of exemplary thin films was measured by interference microscopy (IM) with a Veeco NT 110 instrument (W lamp, phase shift interference mode). After curvature & tilt correction of the 3D height data, the maximum area roughness S_z was obtained as the width of the height histogram. It corresponds to the vertical distance between the lowest valley and the highest peak on the investigated surface area.

The lateral dimensions of the thin films, especially of the patterned top electrodes, were obtained primarily with a calibrated optical microscope (Motic PSM-1000), which belongs to the probe station used for the I-V measurements (cf. sec. 3.3). In case of transparent layers or low optical contrast, profilometry and IM were utilized instead.

3.2.1 Electron Microscopy (EPMA, SAED and EELS)

Several variants of electron microscopy were used to gain insight into the structure and composition of the a-GaO_x thin films. They all rely on the interaction between an incident electron beam and the sample, but differ in the investigated sample region and in the detected effects.

Electron probe microanalysis (EPMA) was used to obtain the O/Ga ratio x in the bulk of the thin films. It belongs to the scanning electron microscopy (SEM) techniques, where the surface of the sample is scanned with a focused electron beam. The incident electrons scatter with core electrons in a teardrop-shaped interaction volume, which can reach several µm into the material (depending on the kinetic energy of the beam electrons). The extent of this interaction region determines the information depth of EPMA. The scattered core electrons are ejected as secondary electrons and the re-filling of their electronic states by valence electrons causes the emission of characteristic X-ray radiation. The latter is analyzed with wavelength-dispersive X-ray detectors, yielding the O/Ga ratio in the sample by comparison with a Ga_2O_3 standard.

The EPMA measurements were carried out by S. Richter at the Central Facility for Electron Microscopy (GFE) of RWTH Aachen University. A JEOL JXA-8530F instrument was used, which achieves relative accuracies of 1% to 5% for composition quantification. Typically, the O/Ga ratio was determined as average of multiple measurement points on the same sample. EPMA error bars

refer to the respective sample standard deviation. Only samples with ITO substrate were used for EPMA to prevent charging. Additionally, carbon coating was employed when needed. Rather thick a-GaO_x films and low electron beam acceleration voltages were necessary to avoid interfering oxygen signals from the substrate. Due to these constraints, only exemplary samples were measured. They were taken as representative for all other samples with equal PLD conditions and temporal proximity of the deposition. The actually measured O/Ga ratios will be written as "x = …", while those estimated from the EPMA results will be written as "$x \approx$ …".

Selected area electron diffraction (SAED) and electron energy loss spectroscopy (EELS) were used to determine the amorphicity and spatially resolved composition, respectively. Both methods belong to the transmission electron microscopy (TEM) techniques, where the electron transmission through an ultra-thin, cross-sectional lamella cut from the material is investigated. The electrons are diffracted upon traversal trough the lamella, leading to interference of the wave functions and an angle-dependence of the detected electron count. In a single crystal, the result is diffraction pattern with regular spots, while in a polycrystalline material the dots merge to discrete rings around the beam axis. In an amorphous material, there are only blurred rings, due to the widely varying inter-atomic distances. Electron diffraction thus yields information about the structure of the material. It is called SAED when an aperture is used to restrict the analyzed region on the lamella. EELS, in contrast, does not measure the diffraction (= elastic scattering) but the energy loss of the electrons (= inelastic scattering) after transmission through the material. The detected shifts of the electron wavelengths are characteristic for each element, since the atom-specific energy levels determine the energy that can be transferred during scattering. EELS therefore provides information about the composition of the material. It is especially sensitive for light elements like oxygen. By scanning the lamella with the electron beam, which is called scanning transmission electron microscopy (STEM), it was possible to record EELS line scans and area maps of the O/Ga ratio within the prepared thin films.

The SAED and EELS measurements were carried out by J. Ryu and S. Lee at the Electron Microscopy & Spectroscopy Lab (EMSL) of Seoul National University (SNU, Republic of Korea). A Thermo Fisher Themis Z STEM microscope was used for the measurements, with an acceleration voltage of 200 kV and an electron beam current of 50 pA. The required lamella were cut from the investigated samples with a focused Ga ion beam (FIB, Thermo Fisher Helios G4) and thinned by ion beam milling (Fischione NanoMill 1040, using Ar ions). The SAED images were recorded from circular regions with about 100 nm diameter. In the EELS measurements, a nano-parallel electron beam (probe diameter: 1 nm) was used for large-scale scans of the bulk, while a convergent electron beam was employed for high-resolution scans near the surfaces/interfaces of the lamella. Plural scattering was removed from the EELS spectra. The areal densities of gallium and oxygen were calculated from the Ga L edge and O K edge signals (core-loss EELS), using the Hartree-Slater cross sections included in the Gatan EELS software package. Both SAED and EELS were accompanied by annular dark-field imaging (ADF-STEM) to visualize the investigated lamella regions.

3.2.2 Photoelectron Spectroscopy (HAXPES and UPS)

Photoelectron spectroscopy was used mainly to analyze the O/Ga ratio x_{surf} at the thin films' surface and to obtain the work function W of a-GaO_x. It is based on the photoexcitation of electrons through monochromatic electromagnetic radiation – essentially the reverse process of EPMA. The electrons must be excited at least to the vacuum level in order to leave the sample, which means that the photon energy $h\nu$ must exceed the sum of the electron binding energy E_{bin} and the sample work function W_s. Excess photon energy becomes kinetic energy of the photoelectron. After leaving the sample, the electrons are accelerated further in the electric field between sample and detector, which originates from their different work functions.[60] Thereby, the electrons gain kinetic energy equal to $W_s - W_d$, where W_d is the detector work function.[61] The kinetic energy E_{kin} of the photoelectrons measured at the detector thus corresponds to

$$E_{kin} = h\nu - (E_{bin} + W_s) + (W_s - W_d) = h\nu - E_{bin} - W_d \tag{3.1}$$

$$= E_{cal} - E_{bin} \quad \text{with} \quad E_{cal} = h\nu - W_d \tag{3.2}$$

After calibration, photoelectron spectroscopy thus allows to determine the energy levels of the electronic states in the material. $E_{bin} = 0$ corresponds to the Fermi level. The peak positions in the photoelectron spectrum are not only characteristic for the elements, but also depend on the chemical environment of the atoms. Therefore, photoelectron spectroscopy provides rich information about the material. Note that a significant part of the generated photoelectrons already loses kinetic energy through inelastic scattering before leaving the sample. These so-called secondary electrons are responsible for a rather uniform background signal after each peak in the resulting spectrum.

In this work, two different variants of photoelectron spectroscopy were employed: Hard X-ray photoelectron spectroscopy (HAXPES), which uses hard X-ray radiation with a typical photon energy of 3–15 keV, and ultraviolet photoelectron spectroscopy (UPS), which uses UV light with photon energies of 3–124 eV.

The HAXPES measurements were carried out by S. Nemsak (PGI-6, FZ Jülich) at the DESY synchrotron facility (PETRA III, beamline 09), using a SPECS Phoibos electron analyzer with an energy resolution of 50 meV and a surface-detector angle of 90°. Hard X-rays with a photon energy of 5950 eV were used. Compared to conventional X-ray photoelectron spectroscopy (XPS) with lower photon energies, this leads to an increased information depth in the order of 10 nm and improved photon-electron scattering cross-sections.[181–183] Note that the information depth does not depend on the penetration depth of the X-rays, which is typically some µm, but on the escape depth of

[60] Sample and detector are electrically connected, such that the equilibration of the electrochemical potential of electrons yields an electric potential difference, i.e., an electric field.

[61] An additional voltage, which modifies the $W_s - W_d$ term (not included in the equations for sake of brevity), may be applied between sample and detector. This shifts the measured kinetic energies away from $E_{kin} = 0$, such that the spectrum does not overlap with secondary electrons that are generated through scattering at the detector. This is necessary when the electrons with lowest E_{kin} are of interest, e.g., for determination of the work function.

the generated photoelectrons (i.e., their inelastic mean free path). This makes the method surface-sensitive, in contrast to EPMA. It also means that electrons with higher kinetic energy E_{kin}, which originate from states with lower binding energy E_{bin}, have a higher information depth than those with lower kinetic energy (e.g., the Ga 3d peak has larger information depth than Ga 2p).[184]

Only thin films deposited on conductive ITO substrates were measured by HAXPES, in order to prevent charging through the photoemission process. No sample damage was observed after HAXPES. In addition to a-GaO_x and a-$GaO_{1.5}$ samples deposited in Ar and in O_2 atmosphere, respectively, also Au|a-GaO_x samples with a full Au coating were investigated (estimated Au thickness: 8 nm). The Au deposition was carried out either directly after a-GaO_x, without leaving the vacuum conditions, or after storage in air for different durations (1 h or 3 d). The influence of post-annealing in air was investigated as well. In the 2–3 weeks between preparation and HAXPES, the samples were stored in vacuum bags with a pressure of about 0.1 bar.

The resulting HAXPES spectra were calibrated with peak positions from literature, using eq. 3.2. The Au|a-GaO_x samples were adjusted to an Au $4f_{7/2}$ peak position of E_{bin} = 84.0 eV,[185, 186] giving equal E_{cal} = 5945.3 eV for all Au-coated samples. This value of E_{cal} also proved viable for the non-coated a-GaO_x samples deposited in Ar, yielding identical E_{bin} = 285.9 eV for the C 1s contamination peak of all samples, as well as gallium and oxygen peak positions in accordance with literature values (cf. tbl. C.1). The insulating a-$GaO_{1.5}$ sample deposited in O_2 showed the same C 1s position but deviations for the Ga and O peaks when using the same E_{cal}, due to the known issue of charging-induced peak shifts to higher E_{bin}. It has also been reported before that the C 1s signal is not a reliable reference for Ga_2O_3.[184] Instead, as recommended in [184], the a-$GaO_{1.5}$ sample was calibrated to a Ga 3d position of E_{bin} = 20.7 eV,[10, 115, 184] resulting in E_{cal} = 5944.7 eV and matching peak positions.

The slightly different chemical shifts of the gallium peak positions for Ga^{3+} and Ga^{+} ions were used to determine the O/Ga ratio at the surface of the thin films. Both the Ga $2p_{3/2}$ and the Ga 3d signal were deconvoluted into two sub-peaks for Ga^{3+} and Ga^{+}, yielding complementary x_{surf} (with higher information depth for Ga 3d). This was done by fitting each signal with a Shirley baseline and two Voigt functions (equal Gaussian and Lorentzian contributions enforced for the sub-peaks). From the areas A^{3+} and A^{+} of the respective Ga^{3+} and Ga^{+} Voigt sub-peaks, the surface stoichiometry was calculated according to

$$x_{surf} = \frac{1}{2} + \frac{A^{3+}}{A^{3+} + A^{+}} \tag{3.3}$$

For the a-$GaO_{1.5}$ sample, just a single Voigt peak was used – corresponding to Ga^{3+} (x_{surf} = 1.5), since adding a second sub-peak did not improve the fits.

The UPS measurements were carried out by K. Hong (and colleagues) at the Advanced in situ Surface Analysis System (AiSAS) of the Korea Basic Science Institute (KBSI) in Daejeon (Republic of

Korea). A Kratos AXIS-Nova photoelectron spectrometer and a He I UV light source with a photon energy of 21.2 eV were employed. The low photon energy limits the observable electronic states to the valence band, but allows to determine the material's work function from the photoelectrons with lowest E_{kin} (i.e., highest E_{bin}). These are the secondary electrons that leave the sample with approximately zero kinetic energy due to inelastic scattering. When reaching the detector, their measured kinetic energy therefore corresponds to $W_s - W_d$. The high-E_{bin} cut-off in the UPS spectra calibrated according to eq. 3.1 thus equals

$$E_{bin}^{max} = h\nu - W_d - E_{kin}^{min} = h\nu - W_d - (W_s - W_d) = h\nu - W_s \tag{3.4}$$

The difference between the photon energy and the binding energy cut-off therefore yields the work function of the material.

Three a-GaO_x thin films, which were deposited on ITO substrates in Ar atmosphere with slightly different laser fluence (0.72–0.81 J/cm^2, in chamber #1), were investigated by UPS. Before the measurement, one of them was annealed at 600 °C for 10 min in O_2 atmosphere with a rapid thermal annealing (RTA) system from Rocky Mountain Vacuum Tech (heating rate: 100 °C/s). During the UPS measurements, a constant bias of 15 V was applied between sample and detector to reduce the influence of scattering at the detector. A cleaned Au thin film was used as standard for calibration of the binding energy axis. There are two different methods for determining the cut-off value from the spectra, either as the inflection point of the high-E_{bin} boundary, or by extrapolating the boundary to the baseline. Here, following [159, 187], the extrapolation to the baseline was carried out with the UPS measurement & analysis software. The typical accuracy of the UPS energy axis, and of the resulting work function W, is about $4k_B T \approx 0.1$ eV due to thermal broadening (width of the Fermi function).[160]

3.2.3 Scanning Kelvin Probe Microscopy (SKPM)

The work function of a-GaO_x thin films was also measured more extensively by scanning Kelvin probe microscopy (SKPM), which is a variant of conductive atomic force microscopy (C-AFM). Like UPS, it relies on the electric field that exists in the space between two electrically connected materials with different work function. However, the measurement technique is completely different.

In SKPM, a conductive AFM tip, mounted to an oscillating cantilever, is moved across the sample surface in non-contact mode. The average tip-sample distance is kept in a region where two major forces act on the tip: van der Waals attraction (induced dipole-dipole interaction) and electrostatic force (due to different electric potential in tip and sample). The van der Waals attraction, which has a much stronger distance-dependance than the electrostatic force, is utilized to measure the sample's topography. Since it is an attractive force, it has a dampening effect on the tip's oscillation and causes a distance-dependent decrease of the amplitude. This is detected with a diode laser that

is reflected from the cantilever onto a four-quadrant photodetector. A lock-in amplifier, which is set to the cantilever's oscillation frequency, extracts the change in amplitude from the detector signal. A feedback loop then automatically adjusts the sample's vertical position to maintain a constant oscillation amplitude, i.e., a constant average distance between tip and sample surface. This vertical adjustment of the sample is recorded while scanning the selected area with the tip, resulting in a topography image. Simultaneously, the work function difference $W_s - W_t$ between sample surface and tip is mapped through the electrostatic force. This force depends on the difference of φ, i.e., on the voltage V, between sample and AFM tip. Since both are connected through an electric circuit, the chemical potential difference $\mu_t - \mu_s = W_s - W_t$ (eq. 2.55) causes a surface potential $(W_s - W_t)/e = -V_s$ that depends on the local work function W_s under the tip. Additionally, both a DC voltage V_{dc} and an AC voltage $V_{ac}\sin(\omega t)$ with angular frequency ω are applied between sample and tip. The total voltage V thus has three components:

$$V = V_{dc} - V_s + V_{ac}\sin(\omega t) \tag{3.5}$$

The electrostatic force acting on the tip in this capacitor-like situation then equals[62]

$$\frac{1}{2}\frac{\partial C}{\partial d}V^2 = \frac{1}{2}\frac{\partial C}{\partial d}\left[(V_{dc} - V_s)^2 + \frac{1}{2}V_{ac}^2 + 2(V_{dc} - V_s)V_{ac}\sin(\omega t) - \frac{1}{2}V_{ac}^2\cos(2\omega t)\right] \tag{3.6}$$

with capacitance C and sample-tip distance d. The detector signal also contains the vertical tip movements induced by this force. They have different frequencies than the mechanically produced cantilever oscillations, because a sufficiently different AC frequency ω is used. With a second lock-in amplifier, which is set to ω, the tip oscillations due to the third electrostatic force term

$$\frac{\partial C}{\partial d}(V_{dc} - V_s)V_{ac}\sin(\omega t) \tag{3.7}$$

can be extracted from the detector signal (cf. eq. 3.6). This term becomes zero when V_{dc} is set equal to V_s, i.e., when the electric field due to the work function difference is compensated by the applied DC voltage. In this case, there are no tip oscillations with angular frequency ω and the output amplitude of the second lock-in amplifier becomes zero. While scanning the selected sample area with the tip, a second feedback loop automatically adjusts V_{dc} to always maintain this zero amplitude situation. This results in a map of the relative work function $W_s - W_t = -eV_{dc}$ for the investigated surface area. Absolute values for the work function W_s can be obtained by determining W_t with a standard sample whose work function is precisely known.

A commercial atomic force microscope (XE-100, Park Systems), with an additional lock-in amplifier (SR830, Stanford Research Systems) for the V_{dc}-feedback loop, was employed for the SKPM measurements. The AFM instrument was configured to "EFM(ext)" mode. An AC voltage with $V_{ac} = 3.5\,\mathrm{V}$ and $\omega = 17.5\,\mathrm{kHz}$ was applied to the Au-coated conductive AFM probe tips (PPP-

[62] Eq. 3.6 is the classical force in a parallel-plate capacitor with homogeneous electric field, which is a common approximation for the C-AFM situation.

NCSTAu, Nanosensors). Highly oriented pyrolytic graphite (HOPG) with a work function of W = 4.65 eV was used as calibration standard. The obtained data was analyzed with XEI and Gwyddion.

Twelve a-GaO_x thin films, which were deposited in Ar atmosphere with different PLD fluence (0.36–1.21 J/cm^2, in both PLD chambers), were investigated by SKPM. Several samples also had Pt or Au electrodes, which were prepared by PLD with shadow masks (cf. sec. 3.3). These metallic layers and the ITO substrates were measured by SKPM as well. Additionally, the effects of different post-annealing procedures on the work functions were investigated. Some of the samples were annealed, either in a tube furnace with H_2(5%)/Ar atmosphere at 200 °C or 300 °C for 1 h, or by RTA (cf. sec. 3.2.2) in O_2 at 300 °C or 600 °C for 10 min, and then measured by SKPM again.[63]

In the SKPM measurements, topography and surface potential images were obtained by scanning rectangular surface regions with sizes of 5–40 µm^2. At least two different regions were measured on each investigated layer and each selected area was scanned twice. All scans from the same layer were stitched together for collective statistical evaluation. For each layer, the area average and the area standard deviation (S_q, root mean square roughness) of the combined surface potential image were calculated. The same was done for the multiple scans of the HOPG standard. From the resulting relative work functions, the absolute work function W_s of each investigated material layer was calculated as explained further above. The corresponding error bar was obtained from the area standard deviations through Gaussian error propagation. Further aggregation of the SKPM results, e.g., of all comparable a-GaO_x layers or all ITO layers, was done by weighted averaging and corresponding error propagation (as explained in sec. A.3).

3.3 Current-Voltage Measurements

The resistive switching behavior of the a-GaO_x thin films was characterized by current-voltage measurements, which is the standard method for detecting voltage-induced changes of the device resistance (cf. sec. 2.4). The sample layout for these experiments followed the classical capacitor-like structure of resistive switching devices, with an a-GaO_x layer vertically sandwiched between a bottom and a top electrode (BE and TE). The respective notation TE|a-GaO_x|BE was already introduced in sec. 2.4.

In most cases, the commercial ITO substrates were used as BE, with a thickness d of 120–160 nm of the conductive layer. Furthermore, PLD-made BEs of c-ITO ($d \approx 1.2$ µm), Au (d = 540 nm) or Pt (d = 20–30 nm) on SiO_2 substrates were employed in order to investigate the influence of the BE material. It was verified that the different BEs had negligible resistance (≤ 33 Ω from corner to corner). The a-GaO_x thin films were deposited on top of the continuous BE layers with a square shadow mask, such that a border of the BE remained accessible for contacting. The PLD fluence and

[63] The work function of the a-GaO_x film annealed at 600 °C in O_2 was also measured by UPS (after the SKPM experiments). It corresponds to the sample mentioned in sec. 3.2.2, allowing a direct comparison of the results.

the thickness of the a-GaO_x deposition varied from sample to sample, as will be specified alongside the respective results. The TEs were made of platinum, in the form of either PLD-made Pt (d = 20–30 nm), sp-Pt (d = 15–60 nm) or eb-Pt ($d \approx 30$ nm). Structuring of the TEs into small electrode pads was usually achieved by using shadow masks with differently sized, circular holes during the deposition (e.g., fig. 3.1a,b). The hole diameters of these masks ranged from 500 µm to 100 µm. Alternatively, in explicitly stated cases, the TE pads were created by photo-lithographic etching after PLD deposition of a continuous Pt layer (e.g., fig. 3.1c,d). This resulted in square TEs with sizes from 100×100 µm^2 down to 5×5 µm^2. The smallest investigated TE size was 10×10 µm^2, though, because the smaller ones could not be contacted in a reliable way. It was verified that the actual TE sizes approximately corresponded to the nominal sizes specified by the masks. Each individual TE pad, together with the a-GaO_x and BE regions underneath, represents a single resistive switching device. The lateral dimensions of this device essentially equal the TE area, because the a-GaO_x thickness is more than two orders of magnitude lower than the TE diameters. Each prepared sample thus contained many independent RS devices. Due to the vertical structure, the a-GaO_x region of each device is shielded from light by the opaque TE.

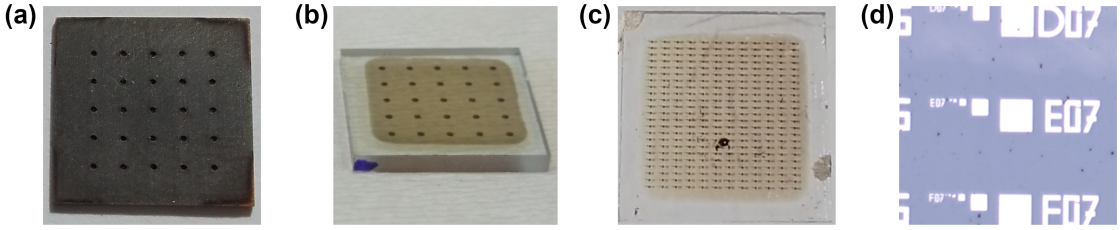


Figure 3.1: Sample setup for the I-V measurements. (a) Drilled shadow mask for circular TEs, in this case with equal diameters of 300 µm. (b) Example of a Pt|a-GaO_x|ITO sample, where the Pt-TEs were prepared by DC sputtering with a shadow mask. (c) Example of a Pt|a-GaO_x|ITO sample with a dense pattern of lithographically etched Pt-TEs. (d) Microscope image of a set of square Pt-TEs on the sample shown in (c), with pad sizes of 100×100 µm^2, 50×50 µm^2, 20×20 µm^2, 10×10 µm^2, 8×8 µm^2 and 5×5 µm^2.

After device preparation, the I-V experiments were carried out with an MPS150FA probe station (Cascade Microtech). The probe station is equipped with an optical microscope (Motic PSM-1000 with a Moticam 3M camera; 20–400× magnification) and stands on a vibration-isolated table (Cascade Microtech VIT801). The sample is kept in place by vacuum suction. Two micropositioners (Cascade Microtech DPP220), each with a probe arm that holds a metallic probe tip, were used for the precise contacting of the electrodes. Via the coaxial cables of the probe arms, the probe tips were connected to a Keithley Sourcemeter 2636B, which served both as voltage source and ammeter. In some early I-V experiments – before the construction of the probe station – a more simple measurement setup without a microscope and without vibrational damping was used. It consisted of a single micropositioner, a plain cable with stripped end for the second contact and a Keithley Sourcemeter 2611A. The contacting of the devices, i.e., the connections between the electrodes and the Sourcemeters, turned out as a major issue in the I-V experiments. Different contact methods were tried, which are summarized in the following together with their strengths and weaknesses.

Tungsten cat-whisker: These commercial probe tips (T-4-22, GGB Industries; 7S, Micromanipulator) consist of a thin tungsten wire (wire diameter: 22 µm and 127 µm, tip radius: < 1 µm).

+ suitable for smallest TEs
+ very flexible: low contact force and no device damage

– small contact area causes rather high contact resistance
– contact sensible to vibrations
– formation of an insulating WO_x surface layer over time[64]
– abrasive cleaning from WO_x or adsorbed contaminants not possible

Tungsten needle: These commercial probe tips (PTT-12-25, Cascade Microtech) consist of a rigid tungsten needle with a sharp tip (tip radius: 0.6 µm).[65]

+ proper contact to ITO-BE, since it allows high contact pressure and abrasive movements
+ less sensitive to vibrations

– small contact area causes rather high contact resistance
– damages metal layers upon touchdown if not careful
– formation of an insulating WO_x surface layer over time[64]

Etched gold wire: These self-made probe tips of pure Au were prepared by chemical etching of gold wire (diameter: 50 µm) with aqua regia (HNO_3:HCl, 1:3) for about 2 h, until a thin tip was obtained (tip radius: < 5 µm).

+ inert material (no oxidation)
+ very flexible: low contact force and no device damage

– irregular tip shape makes it difficult to hit small TEs
– small contact area causes rather high contact resistance
– contact sensible to vibrations
– abrasive cleaning from adsorbed contaminants not possible

Gold wire attached by silver paste: Instead of a using a probe tip, Au wire (diameter: 50 µm) was connected to the electrode by Ag paste (LS200N, Hans Wolbring). The paste was dried at room temperature and it was verified that this leads to negligible paste resistance. The other end of the Au wire was coiled around the stripped end of a cable to the Sourcemeter. This method was used just for contacting ITO-BEs.

[64] Removal of the WO_x layer by abrasive cleaning, chemical etching with NaOH or sublimation through Joule heating only led to minor improvements.

[65] Needles made of palladium alloy, which are inert towards oxidation (Paliney 7, SE-P, Signatone, tip radius: 12 µm), and of beryllium copper alloy, which is a rather soft material (forms native oxide like tungsten; SE-BC, Signatone, tip radius: 12 µm), were tested as well, but brought no significant improvements in comparison to the tungsten needles.

+ durable contact
+ essentially no contact resistance
+ no influence of vibrations
– only suitable for BE (low viscosity makes it difficult to attach Au wire to small TEs)
– brittle

Gold wire attached by platinum paste: Instead of a using a probe tip, Au wire (diameter: 50 µm) was connected to the electrode by Pt paste (LPA88-11S, Heraeus). The paste was dried on a hot plate at approx. 150 °C for a few minutes and it was verified that this leads to negligible paste resistance. In order to prevent later movements of the Au wire (e.g., tearing out or contact with other wires), both the sample and the other end of the Au wire were fixated on a larger glass slide. Thereon, the Au wire was connected to a Cu wire that led to the Sourcemeter. In the case of TE contacts, the Pt paste dot exceeded the deposited TE area significantly (approx. diameter: 700–900 µm). The area of the resulting, enlarged TE was measured with the calibrated microscope of the probe station.

+ durable contact
+ essentially no contact resistance
+ no influence of vibrations
– only suitable for BE and large TEs
– requires short sample heating
– additional Pt paste|a-GaO_x interface at TEs

These contact methods were used in different combinations for the TE and BE of the devices, together with either of the aforementioned measurement stations. Throughout this work, these combinations will be identified by "contact type" labels ("CT"), as defined in tbl. C.2.

The underlying reason, why contact resistances are relevant for the I-V experiments, is the generally low resistance of the a-GaO_x-based resistive switching devices. The resistance of a probe tip-to-electrode contact can influence the results, despite the high conductivity of the involved metals, because on a microscopic scale the actual contact area is extremely small (much less than the tip radius). This applies all the more, when corrosion or contamination of the tip is involved. Reducing the device size, i.e., the TE area, can partly alleviate this issue, as it increases the overall device resistance in comparison to the contact resistances. When conductive paste is used instead of a probe tip, the contact area is much higher and the electrode-to-circuitry resistance becomes negligible. Furthermore, attaching Au wire with Ag or Pt paste ensures a reliable contact right from the start of the measurements and prevents changes of the contact over time. This is crucial for observing how the resistive switching behavior starts and develops throughout repeated measurements. The major downside of the wired sample setup, however, is the inevitably large device size, which is beyond the usual size in applications.

Another aspect concerning the contact circuitry is the polarity of the measurements, which is important both for the distinction between figure-8 and counter-8 resistive switching (cf. sec. 2.4) and for the deduction of the underlying switching processes. The two terminals of the Sourcemeters are called HI and LO, and the voltage V corresponds to the electric potential difference $\varphi_{\mathrm{HI}} - \varphi_{\mathrm{LO}}$ between them. Since the chemical potential μ of the electrons in the metallic terminals is equal, the

voltage also specifies the electrochemical potential difference between LO and HI:

$$V = \varphi_{\mathrm{HI}} - \varphi_{\mathrm{LO}} \quad \Rightarrow \quad eV = \tilde{\mu}_{\mathrm{LO}} - \tilde{\mu}_{\mathrm{HI}} \tag{3.8}$$

For a positive voltage $V > 0$, HI is the plus pole, LO is the minus pole and the positive conventional current I flows from HI to LO. Since the conventional current is the flux of positive charge, for $V > 0$ (and $I > 0$) the electrons flow from LO to HI. For a negative voltage $V < 0$, the opposite is the case. In all I-V experiments, the TE was connected to HI and the BE was connected to LO. Additionally, LO was connected to ground. The voltage V thus was applied to the TE, relative to the grounded BE. When the contact and wire resistances are negligible, there is electrochemical equilibrium between the electrons at the terminals and the corresponding electrodes ($\tilde{\mu}_{\mathrm{HI}} = \tilde{\mu}_{\mathrm{TE}}$ and $\tilde{\mu}_{\mathrm{LO}} = \tilde{\mu}_{\mathrm{BE}}$). Therefore, under ideal conditions, the applied voltage enforces an electrochemical potential difference of

$$\tilde{\mu}_{\mathrm{BE}} - \tilde{\mu}_{\mathrm{TE}} = eV \tag{3.9}$$

between BE and TE. This electrochemical potential difference and the electric field it generates are the basis for the electron and ion movement within the resistive switching devices.

At the beginning of each I-V measurement session, the contact quality was checked by applying low voltages (e.g., 0 V → 0.1 V). The contacts were adjusted until these tests resulted in a smooth I-V curve. After that, the actual resistive switching experiments were carried out. By applying voltage sweeps or pulses $V(t)$, typically in the form displayed in fig. 3.2, changes of the device resistance were triggered that were monitored by measuring the resulting current $I(t)$. The voltage

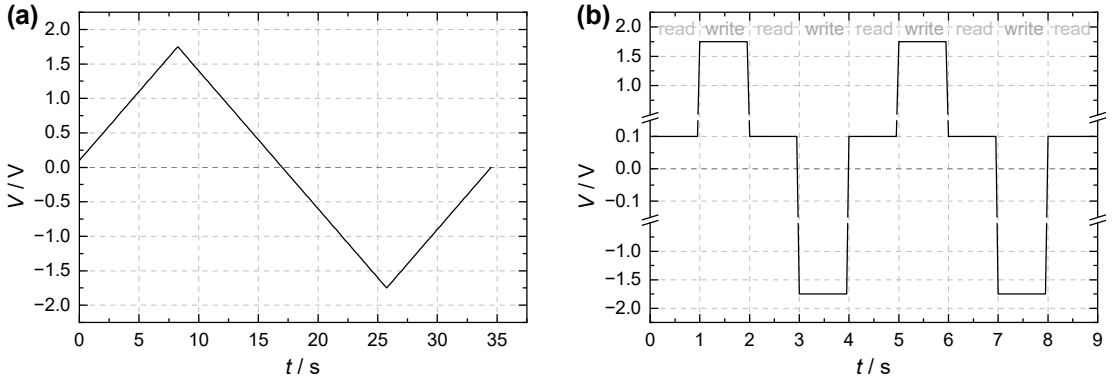


Figure 3.2: Examples for the applied voltages $V(t)$ in the I-V experiments. (a) Triangular voltage sweep with $V_{\mathrm{max}} = \pm 1.75\,\mathrm{V}$ and $v_{\mathrm{sw}} = 0.2\,\mathrm{V/s}$. (b) Voltage pulses with $V_{\mathrm{write}} = \pm 1.75\,\mathrm{V}$, $V_{\mathrm{read}} = +0.1\,\mathrm{V}$ and $\Delta t = 1\,\mathrm{s}$.

sweeps depicted in fig. 3.2a were carried out with different voltage amplitudes V_{max} and different sweep rates $v_{\mathrm{sw}} = \partial V / \partial t$. The measurements on a certain device usually were started with step-wise increasing V_{max} at low v_{sw}, until RS behavior was observed. Then, the influence of v_{sw}-variations and of further changes of V_{max} were investigated. The voltage pulses displayed in fig. 3.2b were used

to explore the effect of specific write voltages V_{write}, applied for a certain pulse length Δt, on the current measured at a low read voltage V_{read} afterwards. While the write pulses alternated between positive and negative voltage, the read voltage always was positive to eliminate eventual resistance changes through pure polarity reversal (diode effect). In most cases, symmetric sweeps or write pulses were used, where the voltage amplitude was equal for both polarities (e.g., $V_{max} = \pm 2.0\,V$). Sometimes, the voltage amplitude was asymmetric (e.g., $V_{max} = +1.8\,V|{-}2.0\,V$), to compensate for different switching at positive and negative polarity. Besides these standard sweep- and pulse-type measurements, also more specialized voltage schemes $V(t)$ were employed (e.g., sweeps with only positive or only negative voltage), which will be displayed alongside the respective results. The measurements were carried out with a compliance current of 10 mA, which was increased to 100 mA when the limit of 10 mA was exceeded without a breakdown of the device. From the measured current I, the current density $j = I/A$ was calculated with the nominal TE area A of the device.

3.4 Conductivity Measurements (van-der-Pauw)

The bulk conductivity σ of a-GaO_x thin films was measured with the van-der-Pauw technique. This 4-point probe method, which has been described for the first time in [188], requires four electrodes placed on the perimeter of a thin film sample. For concise notation, these contacts shall be labeled as A, B, C and D in counter-clockwise direction. The basic principle of the van-der-Pauw method is to apply a current between two neighboring contacts while measuring the voltage between the other two. This is done in two different configurations:

1. The current I_{AB} is injected at A and leaves the sample at B,[66] while the electric potential difference $V_{CD} = \varphi_D - \varphi_C$ is measured.
2. The current I_{BC} is injected at B and leaves the sample at C, while the electric potential difference $V_{DA} = \varphi_A - \varphi_D$ is measured.

Since the voltages V_{CD} and V_{DA} are measured without a current flow through the respective electrodes, there is no influence of contact resistances or Schottky barriers on the resulting values.[67] This is the major advantage of all 4-point probe methods. The van-der-Pauw technique, specifically, has also another benefit: In principle, the sample shape and the relative positions of the electrodes are arbitrary and do not need to be determined. It is just required that the electrodes are located on the circumference of the thin film and that they are sufficiently small in comparison to the sample area. The sample must be a thin film, i.e., much thinner than wide, with approximately homogeneous,

[66] The notation "I_{AB}" refers to a conventional current from A to B, which means that electrons flow from B to A.

[67] Voltmeters actually measure the electrochemical potential difference. The result is expressed as a voltage, because the sensing wires have equal chemical potential. In the ideal case of no current flow in the voltage measurement circuit, the electrochemical potential is constant across each contact region (no voltage drop), even when Schottky barriers or other contact resistances lead to local variations of μ and φ. This is why current-less voltage sensing is not affected by these interface effects. However, typical modern voltmeters always require a tiny current flow. The zero current condition can only be achieved approximately. In the measurements presented here, the current through the voltmeter is negligible, because the employed instrument has a large input resistance of over 200 TΩ.

known thickness d. The measured voltages from the two configurations together then implicitly contain all further information about the sample geometry. As has been derived in [188], they must obey the relation

$$\exp(-\pi d R_1 \sigma) + \exp(-\pi d R_2 \sigma) = 1 \tag{3.10}$$

$$\text{with} \quad R_1 = R_{\mathrm{AB,CD}} = \frac{V_{\mathrm{CD}}}{I_{\mathrm{AB}}}, \quad R_2 = R_{\mathrm{BC,DA}} = \frac{V_{\mathrm{DA}}}{I_{\mathrm{BC}}} \tag{3.11}$$

which links them to the conductivity σ of the material. R_1 and R_2 are also called van-der-Pauw resistances. Eq. 3.10 can be solved for σ [188], which yields

$$\sigma = \left(\frac{\pi d}{\ln(2)} \frac{R_1 + R_2}{2} f \right)^{-1} \tag{3.12}$$

with the geometry correction factor f that can be obtained numerically from the relation

$$\frac{R_1 - R_2}{R_1 + R_2} = f \operatorname{arccosh}\left(\frac{\exp\left(\frac{\ln(2)}{f}\right)}{2} \right) \tag{3.13}$$

Note that $f = 1$ when $R_1 = R_2$. When the thin film and the arrangement of the electrodes are close to symmetric (C_4 axis through the center of the sample), then $R_1 \approx R_2$ and the factor f can be approximated by [188]

$$f = 1 - \frac{\ln(2)}{2}\left(\frac{R_1 - R_2}{R_1 + R_2}\right)^2 - \left(\frac{\ln(2)^2}{4} - \frac{\ln(2)^3}{12}\right)\left(\frac{R_1 - R_2}{R_1 + R_2}\right)^4 \tag{3.14}$$

For a suitable, symmetric sample design, eqs. 3.12 and 3.14 yield the conductivity σ from the applied currents, the measured voltages and the film thickness d.

Accordingly, the a-GaO_x samples for the van-der-Pauw measurements were constructed symmetrically, as shown in fig. 3.3. Four contact pads, consisting of either a-ITO, c-ITO or Pt (cf. sec. 3.1), were deposited in the corners of insulating Al_2O_3 substrates by PLD with circular or triangular shadow masks (fig. 3.3a-c). It was verified that these electrode layers are highly conductive. Afterwards, a-GaO_x was deposited on top, using a square shadow mask (fig. 3.3a,b), such that the gallium oxide had appropriate overlap with the four electrodes. The layer depositions were carried out in this order to prevent intermediate exposure of the metal|a-GaO_x interfaces to air. The triangular electrode shape was introduced to improve the reproducibility of the contact overlap. Al_2O_3 instead of SiO_2 substrates had to be used, because the latter resulted in film delamination during the van-der-Pauw measurements. Fig. 3.3d,e shows exemplary samples from the van-der-Pauw experiments. A wide range of different a-GaO_x thin films was measured, all prepared in chamber #2. The influence of the PLD fluence and of the deposition atmosphere, varied between pure Ar and H_2/Ar, was investigated. The sample-to-sample reproducibility of the conductivity was examined on the basis of several similarly prepared samples. The different electrode materials were tested in combi-

nation with formally equal a-GaO_x layers. Even a sample without deposited electrodes, which was simply contacted in four corners by Pt paste, was measured for comparison.

The van-der-Pauw measurements were carried out in a custom built tube furnace, which is described comprehensively in [97]. It features a sample holder with four Pt contact pins and a thermocouple next to the sample. Gold wire was used to connect the sample's electrode pads to the respective contact pin. The Au wire was attached to the deposited electrodes with Ag paste (LS200N, Hans Wolbring) or Pt paste (LPA88-11S, Heraeus) and coiled around the Pt pins. The Ag paste was dried at room temperature and the Pt paste was dried on a hot plate at approximately 150 °C for few minutes. It was checked that this leads to contacts with negligible paste resistance. Pt paste was used in most cases because it is inert and less fragile, even though it involves a short heating of the sample.

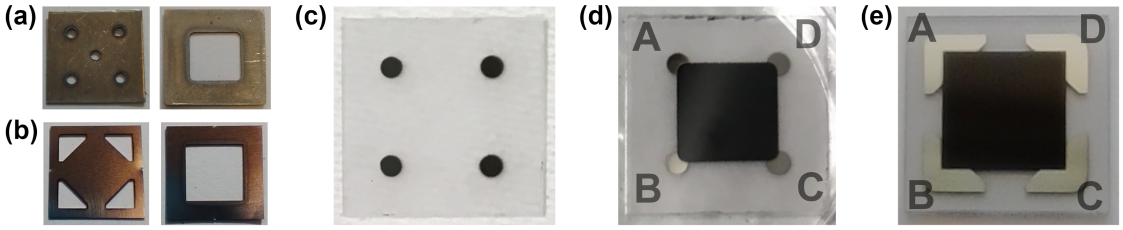


Figure 3.3: Sample setup for the van-der-Pauw measurements. (a) Drilled mask pair for circular contacts. The central hole was covered by Al foil to obtain only four dots. (b) Laser-cut mask pair for triangular contacts. (c) Circular Pt electrodes on Al_2O_3 substrate. (d) Complete sample with a-GaO_x thin film and circular Pt electrodes. (e) Complete sample with a-GaO_x thin film and triangular Pt electrodes. The labels A, B, C and D refer to the contact nomenclature that was introduced in the text.

Electric current was applied between the respective contact pins with a Keithley Sourcemeter 2400, while the voltage between the other two was measured with a Keithley Electrometer 6514. A relais allowed to switch between the two aforementioned configurations. Simultaneously, the temperature T was measured with the calibrated thermocouple in direct vicinity of the sample. The measurements were executed automatically with a LabVIEW program that controlled the Keithley instruments, the relais and the furnace. For each configuration, the applied current was varied from $-I_{\mathrm{max}}$ to $+I_{\mathrm{max}}$ with 7 evenly distributed current values. The maximum forced current I_{max} was adjusted in the range of 0.005–2.0 µA, such that the associated voltages V_{AB} and V_{BC} just not exceeded 5 V. Each individual point on the resulting van-der-Pauw I-V curves for I_{AB} vs. V_{CD} and for I_{BC} vs. V_{DA} was averaged over 5 successive measurements. Both I-V relations were checked for linearity and fitted with a linear function, yielding R_1 and R_2 as the inverse of the slopes. The respective standard deviations s_{R1} and s_{R2} were obtained from the fit errors through Gaussian error propagation. The conventional I-V relations for the current-driven contacts (i.e., I_{AB} vs. V_{AB} and I_{BC} vs. V_{BC}) were always checked for linearity as well, in order to rule out contact problems. The conductivity σ was calculated from R_1, R_2 and the film thickness d with the eqs. 3.12 and 3.14. Its standard deviation s_σ was obtained from s_{R1}, s_{R2} and s_d by Gaussian error propagation. Typically,

s_σ was dominated by the standard deviation s_d of the film thickness, due to the slight, PLD-inherent curvature of the a-GaO_x films (cf. sec. 3.2).

The conductivity $\sigma \pm s_\sigma$ was measured for 3-fold repeated temperature ramps of the form 300 K → T_{max} → 300 K, with a heating rate of 5 K/min and passive cooling. The highest used T_{max} was 575 K, but typically a maximum temperature in the range of 425–475 K was selected. In order to prevent oxidation, the van-der-Pauw measurements were conducted under inert atmosphere, established by a continuous Ar flow of 50 ml/min in the tube furnace. The samples were completely sealed off from light within the furnace to avoid photoexcitation effects. The resulting, T-dependent conductivity data was analyzed with regard to mobility edge conduction (MEC, sec. 2.2.2) and variable-range hopping (MEC, sec. 2.2.3) by comparing the respective $\ln(\sigma)$ vs. T^{-1} and $\ln(\sigma)$ vs. $T^{-1/4}$ plots.

Additionally to the measurements at elevated temperatures, the conductivity was also measured at temperatures below RT. These further measurements were carried out in cooperation with B. Thöner at the Institute of Physics (IA, RWTH Aachen University), using a commercial Physical Properties Measurement System (PPMS DynaCool, Quantum Design) with liquid He cooling. Due to the limited availability of the instrument, only two samples were measured in this way, with qualitatively equal results. The measurements were done in He atmosphere (3x purged, approx. 0.5–25 mbar) and without exposure to light. In contrast to the procedure described above, the PPMS applies constant currents I_{AB} and I_{BC} and records the resulting van-der-Pauw voltages V_{CD} and V_{DA} continuously in short intervals of about 1 s. R_1 and R_2 thus were obtained from the respective V/I ratio (eq. 3.11), instead of the aforementioned I-V fits. As a side effect of this, s_σ was calculated just from the thickness error s_d. The applied currents were adjusted in the range of 0.8–1.0 µA to keep the associated voltages V_{AB} and V_{BC} below 1 V.

A major difference between the PPMS measurements and the other van-der-Pauw experiments is the contacting method. Instead of the Ag or Pt pastes, wire-bonding was used to solder Au wire to the Pt electrode pads of the samples. Therefore, the amorphous thin films did not experience any heating before the PPMS measurements. This allowed to investigate their as-prepared state by measuring a cooling ramp at first. The temperature was varied according to 300 K → 150 K → 400 K → 150 K → 300 K, with cooling and heating rates of 2 K/min. At each turning point of these ramps, the temperature was held stable for about 1 h. These periods with constant T served two purposes. In the first 20 min of each, it was monitored whether σ changes over time at the respective temperature. Afterwards, as described in the following, an additional magnetic field was applied perpendicular to the sample (direction: bottom to top of the thin film), in order to investigate the magnetoresistive effect and the Hall effect for a-GaO_x.

The magnetoresistance $\Delta R/R_0$ is defined as the relative change of a sample's resistance when a magnetic field is applied. This effect can also be expressed by a magnetoconductivity $\Delta\sigma/\sigma_0$, which describes the change $\Delta\sigma$ of the conductivity due to a magnetic field, relative to the conductivity σ_0 without magnetic field. This quantity was obtained by measuring the van-der-Pauw conductivity as described before, while varying the magnetic field strength H according to 0 → 5.6×10^6 A/m →

-5.6×10^6 A/m $\rightarrow$ 0. Simultaneously, a third I-V configuration was measured to determine the Hall effect: The current I_{BD} = 0.5 µA was applied between the diagonal electrodes B and C, while measuring the Hall voltage $V_H = V_{AC}$ between the other two. In the classical case of crystalline semiconductors and for mobility edge conduction (MEC) in amorphous semiconductors, such a perpendicular voltage results from the deflection of the flowing electrons in the magnetic field (Lorentz force). From this effect, the charge carrier density and the majority charge carrier type can be determined. However, as mentioned in sec. 2.3.2, this is not the case for variable-range hopping (VRH). The individual tunneling events are not influenced by a magnetic field and only quantum interference between different hopping paths may lead to a Hall effect that is typically extremely small.[31, 153]

The Hall coefficient $R_H = V_{AC}d/(I_{BD}B)$ was not calculated, because the magnetic permeability that relates H to the magnetic flux density B is not known for a-GaO_x. Instead, just the Hall voltage V_H is reported, after the following asymmetry correction: Since the contacts form no perfect cross, the raw V_H always contains minor contributions from the voltage V_{BD} between the current-driven electrodes. When I_{BD} and σ are constant, then this additive contribution is also constant with regard to H. In this case, the raw V_H at $H = 0$ would have to be subtracted from all measured Hall voltages, because there should be no Hall voltage without magnetic field. However, when σ varies due to the magnetoresistive effect, then the relation of V_H vs. H is altered in a non-constant way by the changing contributions. This can be corrected for by realizing that the magnetoresistive effect does not depend on the sign of H, while the Hall effect is antisymmetric with regard to $H = 0$. For a-GaO_x, removing the deviations turned out to be especially simple, because, as will be shown in sec. 5.3.4, the magnetoconductivity $\Delta\sigma/\sigma_0$ of a-GaO_x depends linearly on $|H|$. Therefore, the raw V_H vs. H data was fitted with a linear absolute value function, which then was subtracted from the data to obtain the corrected Hall voltages.

4

Simulations

4.1 Density Functional Theory (DFT)

The experimental work on a-GaO_x was complemented by theoretical calculations, using three major methods that build upon each other. The first is density functional theory (DFT), which yielded the density of states (DOS) of a-GaO_x model structures. The DOS is the basis for the ensuing kinetic Monte-Carlo (KMC) calculations described in sec. 4.2, which provided the bulk conductivity of a-GaO_x in dependence of parameters like the temperature or the electrons' chemical potential. The conductivity results finally were used to model Pt|a-GaO_x|ITO devices with the finite element method (FEM), as will be explained in sec. 4.3.

The DFT method basically approximates the ground state electron density $n(\vec{r})$ of a quantum mechanical many-body system with N electrons by iteratively solving N Kohn-Sham equations

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(\vec{r})\right]\phi_i(\vec{r}) = E_i\phi_i(\vec{r}) \tag{4.1}$$

which are coupled by the effective potential V_{eff} (m is the electron mass). The Kohn-Sham wave functions ϕ_i with state energy E_i together form the electron density

$$n(\vec{r}) = \sum_i^N |\phi_i(\vec{r})|^2 \tag{4.2}$$

which in turn determines the effective potential V_{eff} in eq. 4.1 that a single electron i experiences:

$$V_{\text{eff}}(\vec{r}) = V_{\text{ext}}(\vec{r}) + \frac{e^2}{4\pi\epsilon_0}\int \frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|}\,\mathrm{d}^3\vec{r}' + V_{\text{XC}}[n(\vec{r})] \tag{4.3}$$

Here, V_{ext} contains the Coulomb interaction with the atomic cores, while the second term describes the Coulomb repulsion by the other electrons. The exchange-correlation term V_{XC} is a functional of the electron density and approximates the quantum mechanical interaction of the electron wave functions in the many-electron system. The ground state electron density, which results from the self-consistent, iterative refinement of the ϕ_i, allows to calculate quantum mechanical observables

like the ground state energy of the system. This enables for example the optimization of the atomic structure by minimizing the system's energy.[68] In solid state DFT, the Kohn-Sham wave functions are typically constructed from periodic Bloch functions with a wave vector $\vec{k}$ (cf. sec. 2.2.1), which leads to the notation ϕ_{nk} with the band number n and the k-point number k as indices. There is no mathematical guarantee that the Kohn-Sham wave functions ϕ_{nk} resemble the actual single-electron wave functions ψ_{nk} of the system, but the empirical approximation $\phi_{nk} \approx \psi_{nk}$ is commonly used nonetheless. The density of states $g(E)$ of the system can then be obtained as the interpolated distribution function of the Kohn-Sham state energies E_{nk} on the energy axis.

4.1.1 Amorphous Model Structures

The first step towards the calculation of the DOS for a-GaO_x was the generation of suitable amorphous model structures. Fortunately, it was possible to utilize the earlier work of Fabian Draber, a colleague in the same workgroup. He created a-GaO_x DFT cells with the ab-initio molecular dynamics (AIMD) procedure that is described in the following, and he also already calculated the DOS for the final structures. His unpublished[69] DFT parameters and results were reused in this work with his explicit, generous permission. The amorphous structures were adopted as obtained from him, but the DFT-DOS calculations were repeated and expanded to allow for the additional analysis that will be described in sec. 4.1.2 and 4.1.3.

The AIMD method uses DFT-derived forces to calculate the trajectories of the atoms at the specified temperature. In each timestep of an AIMD simulation, the atoms are moved according to these trajectories and then the forces are recalculated by DFT. This allows to emulate a melt-quench process to amorphize DFT cells. The a-GaO_x structures used in this work were generated by F. Draber starting from crystalline 2×2×2 supercells of β-Ga_2O_3, which were forced into a tetragonal cell shape. The composition of the cells was adjusted by randomly removing oxygen atoms or adding gallium atoms, which resulted in $Ga_{64}O_{96}$ (x = 1.5), $Ga_{64}O_{78}$ (x = 1.2), $Ga_{64}O_{64}$ (x = 1.0) and $Ga_{72}O_{58}$ (x = 0.8). In multiple sequential AIMD runs with timesteps of 3 fs (NVT ensemble),[70] these four initial structures were virtually melted at 5000 K for a total time of approx. 15 ps, followed by cooling them down to 300 K at rates of 0.5–2 K/fs. Lastly, the structures were artificially tempered at a constant temperature of 300 K for about 15 ps. In-between these steps and intermediately of the tempering step, the cell volume was optimized by DFT until a maximum residual force smaller than 0.05 eV/Å was reached, while keeping the fractional coordinates of the atoms and the cell shape fixed.[71]

[68] For a more thorough introduction to DFT in general and DFT in the solid state in particular, the reader is referred to the respective textbooks (e.g., [189, 190]).
[69] The combination of Fabian Draber's DFT calculations and the results described throughout this work led to the joint publication [40].
[70] Major VASP settings: IBRION = 0, MDALGO = 0, ISIF = 2, PREC = low, SMASS = –1, NBLOCK = 2.
[71] Major VASP settings: IBRION = 2, ISIF = 7, PREC = normal.

A comparable AIMD approach to generate structural models for a-GaO_x has also been applied in [89], for the DFT-DOS calculations that were already discussed in sec. 2.3.1. However, the calculations in [89] only cover O/Ga ratios above 1.2 and therefore they are insufficient for comparison with the experimental results reported here.

The AIMD simulations and the volume relaxations were carried out with the VASP software [191–194], using PAW GW basis sets [195, 196] and the GGA-PBE exchange-correlation functional.[197] The cut-off energy for the basis functions was set to 300 eV for AIMD and to 500 eV for the volume relaxations. The wave vectors from a 4×4×2 Monkhorst-Pack k-point grid were used for the Kohn-Sham wave functions. The energy convergence criterion for the self-consistency of the Kohn-Sham states in the DFT algorithm was 10^{-3} eV for the AIMD runs and 10^{-4} eV for the volume optimizations.

The density of the created a-GaO_x cells is 5.24 g/cm^3 for x = 1.5, 5.05 g/cm^3 for x = 1.2, 4.91 g/cm^3 for x = 1.0 and 5.66 g/cm^3 for x = 0.8. These values are within the wide range of 4.5–6.0 g/cm^3 that has been reported in experiments (cf. sec. 2.3). The decrease of the density from x = 1.5 to 1.0 is plausible from the effective loss of oxygen that is not fully compensated by the simultaneous shrinking of the DFT cells. The same trend has been found experimentally in [107]. The density's increase from x = 1.0 to 0.8 is the result from more efficient packing, with an about 6% smaller DFT cell for $Ga_{72}O_{58}$ than for $Ga_{64}O_{64}$, probably because the composition becomes closer to the other stoichiometric variety of gallium oxide, Ga_2O.

4.1.2 Density of States (DOS)

The density of states (DOS) of a-GaO_x was investigated by DFT with the amorphous model structures for $Ga_{64}O_{96}$ (x = 1.5), $Ga_{64}O_{78}$ (x = 1.2), $Ga_{64}O_{64}$ (x = 1.0) and $Ga_{72}O_{58}$ (x = 0.8). The VASP software [191–194] was employed also for the calculations of the DOS,[72] using the same PAW GW pseudopotentials [195, 196] and the GGA-PBE exchange-correlation functional [197] like in the preparation of the amorphous cells.[73] The cut-off energy and the energy convergence condition were set to 500 eV and 10^{-4} eV, respectively, and Gaussian smearing with a width of 0.05 eV was used. For each of the structures, the DOS was calculated both with 8×8×4 and with 16×16×8 Monkhorst-Pack k-points in order to evaluate the convergence with regard to the k-point density.[74] The difference between the DOS for the two k-point sets is minimal (fig. B.14), indicating

[72] Major VASP settings: IBRION = –1, ISIF = 2, PREC = normal, LORBIT = 11.

[73] Calculation of the DOS by hybrid DFT with the HSE06 functional was also tried extensively. This variant of DFT typically has a higher accuracy of the electronic structure, especially for the band gap, but at the expense of a drastically increased computational demand. Unfortunately, the respective calculations either did not converge or failed with cryptic errors.

[74] The original data from F. Draber only comprised GGA-DOS calculations with 16×16×8 k-points for x = 1.0 and with 8×8×4 k-points for the other compositions, without further analysis. Everything beyond that was done anew here in this work.

that 16×16×8 k-points are sufficient. All further analysis and later discussion refers to the DFT-DOS obtained for 16×16×8 k-points.

The position of the Fermi level E_F is of course crucial for the interpretation of the obtained DOS $g(E)$. Since the Fermi-Dirac distribution depends on the temperature (eq. 2.16), so does the Fermi level. Therefore, the E_F for $T = 0\,K$ that was obtained from VASP was adjusted to room temperature. This was achieved with a Python script by numerically solving the constant electron density condition

$$\int_{-\infty}^{E_F(0\,K)} g(E)\,dE = \int_{-\infty}^{\infty} g(E)\left[\exp\left(\frac{E-E_F}{k_B T}\right)+1\right]^{-1} dE \tag{4.4}$$

for E_F with T set to 300 K. This approach assumes that the DOS $g(E)$ is approximately independent of the temperature. The shift from $E_F(0\,K)$ to $E_F(300\,K)$ is negligible for a-$GaO_{0.8}$ (+0.6 meV) and a-$GaO_{1.0}$ (−3 meV), but more relevant for a-$GaO_{1.2}$ (+47 meV) and a-$GaO_{1.5}$ (+179 meV). In a similar way, the Fermi level was also calculated for other temperatures (e.g., for 150–2000 K).

For the comparison of the DOS for the different O/Ga ratios, a common energy reference is beneficial. Since the DOS all exhibit a steep edge at the top of the valence band, the valence band maximum (VBM) was chosen for this. Its position E_{VBM} was determined by linearly extrapolating the steep VB edge to the baseline.

Besides the DOS calculations, DFT was also used to assess the relative permittivity ϵ_r of a-GaO_x (cf. eq. 2.13). Several different VASP parameter sets were tried. The standard procedure using density function perturbation theory, either with self-consistent or with finite differences optimization,[75] only worked for a-$Ga_{64}O_{96}$ ($x = 1.5$). The resulting macroscopic static dielectric tensor was equal for 4×4×2 and 8×8×4 k-points. The calculation of the response to finite electric fields[76] gave reasonable ϵ_r for a-$Ga_{64}O_{96}$ ($x = 1.5$) and a-$Ga_{64}O_{78}$ ($x = 1.2$), but failed for the two more oxygen-deficient cells with the notification that the material must be insulating and that this does not seem to be the case. Including ionic motion in any of the methods did not work because the underlying phonon calculations did not converge for the amorphous structures.[77]

4.1.3 Localization of States (LOS)

For the interpretation of the calculated DOS with regard to electron transport, it is important to discern whether the electronic states are extended or localized. Therefore, in a first step, the partial

[75] Major VASP settings: LEPSILON = .TRUE., LRPA = .FALSE., IBRION = −1, optionally with LPEAD = .TRUE. for the finite differences approach.

[76] Major VASP settings: LCALCEPS = .TRUE., EFIELD_PEAD = 0.001 0.001 0.001 (in eV/Å), LRPA = .FALSE., IBRION = −1.

[77] Major VASP settings: IBRION = 8 for phonon calculation with perturbation theory and IBRION = 5, NFREE = 2, POTIM = 0.015 (in Å) for phonons from finite differences.

charge density of exemplary Kohn-Sham states was visualized with the VESTA software.[78] As will be discussed in sec. 5.4.2 in detail, they exhibit either rather homogeneously distributed charge density or a concentration in certain regions of the cell. Based on this observation, a metric for the localization of states (LOS) was introduced that will be explained in the following.

Fig. 4.1 shows the charge density $\rho(\vec{r}) = -e|\phi(\vec{r})|^2$ for an electronic state near the Fermi level in a-$Ga_{64}O_{64}$ ($x = 1.0$). In this example, the charge density appears localized. This is visible especially in the direction of the long c-axis of the tetragonal DFT cell. When separating the DFT cell into

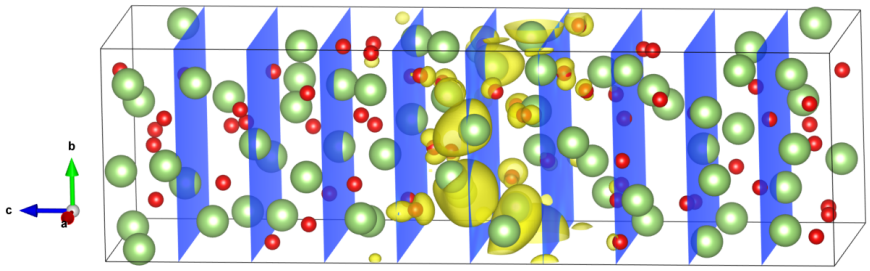


Figure 4.1: DFT cell of a-$Ga_{64}O_{64}$ ($x = 1.0$) with the charge density $\rho(\vec{r}) = -e|\phi_{nk}(\vec{r})|^2$ of a Kohn-Sham state ϕ_{nk} with $E_{nk} - E_F(300\,\mathrm{K}) \approx 0.05\,\mathrm{eV}$. The gallium and oxygen atoms are shown in green and red, respectively, and the charge density is visualized for an isovalue of $\rho = -0.17\,\mathrm{C/cm^3}$. The blue planes indicate the cell slices used in the LOS analysis.

N regular slices along the c-direction, as indicated in fig. 4.1, then some contain a high and some a low percentage of the total charge density. In an extended state, in contrast, the charge density is rather evenly distributed across the DFT cell. The standard deviation of the charge density among the N slices is thus a measure for the localization of the regarded state. This is the essence of the LOS analysis. The approach is quite comparable to the inverse participation ratio (IPR) mentioned in sec. 2.3.1, but it is easier to implement. Instead of explicitly integrating the charge density, the site-projected wave function characters $|\langle Y^a_{lm}|\phi_{nk}\rangle|^2$ provided by VASP were used.[79] They are the projections of each Kohn-Sham wave function ϕ_{nk} on the different spherical harmonics Y^a_{lm} with angular momentum quantum l and magnetic quantum m, for each atom a in the DFT cell. Taking their sum over l and m yields a mapping of the wave function on the atoms, which can be expressed as the relative projection

$$P^a_{nk} = \sum_{l,m} |\langle Y^a_{lm}|\phi_{nk}\rangle|^2 \Big/ \sum_{a,l,m} |\langle Y^a_{lm}|\phi_{nk}\rangle|^2 \tag{4.5}$$

The sum of the P^a_{nk} for the atoms $a \in \mathbb{G}_s$, where $\mathbb{G}_s$ is the group of atoms in a certain slice s of the DFT cell, is then a measure for the relative amount of charge density within this slice. The respective

[78] The required PARCHG files, which contain the charge density data, were generated from the WAVECAR files of the DOS calculations through additional VASP runs with LPARD = .TRUE., ISTART = 1, LSEPB = .TRUE. and LSEPK = .TRUE., with state selection by the EINT and KPUSE parameters.

[79] VASP DOS calculations with LORBIT = 11 yield these projections in the PROCAR file.

standard deviation

$$\sigma_{\mathrm{L},nk} = \sqrt{\frac{1}{N}\sum_{s=1}^{N}\left(\sum_{a\in\mathbb{G}_s} P_{nk}^{a} - \frac{1}{N}\right)^2} \tag{4.6}$$

between the slices s thus quantifies the spatial inhomogeneity of the wave function ϕ_{nk}. Localized states have high σ_L, while a low σ_L indicates that a state is extended. The grouping of the projections into slices is more appropriate than taking the standard deviation across all individual atoms, because, as fig. 4.1 shows, the charge density is not necessarily centered at a certain atom.

The calculation of the localization standard deviation σ_L for the thousands of electronic states was done with a Python script. The resulting LOS plots of $\sigma_{\mathrm{L},nk}$ vs. E_{nk} were compared with the DOS to identify the regions of localized and extended states. The lower envelope of $\sigma_\mathrm{L}(E)$ was used for this assessment, because the mobility gap is defined by the absence of extended states. The LOS analysis was carried out for N = 4, 6, 8 and 10 slices and the outcome was essentially equal except for a scaling on the σ_L-axis. The conclusions with regard to the localization of the states were not affected by the number of slices. The results shown later in sec. 5.4.2 are those for N = 10 slices.

4.2 Kinetic Monte Carlo (KMC)

The DFT calculations are valuable for the interpretation of the experiments, but a critical point remains far from obvious: Since Mott's law for VRH has been derived for $g(E)$ = const. (cf. sec. 2.2.3), what temperature-dependence of the VRH conductivity is expected for the more realistic DFT-DOS of a-GaO_x? For example, the analytic derivations for a parabolic or exponential DOS around E_F, which were mentioned in sec. 2.2.3, yielded an exponential dependence on $T^{-1/2}$ or T^{-1}, respectively, for the hopping conductivity. This demonstrates that the shape of the DOS near the Fermi level can have a significant influence. Since the interpretation of the conductivity measurements in this work relies heavily on the assessment of the T-dependence, it is important to derive the VRH conductivity $\sigma(T)$ from the obtained ab-initio DOS.

The kinetic Monte Carlo (KMC) method is well suited for the simulation of variable-range hopping, since it simulates the development of a system over time on the basis of transition probabilities. These probabilities are readily available for the VRH process through the Miller-Abrahams rate equation (eq. 2.46), which was introduced in sec. 2.2.3. Therefore, KMC is the primary method for the numerical investigation of 3-dimensional hopping transport. There are numerous 3D KMC-VRH simulations in the literature, as reviewed for example in [198], but of course none for a-GaO_x. These previous studies all used parametrized DOS models $g(E)$ with a mathematically defined E-dependence (e.g., constant, exponential, parabolic or Gaussian in [74, 198–203]). They essentially confirmed the respective analytical VRH derivations. Until now, however, to the best of the author's

knowledge, there have been no 3D KMC-VRH simulations with an ab-initio DOS of an amorphous material.

In order to simulate the VRH conductivity of a-GaO_x, based on the obtained ab-initio DOS, the new KMC software "VHoppAS" – short for "Variable-range Hopping in Amorphous Solids" – was developed. It is written in C++ and freely available under open-source license, either from the VHoppAS project on GitHub or in archived, citable form via [204]. These sources contain not only the program's code but also the manual and further documentation. VHoppAS can take any user-supplied DOS as input, not only idealized DOS models but also the DOS from ab-initio methods. This also means that it is not designed specifically for a-GaO_x, but can be used for simulating 3D-VRH in any material whose DOS is known. Moreover, VHoppAS employes the correct effective charge carrier density (cf. sec. 2.1.1) for the calculation of the mobility and the Fickian diffusion coefficient of the electrons, in contrast to the previous KMC studies that simply treated all electrons as charge carriers. It also respects that electrons are fermions with a spin, which typically has not been taken into account. Other advantages of VHoppAS are for example its speed, ease of use and the extensive statistics it provides (e.g., detailed histograms). These and further features, the underlying principles and the implementation of the KMC-VRH algorithm will be described in the following sections.

4.2.1 Concept of the KMC Simulations

As mentioned above, the KMC-VRH simulations are based on the Miller-Abrahams rate equation (eq. 2.46), which yields the transition rate ν_{ij} from an occupied localized state i to an unoccupied localized state j. The first step of each KMC run therefore is to create a simulation cell that contains a certain number of localized states with positions $\vec{r}_i$ and energies ε_i, with some of them occupied by electrons. How this was done, with the DFT-DOS as input, will be explained in sec. 4.2.2 and 4.2.3. Within this cell, the many possible hopping transitions do not occur with the exact frequencies ν_{ij}, of course, since the result of the Miller-Abrahams equation is just the average rate that is expected for each $i \to j$ path. Accordingly, its inverse, the time-to-hop expectancy value $\tau_{ij} = \nu_{ij}^{-1}$, is only the average time until the respective hop occurs. The actual time between successive $i \to j$ hops scatters around this average.

The KMC simulations use random numbers to reflect this variability and to decide how the system evolves over time. For each possible hopping path $i \to j$, a uniform random number $P \in [0,1)$ is used to calculate the randomized time-to-hop

$$\tau_{ij}^{*} = -\ln(P)\,\tau_{ij} \tag{4.7}$$

which it takes until the respective hop would occur next. Impossible transitions, i.e., when i is unoccupied or j is occupied, have $\tau_{ij}^{*} = \infty$. The required random numbers are obtained from a MT 19937

Mersenne-Twister pseudo-random number generator,[80] which is initialized with a different, unique seed for each individual simulation. Eq. 4.7 for τ_{ij}^* originates from the common assumption that the number of KMC transitions obeys the Poisson distribution of rare events, which means that the waiting time in-between follows the exponential distribution.[81] Each τ_{ij}^* describes how long it takes until the electron currently residing at state i would hop to state j. Therefore, the overall first hopping transition that takes place in the cell is the one with the lowest time-to-hop τ_{ij}^*. Moving forward on the time axis, the hops occur in the order of their τ_{ij}^*, but after each transition all affected paths – e.g., those newly available to the electron that just hopped – must be updated with new τ_{ij}^* from eq. 4.7. This is roughly the principle of the KMC algorithm, which was adopted from [203].[82] A more elaborate description follows in sec. 4.2.4. According to the law of large numbers, the resulting electron movement resembles the Miller-Abrahams rates and thus becomes physically meaningful after a sufficiently high number of hops.

The displacement of each electron is traced throughout the whole simulation. This is an important difference to the situation in experiments, where the electrons are indistinguishable and where just the overall flow of electrons is measured, without information about how much an individual electron contributed to the flux. The theory of tracer statistics, which is explained in condensed form in sec. A.4, describes how the conductivity σ, the mobility u and the Fickian diffusion coefficient D can be calculated from the displacements. This requires that the KMC cell is macroscopically homogeneous and time-invariant with periodic boundary conditions. The tracer theory was adapted to electrons, using the effective charge carrier model introduced in sec. 2.1.1, to account for the fact that not all electrons contribute equally to the macroscopic observables. Rather trivially, the total displacement of all electrons in the direction of an applied electric field can be used to calculate the conductivity. For the per-particle properties u and D, however, the effective charge carrier density n_{eff} must be used to obtain well-defined values. This seems to be a new aspect for KMC simulations of variable-range hopping in general, since it was not found in any respective literature. The final equations that were used for the calculation of the simulation results will be described in sec. 4.2.5. Further details, like the detection of oscillations or the optimization of the KMC algorithm, will be explained in sec. 4.2.6.

4.2.2 Physical and Technical Parameters

The KMC simulations are controlled by a number of input values, which can be divided into two categories: "Physical" parameters represent actual physical properties of the material and thus have

[80] Random number generator: std::mt19937_64, combined with std::uniform_real_distribution for probabilities and state positions, std::piecewise_linear_distribution for state energies and std::uniform_int_distribution for state or electron selection (C++ standard library).

[81] Eq. 4.7 results from the conversion of a uniform random number into an exponentially distributed random number.

[82] Only the basic principle of the time-based algorithm was adopted from [203], where the VRH-diffusion of 1–100 electrons has been simulated for an exponential DOS (with a fixed total state density of $1\,\mathrm{nm}^{-3}$ and without electric field). Everything else concerning the KMC-VRH simulations was newly developed in this work.

an influence on the finally obtained quantities like the conductivity. "Technical" parameters, in contrast, are KMC-specific settings that must not affect the simulation results. However, they must be adjusted and validated to ensure that the simulations are correct. All parameters are spatially and temporally constant, of course.

The following list gives an overview of the physical input parameters:

- Density of localized states: $g(\varepsilon)$
- Attempt frequency: ν_0 (default: $3.073{\times}10^{15}$ Hz)
- Localization radius: α (default: 0.1512 nm)
- Material temperature: T (default: 300 K)
- Electric field strength: F (default: 0.0635 MV/cm)
- Relative chemical potential: μ (default: 0 eV)

The default values were used in the simulations unless otherwise stated. Most of these parameters are part of the Miller-Abrahams rate equation (eq. 2.46). The frequency ν_0 of the hop attempts, the localization radius α of the electronic states and the phonon temperature T enter the calculation of the time-to-hop expectancy values τ_{ij} directly. The electric field strength F defines the constant electric field $\vec{F}$, which – with the three spatial axes of the KMC cell denoted by x, y and z – is oriented in $-x$-direction according to

$$\vec{F} = -\nabla\varphi = \begin{pmatrix} -F \\ 0 \\ 0 \end{pmatrix} \tag{4.8}$$

Including this definition in the Miller-Abrahams equation yields for the electric field contribution $e\vec{r}_{ij}\vec{F} = -ex_{ij}F$, which facilitates hops in $+x$-direction for $F = |\vec{F}| > 0$. Synonymously, the electrochemical potential of the electrons, i.e., the Fermi level, varies across the KMC cell with

$$E_\mathrm{F}(x) = \tilde{\mu}(x) = \mu - exF \tag{4.9}$$

Note that VHoppAS also allows $F = 0$, i.e., simulations with pure diffusion, which was used for validation purposes.

The only two physical parameters that are not directly linked to the Miller-Abrahams equation are the DOS $g(\varepsilon)$ and the chemical potential μ of the electrons. Instead, they control the spatial and energetic distribution of the states in the KMC cell and the number of electrons placed therein, as will be explained in the next section. The DFT-DOS of a-$GaO_{0.8}$, a-$GaO_{1.0}$ or a-$GaO_{1.2}$ (cf. sec. 4.1) were supplied as parameter $g(\varepsilon)$ to the KMC simulations. Their energy axis ε refers to the energy of the electronic states without the potential energy from the electric field $\vec{F}$, like introduced in sec. 2.1.1 and 2.2.3. It is defined relative to the Fermi level of the charge-neutral material, i.e., the

Fermi level for the natural amount of valence electrons, which is located at $\varepsilon = 0$. For this energy reference, VHoppAS automatically adjusts the charge-neutral Fermi level obtained from DFT to the simulation temperature T, essentially with the same approach as described sec. 4.1.2 (eq. 4.4 solved with a fail-safe bisection search algorithm). Using this F-independent energy axis for the localized states is beneficial, since the total state energy

$$E(x) = \varepsilon - exF \tag{4.10}$$

varies in x-direction and depends on whether a hop occurs through a periodic boundary of the KMC cell. In contrast to $E(x)$, ε uniquely defines a state's energy in relation to the DOS – irregardless of the state's position. Congruently, from eqs. 4.9 and 4.10 follows that the energy difference of an electronic state to the Fermi level is spatially constant and given by

$$E(x) - E_\mathrm{F}(x) = \varepsilon - \mu \tag{4.11}$$

This means that the chemical potential μ describes the position of the Fermi level on the ε-axis. Since ε is defined relative to the charge-neutral Fermi level at $\varepsilon = 0$, so is the physical parameter μ.[83] For $\mu = 0$, the actual Fermi level in the simulation corresponds to the charge-neutral Fermi level and the KMC cell is populated with the natural amount of electrons. When $\mu \neq 0$ is supplied as input, then the Fermi level is shifted to $\varepsilon = \mu$ and a charged material with either positive or negative charge density is simulated. This feature was used to investigate how the VRH conductivity changes in a space-charge zone like a Schottky barrier.

The following list gives an overview of the major technical parameters:

- DOS energy range: $[\varepsilon_\mathrm{min}, \varepsilon_\mathrm{max}]$ (default: $[-0.5, +0.5]$ eV)
- Number of localized states: N_st (default: 1.2×10^5)
- Spatial cut-off for paths: r_cut (typically 1.7–2.2 nm)
- Energetic cut-off for paths: ε_cut (typically 0.3–0.8 eV)
- Number of hops during pre-equilibration: $N_\mathrm{h,pre}$ (default: 2×10^7)
- Number of hops during equilibration: $N_\mathrm{h,eq}$ (default: 3×10^8)
- Number of hops during main simulation: $N_\mathrm{h,sim}$ (default: 3×10^8)

Further parameters, like switches to enable or disable certain features of VHoppAS, are described in the documentation (cf. sec. 4.2). The energy range $[\varepsilon_\mathrm{min}, \varepsilon_\mathrm{max}]$ defines which region of the DOS is regarded in the simulation. It must be sufficiently wide around the Fermi level at $\varepsilon = \mu$ to include all states that contribute to VRH transport. The electronic states in this energy range are assumed to be localized. The parameter N_st specifies the total number of localized states that are to be placed

[83] Strictly speaking, the parameter μ therefore does not correspond to an absolute chemical potential as defined in eq. 2.2 but to a relative chemical potential $\mu - \mu_0$, which is nonetheless written as "μ" because it only occurs in the context of "$\varepsilon - \mu$".

in the KMC cell. Too low N_{st} yields poor statistics, both with regard to the sampling of the DOS and the number of electrons, while too high N_{st} increases the computational demand and requires longer simulations to obtain meaningful displacements. The path cut-offs r_{cut} and $\varepsilon_{\mathrm{cut}}$ determine which hopping paths are considered in the simulation.[84] Since the Miller-Abrahams rate decreases exponentially with increasing distance $|\vec{r}_{ij}|$ and energy difference ε_{ij}, it is not necessary to include all possible paths. Instead, only those with $|\vec{r}_{ij}| < r_{\mathrm{cut}}$ and $|\varepsilon_{ij}| < \varepsilon_{\mathrm{cut}}$ are taken into account.[85] For each parameter set employed in the simulations, both cut-offs were adjusted to not exclude any relevant paths. The parameters $N_{\mathrm{h,pre}}$, $N_{\mathrm{h,eq}}$ and $N_{\mathrm{h,sim}}$ are the total hop counts of the respective simulation stages, which will be explained in sec. 4.2.4.

The exact meaning of the technical parameters will also become clearer in the next sections. In many situations, other values than the defaults were used. For example, the energy range $[-0.5, +0.5]$ eV was shifted accordingly when the chemical potential μ was varied. In the simulations for a-$GaO_{1.2}$, where E_{F} is in a small band gap, a wider energy range of $[-0.8, +0.8]$ eV had to be used, together with $\varepsilon_{\mathrm{cut}} = 1\,\mathrm{eV}$, $N_{\mathrm{st}} = 1.6{\times}10^5$, $N_{\mathrm{h,eq}} = 0$ and $N_{\mathrm{h,sim}} = 6{\times}10^9$ to $60{\times}10^9$ (depending on the temperature). In simulations with high electric field strength F, the DOS energy range and the cut-offs had to be enlarged successively to account for the increasing $ex_{ij}F$ term in the rate equation. These adjustments are the result of the parameter validation that will be discussed in sec. 5.5.1.

4.2.3 Setup of the KMC Cell

Prior to the actual hopping algorithm, the KMC cell must be prepared. This involves the following steps:

- Generate the localized states
- Find all hopping paths within the cut-offs
- Calculate the time-to-hop expectancy values
- Populate the localized states with electrons
- Initialize the electrons with randomized hop times

The KMC cell is cubic and has periodic boundary conditions. Its volume V_{cell} follows directly from the specified total number of states N_{st} and the DOS energy range $[\varepsilon_{\mathrm{min}}, \varepsilon_{\mathrm{max}}]$:

$$V_{\mathrm{cell}} = N_{\mathrm{st}} \left[\int_{\varepsilon_{\mathrm{min}}}^{\varepsilon_{\mathrm{max}}} g(\varepsilon)\, \mathrm{d}\varepsilon \right]^{-1} \tag{4.12}$$

[84] As an alternative to specifying r_{cut} manually, VHoppAS also provides the option to set the minimum number of outgoing paths from each state, which is then used internally to find a suitable r_{cut} value.

[85] The energy cut-off condition also excludes paths with $\varepsilon_{ij} < -\varepsilon_{\mathrm{cut}}$, even though they are associated with a high Miller-Abrahams rate. The reason for this is that hops with $\varepsilon_{ij} \ll 0$ can only occur on a regular basis when equivalent hops with $\varepsilon_{ij} \gg 0$ take place as well.

The edge length d_{cell} of the KMC cell, henceforth called "cell size", is thus given by

$$d_{\text{cell}} = V_{\text{cell}}^{1/3} = N_{\text{st}}^{1/3} \left[\int_{\varepsilon_{\min}}^{\varepsilon_{\max}} g(\varepsilon)\, \mathrm{d}\varepsilon \right]^{-1/3} \tag{4.13}$$

This is the first important step to use an arbitrary DOS in the VRH simulations, since it defines the scale for the distances between the localized states. The resulting cell size typically was in the range of 20–25 nm.

In principle, this KMC cell now must be filled with localized states that have a randomized position and energy. It would be wrong, though, to simply place N_{st} different states in the cell, because the electrons' spin must be considered. Like the DFT calculations, the KMC-VRH simulations with VHoppAS are based on the assumption of spin-degeneracy, which means that localized states do not differ depending on the electron spin. The supplied DOS $g(\varepsilon)$ is the sum of the spin-up and the spin-down DOS, which are equal. Likewise, N_{st} is the sum of spin-up and spin-down single-electron states. In consequence, the N_{st} localized states would have to have pairwise equal position and energy. However, with the assumption that spin-up and spin-down electrons behave equally, it is sufficient to simulate only one spin-type and take it as representative for the other. VHoppAS therefore internally places only half of the states and half of the electrons in the KMC cell and counts each of them twice for the statistics and the results.[86] This handling of the spin is responsible for the occasional factors 1/2 or 2 that will appear in the remainder of this sec. 4.2. The respective internal quantities will be symbolized by a circumflex accent (e.g., $\hat{N}_{\text{st}} = N_{\text{st}}/2$ for the internal number of states). By convention, all input and output values of VHoppAS refer to both spin-types together. The same applies to all KMC-related quantities and graphs in the later results chapter 5.

The setup of the KMC cell thus commences with the generation of $\hat{N}_{\text{st}} = N_{\text{st}}/2$ localized states. Each state i is assigned a random position $\vec{r}_i$ with uniformly random coordinates $x_i, y_i, z_i \in [0, d_{\text{cell}})$ and a DOS-weighted random energy $\varepsilon_i \in [\varepsilon_{\min}, \varepsilon_{\max}]$. In order to reflect the supplied DFT-DOS in the distribution of the state energies, the ε_i are assigned randomly with the normalized DOS

$$g(\varepsilon_i) \left[\int_{\varepsilon_{\min}}^{\varepsilon_{\max}} g(\varepsilon)\, \mathrm{d}\varepsilon \right]^{-1} \tag{4.14}$$

as distribution function. This is the key step to achieve KMC-VRH simulations with an ab-initio DOS.

Afterwards, for each localized state i the paths to all other states j are screened to find those that satisfy the cut-off conditions. The hopping vector $\vec{r}_{ij} = \vec{r}_j - \vec{r}_i$ for each path $i \to j$ is calculated as the shortest connection from i to j, either within the KMC cell or through any of its periodic boundaries. The energy difference between the states is given by $\varepsilon_{ij} = \varepsilon_j - \varepsilon_i$. All outgoing paths

[86] A workaround for simulating materials without spin-degeneracy can be found in the documentation of VHoppAS.

with $|\vec{r}_{ij}| < r_{\text{cut}}$ and $|\varepsilon_{ij}| < \varepsilon_{\text{cut}}$ are assigned to the state i as valid hopping paths. Note that this procedure guarantees that each path can be traversed in both directions. It is checked whether each localized state has at least two outgoing paths in order to prevent isolated states or dead ends. If this is not the case, then the cut-offs are automatically increased until this auxiliary condition is met.

Since the positions and energies of the localized states do not change during the simulation, it is possible to calculate the time-to-hop expectancy values τ_{ij} of all paths beforehand. With $\vec{r}_{ij}$ and ε_{ij} as defined above, they are obtained as the inverse of the Miller-Abrahams rate equation (eq. 2.46) according to

$$\tau_{ij} = \nu_{ij}^{-1} = \begin{cases} \nu_0^{-1} \exp\left(\frac{2|\vec{r}_{ij}|}{\alpha}\right) \exp\left(\frac{\varepsilon_{ij} - e x_{ij} F}{k_\text{B} T}\right) & \text{for } \varepsilon_{ij} - e x_{ij} F > 0 \\ \nu_0^{-1} \exp\left(\frac{2|\vec{r}_{ij}|}{\alpha}\right) & \text{for } \varepsilon_{ij} - e x_{ij} F \leq 0 \end{cases} \tag{4.15}$$

As explained in sec. 4.2.1, the τ_{ij} correspond to the average time it takes until an electron on state i would hop to the unoccupied state j.

After the creation of the localized states and the paths between them, the next step is to populate the states with electrons. The number of electrons in the KMC cell is controlled by the relative chemical potential μ, with

$$\hat{N} = \sum_{i=1}^{\hat{N}_\text{st}} f_\text{FD}(\varepsilon_i) = \sum_{i=1}^{\hat{N}_\text{st}} \left[\exp\left(\frac{\varepsilon_i - \mu}{k_\text{B} T}\right) + 1 \right]^{-1} \tag{4.16}$$

as the internal and $N = 2\hat{N}$ as the total number of simulated electrons. The $\hat{N}$ electrons can be distributed among the localized states by two different methods, either according to the Fermi-Dirac distribution or as a step function. In the first case, an electron is placed on each state i if a uniform random number $P \in [0,1)$ is smaller than the respective Fermi-Dirac probability $f_\text{FD}(\varepsilon_i)$. If the resulting number of electrons in the cell differs from $\hat{N}$, then randomly chosen states are re-evaluated with $P < f_\text{FD}(\varepsilon_i)$ until $\hat{N}$ is reached. In the second case, the localized states are simply filled with $\hat{N}$ electrons from lowest to highest state energy, which results in a step function as initial electron distribution. In this work, most simulations were started with Fermi-Dirac-distributed electrons to save equilibration time. The second distribution method was used to verify whether the electrons also distribute according to $f_\text{FD}(\varepsilon_i)$ on their own over the course of the hopping simulations.

The final step of the preparations is the initialization of the randomized hop times τ_{ij}^*, which were introduced in sec. 4.2.1. Each electron's possible hopping paths are known from the state it occupies. For each of these outgoing paths, either a randomized time-to-hop τ_{ij}^* is calculated from eq. 4.7 when the target state is unoccupied, or τ_{ij}^* is set to ∞ when the state j is currently occupied by another electron. The shortest time-to-hop τ_{ij}^* for each electron k is denoted as $\tau_{\text{min},k}$. It represents the path that this electron would take next, after the time $\tau_{\text{min},k}$ has passed, unless a different electron hops to the respective target state before and blocks this path.

4.2.4 KMC Algorithm

After the above preparations, the movement of the electrons is simulated. The most important quantities that are recorded during the KMC simulation are the displacement vectors (X_k, Y_k, Z_k) of the electrons and the simulated time span τ, which describes the progression of time during the hopping process. They are initialized to zero at the start of the simulation. Various other statistics are recorded as well (e.g., how often a certain path is used or the number of hops of each electron), but including them here would distract from the algorithm. An overview can be found in the user manual of VHoppAS.

The following pseudo-code explains the principle of the employed KMC algorithm:

1. Find the electron that hops next, i.e., the electron with overall lowest time $\tau_{\text{hop}} = \min(\tau_{\text{min},k})$
2. Execute the hop that belongs to τ_{hop}:
   - Transfer the electron to the target state
   - Add the path vector to the electron's displacement vector
3. Advance the time by τ_{hop}:
   - Add τ_{hop} to the simulated time span τ
   - Subtract τ_{hop} from all other τ_{ij}^* (incl. all $\tau_{\text{min},k}$)
4. Update all affected electrons:
   - Electron that just hopped: Assign new τ_{ij}^* for all paths outgoing from its new location
   - Electrons with a path to the hop's initial state (not blocked anymore): Assign new τ_{ij}^* for the respective path
   - Electrons with a path to the hop's target state (now blocked): Set $\tau_{ij}^* = \infty$ for the respective path
   - Find the new $\tau_{\text{min},k}$ of these affected electrons (if necessary)
5. Repeat from 1. until the specified number of hops is reached

The actual KMC algorithm implemented in VHoppAS is more complicated due to several optimizations, which will be described separately in sec. 4.2.6.

Each simulation is structured into three sequential stages. In each of them, the electrons move according to the above algorithm, but the stages differ in the post-processing of the resulting quantities:

1. Pre-equilibration: The KMC algorithm is executed until $\hat{N}_{\text{h,pre}} = N_{\text{h,pre}}/2$ hops are reached. Then, all recorded statistics like the electron displacements and the simulated time span are reset to zero. The purpose of this simulation stage is that the electrons can redistribute into their steady-state distribution with regard to energy.

2. Equilibration: The KMC algorithm is executed for $\hat{N}_{\mathrm{h,eq}} = N_{\mathrm{h,eq}}/2$ more hops, but afterwards no quantities are reset.[87] This serves the purpose that the electron displacements X_k, Y_k and Z_k can reach continuously developing distributions, which is the basis for linear displacement vs. time relations in the following main simulation. At the end of this equilibration stage, the resulting time span τ and the displacement averages that are defined in the next sec. 4.2.5 are saved as reference values for the final stage.
3. Main simulation: The KMC algorithm is finally executed for $\hat{N}_{\mathrm{h,sim}} = N_{\mathrm{h,sim}}/2$ more hops. From the electrons' movement during this stage, relative to their displacement at the end of the previous stage, the simulation results are calculated as described in the next sec. 4.2.5.

The first two stages are optional and can be omitted by setting $N_{\mathrm{h,pre}} = 0$ or $N_{\mathrm{h,eq}} = 0$, respectively.

4.2.5 Calculation of the Results

The basis for the calculation of the KMC results is the tracer theory, which is explained in sec. A.4. It requires that only those electrons are considered as tracers which actually moved via drift and diffusion. Therefore, before any results can be calculated, a closer examination of each electron's behavior over the course of the simulation is required. Electrons which did not hop at all are obviously no suitable tracers. Furthermore, it is noticeable that typically a large proportion of each electron's hops are oscillations where the electron moves back-and-forth between two localized states. VHoppAS identifies such oscillations by keeping track of each electron's previous location. Each time an electron hops from one state to another and its next hop brings it back to the previous state, this counts as an oscillation (even when other electrons moved in the meantime). Both of these hops then are called "oscillatory hops". They cancel each other with regard to the displacement of the electron. Regular hops which are not part of an oscillation are called "non-oscillatory hops". In consequence, the electrons can be distinguished into three major categories:

- "stationary": The electron has zero hops.
- "oscillating": The electron was not stationary but has only oscillatory hops.
- "mobile": The electron was neither stationary nor purely oscillating, which means it has a non-zero number of non-oscillatory hops.

Only the mobile electrons in the KMC cell, which are counted as $\hat{N}_{\mathrm{mob}}$, are considered as tracers for the subsequent evaluation of the KMC simulation. Since each electron counts twice, the total number of mobile electrons is $N_{\mathrm{mob}} = 2\hat{N}_{\mathrm{mob}}$, which corresponds to the number of tracers N^* in sec. A.4. Note that both the stationary and the oscillating electrons always have $\vec{R}_k = 0$, but occasionally a mobile electron might have zero displacement as well.

[87] A minor consequence of no reset between equilibration and main simulation is that the KMC histograms shown later refer to equilibration and main simulation together.

Next, the corresponding number of effective charge carriers N_{eff}^{*} must be determined. It equals the total number of effective charge carriers N_{eff}, since only electrons without any contribution to charge transport are excluded from N^{*}. According to eq. 2.23, N_{eff} depends on the chemical potential and the temperature of the electrons, which normally correspond to the relative chemical potential μ and the material temperature T, respectively. The later results in sec. 5.5.4 will demonstrate, however, that at high electric field strength the electrons adopt a different Fermi-Dirac-distribution with a higher effective temperature $T_{\mathrm{eff}} > T$. Due to the temperature-dependance of the Fermi level, this also causes a differing effective chemical potential $\mu_{\mathrm{eff}} \neq \mu$ of the electrons. These effective values must be used for the calculation of N_{eff}, because the derivation of n_{eff} in sec. 2.1.1 is based on the electrons' actual distribution. VHoppAS determines T_{eff} and μ_{eff} by fitting the steady-state electron distribution $\{\varepsilon_k\}$ with the Fermi-Dirac function[88]

$$f_{\mathrm{FD}}(\varepsilon, \mu_{\mathrm{eff}}, T_{\mathrm{eff}}) = \left[\exp\left(\frac{\varepsilon - \mu_{\mathrm{eff}}}{k_{\mathrm{B}} T_{\mathrm{eff}}}\right) + 1\right]^{-1} \tag{4.17}$$

For simulations with a weak electric field, this fitting procedure is omitted and $T_{\mathrm{eff}} = T$ and $\mu_{\mathrm{eff}} = \mu$ are used. With the above distribution function follows from eq. 2.23 the total number of effective charge carriers

$$N_{\mathrm{eff}} = 2\hat{N}_{\mathrm{eff}} = 2\sum_{i}^{\hat{N}_{\mathrm{st}}} f_{\mathrm{FD}}(\varepsilon_i, \mu_{\mathrm{eff}}, T_{\mathrm{eff}})\left[1 - f_{\mathrm{FD}}(\varepsilon_i, \mu_{\mathrm{eff}}, T_{\mathrm{eff}})\right] \tag{4.18}$$

and the effective charge carrier density $n_{\mathrm{eff}} = N_{\mathrm{eff}}/V_{\mathrm{cell}}$ in the simulation.

After determining N_{mob} and N_{eff}, the displacement averages are calculated. The collective motion of the mobile electrons is re-interpreted as the movement of effective charge carriers, as derived in sec. A.4. The displacement vector (X_k, Y_k, Z_k) of each internal mobile electron k is counted twice to account for the electrons' spin. The mean displacement $\langle X \rangle$, the mean squared displacement $\langle X^2 \rangle$ and the displacement variance $\mathrm{Var}(X)$ in x-direction are thus given by

$$\langle X \rangle = \frac{2}{N_{\mathrm{eff}}} \sum_{k=1}^{\hat{N}_{\mathrm{mob}}} X_k = \frac{1}{\hat{N}_{\mathrm{eff}}} \sum_{k=1}^{\hat{N}_{\mathrm{mob}}} X_k \,, \quad \langle X^2 \rangle = \frac{2}{N_{\mathrm{eff}}} \sum_{k=1}^{\hat{N}_{\mathrm{mob}}} X_k^2 = \frac{1}{\hat{N}_{\mathrm{eff}}} \sum_{k=1}^{\hat{N}_{\mathrm{mob}}} X_k^2 \tag{4.19}$$

$$\text{and} \quad \mathrm{Var}(X) = \langle X^2 \rangle - \frac{N_{\mathrm{eff}}}{N_{\mathrm{mob}}} \langle X \rangle^2 = \langle X^2 \rangle - \frac{\hat{N}_{\mathrm{eff}}}{\hat{N}_{\mathrm{mob}}} \langle X \rangle^2 \tag{4.20}$$

The respective quantities in y- and z-direction are obtained analogously. The values of these displacement averages and the elapsed time τ at the end of the equilibration stage (denoted by "eq") are taken as reference for the final simulation stage. This leads to the relative values for the main

[88] This fitting procedure again employes the bisection search algorithm with guaranteed convergence for finding the appropriate T_{eff}. This is the only degree of freedom for this fit, because the relation between T_{eff} and μ_{eff} is implicitly known from the DOS: For each trial temperature, the respective chemical potential is determined with the same method as the aforementioned T-adjustment of the Fermi level.

simulation (denoted by "sim")

$$\langle X\rangle_{\mathrm{sim}} = \langle X\rangle - \frac{N_{\mathrm{eff}}^{\mathrm{eq}}}{N_{\mathrm{eff}}}\langle X\rangle_{\mathrm{eq}}\,, \quad \langle X^2\rangle_{\mathrm{sim}} = \langle X^2\rangle - \frac{N_{\mathrm{eff}}^{\mathrm{eq}}}{N_{\mathrm{eff}}}\langle X^2\rangle_{\mathrm{eq}}\,, \tag{4.21}$$

$$\mathrm{Var}(X)_{\mathrm{sim}} = \mathrm{Var}(X) - \frac{N_{\mathrm{eff}}^{\mathrm{eq}}}{N_{\mathrm{eff}}}\mathrm{Var}(X)_{\mathrm{eq}} \quad \text{and} \quad \tau_{\mathrm{sim}} = \tau - \tau_{\mathrm{eq}} \tag{4.22}$$

Again, analog equations apply to the y- and z-directions. $N_{\mathrm{eff}}^{\mathrm{eq}}$ and N_{eff} are the effective charge carriers determined at the end of the equilibration and the main simulation, respectively. At high electric field strength, they may differ slightly due to fluctuations in the electron distribution and their ratio just serves as correction factor to adjust the "eq"-averages to the same N_{eff}. In the weak-field regime, this is not necessary since always $T_{\mathrm{eff}} = T$ and $\mu_{\mathrm{eff}} = \mu$. When the equilibration is disabled ($N_{\mathrm{h,eq}} = 0$), then the above quantities for the main simulation simply correspond to the total values (e.g., $\langle X\rangle_{\mathrm{sim}} = \langle X\rangle$).

Finally, the major result values are calculated. According to eqs. A.45 and A.49, the conductivity σ and the mobility u are given by

$$\sigma = e n_{\mathrm{eff}} u \quad \text{and} \quad u = \frac{\langle X\rangle_{\mathrm{sim}}}{F\tau_{\mathrm{sim}}} \tag{4.23}$$

From eqs. A.31 and A.46 follows for the Fickian diffusion coefficients

$$D = D_{yz} = \frac{\langle Y^2\rangle_{\mathrm{sim}} + \langle Z^2\rangle_{\mathrm{sim}}}{4\tau_{\mathrm{sim}}} \quad \text{and} \quad D_x = \frac{\mathrm{Var}(X)_{\mathrm{sim}}}{2\tau_{\mathrm{sim}}} \tag{4.24}$$

The later analysis in sec. 5.5.1 will reveal that the true parallel Fickian diffusion coefficient D_x cannot be reliably determined in this way. Therefore, the transverse Fickian diffusion coefficient D_{yz} alone represents the isotropic Fickian diffusion coefficient D. However, when no electric field is employed in the simulation ($F = 0$), then D is given by

$$D = D_{xyz} = \frac{\langle X^2\rangle_{\mathrm{sim}} + \langle Y^2\rangle_{\mathrm{sim}} + \langle Z^2\rangle_{\mathrm{sim}}}{6\tau_{\mathrm{sim}}} \tag{4.25}$$

Another special case that will be justified later is the calculation of D_{yz} at high electric field strength. Due to random imbalances between the available hopping paths, it must be determined from the displacements variances $\mathrm{Var}(Y)$ and $\mathrm{Var}(Z)$ when F exceeds the weak-field region:

$$D = D_{yz} = \frac{\mathrm{Var}(Y)_{\mathrm{sim}} + \mathrm{Var}(Z)_{\mathrm{sim}}}{4\tau_{\mathrm{sim}}} \tag{4.26}$$

All the above output values of VHoppAS are not only determined at the end but also in 2.5% intervals of $\hat{N}_{\mathrm{h,eq}}$ and $\hat{N}_{\mathrm{h,sim}}$ during the equilibration and the main simulation.[89] This also includes

[89] During the equilibration, the temporary values of σ, u, etc. are of course calculated with the total displacement averages (e.g., with $\langle X\rangle$ instead of $\langle X\rangle_{\mathrm{sim}}$).

the determination of N_{mob} and N_{eff},[90] as well as several other quantities that are omitted here for brevity. This intermediary data is vital, for example to assess the linearity of $\langle X \rangle$ vs. τ or to check the convergence of the simulations.

Apart from explicit exceptions, the KMC results that will be reported later in sec. 5.5 are always the arithmetic average of 5 individual simulations with equal parameters. Only the seed for the random number generator is varied between these repetitions to ensure that the generated localized states and the movements of the electrons are different. The error bars of the KMC output values correspond to the sample standard deviation of these 5 repetitions.

4.2.6 Implementation Details

The KMC software VHoppAS features several optimizations of the algorithm, which shall be outlined in this section. For a more technical description, see the VHoppAS documentation and the heavily annotated source code of the KMC loop (cf. sec. 4.2).

The quality of the KMC results depends mainly on how well the DOS is represented by the localized states and how thorough the VRH process is sampled by the random walk of the electrons. Accordingly, the number of states and the number of hops should be as high as possible within the limits of the available computational resources. Since the memory consumption of VHoppAS is already rather low, with almost no room for improvement, the optimization efforts focused on the speed of the algorithm.

The most time-consuming parts of the KMC loop described in sec. 4.2.4 are the search for the lowest time-to-hop and the updating of the path times, because these tasks require to iterate over a huge number of electrons or paths. Three basic optimizations were implemented to reduce the general impact of these iterations:

- In comparison to the classical search of the shortest τ_{ij}^{*} in the complete catalog of several million paths, the usage of an individual $\tau_{\mathrm{min},k}$ for each electron is already an improvement. In consequence, the global minimum τ_{hop} must be searched just among thousands of electrons. Since only few $\tau_{\mathrm{min},k}$ need to be determined anew after each hop, this greatly reduces the total number of paths to be examined in each KMC cycle.
- The subtraction of τ_{hop} from all τ_{ij}^{*} and all $\tau_{\mathrm{min},k}$ in the 3. step of the algorithm was omitted altogether. Instead, each electron keeps record of the last time it hopped, which is then taken into account when comparing the $\tau_{\mathrm{min},k}$ of different electrons.
- An automatic adjustment of the cut-offs after the equilibration was introduced. The highest $|\vec{r}_{ij}|$ and $|\varepsilon_{ij}|$ of the hops executed during the equilibration are determined and the cut-offs are set

[90] The eventual fluctuations of N_{eff} have no effect on the other intermediary values because the latter get re-adjusted to the final N_{eff} when the simulation is complete.

to a certain percentage of these maximum values (typically 110% for r_{cut} and 125% for ε_{cut}). The unused paths that exceed these new cut-offs are removed before the main simulation.

A special feature of VHoppAS is the efficient handling of the oscillations that were introduced in the previous section. Typically, more than 90% of the hops correspond to such back-and-forth transitions between the same states. Since they do not yield any electron displacement, one might be tempted to exclude the respective paths from the simulation.[64, 205] However, this can at most be justified as a phenomenological approximation, by verifying that the conductivity remains equal. In general, the oscillatory hops are an inherent part of the conduction process and cannot be neglected. Removing the respective paths would inhibit the following two events that were frequently observed during the simulations:

- Oscillating electrons escape to other states from both states they oscillated between.
- While an electron oscillates between two states, other electrons hop across both of these states.

Instead of preventing the oscillations, the KMC algorithm therefore was optimized to reduce their computational cost. The oscillations are especially detrimental through the iterations over electrons or paths that they trigger. Suppose, for example, that an electron hops back-and-forth between two states i and j several times in row (e.g., due to $|\vec{r}_{ij}| \approx 0$ and $|\varepsilon_{ij}| \approx 0$). Each time, τ_{hop} is searched among all electrons, even though the KMC cell essentially did not change. Moreover, if i or j are the destination of any $\tau_{\text{min},k}$ of surrounding electrons, then each oscillation is accompanied by an unnecessary search of the new $\tau_{\text{min},k}$ for these electrons. But how to avoid this waste of computational resources without altering the VRH process? The key is to not only obtain the shortest hop times but also the second shortest times when iterating over electrons or paths.[91] This costs almost no additional computational time, but allows to omit subsequent searches in many cases. As described in the following, this idea is applied both on the level of the individual electron and globally.

Each search for $\tau_{\text{min},k}$ among the $\{\tau^*_{ij}\}$ of an electron k, during the preparations or in the actual KMC loop, also yields the second lowest time-to-hop $\tau_{\text{2nd},k}$ of the electron. These two values essentially form a short queue. Whenever $\tau_{\text{min},k}$ gets blocked because another electron hops to the respective destination state, then $\tau_{\text{2nd},k}$ becomes the new $\tau_{\text{min},k}$ of the electron k (and $\tau_{\text{2nd},k}$ is set to "undefined"), without iterating over the electron's paths. Conversely, when any path of the electron gets a new τ^*_{ij} because the respective destination becomes unoccupied, this new value is a candidate for $\tau_{\text{min},k}$ or $\tau_{\text{2nd},k}$: If the new τ^*_{ij} is smaller than $\tau_{\text{min},k}$ it becomes the new $\tau_{\text{min},k}$ (and $\tau_{\text{2nd},k}$ becomes the old $\tau_{\text{min},k}$). Instead, if the new τ^*_{ij} is between $\tau_{\text{min},k}$ and $\tau_{\text{2nd},k}$ it becomes the new $\tau_{\text{2nd},k}$. As a result, a full search for $\tau_{\text{min},k}$ (and $\tau_{\text{2nd},k}$) only becomes necessary when this short queue runs empty (i.e., when $\tau_{\text{min},k}$ would become "undefined") or after the electron itself made a hop. Specifically, this approach solves the following situation: Suppose that the destination state j of $\tau_{\text{min},k}$ is one of the states that a nearby electron oscillates between. When j gets occupied by the oscillating electron, then $\tau_{\text{min},k}$ becomes temporarily replaced by $\tau_{\text{2nd},k}$. When the oscillating electrons leaves

[91] Note that keeping all electrons or paths fully sorted would yield no optimization.

the state j, then the respective new τ_{ij}^{*} for the electron k is likely smaller than the replacement and it again becomes $\tau_{\mathrm{min},k}$. This restores the original paths that the electron k would take next, without searching them anew.

On the global level, two new variables – τ_{min} and τ_{2nd} – are introduced additional to τ_{hop}. While τ_{hop} refers to the hop that is currently executed, τ_{min} denotes the hop that occurs next and τ_{2nd} is the hop that would take place thereafter. They also work like a short queue: τ_{min} acts as supplier for τ_{hop} and τ_{2nd} is the backup for τ_{min}. At the beginning of the simulation, they are all "undefined". The 1. step of the KMC loop gets replaced by the following:

1. When τ_{min} is not already known (i.e., when it is "undefined"), then the lowest and second lowest $\tau_{\mathrm{min},k}$ are searched among all electrons, and τ_{min} and τ_{2nd} are set accordingly.
2. Then, τ_{min} is used as executed hop τ_{hop}, τ_{2nd} becomes the new τ_{min} and τ_{2nd} is set to "undefined".

Throughout the remainder of the KMC cycle, while processing the current hop τ_{hop}, τ_{min} is the candidate for the next hop. Whenever the $\tau_{\mathrm{min},k}$ of any electron gets a new value, it is compared to τ_{min} and becomes the new next hop if it is smaller (and τ_{2nd} is set to the old τ_{min}).[92] This applies both to the currently hopping electron and to the other electrons that are updated. Further special cases are described in the documentation of VHoppAS (e.g., when τ_{min} gets blocked by the current hop). In case of a simple oscillation, the following scenario unfolds: In the 4. step of the algorithm, the currently hopping electron gets its new lowest time-to-hop $\tau_{\mathrm{min},k}$, which might correspond to the reverse path, back to the localized state it just came from. This is quite likely when its current hop had $\varepsilon_{ij} > 0$, because the reverse direction with $\varepsilon_{ji} < 0$ then would have an even higher hopping probability. If the new $\tau_{\mathrm{min},k}$ is smaller than τ_{min}, then the currently hopping electron also becomes the hop candidate for the next KMC cycle. In this exemplary case, the electron would return to its previous location without a new search for τ_{hop} (or τ_{min}) among the electrons.

4.3 Finite Element Modeling (FEM)

When the density of localized states near the Fermi level is sufficiently high to enable variable-range hopping in the bulk of a-GaO_x, then the classical Schottky barrier models described in sec. 2.2.4 are not appropriate for metal|a-GaO_x interfaces. There are still Schottky barriers due to the different work functions, of course, but excitation over or tunneling through them is not necessary when enough localized states are present in the former band gap. Instead, the electrons may traverse these space charge regions by variable-range hopping. The electron transport in a TE|a-GaO_x|BE device then only depends on the local hopping conductivity and the electrochemical potential gradient of the electrons throughout the a-GaO_x layer. This idea was investigated by finite element modeling (FEM), which is a numerical method for solving partial differential equations over a spatial domain.

[92] When there is nothing to compare to (i.e., when τ_{min} is "undefined"), then there is no candidate for the next hop and the next KMC cycle definitely begins with a search for τ_{min} and τ_{2nd}.

The FEM simulations were carried out for a Pt|a-$GaO_{1.0}$|ITO device with a spatially and timely constant O/Ga ratio of $x = 1.0$ in the whole a-GaO_x layer. Oxygen and gallium migration were not included in the models. On the one hand, the required input data from DFT and KMC for the electrons is too far apart for a reasonable interpolation on the O/Ga ratio scale. On the other hand, there is also a severe lack of data for the ion transport (e.g., ionic conductivities and activity coefficients for different O/Ga ratios). For the time being, the FEM models therefore only describe an ideal as-prepared state of the device, without any surface oxidation, resistive switching or reorganization of the ions in the space charge zones. Joule heating was excluded as well and the simulations were carried out for a constant material temperature of $T = 300\,\mathrm{K}$.

The Pt|a-$GaO_{1.0}$|ITO device was modeled as one-dimensional, with z as the perpendicular axis from the TE to the BE (TE interface at $z = 0$, BE interface at $z = d$, $\nabla = \partial/\partial z$). This implies a parallel electric field and lateral homogeneity of all quantities in the xy-plane, which is a justified approximation because the TE diameter of the prepared RS devices in this work is orders of magnitude larger than the a-GaO_x film thickness d.

4.3.1 Zero Voltage Model

At first, the electron equilibrium without applied voltage was simulated. In this case, there is no electron flux and the chemical potential μ of the electrons and the electric potential φ are convertible into each other (cf. eq. 2.7):

$$J_\mathrm{e}(z) = -\frac{\sigma(z)}{e^2}\nabla\tilde{\mu}(z) = 0 \quad \Rightarrow \quad \nabla\tilde{\mu}(z) = 0 \quad \Rightarrow \quad \nabla\varphi(z) = \frac{1}{e}\nabla\mu(z) \tag{4.27}$$

In combination with the Poisson equation (eq. 2.13), this yields

$$\nabla^2\varphi(z) = -\frac{\rho(z)}{\epsilon_0\epsilon_\mathrm{r}} \quad \Rightarrow \quad \nabla^2\mu(z) = -\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\rho(z) \tag{4.28}$$

In the absence of ion transport, the local charge density $\rho(z)$ only depends on how far the electron density $n(z)$ deviates from its value in charge-neutral a-$GaO_{1.0}$. In other words, $\rho(z)$ is a unique function of the local chemical potential $\mu(z)$ and the constant temperature T. Like in the KMC simulations, μ shall be a relative chemical potential, with $\mu = 0$ for charge-neutral a-$GaO_{1.0}$. The charge density for a certain chemical potential is thus given by

$$\rho(\mu) = -e\,[n(\mu, T) - n(0, T)] \quad \text{with} \quad n(\mu, T) = \int_{-\infty}^{\infty} g(\varepsilon)\left[\exp\left(\frac{\varepsilon - \mu}{k_\mathrm{B}T}\right) + 1\right]^{-1} \mathrm{d}\varepsilon \tag{4.29}$$

Before the actual FEM simulations, this expression was used to calculate $\rho(\mu)$ by numerical integration of the previously obtained DFT-DOS $g(\varepsilon)$ of a-$GaO_{1.0}$. Thereby, it was assumed that the DOS is independent of the electron number. With the resulting data set, the local charge density in

the Poisson equation can be expressed as $\rho(\mu(z))$. The only spatially varying unknown in eq. 4.28 then is the chemical potential $\mu(z)$:

$$\nabla^2\mu(z) = -\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\rho(\mu(z)) \tag{4.30}$$

This ordinary differential equation was solved by FEM with the chemical potentials $\mu_\mathrm{TE} = \mu(z{=}0)$ and $\mu_\mathrm{BE} = \mu(z{=}d)$ as Dirichlet boundary conditions. Due to the equilibration at the interfaces, their values relative to charge-neutral a-$GaO_{1.0}$ can be approximated by the work function difference of the involved materials (cf. eq. 2.55):

$$\mu_\mathrm{TE} = W_\mathrm{GaO} - W_\mathrm{TE} \tag{4.31}$$

$$\mu_\mathrm{BE} = W_\mathrm{GaO} - W_\mathrm{BE} \tag{4.32}$$

Besides $\rho(\mu)$, the FEM model for $V = 0$ thus possesses five physical parameters – the film thickness d, the relative permittivity ϵ_r and the work functions of a-$GaO_{1.0}$, Pt and ITO. The thickness was set to $d = 40\,\mathrm{nm}$. The permittivity of a-$GaO_{1.0}$ is unknown and may differ from $\epsilon_\mathrm{r} \approx 10$ for a-$GaO_{1.5}$ (cf. sec. 2.3.3 and 4.1.2). Therefore, the FEM simulations were carried out for $\epsilon_\mathrm{r} = 10$, 100 and 1000. The work function of a-$GaO_{1.0}$ was approximated by $W_\mathrm{GaO} = 3.7\,\mathrm{eV}$, based on the UPS results discussed later in sec. 5.1.4. For the work functions of Pt and ITO, the literature values $W_\mathrm{TE} = 5.6\,\mathrm{eV}$ and $W_\mathrm{BE} = 4.4\,\mathrm{eV}$ were used (cf. sec. 2.3.3).

The primary result of this FEM simulation is the chemical potential $\mu(z)$, which characterizes the equilibrium distribution of the electrons. The calculation of the electric potential $\varphi(z)$ then requires just another reference definition: Since the BE was grounded in the I-V experiments, the electrochemical potential at $z = d$ shall be $\tilde{\mu}_\mathrm{BE} = 0$.[93] This implies $\tilde{\mu}(z) = \mu(z) - e\varphi(z) = 0$ in the whole a-$GaO_{1.0}$ layer, because $\tilde{\mu}$ is constant for $V = 0$ ($\nabla\tilde{\mu}(z) = 0$, cf. eq. 4.27). The electric potential is then simply given by

$$\varphi(z) = \frac{1}{e}\mu(z) \tag{4.33}$$

From the electric potential, the electric field $\vec{F}(z) = -\nabla\varphi(z)$ is accessible through numerical differentiation. The electric field strength $F(z) = |\nabla\varphi(z)|$ is generally quite high in the interface regions, even without an applied voltage. As mentioned already in the explanation of the KMC simulations (sec. 4.2.5), this may lead to an F-dependency of the electron temperature T_eff and the hopping conductivity σ. While the conductivity is irrelevant at $V = 0$, the elevated T_eff has a minor influence on the electron density for a certain chemical potential (also called "effective" chemical potential in this context). Therefore, the FEM simulations for $V = 0$ were done in two variants: The model "FEM-1" ignores high-field effects, while "FEM-2" takes them into account. The only difference between these variants is that the relation between the charge density and the chemical potential

[93] This arbitrary choice for $\tilde{\mu}_\mathrm{BE}$ is an indirect way to define $\varphi_\mathrm{BE} = \varphi(z{=}d) = (W_\mathrm{GaO} - W_\mathrm{BE})/e$ as reference point for the electric potential.

becomes F-dependent in FEM-2:[94]

$$\rho(\mu, F) = -e\left[n(\mu, T_{\text{eff}}(F)) - n(0, T)\right] \tag{4.34}$$

This 2-dimensional data set was calculated from the DFT-DOS in combination with the $T_{\text{eff}}(F)$ values obtained in the KMC simulations.

4.3.2 Current-Voltage Model

Afterwards, the FEM models for $V = 0$ were extended to obtain the steady-state current density j for an applied voltage $V \neq 0$. Instead of $\nabla\tilde{\mu}(z) = 0$, the condition $\nabla\left[\sigma(z)\nabla\tilde{\mu}(z)\right] = 0$ applies in this case according to the current continuity equation (cf. eq. 2.14):

$$\frac{\partial\rho(z)}{\partial t} = -\nabla j(z) = e\nabla J_{\text{e}}(z) = -\nabla\left[\frac{\sigma(z)}{e}\nabla\tilde{\mu}(z)\right] = -\nabla\left[\frac{\sigma(z)}{e}\nabla\left(\mu(z) - e\varphi(z)\right)\right] = 0 \tag{4.35}$$

The hopping conductivity as a function of μ was obtained from the KMC simulations, such that $\sigma(z)$ can be replaced by $\sigma(\mu(z))$. As a result, the above equation contains two spatially unknown quantities – the chemical potential $\mu(z)$ and the electric potential $\varphi(z)$. The second differential equation that links these quantities is the Poisson equation

$$\nabla^2\varphi(z) = -\frac{\rho(z)}{\epsilon_0\epsilon_{\text{r}}} \tag{4.36}$$

where $\rho(z)$ depends on $\mu(z)$ like in the zero voltage case. The resulting equation system with

$$\nabla\left[\frac{\sigma(\mu(z))}{e}\nabla\mu(z)\right] - \nabla\left[\sigma(\mu(z))\nabla\varphi(z)\right] = 0 \quad \text{and} \quad \nabla^2\varphi(z) = -\frac{1}{\epsilon_0\epsilon_{\text{r}}}\rho(\mu(z)) \tag{4.37}$$

was solved by FEM with Dirichlet boundary conditions for μ and φ. The chemical potentials μ_{TE} and μ_{BE} possess the same values as for $V = 0$ (eqs. 4.31 and 4.32). Their respective counterparts $\varphi_{\text{TE}} = \varphi(z{=}0)$ and $\varphi_{\text{BE}} = \varphi(z{=}d)$ follow from the electrochemical potential. The BE is again defined as reference electrode with $\tilde{\mu}_{\text{BE}} = 0$. Since the applied voltage corresponds to the electrochemical potential difference between the electrodes according to $\tilde{\mu}_{\text{BE}} - \tilde{\mu}_{\text{TE}} = eV$, the electrochemical potential at the TE interface is $\tilde{\mu}_{\text{TE}} = -eV$. The boundary conditions for the electric potential are therefore given by

$$\mu_{\text{TE}} - e\varphi_{\text{TE}} = -eV \quad \Rightarrow \quad \varphi_{\text{TE}} = \frac{1}{e}\left(W_{\text{GaO}} - W_{\text{TE}}\right) + V \tag{4.38}$$

$$\mu_{\text{BE}} - e\varphi_{\text{BE}} = 0 \quad \Rightarrow \quad \varphi_{\text{BE}} = \frac{1}{e}\left(W_{\text{GaO}} - W_{\text{BE}}\right) \tag{4.39}$$

[94] In principle, a spatially varying F could also cause a thermoelectric flux due to ∇T_{eff} (cf. eq. 2.7). However, as discussed already in sec. 2.1, it is $\nabla\varphi \gg S_{\text{e}}\nabla T_{\text{eff}}$ because of $S_{\text{e}} \ll 1\,\text{V/K}$ [40] and because ∇T_{eff} depends on $\nabla^2\varphi$. Therefore, the thermoelectric effect was omitted in the FEM simulations.

Besides the data set for $\sigma(\mu)$, the only new physical parameter in this model is the applied voltage V. The simulations were carried out for different voltages between −1 V and +1 V in 0.1 V steps. The other physical parameters were set to the same values as for $V = 0$ (cf. sec. 4.3.1).

At the end of each simulation, the current density j was calculated from the converged results for $\mu(z)$ and $\varphi(z)$ (cf. eq. 2.7):

$$j(z) = -eJ_e(z) = \frac{\sigma(\mu(z))}{e}\left[\nabla\mu(z) - e\nabla\varphi(z)\right] \tag{4.40}$$

It was verified that $j(z)$ is spatially constant, as required by eq. 4.35, except for negligible numerical deviations. The final value for j was then obtained as spatial average of $j(z)$.

Like for $V = 0$, also the simulations for $V \neq 0$ were carried out in two variants: The "FEM-3" model ignores high-field effects, whereas "FEM-4" respects the F-dependence of the charge density and the conductivity. The model equations are similar, but $\rho(\mu)$ and $\sigma(\mu)$ become $\rho(\mu, F)$ and $\sigma(\mu, F)$ in FEM-4. For $\rho(\mu, F)$, the same data set was used as in the FEM-2 model (cf. eq. 4.34). The conductivity was described by the 2-dimensional $\sigma(\mu_{\text{eff}}, F)$ data set that was obtained from the KMC-VRH simulations.[95]

4.3.3 FEM Settings

The FEM simulations for the four models were carried out with the FEniCS library for the Python programming language.[206–215] The one-dimensional mesh for the a-$GaO_{1.0}$ layer from $z = 0$ to $z = d$ was divided into 1300 linear Lagrange elements (P1), with an increased node density near the boundaries. This means that $\mu(z)$ and $\varphi(z)$ were approximated as piece-wise linear functions.

To enable this discretization of the simulated geometry into finite elements, the non-linear partial differential equations of the models were transformed into their variational forms (also called "weak formulation"). The resulting expressions are listed in sec. A.5. The test functions introduced in these equations belonged to the same function space as the searched functions $\mu(z)$ and $\varphi(z)$ (Galerkin-FEM).

The variational expressions were solved iteratively with the Newton method, to circumvent the non-linearity that originates from the μ- and F-dependence of ρ and σ. This approach required the so-called Jacobian, which is essentially the derivative of the variational formulation with regard to the searched functions. It describes how the solution changes in each iteration step, based on the known ρ and σ values of the previous step. The derived Jacobians for the four models are given in sec. A.5. In each Newton iteration, the change of μ and φ was obtained from the Jacobian with the default Krylov solver of FEniCS. The specific ρ and σ values needed in this algorithm were

[95] Note that the chemical potential μ in the models FEM-2 and FEM-4 corresponds to the effective chemical potential μ_{eff} introduced in sec. 4.2.5, because μ_{eff} is the actual chemical potential of the electrons beyond the weak field regime (cf. eq. 4.17).

calculated dynamically from the supplied data sets by linear interpolation on the μ- and F-axis. Linear extrapolation was used when $F(z) = |\nabla\varphi(z)|$ exceeded the maximum electric field strength of 5.56 MV/cm in the underlying KMC data.

Convergence was reached when the variational expression differed from zero by less than 10^{-8} for the models FEM-1 and FEM-2 and by less than 10^{-5} for FEM-3 and FEM-4. This was also verified by inserting the final results for $\mu(z)$ and $\varphi(z)$ into the original differential equations.

5

Results and Discussion

5.1 Material Characterization

The amorphous gallium oxide thin films, which were deposited in Ar or H_2/Ar atmosphere, all have a brown color that darkens with increasing film thickness – from light brown and semi-transparent to dark brown and opaque (fig. 5.1a,b). Decreasing PLD fluence (fig. 5.1a,c,d) or the presence of H_2 in the deposition atmosphere (fig. 5.1e,f) also tends to yield darker thin films. In contrast, the gallium oxide thin films deposited in O_2 atmosphere are colorless and fully transparent. No change of the samples' color was observed even after years of storage in air, which is an argument for the long-term stability of the material.

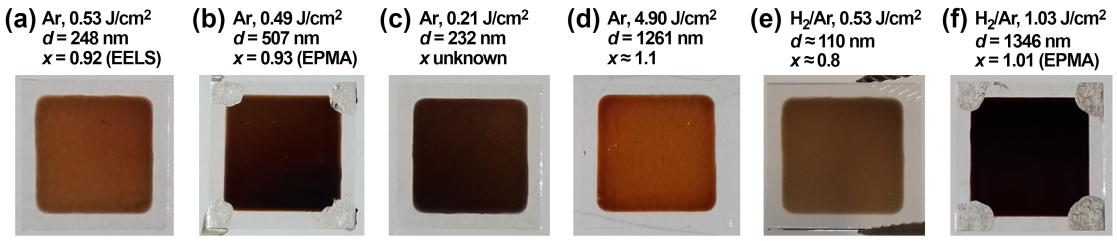


Figure 5.1: Exemplary photos of a-GaO_x thin films, which differ in the used PLD atmosphere, the PLD fluence, the film thickness d and the O/Ga ratio x.

These observations for the color of a-GaO_x are in line with the expectations from sec. 2.3.1. The dark color of the a-GaO_x thin films originates from the absorption of light in the whole visible spectrum. Accordingly, the band gap of the a-GaO_x films cannot exceed 1.6 eV, which is the lower limit of the visible spectrum. The optical absorption becomes gradually less, though, towards the low-energy end of the visible spectrum (e.g., for red and orange light, approx. energy range: 1.6–2.1 eV), leading to the brown color when the films are thin enough.[96] This indicates an at least about 2 eV wide region between valence and conduction band, where the DOS is low but not always zero (i.e., a

[96] UV-Vis absorption spectroscopy was also tested in this work, but the Tauc plots had no well-defined, linear edge for the optical absorption, in accordance with the presence of a large number of defect states in the former band gap. The absorption spectra showed that the a-GaO_x thin films absorb light in the whole visible spectrum, with a gradual decrease of the absorbance towards higher wavelengths (lower energies).

pseudo-gap). These conclusions fit to the DFT-DOS from literature (sec. 2.3.1), which showed a clear trend of more and more defect states in the band gap with decreasing O/Ga ratio. Conversely, this implicates that the darker color for lower fluence or for H_2/Ar atmosphere signals a decreasing O/Ga ratio in these cases. This hypothesis will be evaluated further in sec. 5.1.2, where the EPMA and EELS results will be discussed.

The growth rates of the a-GaO_x thin films also follow plausible trends with regard to the PLD conditions. Higher laser fluence leads to a higher growth rate, which can be readily understood from the increased ablation from the PLD target. The same has also been reported in [123]. Typical values for the growth rate in Ar atmosphere are 0.37 Å/pulse for 0.49 J/cm^2 and 0.61 Å/pulse for 0.77 J/cm^2 (averaged over several samples). Note that the growth rates have quite large scatter in the order of ±0.1 Å/pulse for equal PLD conditions. In H_2/Ar atmosphere, the growth rates tend to be substantially lower – almost halved in comparison to pure Ar with equal pressure. This is most probably due to hydrogen forming volatile molecules with the ions in the PLD plasma plume (or with already adsorbed material on the film). Deposition in O_2 atmosphere, in contrast, leads to significantly faster film growth, e.g., 1.18 Å/pulse on average for 0.79 J/cm^2. This can be explained with the enhanced incorporation of oxygen during the film growth. The additional O_2 shifts the equilibrium between volatile Ga_2O and solid Ga_2O_3 towards the latter, thereby reducing the amount of gallium lost to the gas phase.[216, 217]

The deposited a-GaO_x thin films have rather smooth surfaces, with a maximum area roughness S_z below 10 nm on average (S_z-range of all IM measurements: 5–13 nm). The thickness of the a-GaO_x thin films, however, is not homogeneous over the whole sample area. It decreases towards the borders of the film due to the shadow effect from the sample holder, which blocks a part of the deposition trajectories towards the substrate. The thickness difference between the center and the border of the a-GaO_x thin films is typically in the order of 10%. A measure for this effect is the thickness standard deviation, which is used for example to obtain reasonable conductivity error bars in the van-der-Pauw experiments (sec. 5.3).

5.1.1 Amorphicity

From the previous studies on PLD-made gallium oxide thin films (sec. 2.3), there is no doubt that those deposited at room temperature – as done in this work – are amorphous. Nevertheless, the amorphicity of exemplary samples was confirmed by electron diffraction. Fig. 5.2 shows the respective SAED patterns of three different thin films. As expected for amorphous samples, they exhibit only broad, blurred rings and no discrete rings or spots. Variation of the PLD fluence or the PLD atmosphere has no visible effect on the patterns.

The STEM-SAED measurements refer to the as-prepared state of the samples. The influence of post-deposition annealing on the structure of similarly prepared a-GaO_x has already been investigated by others in [33, 97], using the same TEM method. As already discussed in sec. 2.3, they

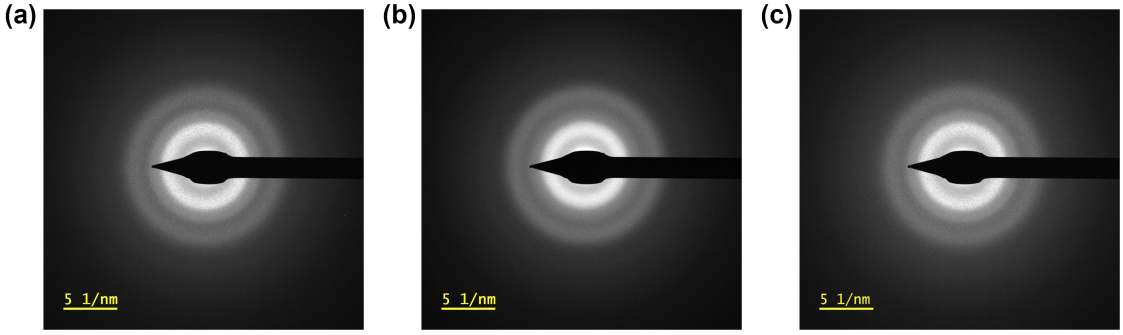


Figure 5.2: Selected area electron diffraction patterns (STEM-SAED) of three exemplary a-GaO_x films that were deposited with different PLD conditions: (a) in Ar with 0.49 J/cm^2, (b) in Ar with 2.94 J/cm^2 and (c) in H_2/Ar with 0.49 J/cm^2. The O/Ga ratios of these samples, as measured by EPMA and EELS, are shown in fig. 5.4.

observed neither full nor partial crystallization for temperatures up to 320 °C. Since the highest sample temperature reached within this work is 300 °C, except for RTA treatments were crystallization was intended (sec. 3.2.2 and 3.2.3), it is safe to assume that all investigated a-GaO_x thin films were amorphous at all times.

The SAED patterns shown in fig. 5.2 were measured several years after the respective thin films were deposited. The TEM results for these quite old samples therefore also illustrate that the a-GaO_x thin films stay amorphous during storage at room temperature. This long-term stability of the amorphous structure is a sign that the ions within a-GaO_x are not very mobile. Otherwise, the meta-stable structure would convert into the energetically more favorable crystalline structure, accompanied by a phase separation as observed at higher temperatures (cf. sec. 2.3).

5.1.2 Non-Stoichiometry

The stoichiometry x of the a-GaO_x thin films, and the influence of the PLD conditions thereon, was investigated by EPMA. Fig. 5.3 shows the results for all measured samples. Overall, there is an approximate trend of increasing O/Ga ratio x, i.e., lower oxygen-deficiency, for increasing PLD fluence. This relationship, which was already hypothesized based on the literature data in sec. 2.3, becomes plausible when taking the scattering processes in the PLD chamber into account: Higher laser fluence increases the kinetic energy of the ablated material, thereby increasing the mean free path of the particles on their way to the substrate. This reduction of the scattering with the gas atmosphere is especially important for lighter species like oxygen, which in general are more susceptible to scattering away from the substrate. Therefore, more oxygen can reach the substrate when the PLD fluence is higher.

While the influence of the fluence is clearly visible in fig. 5.3, further trends are hard to identify. Since only two of the measured samples were deposited in chamber #1, it is difficult to judge whether the two chambers yield comparable O/Ga ratios for equal PLD parameters. It seems like chamber #1

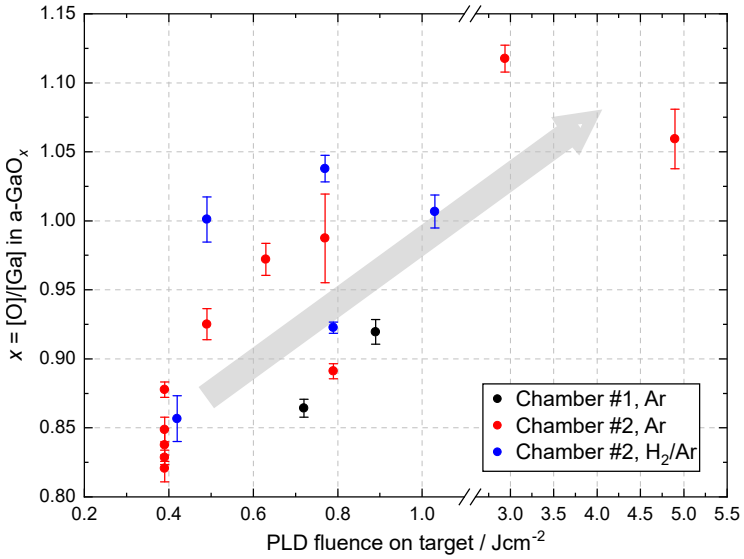


Figure 5.3: Fluence-dependence of the bulk stoichiometry of a-GaO_x thin films, as measured by EPMA. The samples were deposited in either chamber #1 or #2 and in Ar or H_2/Ar atmosphere.

produces slightly more non-stoichiometric films with lower x, but the difference is not significant. There is also no clear change of the O/Ga ratio when using H_2/Ar gas instead of pure Ar in the PLD chamber, at least according to the limited amount of EPMA measurements. The EELS results, which will be discussed further on, yield a bit more insight into this, but the EPMA data already tells that 5% H_2 in the deposition atmosphere has no drastic effect on the stoichiometry. This is in stark contrast to the consequences for the growth rate, which, as mentioned in sec. 5.1, is almost halved by using H_2/Ar. An explanation could be that H_2 not only scavenges oxygen as H_2O (g) but also gallium as volatile gallane (GaH_3). In principle, the ratio between H_2O (g) and GaH_3 (g) formation then would influence the O/Ga ratio. In the present case, the effects from both reactions probably almost cancel each other, leading to a rather weak influence of H_2 on the stoichiometry.

Fig. 5.3 shows that the sample-to-sample reproducibility of the stoichiometry x for formally equal PLD conditions is quite poor. Therefore, care must be taken when estimating the stoichiometry of samples which were not explicitly measured by EPMA. Additional to the equality of the PLD conditions, also temporal proximity of the deposition is a requirement for this. However, there is substantial scatter even for films prepared immediately after each other, e.g., for those with a fluence of 0.39 J/cm^2 in fig. 5.3. Despite many efforts, it was not possible to find the underlying cause for the deviations between formally equal PLD depositions. For example, repeated fluence calibration, precise control of the chamber atmosphere (incl. the background pressure) and evaluation of the flushing procedure brought no improvement. It is also unlikely that the unavoidable variations of the room temperature or the target's surface have such a big influence. The same goes for potential leakages of the chambers, which might raise the p_{O_2} beyond what is expected from the Ar gas purity, though presumably to the same extent for each deposition. Essentially, to clarify this problem, the conditions inside the chamber would have to be analyzed simultaneously to the deposition process.

The used PLD workstation, however, has no such equipment. The only way to compensate this reproducibility issue would be to measure the O/Ga ratio of each investigated sample explicitly. Unfortunately, this was not possible due to the limited availability of the respective measurement capabilities. Ultimately, based on the variations seen in fig. 5.3, there remains no other option than to assign an uncertainty of $\Delta x \approx 0.1$ to any estimated O/Ga ratios.

The bulk stoichiometry of selected a-GaO_x thin films was analyzed further through STEM-EELS measurements. Fig. 5.4 shows three of the resulting, depth-resolved composition profiles. The majority of the EELS profiles exhibits a quite homogeneous O/Ga ratio throughout the bulk, with $\Delta x < 0.03$ within the individual profiles (e.g., fig. 5.4a,c). Some profiles, however, contain larger stoichiometry variations, for example $\Delta x \approx 0.08$ in fig. 5.4b. Measurements in different regions of the same sample also differed more significantly, typically with $\Delta x \approx 0.06$. This suggests that a part of the EPMA reproducibility issue originates from measurable, yet not dramatic lateral inhomogeneity.[97] This cannot be the full explanation, however, because the later conductivity measurements (sec. 5.3), which essentially probe each thin film as a whole, yield large sample-to-sample variability as well.

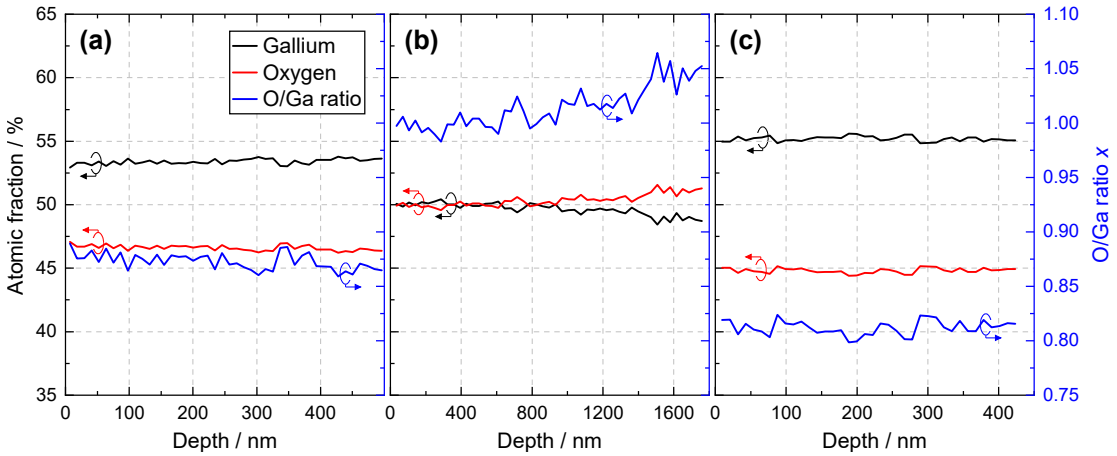


Figure 5.4: Electron energy loss spectroscopy (EELS) depth profiles of three exemplary a-GaO_x films that were deposited with different PLD conditions: (a) in Ar with 0.49 J/cm^2, (b) in Ar with 2.94 J/cm^2 and (c) in H_2/Ar with 0.49 J/cm^2 (same samples as in fig. 5.2). The profiles show the atomic fractions of Ga and O (left axis) and the stoichiometry x (right axis) in the bulk, from the surface towards the substrate. According to EELS, the three thin films have the following O/Ga ratios: (a) 0.82–0.89, (b) 0.98–1.06 and (c) 0.80–0.86 (incl. further profiles not shown here). These values are lower than the stoichiometries $x = 0.93 \pm 0.02$ for (a), $x = 1.12 \pm 0.01$ for (b) and $x = 1.00 \pm 0.02$ for (c), which were measured by EPMA on the same samples.

The EELS results confirm the trend that higher PLD fluence leads to a higher O/Ga ratio x of the gallium oxide thin films (fig. 5.4a,b). Furthermore, they indicate a slight reduction of the stoichiometry x by using H_2/Ar gas instead of pure Ar (fig. 5.4a,c). This latter observation could be a coincidence for these two specific samples, but it fits to the general tendency of darker color for

[97] The EPMA results were averaged over multiple measurement spots on each sample, but they typically cover only an about 0.1 mm wide region of the 8 × 8 mm^2 samples.

such thin films (cf. sec. 5.1). In any case, like EPMA, also the EELS measurements detect at most a minor effect of H_2 on the non-stoichiometry of a-GaO_x.

A direct comparison between EPMA and EELS is possible as well, since the three shown EELS profiles were obtained after EPMA measurements on the same samples (see description of fig. 5.4). The O/Ga ratios from EPMA are all higher than the respective, later EELS results, even though the samples were stored in air for over a year in between the measurements. There is no increase of the O/Ga ratio in the bulk over time, despite the high oxygen-deficiency of the samples. This long-term stability of the bulk against oxidation is an important point for potential applications. It is also most unlikely that the discrepancy between EELS and EPMA originates from a real loss of oxygen in the meantime. Instead, it seems like either EPMA overestimates or EELS underestimates the oxygen content of the thin films, probably due to the different calibration procedures of the two methods (cf. sec. 3.2.1). This could be a suitable explanation for the samples of fig. 5.4a and b, because for them the O/Ga ratios from both methods are quite close and differ approximately by the same factor. For the sample of fig. 5.4c, however, the deviation is rather large. The respective EPMA data point at 0.49 J/cm^2 in fig. 5.3 also looks like an outlier in comparison to the other EPMA results. Apparently, EPMA yields a too high O/Ga ratio for this specific sample. This could be connected to the fact that – due to the combination of low fluence and H_2/Ar gas – this sample has a much lower film thickness (502 nm) than the others that were sent to the external EPMA measurements at the same time. The adjustment of the EPMA electron beam parameters, in order to ensure that the excitation bulb does not include the ITO substrate, probably was done just with one of the thicker samples. The EPMA-derived stoichiometry for the rather thin sample of fig. 5.4c thus probably contains an additional oxygen signal from the substrate, which would explain the too high O/Ga ratio. Taking into account the EPMA workflow and the other sample thicknesses, this is likely the only EPMA sample where a major misadjustment of this kind occurred.

5.1.3 Surface Oxidation

While the previous section showed that the bulk stoichiometry does not change during storage in air, the opposite is true for the surface of the a-GaO_x thin films. As described in the following, both surface-sensitive HAXPES spectra and high-resolution EELS profiles of the surface regions yield a substantially increased O/Ga ratio x_{surf} after air exposure.

In the measured HAXPES spectra, all peaks could be identified (see full example spectra in fig. B.1). Additional to the expected signals from gallium and oxygen (and gold for the Au-coated samples), only an adventitious carbon impurity peak is observed. The adsorption of carbon species on the surface is unavoidable for samples stored in air. No other elements are visible in the broad binding energy range of the spectra. This indicates a good purity of the a-GaO_x thin films.

The Ga $2p_{3/2}$ peak of the non-stoichiometric a-GaO_x samples has a clearly visible shoulder (fig. 5.5a), which is absent for the a-$GaO_{1.5}$ thin film that was deposited in O_2 (fig. 5.5b). As explained in

sec. 3.2.2, the shoulder was analyzed by deconvolution of the Ga $2p_{3/2}$ signal into two sub-peaks, which – following the peak assignment in [90, 184] – were attributed to the formal oxidation states Ga^{3+} and Ga^{+}. For the a-$GaO_{1.5}$ sample, a single Ga^{3+} sub-peak was sufficient to describe the whole Ga $2p_{3/2}$ signal, resulting in $x_{surf} = 1.5$ for this sample. The Ga 3d peak of the samples also exhibits a shoulder (fig. B.2), which – after subtraction of the overlapping O 2s peak – was fitted by two sub-peaks for Ga^{3+} and Ga^{+} as well. The O 1s peak of the gallium oxide thin films, in contrast, seems to be mostly unaffected by non-stoichiometry (fig. B.3). For all samples, including a-$GaO_{1.5}$, the O 1s signal consists of a major peak for O^{2-} in the solid and a minor sub-peak at higher binding energy, which according to [90, 92] most probably originates from adsorbed carbon dioxide (CO_2).

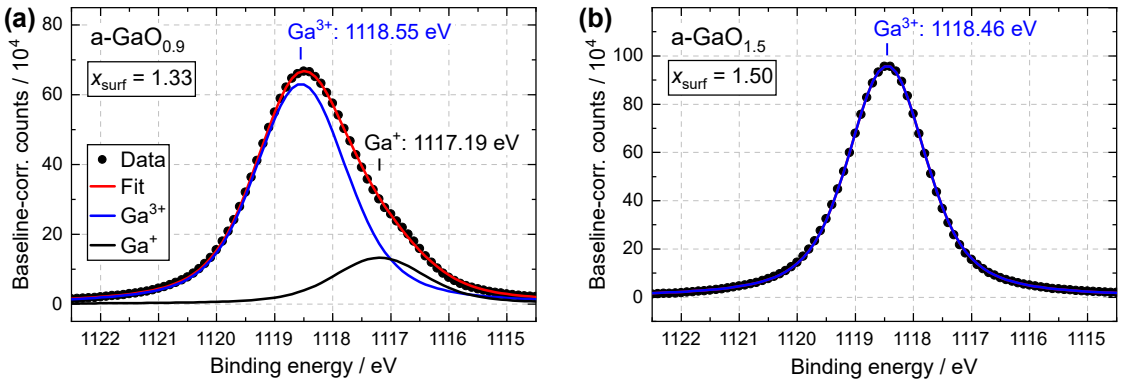


Figure 5.5: HAXPES Ga $2p_{3/2}$ peaks of (a) a-GaO_x with $x_{bulk} \approx 0.9$ (PLD with 0.79 J/cm² in Ar) and (b) a-$GaO_{1.5}$ as reference (PLD in O_2). The sub-peaks were identified by comparison with [90, 184]. The O/Ga ratio x_{surf} in the first few nanometers of the thin films was calculated from the sub-peak area ratio with eq. 3.3. The corresponding full HAXPES spectrum, the Ga 3d signal and the O 1s signal of sample (a) are shown in figs. B.1a, B.2 and B.3, respectively.

Due to the employed hard X-rays, reliable peak fitting was possible even for the Au-coated a-GaO_x films. Despite lower absolute intensity, the gallium and oxygen peaks measured through the approximately 8 nm thick Au layer still had a high signal quality. The resulting peak positions of all HAXPES measurements, which are summarized in tbl. C.1, agree well with previous gallium oxide studies.[10, 90–92, 104, 115, 134, 184, 185]

The O/Ga ratio x_{surf} at the surface of the thin films was calculated from the area ratio of the fitted Ga^{3+} and Ga^{+} sub-peaks (eq. 3.3). The final results for all measured samples, both from the Ga $2p_{3/2}$ and the Ga 3d signal, are listed in tbl. 5.1. For each of the a-GaO_x thin films, the surface stoichiometry x_{surf} is much higher than the O/Ga ratio $x \approx 0.9$ in the bulk. This confirms the expected, yet never before proved surface oxidation of the highly non-stoichiometric material. Due to the large oxygen deficit and the high amount of compensating electrons, the freshly prepared a-GaO_x surface is highly reactive when exposed to ambient atmosphere after the PLD process. The reaction with O_2 (g) leads to the observed increase of the O/Ga ratio (reversal of eq. 2.68). In humid atmosphere, also the reaction

$$H_2O\,(g) + O^{2-}\,(s) \rightleftharpoons 2\,OH^{-}\,(s) \tag{5.1}$$

could take place, which also would increase x_{surf} and may result in the incorporation of at least small amounts of hydrogen into the material.

In tbl. 5.1, the samples covered by an Au layer have lower x_{surf} than those without such a protective layer. It seems, however, quite irrelevant for the degree of surface oxidation whether Au was deposited directly, after 1 h or after 3 days of air exposure. Even keeping the sample in ultra-high vacuum in-between the two PLD depositions did not prevent surface oxidation. Apparently, the thin Au layer is not air-tight and most of the oxidation happened after the Au deposition. This is also demonstrated by the results for the sample that was annealed in air after the Au deposition. It exhibits a significantly higher x_{surf} than the other a-GaO_x thin films, even higher than those without Au layer, because of thermally accelerated oxidation through the pores of the Au layer.

Table 5.1: O/Ga ratio at the surface of different a-GaO_x thin films, according to the HAXPES measurements. The second column shows the bulk stoichiometry, as estimated from EPMA reference samples, while the fourth and fifth columns contain the respective surface stoichiometry obtained from Ga $2p_{3/2}$ and Ga 3d analysis. The description "none" for air exposure means that the respective sample was continuously kept in the evacuated PLD chamber until it was covered by the Au layer.

			x_{surf}		
PLD conditions	x_{bulk}	Au layer	Ga $2p_{3/2}$	Ga 3d	Air exposure during preparation
0.72 J/cm^2, Ar, #1	0.9	no	1.28	1.19	always because no Au layer
0.79 J/cm^2, Ar, #2	0.9	no	1.33	1.23	always because no Au layer
0.79 J/cm^2, Ar, #2	0.9	yes	1.25	1.17	3 days in air before Au layer
0.72 J/cm^2, Ar, #1	0.9	yes	1.18	1.13	less than 1 h in air before Au layer
0.79 J/cm^2, Ar, #2	0.9	yes	1.21	1.14	none
0.79 J/cm^2, Ar, #2	0.9	yes	1.40	1.24	none, post-annealed in air (573 K, 1 h)
0.72 J/cm^2, O_2, #1	1.5	no	1.50	1.33	always because no Au layer

The x_{surf} values obtained from the Ga $2p_{3/2}$ signal are always higher than those from the respective Ga 3d analysis. Since the latter peak has a higher information depth, this fits to the expected gradual decrease of the O/Ga ratio from the surface to the bulk. Since even the Ga 3d signal only yields an average x_{surf} from the first several nanometers of the sample, the significant difference between the results from Ga $2p_{3/2}$ and Ga 3d indicates that the surface oxidation layer does not reach far into the material.

This was confirmed by high-resolution EELS profiles in the surface region of non-coated a-GaO_x films, which were stored in air for more than a year. Both the EELS maps and their linear plots in fig. 5.6 clearly show an absolute increase of the oxygen content at the surface of a-GaO_x. The maximum $x \approx 1.3$ of the EELS profile also fits to the typical values obtained by HAXPES. Even after such long air exposure, the surface oxidation only extends approximately 10–15 nm into the sample. This gives an idea of what to expect for diffusive ion transport in a-GaO_x, because the gradient in x is still remarkable even after more than a year. If oxygen ion transport at room temperature would be fast for such steep oxygen concentration gradients, then the surface oxidation should extend

further into the material. Resistive switching processes solely based on oxygen ion diffusion in a-GaO_x seem rather unlikely according to these results (cf. sec. 2.4.3). Low ionic conductivity is, however, good news for the retention of RS states.

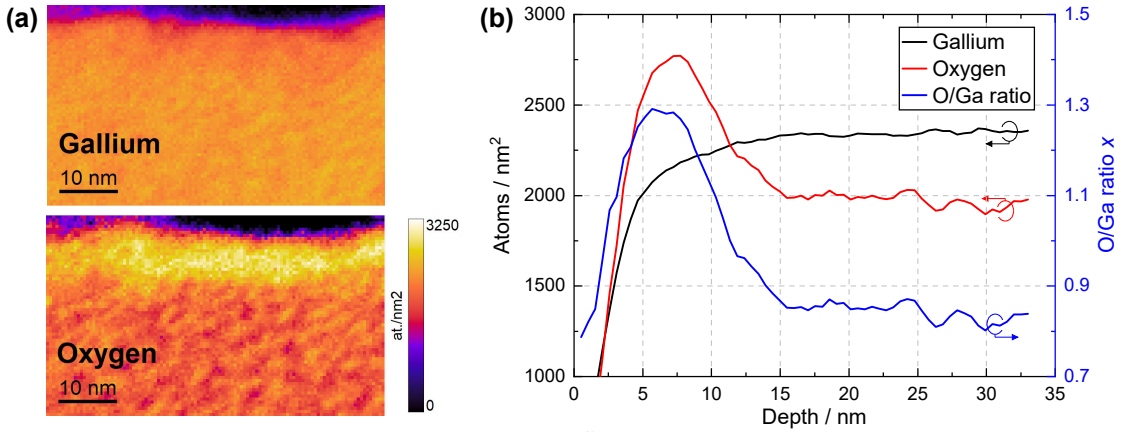


Figure 5.6: EELS surface analysis of a-GaO_x, deposited with 0.49 J/cm^2 in Ar, with x_{bulk} of 0.82–0.95 according to EPMA and bulk EELS measurements. (a) Atomic density maps of gallium and oxygen. The air-exposed surface of the thin film is at the top. (b) Line plots of the map data, with the depth axis pointing from the top to the bottom of the EELS maps.

Besides the surface, also the a-GaO_x|ITO interface was investigated by EELS (fig. B.4). Even though a minor oxygen transfer from ITO to a-GaO_x cannot be ruled out, there are no visible effects that exceed the roughness of the interface.[98]

Overall, the HAXPES and EELS results indicate that the surface oxidation of a-GaO_x is a slow, diffusion-limited process after the initial, probably fast oxidation of the first atomic layer(s). A comparable passivation effect is well-known for liquid gallium metal, which in air rapidly forms a 1–5 nm thick skin of a-$GaO_{1.5}$ but stays unaffected on the inside.[184, 218, 219] For this reason, oxidizing Ga metal to gallium oxide powder is rather difficult and requires for example ultrasonication to break the passivation layer.[220, 221]

The surface oxidation of a-GaO_x is especially relevant for ultra-thin films, where the oxidation occurs in a large percentage of the layer, or for applications where the surface plays a dominating role. It thus can be relevant for interface-type resistive switching (cf. sec. 2.4.2), when the a-GaO_x layer gets in contact with air before the metal|a-GaO_x interface is formed. For example, in [28] it has been proposed that the first RS transition of a-GaO_x interface switching required higher voltage because of a surface oxidation layer with high resistance. The work function of a-GaO_x, which will be discussed in the following, is another way how surface oxidation may influence resistive switching.

[98] SIMS measurements on a Pt|a-GaO_x|ITO sample with ^{18}O-enriched ITO also showed no measurable ^{18}O migration from ITO to a-GaO_x – neither in the as-prepared state, nor after applying −2 V to the Pt-TE for 2 min or after annealing at 150 °C for 30 min.

5.1.4 Work Function

The work functions of the a-GaO_x thin films, and of the different electrode materials used in this work, were measured by SKPM and UPS. The results are shown in tbl. 5.2. Naturally, both of these methods are surface-sensitive, which means that the results are heavily affected by the surface oxidation that was discovered in the previous section. This also explains the first major observation from these measurements, namely the approximately equal work functions for all non-annealed a-GaO_x samples, irregardless of their differing PLD parameters. Since all the samples were stored in air, they likely developed a more or less equal surface composition $x_{surf} \approx 1.3$ (cf. tbl. 5.1, fig. 5.6). The measured work functions thus refer to this rather uniform, elevated O/Ga ratio instead of the bulk stoichiometries of the samples. Removal of the oxidation layer by sputtering was no option, unfortunately, because it would have altered the O/Ga ratio artificially due to the different sputtering rates of oxygen and gallium.

Table 5.2: Work functions from the SKPM and UPS measurements. For the a-GaO_x thin films, which were deposited in Ar atmosphere, no influence of the laser fluence (0.36–1.21 J/cm^2) or the used PLD chamber (#1 or #2) was observed. Therefore, no distinction is made with regard to these parameters and the individual results were aggregated. The "datasets" column shows the number of samples and the total number of measurements that are included in each average work function.

Method	Material	Annealing	Datasets	Work function (eV)
SKPM	a-GaO_x		12, 52	3.96 ± 0.10
SKPM	a-GaO_x	200 °C, 1 h, in H_2(5%)/Ar	2, 12	3.92 ± 0.04
SKPM	a-GaO_x	300 °C, 1 h, in H_2(5%)/Ar	2, 12	4.19 ± 0.05
SKPM	a-GaO_x	300 °C, 10 min, in O_2	2, 12	4.12 ± 0.04
SKPM	a-GaO_x	600 °C, 10 min, in O_2	1, 12	4.44 ± 0.04
SKPM	Pt		3, 14	4.82 ± 0.06
SKPM	Au		1, 4	4.84 ± 0.05
SKPM	ITO		9, 22	4.56 ± 0.06
UPS	a-GaO_x		2, 2	3.66 ± 0.11
UPS	a-GaO_x	600 °C, 10 min, in O_2	1, 1	4.11

The average work function $W = (3.96 \pm 0.10)$ eV from the SKPM measurements of all non-annealed a-GaO_x films is still much lower than the literature value of 4.68 eV for a-$GaO_{1.5}$.[157] This fits to the expected trend of decreasing work function for lower O/Ga ratio, as indicated by the shift of the Fermi level in previous DFT calculations (cf. fig. 2.2). The work function of (3.66 ± 0.11) eV, which was measured by UPS, is even lower, probably for the two following reasons combined. On the one hand, due to the cut-off method (sec. 3.2.2), UPS in general yields the lowest work function on the whole surface, while SKPM gives an areal average.[160, 163] On the other hand, the two a-GaO_x samples measured by UPS were stored in vacuum bags (approx. 0.1 bar) for almost the full time between their preparation and the measurement. It was seen in tbl. 5.1 that this does not prevent surface oxidation, but of all the analyzed samples, they likely have the lowest x_{surf}.

In the annealing experiments, it was found that prolonged exposure to H_2(5%)/Ar gas at 200 °C has no effect, whereas at 300 °C an almost similar increase of the work function is obtained for 1 h in H_2/Ar and for 10 min in O_2. There seems to be no reduction of x_{surf} by the hydrogen gas. In contrast, a change of the amorphous structure might be responsible for the higher work function. In O_2, the work function likely increases faster due to an additional, slight increase of x_{surf}. Since the sample annealed in O_2 at 300 °C stayed brown and opaque, it can be concluded that 10 min is too short to oxidize the bulk to transparent a-GaO_x with $x \geq 1.3$. This is another sign of slow chemical diffusion of oxygen in a-GaO_x. Even the a-GaO_x thin film that was annealed at 600 °C in O_2 for 10 min did not become transparent. The higher annealing temperature though led to an even stronger increase of the SKPM-derived work function, almost reaching the aforementioned literature values for a-$GaO_{1.5}$ and β-Ga_2O_3 (sec. 2.3.3). According to previous post-annealing studies (e.g., [121]), the short heating period likely is not sufficient to fully crystallize the sample. Structural changes and even partial crystallization, however, are to be expected and probably play a role in the increase of the work function, in addition to a further increase of x_{surf} towards 1.5. The UPS measurement on this very sample yields a much lower work function than SKPM, most likely because of the accumulation of photoelectrons on the insulating surface. The same shift to higher binding energy due to charging was also seen in the HAXPES measurements on a-$GaO_{1.5}$ (sec. 3.2.2), but the short UPS spectra ($E_{bin} < 18$ eV) feature no peaks that would allow for a correction of the energy axis.

The measured work functions for the Pt and Au electrodes in tbl. 5.2 differ largely from the literature values of 5.5–5.6 eV and 5.3–5.5 eV, respectively.[158–160] The obtained value for Au, however, matches the work function of 4.9 eV that is reported in [160] for air-exposed – i.e., dirty – gold thin films. In the same study, 5.3 eV was measured for Ar^+ sputter-cleaned Au. Therefore, the deviations observed here for Pt and Au must be attributed to adsorbed contaminants (hydrocarbons, water, etc.), which alter the metal surface potential. The respective literature values are presumably more appropriate for the deposited Pt and Au electrodes in this work. The measured work function for the ITO substrates in tbl. 5.2, on the contrary, agrees very well with the literature values of 4.3–4.5 eV.[159, 161–163] Apparently, the work function of ITO is less sensitive to the inevitable surface contamination. While this could mean that the gallium oxide work functions are also less affected than those of the metals, it cannot be ruled out that there is a significant influence.

In summary, the SKPM and UPS measurements yield some first work functions for amorphous, non-stoichiometric gallium oxide, which might be especially relevant for metal|a-GaO_x interfaces that were exposed to air during their preparation. Reliable work functions that refer to the bulk composition of the a-GaO_x thin films require more elaborate experiments, where surface oxidation and contamination are avoided by keeping the samples in high vacuum or inert atmosphere all the time.

5.2 Resistive Switching

The resistive switching of a-GaO_x was investigated by I-V measurements, mainly on Pt|a-GaO_x|ITO devices. The main goals of these experiments were the identification of characteristic behavior and the collection of clues to the underlying RS mechanism(s). Like in the literature summarized in sec. 2.4, the a-GaO_x devices exhibited a broad variety of resistive switching patterns:

- **Counter-8 bipolar RS**: LRS → HRS at $V > 0$ and HRS → LRS at $V < 0$.
- **Figure-8 bipolar RS**: HRS → LRS at $V > 0$ and LRS → HRS at $V < 0$.
- **Abnormal bipolar RS**: Similar transitions at $V > 0$ and $V < 0$ (e.g., LRS → HRS), in combination with an apparent reversal at/near $V = 0$ (e.g., HRS → LRS).

The respective measurements each will be described in a dedicated section, followed by a collective analysis. Strictly speaking, most presented results will be j-V curves with the current density j, but the more common term "I-V curve" will be used synonymously.

The I-V results discussed in the following are only a selection of the large number of conducted experiments. In many cases, it proved difficult to obtain RS behavior due to the contacting issues already mentioned in sec. 3.3. Likewise, once resistive switching was observed, a further investigation with subsequent I-V measurements often was prevented by contact alterations or contact loss of the probe tips. The contact quality typically had an influence on the curve shape, since the a-GaO_x thin film devices themselves have rather low resistance. Re-contacting of a certain device usually led to the same basic switching pattern, but not to congruent I-V curves (e.g., fig. B.5a). The same or even more substantial variations were seen between RS devices on the same sample, i.e., between different TEs on the same a-GaO_x film. In most cases, they produced the same switching behavior, albeit with more or less differing curve shape (e.g., fig. B.5b). Sometimes even the RS pattern differed, as will be seen later. Since the as-prepared state of devices on the same sample can be assumed as identical, these variations certainly are related to the contact quality. The higher the contact resistance, the lower the fraction of the voltage that was actually applied to the RS device. The resulting uncertainty in the voltage history of each device is a significant obstacle for the comparison and interpretation of the I-V curves.

Besides these contact-related complications, there often were gradual variations between successive I-V sweeps due to asymmetric switching. When the HRS → LRS and LRS → HRS transitions do not completely undo each other on the atomic level, the result is a steady increase or decrease of the overall current density from cycle to cycle (e.g., fig. B.6). In this case, the internal state of the device is not the same before and after a full cycle, such that the device stays a bit more LRS'ish or HRS'ish with each further I-V sweep. Often, a form of self-balancing was observed, where the cycles became more and more congruent during these repeated I-V sweeps (e.g., fig. B.6a). This is a plausible effect, since the farther the deviation of the device from equilibrium either towards LRS or HRS, the stronger the transition in the opposite direction should be for a certain applied voltage.

5.2.1 Counter-8 Bipolar Switching

Counter-8 bipolar resistive switching was observed for a variety of Pt|a-GaO_x|ITO devices. Fig. 5.7 shows an overview of the respective I-V curves measured on four different samples. They each exhibit clearly visible counter-8 bipolar RS, with a gradual LRS → HRS transition at positive voltage and a gradual HRS → LRS transition at negative voltage. The same gradual RS processes were obtained in pulsed I-V measurements (fig. B.7). The I-V sweeps are quite symmetrical, even though the materials of TE and BE are not the same. The consecutive I-V curves are congruent, which demonstrates that the RS transitions are reversible and reproducible. While the four samples yield

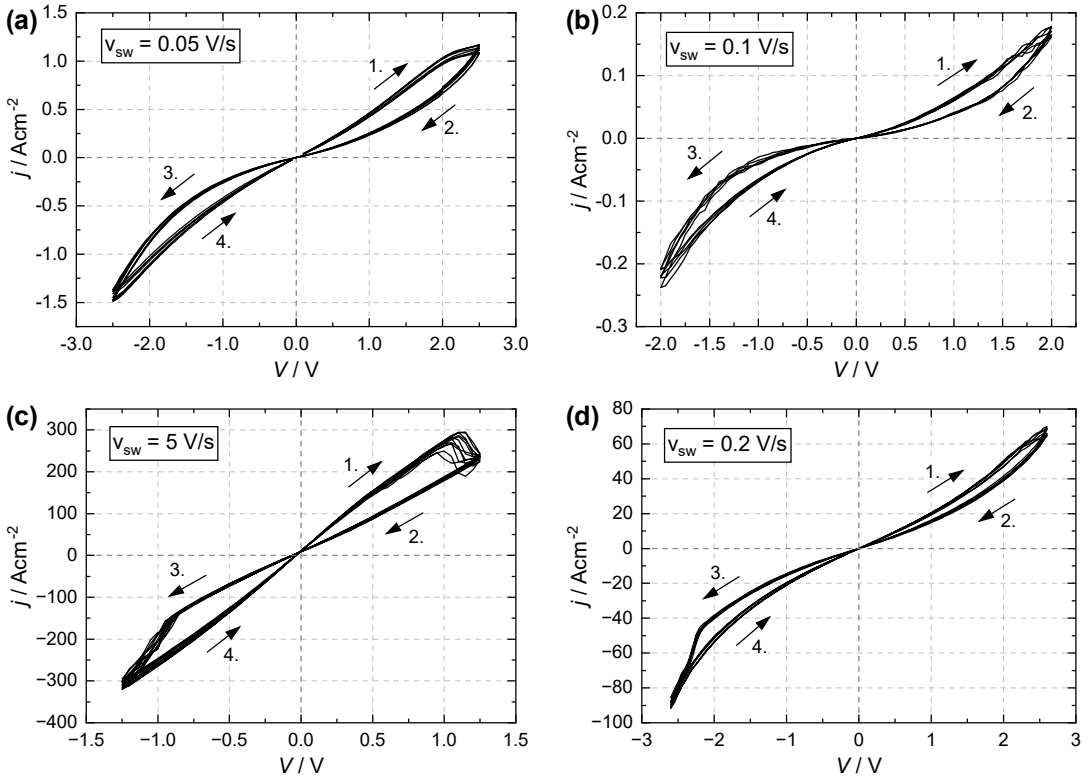


Figure 5.7: Counter-8 bipolar resistive switching of Pt|a-GaO_x|ITO devices on four different samples (contacted by either W or Au probe tips): (a) $x \approx 0.9$, $d \approx 60$ nm (PLD with 0.80 J/cm^2 in Ar, chamber #1), with sp-Pt-TE (TE-∅: 400 µm, CT1, $V_{max} = \pm 2.5$ V). (b) $x \approx 0.9$, $d \approx 90$ nm (PLD with 0.78 J/cm^2 in Ar, chamber #1), with sp-Pt-TE (TE-∅: 300 µm, CT3, $V_{max} = \pm 2.0$ V). (c) $x \approx 0.9$, $d \approx 100$ nm (PLD with 0.94 J/cm^2 in Ar, chamber #1), with c-ITO-BE (TE-∅: 100 µm, CT2, $V_{max} = \pm 1.25$ V). (d) $x \approx 1.0$, $d \approx 120$ nm (PLD with 0.79 J/cm^2 in Ar), with lithographic Pt-TE (TE: 20 × 20 µm^2, CT4, $V_{max} = \pm 2.6$ V). Each graph features several consecutive I-V sweeps (5–12 cycles). The numbered arrows indicate the direction of the switching behavior.

the same switching behavior, their I-V curves have differences. Fig. 5.7a and b look quite alike, which fits to the comparable preparation of these devices (e.g., sputtered Pt-TE). The kink shortly before reaching $+V_{max}$ in fig. 5.7a, however, is otherwise only visible in fig. 5.7c, where it is more pronounced and less gradual. Furthermore, fig. 5.7c also has a kink before $-V_{max}$, which is only reproduced in fig. 5.7d. These kinks likely mark the onset of the respective RS transitions and their

visibility seems to correlate with the extent of the switching processes. Fig. 5.7c also stands out due to its particularly linear I-V curves and the higher current density. The other three samples have non-linear I-V relations and orders of magnitude lower current density. The major difference of the device of fig. 5.7c to the others is its PLD-made c-ITO-BE, probably together with less interface contamination in comparison to the commercial ITO substrates, but it is unclear how that could have such an influence. The device of fig. 5.7d differs from the others in that a-GaO_x was deposited in chamber #2, interface oxidation was prevented by lithographic Pt-TE and inert Au probe tips were used for contacting. These differences apparently had no large impact on its RS curve.

The bipolarity of the counter-8 RS behavior was confirmed by repeated I-V sweeps with only positive or only negative voltage. Fig. 5.8 shows an example, measured on the same device as fig. 5.7d. In fig. 5.8a, the first triangular 0 V → +2.25 V → 0 V cycle led to a proper LRS → HRS transition, while the following repetitions essentially just reproduced the HRS curve, almost without further switching. Upon reversal of the voltage polarity, in fig. 5.8b, the first 0 V → −2.25 V → 0 V cycle triggered the expected HRS → LRS transition and the following cycles just yielded the LRS curve. These measurements demonstrate that continuous back-and-forth switching requires both voltage polarities.

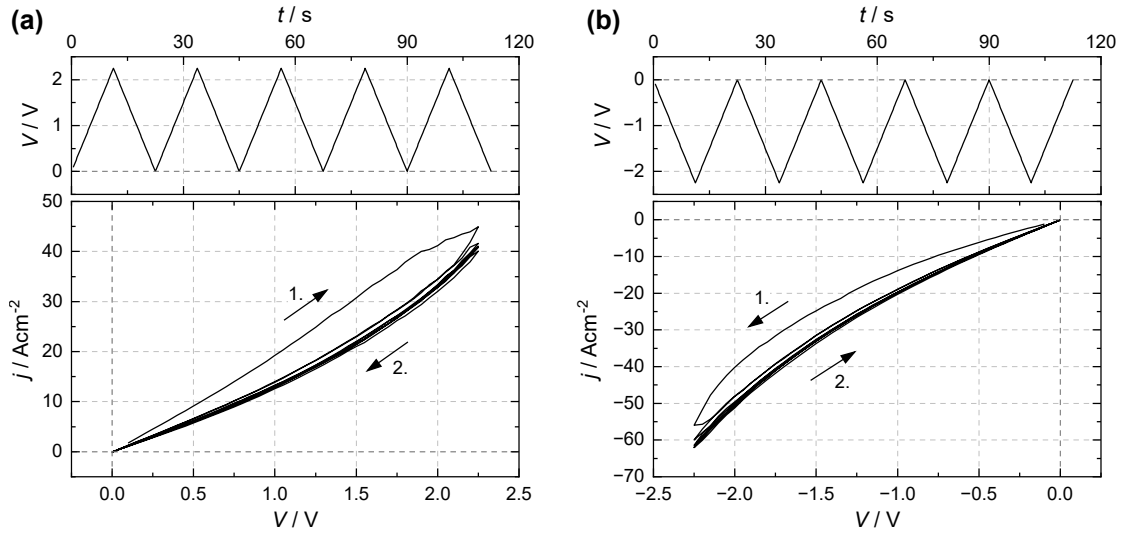


Figure 5.8: Bipolarity check for counter-8 RS of a Pt|a-GaO_x|ITO device with $x \approx 1.0$ and $d \approx 120$ nm (same as in fig. 5.7d; PLD with 0.79 J/cm^2 in Ar, lithographic Pt-TE: 20 × 20 µm^2, CT4). The top graphs show the applied voltage $V(t)$, while the bottom graphs display the resulting I-V curve. (a) Triangular I-V sweep with only positive voltage (V_{max} = +2.25 V, ν_{sw} = 0.2 V/s). (b) Triangular I-V sweep with only negative voltage (V_{max} = −2.25 V, ν_{sw} = 0.2 V/s). In each graph, the numbered arrows refer to the first 0 V → V_{max} → 0 V cycle, which is the only one with a notable switching effect.

The results discussed so far were obtained with tungsten and gold probe tips as contacting method (CT1–4). Due to the aforementioned difficulties in reaching steady RS with the probe tips, the I-V curves generally do not represent the as-prepared state of the devices. Instead, they typically have a voltage history with contact variations, which makes it difficult to interpret how the RS behavior

starts. What can be said, at least, is that there was no classical electroforming with an abrupt soft-breakdown, as it would be expected for a filamentary RS mechanism. A clearer account of the development of counter-8 bipolar RS from the as-prepared state was obtained by contacting the RS devices with Au wires (CT5). This has the advantage of reliable, immutable contacts right from the beginning of the measurements. Fig. 5.9 shows the first 16 *I*-*V* cycles of such a wired Pt|a-GaO_x|ITO device. The *I*-*V* sweeps up to V_{max} = ±1.25 V just produced S-shaped, congruent curves without

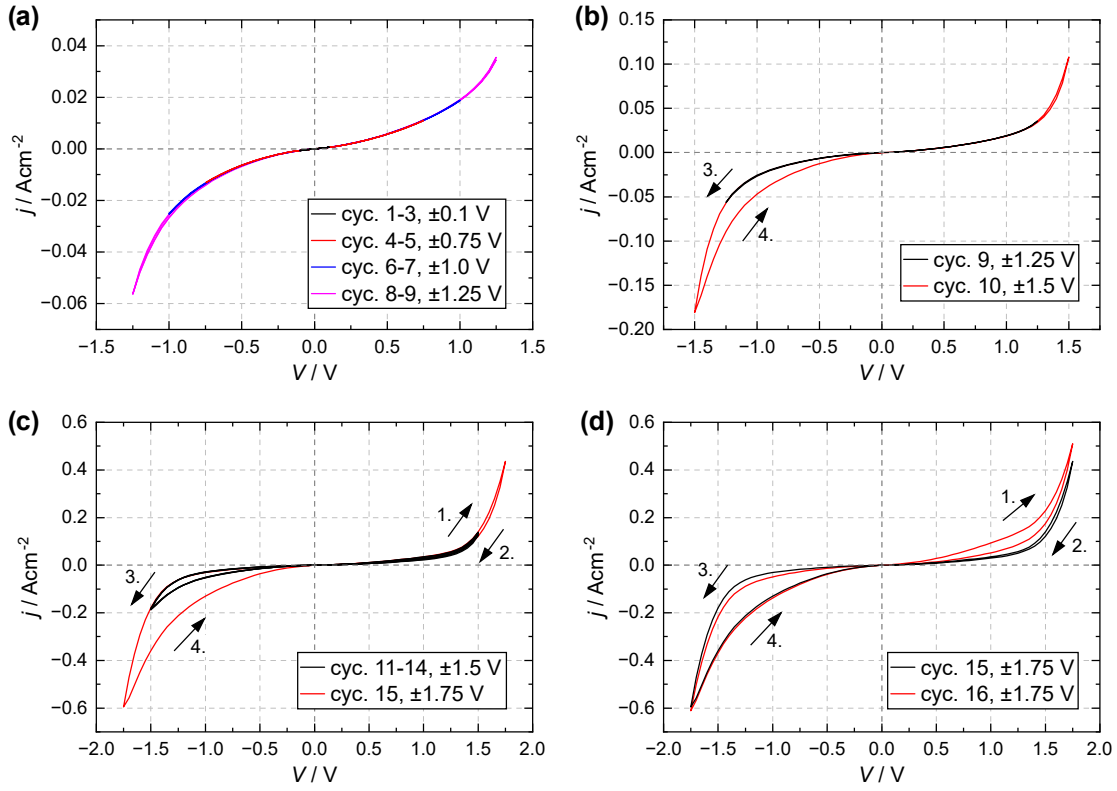


Figure 5.9: Development of counter-8 bipolar RS from the as-prepared state of a Pt|a-GaO_x|ITO device with $x \approx 0.9$ and $d \approx 90$ nm (same sample as fig. 5.7b; PLD with 0.78 J/cm² in Ar, chamber #1) that was contacted by Au wires (CT5, TE area: 0.592 mm², v_{sw} = 0.2 V/s). (a) *I*-*V* cycles with successively increasing V_{max} from ±0.1 V to ±1.25 V. (b) The last *I*-*V* cycle with V_{max} = ±1.25 V and the next cycle with higher V_{max} = ±1.5 V. (c) Congruent *I*-*V* cycles with V_{max} = ±1.5 V, together with the next cycle with even higher V_{max} = ±1.75 V. (d) The first and second *I*-*V* cycles with V_{max} = ±1.75 V.

resistive switching (fig. 5.9a). The first sweep with V_{max} = ±1.5 V (cyc. 10) still did not yield any switching at positive voltage and followed the previous cycle exactly until reaching −1.25 V, but then a pronounced HRS → LRS transition occurred at around −1.5 V (fig. 5.9b). The next sweeps with V_{max} = ±1.5 V (cyc. 11–14) then also featured a weak LRS → HRS transition at positive voltages, together with the strong HRS → LRS transition at negative voltage that was already present in cyc. 10 (fig. 5.9c). Equivalent observations were made after V_{max} was increased to ±1.75 V. The first sweep with V_{max} = ±1.75 V (cyc. 15) only had a weak transition at positive voltage, just like the previous sweeps with V_{max} = ±1.5 V, and again an equal HRS curve until −1.5 V (fig. 5.9c). Then, in the range between −1.5 V and −1.75 V, an even more pronounced HRS → LRS transition occurred than

observed for V_{max} = ±1.5 V (fig. 5.9c). The second sweep with V_{max} = ±1.75 V (cyc. 16) then also produced a quite strong transition at positive voltage (fig. 5.9d). This measurement sequence reveals that the counter-8 bipolar RS starts with an HRS → LRS transition at negative voltage, without a prior electroforming process. The switching at negative voltage is the prerequisite for the opposite LRS → HRS transition at positive voltage, which just reverses what happened at negative voltage. Accordingly, the LRS curve depends on V_{max}, while the HRS curve stays almost the same. The final cycle 16 in fig. 5.9d has approximately a similar shape as fig. 5.7b, where a tip-contacted device on the same sample was measured. This means that the short annealing at 150 °C, which is required to attach the Au wires, had no decisive influence on the curve shape or the switching behavior. It is possible, though, that the annealing process prevented some initial curve variations that otherwise would have resulted for example from Joule heating.[99]

Precisely the same development of counter-8 RS from the as-prepared state was also observed for a second wired device on the same sample, which confirms that this is a reproducible finding. Afterwards, as shown in fig. 5.10, both devices exhibited steady counter-8 bipolar resistive switching at V_{max} = ±1.75 V, with remarkably equal curve shape and rather similar current density.

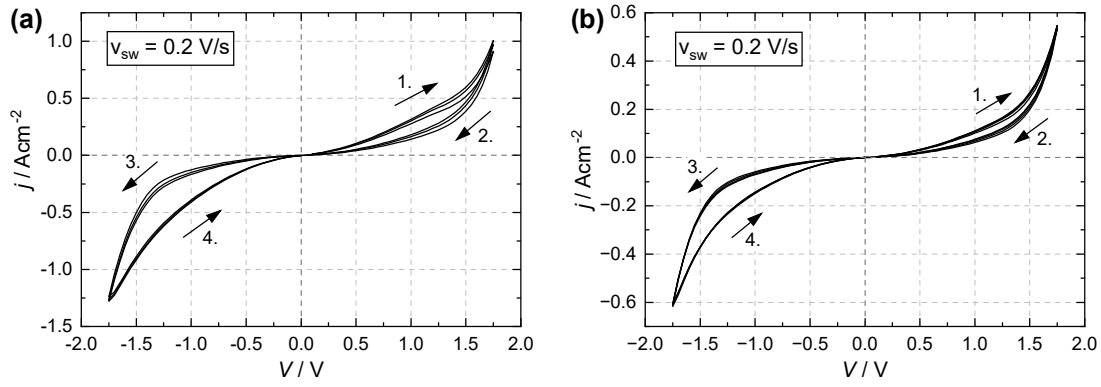


Figure 5.10: Reproducible counter-8 bipolar RS of two Pt|a-GaO_x|ITO devices on the same sample ($x \approx 0.9$, $d \approx 90$ nm, PLD with 0.78 J/cm² in Ar, chamber #1; CT5, TE areas: 0.465 mm² for (a) and 0.592 mm² for (b), V_{max} = ±1.75 V). The device measured in (b) is the same as in fig. 5.9.

The influence of the sweep rate v_{sw} on counter-8 bipolar switching was investigated next. Fig. 5.11a shows that increasing v_{sw} has a significant effect on the *I-V* sweeps. The higher v_{sw}, the less pronounced are the RS transitions, as expected. At v_{sw} = 10 V/s, the counter-8 bipolar RS is barely visible anymore, while v_{sw} = 450 V/s yields no resistive switching at all. The same was also observed for the counter-8 switching of other Pt|a-GaO_x|ITO samples. Pulsed switching experiments yielded a comparable time-dependence, as shown in fig. 5.11b. The longer the pulse length, the more pronounced is the counter-8 switching behavior, i.e., the more differ the read currents after positive and negative write pulses. The shape of the write pulses also gives further insight into the switch-

[99] The discussion of thermal relaxation later in sec. 5.3.1 will show that short heating causes slight changes of the amorphous structure of a-GaO_x, which reduces the bulk conductivity and probably also affects other properties like the work function or interfacial defect states.

ing processes, since they characterize the transitions at constant voltage. During the write pulses in fig. 5.11b, the current changes approximately exponentially with time. This could mean that the ionic motion in the underlying RS mechanism has an exponential time dependence. It is more probable, though, that the ion transport changes some property of the sample that has an exponential effect on the measured current (e.g., φ_B for thermionic field emission, ΔE for mobility edge conduction or T_0 for variable-range hopping). The exponential shape of the write pulses also underlines the gradual nature of the RS transitions. Furthermore, their steep initial slope could explain the kinks, where the resistance changed rather fast in the previously discussed I-V sweeps (e.g., fig. 5.7c).

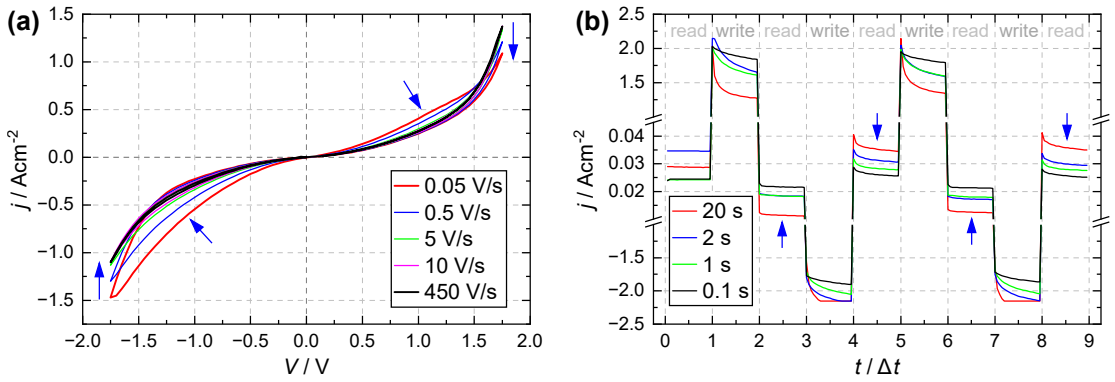


Figure 5.11: Influence of the sweep rate or the pulse length on the counter-8 bipolar RS of Pt|a-GaO_x|ITO (same device as in fig. 5.10a; $x \approx 0.9$, $d \approx 90$ nm, PLD with 0.78 J/cm^2 in Ar, chamber #1; CT5, TE area: 0.465 mm^2). (a) I-V sweeps with $V_{max} = \pm 1.75$ V and different sweep rates v_{sw} from 0.05 V/s up to 450 V/s. (b) I-V pulse sequences with $V_{write} = \pm 1.75$ V, $V_{read} = +0.1$ V and different pulse lengths Δt of 20 s, 2 s, 1 s and 0.1 s. The negative write pulses are truncated at $j = -2.15$ A/cm^2 because the compliance current of 10 mA was reached. The blue arrows indicate the trends w.r.t. increasing v_{sw} and decreasing Δt.

The shape of the read pulses in fig. 5.11b already indicates that HRS and LRS remain sufficiently stable at low or no voltage. This state retention was investigated further by monitoring the resistance states after different waiting times. First, a device was switched into an LRS by applying –1.75 V for 100 s (fig. 5.12a). Then, I-V sweeps with $v_{sw} = 400$ V/s were carried out both immediately and after different waiting times of up to 3 min (fig. 5.12b). As discussed above, these fast sweeps do not induce resistive switching. Instead, they reflect the I-V characteristics of the momentary device state without further switching. The respective curves for the LRS in fig. 5.12b are all congruent, which means that the resistance state did not change during the waiting times where the voltage was disabled. Afterwards, the same procedure was repeated for an HRS, by applying +1.75 V for 100 s (fig. 5.12a), followed by fast I-V sweeps with $v_{sw} = 400$ V/s (fig. 5.12b). Again, the cycles are congruent and no change of the resistance state is observed. This confirms that the LRS and HRS are stable states and not the result of a purely electronic, capacitive charging. The elongated write pulses in fig. 5.12a have the same exponential shape as those in fig. 5.11b. The fast I-V sweeps without superimposed resistive switching in fig. 5.12b are highly symmetric, with $j(V) = -j(-V)$ in the voltage range from approximately –1.5 V to +1.5 V (fig. B.8). This is quite unexpected for such asymmetric devices where different materials were used for the two electrodes.

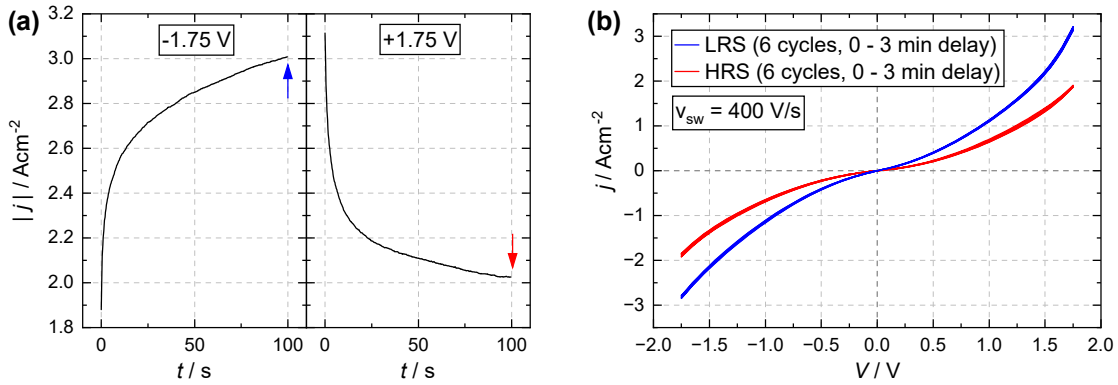


Figure 5.12: State retention testing by fast I-V sweeps, which were recorded with up to 3 min delay after triggering counter-8 RS transitions in Pt|a-GaO_x|ITO (same device as in fig. 5.11; $x \approx 0.9$, $d \approx 90$ nm, PLD with 0.78 J/cm^2 in Ar, chamber #1; CT5, TE area: 0.465 mm^2). (a) HRS → LRS transition at a constant applied voltage of −1.75 V for 100 s (left) and LRS → HRS transition at +1.75 V for 100 s (right). (b) I-V sweeps with $V_{max} = \pm 1.75$ V at a high sweep rate of 400 V/s, which visualize the pure electron transport characteristics of the states reached in (a), without further switching (18 ms for a full I-V cycle). The colored arrows in (a) point to the LRS and HRS that were investigated in (b). The fast I-V sweeps were measured immediately after the respective transition in (a) and then repeated after different waiting times of up to 3 min (voltage disabled in-between).

Another way to test the state retention is to interrupt an I-V sweep and resume it after a certain waiting time Δt_{off}. Fig. 5.13a shows measurements, where the voltage was turned off for 10 min after the LRS → HRS transition at positive voltage. When the sweep continued after the waiting time, the current was essentially equal to before the pause, which shows that the HRS did not relax towards the previous LRS. In fact, the HRS current was even slightly lower after the pause. Measurements with different waiting times yielded a significant trend of more negative Δj for longer Δt_{off} (fig. 5.13b). This is not a further switching effect, since the voltage is zero during the interruption. Most likely, the HRS current before the pause contains a small contribution from a capacitive discharge current, which wears off over time even without an applied voltage.

Besides the Pt|a-GaO_x|ITO devices, also samples with Au or Pt as bottom electrode were investigated. Fig. 5.14 shows the I-V curves of a Pt|a-GaO_x|Au and a Pt|a-GaO_x|Pt device, which were both prepared like the one with c-ITO-BE in fig. 5.7c. The Pt|a-GaO_x|Au variant in fig. 5.14a exhibits pronounced counter-8 bipolar RS. Its transitions look like those of the Pt|a-GaO_x|ITO device in fig. 5.7c and also the overall current density is quite comparable. The HRS curve is less linear, though, for the Au-BE, which could be related to Au's higher work function. Moreover, the Au-BE seems to enable faster counter-8 switching that is clearly visible at $v_{sw} = 25$ V/s, whereas for ITO the limit seems to be at around 10 V/s. A further peculiarity of the Pt|a-GaO_x|Au device is the difference between the subsequent I-V cycles. The measurements were carried out in sets of 3 directly consecutive cycles. The first cycle of each set differs from the second and the third in the positive voltage range, where it has an earlier LRS → HRS transition, a significantly lower LRS current and a small loop near the maximum voltage (fig. 5.14a). This was not observed for the devices with ITO-BE. In the short time between the sets, something seems to change within the Pt|a-GaO_x|Au

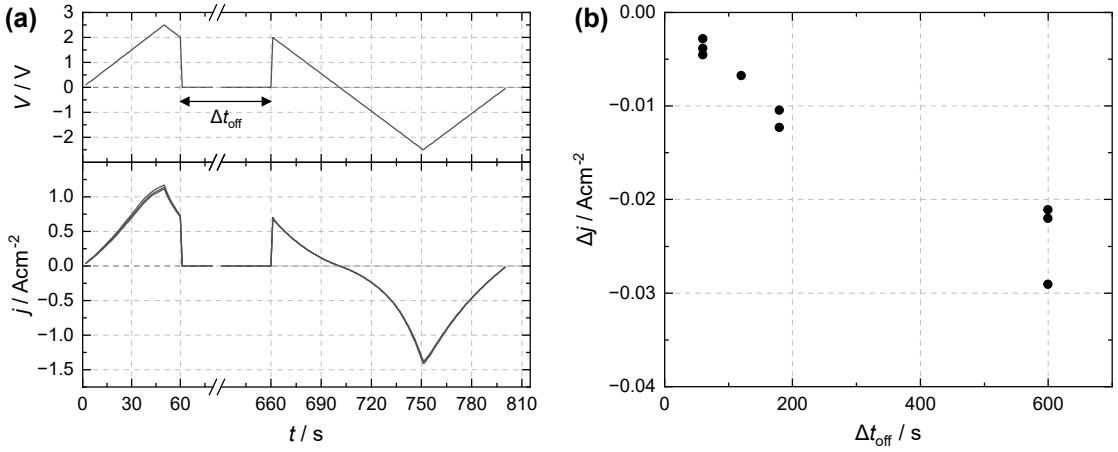


Figure 5.13: State retention testing by interrupting I-V sweeps of Pt|a-GaO_x|ITO at +2.0 V for different times Δt_{off}, in which the voltage was disabled (same device as in fig. 5.7a; $x \approx 0.9$, $d \approx 60$ nm, PLD with 0.80 J/cm^2 in Ar, chamber #1, sp-Pt-TE-∅: 400 µm; CT1, $V_{max} = \pm 2.5$ V). (a) Applied voltage $V(t)$ (top) and measured current density $j(t)$ (bottom) for a measurement with $\Delta t_{off} = 10$ min (3 cycles). (b) Difference Δj of the current density at +2.0 V before and after the interruption, for varied Δt_{off} ($\Delta j = j_{after} - j_{before}$).

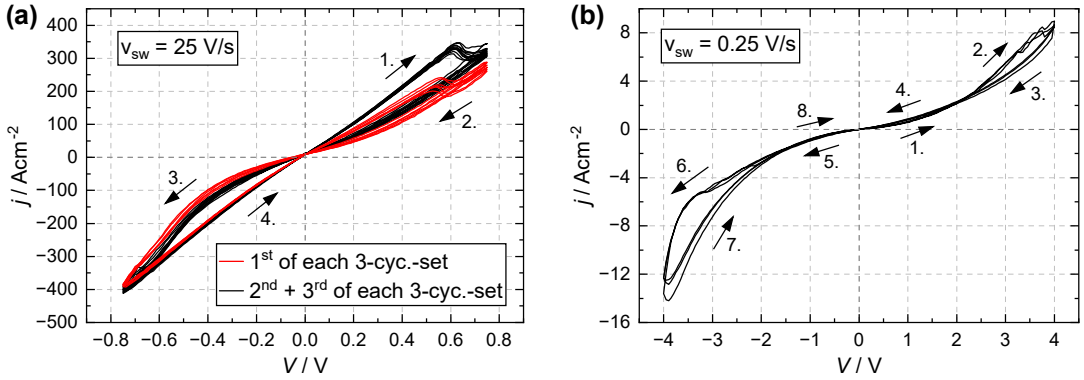


Figure 5.14: Counter-8 bipolar resistive switching of devices with a different BE material but otherwise equal preparation conditions as the Pt|a-GaO_x|ITO device in fig. 5.7c ($x \approx 0.9$, $d \approx 100$ nm, PLD with 0.94 J/cm^2 in Ar, chamber #1, Pt-TE-∅: 100 µm, CT2). (a) I-V sweeps of Pt|a-GaO_x|Au with $V_{max} = \pm 0.75$ V (6 sets of 3 consecutive cycles, approx. 10–20 s delay between sets). (b) I-V sweeps of Pt|a-GaO_x|Pt with $V_{max} = \pm 4.0$ V, which have crossing points at approx. +2 V and −1.5 V.

device. The original curve shape is always recovered after the first cycle, which means that this is just a temporary effect. It could be due to a capacitive component that is charged at negative voltage and that discharges at positive voltage. This would increase the positive current if it is directly preceded by negative voltage. However, the deviation of the first cycle is also present at rather low sweep rates (e.g., 2.5 V/s), which makes a purely electronic discharge less likely. Instead, it seems like a part of the underlying ion movement during the HRS → LRS transition at negative voltage reverses itself quite fast.

The Pt|a-GaO_x|Pt device in fig. 5.14b also yielded counter-8 bipolar RS, albeit with a more complex shape. It not only has much lower current density, but it also requires high voltages to trigger resis-

tive switching (no RS for $V_{max} < \pm 3$ V). This could mean that in this case the BE-interface primarily acts as a high series resistance, probably because of Pt's high work function. Furthermore, the *I-V* sweeps have subtle crossing points at about +2 V and −1.5 V, which result from changes of the *I-V* curves' non-linearity. It is thus debatable, whether this is true counter-8 bipolar RS. For example, monitoring the ±4 V switching at a read voltage below 1.5 V would yield an HRS → LRS transition after +4 V and an LRS → HRS transition after −4 V for the read current, which corresponds to figure-8 bipolar RS. Interestingly, the Pt|a-GaO_x|Pt device also exhibits the least symmetric *I-V* curves from the three tested BE variants. This could be due to the different areas of TE and BE, though the latter should not be relevant due to the low a-GaO_x thickness. Most likely, the device is actually asymmetric because the Pt-TE interface was exposed to air before the TE deposition. Judging from the discussion of surface oxidation in sec. 5.1.3, this could have led to substantially different interfaces. The results for Pt|a-GaO_x|Pt thus indicate that surface oxidation could be relevant even for Pt|a-GaO_x|ITO devices – in contrast to what was concluded from fig. 5.7d. All in all, the observed differences for the three variants demonstrate that the BE material certainly has an influence on the RS curves. However, whether the a-GaO_x|BE interface is actively involved in the counter-8 RS processes remains unclear at this point, since the results for Au- and c-ITO-BE are rather similar and since the Pt-BE probably just causes a non-linear series resistance.

5.2.2 Figure-8 Bipolar Switching

Several Pt|a-GaO_x|ITO samples exhibited well-defined figure-8 bipolar resistive switching, instead of the counter-8 behavior described above. Fig. 5.15 shows an overview of four devices located on two different samples. Each of them clearly features an HRS → LRS transition at positive voltage and an LRS → HRS transition at negative voltage. The figure-8 behavior thus was reproducible both for different devices on the same sample (e.g., fig. 5.15a vs. b) and across different samples (fig. 5.15a,b vs. c,d). The general shape of the figure-8 *I-V* curves is quite comparable to the previously discussed counter-8 curves. They are rather symmetric and the transitions are gradual except for some kinks. The figs. 5.15a–c have a characteristic kink at negative voltage that looks remarkably similar to the one observed at positive voltage in some of the counter-8 *I-V* curves (e.g., fig. 5.7c). Except for some minor difference in curvature, fig. 5.15b even appears like a mirrored version of fig. 5.14a. One could get the impression that the electrical connections were simply swapped by accident, but that was ruled out. Instead, the common features of the figure-8 and counter-8 *I-V* curves are a hint that the responsible RS mechanisms follow the same principles but have opposite orientation.

Deducing the cause for the reversed switching direction is not simple, though, because there is no unique property that sets the samples with figure-8 switching apart from those with counter-8 RS. Neither the bulk O/Ga ratio, nor the device dimensions or the voltage history seem to be decisive. It is noticeable, however, that almost all of the samples with counter-8 RS were exposed to air before the TE deposition, while those with figure-8 RS mostly have lithographic TEs without interface oxidation (incl. those not yet discussed). There are only two exceptions to this. The lithographically

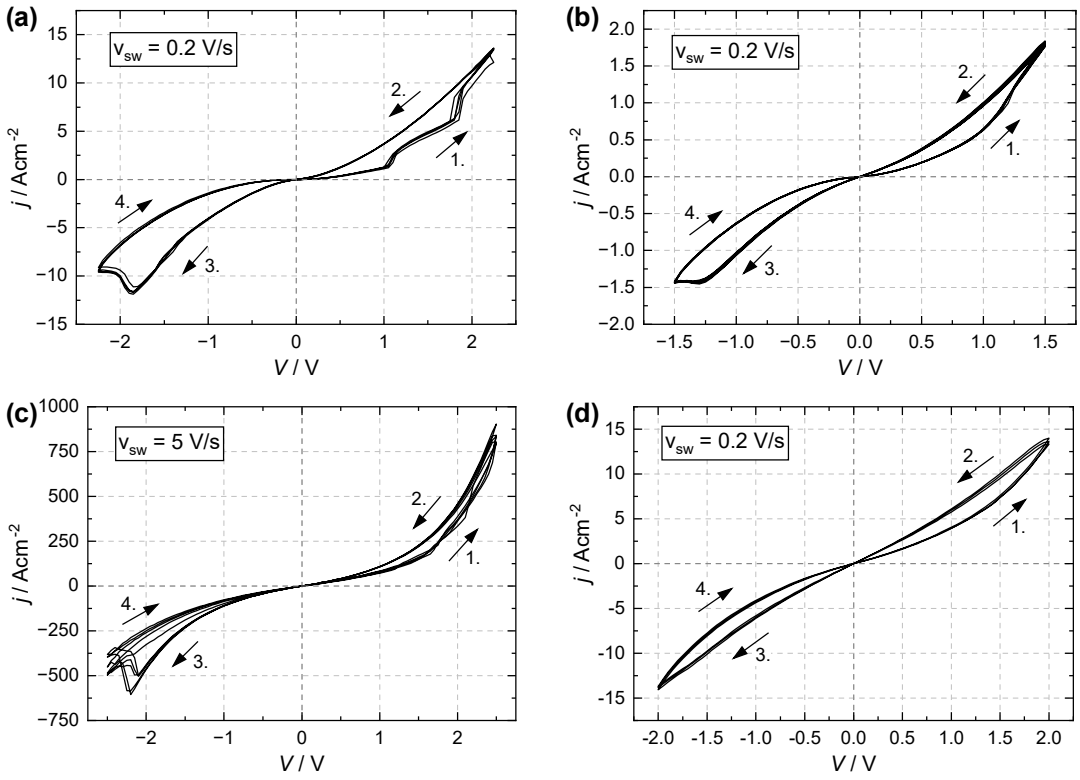


Figure 5.15: Figure-8 bipolar resistive switching of two Pt|a-GaO_x|ITO devices on each of two different samples (contacted by either W or Au probe tips): (a,b) $x \lesssim 0.8$, $d \approx 50$ nm (PLD with less than 0.20 J/cm^2 in Ar, chamber #1), with eb-Pt-TEs (TE-∅: 100 µm for (a) and 200 µm for (b), CT3, $V_{max} = \pm 2.25$ V for (a) and ±1.5 V for (b)). (c,d) x in the range of 0.98–1.07 according to EELS, $d \approx 60$ nm (PLD with 0.79 J/cm^2 in Ar), with lithographic Pt-TEs (TE: 20 × 20 µm^2 for (c) and 50 × 50 µm^2 for (d), CT4, $V_{max} = \pm 2.5$ V for (c) and ±2.0 V for (d)). Each graph features several consecutive I-V sweeps (3–10 cycles).

patterned sample of fig. 5.7d showed counter-8 bipolar switching, although it was prepared exactly like the figure-8-switching sample of fig. 5.15c,d, only with doubled a-GaO_x thickness. Yet, for this specific sample, it could be that the voltage history is more relevant than the preparation, because it did not exhibit counter-8 RS right from the start but only after a lot of inconclusive I-V cycles (incl. contact variations and re-contacting). The second exception to the above hypothesis is the sample of fig. 5.15a,b, which showed figure-8 RS even though the eb-Pt-TEs were prepared with a shadow mask. The sample was sealed in a vacuum bag directly after the a-GaO_x deposition and until it was placed in the e-beam chamber (approx. 2 months), but the HAXPES results in sec. 5.1.3 showed that this does not prevent surface oxidation. In summary, it is not entirely certain but it seems like the TE interface oxidation is the property that mainly determines the switching direction.

The bipolarity of the figure-8 RS behavior was confirmed in the same way as for the counter-8 RS (cf. fig. 5.8). Fig. 5.16 shows a respective measurement for the device of fig. 5.15d, while a likewise example for the device of fig. 5.15a is displayed in fig. B.9. In each case, only the first triangular cycle at positive voltage produced a significant HRS → LRS transition, while the subsequent cycles essentially just reproduced the LRS curve. Similarly, only the first triangular cycle at negative

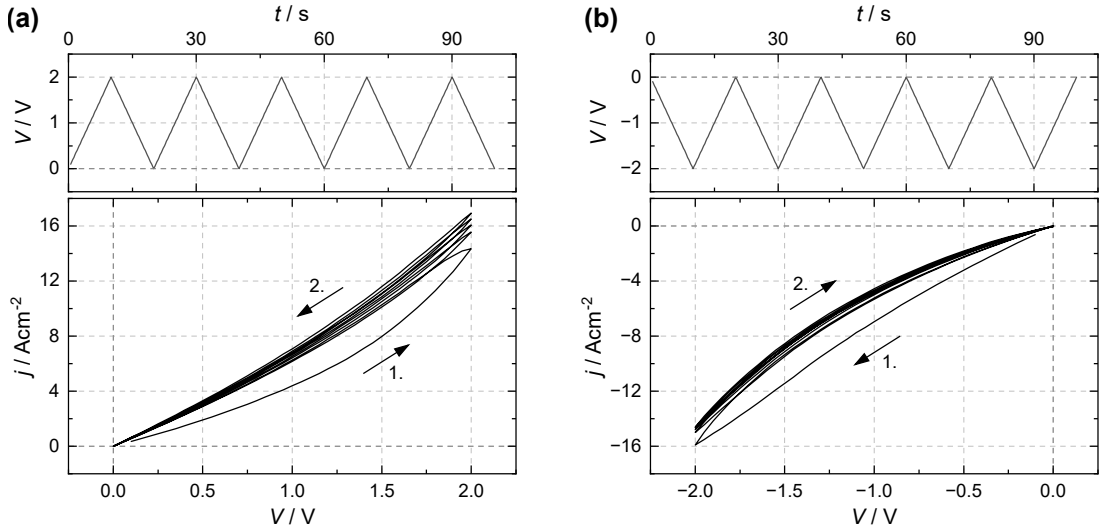


Figure 5.16: Bipolarity check for figure-8 RS of a Pt|a-GaO_x|ITO device with x in the range of 0.98–1.07 (EELS) and $d \approx 60$ nm (same as in fig. 5.15d; PLD with 0.79 J/cm^2 in Ar, lithographic Pt-TE: 50 × 50 µm^2, CT4). The top graphs show the applied voltage $V(t)$, while the bottom graphs display the resulting I-V curve. (a) Triangular I-V sweep with only positive voltage (V_{max} = +2.0 V, ν_{sw} = 0.2 V/s). (b) Triangular I-V sweep with only negative voltage (V_{max} = −2.0 V, ν_{sw} = 0.2 V/s). In each graph, the numbered arrows refer to the first 0 V → V_{max} → 0 V cycle, which is the only one with a notable switching effect.

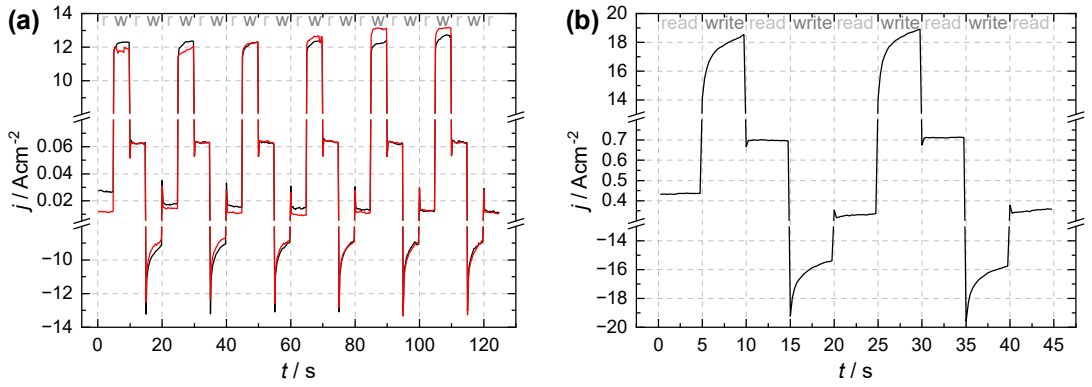


Figure 5.17: Pulsed figure-8 bipolar RS of Pt|a-GaO_x|ITO devices on two different samples (contacted by W or Au probe tips): (a) $x \lesssim 0.8$, $d \approx 50$ nm (PLD with less than 0.20 J/cm^2 in Ar, chamber #1), with eb-Pt-TE (same device as in fig. 5.15a, TE-∅: 100 µm, CT3, V_{write} = ±2.25 V, V_{read} = +0.1 V, Δt = 5 s). The graph shows two consecutive pulse sequences (black = 1st, red = 2nd). (b) x in the range of 0.98–1.07 (EELS), $d \approx 60$ nm (PLD with 0.79 J/cm^2 in Ar), with lithographic Pt-TE (same device as in fig. 5.15d, TE: 50 × 50 µm^2, CT4, V_{write} = ±2.0 V, V_{read} = +0.1 V, Δt = 5 s).

voltage yielded an LRS → HRS transition. This confirms that the reversal of the voltage polarity is mandatory also for continuous figure-8 RS, just like observed for counter-8 switching.

The figure-8 bipolar RS was also investigated by pulsed I-V measurements, as shown in fig. 5.17. The write pulses yielded gradual, exponentially decaying transitions, in accordance with respective I-V sweeps (cf. fig. 5.15a,d). The steep slope at the beginning of the negative write pulses in fig. 5.17a, for example, fits to the characteristic kink observed in fig. 5.15a. The read currents of the HRS and

the LRS at V_{read} = +0.1 V differ by a factor of 2–3. There is no visible decay of the read currents towards each other, which demonstrates the stability of the resistance states.

So far, the figure-8 and counter-8 switching behaviors appeared rather similar except for the reversed direction. However, a major difference was observed for the switching speed. Fig. 5.18 shows fast *I*-*V* sweeps of the same devices as in the already discussed figure-8 overview. Those in fig. 5.18a,c exhibit figure-8 bipolar switching even at a sweep rate of 450 V/s, which is the highest that the instruments allow. This is much faster than all counter-8 results, where the highest

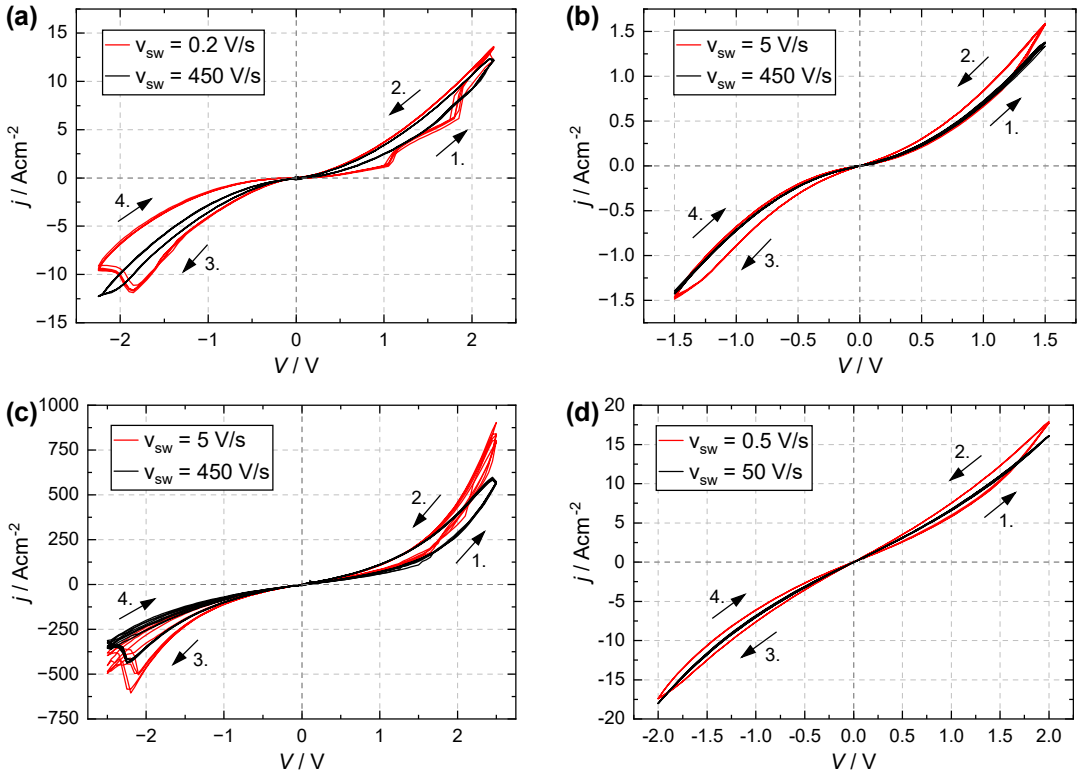


Figure 5.18: Influence of the sweep rate on the figure-8 bipolar RS of four Pt|a-GaO_x|ITO devices (same as in fig. 5.15, in the same order). (a,b) $x \lesssim 0.8$, $d \approx 50$ nm (PLD with less than 0.20 J/cm^2 in Ar, chamber #1), with eb-Pt-TEs (TE-∅: 100 µm for (a) and 200 µm for (b), CT3, V_{max} = ±2.25 V for (a) and ±1.5 V for (b)). (c,d) x in the range of 0.98–1.07 (EELS), $d \approx 60$ nm (PLD with 0.79 J/cm^2 in Ar), with lithographic Pt-TEs (TE: 20 × 20 µm^2 for (c) and 50 × 50 µm^2 for (d), CT4, V_{max} = ±2.5 V for (c) and ±2.0 V for (d)).

sweep rate with visible switching was 10 V/s (e.g., fig. 5.11a). In fig. 5.18c, increasing the sweep rate by about two orders of magnitude even had a remarkably weak impact on the shape of the figure-8 *I*-*V* curve, in contrast to what was observed in fig. 5.11a for counter-8 RS. The *I*-*V* sweeps in fig. 5.18b,d demonstrate that figure-8 switching is not always that fast. For the device in fig. 5.18b, the highest v_{sw} with slightly visible RS was 50–100 V/s, while for fig. 5.18d the limit was even lower, at about 5 V/s. In comparison to their faster counterparts in fig. 5.18a,c, which are located on the same respective samples, these slower devices have a larger TE area and their sweeps were carried out with lower V_{max}. Moreover, they yielded much lower current densities than the fast devices.

Since the applied voltage drives the switching processes, it is plausible that the lower V_{max} led to the reduced switching speed. This notion is supported by the kinks in fig. 5.18a,c, which indicate that certain voltages must be exceeded to trigger fast switching. The contact quality may have affected the RS speed as well, since a high series resistance would decrease the voltage drop over the actual device. It is unlikely, though, that the lower V_{max} and eventual contact variations are the sole reasons for the largely different current densities. These are likely related to the increase of the TE areas by the factors 4 and 6.25. In presence of a significant series resistance due to the probe tips, a larger TE reduces the device resistance and thus increases the fraction of the applied voltage that drops over the series resistance. Since the latter does not scale with the TE area, this leads to an artificial decrease of the current density.

The influence of the TE size was also investigated more systematically. Fig. 5.19 shows the results of four devices on one sample, each with a different TE size. They all exhibit figure-8 bipolar switching at v_{sw} = 450 V/s. The curves are almost equal in comparison to v_{sw} = 1 V/s (fig. B.10), only with a slightly reduced separation between HRS and LRS. Two clear trends are visible, in accordance with the previous fig. 5.18: The smaller the TE, the higher is the absolute current density and the more S-shaped are the *I*-*V* curves. Ideally, the current density and the curve shape should not depend on the TE area for non-filamentary resistive switching, but here a systematic influence is observed. The current is not equal either, in contrast to what would be expected for a filamentary RS mechanism (fig. B.10a). Instead, the current decreases with decreasing TE area, albeit not enough for the current density to stay equal. The most probable explanation for this is the aforementioned effect of an area-independent series resistance due to the contacts.[100] For a large TE, the resistance of the device is low and the *I*-*V* curve resembles the linear characteristics of the series resistance. Reducing the TE size then reveals more and more the S-shaped *I*-*V* relation of the Pt|a-GaO_x|ITO devices.

A comparison between figs. 5.18c,d and 5.19 is interesting since the two samples essentially just differ in the PLD fluence (0.79 J/cm^2 vs. 4.90 J/cm^2) and in the a-GaO_x thickness (60 nm vs. 130 nm). Despite these differences, the two device pairs with equal TE sizes have remarkably similar curve shapes and switching behaviors (figs. 5.18c vs. 5.19d and 5.18d vs. 5.19b). This is another example for the reproducibility of the figure-8 bipolar switching and it indicates that variations of the O/Ga ratio and the film thickness in this range are of minor importance for the RS mechanism.

At the beginning of this section, it was suspected that the TE interface oxidation plays a role in whether counter-8 or figure-8 switching occurs. This hypothesis was tested further by preparing a Pt|a-$GaO_{1.5}$|a-GaO_x|ITO device, where the approximately 15 nm thin a-$GaO_{1.5}$ layer represents an artificial surface oxidation. The thickness of the a-$GaO_{1.5}$ layer is approximately similar to the surface oxidation length observed in the EELS measurements (sec. 5.1.3). However, whereas the

[100] Note that the diameter even of the smallest TE is still about two orders of magnitude larger than the a-GaO_x film thickness. Therefore, the electric field below the TE should be approximately homogeneous in lateral direction and the area ratio between TE and BE should have no influence on the switching behavior (in contrast to the FEM simulations in [30]).

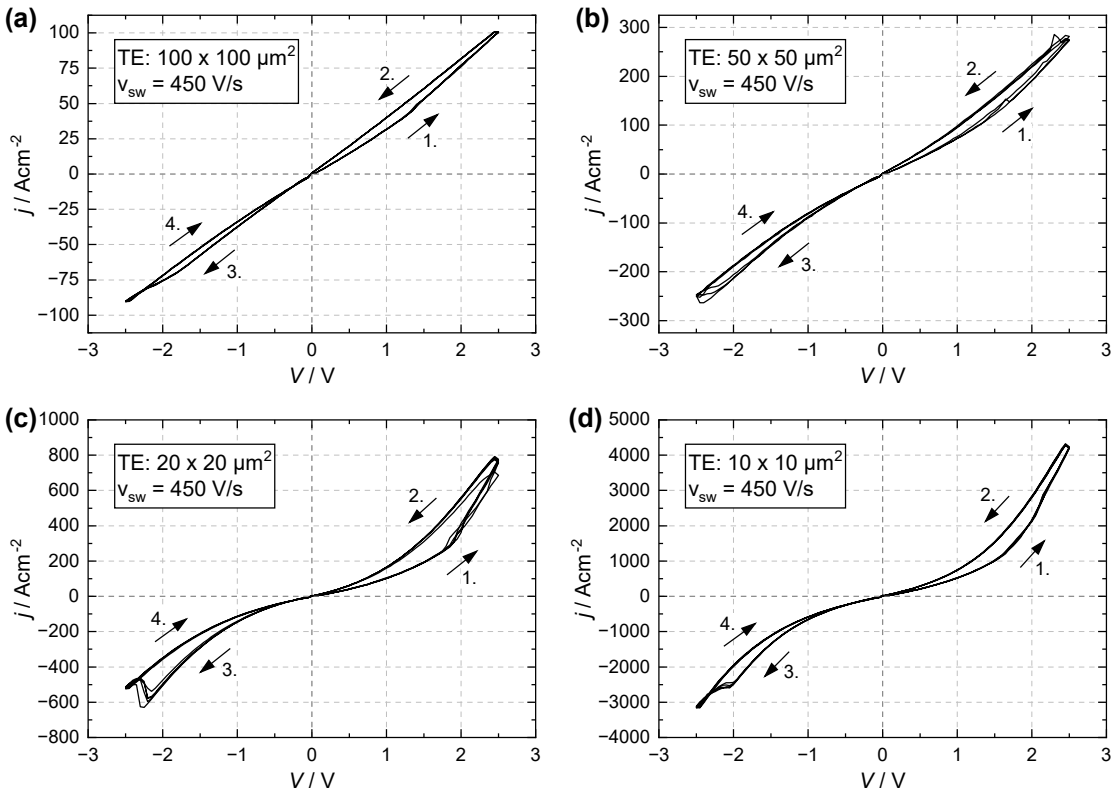


Figure 5.19: Comparison of the figure-8 bipolar RS of four Pt|a-GaO_x|ITO devices with different TE sizes on the same sample ($x \approx 1.1$, $d \approx 130$ nm, PLD with 4.90 J/cm² in Ar; lithographic Pt-TEs, CT2, $V_{max} = \pm 2.5$ V).

sample in this experiment has $x \approx 1.5$ over this whole length, the devices with natural interface oxidation have a gradual decay of x towards the bulk value (cf. fig. 5.6). This is most likely the reason why the I-V curves of the Pt|a-$GaO_{1.5}$|a-GaO_x|ITO device in fig. 5.20a differ from all previously discussed results. Instead of the symmetric S-shape, they are more diode-like with a convex curvature over the whole voltage range. Furthermore, the RS behavior is not counter-8 but figure-8 bipolar, in contrast to the expectation. Due to the differing curve shape, this experiment is likely not a valid counterexample for the influence of interface oxidation on the RS direction. Apparently, the Pt|a-$GaO_{1.5}$|a-GaO_x interface acts like a classical metal-insulator-semiconductor diode and dominates the I-V relation. It is unclear, though, how figure-8 transitions could occur in this region. These did not result from a complex voltage history, at least, but were visible right from the beginning of the I-V measurements (fig. B.12). The first cycle with $V_{max} = \pm 1.0$ V already triggered a strong HRS → LRS transition at positive voltage – in contrast to the counter-8 RS where the first transition occurred at negative voltage (cf. fig. 5.9).

Increasing the voltage amplitude from ±1.5 V to ±2.0 V for the Pt|a-$GaO_{1.5}$|a-GaO_x|ITO device led to an abrupt increase of the current density at about −1.7 V (fig. 5.20b). This sudden transition happened in the first cycle after increasing V_{max}. The following sweeps with $V_{max} = \pm 2.0$ V again produced figure-8 bipolar RS, but with an almost linear curve shape. The RS behavior both before and after the change of V_{max} was definitely bipolar (e.g., fig. B.11) and slightly visible even at

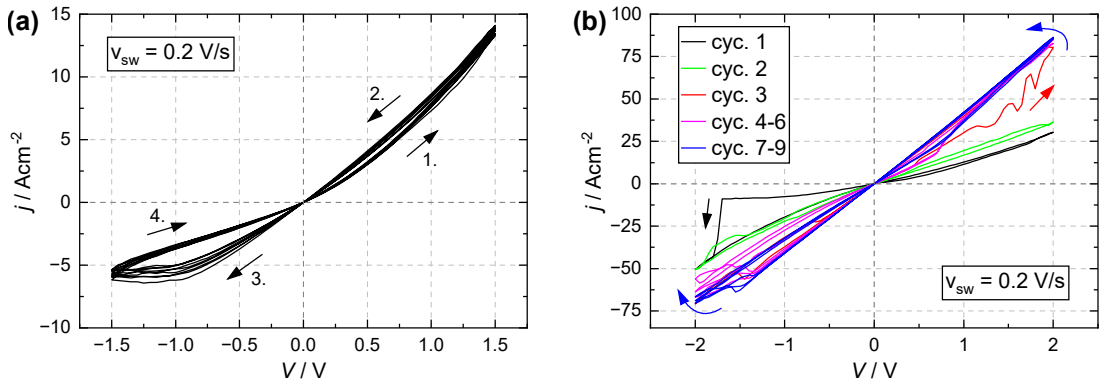


Figure 5.20: Figure-8 bipolar RS of a Pt|a-$GaO_{1.5}$|a-GaO_x|ITO device with $d \approx 15$ nm for the a-$GaO_{1.5}$ layer and $d \approx 55$ nm for the a-GaO_x layer ($x \approx 1.0$, PLD with 0.79 J/cm^2 in Ar, lithographic Pt-TE: 100 × 100 µm^2, CT4). (a) *I*-*V* sweeps with $V_{max} = \pm 1.5$ V. (b) *I*-*V* sweeps after increasing V_{max} to ±2.0 V (cycle 1 is the first at ±2.0 V). The colored arrows indicate the sweep direction in the respective curve regions.

v_{sw} = 450 V/s. The final switching curves in fig. 5.20b are very similar to those in fig. 5.19a, which indicates that the abrupt transition at −1.7 V virtually transformed the Pt|a-$GaO_{1.5}$|a-GaO_x|ITO device into a Pt|a-GaO_x|ITO device. The most likely explanation for this is the growth of non-stoichiometric filaments through the a-$GaO_{1.5}$ layer at negative voltage (oxygen migration from a-$GaO_{1.5}$ to a-GaO_x). Once they reached the Pt-TE, the a-$GaO_{1.5}$ layer abruptly and irreversibly became irrelevant. The observed switching behavior is rather similar before and after the abrupt transition, except for the differing curve shape. There are no signs of filamentary switching. This could mean that the a-$GaO_{1.5}$ layer mainly acted as a non-linear series resistance, without a direct involvement in the figure-8 RS mechanism.

5.2.3 Abnormal Bipolar Switching

Besides counter-8 and figure-8 bipolar switching, also abnormal bipolar RS behavior was observed in several cases. Fig. 5.21 shows an overview of the respective *I*-*V* curves of four devices on different samples. The switching processes occurred either in clockwise direction – with LRS → HRS transitions at positive and negative voltage (fig. 5.21a–c) – or in counter-clockwise direction – with HRS → LRS transitions at both polarities (fig. 5.21d). The switching behavior is not unipolar, though, since *I*-*V* sweeps with only positive or only negative voltage just yielded a transition in the first cycle (e.g., fig. 5.22). Similar to figure-8 and counter-8 RS, the reversal of the voltage polarity is mandatory for switching back and forth. Here, this observation is even more important, since it proves that there is no switching process at or near zero voltage. Otherwise, each of the repeated cycles with the same voltage polarity would exhibit switching. But where are the transitions opposite to the obvious switching processes near V_{max}? Fig. 5.21 already contains some hints concerning this question. The devices of fig. 5.21c,d both exhibited counter-8 bipolar switching at lower V_{max} (cf. fig. 5.7c,d), which changed into abnormal bipolar RS after increasing the voltage amplitude. Furthermore, the

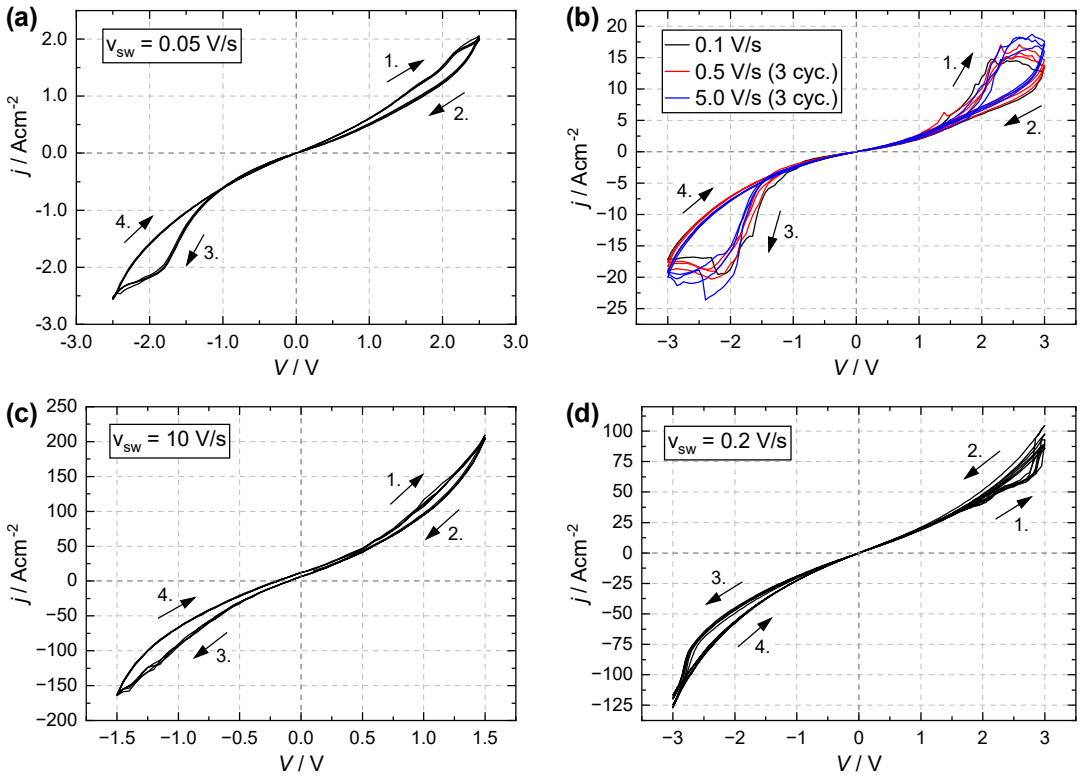


Figure 5.21: Abnormal bipolar resistive switching of Pt|a-GaO_x|ITO devices on four different samples (contacted by either W or Au probe tips): (a) $x \approx 0.9$, $d \approx 60$ nm (same sample as fig. 5.7a; PLD with 0.80 J/cm^2 in Ar, chamber #1), with sp-Pt-TE (TE-∅: 400 µm, CT1, $V_{max} = \pm 2.5$ V). (b) $x \approx 0.9$, $d \approx 90$ nm (PLD with 0.86 J/cm^2 in Ar, chamber #1), with sp-Pt-TE (TE-∅: 200 µm, CT1, $V_{max} = \pm 3.0$ V). (c) $x \approx 0.9$, $d \approx 100$ nm (PLD with 0.94 J/cm^2 in Ar, chamber #1), with c-ITO-BE (same device as fig. 5.7c; TE-∅: 100 µm, CT2, $V_{max} = \pm 1.5$ V). (d) $x \approx 1.0$, $d \approx 120$ nm (PLD with 0.79 J/cm^2 in Ar), with lithographic Pt-TE (same device as fig. 5.7d; TE: 20 × 20 µm^2, CT4, $V_{max} = \pm 3.0$ V). Each graph contains several I-V sweeps (4–7 cycles).

abnormal bipolar RS curves frequently have crossing points, e.g., approx. at −1 V in fig. 5.21a and at ±1.5 V in fig. 5.21b. The respective outer regions of those I-V sweeps look like separate loops.

Such outer loops are a direct result of high V_{max}, as shown in fig. 5.23. They may form at both voltage polarities independently. In the first example in fig. 5.23a, a crossing point at negative voltage was present for $V_{max} = \pm 3$ V. Decreasing the negative amplitude to −2.5 V prevented the crossing point and led to the usual counter-8 bipolar switching. The device in fig. 5.23b yielded figure-8 bipolar switching at $V_{max} = \pm 2$ V, while increasing V_{max} to ±3 V led to a crossing point at positive voltage. These examples reveal where the reverse transitions in the abnormal bipolar RS occur. The non-linear increase of the current before and at the crossing points is already an HRS → LRS transition and not just the usual S-shape of the I-V curves (3rd arrow in fig. 5.23a, 1st arrow in fig. 5.23b). When V_{max} is set rather low, then only this transition is observed. Higher V_{max} leads to an additional LRS → HRS transition for the same voltage polarity, which is visible as outer loop (4th to 5th arrow in fig. 5.23a, 2nd to 3rd arrow in fig. 5.23b). Even though these two transitions have opposite direction, they are not the reverse switching processes of each other. Otherwise, the

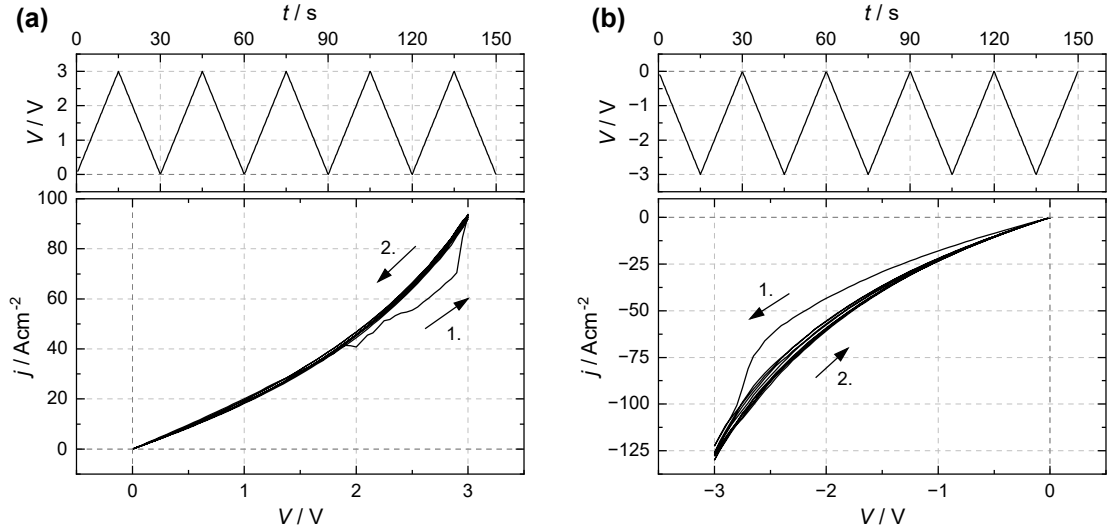


Figure 5.22: Bipolarity check for abnormal RS of a Pt|a-GaO_x|ITO device with $x \approx 1.0$ and $d \approx 120$ nm (same as in fig. 5.21d; PLD with 0.79 J/cm^2 in Ar, lithographic Pt-TE: 20 × 20 µm^2, CT4). The top graphs show the applied voltage $V(t)$, while the bottom graphs display the resulting I-V curve. (a) Triangular I-V sweep with only positive voltage (V_{max} = +3.0 V, v_{sw} = 0.2 V/s). (b) Triangular I-V sweep with only negative voltage (V_{max} = −3.0 V, v_{sw} = 0.2 V/s). In each graph, the numbered arrows refer to the first 0 V → V_{max} → 0 V cycle, which is the only one with a notable switching effect.

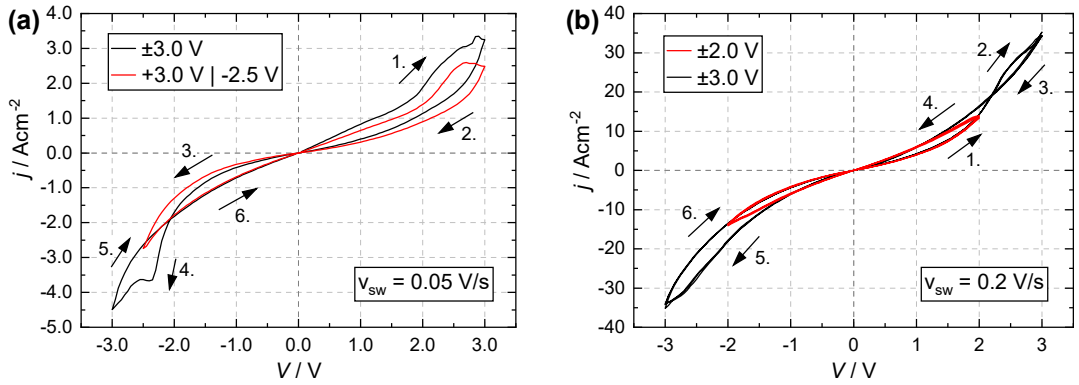


Figure 5.23: Loop formation in the I-V curves of Pt|a-GaO_x|ITO devices at high V_{max}. The displayed I-V sweeps were measured directly after each other, in the order indicated by the legends. (a) In this case, a loop at negative voltages is observed for V_{max} = ±3.0 V, while V_{max} = +3.0 V|−2.5 V led to counter-8 bipolar RS (same device as fig. 5.7a; $x \approx 0.9$, $d \approx 60$ nm, PLD with 0.80 J/cm^2 in Ar, chamber #1; sp-Pt-TE-∅: 400 µm, CT1). (b) Here, figure-8 bipolar RS occurred for V_{max} = ±2.0 V, while V_{max} = ±3.0 V resulted in a loop at positive voltages (3 cycles each; same device as fig. 5.15d; x in the range of 0.98–1.07 according to EELS, $d \approx 60$ nm, PLD with 0.79 J/cm^2 in Ar; lithographic Pt-TEs: 50 × 50 µm^2, CT4).

previously discussed bipolarity checks would have yielded continuous switching (e.g., fig. 5.22). This leads to the following conclusion: Both for positive and for negative voltage two transitions may take place, one at intermediate voltage and one at high voltage. The transitions at intermediate voltage are the reverse processes of those at high voltage with opposite polarity. The best example for this is fig. 5.21b, where all four transitions are clearly visible.[101] This indicates that the devices

[101] Fig. B.13 shows how the device of fig. 5.21b gradually developed from counter-8 to abnormal bipolar switching.

have at least two different regions where switching may take place – a situation that is commonly called "complementary resistive switching".

The I-V curve with stepwise increasing voltage in fig. 5.24 confirmed the two transitions at the same polarity. In contrast to the usual I-V sweeps, this type of measurement allows to discriminate the current variations due to switching from those due to pure I-V non-linearity. At V = +2.0 V and +2.1 V, the current increased over time due to an HRS → LRS transition. In contrast, the current decreased due to an LRS → HRS transition at $V \geq +2.3$ V. This fits to the previous conclusions about two transitions at the same polarity. In-between, at V = +2.2 V, both transitions even occurred simultaneously and the resulting $j(t)$ curve looks like a superposition of the transitions at lower and higher voltage. The voltage range where these transitions took place is comparable to the typical I-V sweeps. It is thus plausible that each of the transitions may occur in RS curves, depending on the actual voltage drop over the Pt|a-GaO_x|ITO devices and the local electric field within them.

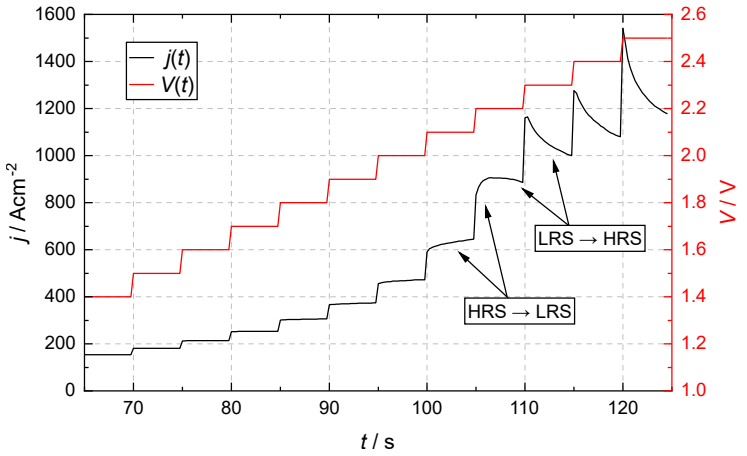


Figure 5.24: I-V measurement with stepwise increasing voltage from 1.4 V to 2.5 V, on a Pt|a-GaO_x|ITO device with x in the range of 0.98–1.07 (EELS) and $d \approx 60$ nm (same sample as fig. 5.15c,d; PLD with 0.79 J/cm^2 in Ar; lithographic Pt-TEs: 20 × 20 µm^2, CT4).

Depending of the relative extent of the transitions, there may be no crossing points in the I-V sweeps.[102] This can be seen in fig. 5.23a: Both sweeps have slight kinks at approx. +2 V, which correspond to a minor HRS → LRS transition. The subsequent LRS → HRS transition at about +2.8 V is much stronger, though, and decreases the current to such an extent that it does not cross the LRS curve. For the counter-8 switching at lower V_{max}, the minor transition at +2 V is essentially irrelevant since the other transition pair dominates the RS behavior. This indicates that the abnormal bipolar RS is actually the general case of the counter-8 and figure-8 bipolar switching. At low V_{max}, either counter-8 or figure-8 switching is observed, depending on which transition pair dominates, while at high V_{max} both may become visible. Fig. 5.25 illustrates this hypothesis, which is an important step in understanding the diverse switching behavior of a-GaO_x. The relative intensity of the four transitions determines which RS behavior is measured. In general, this depends on where

[102] The extent of the transitions is also influenced by the sweep rate, for example.

the highest electric fields occur within the device and how strong the induced local changes in the material affect the current.

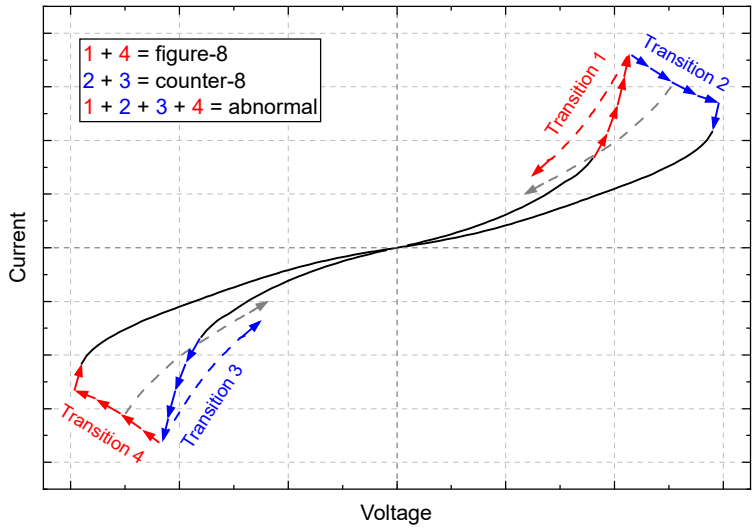


Figure 5.25: Schematic explanation of the connection between figure-8, counter-8 and abnormal bipolar switching. The transitions 1 and 3 take place in a certain region of the device and revert each other on the atomic level, at least approximately (red arrows). Likewise, the transitions 2 and 4 are an interdependent transition pair (blue arrows), but the related ion migration occurs in a different region than for the other pair. Depending on V_{max}, v_{sw} and the relative resistances of the two responsible regions within the device, only one pair might be visible (as indicated by the colored dashed lines). In this case, figure-8 or counter-8 bipolar switching is observed. When all four transitions take place with sufficient intensity then abnormal bipolar RS is measured. It may be clockwise or counter-clockwise, depending on whether the transitions 1 and 3 or 2 and 4 are more visible. When the second transition for a certain polarity is rather weak (e.g., due to intermediate V_{max}), then an outer loop may be observed (grey dashed lines). Since the transition pairs occur in different regions, their order is probably not fixed – especially when one of them has much higher local resistance (e.g., transition 2 at lower V than transition 1).

5.2.4 Interpretation of the Resistive Switching Behavior

The three previous sections displayed the counter-8, figure-8 and abnormal bipolar resistive switching of Pt|a-GaO_x|ITO devices with O/Ga ratios in the range of 0.8–1.1. It is not without a certain irony that the "abnormal" curves led to a new idea how these switching types could be related. In this section, the implications for the underlying RS mechanisms shall be discussed. The mechanisms will be described on the basis of oxygen ion migration, but they could be formulated analogously with gallium ion transport (cf. sec. 2.3.4).

Several aspects of the measured I-V curves speak against a filamentary switching mechanism. The observed transitions were gradual instead of abrupt and produced just rather low HRS/LRS resistance ratios. Furthermore, no electroforming was required to induce resistive switching and the measured currents were not independent of the TE size. The devices generally were well conductive,

which is another argument against filament formation. There is no incentive for a soft breakdown when the electrons already flow easily in the as-prepared state. As discussed in sec. 2.4.1, this might be different for less oxygen-deficient a-GaO_x (e.g., $x \approx 1.5$). For the devices investigated in this work, however, the above characteristics rule out a filamentary mechanism with some certainty. Instead, they would agree with interface- or bulk-type switching.

Interface switching at the Pt|a-GaO_x interface would be a plausible explanation for the counter-8 bipolar RS. Due to the high work function of Pt, a Schottky barrier should be present in this region. As discussed in sec. 2.4.2, it is not entirely clear how the barrier's resistance depends on the local O/Ga ratio. However, the more probable scenario is that high x leads to an increase of the barrier height φ_B and the barrier width x_d, which would increase the resistance both with regard to thermionic emission and tunneling. This would fit to the observed trend that TE interface oxidation leads to counter-8 switching. The increased O/Ga ratio beneath the TE would cause a rather high Schottky barrier resistance in the as-prepared state. The TE interface thus would dominate the device resistance and the highest voltage drop would occur in this region. Initially, a positive voltage would not result in a transition, since oxygen is already accumulated below the TE. At negative voltage, where the counter-8 switching started, oxygen would migrate away from the interface and the current through this Schottky barrier would increase (HRS → LRS). Thereafter, applying a positive voltage would attract oxygen back to the TE, which would restore the originally high resistance (LRS → HRS).

Another indication that the counter-8 switching takes place in the oxidized region near the Pt-TE is its inherently slower switching speed in comparison to the figure-8 behavior. Since the oxygen ion conductivity likely decreases with increasing x (cf. sec. 2.3.4), it would be plausible that the oxygen migration is slower in the interface oxidation layer than in the rest of the device.

The non-linearity of most measured I-V curves also would agree with a Schottky barrier as the dominating resistance, because thermionic emission and tunneling have an exponential voltage dependence (cf. eqs. 2.61 to 2.66). Moreover, this would explain why increasing V_{max} led to abnormal bipolar switching. Once the current through the dominating Schottky barrier becomes very high, i.e., its resistance sufficiently low, there could be a significant voltage drop also in another region of the device.

The second switching region could be the a-GaO_x|ITO interface. Since it is oriented opposite to the Pt|a-GaO_x interface, an analogous interface switching mechanism in this region would yield figure-8 bipolar RS. For a positive voltage applied to the TE, the a-GaO_x|ITO interface is reverse-biased and thus might dominate the device resistance despite the low work function of ITO – especially when there is no TE interface oxidation. This would explain why figure-8 switching occurs when the a-GaO_x surface was not exposed to air. Oxygen migration away from the ITO interface then would lead to an HRS → LRS transition at $V > 0$. It is, however, not plausible that this process is simply reverted at negative voltage. The a-GaO_x|ITO interface would have low resistance both due to the previous transition at $V > 0$ and because it is forward-biased at $V < 0$. Under such circumstances

the reverse-biased Pt|a-GaO_x interface should clearly dominate at $V < 0$, even without interface oxidation. Therefore, there would be a much lower voltage drop at the a-GaO_x|ITO interface and thus no oxygen transport towards ITO. In consequence, it is difficult to explain the figure-8 bipolar switching with interface switching near the BE.

An alternative would be figure-8 switching of the Pt|a-GaO_x interface, which would require the Schottky barrier resistance to decrease for increasing O/Ga ratio. While this could be justified with a lower work function difference (cf. sec. 2.4.2), it would be difficult to reconcile with simultaneous counter-8 switching at the same interface. Since interface oxidation seems to favor counter-8 behavior, this would imply the following complex scenario: At rather low x, the TE barrier resistance would have to decrease for increasing x, while at rather high x a further increase would have to increase the resistance. These trends would have to apply equally to forward and reverse bias. Whether this is realistic is difficult to judge without additional evidence. A strong argument against this model is the frequent congruency between counter-8 and figure-8 I-V curves, which is particularly noticeable in some abnormal bipolar RS curves. In fig. 5.21b, for example, the HRS parts from +3 V to 0 V and from −3 V to 0 V look almost similar. However, if all transitions would occur at the TE interface, then the O/Ga ratio would have to be especially large after high positive voltage and especially low after high negative voltage. In this case, the two HRS parts should differ much more than observed in the measurements.

Besides these inconsistencies with regard to figure-8 switching, there are also some general issues with the interface switching mechanism. The I-V curves are often highly symmetric (e.g., fig. B.8), which is unexpected both for a single Schottky barrier and for two back-to-back Schottky barriers with different metals. The Schottky barrier models in sec. 2.2.4 yield no symmetric I-V curves, even for pure tunneling. Furthermore, the I-V curves were quite linear in several cases (e.g., figs. 5.7c and 5.19a), which is even more incompatible with a Schottky barrier as dominating resistance. The usual explanation would be a pronounced series resistance that governs the shape of the I-V curve. It is unclear, though, how interface switching could produce a substantial hysteresis between HRS and LRS under such circumstances. The voltage drop across the Schottky barrier and thus the driving force for ion transport would be significantly reduced. If the interface region still would be responsible for the different resistance states, then at least the HRS should be consistently less linear than the LRS. But that is not always the case in the measurements (e.g., fig. 5.7c).

The third classical RS mechanism is the bulk switching model, where the conductivity σ within the a-GaO_x layer is modulated by ion transport. The switching transitions essentially correspond to the generation and dissolution of a region with elevated O/Ga ratio, typically near one of the ion-blocking interfaces (cf. [30]). This region with high x and therefore low σ has more impact on the device resistance than the simultaneously formed regions with decreased x because of two reasons: The relation between σ and x is non-linear with $\partial^2\sigma/\partial x^2 < 0$ (cf. fig. 2.3) and the total bulk resistance is essentially an integral over σ^{-1}. In counter-8 bulk switching, this region of high x would emerge near the TE. Positive voltage would attract oxygen ions towards the TE, resulting in

an increase of the O/Ga ratio in this region and an overall higher device resistance (LRS → HRS). When x is already increased in this region due to interface oxidation, then this transition would be absent in the first I-V cycle – as observed in the measurements. At negative voltage, oxygen would migrate away from the TE, leading to a more homogeneous oxygen distribution and a lower bulk resistance (HRS → LRS). In accordance with the experiments, TE interface oxidation would promote this counter-8 bulk switching mechanism because the oxidized region near the TE would be the dominating resistance right from the start.

Another argument for a bulk-dominated switching mechanism is the frequent symmetry of the I-V curves. In contrast to the previously discussed Schottky barrier currents, the bulk electron transport does not depend on the direction.

Figure-8 bulk switching would have to take place in the vicinity of the BE, where $V > 0$ would reduce the O/Ga ratio (HRS → LRS) and where $V < 0$ would cause oxygen enrichment (LRS → HRS). In combination with the above counter-8 mechanism, this would also explain the abnormal bipolar switching: High positive and negative voltages would lead to oxygen accumulation near the TE and the BE, respectively (LRS → HRS). Then, at intermediate voltages the major voltage drop would be in these regions with high x, thus driving oxygen away from the BE at $V > 0$ and away from the TE at $V < 0$ (HRS → LRS). This model would also agree with the symmetry of the abnormal bipolar curves (e.g., fig. 5.21b), since the two regions should behave quite similar.

The figure-8 bulk switching mechanism has an important prerequisite, though. It requires that the O/Ga ratio becomes substantially higher near the BE than in the rest of the bulk. Otherwise, this region could not dominate the device resistance. However, the I-V sweeps typically started with positive voltages, which should lead to oxygen accumulation near the TE. Therefore, even without interface oxidation, the bulk near the TE immediately would become the dominating region. It is unclear whether this could change during subsequent I-V sweeps. If there is no initial TE interface oxidation and if V_{max} is quite high, then at negative voltage the O/Ga ratio near the TE might return to normal and oxygen might actually accumulate near the BE. But why should this lead to continuous figure-8 switching? With the same V_{max}, the next semi-cycle with positive voltage should simply revert the situation and again should cause oxygen accumulation at the TE. It is unlikely that the O/Ga ratio near the BE would remain elevated in this case. A possible solution to this issue would be the transfer of oxygen from the ITO-BE to a-GaO_x, which would cause interface oxidation at the BE either in the as-prepared state or after positive voltage. This hypothesis was tested by SIMS measurements on devices with ^{18}O-enriched ITO, but the ^{18}O profile did not extend significantly into the a-GaO_x layer after preparation and after positive voltage (cf. sec. 5.1.3). EELS measurements of an a-GaO_x|ITO interface also yielded no evidence for such an effect (fig. B.4). Another alternative explanation for oxygen accumulation near the BE could be the influence of a directional series resistance that is less conductive at positive voltage than at negative voltage. This would reduce the voltage drop across the bulk only for positive voltages, thereby favoring the oxygen enrichment near ITO at negative voltage. The only option for this would be a Schottky barrier

at the ITO interface, which is reverse-biased at positive voltage. However, in this case the Schottky barrier at the Pt interface should reduce the negative voltage across the bulk even further. Additionally, the I-V curves should probably look more diode-like if a directional resistance would have such an effect.

Another major inconsistency concerning a bulk switching mechanism is the non-linearity of the I-V curves. If the bulk would dominate the device resistance, then especially the sweeps without superimposed switching should have linear I-V characteristics. Fig. 5.9a with low V_{max} and fig. 5.12b with high v_{sw} clearly demonstrate that this is not the case. A possible explanation for the non-linearity would be a non-switching Schottky barrier as series resistance. However, this would imply that the HRS should always be more linear than the LRS – in contrast to the experimental observations (e.g., fig. 5.7a).

In summary, the qualitative evaluation of the I-V curves indicates that more complex processes take place than in the classical switching mechanisms. The next step was the quantitative analysis of the j-V relations. Comparing the measured curves with theoretical $j(V)$ models is a common tactic for identifying the current-limiting region within RS devices.

For this approach, the fast I-V sweeps of fig. 5.12b are optimal, where $j(V)$ of individual states was recorded without superimposed resistive switching. These sweeps are also not disturbed by external series resistances because the device was contacted by Au wires (CT5). The LRS and HRS curves could not be described with the Schottky barrier models introduced in sec. 2.2.4, which is not surprising since the underlying assumptions hardly apply to amorphous and highly non-stoichiometric compounds. Afterwards, further j-V relations mentioned in the gallium oxide literature were tested. It was discovered that both the LRS and the HRS curve yield linear $\ln(j/V)$ vs. $V^{1/2}$ plots, as shown in fig. 5.26. The same relation has been reported in [98, 100], albeit with questionable interpretations (cf. sec. 2.4).[103] Therefore, the meaning of this observation is not clear from the literature. It should be kept in mind, though, because an interesting connection to variable-range hopping will emerge further on in sec. 5.5.4.

Besides these single-state I-V curves, also some of the RS curves with hysteresis were investigated by curve fitting. Similarly, they did not conform to the classical Schottky barrier models. However, they also did not agree with the above $\ln(j/V) \propto V^{1/2}$ relation. The reason for this could be either a real difference in the I-V characteristics or simply the influence of overlaid switching processes and contact resistances. In two cases, these fitting attempts revealed other I-V relations. Fig. 5.27a visualizes that the LRS and the HRS curves of the figure-8 bipolar RS in fig. 5.15b can be described by $j \propto V^2$. This expression is well-known as Mott-Gurney law for space charge-limited current and it has been reported for a-GaO_x devices in [18, 169, 176]. However, as mentioned in sec. 2.4.1, it is unlikely that space-charge limited current occurs in this well-conductive, amorphous solid. Furthermore, the Mott-Gurney model predicts $j/V \to 0$ for $V \to 0$, which is not the case in fig. 5.27a.

[103] The related expression $\ln(j) \propto V^{1/2}$, which has been reported in [9, 26, 27, 177], is less appropriate for the data in fig. 5.26.

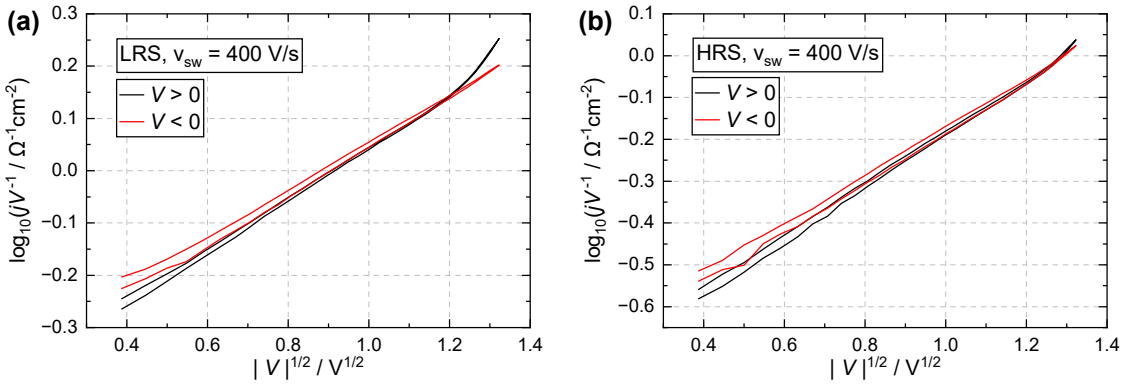


Figure 5.26: Analysis of the fast I-V sweeps from fig. 5.12b, which represent the electron transport in an LRS and an HRS of Pt|a-GaO$_x$|ITO without further counter-8 resistive switching ($x \approx 0.9$, $d \approx 90$ nm, PLD with 0.78 J/cm^2 in Ar, chamber #1; CT5, TE area: 0.465 mm^2, $V_{max} = \pm 1.75$ V, same measurements as in fig. B.8).

The other I-V relationship that was discovered by chance is $\ln(j/V) \propto V^2$, as shown in fig. 5.27b. It describes remarkably well the HRS of the fast figure-8 RS curve in fig. 5.18c. Unfortunately, the interpretation is unclear in this case as well, because no electron transport model with such an I-V relation was found in the literature.[104]

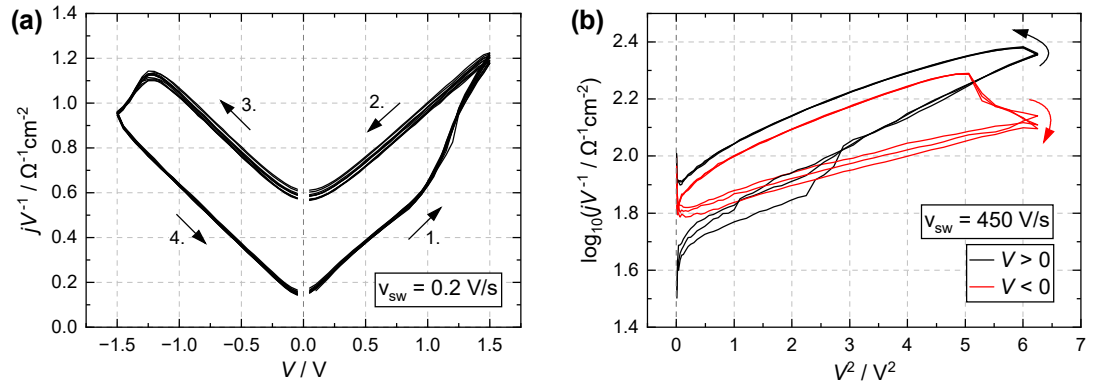


Figure 5.27: Analysis of the figure-8 bipolar I-V sweeps in figs. 5.15b and 5.18c. (a) Pt|a-GaO$_x$|ITO with $x \lesssim 0.8$ and $d \approx 50$ nm (same measurement as in fig. 5.15b; PLD with less than 0.20 J/cm^2 in Ar, chamber #1; eb-Pt-TEs-∅: 200 µm, CT3, $V_{max} = \pm 1.5$ V). (b) Pt|a-GaO$_x$|ITO with x in the range of 0.98–1.07 (EELS) and $d \approx 60$ nm (same measurement as in fig. 5.18c; PLD with 0.79 J/cm^2 in Ar; lithographic Pt-TEs: 20 × 20 µm^2, CT4, $V_{max} = \pm 2.5$ V).

Additional to the fitting attempts with single $j(V)$ expressions, it was also tried to model the RS curves with series circuits of a linear bulk resistance and one or two Schottky barriers. Most of these combined models could not describe the measured curves at all. In the rare cases that reached sufficient convergence, the obtained fit parameters were either not plausible or changed in an unreasonable way between LRS and HRS.

[104] Note that [80] mentions $\ln(j/F) \propto F^2$ as a rough approximation for thermionic field emission through a Schottky barrier, but the electric field F at the interface depends on the voltage approx. via $F \propto (\varphi_{bi} - V)^{1/2}$ (cf. eqs. 2.57 and 2.59), not $F \propto V$.

All in all, neither the qualitative nor the quantitative analysis of the switching curves led to a clear picture of the underlying RS mechanism(s). Evidently, the sole interpretation of the I-V curves is not a reliable approach for identifying the switching regions within the devices. Too much about the electron transport in the devices is unknown at this point. However, the obtained I-V curves and the primary conclusions drawn from them are still valuable for the understanding of a-GaO_x-based resistive switching devices. They serve both as guideline and as benchmark for developing a proper switching model. Essentially, a plausible RS mechanism must be devised from further evidence and then tested whether it could explain the observations in the I-V experiments.

5.3 Experimental Bulk Conductivity

The quite puzzling observations in the resistive switching experiments called for a more thorough investigation of the electron transport in the bulk and at the interfaces of a-GaO_x. As discussed in sec. 2.3, the literature is not clear, both with regard to bulk conduction – MEC vs. VRH – and concerning the proper description of Schottky barriers in the amorphous material. In the following, the results of the van-der-Pauw bulk conductivity measurements will be discussed. The also attempted, experimental analysis of metal|a-GaO_x Schottky barriers by impedance spectroscopy (C-V) unfortunately failed, because the maximum available AC frequency of 1 MHz was too low to observe Cole-Cole semicircles for the well conductive samples. The investigation of metal|a-GaO_x interfaces is therefore limited to the FEM simulations, which will be presented later in sec. 5.6.

5.3.1 Thermal Relaxation

The as-prepared a-GaO_x thin films have a meta-stable structure that is governed by the PLD conditions, which of course includes the deposition temperature of $T \approx 300$ K. Heating of the samples to temperatures above RT thus provides additional thermal energy in comparison to the PLD process, enabling the rearrangement of the atoms into a more stable, yet still amorphous structure. Such a structural relaxation at elevated temperatures is well-known for amorphous materials.[38, p. 506] The consequence of this is a thermal relaxation of the material's properties, which was also observed in the van-der-Pauw measurements of this work. Fig. 5.28a shows an exemplary conductivity measurement, where the temperature was varied between 300 K and 425 K three times. Towards the end of the first heating ramp, there is already a pronounced downwards bend of the $\sigma(T)$ curve. The conductivity decreases irreversibly with each successive T-cycle, but the relaxation effect gradually ceases as the material slowly adapts to $T_{\max}$. The higher the employed temperatures, the stronger is the decrease of the conductivity (cf. fig. 5.32a). A similar thermal relaxation of the conductivity of a-GaO_x has also been reported in [33].

The observed relaxation is due to subtle structural changes, not due to crystallization, because the temperatures employed in all the van-der-Pauw measurements are below the limit where first crys-

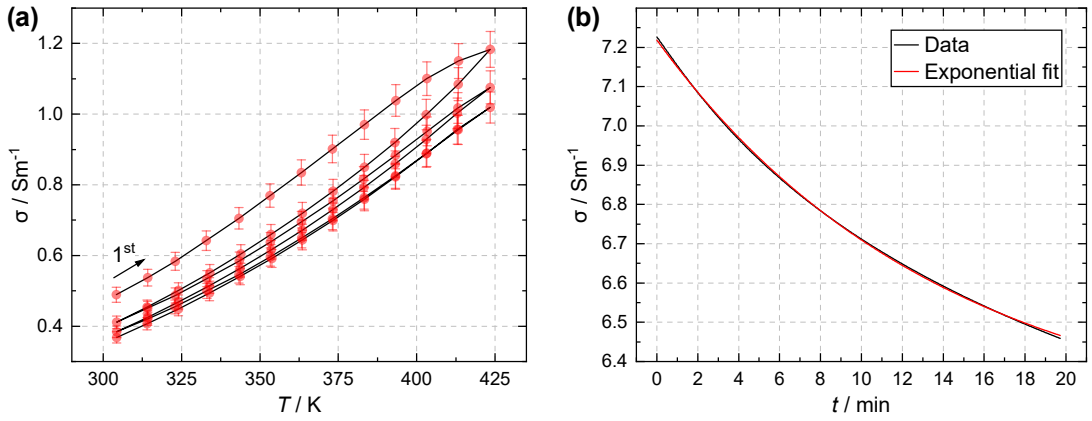


Figure 5.28: Irreversible decay of the bulk conductivity σ at elevated temperatures. (a) Conductivity measurement of a-GaO_x with $x \approx 1.0$ (PLD with 0.49 J/cm^2 in H_2/Ar), where the temperature was varied in 3 cycles with 300 K → 425 K → 300 K. The arrow indicates the first cycle. (b) Continuous conductivity measurement of a-GaO_x with $x \approx 0.8$ (PLD with 0.39 J/cm^2 in Ar) at a constant temperature of 400 K. Due to the prior heating from RT to 400 K (approx. 50 min), $t = 0$ does not refer to the absolute beginning of thermal relaxation. The conductivity approximately follows an exponential decay, as indicated by the fitted curve $\sigma = 0.99\,\mathrm{S\,m^{-1}} \exp[-t/(13.9\,\mathrm{min})] + 6.23\,\mathrm{S\,m^{-1}}$.

tallites have been detected in comparable samples by TEM (cf. sec. 2.3). It is also important to note that the conductivity measurements were carried out in inert atmosphere. Therefore, surface oxidation, which in principle reduces the effective film thickness by rendering a surface layer less conductive, is not involved in the irreversible decrease of σ. In fact, surface oxidation is not relevant for the measured conductivities altogether, even though the samples were stored in air before the measurements. The van-der-Pauw samples are about 10–100 times thicker than the 10–15 nm wide oxidation layer that was detected by EELS after much longer storage in air at RT (cf. fig. 5.6).[105] The systematic shift of the conductivities by the effective thickness reduction is thus negligible.

The kinetics of the thermal relaxation were investigated by keeping a sample at a constant temperature of 400 K while measuring the conductivity. Fig. 5.28b shows that the conductivity decreased exponentially with time, resulting in an extrapolated total decay by 14% between $t = 0$ and $t \to \infty$. In assessing this, it must be taken into account that $t = 0$ marks the time when the sample reached 400 K. The curve thus excludes any thermal relaxation during the prior heating ramp. The actual difference between as-prepared and fully relaxed state could be larger and of course depends on the temperature. In any case, the exponential time-dependence indicates that the a-GaO_x thin films are sensible even to short heating, which may impair the quantitative reproducibility of experiments. This could even be relevant for the resistive switching experiments, even though they were carried out at room temperature. The current through the RS devices could cause elevated temperatures within the material by Joule heating. Thermal imaging with an infrared light camera, which was attempted to address this issue, did not show significant heating during RS operation, but the area immediately below the probe tip was blocked from the view. If the voltage drop and the current

[105] Furthermore, for all van-der-Pauw samples the expected surface oxidation is less than the thickness standard deviation from profilometry (which enters into the error bars of σ).

would be confined to a small region of the device, there could be strong local heating. Since the thermal conductivity of amorphous materials is often several orders of magnitude lower than for crystals,[31, p. 307] and since β-Ga_2O_3 already has rather weak heat transport to begin with,[11] it could be that the thermal energy dissipates rather slowly. The resulting thermal relaxation could be involved in the often observed decrease of the current with each successive I-V cycle in the RS experiments (e.g., fig. B.6b), in addition to the asymmetric switching effect already discussed in sec. 5.2. Stable resistive switching with congruent I-V curves then could be a sign of rather complete thermal relaxation. An argument against this interpretation is the generally high electronic conductivity of the a-GaO_x thin films, which could improve the thermal conductivity and reduce the heat accumulation in the material. For further clarification, the question of local heating during resistive switching of a-GaO_x should certainly be addressed in future studies.

5.3.2 Variable-Range Hopping

The T-dependence of the bulk conductivity is the central criterion for discerning mobility edge conduction (MEC) and variable-range hopping (VRH) in amorphous materials. As explained in detail in sec. 2.2.2 and 2.2.3, for MEC a linear relation between $\ln(\sigma)$ and T^{-1} is expected, while for VRH $\ln(\sigma)$ depends linearly on $T^{-1/4}$ instead. It was demonstrated in fig. 2.3 on p. 40 that the a-GaO_x conductivity data in the literature is too coarse to identify the electron transport mechanism, yet it has always been assumed that MEC prevails.

Fig. 5.29 shows side by side the VRH plot according to eq. 2.54 and the MEC plot according to eq. 2.33 for a van-der-Pauw measurement of an exemplary a-GaO_x thin film. With the naked eye, it is difficult to decide which fit curve looks better, but the fit residuals give a clear answer. The VRH residuals have almost no systematic error while the MEC residuals have significant curvature, both for the first T-ramp – in the part before the visible onset of thermal relaxation – and for the last T-ramp where thermal relaxation is rather complete. The conductivity of the sample definitely exhibits the characteristic temperature-dependence of variable-range hopping from 300 K at least up to 425 K, which disproves the common assumption of dominating MEC around room temperature.

The same was observed for all measured samples (cf. fig. 5.32). For each of the a-GaO_x thin films, the VRH fit was superior to the MEC model – also for those deposited in the presence of H_2. Fig. 5.30 shows a representative conductivity measurement for a-GaO_x deposited in H_2/Ar atmosphere. Again, the residuals of the VRH fits are flat, while those for the MEC fits exhibit systematic curvature. Evidently, the deposition in H_2/Ar instead of pure Ar does not change the dominating electron transport mechanism. This could mean that n-type doping by hydrogen incorporation, which has been proposed for β-Ga_2O_3 [14, 112, 146] and which would favor MEC by shifting the Fermi level closer to the CB mobility edge, does not take place during PLD in H_2/Ar. How the PLD conditions affect the conductivity values will be discussed in the next section 5.3.3.

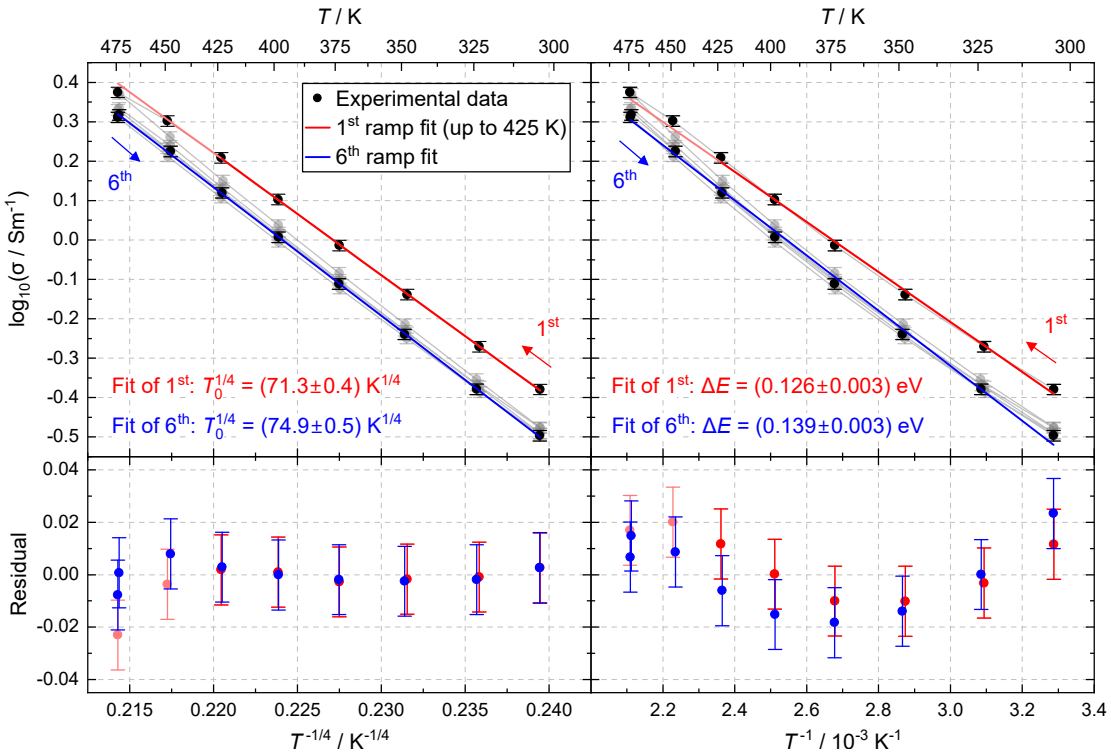


Figure 5.29: Conductivity measurement of a-GaO_x with $x \approx 1.0$ (PLD with 0.77 J/cm^2 in Ar) at elevated temperatures (3x: 300 K → 475 K → 300 K). The van-der-Pauw results are shown as VRH plot on the left and as MEC plot on the right side. Fitted are the first T-ramp (heating; fit up to 425 K) and the last T-ramp (cooling; fit over full range) by linear functions according to eqs. 2.54, 2.33. The intermediate T-ramps are shown in gray. The semi-transparent red lines and data points refer to the $T > 425$ K region that is excluded from the first fit due to interfering thermal relaxation. The $T_0^{1/4}$ and ΔE values from the fits are displayed in the respective graphs. The fit residuals, which are essential for assessing the quality of the fits, are plotted in the two lower graphs.

Since the van-der-Pauw experiments at elevated temperatures are bound to change the amorphous structure during the measurement, they do not refer exactly to as-prepared a-GaO_x. This is even more so for the samples contacted by Pt paste, which required drying at about 425 K for a few minutes during their preparation. The two samples discussed so far (figs. 5.29 and 5.30) belong into this category, since they both have a-ITO electrodes that were contacted by Pt paste. This raises the question whether pristine a-GaO_x exhibits variable-range hopping as well. In fact, it does, as demonstrated by the PPMS measurements that were carried out to tackle the aforementioned issues. In these experiments, the electrodes were contacted by wire-bonding without heating and the conductivity was measured below room temperature at first. Since the initial structure is adjusted to RT, it does not change at the lower temperatures. Fig. 5.31 shows the respective results, which prove that as-prepared a-GaO_x follows Mott's VRH law from 300 K down to 150 K. On this temperature range, the MEC plot not even looks linear anymore, while the VRH plot exhibits a perfectly straight line. The congruence between the experimentally determined conductivities and the predictions from VRH theory is remarkably exact.

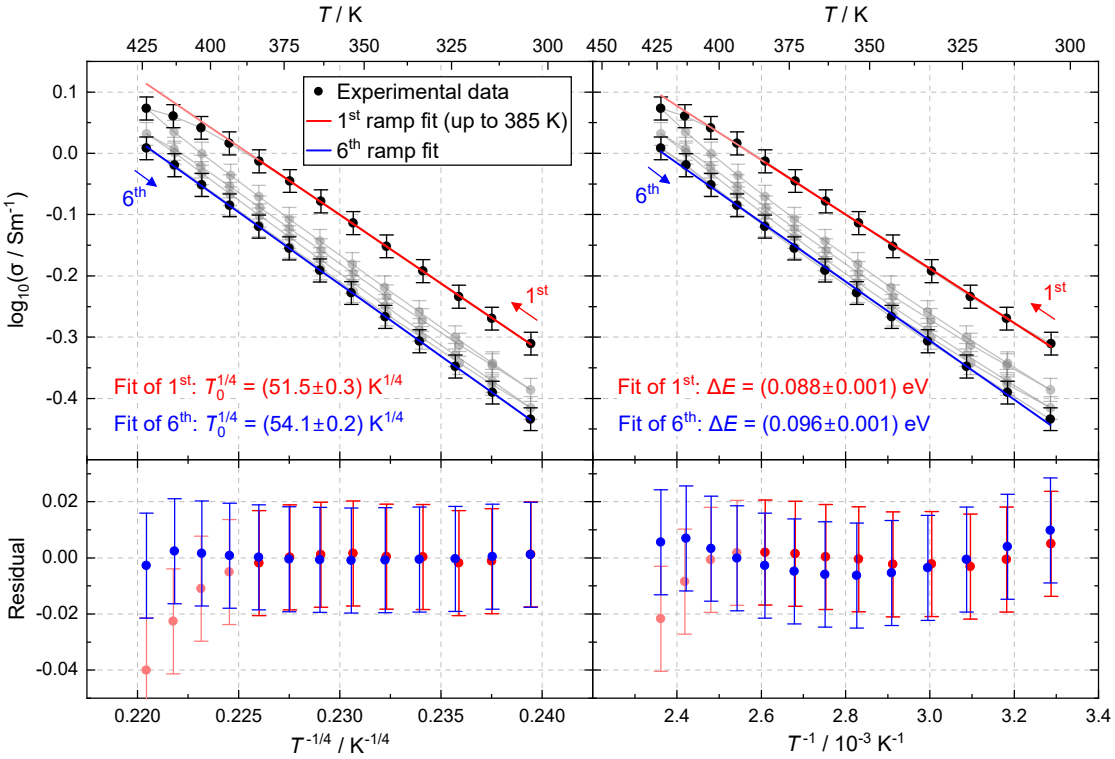


Figure 5.30: Conductivity measurement of a-GaO_x deposited in H_2/Ar gas ($x \approx 1.0$, PLD with 0.49 J/cm^2), with 3 T-cycles according to 300 K → 425 K → 300 K. The van-der-Pauw data is analyzed and plotted in the same fashion as in fig. 5.29, though with the first heating fitted just up to 385 K.

At T > RT, the comparison between the heating and cooling ramps in fig. 5.31 shows that the T-dependence according to Mott's VRH law is only visible after there is no longer a disturbance through simultaneous thermal relaxation. However, the heating ramp also shows an upwards deviation from the VRH fit at high temperatures, before the relaxation decay becomes visible at about 375 K. This was only observed for the two PPMS samples, not for the others, probably because the latter already experienced more structural relaxation before the respective data points were measured. For example, the time between the measurements at 350 K and 400 K is 25 min for the PPMS samples but about 2 h for the others. The underlying reason for the convex deviation in the PPMS measurements in unclear, though. Since VRH normally dominates at low temperatures while MEC typically prevails at high temperatures (cf. sec. 2.2.2 and 2.2.3), this observation probably corresponds to a gradual transition from VRH- to MEC-dominated electron transport. In the MEC plots of figs. 5.29, 5.30 and 5.31, the VRH conductivity increases with temperature roughly with a slope equivalent to $\Delta E \approx 0.1$ eV. In view of the DFT-DOS from literature (sec. 2.3.1), it is well possible that MEC has a steeper slope with $\Delta E > 0.1$ eV. Despite the curvature of the MEC plots, it is thus plausible that at some elevated temperature the MEC conductivity would become higher than the VRH conductivity. If the deviation in the PPMS results would originate from this transition, the question remains why it is no longer visible after thermal relaxation (e.g., 2nd vs. 3rd T-ramp in fig. 5.31). This would imply that structural relaxation has more impact on MEC than on VRH, thus

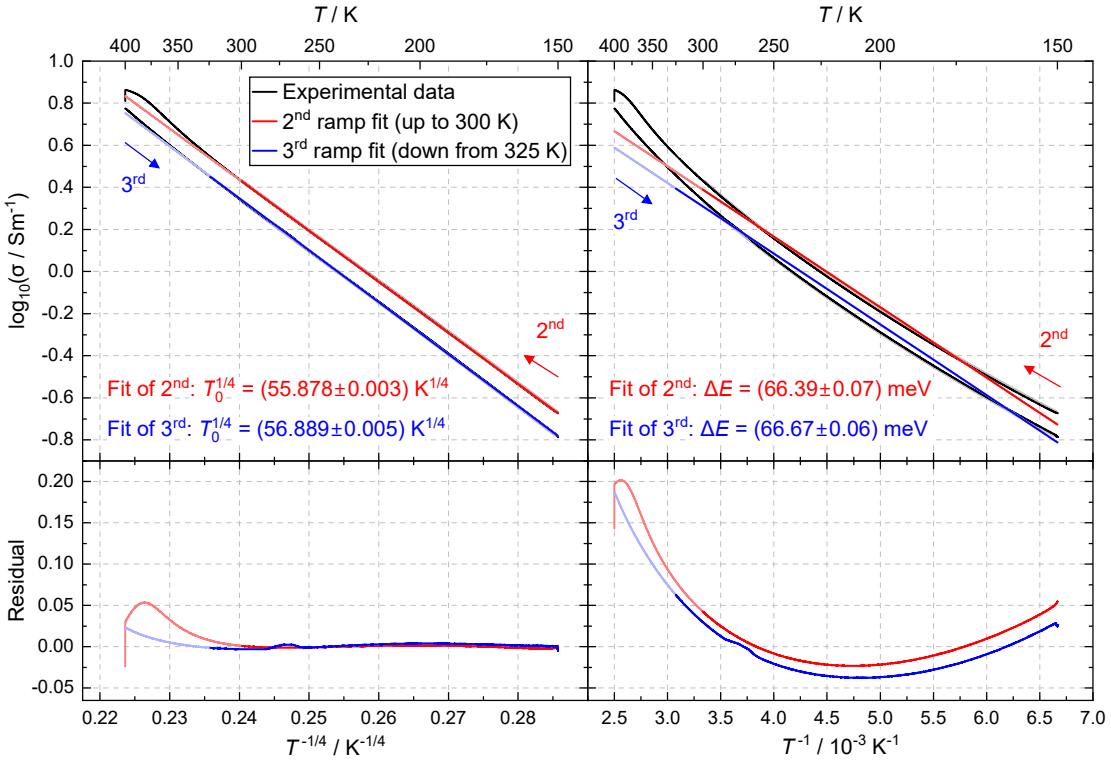


Figure 5.31: PPMS conductivity measurement of a-GaO_x with $x \approx 0.8$ (PLD with 0.39 J/cm^2 in Ar) at low temperatures (300 K → 150 K → 400 K → 150 K → 300 K). The van-der-Pauw data is displayed like in fig. 5.29, yet here with fits of the second T-ramp (heating; fit up to 300 K) and of the third T-ramp (cooling; fit down from 325 K). The first and fourth T-ramps are plotted in gray, but they are not visible because they are virtually identical to the two analyzed ramps. The error bars with $\Delta \log_{10}(\sigma/(\mathrm{S\,m^{-1}})) = \pm 0.04$ are not shown for clarity. At each turning point of the thermal cycling, the sample was held at constant temperature for about 1 h for thermal relaxation and Hall measurements (cf. figs. 5.28b and 5.33).

shifting the transition to higher temperatures. So far, this interpretation of a small deviation in the VRH plots is utterly speculative, yet it is certainly worth further investigation in the future.

The VRH fits in figs. 5.29, 5.30 and 5.31 also give more fine-grained information how variable-range hopping is affected by the structural relaxation at elevated temperatures. Comparison of the initial and final T-ramps shows that thermal relaxation causes an increase of $T_0^{1/4}$. Since also σ_0 from the fits tends to increase (not shown), the overall decrease of σ must be due to the change of T_0. Its increase, in turn, becomes plausible when considering Mott's estimate $T_0 \propto 1/g(E_\mathrm{F})$ from eq. 2.54. The relaxation into a more stable structure probably leads to a "healing" of electronic defects and thus to a reduction of the density $g(E_\mathrm{F})$ of localized states at (or rather around) the Fermi level. This would explain the observed changes of T_0 and σ during thermal relaxation.

The observation of VRH up to and around room temperature in as-prepared and annealed a-GaO_x is, of course, highly relevant for a-GaO_x-based resistive switching devices, which are typically operated in this temperature range. The plethora of switching mechanisms postulated for a-GaO_x may be related to the fact that variable-range hopping was not taken into account. Naturally, VRH concerns the proposed bulk switching mechanism, which relies on changes of the bulk conductivity,

but it would also be relevant for the nature of filaments as hopping percolation paths. The classical Schottky barrier models employed in interface switching mechanisms would become invalid when electron hopping through the barrier region takes place, as discussed in sec. 2.2.4. In short, the consequences for the RS mechanisms are potentially severe. They will be discussed in the FEM modeling section 5.6, where the first steps towards a new VRH-based switching mechanism will be presented. Before, however, in the sections to come, much is to be learned for example about the trends of the VRH conductivity with regard to the stoichiometry x, the chemical potential μ of the electrons or the electric field strength F.

5.3.3 Influence of the Stoichiometry

The main characteristic of a-GaO_x thin films is their O/Ga ratio x, which, of course, has a major influence on the conductivity (cf. sec. 2.3.2). The two central questions are how the VRH transport depends on the stoichiometry x and whether VRH dominates for all $x < 1.5$. In order to investigate these questions, a wide range of different a-GaO_x thin films were measured with the van-der-Pauw method. The PLD laser fluence was varied between the samples, since the EPMA and EELS results in sec. 5.1.2 demonstrated that the O/Ga ratio increases with higher fluence. The composition of the deposition atmosphere was varied as well, using either pure Ar or H_2/Ar gas. Since the stoichiometry measurements in sec. 5.1.2 yielded – rather unexpectedly – only a weak influence of additional H_2, it was tested whether the same is observed for the conductivity.

Fig. 5.32 shows an overview over the van-der-Pauw results of the different a-GaO_x thin films. For all measured samples, the VRH law (eq. 2.54) fits well and is superior to the MEC model (eq. 2.33). Both the thin films deposited in Ar (fig. 5.32a) and those deposited in H_2/Ar (fig. 5.32b) exhibit an approximate trend of increasing VRH conductivity for decreasing PLD fluence. In conjunction with the aforementioned relation between fluence and stoichiometry, this reproduces the nonlinear trend of higher σ for lower O/Ga ratio x that was already indicated by the previous works discussed in sec. 2.3.2.[32, 96, 97] Furthermore, fig. 5.32 shows that the T-dependence of σ becomes weaker, i.e., the slopes get shallower, for lower fluence and thus for lower x. This is also in accordance with the collected literature results in fig. 2.3. What is new, however, is that these trends are driven by a change of VRH transport, and not by the decrease of $\Delta E = E_C - E_F$ for lower x that has been proposed for MEC in the past. The observed reduction of the slope $T_0^{1/4}$ for decreasing O/Ga ratio becomes understandable when combining the approximate relation $T_0 \propto 1/g(E_F)$ from eq. 2.54 with the trends seen in the DFT-DOS of sec. 2.3.1. The lower x, the higher is the density of the – presumably localized – defect states close to E_F in the former band gap. This is synonymous to an increase of $g(E_F)$ in Mott's VRH model, which leads to lower T_0 according to the relation mentioned above. The same effect is also responsible for the overall trend of higher σ for lower x. Since σ_0 of the VRH fits in fig. 5.32 decreases for lower x (not shown), the simultaneous increase of σ must originate from the reduction of T_0 and thus from the growing number of localized states. The

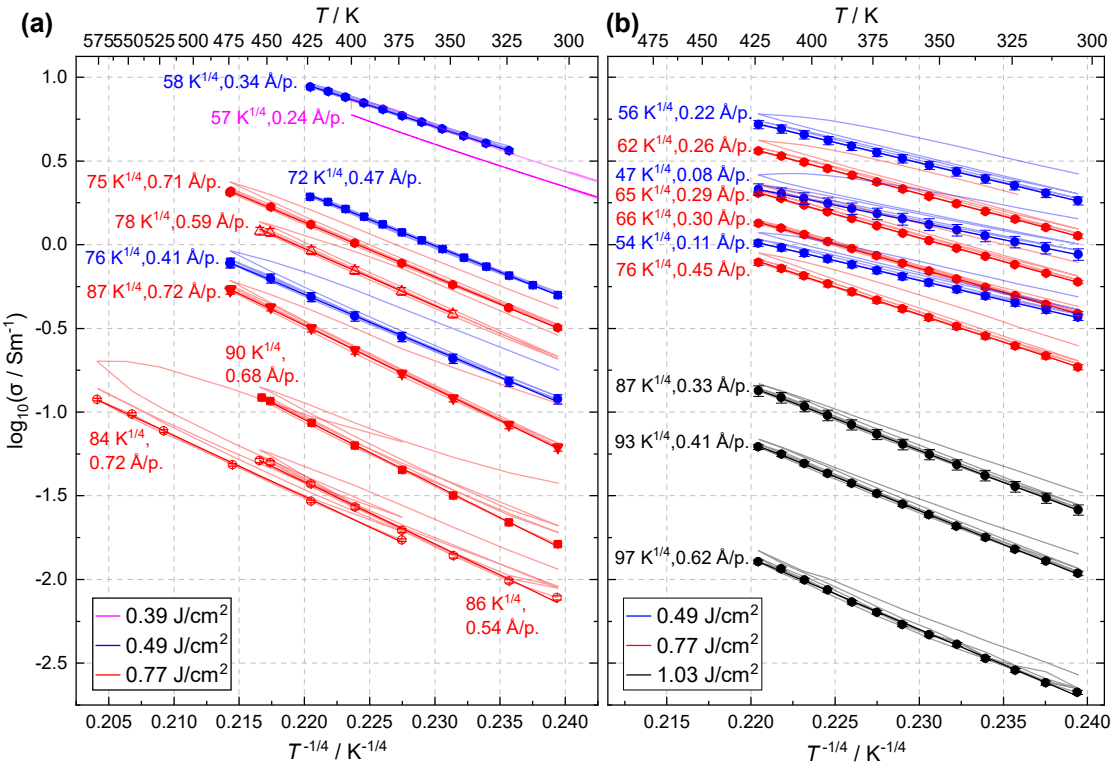


Figure 5.32: Comparison of all conductivity measurements for a-GaO_x deposited with different PLD laser fluence (in J/cm^2), either in (a) pure Ar or (b) H_2(5%)/Ar atmosphere. The opaque symbols and linear fits refer to the last cooling ramp and the resulting $T_0^{1/4}$ values (in K$^{1/4}$) are displayed nearby, together with each sample's growth rate (in Å/pulse). The other T-ramps are shown as semi-transparent lines. The symbol shapes identify the used electrode material: a-ITO = circle, c-ITO = upward triangle, Pt = box, none = downward triangle. The symbol type indicates the metallic paste that was used for attaching the Au wires: filled = Pt, empty = Ag. The opaque line without symbols refers to the continuous PPMS measurement, which had Pt electrodes contacted by wire bonding.

more concentrated the localized states, the closer together they are and the higher are the hopping probabilities (cf. eq. 2.46).

A further trend visible in fig. 5.32 is that thin films deposited with lower PLD fluence typically exhibit stronger thermal relaxation. This might be related to the lower a-GaO_x density that has been reported for lower PLD fluence.[123] The less efficient arrangement of the atoms in the thin films deposited at lower fluence probably leaves more room for improvement by structural changes.

Comparison between the samples deposited in Ar (fig. 5.32a) and those deposited in H_2/Ar atmosphere (fig. 5.32b) with equal PLD fluence yields an approximate trend of higher conductivity and lower T_0 for those prepared in H_2/Ar gas. The effects are not huge, though, which is in accordance with the weak influence of H_2 on the O/Ga ratio that was discussed in sec. 5.1.2. Since more H_2/Ar samples were measured by van-der-Pauw than by EPMA, the here observed trends might be taken as more serious evidence for a lowering of the O/Ga ratio by the additional H_2. One cannot be certain, however, that the conductivity is not influenced by hydrogen incorporation, instead. Further experimental evidence is also inconclusive on this point (cf. [40]) and more research on the role of

hydrogen in a-GaO_x is required. Note that the reduction of the growth rates by using H_2/Ar gas, which was already discussed in sec. 5.1.2, is clearly visible in fig. 5.32 as well.

The van-der-Pauw results in fig. 5.32 have substantial scatter for formally equal PLD conditions, which is no surprise considering the variations seen in the EPMA measurements (cf. sec. 5.1.2). Since both methods exhibit such deviations, it is unlikely that they originate from measurement errors. Differences in thermal relaxation, e.g., due to varying Pt paste drying durations, also cannot be the cause for the scatter in fig. 5.32, because then the last T-ramps should still be comparable. The same goes for enhanced surface oxidation by Pt paste drying, because then those samples contacted by Ag paste should have higher conductivity than those with Pt paste. Instead, the as-deposited material actually seems to vary, even though the same PLD parameters were used. As already discussed for EPMA in sec. 5.1.2, there a several ways how the internal deposition conditions could fluctuate (e.g., beamline loss, laser output, temperature, chamber pressure and purity, target surface), but none of them is particularly probable. The van-der-Pauw results give a bit more insight on this, because the growth rates in fig. 5.32 can be expected to change systematically with the internal PLD conditions as well. Overall, the growth rates increase with higher PLD fluence, as expected and observed before.[123] For equal fluence setting, however, higher growth rate is also correlated with lower conductivity. This is a clear indication that varying growth conditions are responsible for at least most of the scatter. Furthermore, fluctuations of the actual laser fluence can be eliminated from the list of possible reasons, because σ does not decrease steadily with increasing growth rate across all samples in fig. 5.32. The correlation between the conductivity and the growth rate is especially visible for the samples deposited in H_2/Ar (fig. 5.32b), while those deposited in pure Ar exhibit additional and overall stronger scatter (fig. 5.32a). While this could hint at an improved reproducibility of PLD in H_2/Ar atmosphere, e.g., due to the reaction between H_2 and impurities in the gas phase, it also might be related to the different contacting of the samples. Those deposited in H_2/Ar all have triangular a-ITO electrodes contacted with Pt paste, while the samples deposited in pure Ar have different circular electrodes in combination with either Pt or Ag paste (as indicated by the different symbols in fig. 5.32). Since the van-der-Pauw method prevents the influence of contact resistances and since the electrodes and pastes were checked to be highly conductive, it is unlikely that the variation of the electrode materials or the pastes had an influence. Instead, the circular electrode shape might be responsible for the additional scatter of the samples deposited in Ar. The circular masks allow more variation of the contact area than the new, laser-cut triangular masks that were used for the samples prepared in H_2/Ar (cf. fig. 3.3).

In total, the measured thin films in fig. 5.32 cover an approximate stoichiometry range of 0.8–1.0 (cf. fig. 5.3) and a conductivity variation over 3 orders of magnitude within the examined fluence range. In this region, variable-range hopping definitely dominates. Due to the fluence limit of the laser, it was not possible to extend the investigation to significantly less oxygen-deficient a-GaO_x with $1.2 \leq x < 1.5$. For this upper stoichiometry range, the chosen approach of fluence variation is probably not well suited. It is also unlikely that the trend of increasing O/Ga ratio for higher PLD fluence continues up to $x \approx 1.5$. The systematic variation of other PLD parameters, e.g., of the

chamber pressure,[106] is entrusted to future studies. As discussed already in sec. 2.3, there is some evidence that a transition to MEC should occur for $x \to 1.5$. According to the DFT-DOS in sec. 2.3.1, the localized defect states in the former band gap vanish with increasing oxygen content. VRH is effectively suppressed when they become too dilute, such that MEC inevitably prevails even though $\Delta E = E_C - E_F$ likely increases for higher x as well. This is even more probable when MEC is enhanced by the unintentional n-type impurity doping that is typical for gallium oxide (cf. sec. 2.3.2). The conductivity measurements reported in [123] likely correspond to this case. Therein, almost stoichiometric, transparent a-GaO_x, deposited by PLD at RT in O_2 atmosphere (0.05 mbar), has been investigated. Not only are the reported conductivities of up to about 10^{-2} S/m at RT unusually high, but σ also varies in the opposite way – it decreases for lower PLD fluence. This contrary trend has been attributed to an increasing number of localized states that act as electron traps and hinder MEC band transport.[123]

All in all, it is thus quite clear that the dominating electron transport mechanism in a-GaO_x depends on the stoichiometry x – VRH for x of about 0.8–1.0, MEC for $x \approx 1.5$ and a transition somewhere in-between. In the region of severe oxygen-deficiency, where VRH dominates, the conductivity is higher the lower the O/Ga ratio x. The same is expected for MEC in almost stoichiometric a-GaO_x of high purity, but the opposite might be the case for MEC when impurities are present. Naturally, these trends are important for the resistive switching of a-GaO_x, as will be discussed in sec. 5.6. The possibility of a transition between VRH and MEC by variation of the local O/Ga ratio in the material brings more complexity, yet also a completely new option, into the question of the RS mechanism.

5.3.4 Magnetoresistance and Hall Effect

The PPMS measurements with additionally applied magnetic field, which are shown in fig. 5.33, substantiate even further the finding that VRH dominates in a-GaO_x at room temperature. The magnetoconductivity in fig. 5.33a is extremely small and positive, which means that the conductivity is slightly higher the stronger the magnetic field. This is also called negative magnetoresistance. The relation between σ and the magnetic field strength H is approximately linear, as indicated by the fit with $\Delta\sigma/\sigma_0 = 2.93\times10^{-10}|H|/\mathrm{A\,m^{-1}}$ in fig. 5.33a. The conductivity increases for example by a mere 0.03% at a typical magnetic field strength of $H = 10^6$ A/m.[107] This subtle change is symptomatic of the fact that the electron tunneling in the VRH mechanism is not influenced by the Lorentz force that classically affects moving charges in a magnetic field.[31, 153, 154] Instead, a positive and linear magnetoconductivity – exactly like observed here for a-GaO_x – has been predicted for VRH due to quantum interference with three-site hopping paths.[222]

[106] Recently published results for a-GaO_x deposited by RF magnetron sputtering indicate that varying the deposition pressure might cause changes of the O/Ga ratio on a broader scale.[87]

[107] The magnetic permeability of a-GaO_x is unknown, but the permeability of free space can be used to get a better idea of the magnitude of the employed magnetic fields. In vacuum, a magnetic field strength of 10^6 A/m corresponds to a magnetic flux density of 1.26 T.

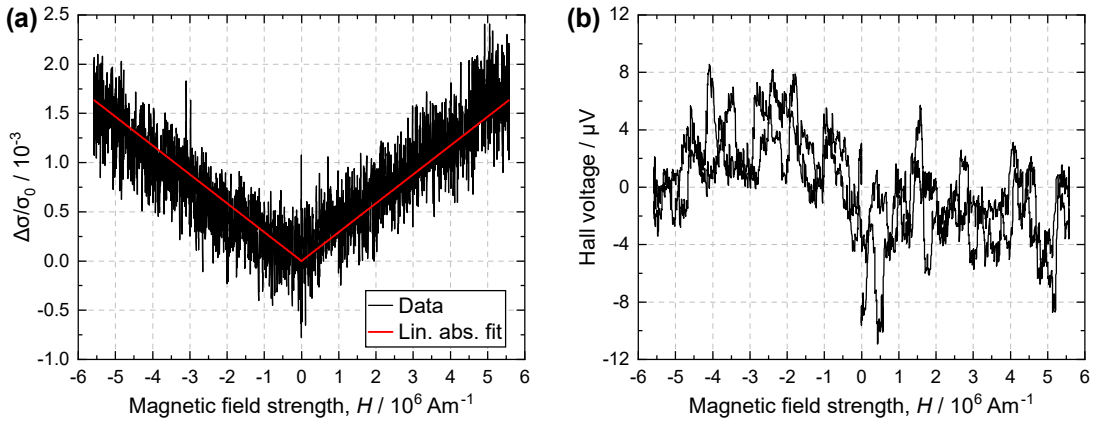


Figure 5.33: Influence of a magnetic field with strength H ($0 \rightarrow 5.6{\times}10^6$ A/m $\rightarrow -5.6{\times}10^6$ A/m $\rightarrow 0$) for a-GaO_x with $x \approx 0.8$ (PLD with 0.39 J/cm^2 in Ar), measured at $T = 300$ K. (a) Magnetoconductive effect: The van-der-Pauw conductivity increases approximately linear with H. The fit with a linear absolute value function yields $\Delta\sigma/\sigma_0 = 2.93{\times}10^{-10}|H|/\mathrm{A\,m^{-1}}$. (b) Hall effect: The Hall voltage V_H, measured for a perpendicular current of 0.5 A and a sample thickness of $(1.09{\pm}0.09)$ µm, is lower than 10 µV and cannot be distinguished from noise.

Fig. 5.33b shows that a-GaO_x exhibits no measurable Hall effect at room temperature. In the explored range of magnetic field strengths, the Hall voltage is below the noise level, with $|V_H| < 10$ µV. This, again, is in accordance with VRH theory and with observations for other materials where VRH dominates.[31, 153] Since the magnetic field influences VRH transport not by the Lorentz force but by rare higher-order hopping transitions, the Hall effect is expected to be very small and its sign is no indicator of any majority charge carrier (cf. sec. 2.3.2).

5.4 Localized Electronic States

The experimental results presented so far clearly suggest that VRH dominates in a-GaO_x even for temperatures around RT. This is a remarkable and quite surprising finding, given that VRH is typically observed only at temperatures well below room temperature, when the amount of electrons excited to delocalized states above the CB mobility edge is greatly reduced. The reason for this is that electrons are usually more mobile in delocalized states because of the extended wave functions and that the normally rather low density of the localized states in the band gap hinders the hopping process. The DFT-DOS from previous studies, which were discussed in sec. 2.3.1 and which were used in the interpretation of the experimental trends so far, already indicate that the situation might be different for a-GaO_x. The HAXPES and DFT results presented in the following confirm this and demonstrate that the severe non-stoichiometry of a-GaO_x indeed creates a large amount of localized states near the Fermi level.

5.4.1 Defect States from Oxygen-Deficiency

Experimental clues about the electronic states close to the Fermi level were obtained from the HAXPES measurements. Fig. 5.34 shows the photoemission from states below E_F for all measured thin films without Au coating.[108] Besides the already discussed Ga 3d and O 2s peaks (cf. sec. 5.1.3), also the valence band (VB) of the samples is visible. It consists mainly of a Ga 4s peak with a broad Ga 4p shoulder towards lower E_{bin}. For $E_{bin} > 5$ eV, the VB looks quite similar for the non-stoichiometric a-GaO_x thin films and the stoichiometric a-$GaO_{1.5}$ reference sample, but for $E_{bin} < 5$ eV there is a remarkable difference. The a-GaO_x spectra exhibit defect states which spread through the former band gap, up to the Fermi level, while a-$GaO_{1.5}$ only shows the band tailing expected in amorphous materials but still has a clear band gap. This is proof that the oxygen-deficiency of a-GaO_x leads to a large number of defects states near the Fermi level. Both measured a-GaO_x samples show the same effect, which rules out a coincidence.

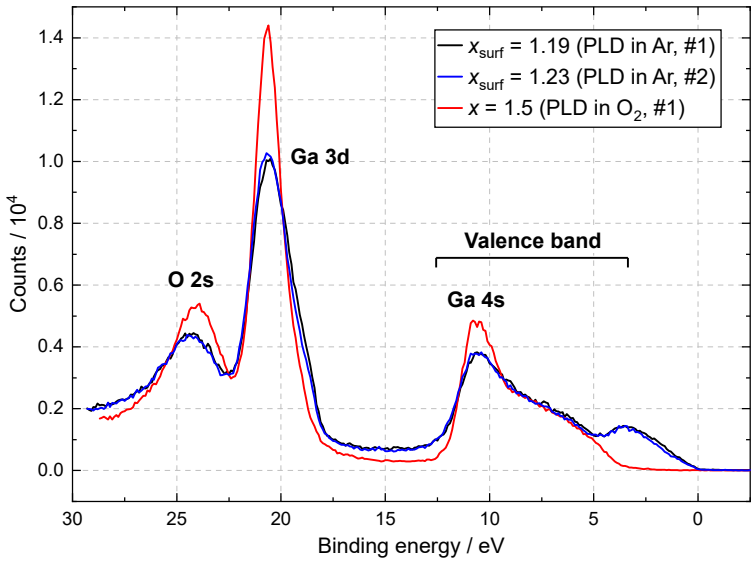


Figure 5.34: HAXPES spectra of the states near the Fermi level, for two non-stoichiometric a-GaO_x thin films with $x_{surf} \approx 1.2$ (PLD with 0.79 J/cm^2 in Ar, $x_{bulk} \approx 0.9$) in comparison to a stoichiometric a-$GaO_{1.5}$ sample (PLD in O_2). The films were deposited either in PLD chamber #1 or #2, as indicated by the legend. The binding energy is measured relative to the Fermi level. The displayed x_{surf} values are those determined from the Ga 3d peak (cf. tbl. 5.1), since they have similar information depth.

The emerging of defect states above the VB edge in HAXPES spectra of a-GaO_x has also been reported in [123] and [104]. In [123], the spectrum of a sample prepared by PLD at RT in O_2 atmosphere with a pressure of 0.01 mbar contains such defect states, while a thin film deposited at 0.06 mbar has none. Essentially the same has been observed in [104], with defect states for PLD at 300 °C in O_2 (0.001 mbar), and none for deposition at 500 °C. However, in both cases the defect states do not reach the Fermi level, most probably because of low oxygen-deficiency.

[108] For the a-GaO_x thin films with Au layer, the valence band region is dominated by the Au states.

Since HAXPES is surface-sensitive, the defect states in fig. 5.34 refer to an O/Ga ratio of $x_{surf} \approx 1.2$. In the bulk of the samples with $x_{bulk} \approx 0.9$, even more defect states can be expected to fill the former band gap. Furthermore, HAXPES does not reveal the complete DOS of a-GaO_x. It only shows a variant weighted with photoionization cross-sections (PICS), which differ between the electronic states. For HAXPES, the PICS for Ga 4s and Ga 4p are 1–2 orders of magnitude higher than for O 2p, while for XPS it is essentially the other way round.[106, 183] The O 2p contributions to the valence band are therefore mostly invisible in HAXPES.[109] Likewise, the signal from the defect states in fig. 5.34 might underestimate their true density.

HAXPES cannot reveal the unoccupied states above the Fermi level, but at least a rough idea can be inferred from the position of E_F in fig. 5.34 (at $E_{bin} = 0$). Due to the symmetric nature of the Fermi-Dirac function, in conjunction with the decrease of the defect density towards E_F, a large band gap above E_F is unlikely. Instead, shortly above the Fermi level, the conduction band, or rather defect states below the CB edge, should commence.

It is remarkable that the HAXPES spectra of the two a-GaO_x thin films in fig. 5.34 are almost identical, even though they were deposited in different PLD chambers. This is in stark contrast to the scatter between different samples that was observed in most other experimental results. Apparently, surface oxidation compensates eventual differences in the bulk and yields quite similar surface properties after storage under equal conditions.

Complementary to these photoemission experiments, the electronic structure of a-GaO_x was investigated through the DFT calculations. Fig. 5.35 shows the resulting DOS for the different, amorphized DFT cells with O/Ga ratios of 1.5, 1.2, 1.0 and 0.8. It is clearly visible that more and more defect states emerge in the former band gap for lower x, just like observed in the HAXPES spectra. They appear predominantly in the region close to and above the VB edge, but also below the CB edge. The band gap is successively filled and already for $x = 1.2$ there is almost no true band gap anymore. This is also in accordance with the HAXPES results for $x_{surf} \approx 1.2$ (fig. 5.34). The trend continues further, with only a minimal band gap for a-$GaO_{1.0}$ and with the whole region between VB and CB covered by defect states in the DOS of a-$GaO_{0.8}$.

Both the total DOS and the site-projected DOS, which are also shown in fig. 5.35, agree quite well with those previously published in [89] (cf. fig. 2.2b,d). The trends and the proportions of the DOS are equal, only the energy scale differs moderately (e.g., $E_F - E_{VBM}$ by about 1 eV or less). This is expected, since hybrid DFT has been used in these previous studies. It is known that the GGA method employed here tends to underestimate the band gap, e.g., for β-Ga_2O_3,[113] which can be compensated by hybrid DFT. For a-$GaO_{1.5}$, this is certainly the case, since its DOS in fig. 5.35 has a much smaller band gap than reported in the literature (cf. sec. 2.3). It is more difficult to judge for the severely oxygen-deficient a-GaO_x, though, because no reliable experimental band gap data is available. Own UV-Vis experiments did not yield a well-defined absorption edge because of the

[109] This also explains why the HAXPES spectra have no steep VB edge, in contrast to the DFT-DOS (e.g., fig. 2.2).

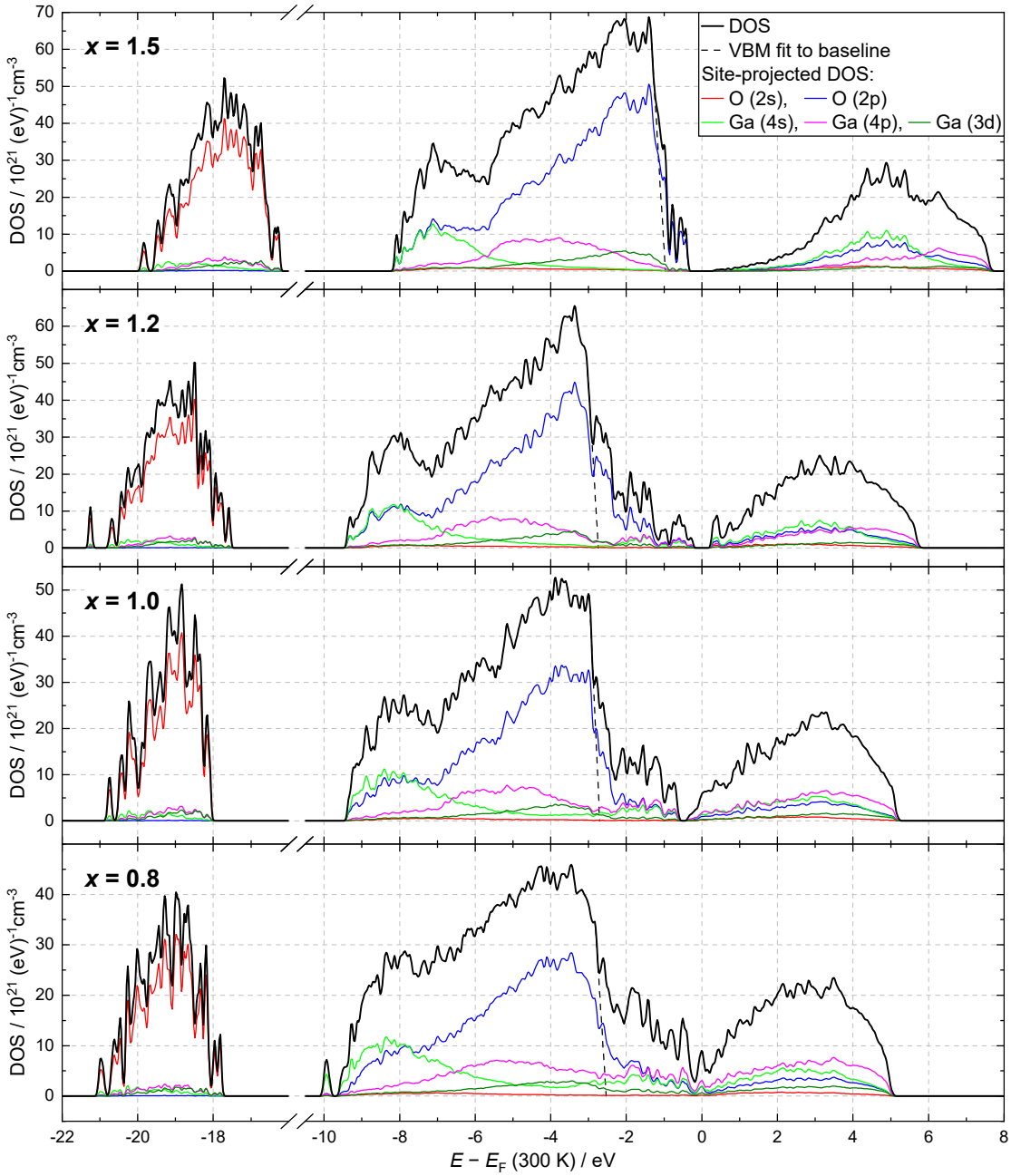


Figure 5.35: DFT-DOS of a-$Ga_{64}O_{96}$ (x = 1.5), a-$Ga_{64}O_{78}$ (x = 1.2), a-$Ga_{64}O_{64}$ (x = 1.0) and a-$Ga_{72}O_{58}$ (x = 0.8), including the site-projected DOS. The dashed line corresponds to the determination of the valence band maximum E_{VBM} by linear extrapolation of the steep valence band edge.

defect states. The hybrid DFT-DOS in fig. 2.2 indicate that it is not the region with $g(E) \approx 0$ that is underestimated in this case, but the defect states are simply spread over a slightly wider energy interval. Furthermore, the HAXPES results can again be used for comparison. In the spectra for $x_{surf} \approx 1.2$, the difference between the Fermi level and the extrapolated VB edge is approximately 3 eV (fig. 5.34). Similarly, the DFT-DOS for x = 1.2 in fig. 5.35 exhibits a difference of 3 eV between E_F and the upper edge of the Ga 4p projection in the valence band. Together with the equal shape

of the HAXPES VB and the Ga-projected DOS, this indicates that the GGA-DOS are quite realistic and can be considered as a reasonable approximation of the actual DOS.

The site-projected DOS in fig. 5.35 also reveal that the defect states have rather equal contributions from gallium (4s, 4p, 3d) and oxygen (2p) orbitals. As expected from the discussion in sec. 2.3, an interpretation of the defects in a-GaO_x as the classical oxygen vacancies or Ga lone pairs therefore seems incorrect. Their shape is more complicated, as the following visualization of the electronic states will demonstrate.

5.4.2 Localized States from Amorphicity

With the existence of defect states near E_F verified beyond doubt, the next question is whether they are localized or extended states. It was tackled by investigating the spatial distribution of the partial charge density $\rho(\vec{r}) = -e|\phi(\vec{r})|^2$ for individual electronic states from the DFT calculations. Four examples for different regions of the DOS are displayed in Fig. 5.36. The two visualized electronic states from within the valence band and the conduction band (fig. 5.36b,e) have a rather homogeneous distribution of the charge density, which means that they are extended states, as expected. The defect state in fig. 5.36c, in contrast, has high ρ in the center of the DFT cell and almost none on the left and the right side. This is a localized state. The charge density appears clustered just with regard to the long c-axis of the tetragonal cell, but, since amorphous materials are isotropic, the same would be observed in a- and b-direction if the DFT cell would be larger. The electronic state at the Fermi level, which is shown in fig. 5.36d, has a localized charge density as well. This already fits nicely to the conclusion from the conductivity measurements that variable-range hopping via localized states near E_F takes place.

To ensure that all states around the Fermi level are localized, it is unfeasible to look at them all individually. Instead, the inhomogeneity of each Kohn-Sham wave function was quantified by the LOS analysis described in sec. 4.1.3. The resulting localization standard deviation σ_L, which was computed for each of the 3×10^5 DFT states in the analyzed energy range, is high for localized states and low for extended states. As an example, fig. 5.36a shows – in red – the final $\sigma_L(E)$ plot for all individual electronic states of a-$Ga_{64}O_{64}$ (x = 1.0). It is clearly visible that the values in the former band gap are overall higher than those in the valence and conduction band, indicating a mobility gap with localized states that includes the VB tail, the emerging defect states and the CB tail. For determining the mobility edges E_C and E_V that delimit this region, at least approximately, the lower envelope of $\sigma_L(E)$ must be considered (fig. 5.36a, in blue), because the mobility gap is characterized by the absence of extended states.

Fig. 5.37 shows the σ_L-envelopes for the four different O/Ga ratios that were investigated by the DFT calculations (analog plot with all states: fig. B.15). In accordance with the Anderson localization theory in sec. 2.2.1, there are already localized states in the VB and CB tails of the amorphous, stoichiometric a-$GaO_{1.5}$. It is only logical that the defect states, which appear in the band gap between

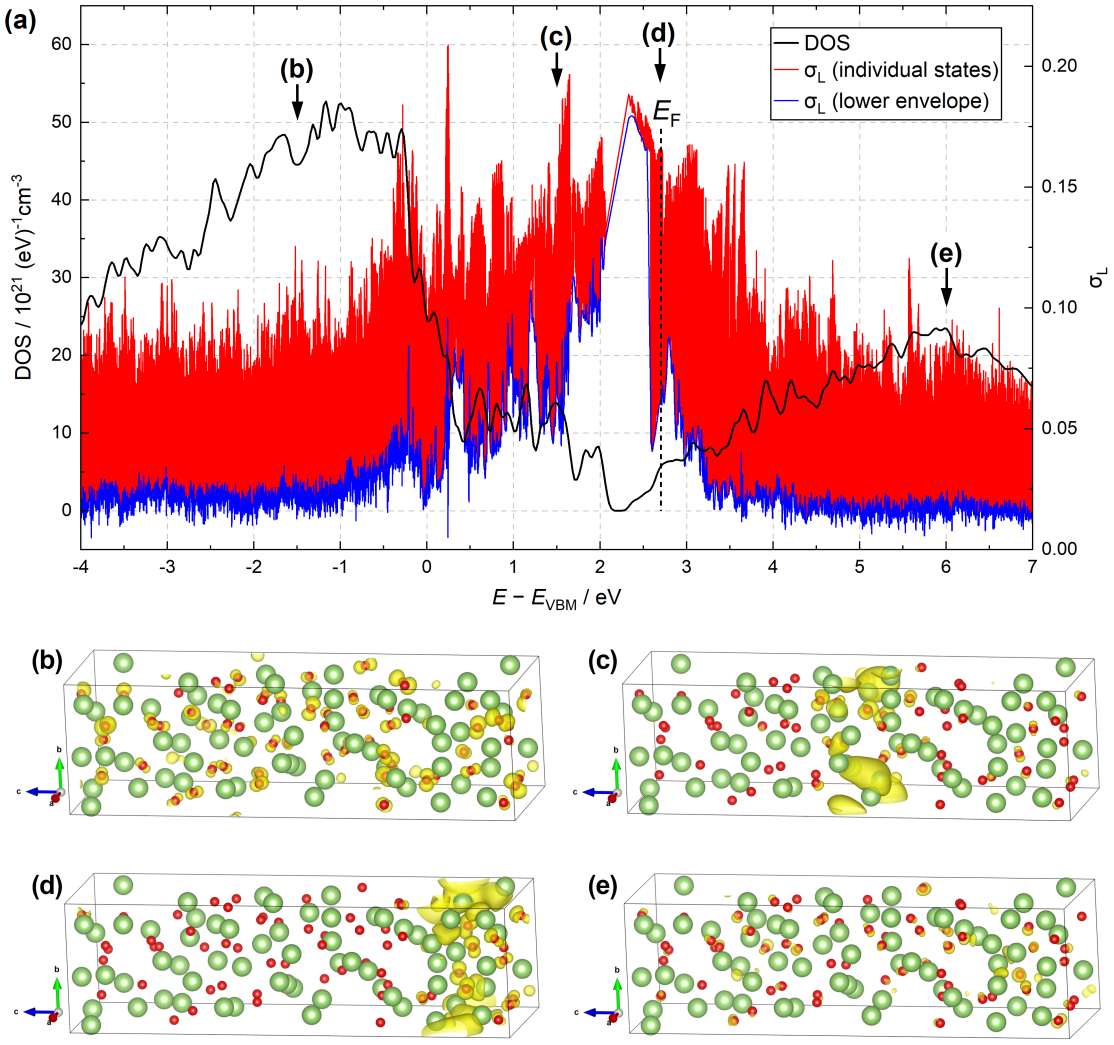


Figure 5.36: Exemplary electronic states and LOS analysis for a-$Ga_{64}O_{64}$ (x = 1.0). (a) DFT-DOS with markers for E_F (at T = 300 K) and for the four electronic states visualized below, overlaid with the localization standard deviation $\sigma_L(E)$ both for all Kohn-Sham states (red) and as the lower envelope (blue). (b-e) Partial charge density of individual Kohn-Sham states, in yellow for an isovalue of −0.17 C/cm^3 (Ga atoms in green, O atoms in red). They represent different DOS regions: (b) Valence band ($E - E_F$ = −4.2 eV). (c) Defect band ($E - E_F$ = −1.2 eV). (d) Fermi level ($E - E_F$ = 0.0 eV). (e) Conduction band ($E - E_F$ = 3.3 eV).

them with increasing oxygen deficit, are localized as well. For each investigated stoichiometry, especially the electronic states close to the Fermi level have much higher σ_L than those within the valence and conduction bands. These DFT results nicely confirm the existence of localized states at and around E_F that was suspected from the conductivity measurements. Comparison with fig. 5.34 also shows that the defect states observed in the HAXPES experiments are in the mobility gap and thus must be localized states.

The mobility edges can be estimated as the energies in fig. 5.37 where σ_L deviates from the rather constant CB and VB plateaus. An exact assignment is probably beyond the accuracy of the LOS method, but for each of the O/Ga ratios the difference $E_C - E_F$ seems to be in the range of 0.5–1.0 eV.

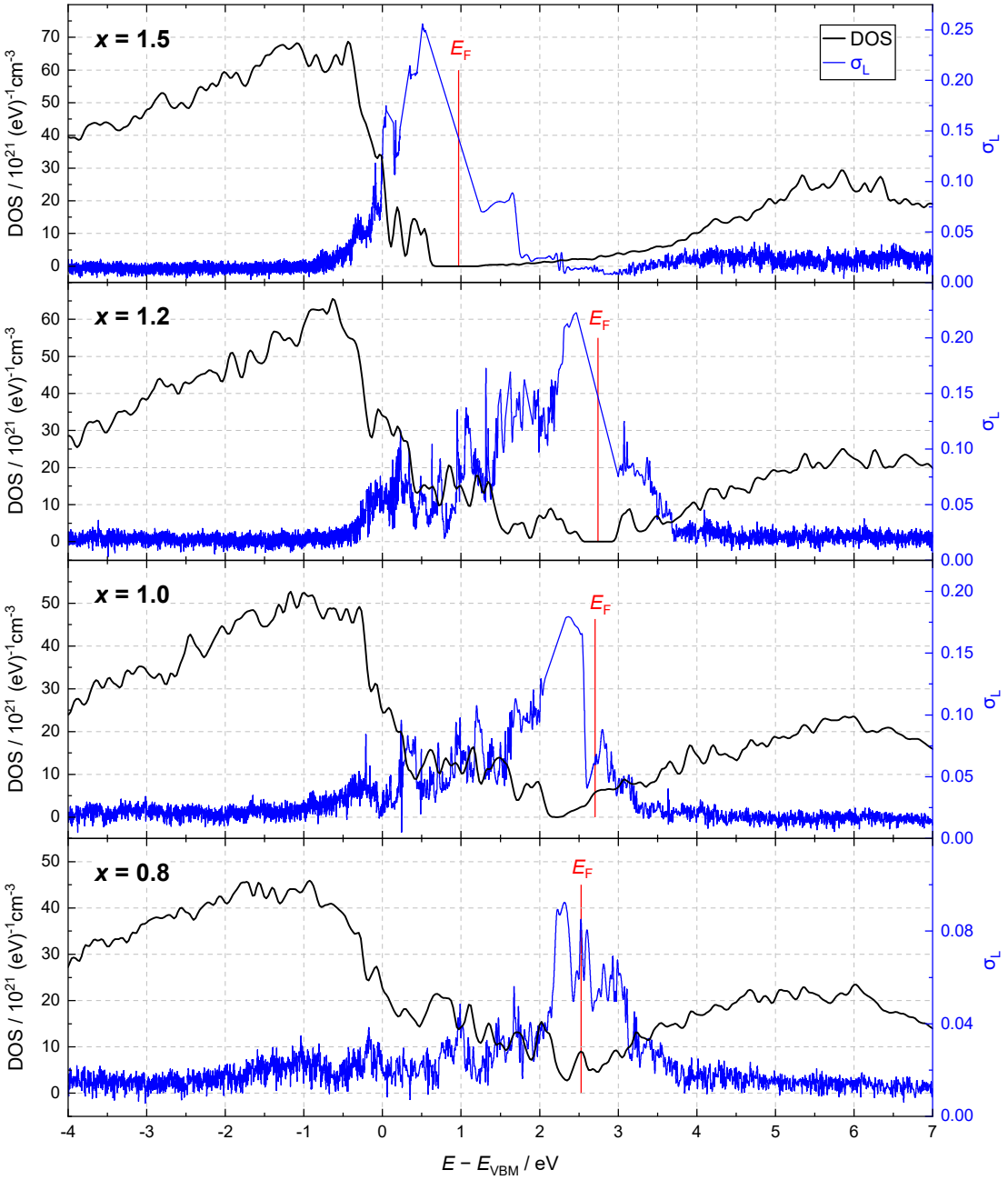


Figure 5.37: DFT-DOS in combination with the LOS results in form of the lower envelope of $\sigma_L(E)$, for a-$Ga_{64}O_{96}$ (x = 1.5), a-$Ga_{64}O_{78}$ (x = 1.2), a-$Ga_{64}O_{64}$ (x = 1.0) and a-$Ga_{72}O_{58}$ (x = 0.8). The "E_F" lines mark the Fermi level for T = 300 K.

This is nowhere near $\Delta E \approx 0.1$ eV that would be required to interpret the conductivity experiments for $x \leq 1.0$ in sec. 5.3.2 with the MEC model. Only the value of $\Delta E \approx 0.5$ eV that has been published in [32] for a-GaO_x with $x \approx 1.2$ might be realistic for MEC (cf. fig. 2.3a). Band conduction by electron holes below the VB mobility edge is no relevant option, since $E_F - E_V$ is approximately in the range of 1.5–3.5 eV according to fig. 5.37 and thus even larger than $E_C - E_F$. Furthermore, fig. 5.37 yields no clear trend for $E_C - E_F$ with regard to the stoichiometry. In contrast to the common

model of donor doping by oxygen-deficiency, most defect states appear far below E_F, which makes them irrelevant for $E_C - E_F$. Together with the increasing density of states in the CB tail, this seems to suppress the expected shrinking of $E_C - E_F$ for lower x. This might be another aspect of why VRH dominates in severely oxygen-deficient a-GaO_x.

Overall, the trends with regard to VRH transport are as expected from the previous discussion in sec. 2.3.1 and 5.3.3: The electronic states around E_F are localized due to the amorphous structure and their density successively increases the more a-GaO_x deviates from stoichiometric composition. This makes a-GaO_x a "Fermi glass", which is an amorphous material with localized states at the Fermi level as defined by Anderson.[31, 47] According to the VRH law (eq. 2.54), the above trend for the density of localized states implies an increase of σ and a decrease of T_0 for lower O/Ga ratio x, as observed in the conductivity measurements in sec. 5.3.3. However, the VRH theory by Mott was derived for a constant DOS (w.r.t. E) around E_F, which is not the case for a-GaO_x: The Fermi level is in a small band gap for a-$GaO_{1.2}$, in the CB tail with $\partial g(E)/\partial E > 0$ for a-$GaO_{1.0}$ and at a local DOS maximum for a-$GaO_{0.8}$. Therefore, it is questionable whether Mott's theory is still valid under these circumstances. The next section approaches this issue by simulating the VRH conductivity for the obtained ab-initio DOS of a-GaO_x.

5.5 Simulated Variable-Range Hopping

At the end of the previous section, an important question was raised: Does a non-constant DOS still lead to an exponential $T^{-1/4}$-dependence of the conductivity according to Mott's VRH law? Otherwise, the conclusions drawn from the conductivity measurements would essentially be invalid. To answer this question, the VRH process was simulated with the above DFT-DOS as distribution function for the localized states, using the KMC method. Since the experimentally investigated samples in this work have O/Ga ratios in the range of 0.8 to 1.1 (cf. sec. 5.1.2), the DOS of a-$GaO_{1.0}$ was used primarily in these KMC-VRH simulations. Additional to the temperature-dependence of the VRH conductivity, the simulations also yield further data that otherwise would be difficult to obtain. This includes for example the relations of σ to the chemical potential μ of the electrons or to the electric field strength F, as well as the investigation of electron diffusion.

5.5.1 Validation of the KMC Simulations

Before the discussion of the major KMC results in sec. 5.5.2 to 5.5.6, this section is dedicated to the validation of the simulations. A lot of aspects must be checked to ensure that the obtained results are physically meaningful. The representative examples shown in the following refer to a-$GaO_{1.0}$ with the default KMC parameters (if not stated otherwise), but the same analysis was also carried out for the other O/Ga ratios and the different parameter combinations employed. Specifically, the validations were done for all extrema of the used physical parameters, e.g., for the highest and lowest

temperature T, for the highest electric field strength F or for the chemical potential μ shifted to the small band gap of a-$GaO_{1.0}$.

Fig. 5.38a illustrates that the distribution of the localized states with regard to the state energy ε mirrors the DOS $g(\varepsilon)$ (cf. fig. 5.37), as required. The distribution of the electrons among the localized states, which is shown in fig. 5.38a as well, corresponds to the expected Fermi-Dirac distribution. This is confirmed by the histogram of the occupation ratio in fig. 5.38b, which accurately equals the Fermi-Dirac function $f_{FD}(\varepsilon)$.

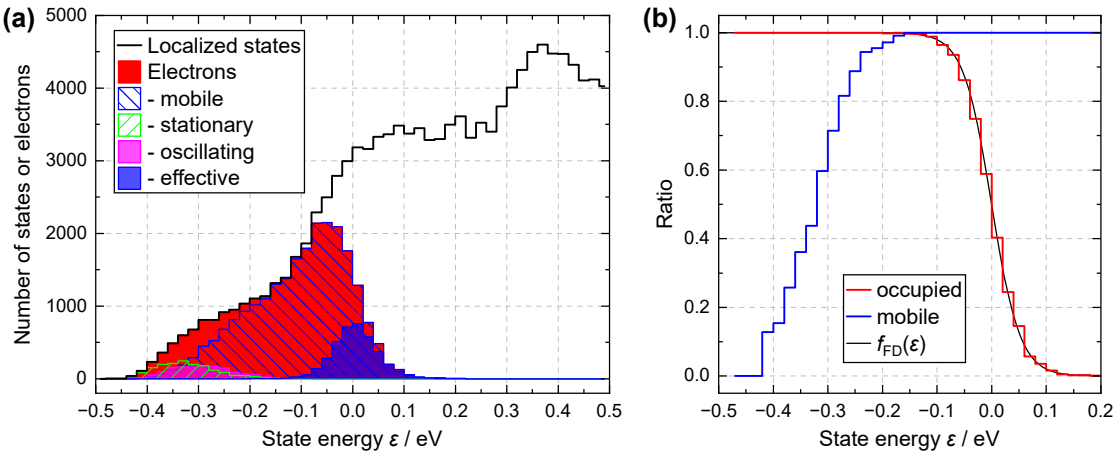


Figure 5.38: Distribution of the electrons at the end of a KMC simulation for a-$GaO_{1.0}$ at T = 300 K (DOS energy range: $[-0.5, +0.5]$ eV, bin size: 20 meV, $N_{h,sim}$ = 2×10^{10}, V_{cell} = 2.64×10^{-17} cm^3). (a) Histograms of the localized states and the electrons as a function of the state energy ε. The electrons are classified as mobile (88.4%), stationary (6.2%) or oscillating (5.4%), as defined in sec. 4.2.5. Additionally, the theoretical distribution of the effective charge carriers according to eq. 4.18 is shown (N_{eff}/N = 14.5%). (b) Histograms of the occupation ratio N/N_{st} (= fraction of the states that is occupied) and the mobile ratio N_{mob}/N (= fraction of the electrons that is mobile). The occupation ratio matches the Fermi-Dirac function $f_{FD}(\varepsilon)$, with the Fermi level at $\varepsilon = \mu = 0$ eV.

Fig. 5.38a also demonstrates the different behavior of the electrons. Towards the lower end of the specified DOS energy range, there are stationary electrons, which did no hop during the simulation, and oscillating electrons, which did not move except for oscillations. These electrons did not contribute to charge transport at all. The majority of the electrons in the KMC cell was mobile in the sense that they made one or more non-oscillating hops. These are the electrons that acted as charge carriers during the simulation. This does not mean, however, that each mobile electron moved for the whole time. Fig. B.16 shows that the individual number of hops per electron may differ vastly. For example, occasionally an electron from deep below the Fermi level hops to a higher state in the region of persistent charge transport around $\varepsilon = \mu$, while another electron hops to the lower state and becomes temporarily trapped. In this simplified case, both artificially distinguished electrons moved, but effectively only one of them took part in conduction at the same time. Consequently, the number of mobile electrons increases and the stationary or oscillating electrons become less over the course of the simulation, as shown in fig. B.17. In reality, no clear distinction can be made between a localized state that acts as trap and a state that is involved in charge transport, since the

energetic distribution of the localized states is continuous. As a result, each mobile electron generally has a certain fraction of the total simulated time span where its movement was impeded by the other electrons.[110] In other words, in each moment only a part of the mobile electrons contributes to the charge transport. While it is quite impossible to determine which of the artificially distinguished electrons belongs to this group in any given moment, their amount is readily available through the effective charge carrier model, which was introduced in sec. 2.1.1. The theoretical distribution of the effective charge carriers according to eq. 4.18 is displayed in fig. 5.38a as well. It is a rather symmetric peak around the Fermi level and much narrower than the histogram of the mobile electrons, as expected.

In order to obtain correct KMC results, the simulated movement of the electrons must be reinterpreted as a movement of these effective charge carriers, as explained in sec. 4.2.5 and A.4. Considering all N electrons as charge carriers in the evaluation of the simulations, as it has been done in the KMC-VRH literature mentioned in sec. 4.2, would lead to the following ambiguities in the general case of an arbitrary DOS: In fig. 5.38, lowering $\varepsilon_{\mathrm{min}}$ would add more and more stationary electrons to the simulation (cf. fig. 5.37).[111] The number of electrons N would increase but for example the sum of all displacements X_k or the sum of all squared displacements X_k^2 would not change. As a result, it would be possible to vary the respective averages $\langle X \rangle_N$ and $\langle X^2 \rangle_N$ at will by changing $\varepsilon_{\mathrm{min}}$. This would make the KMC results, namely the mobility and the Fickian diffusion coefficient, dependent on the chosen DOS energy range, which would be completely unphysical. Using instead only the mobile electrons N_{mob} as charge carriers would lead to comparable problems. On the one hand, N_{mob} changes over the course of the simulation (fig. B.17), which makes for example the relation between the N_{mob}-based average displacement $\langle X \rangle_{\mathrm{mob}}$ and the time τ non-linear. On the other hand, N_{mob} converges to N (after extremely long simulation time, cf. fig. B.17), and not to any lower value, which leads to the same ambiguities as when using N directly. In the aforementioned KMC literature, these consequences have been avoided either by regarding only the conductivity – where the number of charge carriers cancels from the final result (eq. A.49) – or by using DOS models that have no localized states below the Fermi level ($N_{\mathrm{eff}} \approx N$). In the general case, however, N_{eff} must be used for the charge carrier definition in KMC-VRH simulations.

The fact that charge transport predominantly occurs in a narrow region around the Fermi level is also illustrated by the distribution of the hops that occurred during the KMC simulation. Fig. 5.39 shows the number of outgoing hops as a function of the state energy where each hop started. Both the complete hop histogram in fig. 5.39a and the one for the non-oscillating hops in fig. 5.39b have a similar peak shape as the effective charge carrier distribution in fig. 5.38a, with approximately the same full width at half maximum. Fig. 5.39a also shows that the majority of the hops are oscilla-

[110] There is a notable exception to the general case: When all localized states have $\varepsilon_i \gg \mu$, i.e., when the Fermi level is well below all states, then the electrons are dilute and approximately do not block each others movement. This situation may occur for example in an amorphous, n-type non-degenerate semiconductor with localized CB tail states. In this case, each electron acts as a charge carrier during the whole simulation. This is also reflected in $N_{\mathrm{eff}} \approx N$ (due to $1 - f_{\mathrm{FD}}(\varepsilon_i) \approx 1$, cf. eqs. 4.16 and 4.18).

[111] The KMC histogram for a-$GaO_{0.8}$ shown later in fig. 5.51a probably illustrates this issue even better.

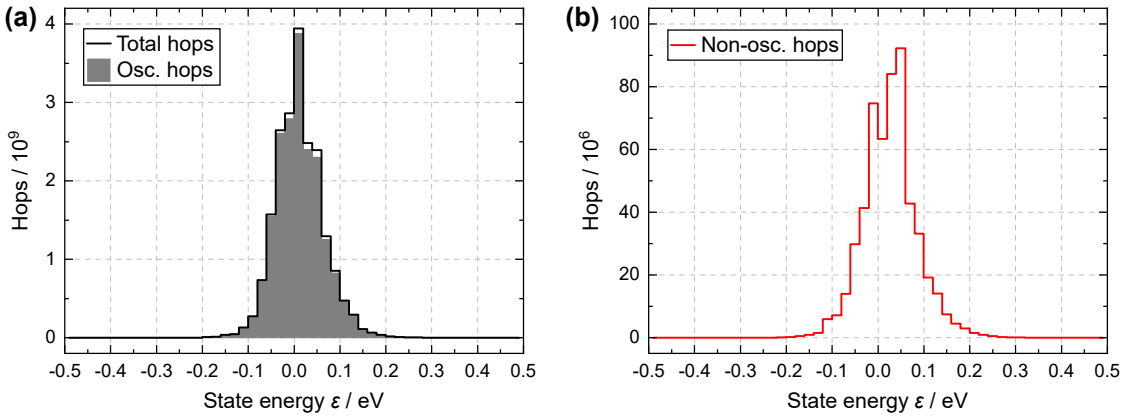


Figure 5.39: Energetic distribution of the hops that took place during a KMC simulation for a-$GaO_{1.0}$ at T = 300 K (DOS energy range: $[-0.5, +0.5]$ eV, bin size: 20 meV, same simulation as fig. 5.38). (a) Histogram of the total hops and the oscillating hops (97.3%) outgoing from a certain state energy ε, for equilibration and main simulation together ($N_{h,eq} = 3\times10^8$, $N_{h,sim} = 2\times10^{10}$). (b) Corresponding histogram of the non-oscillating hops (5.44×10^8, 2.7%). The respective graphs for the incoming hops, where ε is the energy of each hop's destination state, are virtually identical.

tions, which are an inherent part of the VRH process even though they yield no displacement (cf. sec. 4.2.6). The peak of the non-osc. hops in fig. 5.39b is not completely symmetric with regard to the Fermi level at $\varepsilon = \mu = 0$ eV but slightly shifted to higher energy. The reason for this is the increased number of unoccupied states for $\varepsilon > \mu$ due to the positive slope of the DOS ($\partial g(\varepsilon)/\partial\varepsilon > 0$, cf. fig. 5.38a).

The distributions of the electrons and the hops with regard to the state energy ε in figs. 5.38 and 5.39 also indicate that the chosen DOS energy range $[\varepsilon_{min}, \varepsilon_{max}]$ is sufficiently wide. The ratios in fig. 5.38b show that in this example $\varepsilon_{min} = -0.5$ eV is low enough to include all mobile electrons and $\varepsilon_{max} = +0.5$ eV is high enough to not exclude any occupied states above the Fermi level. Both the N_{eff}-histogram in fig. 5.38a and the hop distributions in fig. 5.39 decay to zero long before reaching the boundaries of the energy range. Another aspect to be regarded in this context is which localized states were "used" – i.e., involved in any hop – during the simulation. The respective histogram in fig. 5.40a demonstrates that despite the narrow hop distributions also states far from the Fermi level are sometimes used in the hopping process. It also shows that the specified DOS energy range is large enough to not exclude any potentially used states. The question arises, though, if it is really necessary to include all mobile electrons and all used states in this manner, considering that the distributions of the effective charge carriers and the non-osc. hops are so much narrower. Since the ultimate criterion for a valid DOS energy range $[\varepsilon_{min}, \varepsilon_{max}]$ is that is does not affect the conductivity, it was tested how σ changes when different energy ranges are used. In fig. 5.40b it can be seen that the conductivity stays constant when shrinking the DOS energy range from ±0.5 eV down to ±0.175 eV, while even narrower ranges result in a decreased conductivity. The lower limit of ±0.175 eV roughly corresponds to the visible range of the N_{eff} and hop histograms (figs. 5.38a, 5.39), which makes sense. The steep drop of σ beyond this limit indicates that one must be careful when reducing the energy range. At ±0.1 eV, for example, the conductivity is almost halved, even though only 3% of

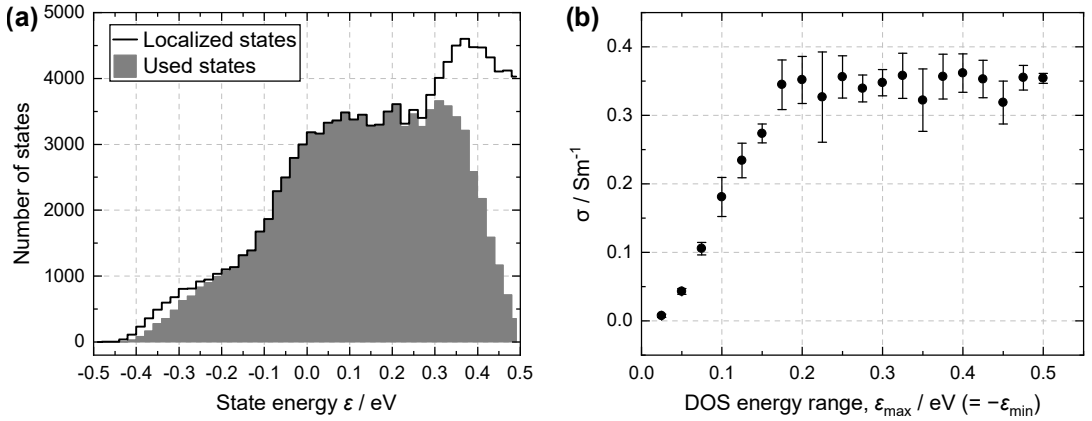


Figure 5.40: Influence of the DOS energy range on the KMC simulations for a-$GaO_{1.0}$ at T = 300 K. (a) Histogram of the used states, i.e., of those where at least one hop started or ended during the simulation, for the default DOS energy range of $[-0.5, +0.5]$ eV (bin size: 20 meV, $N_{h,sim} = 2{\times}10^{10}$, same simulation as fig. 5.38). (b) VRH conductivity obtained with different symmetric DOS energy ranges from $[-0.5, +0.5]$ eV down to $[-0.025, +0.025]$ eV.

the effective charge carriers and 6% of the hops are outside of this range in figs. 5.38 and 5.39. In this work, the DOS energy range generally was adjusted to include approximately all mobile electrons and used states, which is definitely large enough.

The calculation of the number of effective charge carriers N_{eff} through eq. 4.18 requires that the electrons' steady-state distribution equals $f_{FD}(\varepsilon)$. Fig. 5.38b already demonstrated that this is the case when the simulation is started with Fermi-Dirac-distributed electrons. The distribution remains unchanged in this case.[112] However, when the initial electron distribution is a step function, with the localized states filled up to the Fermi level, then the electron distribution adjusts on its own during the hopping process. This can be seen in fig. 5.41a, where the occupation histogram gradually becomes equal to $f_{FD}(\varepsilon)$ with increasing number of hops. It must be stressed that this is not explicitly enforced by the KMC software. The electrons adopt the Fermi-Dirac distribution automatically because they can only hop to unoccupied states (Pauli principle) and hops upwards in energy have an exponentially decaying probability (Miller-Abrahams equation). Monitoring the number of electrons above the Fermi level and the number of unoccupied states below the Fermi level allows to determine when the steady state is reached (fig. 5.41b).

Besides the DOS energy range, the spatial and energetic cut-offs r_{cut} and ε_{cut} are the most important technical parameters to validate. They correspond to the maximum length and energy difference of the considered hopping paths in the KMC cell and choosing them too low would exclude relevant paths from the simulation. Fig. 5.42 shows – in accordance with the Miller-Abrahams equation (eq. 2.46) – that the longer the hop distance r_{ij} and the higher the absolute state energy difference $|\varepsilon_{ij}|$, the fewer respective hops occur during the simulation. The same is observed in fig. 5.42 for the "used" paths, i.e., for those paths that were taken by electrons (in relation to all available paths).

[112] The only exception to this are simulations with high electric field strength, as will be discussed later in sec. 5.5.4.

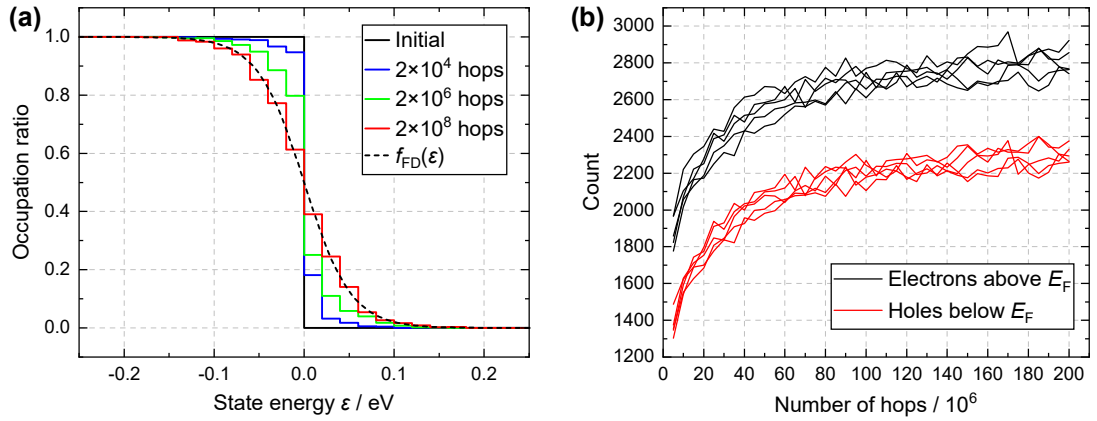


Figure 5.41: Redistribution of the electrons during KMC simulations for a-$GaO_{1.0}$ at T = 300 K (DOS energy range: $[-0.5, +0.5]$ eV, bin size: 20 meV, $N_{h,pre}$ = 0, $N_{h,eq}$ = 0), when initially the electrons are placed according to a step function (i.e., all states occupied up to approx. $\varepsilon = \mu = 0$ eV). (a) Successive development of the histogram for the occupation ratio N/N_{st}, in comparison to the Fermi-Dirac function $f_{FD}(\varepsilon)$. (b) Convergence of the number of electrons above the Fermi level ($\varepsilon_k > \mu$) and the number of unoccupied states below the Fermi level ($\varepsilon_i < \mu$) for 5 repetitions with different seed. In the steady state, the number of electrons above μ is higher than the number of electron holes below μ, since there are more localized states above μ than below (cf. fig. 5.38).

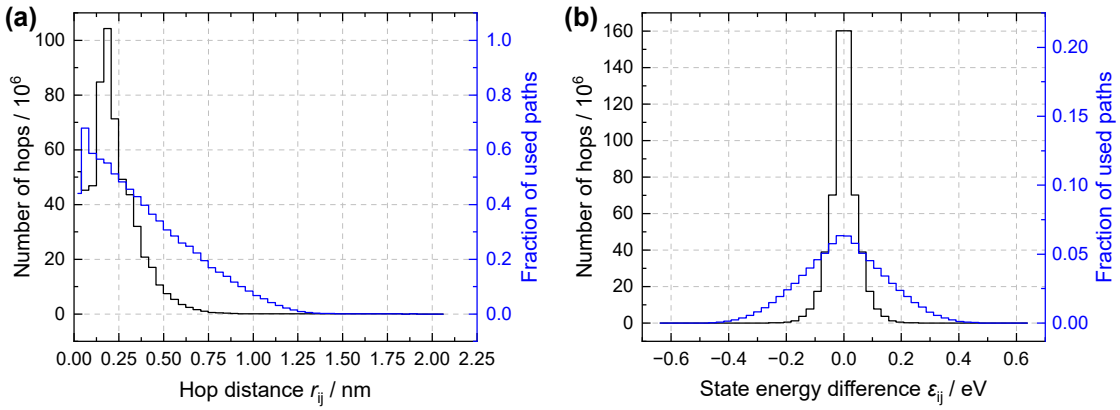


Figure 5.42: Cut-off validation of a KMC simulation for a-$GaO_{1.0}$ at T = 300 K. (a) Histograms of the number of hops that occurred (black) and the fraction of the paths that were used at least once (blue) during the equilibration and the main simulation, as a function of the path length $r_{ij} = |\vec{r}_{ij}|$ (bin size: 0.042 nm). The spatial cut-off is r_{cut} = 2.08 nm and the longest hop had r_{ij} = 1.89 nm. (b) Histograms of the number of hops (black) and the fraction of used paths (blue) with a certain state energy difference ε_{ij} (bin size: 0.026 eV). The energetic cut-off is ε_{cut} = 0.65 eV and the hop with the largest energy difference had ε_{ij} = 0.52 eV. Both cut-offs are the result of the automatic adjustment after the equilibration (previous values: r_{cut} = 2.14 nm, ε_{cut} = 0.75 eV), which reduced the total number of paths by 1.1×10^6 (10%).

The symmetry of the ε_{ij}-distributions in fig. 5.42b originates from the fact that hops with $\varepsilon_{ij} < 0$ still require the target state to be unoccupied and thus depend on equivalent hops with $\varepsilon_{ij} > 0$. Both the histograms for r_{ij} and those for ε_{ij} decay long before reaching the cut-offs, which means that r_{cut} and ε_{cut} were sufficiently large. They actually could have been reduced further until σ is affected, but the simulations were typically carried out with such generous cut-offs. In this example,

the cut-offs led to a total of approx. 10^7 paths considered in the simulation, which amounts to an average of 83 outgoing paths per localized state.

The basic results of the VRH simulation are the displacements X_k, Y_k and Z_k of the mobile electrons after the simulated time span τ. Fig. 5.43 shows that the displacements have bell-shaped distributions, which are symmetric around the origin for Y_k and Z_k, as expected. The X_k-distribution of this particularly long simulation has a clear shift in $+x$-direction due to the applied electric field. For the usually used, lower $N_{\mathrm{h,sim}} = 3{\times}10^8$ this shift is more subtle, but $\langle X \rangle$ still differs significantly from $\langle Y \rangle$ and $\langle Z \rangle$. Overall, the drift in x-direction and the diffusion in y- and z-direction are reflected in the displacements as anticipated. There are, however, first signs of deviations for

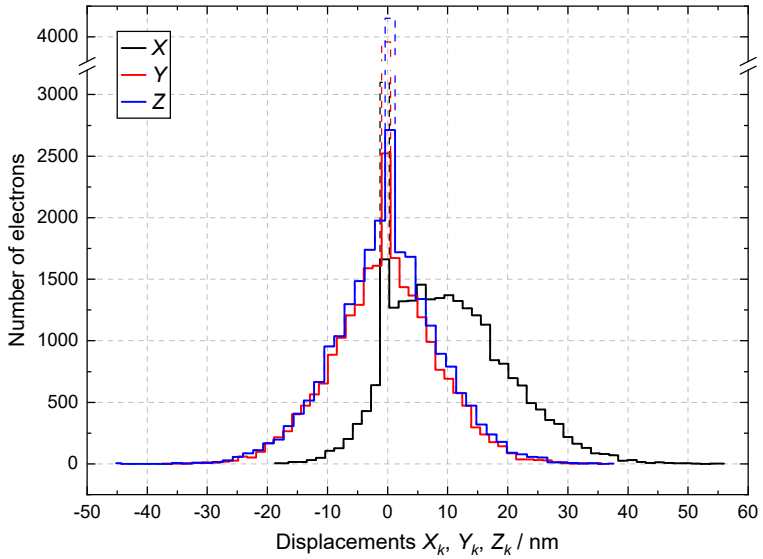


Figure 5.43: Histogram of the electron displacements X_k, Y_k and Z_k at the end of a KMC simulation for a-$GaO_{1.0}$ at $T = 300$ K (same simulation as fig. 5.38). The data originates from the equilibration and the main simulation together, with especially large $N_{\mathrm{h,sim}} = 2{\times}10^{10}$ to produce a visible shift in x-direction. The solid lines refer to the mobile electrons only, while the dashed lines include the oscillating electrons. The stationary electrons are not shown.

the diffusion in x-direction. The X_k-distribution features an additional spike at the origin, which most probably belongs to electrons with loop-like oscillations that involve more than two states. These electrons seem to resist the influence of the electric field, even though they do not belong to the conventional oscillating electrons. For the calculation of the diffusion coefficient in x-direction from the variance $\mathrm{Var}(X)$, such inert electrons must be excluded from the statistics (cf. sec. A.4), but it is difficult to discern multi-center oscillations from a random return of an electron to its origin. The influence of this issue should be rather small, though, since it only concerns about 700 of the $2.7{\times}10^4$ mobile electrons (2.6%), as inferred from the spike in fig. 5.43. Another requirement for the calculation of D_x from $\mathrm{Var}(X)$ is that the X_k-distribution is symmetric around the average displacement (cf. sec. A.4). The X_k-histogram in fig. 5.43, however, has a slight asymmetry that becomes more pronounced the longer the simulation runs. This distortion of the bell-shaped X_k-

distribution is inevitable at finite electric field and it becomes stronger for higher field strength F. Consequently, $\mathrm{Var}(X)$ only reflects the diffusion in x-direction for short simulations ($\tau \to 0$) or vanishing electric field ($F \to 0$). In contrast, the transverse diffusion coefficient D_{yz} obtained from $\langle Y^2 \rangle$ and $\langle Z^2 \rangle$ has no such limited validity. Therefore, since amorphous materials are isotropic, D_{yz} is a better measure for the diffusion in x-direction than the approximate D_x. The isotropic Fickian diffusion coefficient is hence given by $D = D_{yz}$.

The calculation of the simulation results – as described in sec. 4.2.5 and A.4 – not only requires proper displacement distributions but also a linear increase of the different displacement averages $\langle X \rangle$, $\mathrm{Var}(X)$, $\langle Y^2 \rangle$ and $\langle Z^2 \rangle$ with the time τ. Fig. 5.44a,b shows that this latter condition is typically not fulfilled right from the start of the simulations, even when the electrons are already Fermi-Dirac-distributed. Especially the diffusion-related quantities in fig. 5.44b exhibit a system-

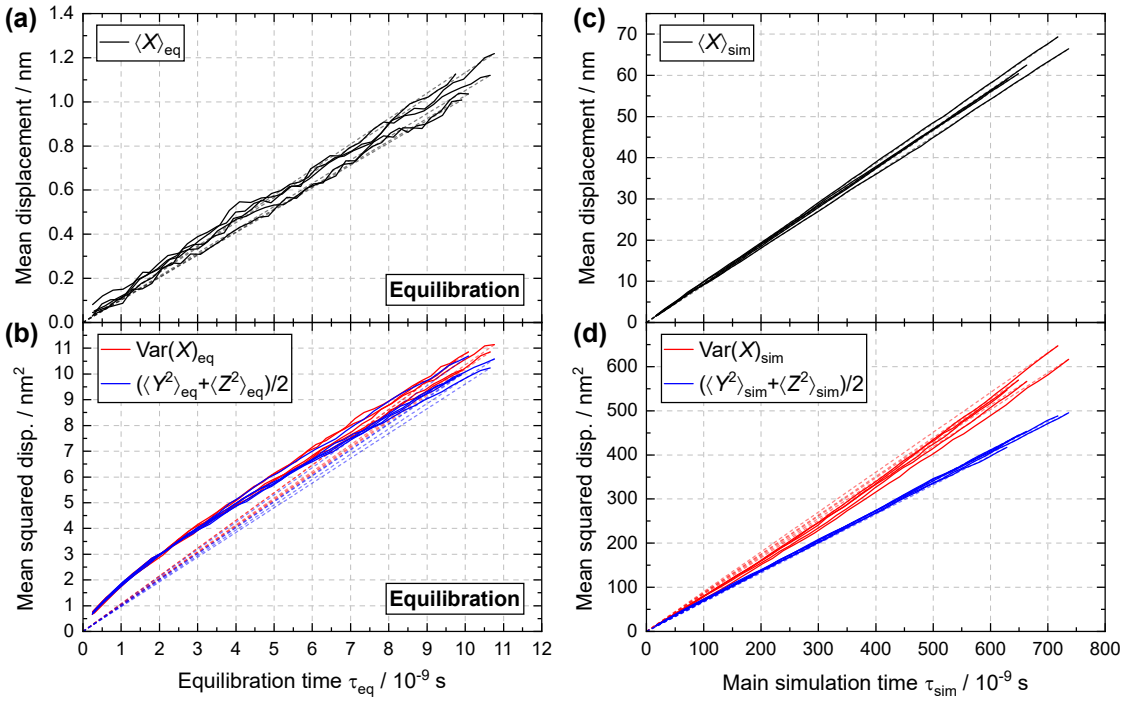


Figure 5.44: Displacement averages vs. simulated time span τ, from the 5 repetitions of a KMC simulation for a-$GaO_{1.0}$ at T = 300 K ($N_{h,eq} = 3{\times}10^8$, $N_{h,sim} = 2{\times}10^{10}$, same simulation as fig. 5.38). (a) Mean displacement in x-direction during the equilibration. (b) Displacement variance in x-direction and mean squared displacement in y- and z-direction during the equilibration. (c,d) Respective plots for the main simulation, relative to the values at the end of the equilibration stage in (a,b). The dashed straight lines are a guide for the eye and point from each line's final point to the origin.

atic curvature because the bell shape of the displacement distributions has yet to form. This is the reason why an equilibration of these distributions is necessary before the actual main simulation (particularly when the latter is kept short). The displacement averages during the main simulation are defined relative to the final values of the equilibration (sec. 4.2.5). Fig. 5.44c,d demonstrates that this leads to linear relations for the main simulation, as required. The only exception to this

is $\mathrm{Var}(X)_{\mathrm{sim}}$ vs. τ_{sim} in fig. 5.44d, which stays slightly non-linear irregardless of the length of the equilibration or the main simulation. The reason for this is the already discussed asymmetry of the X_k-distribution (fig. 5.43), which makes the derivation of the linear relation between $\mathrm{Var}(X)$ and τ in sec. A.4 invalid. Since the asymmetry increases over time, so does the discrepancy between $\mathrm{Var}(X)$ and $(\langle Y^2\rangle + \langle Z^2\rangle)/2$. They only stay equal in simulations without an electric field ($F = 0$).

The linearity of the plots in fig. 5.44c,d indicates that the main simulation can also be kept much shorter than $N_{\mathrm{h,sim}} = 2{\times}10^{10}$. Indeed, the default $N_{\mathrm{h,sim}} = 3{\times}10^{8}$ already yields linear plots and proper convergence of σ and D_{yz}, as demonstrated in fig. 5.45. The slightly time-varying D_x deviates from D_{yz}, of course, due to the already discussed reasons, but for these rather short simulations with a weak electric field the difference is quite small.

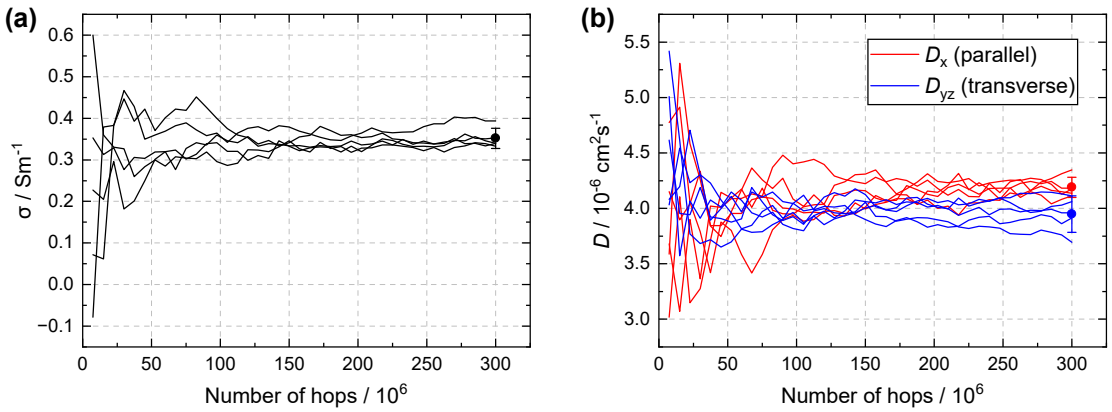


Figure 5.45: Convergence of the conductivity σ, the parallel Fickian diffusion coefficient D_x and the transverse Fickian diffusion coefficient D_{yz} during the main simulation stage of 5 KMC repetitions for a-$GaO_{1.0}$ at $T = 300\,\mathrm{K}$ ($N_{\mathrm{h,eq}} = 3{\times}10^{8}$, $N_{\mathrm{h,sim}} = 3{\times}10^{8}$). The final average values are shown as dot.

The final results should not depend on the number of localized states N_{st} employed in the simulations, since this parameter just determines the size of the investigated KMC cell. In principle, low N_{st} makes the randomly generated localized states less representative of the DOS. This can introduce artificially large variations between the different KMC cells of the repetitions, which increases the statistical error of the averaged results. High N_{st} improves the accuracy of the simulations, but only when the associated slower convergence of all distributions is accounted for by an increase of the simulation time. Otherwise the results may suffer from incomplete convergence. Fig. 5.46a demonstrates that the default $N_{\mathrm{st}} = 12{\times}10^{4}$ is an appropriate choice, because there is not significant variation of the conductivity in a broad range around it.

Furthermore, it must be verified that the default electric field strength F in the simulations is in the weak-field regime where σ does not change with F. Otherwise the results could not be compared to the conductivity measurements. Fig. 5.46b shows that the conductivity indeed stays approximately constant around the default $F = 0.0635\,\mathrm{MV/cm}$. The thorough investigation of the F-dependence of σ later in sec. 5.5.4 will prove this even further.

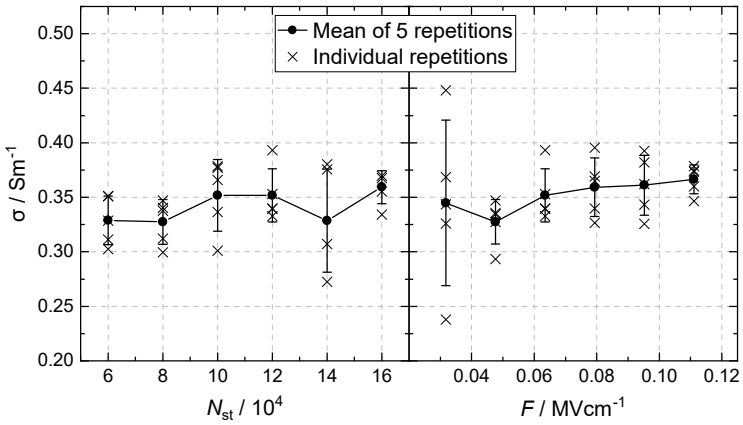


Figure 5.46: Independence of the conductivity σ from the number of states N_{st} and the electric field strength F in KMC simulations for a-$GaO_{1.0}$ at T = 300 K. (a) Variation of N_{st}, with the default F = 0.0635 MV/cm. (b) Variation of F, with the the default $N_{st} = 12{\times}10^4$.

Finally, one aspect of these later simulations with varied F shall already be addressed here, because an unexpected effect emerges that would have gone unnoticed without the careful validation of the simulations. At high electric field strength, it was observed that the mean displacements $\langle Y \rangle_{sim}$ and $\langle Z \rangle_{sim}$ no longer stay close to zero but vary linearly with the time τ_{sim} (fig. 5.47a). Even though the electric field points in $-x$-direction, there is definitely a shift of the Y_k- and Z_k-distributions away from the origin. The direction of the shift varies randomly from repetition to repetition, though, which hints at the underlying cause: In the randomly generated KMC cells, there is always a slight imbalance between the paths in $(+x, +y)$-, $(+x, -y)$-, $(+x, +z)$- and $(+x, -z)$-direction. Since the respective reverse paths with $x_{ji} < 0$ have a much lower probability at high F, the strong electric field amplifies these imbalances and causes artificial drift fluxes perpendicular to $\vec{F}$. This effect is not observed for weak electric fields, because then the reverse paths approximately balance those with positive x-component.[113] In consequence of the random shift of the Y_k- and Z_k-distributions, it is not valid to calculate the Fickian diffusion coefficient D_{yz} from the mean squared displacements $\langle Y^2 \rangle_{sim}$ and $\langle Z^2 \rangle_{sim}$ at high electric field strength. Instead, the displacement variances $\mathrm{Var}(Y)_{sim}$ and $\mathrm{Var}(Z)_{sim}$ must be used in this case, analog to the x-direction (cf. sec. A.4). Fig. 5.47a shows that the variances have slightly better linearity and a significantly different absolute value than the mean squared displacements. Accordingly, D_{yz} in fig. 5.47b differs depending on whether it is obtained from $\langle Y^2 \rangle_{sim}$ and $\langle Z^2 \rangle_{sim}$ or from $\mathrm{Var}(Y)_{sim}$ and $\mathrm{Var}(Z)_{sim}$. Only the latter yields the correct value at high electric field strength. During even longer simulations with high F it also becomes visible that D_{yz} does not converge to a constant value when it is calculated from the mean squared displacements.

[113] For weak electric fields, the ratio of non-oscillating hops with $x_{ij} > 0$ in relation to all non-osc. hops is very close to 50% (e.g., 50.045% for fig. 5.43 with F = 0.0635 MV/cm), while at high F this ratio is far above 50% (e.g., 75% to 78% for fig. 5.47 with F = 5.56 MV/cm).

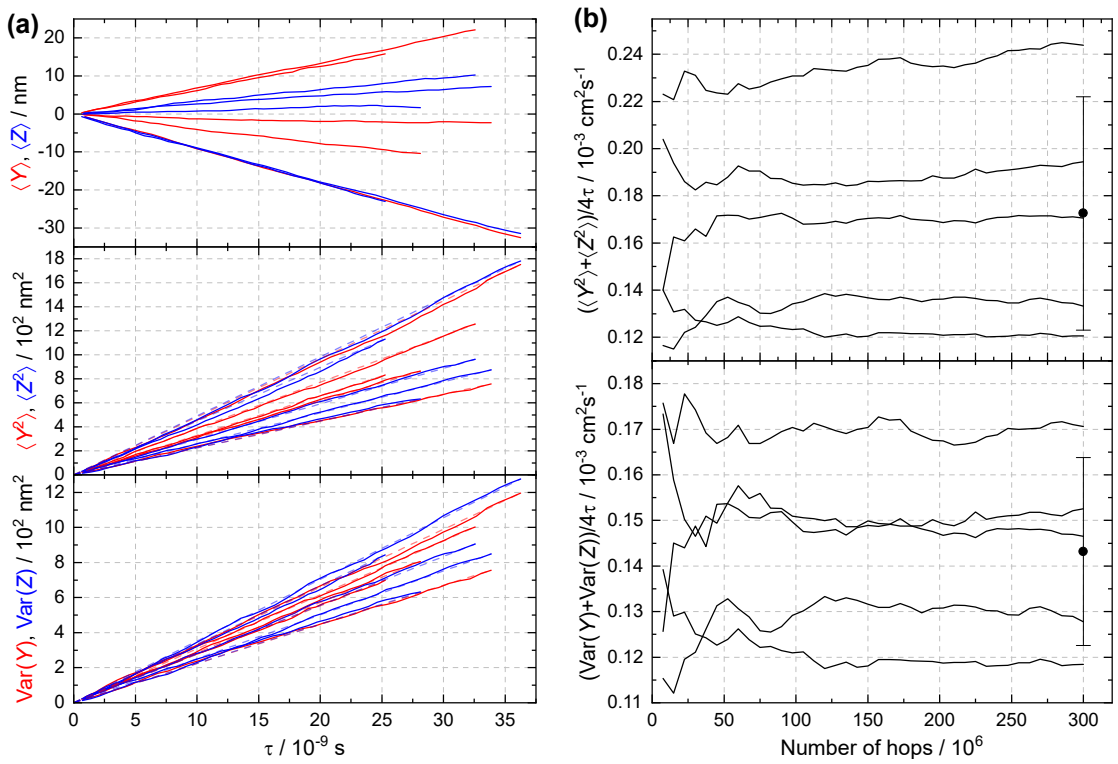


Figure 5.47: Investigation of the diffusion perpendicular to a strong electric field with $F = 5.56\,\mathrm{MV/cm}$ during 5 KMC repetitions for a-$GaO_{1.0}$ at $T = 300\,\mathrm{K}$. All displayed quantities refer to the main simulation (notation "sim" omitted for brevity). (a) Mean displacements (top), mean squared displacements (middle) and displacement variances (bottom) in y- and z-direction. The dashed straight lines are a guide for the eye and connect each plots' last data point to the origin. (b) Convergence of the transverse Fickian diffusion coefficient D_{yz}, calculated either from the mean squared displacements (top) or from the displacement variances (bottom). The final average of the repetitions is shown as dot.

5.5.2 Quantitative Agreement with Experiments

After the above examination of the simulations' inner workings, now the aforementioned question about the T-dependence of the VRH conductivity can be tackled. At first, a broad range of different localization radii α were tested, since this quantity is unknown for a-GaO_x, of course. For the attempt frequency ν_0, which is essentially a scaling factor for the conductivity, the common Debye frequency of 10^{13} Hz was used initially. The resulting temperature series in fig. 5.48a follow Mott's VRH law $\sigma = \sigma_0 \exp[-(T_0/T)^{1/4}]$ (eq. 2.54) for each α, even though the employed DFT-DOS of a-$GaO_{1.0}$ varies around the Fermi level. At extremely low T, this would be expected, because the relevant DOS energy range becomes successively more narrow. For the temperature range investigated here, however, the N_{eff} and hop histograms shown above in figs. 5.38a and 5.39b imply a rather broad relevant ε-range (e.g., ±0.175 eV, cf. fig. 5.40), where the DOS is definitely not constant. Apparently, Mott's VRH law is still applicable for such a moderately varying DOS. This means that the conclusions drawn from the exponential $T^{-1/4}$-dependence of the measured σ in sec. 5.3.2 indeed are valid.

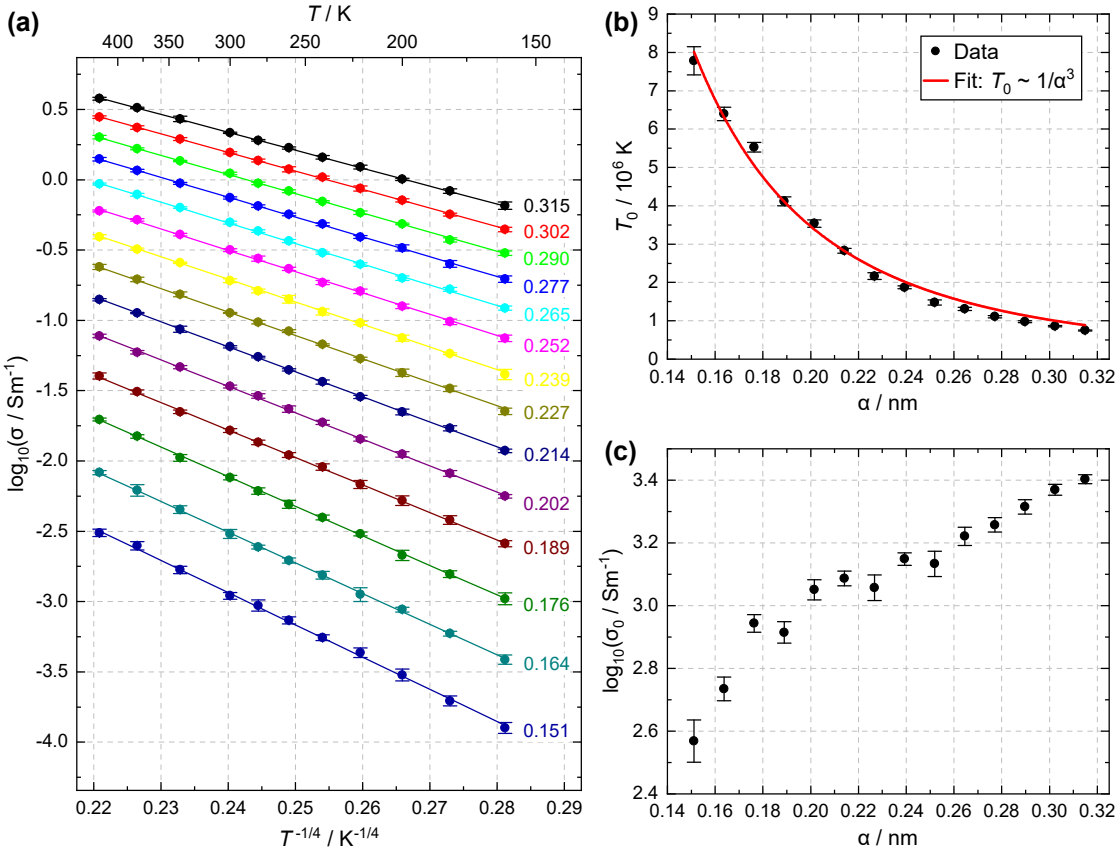


Figure 5.48: KMC simulations for a-$GaO_{1.0}$ with varied temperature T and localization radius α (with constant $\nu_0 = 10^{13}$ Hz). (a) VRH plots for the different, nearby shown α values (in nm), with linear fits according to Mott's VRH law (eq. 2.54). (b) The T_0 values obtained from the fits in (a) as a function of the localization radius, with the fit errors as error bars. The relation between T_0 and α can be approximated by $T_0 = c\alpha^{-3}$ with $c = (2.77 \pm 0.06) \times 10^4$ nm^{3}K. (c) The σ_0 values obtained from the fits in (a) for the different α.

The parameters T_0 and σ_0 obtained from the VRH fits in fig. 5.48a also vary with the localization radius α in the way expected from the classical VRH theory. Fig. 5.48b shows that T_0 decreases approximately according to $T_0 \propto \alpha^{-3}$, as predicted by Mott in eq. 2.46. The pre-exponential factor σ_0 in fig. 5.48c increases with α due to the increasing wave function overlap, but, as expected from the discussion of Mott's method in sec. 2.2.3, it does not follow Mott's estimate $\sigma_0 \propto \alpha^{1/2}$. The trends for T_0 vs. α and σ_0 vs. α together lead to a pronounced decrease of the conductivity for lower α (fig. 5.48a).

In the next step, these observations for the influence of the localization radius were utilized for a systematic adjustment of α and ν_0 in order to reproduce the experimental conductivity data quantitatively. As a proof of principle, this was done for an exemplary a-GaO_x sample with an O/Ga ratio of $x \approx 1.0$, whose conductivity measurements were already shown in fig. 5.30 in sec. 5.3.2. After successive adjustment of the two parameters, quantitative agreement was reached for $\alpha = 0.1512$ nm and $\nu_0 = 3.073 \times 10^{15}$ Hz. Fig. 5.49 shows that the simulated conductivities agree remarkably well with the experimental data, both in terms of the T-dependence and with regard to the absolute

values. In principle, a similar adjustment is possible for all measured samples, but the iterative fine-tuning of α and ν_0 is quite time-consuming. Fig. 5.49 also demonstrates – more precisely than fig. 5.48a – that the KMC simulations closely follow the VRH law without systematic deviations in the fit residuals, while the MEC law is clearly inappropriate as expected for VRH simulations.

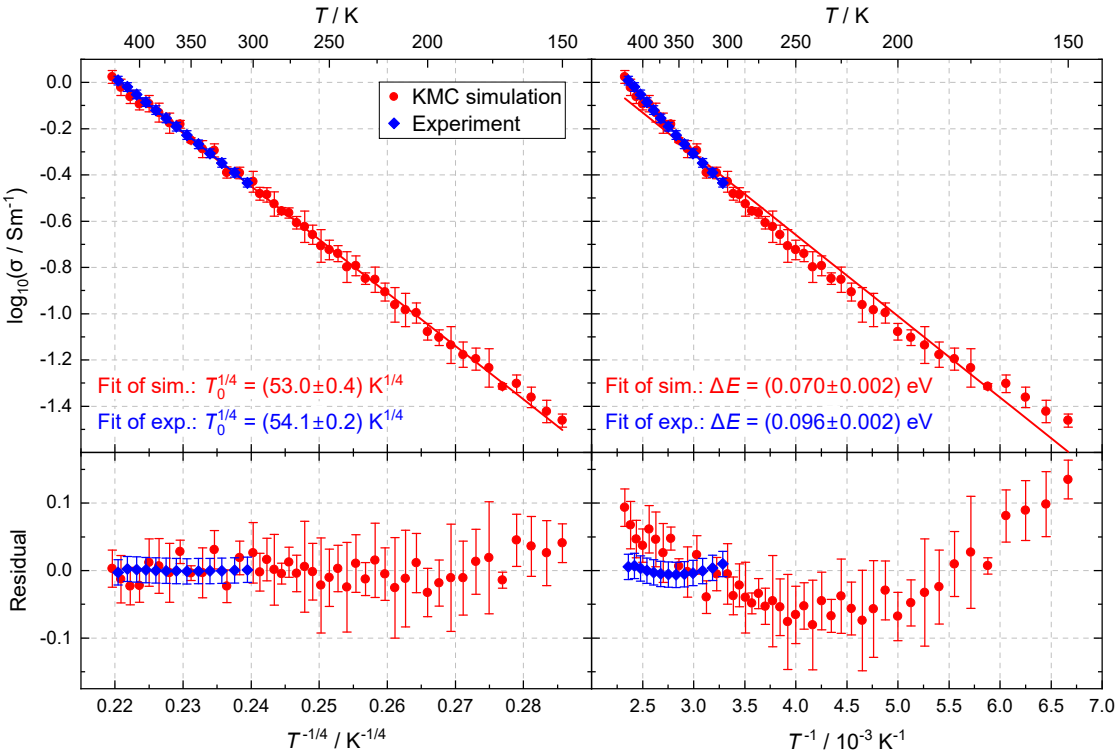


Figure 5.49: Quantitative agreement between an exemplary conductivity measurement (blue) and the KMC-VRH simulations for a-$GaO_{1.0}$ with adjusted localization radius $\alpha = 0.1512\,\text{nm}$ and attempt frequency $\nu_0 = 3.073\times10^{15}\,\text{Hz}$ (red). The experimental data corresponds to the last T-ramp in fig. 5.30. The two upper graphs show the VRH plot (left) and the MEC plot (right) of the results, together with the respective linear fits according to eqs. 2.54 and 2.33, while the lower graphs contain the fit residuals. Besides the $T_0^{1/4}$ displayed in the top left graph, the VRH fits also yield $\log_{10}(\sigma_0/(\text{S m}^{-1})) = 5.08 \pm 0.04$ and 5.19 ± 0.02 for the simulations and the measurements, respectively.

The localization radius and the attempt frequency obtained through this adjustment process are the very first estimates of these quantities for a-GaO_x. Therefore, they were used as default values for the KMC simulations shown in this work – also for the statistics in the previous section. But are these adjusted values even realistic? In order to get an idea of which values of α are reasonable, the localized DFT state at the Fermi level previously shown in fig. 5.36d was fitted with $|\phi_i|^2 \propto \exp(-2|\vec{r} - \vec{r}_i|/\alpha)$, which yielded $\alpha = 0.106\,\text{nm}$ (fig. B.18). This is of course just a random example, but it indicates that the adjusted localization radius of $\alpha = 0.1512\,\text{nm}$ is not unrealistic. The adjusted attempt frequency $\nu_0 = 3.073\times10^{15}\,\text{Hz}$ is quite high in comparison to the usual Debye frequencies of crystalline solids, but, like the DOS, the vibrational density of states (vDOS) of amorphous materials differs from crystals. As mentioned in sec. 2.2.3, the high-frequency region in the vDOS of amorphous solids typically contains localized vibrations, which are called locons.[70, 71,

223] In contrast to the phonons in periodic lattices, the locons are non-propagating and contribute to heat transport only by locon hopping – just like the hopping of the electrons leads to charge transport.[70, 224, 225] This limited relevance of the locons is probably the reason why they have not been studied extensively in the past. It is possible, however, that they play an important role for the VRH conduction of the electrons. The locons might behave more like vibrating molecules, which could allow higher frequencies than usual for periodic lattice vibrations. The localized, high-frequency movement of these atomic clusters then would trigger the frequent hop attempts of the electrons.

A last point to address with regard to the adjusted α and ν_0 is their dependence on the DOS supplied to the KMC simulations. The used DFT-DOS is just an approximation and the comparison with the literature (cf. fig. 2.2) demonstrates that the absolute value of $g(\varepsilon)$ in the narrow relevant range around the Fermi level may differ depending on the ab-initio method/settings employed. Such variations of the DOS would also change the α and ν_0 that reproduce the experiments (e.g., approximately $\alpha \propto g(\mu)^{1/3}$, cf. eq. 2.54), which means that one should not overinterpret the adjusted α and ν_0 as exact material parameters of a-$GaO_{1.0}$. The general reproducibility most probably remains unaffected, though. For example, the KMC simulations with the quite different DOS of a-$GaO_{0.8}$, which will be shown in the next section, still yield the same trends as observed for a-$GaO_{1.0}$. It thus appears plausible that the experimental conductivities can always be reproduced by adjusting α and ν_0 in the KMC-VRH simulations, provided that there is still a significant, rather slowly varying DOS at the Fermi level. Therefore, even under the assumption that the DFT-DOS are inaccurate, the KMC results clearly suggest that the electron conduction in a-GaO_x occurs via variable-range hopping.

The KMC simulations not only yield the conductivity $\sigma = en_{\mathrm{eff}}u$, but also the effective charge carrier density n_{eff} and the mobility u. The T-dependence of the two latter quantities is displayed in fig. 5.50. It visualizes a major difference between mobility edge conduction and variable-range hopping. While for MEC the T-dependence of σ mainly originates from a change of the charge carrier density with T (cf. sec. 2.2.2), for VRH the variation of the charge carrier mobility governs the T-dependence of the conductivity. The effective charge carrier density increases approximately linear with the temperature in fig. 5.50a, just like expected from the relation $n_{\mathrm{eff}} = k_{\mathrm{B}}Tg(\mu)$ for constant or linear DOS (sec. A.2). In contrast, the mobility in fig. 5.50b increases exponentially according to

$$u = u_0 \exp\left[-\left(\frac{T_0}{T}\right)^{1/4}\right] \tag{5.2}$$

as expected from eq. 2.52. In terms of absolute value, the situation is reversed. Fig. 5.50a shows that the effective charge carrier density of a-$GaO_{1.0}$ is many orders of magnitude higher than the approx. 10^{14} cm^{-3} reported for MEC in a-GaO_x with $x \gg 1.0$ at RT.[123, 155, 156] Conversely, the VRH mobility in fig. 5.50a is several orders of magnitude lower than the MEC mobilities of 1.5 to

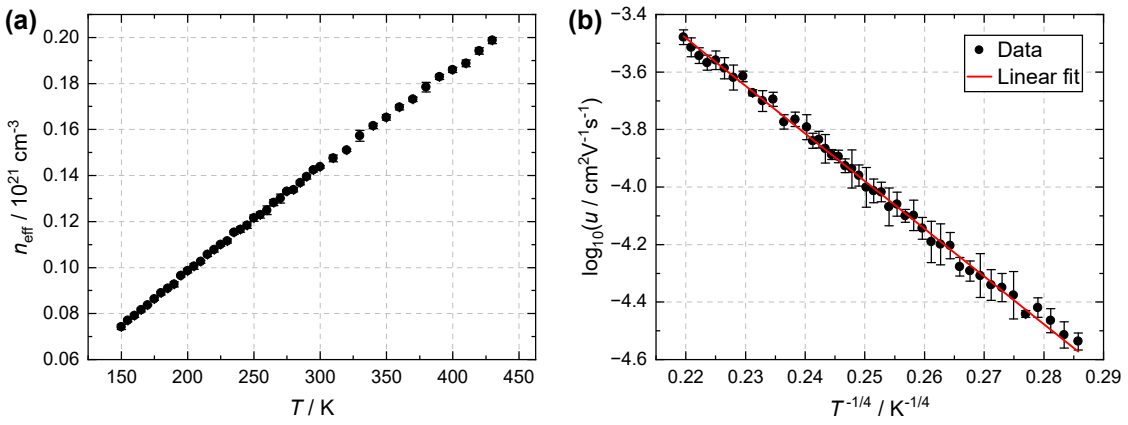


Figure 5.50: Temperature-dependence of the effective charge carrier density n_{eff} and the mobility u, as obtained through KMC-VRH simulations for a-$GaO_{1.0}$ with the adjusted $\alpha = 0.1512\,\mathrm{nm}$ and $\nu_0 = 3.073{\times}10^{15}\,\mathrm{Hz}$ (same simulations as in fig. 5.49). The linear fit of the mobility by eq. 5.2 yields $\log_{10}(u_0/(\mathrm{cm^2V^{-1}s^{-1}})) = 0.16 \pm 0.04$ and $T_0^{1/4}/\mathrm{K}^{1/4} = 38.2 \pm 0.4$, with the latter differing from the T_0 obtained for σ in fig. 5.49 due to the T-dependence of n_{eff}.

33.6 $\mathrm{cm^2/(Vs)}$ that have been published for a-GaO_x with $x \gg 1.0$ at RT.[123, 155, 156] As expected, the electrons are more mobile in the extended states above the CB mobility edge, while the charge carrier density is higher for VRH because the relevant electronic states are closer to the Fermi level.

5.5.3 Influence of the Stoichiometry

Like the DOS of a-$GaO_{1.0}$, also the DFT-DOS of a-$GaO_{0.8}$ and a-$GaO_{1.2}$ feature localized states at the Fermi level or at least rather close to it (cf. fig. 5.37). Therefore, the KMC simulations were also carried out for these other O/Ga ratios, with the respective ab-initio DOS as input. The localization radius and the attempt frequency adjusted for a-$GaO_{1.0}$ in the previous section were used also for these other compositions, in order to investigate the pure influence of the different DOS.

Fig. 5.51 shows the ε-related distributions of the localized states, the electrons and the non-oscillating hops from the KMC simulations for a-$GaO_{0.8}$ and a-$GaO_{1.2}$. Those for $x = 0.8$ in fig. 5.51a are quite comparable to those for $x = 1.0$ in fig. 5.38a and 5.39b, except that the localized states around the Fermi level have a peak-shaped distribution for a-$GaO_{0.8}$ instead of the monotonous increase for a-$GaO_{1.0}$. The histograms for $x = 1.2$ in fig. 5.51b, however, differ completely because the Fermi level is located in a 0.3–0.4 eV wide band gap. There is just a single electron above the band gap and only one unoccupied state below it during most of the simulation time for $x = 1.2$. The reason for this is obviously the $\exp\left[-\varepsilon_{ij}/(k_{\mathrm{B}}T)\right]$ term in the Miller-Abrahams equation that makes hops from below to above the band gap rather rare (typically approx. 0.01% of the non-osc. hops at $T = 300\,\mathrm{K}$).[114] Congruently, the effective charge carriers are not visible in fig. 5.51b due to the low $n_{\mathrm{eff}} = 3.50{\times}10^{16}\,\mathrm{cm^{-3}}$ for a-$GaO_{1.2}$ in comparison to $1.44{\times}10^{20}\,\mathrm{cm^{-3}}$ for a-$GaO_{1.0}$ and

[114] Since these transitions across the band gap still occur, it is not always the same electron that is moving rather freely via the unoccupied states above the band gap.

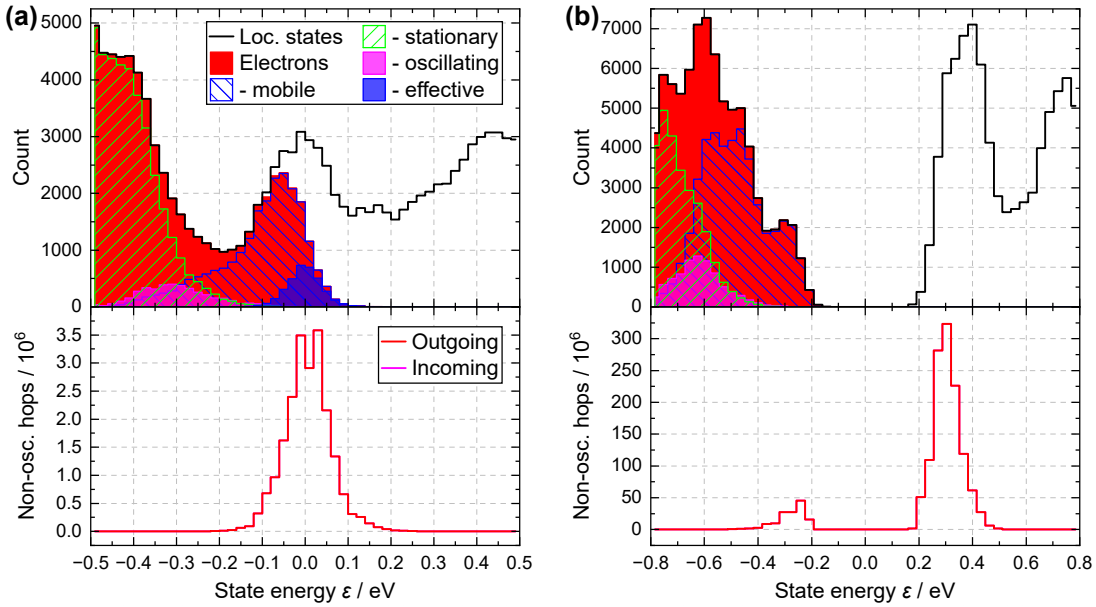


Figure 5.51: Histograms from the KMC-VRH simulations for (a) a-$GaO_{0.8}$ and (b) a-$GaO_{1.2}$ at T = 300 K (bin size: 20 meV for x = 0.8, 32 meV for x = 1.2; V_{cell}: 1.68×10^{-17} cm^3 for x = 0.8, 2.61×10^{-17} cm^3 for x = 1.2). Each top graph displays the distributions of the localized states and the different electron categories at the end of the simulation (like fig. 5.38a for x = 1.0), while the bottom graphs show the non-oscillating hops outgoing from or incoming to a certain state energy (like fig. 5.39b for x = 1.0). The incoming hops are concealed because they are almost identical to the outgoing hops. The N_{eff}-distribution is not visible in (b) because the band gap leads to a low number of effective charge carriers for a-$GaO_{1.2}$. Note that most technical parameters are different for x = 1.2 than for the other two O/Ga ratios (e.g., the energy range and the number of hops), as specified in sec. 4.2.2.

2.07×10^{20} cm^{-3} for a-$GaO_{0.8}$ (at T = 300 K). This already indicates a much lower VRH conductivity for x = 1.2 than for the other two O/Ga ratios.

The majority of both the oscillating[115] and the non-oscillating hops in a-$GaO_{1.2}$ starts and ends above the band gap (typically about 70% of the osc. hops and more than 90% of the non-osc. hops, fig. 5.51b), simply because there are more localized states closely above the band gap than below it. Likewise, most of the electrons' displacement is acquired through hops above the band gap (over 90%), which is why the hopping transport in a-$GaO_{1.2}$ could be called "n-type VRH".

Another characteristic of the KMC simulations for a-$GaO_{1.2}$ is the even higher proportion of oscillating hops (e.g., 97.8% for fig. 5.51b). Due to the low number of available paths in each KMC cycle, there is a high probability to get stuck in configurations that favor oscillations. Since the oscillating hops usually are fast with regard to the simulated time span, this does not necessarily imply a low conductivity, but it severely increases the number of hops required to obtain valid simulation results. Without the oscillation-related optimizations of the KMC algorithm described in sec. 4.2.6, it would not have been possible to simulate the variable-range hopping in a-$GaO_{1.2}$ with the available computational resources.

[115] The oscillating hops are not shown in fig. 5.51 but their distributions have essentially the same shape as the histograms of the non-osc. hops.

Like for a-$GaO_{1.0}$, the VRH conductivity σ of a-$GaO_{0.8}$ and a-$GaO_{1.2}$ was simulated at different temperatures. The results, which are displayed in fig. 5.52 together with the already discussed data for $x = 1.0$, agree quite nicely with the expectations and the experiments. The conductivity decreases and the fitted T_0 increases non-linearly with increasing oxygen content, i.e., with higher x, as observed in the van-der-Pauw measurements (sec. 5.3.3) and in the literature (fig. 2.3). This

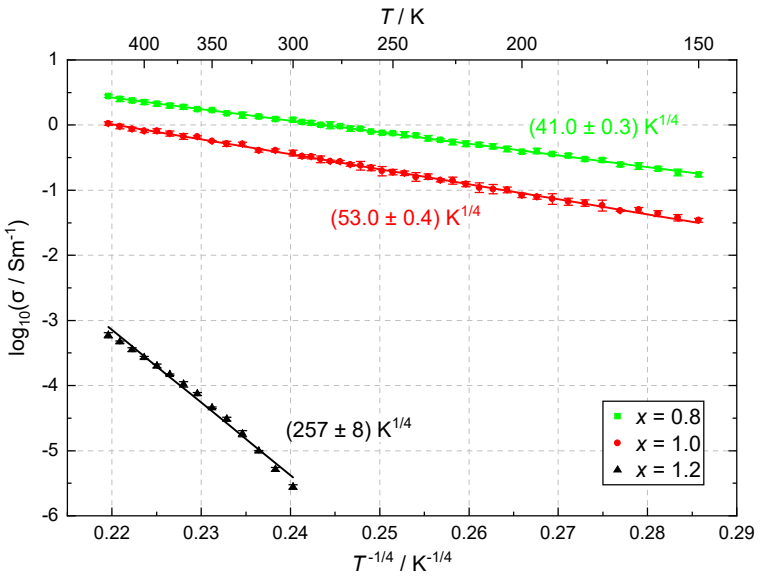


Figure 5.52: Influence of the O/Ga ratio x on the VRH conductivity, as obtained through KMC simulations with the DFT-DOS for $x = 0.8$, 1.0 and 1.2. The simulations were all carried out with the same localization radius $\alpha = 0.1512$ nm and attempt frequency $\nu_0 = 3.073 \times 10^{15}$ Hz, in order to investigate solely the effect of the differing DOS. The lines are the linear fits according to the VRH law (eq. 2.54) and the nearby shown values correspond to the resulting $T_0^{1/4}$.

confirms that variable-range hopping indeed leads to the experimentally observed trends with regard to the stoichiometry. Furthermore, the KMC simulations also elucidate the underlying reasons for this stoichiometry-dependence. Both the effective charge carrier density n_{eff} and the mobility u decrease significantly from $x = 0.8$ to 1.0 and further to 1.2, as can be seen in fig. 5.53. These changes of n_{eff} and u together yield the variation of the conductivity over several orders of magnitude. The decrease of n_{eff} directly follows from the decreasing DOS around the Fermi level according to eq. 4.18. Simultaneously, lower DOS in the energy range with the most non-oscillating hops also causes larger average distances of the involved localized states, which decreases the hopping probabilities (eq. 2.46) and thus leads to a lower mobility of the electrons. This latter effect of the differing DOS is also responsible for the increase of T_0. According to eq. 2.54, T_0 is approximately antiproportional to the magnitude of the DOS in the relevant energy range. Therefore, both the mobility and the conductivity become stronger T-dependent for higher x (figs. 5.52 and 5.53; cf. eq. 2.52).

Concerning the exact form of the T-dependence, the simulated conductivity of a-$GaO_{0.8}$ accurately follows Mott's classical VRH law (eq. 2.54), just like observed for a-$GaO_{1.0}$. As mentioned above, this

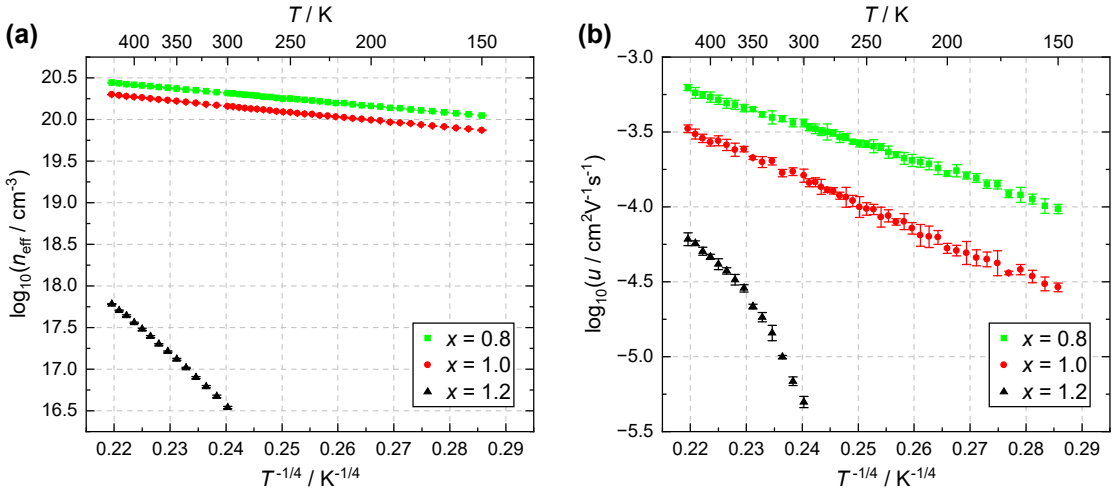


Figure 5.53: Influence of the O/Ga ratio x on (a) the effective charge carrier density n_{eff} and (b) the mobility u, as obtained through the KMC simulations for x = 0.8, 1.0 and 1.2. Both quantities are displayed analog to the VRH law by Mott (eq. 2.54), in order to visualize their contribution to the overall T-dependence of the conductivity $\sigma = e n_{\text{eff}} u$ in fig. 5.52 (same simulations).

demonstrates that the VRH law remains valid as long as the DOS exhibits no drastic variations in the relevant energy range. The limits to the validity of the VRH law are indicated by the KMC results for a-$GaO_{1.2}$, which deviate visibly from the $T^{-1/4}$-dependence.[116] Even though the T-dependence of n_{eff} for a-$GaO_{1.2}$ also differs in comparison to the other two O/Ga ratios (fig. 5.53a), the deviation of σ from the VRH law mainly originates from the mobility. This can be seen in fig. 5.53b, where $\ln(u)$ is a concave function of $T^{-1/4}$ – just like $\ln(\sigma)$ in fig. 5.52. The reason for this is that the hopping mainly occurs in the steep CB tail of a-$GaO_{1.2}$ (cf. fig. 5.51b), which essentially leads to a strong interdependence of ε_{ij} and r_{ij} (high $\varepsilon_{ij} > 0 \rightarrow$ low r_{ij}). Therefore, as expected from the VRH theory (e.g., for exponential or parabolic DOS, cf. sec. 2.2.3), the KMC simulations no longer follow the classical VRH law by Mott when the DOS varies severely near the Fermi level.

5.5.4 Hopping at High Electric Field Strength

The KMC results discussed so far agree well with the VRH theory and the conductivity measurements. Especially the quantitative agreement reached for a-$GaO_{1.0}$ justifies to move on to predictive simulations. The first phenomenon to be investigated with the established KMC-VRH model for a-$GaO_{1.0}$ is the electric field-dependence of the VRH conductivity, which is difficult to assess experimentally without risking sample damage or influences from the contacts. Nevertheless, the relationship between σ and F could be highly relevant for a-GaO_x devices, because strong local electric fields may occur for example in ultra-thin films (e.g., 1 V/10 nm = 1 MV/cm) or in Schottky barrier regions (cf. sec. 2.2.4).

[116] The other popular VRH models mentioned in sec. 2.2.3 – with an exponential dependence on $T^{-1/2}$ (parabolic DOS) or T^{-1} (exponential DOS) – are also not appropriate for a-$GaO_{1.2}$.

The electric field strength F influences the KMC-VRH simulations solely through the $ex_{ij}F$ term in the Miller-Abrahams equation (eq. 4.15). At high F, this term not only makes hops in $+x$-direction more likely, it also enables hops with larger state energy differences $\varepsilon_{ij} > 0$. This leads to a broadening of the steady-state electron distribution in comparison to the weak-field Fermi-Dirac distribution $f_{\mathrm{FD}}(\varepsilon, \mu, T)$, as can be seen in fig. 5.54. The fit of the broadened occupation histogram in

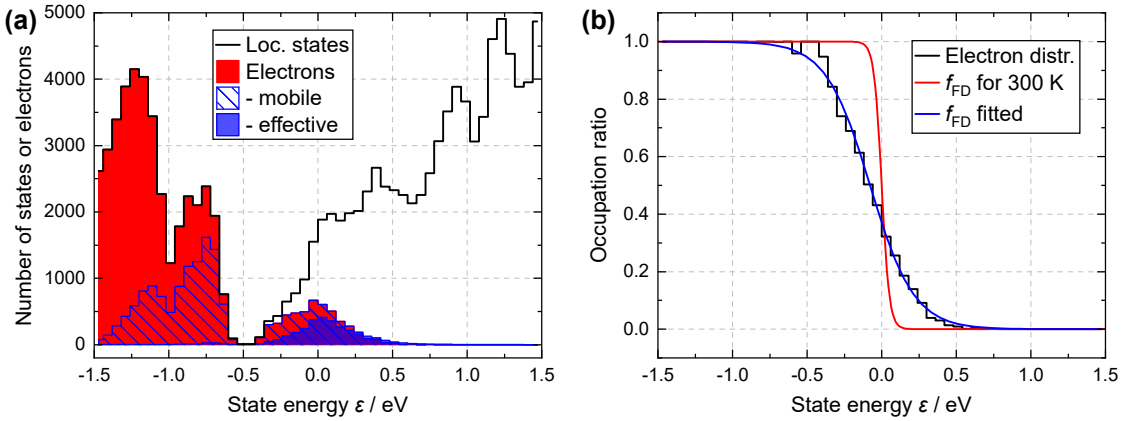


Figure 5.54: Broadened electron distribution in a KMC simulation for a-$GaO_{1.0}$ at high electric field strength $F = 5.56\,\mathrm{MV/cm}$ (9.52 V applied over a cell size of 17.1 nm; $T = 300\,\mathrm{K}$; DOS energy range: $[-1.5, +1.5]$ eV). (a) Histograms of the localized states and the electrons as a function of the state energy ε at the end of the simulation (bin size: 60 meV). They did not change during the main simulation (as monitored by periodic Fermi-Dirac fits), which means they represent the steady state. (b) Histogram of the occupation ratio ($= N/N_{\mathrm{st}}$), together with the Fermi-Dirac function $f_{\mathrm{FD}}(\varepsilon)$ for $T = 300\,\mathrm{K}$ and a fitted Fermi-Dirac function (eq. 4.17). The f_{FD}-fit yielded an effective temperature of $T_{\mathrm{eff}} = 1709\,\mathrm{K}$ and an effective chemical potential of $\mu_{\mathrm{eff}} = -0.078\,\mathrm{eV}$.

fig. 5.54b reveals that the electrons still follow a Fermi-Dirac distribution, albeit with a significantly higher effective temperature $T_{\mathrm{eff}} \gg T$ and a slightly different effective chemical potential $\mu_{\mathrm{eff}} \neq \mu$. The electron distribution at high electric field strength corresponds to $f_{\mathrm{FD}}(\varepsilon, \mu_{\mathrm{eff}}, T_{\mathrm{eff}})$ instead of $f_{\mathrm{FD}}(\varepsilon, \mu, T)$, which is well-known as "effective temperature model".[63, 64, 205] The name "effective temperature model" is somewhat misleading, though, because the electrons really have a higher temperature than the material, i.e., the phonon bath (cf. sec. 2.2.3).[65, 66, 226] Likewise, μ_{eff} is the actual chemical potential of the electrons, not the weak-field value μ that was supplied as input to the simulations. The slight shift of μ_{eff} in comparison to μ is a mandatory consequence of the increasing electron temperature T_{eff} at constant electron density n, analog to the usual T-dependence of the Fermi level.

Fig. 5.55a shows how the effective temperature varies with the electric field strength. In the literature, the relation between T_{eff} and F has been described phenomenologically as

$$T_{\mathrm{eff}}(T, F) = \left[T^{\beta} + \left(\gamma \frac{e\alpha F}{k_{\mathrm{B}}} \right)^{\beta} \right]^{1/\beta} \tag{5.3}$$

with the coefficients $\beta > 1$ and $\gamma > 0$.[63, 64, 205] Fitting the obtained $T_{\mathrm{eff}}(F)$ data with eq. 5.3

failed, however, because this model is restricted to a convex curve shape and a linear asymptote for high F ($T_{\mathrm{eff}} \propto F$). Only the region of low to medium field strength can be approximated by eq. 5.3 (fig. 5.55a). In order to describe the full range of $T_{\mathrm{eff}}(F)$, the model was refined to allow a power-law asymptote:

$$T_{\mathrm{eff}}(T,F) = \left[T^{\beta} + \left(\gamma \frac{e\alpha F}{k_{\mathrm{B}}} \right)^{(\beta-\omega)} \right]^{1/\beta} \tag{5.4}$$

The additional coefficient ω must fulfill the condition $\omega < \beta - 1$. Eq. 5.4 retains the necessary asymptotic behavior of $T_{\mathrm{eff}} \to T$ and $\partial T_{\mathrm{eff}}/\partial F \to 0$ for $F \to 0$. This refined model describes the $T_{\mathrm{eff}}(F)$ data very well, e.g., for the purpose of inter- or extrapolation. The respective fit in fig. 5.55a yields $T_{\mathrm{eff}} \propto F^{(\beta-\omega)/\beta} = F^{0.60\pm0.06}$ as asymptote in the high field regime. The effective chemical potential μ_{eff} does not vary much in the explored F-range (fig. B.19a), by less than 0.1 eV, and it is identical to the respective T-adjusted Fermi level, of course (fig. B.19b, cf. sec. 4.2.5).

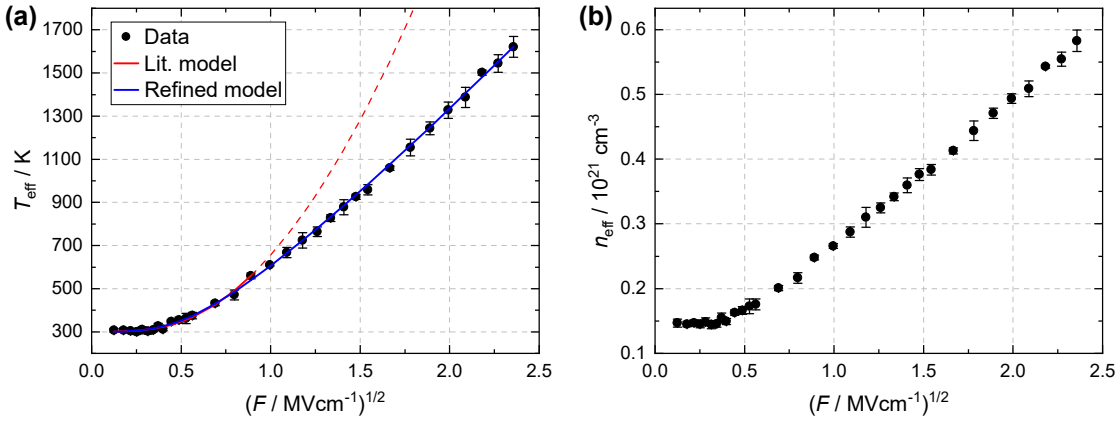


Figure 5.55: Electric field-dependence of the effective temperature T_{eff} and the effective charge carrier density n_{eff} for a-$GaO_{1.0}$ at $T = 300$ K. (a) $T_{\mathrm{eff}}(F)$ with fits according to the literature model (eq. 5.3, fitted up to 0.8 MV/cm and extrapolated as dashed line) and according to the refined model (eq. 5.4, fitted over the full range of F). The latter yielded the coefficients $\gamma = (231 \pm 51)\ \mathrm{K}^{\omega/(\beta-\omega)}$, $\beta = 2.93 \pm 0.28$ and $\omega = 1.17 \pm 0.14$, resulting in an exponent of $(\beta - \omega)/\beta = 0.60 \pm 0.06$ for the power-law asymptote at high F. (b) Increase of the effective charge carrier density n_{eff} at high electric field strength due to the elevated temperature T_{eff} of the electrons.

Since the derivation of the effective charge carrier density is based on the actual energetic distribution of the electrons (sec. 2.1.1), n_{eff} must be calculated using T_{eff} and μ_{eff} instead of T and μ in eq. 2.23 as soon as F exceeds the weak-field regime. Therefore, the histogram of the effective charge carriers in fig. 5.54a is broadened like the electron distribution. An increasing electric field thus not only leads to a higher drift mobility u, but also to a higher effective charge carrier density n_{eff}, even though the total electron density n remains constant. Despite the quite high T_{eff}, n_{eff} still approximately equals $k_{\mathrm{B}} T_{\mathrm{eff}} g(\mu_{\mathrm{eff}})$ (cf. sec. A.2), which explains why the F-dependencies of T_{eff} in fig. 5.55a and of n_{eff} in fig. 5.55b resemble each other closely.

The two major effects of a strong electric field – more hops in $+x$-direction and higher n_{eff} – together lead to an exponential increase of the VRH conductivity σ at high F, as can be seen in fig. 5.56a. The further analysis in fig. 5.56b shows that σ remains approximately constant up to 0.24 MV/cm, which corresponds to the weak-field regime where the conductivity is independent of F. This is another

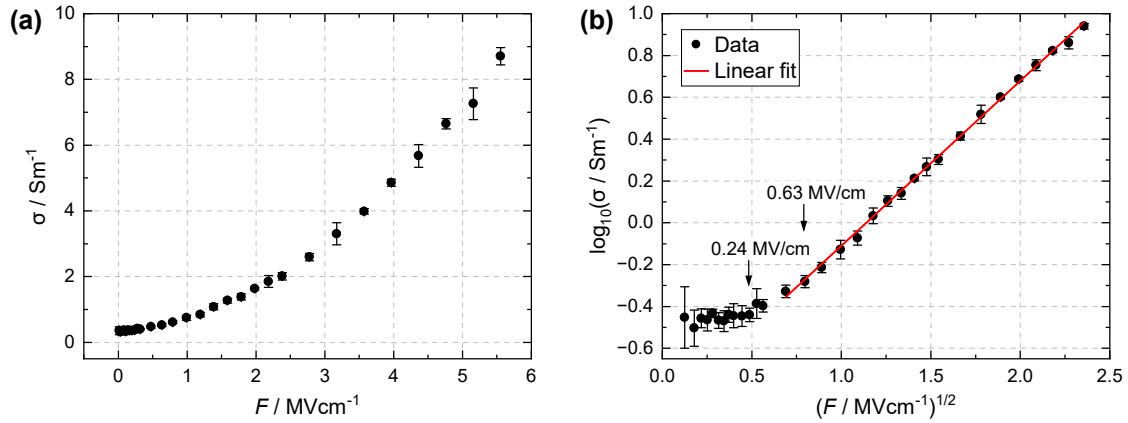


Figure 5.56: Electric field-dependence of the KMC-VRH conductivity of a-$GaO_{1.0}$ at $T = 300$ K, plotted in two different ways. (a) The conductivity σ increases exponentially beyond the weak-field regime. (b) The F-dependence in the high-field regime ($F \geq 0.63$ MV/cm) can be described by a linear relation between $\ln(\sigma)$ and $F^{1/2}$. The weak-field region where $\partial\sigma/\partial F = 0$ is clearly visible as plateau ($F \leq 0.24$ MV/cm).

proof that all previously discussed KMC simulations with the default $F = 0.0635$ MV/cm were carried out under weak-field conditions. In the high-field regime above 0.63 MV/cm, σ increases exponentially according to $\ln(\sigma) \propto F^{1/2}$. Comparison with the analog graphs for the effective charge carrier density and the mobility in fig. B.20 reveals that this characteristic $F^{1/2}$-dependence in the high-F region mainly originates from the mobility $u(F)$, while $n_{\mathrm{eff}}(F)$ dominates in the medium-F region (0.24 MV/cm < F < 0.63 MV/cm). Furthermore, these graphs demonstrate that both aforementioned effects contribute significantly to the exponential increase of the conductivity.

The exponential $F^{1/2}$-dependence of the conductivity obtained here for a-$GaO_{1.0}$ is one of the variants proposed in the VRH literature. The same relation between σ and F has been derived analytically in [227] and numerically in [228], with $g(\varepsilon) = \text{const.}$ assumed in both cases. Like the $T^{-1/4}$-dependency, this again underlines the applicability of such simplified models to real-world materials without a constant DOS. Another analogy to the T-dependence is the multitude of other models that have been proposed for the F-dependence of VRH, depending on the field strength, the DOS or other approximations (e.g., [31, 229, 230]). Without the KMC results presented here it would be difficult to identify which of these model is appropriate for highly oxygen-deficient a-GaO_x.

Yet another model can be evaluated using the $\sigma(F)$ data for a-$GaO_{1.0}$. Originally, the effective temperature model has been developed with the aim to re-express the F-dependence of the conductivity by a temperature-dependence $\sigma(T_{\mathrm{eff}}) \propto \exp[-(T_0/T_{\mathrm{eff}})^{1/4}]$ analog to the classical VRH law for $\sigma(T)$.[54] This is definitely not possible for a-$GaO_{1.0}$, because $\ln(\sigma)$ vs. $T_{\mathrm{eff}}^{-1/4}$ is clearly non-linear

in the high-field regime (fig. B.21). Instead, the aforementioned trends of $T_{\text{eff}} \propto F^{0.60\pm0.06}$ and $\ln(\sigma) \propto F^{1/2}$ suggest that approximately $\ln(\sigma) \propto T_{\text{eff}}$ (for T = const.). The reason why this aspect of the effective temperature model does not work is rather obvious: Even at high F, the temperature in the Miller-Abrahams equation is still the phonon temperature T, not the electron temperature (cf. sec. 2.2.3), which makes a one-to-one replacement of T by T_{eff} rather impossible.

As mentioned at the beginning of this section, it is well possible that strong electric fields occur in a-GaO_x thin films, which includes those in resistive switching devices. In fact, the exponential F-dependence of σ in fig. 5.56a bears a striking resemblance to the I-V curves observed in the resistive switching experiments (cf. sec. 5.2). When a thin a-GaO_x layer dominates the device resistance, then the change of the VRH conductivity with the local electric field strength could be responsible for the shape of the switching curves. This could be the much needed, alternative explanation for the non-linearity of the I-V curves, since the classical Schottky barrier models are invalid when variable-range hopping in the depletion zone becomes possible (cf. sec. 2.2.4). The FEM simulations discussed later in sec. 5.6 will pick up on this idea. For now, it shall be noted that the experimentally observed relation $\ln(I/V) \propto V^{1/2}$ in the resistive switching experiments (cf. fig. 5.26) is consistent with the exponential $F^{1/2}$-dependence of σ, provided that the major voltage drop occurs over a thin region of the a-GaO_x film with rather homogeneous electric field strength. This indicates that VRH might play an important role in the resistive switching of a-GaO_x, as it has recently been proposed for TaO_x and $SrTiO_3$.[231, 232]

5.5.5 Variation of the Chemical Potential

The next property to be investigated with the established KMC-VRH model for a-$GaO_{1.0}$ is the chemical potential-dependence of the VRH conductivity. The simulations discussed so far were carried out for charge-neutral a-GaO_x, with a relative chemical potential of $\mu = 0$ (cf. sec. 4.2.2). A variation of μ shifts the Fermi level relative to $\varepsilon = 0$, which either increases ($\mu > 0$) or decreases ($\mu < 0$) the number of electrons in the KMC cell (cf. eq. 4.16). This allows to simulate the hopping process in negatively or positively charged a-GaO_x, respectively. Determining the μ-dependency of σ experimentally would be more challenging, since it would require measurements on a homogeneously charged sample. Nevertheless, it is an important property, because a non-zero charge density is a common case in electronic devices (e.g., Schottky barrier depletion zone, cf. sec. 2.2.4).

In principle, varying μ just changes the region of the DOS where the variable-range hopping occurs. An interesting case is $\mu = -0.5\,\text{eV}$, which shifts the Fermi level to the tiny band gap in the DOS of a-$GaO_{1.0}$ (cf. fig. 5.35). Fig. 5.57 shows the distributions of the electrons and the non-oscillating hops for this case. Due to the band gap around E_F, the effective charge carrier density is strongly reduced to $n_{\text{eff}} = 2.74\times10^{18}\,\text{cm}^{-3}$ (fig. 5.57a; $n_{\text{eff}} = 1.44\times10^{20}\,\text{cm}^{-3}$ for $\mu = 0$), just like for $x = 1.2$ (fig. 5.51b). In contrast to $x = 1.2$, however, the density of localized states is higher below the gap than above (fig. 5.57a), which is why the hopping transport mainly takes place below the Fermi

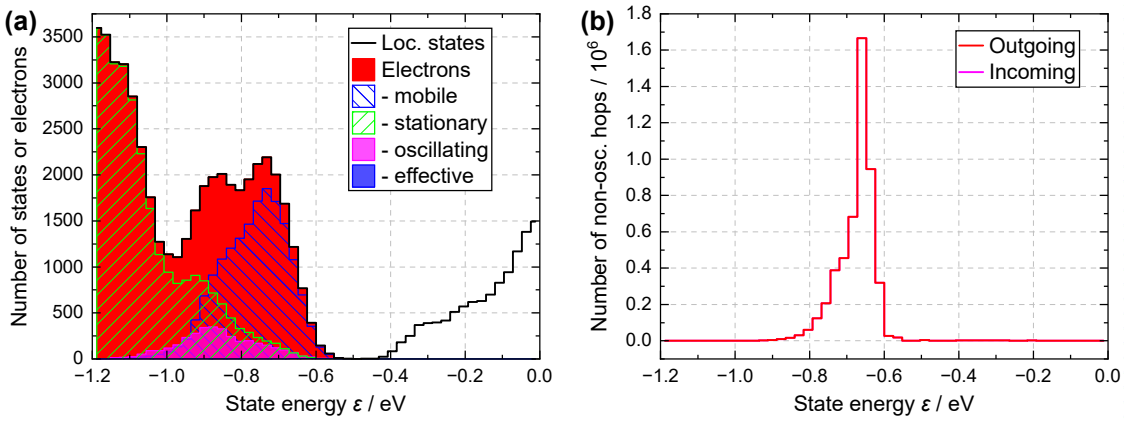


Figure 5.57: KMC histograms for positively charged a-$GaO_{1.0}$ with $\mu = -0.5\,\text{eV}$ ($\rho = -161\,\text{C/cm}^3$, $T = 300\,\text{K}$, DOS energy range: $[-1.2, +0.0]$ eV, bin size: 24 meV, $V_{\text{cell}} = 2.21{\times}10^{-17}\,\text{cm}^3$). (a) Distributions of the localized states and the different electron categories, analog to fig. 5.38a. The electrons above $\varepsilon = -0.5\,\text{eV}$ and the effective charge carriers are not visible because their number is very low ($n_{\text{eff}} = 2.74{\times}10^{18}\,\text{cm}^{-3}$). (b) Distribution of the non-oscillating hops during the simulation, as a function of the energy ε of the initial state (= outgoing) or the final state (= incoming) of the hop. The graph for the incoming hops is not visible because it is virtually equal to the graph of the outgoing hops.

level ("p-type VRH"; fig. 5.57b). The mobility is reduced by about one order of magnitude from $(1.6 \pm 0.2) \times 10^{-4}\,\text{cm}^2/\text{V s}$ for $\mu = 0$ to $(2.3 \pm 0.5) \times 10^{-5}\,\text{cm}^2/\text{V s}$ for $\mu = -0.5\,\text{eV}$. Together with the lower n_{eff}, this leads to a severe drop of the conductivity from $(0.37 \pm 0.04)\,\text{S/m}$ for $\mu = 0$ to $(1.0 \pm 0.2) \times 10^{-3}\,\text{S/m}$ for $\mu = -0.5\,\text{eV}$.

Interestingly, a rather similar electric field-dependence of the conductivity is observed for the positively charged a-$GaO_{1.0}$ with $\mu = -0.5\,\text{eV}$ (fig. 5.58), despite the significantly different DOS around the shifted Fermi level. In the high-field regime, the conductivity still depends exponentially on $F^{1/2}$ (cf. sec. 5.5.4). The slope of the respective fit in fig. 5.58 is higher than for $\mu = 0$, though (approx. doubled; cf. fig. 5.56b), which means that the conductivity increases more strongly for $\mu = -0.5\,\text{eV}$. This is plausible, since the broadening of the electron and hop distributions at high field strength has more impact when the Fermi level is at a local minimum of the DOS.

The influence of the chemical potential was studied in a broad range from $\mu = -2.5\,\text{eV}$ to $+0.5\,\text{eV}$, which is approximately the range where states are localized according to the previous LOS analysis (fig. 5.36). Fig. 5.59 shows that the conductivity approximately follows the DOS, as expected from the VRH theory in sec. 2.2.3. The higher the DOS near the shifted Fermi level, the higher is the conductivity. The minimum of σ does not coincide with the minimum of the DOS, of course, because not only the DOS exactly at the Fermi level but the whole DOS close to it influences the conductivity. At weak electric field strength, the conductivity resembles the DOS rather closely (fig. 5.59), because the relevant energy range around the Fermi level is narrow. Beyond the weak-field regime, it must be respected that μ_{eff} is the actual chemical potential of the electrons, not μ (cf. sec. 5.5.4). With increasing F, $\sigma(\mu_{\text{eff}})$ becomes successively higher and smoother due to the broadening of the relevant energy range (fig. 5.59).

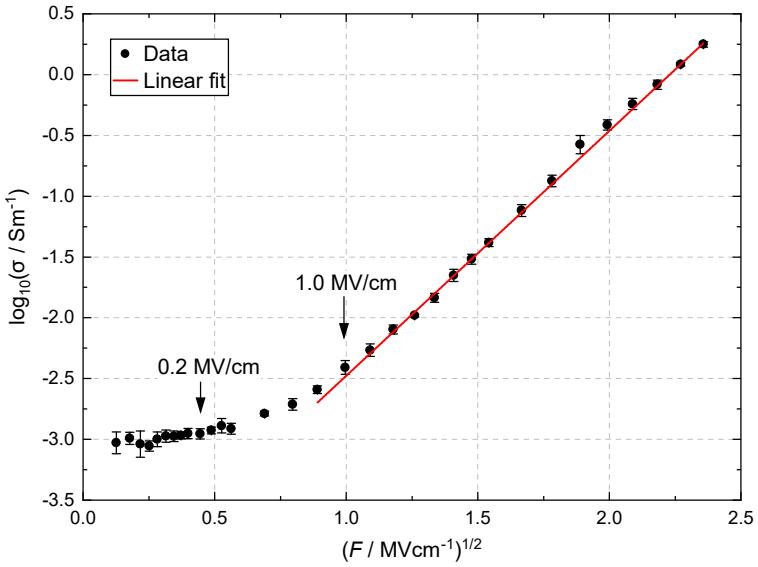


Figure 5.58: Electric field-dependence of the KMC-VRH conductivity of positively charged a-$GaO_{1.0}$ with $\mu = -0.5\,\text{eV}$ ($T = 300\,\text{K}$). The arrows indicate the approximate end of the weak-field regime and the beginning of the high-field regime, respectively.

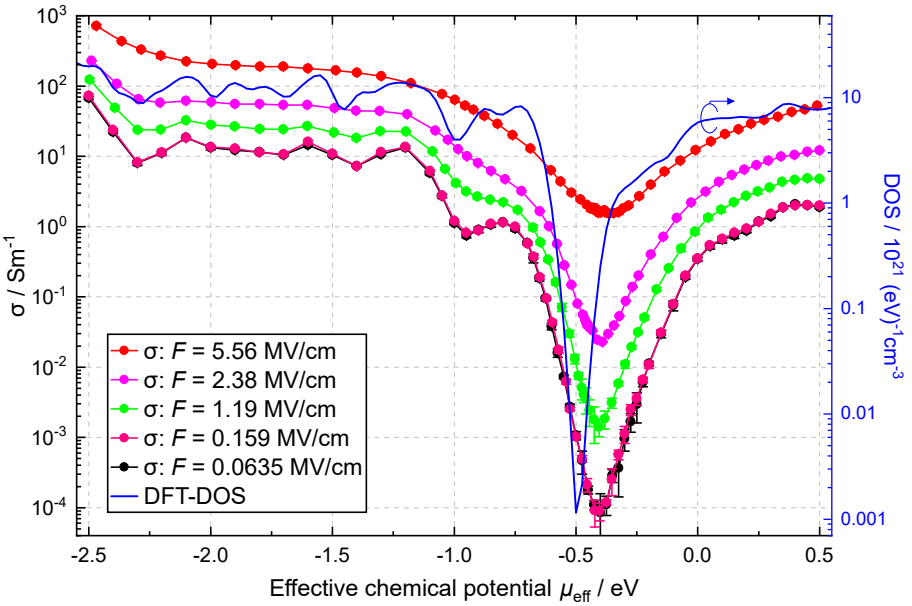


Figure 5.59: Chemical potential-dependence of the KMC-VRH conductivity of a-$GaO_{1.0}$ for different electric field strengths (left axis; $T = 300\,\text{K}$), in comparison to the DFT-DOS $g(\varepsilon{=}\mu_{\text{eff}})$ of a-$GaO_{1.0}$ (right axis). The data is plotted against μ_{eff} instead of the varied input parameter μ, because μ_{eff} is the actual chemical potential of the electrons relative to $\varepsilon = 0$ (cf. sec. 5.5.4). The finite DOS around $\mu_{\text{eff}} = -0.5\,\text{eV}$ results from the Gaussian smearing by VASP (cf. sec. 4.1.2). The DFT calculation with 16×16×8 k-points did not yield any Kohn-Sham states in the range from $\varepsilon = -0.63\,\text{eV}$ to $-0.38\,\text{eV}$, which indicates a true band gap.

The KMC-VRH simulations for a-$GaO_{1.0}$ with varied chemical potential and varied electric field strength ultimately yielded a fine-grained map of σ in relation to μ_{eff} and F, which is displayed in fig. 5.60. For all investigated chemical potentials, an approximately linear relation between $\ln(\sigma)$ and $F^{1/2}$ is observed at high F. Since these simulations cover a quite wide range of different DOS shapes in the vicinity of E_{F}, this relation is probably universal for high-field VRH in a-GaO_x (e.g., for other O/Ga ratios).

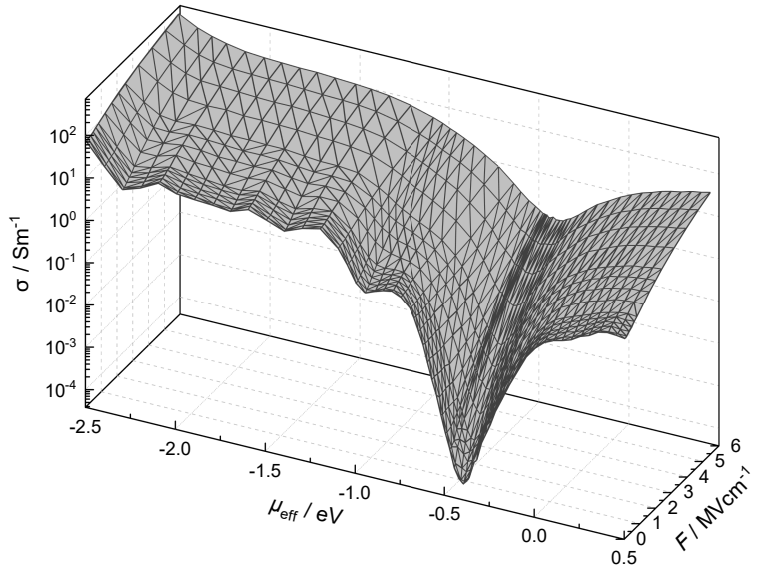


Figure 5.60: 3D plot of the KMC-VRH conductivity of a-$GaO_{1.0}$ at $T = 300\,\text{K}$, as a function of the electric field strength F and the effective chemical potential μ_{eff} of the electrons. Each intersection of the grid corresponds to an individual KMC simulation with 5 repetitions (3990 KMC runs in total). The connecting lines are just a guide for the eye.

From the data in figs. 5.59 and 5.60, two rivaling effects can be expected for hopping transport in a Schottky barrier depletion zone. While the electron depletion leads to a decrease of σ approximately for $-0.7\,\text{eV} < \mu_{\text{eff}} < 0\,\text{eV}$, the high electric field strength in the barrier region causes an increase of σ. The later FEM simulations in sec. 5.6 will elucidate – based on the data of fig. 5.60 – how these two trends influence the overall electron transport.

5.5.6 Electron Diffusion

The VRH conductivity discussed in the previous sections was calculated from the drift flux of the electrons, utilizing the linear relation between the mean displacement $\langle X \rangle$ and the simulated time span τ for a certain electric field in the KMC cell. According to the basic transport theory in sec. 2.1 (eq. 2.7), the resulting σ not only characterizes the drift flux due to $\nabla\varphi$ but also any diffusion flux due to a chemical potential gradient $\nabla\mu$ (or $\nabla\mu_{\text{eff}}$ at high F). But what about the Fickian diffusion

driven by an electron density gradient ∇n (cf. eq. 2.12)? One route to obtain the Fickian diffusion coefficient D is the generalized Nernst-Einstein relation that was introduced together with the effective charge carrier model in sec. 2.1.1 (eq. 2.29). After the findings of sec. 5.5.4, another generalization is necessary, though. To account for a potentially elevated electron temperature T_{eff} beyond the weak-field regime, the Nernst-Einstein relation should be written as

$$D = \frac{u k_{\mathrm{B}} T_{\mathrm{eff}}}{e} \quad \Leftrightarrow \quad \frac{\sigma}{e^2} = \frac{n_{\mathrm{eff}} D}{k_{\mathrm{B}} T_{\mathrm{eff}}} \tag{5.5}$$

with T replaced by T_{eff}, because the temperature term originates from the Fermi-Dirac distribution of the electrons (essentially eq. 2.22). For weak F, the original form is recovered due to $T_{\mathrm{eff}} = T$. Eq. 5.5 would allow to calculate D from σ, i.e., from $\langle X \rangle$ vs. τ. However, the KMC simulations also provide an alternative, more direct route to obtain the Fickian diffusion coefficient, even though the KMC cell has no electron density gradient ($\nabla n = 0$). As explained thoroughly in sec. A.4, the n_{eff}-model enables the calculation of D from the mean squared displacements $\langle Y^2 \rangle$ and $\langle Z^2 \rangle$ perpendicular to the electric field (eq. 4.24).[117] The values resulting from this second approach will be discussed in the following. Furthermore, by calculating the Nernst-Einstein ratio $eD/uk_{\mathrm{B}}T_{\mathrm{eff}}$ – with u calculated from the x-displacements and D obtained from the y- and z-displacements – the validity of eq. 5.5 for VRH will be assessed.

Fig. 5.61a shows the temperature-dependence of D for a-$GaO_{1.0}$, as obtained simultaneously to the weak-field conductivity already discussed in sec. 5.5.2. The diffusion coefficient clearly exhibits an analog T-dependence as σ, which can be described by

$$D = D_0 \exp\left[-\left(\frac{T_0}{T}\right)^{1/4}\right] \tag{5.6}$$

This becomes understandable when regarding the Nernst-Einstein ratio, which is also displayed in fig. 5.61a. The simulations yield $eD/uk_{\mathrm{B}}T \approx 1$, which means that the electrons' behavior actually conforms with the prediction by the Nernst-Einstein relation. The same is also observed in the simulations for a-$GaO_{0.8}$ and a-$GaO_{1.2}$ (fig. B.23). Together with the previously discussed, approximate relation of $n_{\mathrm{eff}} \propto T$ (fig. 5.50a), eq. 5.5 directly leads to the similar T-dependence for σ and D. Accordingly, the fit of $D(T)$ by eq. 5.6 yields $T_0^{1/4} = (49.8 \pm 0.3)\,\mathrm{K}^{1/4}$, which is close to the value of $T_0^{1/4} = (53.0 \pm 0.4)\,\mathrm{K}^{1/4}$ for the conductivity (cf. fig. 5.49). A slight difference remains, though, because the proportionality between n_{eff} and T is not exact (cf. fig. 5.50a).

In fact, it would have been a serious surprise if the weak-field simulations would have disagreed with the Nernst-Einstein relation, since its derivation is not tied to a specific conduction mechanism (cf. sec. 2.1.1). The original formulation with T for the temperature term is also common in the VRH

[117] In contrast, the Fickian diffusion coefficient D_x cannot be reliably obtained from the x-displacement variance parallel to the electric field, as discussed already in sec. 5.5.1 (non-linear $\mathrm{Var}(X)$ vs. τ). For the default number of hops ($N_{\mathrm{h,sim}} = 3{\times}10^8$), the calculated D_x is still rather close to its true value, i.e., to D_{yz}, but there is already a systematic offset (fig. B.22). The higher $N_{\mathrm{h,sim}}$ and the higher F, the stronger deviates $\mathrm{Var}(X)$ vs. τ from linearity and the more inaccurate and non-convergent becomes this approach for obtaining D_x.

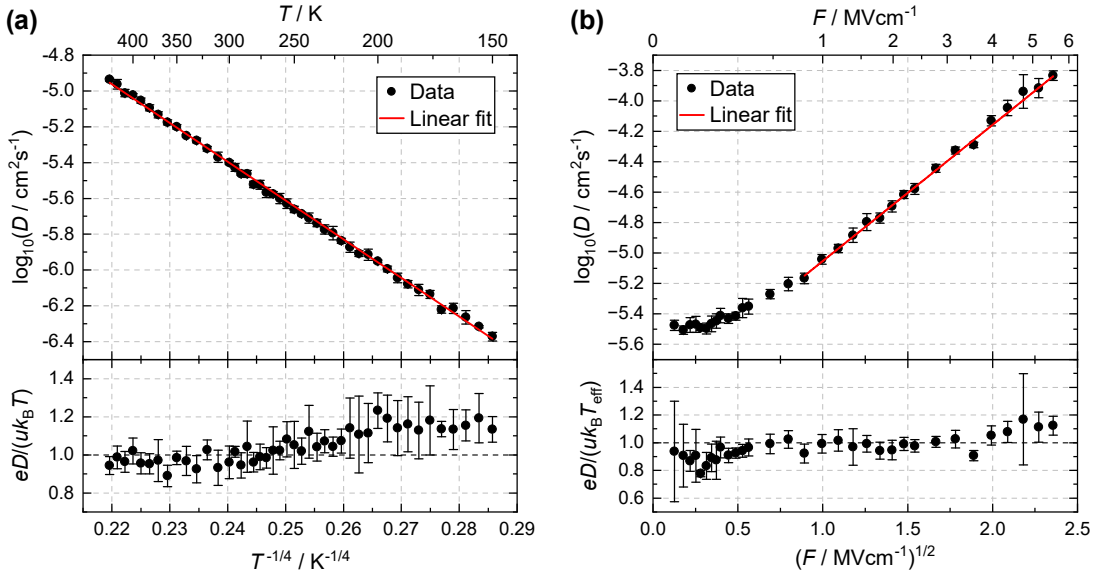


Figure 5.61: Temperature- and electric field-dependence of the Fickian diffusion coefficient D (top) and the corresponding Nernst-Einstein ratio (bottom) for VRH in a-$GaO_{1.0}$. (a) T-dependence of D, from the same simulations as the conductivity data in fig. 5.49. The linear fit with eq. 5.6 yields $T_0^{1/4} = (49.8 \pm 0.3)\,\mathrm{K}^{1/4}$. The minor deviations of $eD/uk_\mathrm{B}T$ from unity at low temperatures are not a real effect but result from slightly insufficient equilibration for D. (b) F-dependence of D at $T = 300\,\mathrm{K}$, from the same simulations as the conductivity data in fig. 5.56.

literature (e.g., [233–235]). Less obvious is the validity of eq. 5.5 at high electric field strength. Both the replacement of T by T_eff and the coupling between u and D at high F are not established knowledge. In the previous discussion of the high-field effects in sec. 5.5.4, it was observed that still approximately $n_\mathrm{eff} \propto T_\mathrm{eff}$ at high electric field strength (cf. fig. 5.55). In combination, eq. 5.5 predicts a similar F-dependence for D as observed for σ. This is exactly what is obtained in the KMC simulations, as shown in fig. 5.61b. In the high-field regime, D increases exponentially with $F^{1/2}$ and the Nernst-Einstein ratio remains equal to unity. The Nernst-Einstein relation in the generalized form stated above is indeed valid beyond the weak-field regime. But why does a strong electric field along the x-axis lead to an exponential increase of $D = D_{yz}$, i.e., of the diffusion in the yz-plane? The increase of n_eff due to the elevated electron temperature cannot be the cause, as D is a per-particle property. The actual reason is the amorphicity of the system. Almost every hopping path $\vec{r}_{ij}$ has a non-zero x_{ij}-component and is thus affected by the $ex_{ij}F$ term in the Miller-Abrahams equation. Therefore, the electron movements parallel and perpendicular to $\vec{F}$ are not independent. Since the state positions $\vec{r}_i$ are uniformly random, the enhancement of the hopping rates is symmetric with regard to $y = 0$ and $z = 0$ (apart from finite size effects, cf. fig. 5.47).

The Fickian diffusion coefficients obtained for varied chemical potential μ, which are shown in fig. 5.62, also resemble the μ-dependence observed for σ (cf. fig. 5.59). Congruently, the Nernst-Einstein ratio is close to unity (fig. 5.62), though it exhibits systematic deviations that demand a closer look: The primary aim of the large number of simulations with varied μ was the investigation of the conductivity. Therefore, they all were carried out with the rather low default hop counts.

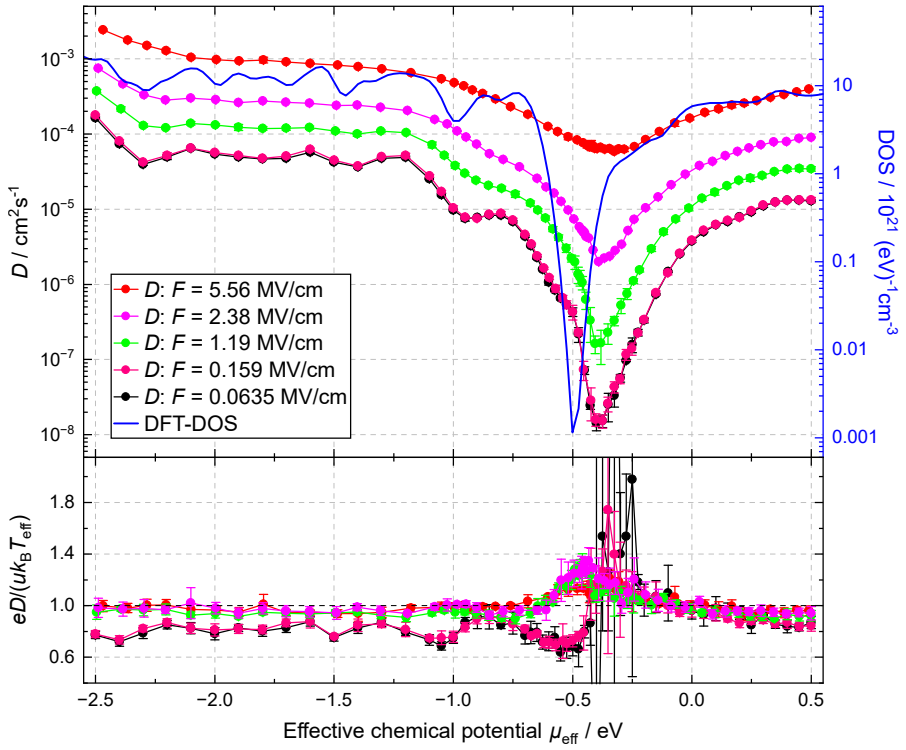


Figure 5.62: Chemical potential-dependence of the Fickian diffusion coefficient D (top; left axis) and the corresponding Nernst-Einstein ratio eD/uk_BT_{eff} (bottom) for VRH in a-$GaO_{1.0}$ at different electric field strengths (T = 300 K; same simulations as in fig. 5.59).

These are completely sufficient for σ, since $\langle X \rangle$ vs. τ converges quickly and does not require an equilibration (cf. sec. 5.5.1). For D, however, the default hop settings were less adequate in two cases: Shifting the Fermi level to a region with especially high DOS leads to an increased number of electrons in the KMC cell, which would require a longer equilibration at least for low electric field strength. Likewise, the diffusion statistics are impaired when the Fermi level is in or near the band gap (around $\mu = -0.5\,\text{eV}$). In this case only few electrons can move simultaneously (cf. fig. 5.57a) and it takes longer until enough electrons have acquired a reasonable displacement to produce bell-shaped Y_k- and Z_k-distributions (irregardless of the electric field strength).[118] The resulting inaccuracies for D are irrelevant on a log-scale, but they show up in the Nernst-Einstein ratio as artificial deviations. It is fair to assume that without them, the Nernst-Einstein ratio would equal unity for all μ. This implicates that eq. 5.5 is valid irregardless of the shape of the DOS.

All in all, the diffusion results confirm that the Nernst-Einstein relation (eq. 5.5) is valid for VRH, even beyond the weak-field regime. This is not only of theoretical interest but also has practical consequences. Future studies could determine D from simulated or even from measured conductivities. In the context of KMC-VRH simulations, this would have the benefit of better accuracy, because – as demonstrated by the discussion above – it is typically more difficult to reach exact

[118] These effects can also be seen when varying the number of states N_{st} in the simulations: While σ is independent of N_{st} due to sufficient hop counts, D shows a minor systematic variation (e.g., for $\mu = -0.5\,\text{eV}$ in fig. B.24).

convergence for $\langle Y^2 \rangle$ and $\langle Z^2 \rangle$ than for $\langle X \rangle$ (cf. fig. 5.44). Moreover, eq. 5.5 may be relevant for other modeling approaches. The FEM simulations in this work, which will be discussed in the next section, use the chemical potential as one of the unknowns to be determined and thus do not need Fickian diffusion coefficients as input. Future FEM studies, however, might choose to use the electron density instead, which would require the Nernst-Einstein relation as connection between drift and diffusion flux.

5.6 Modeling of Pt|a-$GaO_{1.0}$|ITO Devices

The DFT and KMC calculations not only brought a new understanding of the bulk electron transport in a-GaO_x, but also laid the foundation for a theoretical treatment of the metal|a-GaO_x interfaces. These cannot be described by the classical Schottky barrier models of sec. 2.2.4, because the rather high density of localized states in the mobility gap enables electron hopping through the barrier region. Since no model for VRH through a Schottky barrier was found in the literature, it was far from obvious how this affects the I-V characteristics. Therefore, the VRH transport in Pt|a-$GaO_{1.0}$|ITO devices was simulated with the FEM method, using the DFT and KMC results as input. The DFT-DOS provided the relation between the density n and the chemical potential μ of the electrons, which is essential for describing the electron depletion zone with the Poisson equation (eq. 4.28). The KMC simulations yielded the VRH conductivity σ as a function of the chemical potential μ, which allows to obtain the electron flux through the space charge zone from the continuity equation (eq. 4.35). KMC also revealed that the electron temperature and the conductivity depend on the electric field strength F. Accordingly, the FEM calculations were carried out in two variants: Initially, the F-dependency of n and σ was neglected. After it became clear that high electric field strengths occur in the simulations, they were improved by taking the F-dependency into account. In the following, mainly the results of the latter will be shown and discussed, while the corresponding data without high-F effects can be found in the appendix.

The further ingredients of the FEM simulations were the work functions of the involved materials and the relative permittivity ϵ_r of a-$GaO_{1.0}$. The work functions were approximated from the UPS measurements and the literature (W_{GaO} = 3.7 eV, W_{Pt} = 5.6 eV, W_{ITO} = 4.4 eV, cf. sec. 5.1.4). Since the permittivity of a-$GaO_{1.0}$ is unknown, at first ϵ_r = 10 was used, which is the literature value for β-Ga_2O_3 and a-$GaO_{1.5}$ (cf. sec. 2.3.3). Additionally, it was tried to calculate the permittivity with DFT. These calculations only worked for the amorphous model structures of a-$GaO_{1.5}$ and a-$GaO_{1.2}$, as described in sec. 4.1.2. Density functional perturbation theory yielded ϵ_r = 3.5 for x = 1.5, while DFT with superimposed electric fields produced ϵ_r = 4.0 and 4.8 for x = 1.5 and 1.2, respectively. These values are too low in comparison to the aforementioned literature value, but the increase of ϵ_r from x = 1.5 to 1.2 agrees with the trend reported in [92]. The permittivity calculations for x = 1.0 did not converge completely, but they yielded at least the head of the microscopic static dielectric tensor with diagonal elements ranging from ϵ_r = 22 to 115 (independent-particle approximation

without Hartree and local field effects). This indicates that the permittivity may increase substantially for highly oxygen-deficient a-GaO_x, probably due to the localized states near the Fermi level. A similar effect of localized electrons on the permittivity has also been reported for other materials.[166–168] In consequence of these DFT calculations, the FEM simulations were also carried out for higher permittivities, i.e., for $\epsilon_r = 100$ and 1000.

5.6.1 Electron Equilibrium at Zero Voltage

At first, the distribution of the electrons within the a-$GaO_{1.0}$ layer for $V = 0$ was simulated. Fig. 5.63a shows the resulting chemical potential $\mu(z)$ from the Pt-TE interface at $z = 0$ to the ITO-BE interface at $z = 40$ nm. As expected due to the electron depletion, the chemical potential is decreased

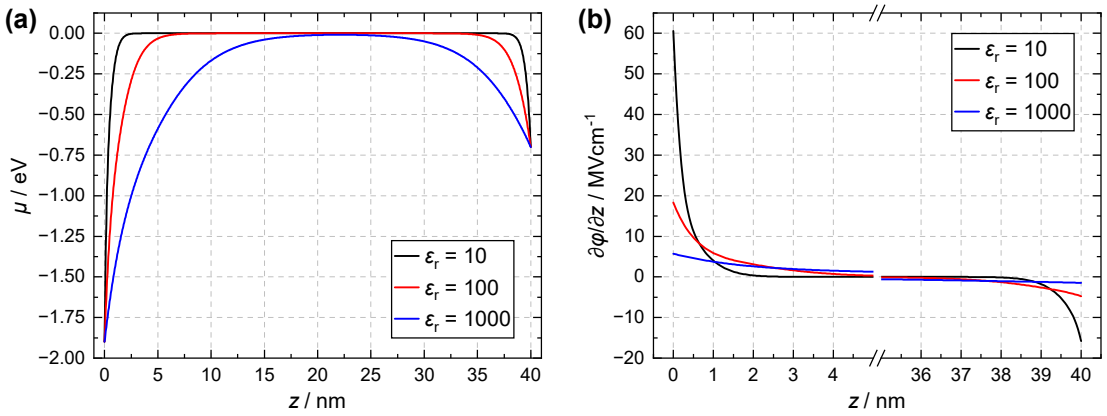


Figure 5.63: Electron equilibrium in Pt|a-$GaO_{1.0}$|ITO with $V = 0$, for different relative permittivities ϵ_r of the a-$GaO_{1.0}$ layer (FEM-2 model with high-F effects; cf. fig. B.25 for the respective FEM-1 results without high-F effects). (a) Chemical potential μ of the electrons, which corresponds to the electric potential according to $\mu = e\varphi$. (b) Electric potential gradient $\partial\varphi/\partial z$ ($= -\vec{F}$).

near both interfaces in comparison to the reference value $\mu = 0$ in the bulk. The Schottky barrier width increases significantly from $\epsilon_r = 10$ to 1000, in accordance with the Poisson equation (eq. 4.28). Correspondingly, the electric field strength in the barrier regions decreases with increasing permittivity, as shown in fig. 5.63b.

Even without an applied voltage, the electric field near the interfaces exceeds the weak-field regime by far (fig. 5.63b). This is why the aforementioned high-F effects were included in these simulations. However, for $V = 0$, there is almost no difference to the simulations where high-F effects were neglected (fig. B.25). Both the chemical potential $\mu(z) = e\varphi(z)$ and its gradient $\nabla\mu(z) = e\nabla\varphi(z)$ are essentially identical for the two model variants. The reason for this is the overall weak effect of the field strength on the electron density. While high F leads to a strong increase of the electron temperature T_{eff} (cf. sec. 5.5.4), this barely affects the electron density due to the symmetry of the Fermi-Dirac function (cf. eq. 4.29, fig. B.19).

Fig. 5.64 shows the charge density $\rho(z)$, i.e., the relative electron density, in the a-$GaO_{1.0}$ layer. The electron depletion near both interfaces is clearly visible. Its shape is quite far from the rectangular full depletion approximation that is normally used to approximate Schottky barriers. This obviously results from the more realistic $n(\mu)$-function in conjunction with the vast difference between the DOS of a-$GaO_{1.0}$ and the DOS of typical n-type semiconductors.

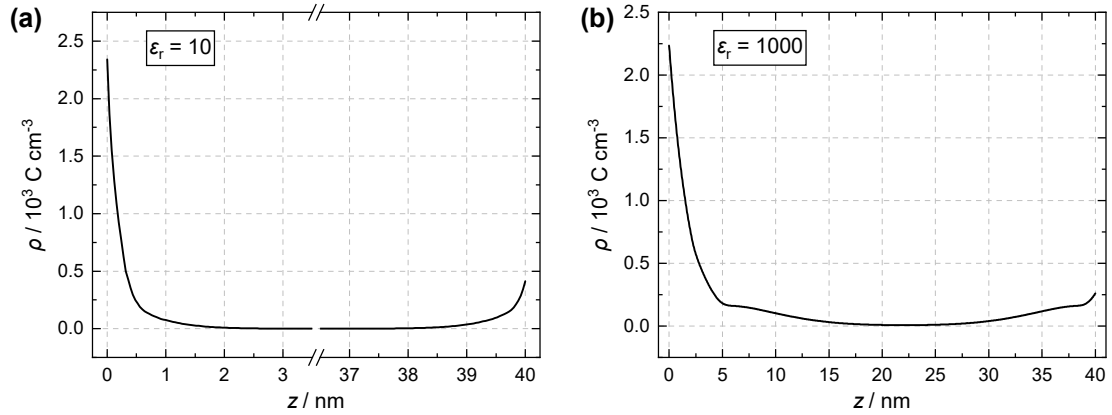


Figure 5.64: Charge density ρ in Pt|a-$GaO_{1.0}$|ITO with $V = 0$, either with (a) $\epsilon_r = 10$ or (b) $\epsilon_r = 1000$ (FEM-2 model with high-F effects).

Some consequences of the Schottky barriers for VRH transport through the device can already be conceived from the band bending near the interfaces. In the space charge zones, the state energies vary according to $E(z) = \varepsilon - e\varphi(z)$, as displayed in fig. 5.65. This causes the Fermi level E_F to cross the small band gap of a-$GaO_{1.0}$, which is located at $\varepsilon \approx -0.5\,\text{eV}$. In consequence, there is a small region with $E_F(z) + e\varphi(z) \approx -0.5\,\text{eV}$ near both interfaces, where the VRH conductivity $\sigma(z)$ is very low due to the lack of localized states near E_F. These regions may dominate the device resistance, when the rest of the a-$GaO_{1.0}$ layer is sufficiently conductive. The comparison between fig. 5.65a for $\epsilon_r = 10$ and fig. 5.65b for $\epsilon_r = 1000$ demonstrates that the extent and the impact of these regions increases with increasing permittivity.

These results just describe the situation in an idealized, pristine Pt|a-$GaO_{1.0}$|ITO device with a spatially and timely constant O/Ga ratio. In a real device, the ions could move, especially when strong electric fields are present. The electric potential gradients in fig. 5.63b indicate that this would happen even without an applied voltage. After the device preparation, the oxygen ions would move away from the electrodes and/or the gallium ions would migrate towards them over time. Since the ion movement most probably requires high gradients of either the electric potential or the chemical potential of the ions (due to their low ionic conductivity, cf. sec. 2.3.4), it is unlikely that this leads to a homogeneous increase of the O/Ga ratio in the bulk. Such an equilibrium for the ions should not be reached on reasonable timescales.[119] Instead, the decrease of the O/Ga ratio x directly at the interfaces should be accompanied by an increase of x near the end of the depletion zones, where the

[119] Otherwise, any resistive switching states would be volatile – in contrast to the observations in the I-V experiments.

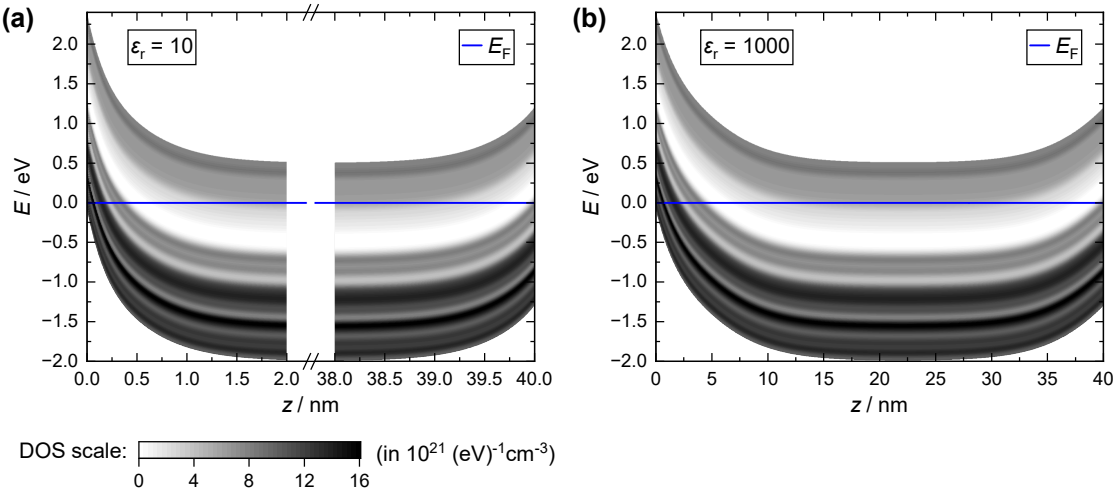


Figure 5.65: Band bending in Pt|a-$GaO_{1.0}$|ITO with $V = 0$, either with (a) $\epsilon_r = 10$ or (b) $\epsilon_r = 1000$ (FEM-2 model with high-F effects). The DOS $g(\varepsilon)$ for each energy $E = \varepsilon - e\varphi$ in the range of $\varepsilon \in [-2, +0.5]$ eV is shown as greyscale plot. The upper limit of the displayed ε-range approximately corresponds to the CB mobility edge, whereas the lower limit was chosen arbitrarily (since the VB mobility edge is even lower; cf. LOS analysis in fig. 5.36).

electric field is weak.[120] The full consequences for electron transport are difficult to predict, since these changes affect the local DOS, the work function, the charge density and the relations of n and σ to μ and F. However, there should still be the regions where the Fermi level crosses the band gap. The most important effect thus probably would be a broadening of the band gap in these regions due to the local increase of the O/Ga ratio over time. On the one hand, this would make it even more likely that these regions dominate the device resistance. On the other hand, this would also provide an additional explanation for the reproducibility issues of the I-V experiments, because the initial state of the devices would depend on the time since their preparation (and on eventual heat treatment, e.g., during lithography or paste drying).

5.6.2 Steady State for Non-Zero Voltages

In the next step, the steady-state electron transport through the Pt|a-$GaO_{1.0}$|ITO device was simulated for different applied voltages.[121] Fig. 5.66 shows the chemical potential $\mu(z)$, the electric potential $\varphi(z)$ and the Fermi level $E_F(z) = \tilde{\mu}(z) = \mu(z) - e\varphi(z)$ for $V = +1$ V and for $V = -1$ V. In fig. 5.66a,b, they were obtained with a permittivity of $\epsilon_r = 10$, which leads to narrow Schottky barriers and a linear variation of E_F through the whole a-$GaO_{1.0}$ layer. Almost the complete voltage drop[122] occurs over the bulk and the Schottky barriers are apparently irrelevant for the device re-

[120] The EELS results in sec. 5.1.3 do not clarify this hypothesis, since the surface was investigated without a TE (fig. 5.6) and since the a-GaO_x|ITO interface is too rough for conclusions (fig. B.4).

[121] V is the voltage at the Pt-TE, relative to the grounded ITO-BE – like in the I-V experiments.

[122] Despite its misleading name, the common term "voltage drop" does not refer to changes of the electric potential inside a device. Instead, it generally refers to variations of the electrons' electrochemical potential $\tilde{\mu}$, which sum up to the applied voltage times the electron charge.

sistance. For $\epsilon_r = 1000$ in fig. 5.66c,d, the situation is different. The Schottky barriers are broader, as expected from the previous results for $V = 0$, and the $E_F(z)$ curve exhibits two major voltage drops near the interfaces. For $V = +1\,V$, the BE barrier dominates with a voltage drop of $\Delta E_F \approx 0.63\,eV$, while the TE barrier has $\Delta E_F \approx 0.32\,eV$ (fig. 5.66c). For $V = -1\,V$, the TE barrier dominates with $\Delta E_F \approx 0.61\,eV$, whereas the BE barrier has $\Delta E_F \approx 0.32\,eV$ (fig. 5.66d). The bulk is almost negligible in both cases, with $\Delta E_F \approx 0.05\,eV$ and $0.07\,eV$, respectively (approx. $12\,nm < z < 30\,nm$). The reverse-biased interface dominates for each polarity, as would be expected from the classical Schottky barrier theory, but there is a remarkable difference: The voltage drops do not occur directly at the interfaces, where $E_F(z)$ is almost constant. Instead, they occur near the end of the depletion regions, as expected from the previous section.

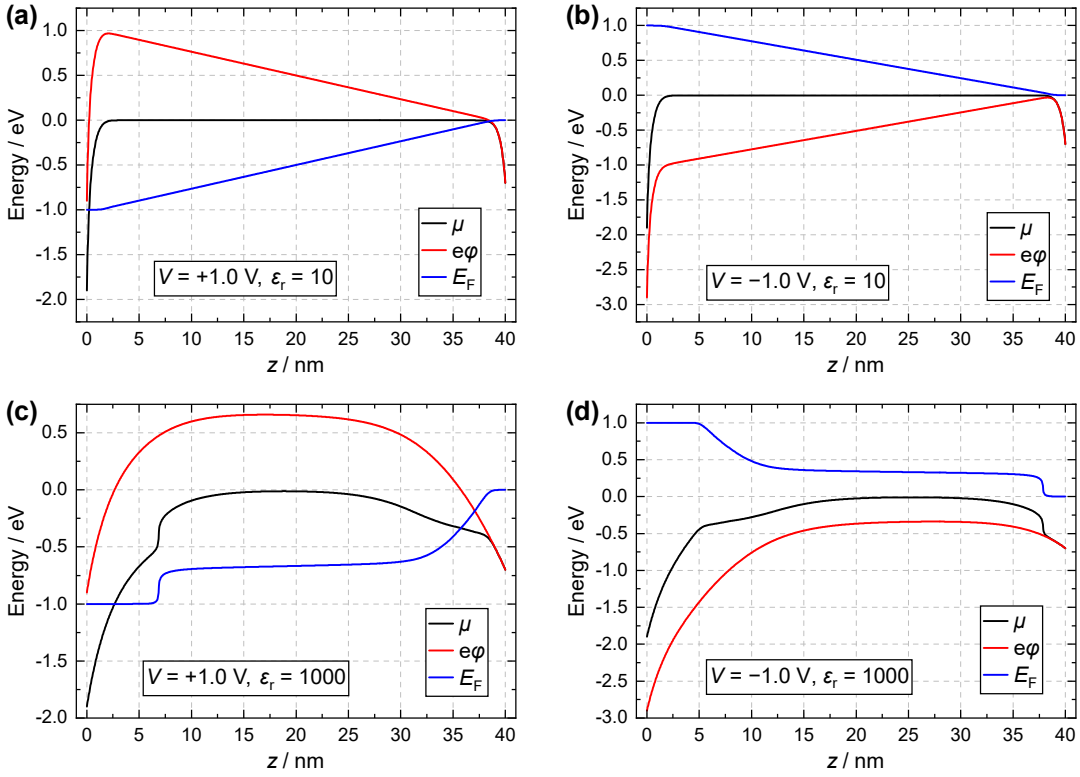


Figure 5.66: Chemical potential μ, electric potential φ and Fermi level E_F (= $\tilde{\mu}$) in Pt|a-$GaO_{1.0}$|ITO with an applied voltage of (a,c) $V = +1\,V$ or (b,d) $V = -1\,V$, for different permittivities of either (a,b) $\epsilon_r = 10$ or (c,d) $\epsilon_r = 1000$ (FEM-4 model with high-F effects; cf. fig. B.30 for the corresponding FEM-3 results without high-F effects).

The corresponding band bending graphs in fig. 5.67 illustrate the location of the voltage drops even better. For $\epsilon_r = 10$, the Fermi level varies almost exclusively over the bulk and it crosses the Schottky barriers as a straight line (fig. 5.67a,b). For $\epsilon_r = 1000$, however, the voltage drops occur precisely in the regions where the Fermi level crosses the small band gap of a-$GaO_{1.0}$ (fig. 5.67c,d). It is also visible why the dominating barrier changes with the voltage polarity for $\epsilon_r = 1000$. The dominating interface is the one where the course of $E_F(z)$ matches the curvature of the band bending $E(z)$, because this enlarges the region where the Fermi level is in or near the band gap.

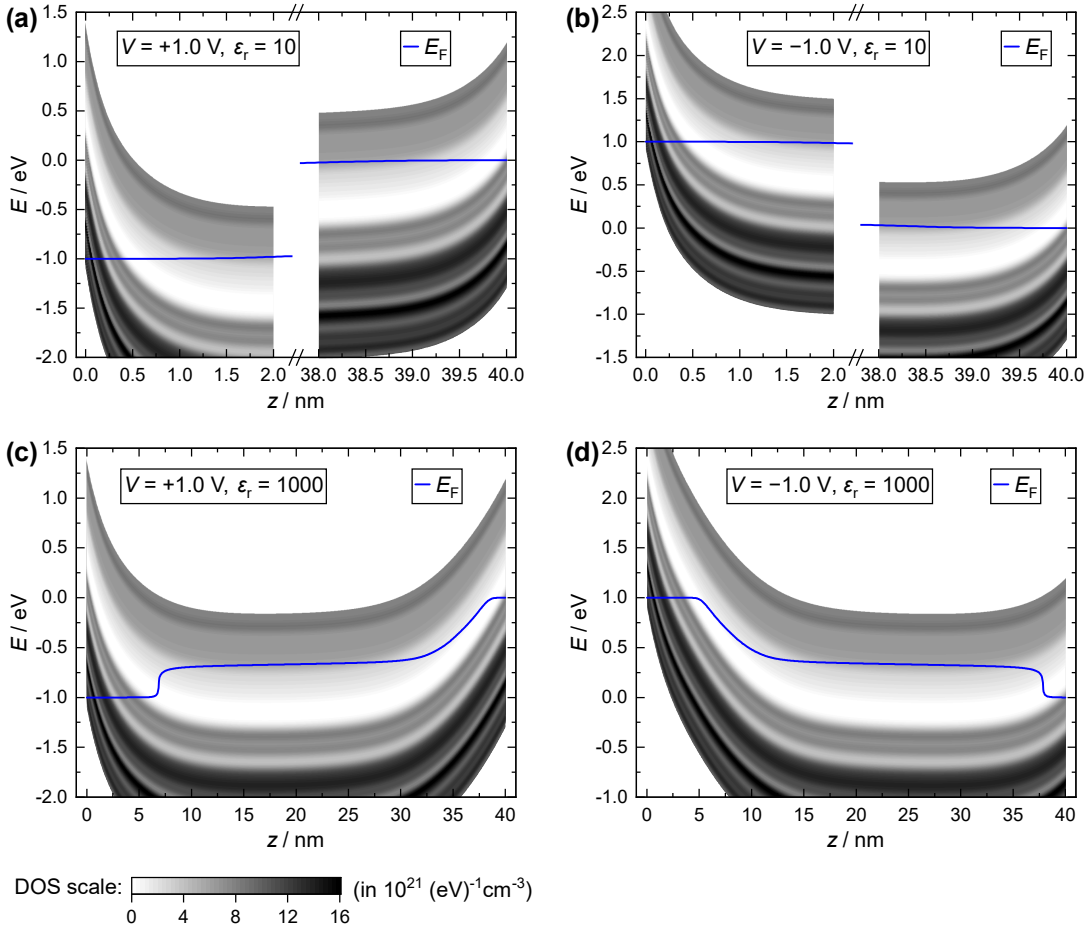


Figure 5.67: Band bending in Pt|a-$GaO_{1.0}$|ITO with an applied voltage of (a,c) $V = +1\,V$ or (b,d) $V = -1\,V$, for different permittivities of either (a,b) $\epsilon_r = 10$ or (c,d) $\epsilon_r = 1000$ (FEM-4 model with high-F effects). The DOS $g(\varepsilon)$ for each energy $E = \varepsilon - e\varphi$ in the range of $\varepsilon \in [-2, +0.5]$ eV is shown as greyscale plot. The upper limit of the displayed ε-range approximately corresponds to the CB mobility edge, whereas the lower limit was chosen arbitrarily (since the VB mobility edge is even lower; cf. LOS analysis in fig. 5.36).

The difference between the results for $\epsilon_r = 10$ and 1000 can be understood by regarding the local VRH conductivity $\sigma(z)$ in the a-$GaO_{1.0}$ layer (fig. 5.68). In contrast to the conclusions from the zero voltage simulations, the conductivity is not significantly reduced within the Schottky barrier regions for $\epsilon_r = 10$, not even where the Fermi level is in the band gap (fig. 5.68a). The reason for this is the high electric field strength $F(z)$ in these regions (fig. B.26a,b), which raises the F-dependent VRH conductivity in the band gap to approximately the same order of magnitude as the value in the bulk. This demonstrates how metal|a-GaO_x Schottky barriers can become irrelevant for the device resistance, despite a high work function difference. For $\epsilon_r = 1000$, the Schottky barriers are broader and therefore the reached field strengths are much lower (fig. B.26c,d). The conductivity $\sigma(z)$ for $\epsilon_r = 1000$ in fig. 5.68b thus exhibits the expected minima near both interfaces. They differ from the bulk value by 2–3 orders of magnitude. These are the regions where the voltage drops were observed in fig. 5.66c,d. The low-σ region is not deeper but broader at the dominating barrier, in accordance with the band bending graph (fig. 5.67c,d).

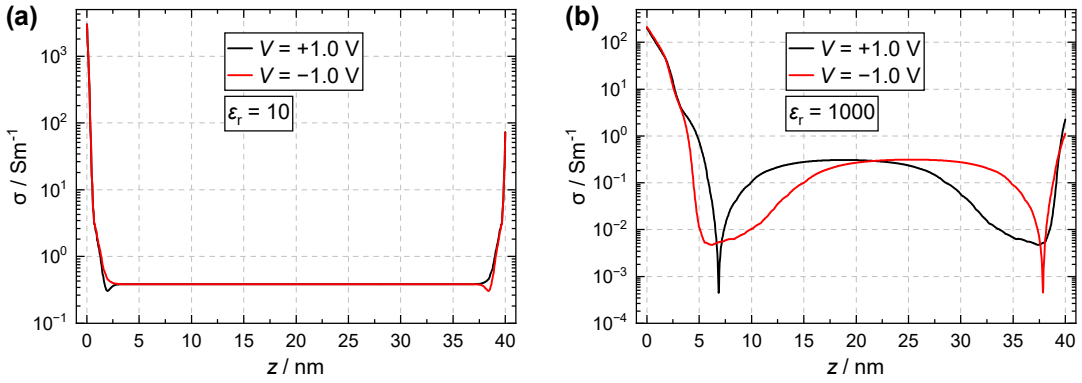


Figure 5.68: Hopping conductivity σ in Pt|a-$GaO_{1.0}$|ITO with an applied voltage of $V = +1\,\mathrm{V}$ or $-1\,\mathrm{V}$, for different permittivities of either (a) $\epsilon_r = 10$ or (b) $\epsilon_r = 1000$ (FEM-4 model with high-F effects).

In order to investigate the consequences for the I-V response of the Pt|a-$GaO_{1.0}$|ITO device, the FEM simulations were also carried out for the voltages between +1 V and −1 V, in 0.1 V steps. The resulting, smooth I-V curves for the different permittivities are shown in fig. 5.69a. The curve for $\epsilon_r = 10$ is almost linear and highly symmetric, as expected for a bulk-dominated device, but it also possesses a slight S-shape (convex for $V > 0$, concave for $V < 0$). This curvature originates directly from the F-dependency of the conductivity in the bulk, which increases gradually with increasing absolute voltage (fig. B.27). The simulated device with $d = 40\,\mathrm{nm}$ is thin enough such that the electric field in the bulk exceeds the weak-field regime (fig. B.26a,b).

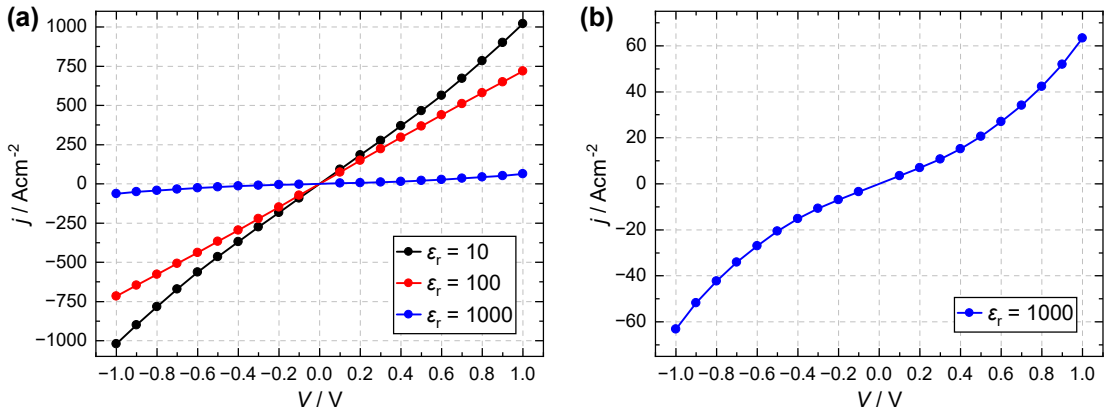


Figure 5.69: Current density j for different voltages V applied to Pt|a-$GaO_{1.0}$|ITO (FEM-4 model with high-F effects; cf. fig. B.32 for the corresponding FEM-3 results without high-F effects). The points represent the individual FEM simulations, which are connected by straight lines as guide for the eye. In (a), the results for the different relative permittivities $\epsilon_r = 10$, 100 and 1000 are displayed together, while (b) shows a close-up plot of the data for $\epsilon_r = 1000$.

The I-V curve for $\epsilon_r = 1000$ has much lower current density than the curve for $\epsilon_r = 10$ (fig. 5.69a), due to the reduced conductivity in the dominating near-interface regions (cf. fig. 5.68). Furthermore, the curve for $\epsilon_r = 1000$ is considerably more S-shaped, as can be seen in fig. 5.69b. The reason for this is again the F-dependency of the conductivity, but now the relevant effects occur in the low-

σ regions instead of the bulk. Fig. 5.70a shows that mainly the voltage drop near the BE increases for higher positive voltage, whereas the bulk and the TE barrier vary only moderately. Congruently, mostly the voltage drop near the TE increases for more negative voltage (fig. 5.70b). These changes within the respective, dominating Schottky barrier govern the shape of the I-V curve. The higher the voltage drop over the barrier, the broader is the region where the Fermi level is in the band gap (cf. fig. 5.67). This causes a broadening of the dominating conductivity minimum with increasing absolute voltage, as shown in fig. 5.70c,d. However, this does not lead to a higher device resistance, because the increasing voltage drop is also coupled to a local increase of the electric field strength (fig. 5.70e). The resulting increase of the conductivity overcompensates the broadening of the minimum (fig. 5.70c,d), such that the total device resistance decreases for higher absolute voltage. This produces the S-shape of the I-V curve. The V-dependent conductivity profiles in fig. 5.70c,d also explain the high symmetry of the I-V curve for $\epsilon_\mathrm{r} = 1000$. Normally, one would expect an asymmetric I-V curve for such a pair of back-to-back Schottky barriers with different electrode materials. Here, however, the decisive low-σ regions are congruent for $V > 0$ and $V < 0$, just inverted with regard to the spatial axis.

The I-V curve for the intermediate permittivity $\epsilon_\mathrm{r} = 100$ in fig. 5.69a is linear with a slight curvature, but the S-shape is opposite in this case (concave for $V > 0$, convex for $V < 0$). This minor increase of the device resistance for increasing absolute voltage seems strange at first sight. The main voltage drop occurs across the bulk (fig. B.28), with negligible variation of the bulk conductivity for increasing voltage (fig. B.29). This is the reason for the linearity of the I-V curve. In contrast to the simulations with $\epsilon_\mathrm{r} = 1000$, the reverse-biased Schottky barriers at the BE for $V > 0$ and at the TE for $V < 0$ are not relevant for $\epsilon_\mathrm{r} = 100$, because their F-enhanced conductivity is already too high (fig. B.29). Instead, there are increasing voltage drops at the other, forward-biased Schottky barriers (fig. B.28), together with a local decrease of the conductivity due to lower electric field strengths (fig. B.29). This change near the TE for $V > 0$ and near the BE for $V < 0$ is responsible for the increasing device resistance and the inverted S-shape of the I-V curve.

The different I-V curves for $\epsilon_\mathrm{r} = 10$, 100 and 1000 are examples for the wide variety of curve shapes that may result from the combination of Schottky barriers with VRH transport. The analysis of the underlying processes revealed that the F-dependence of the hopping conductivity plays a decisive role. This is also demonstrated by the analogous I-V simulations where this F-dependency was neglected (FEM-3 model). In this case, the device resistance would always be dominated by the regions where $E_\mathrm{F}(z)$ crosses the band gap, even for $\epsilon_\mathrm{r} = 10$ (fig. B.30). The conductivity in these regions would no longer be increased by the electric field in the space-charge zone (fig. B.31), in contrast to fig. 5.68. This would lead to completely different I-V curves for all tested ϵ_r values (fig. B.32), which would have an inverted S-shape with approximately linearly increasing current at high voltages (concave for $V > 0$, convex for $V < 0$). The broader the Schottky barriers, i.e., the higher ϵ_r, the lower would be the currents. The different shape of these I-V curves originates from the broadening of the low-σ regions with increasing absolute voltage (fig. B.31). This is the same broadening as already discussed further above for $\epsilon_\mathrm{r} = 1000$ (cf. fig. 5.70c,d). However, when

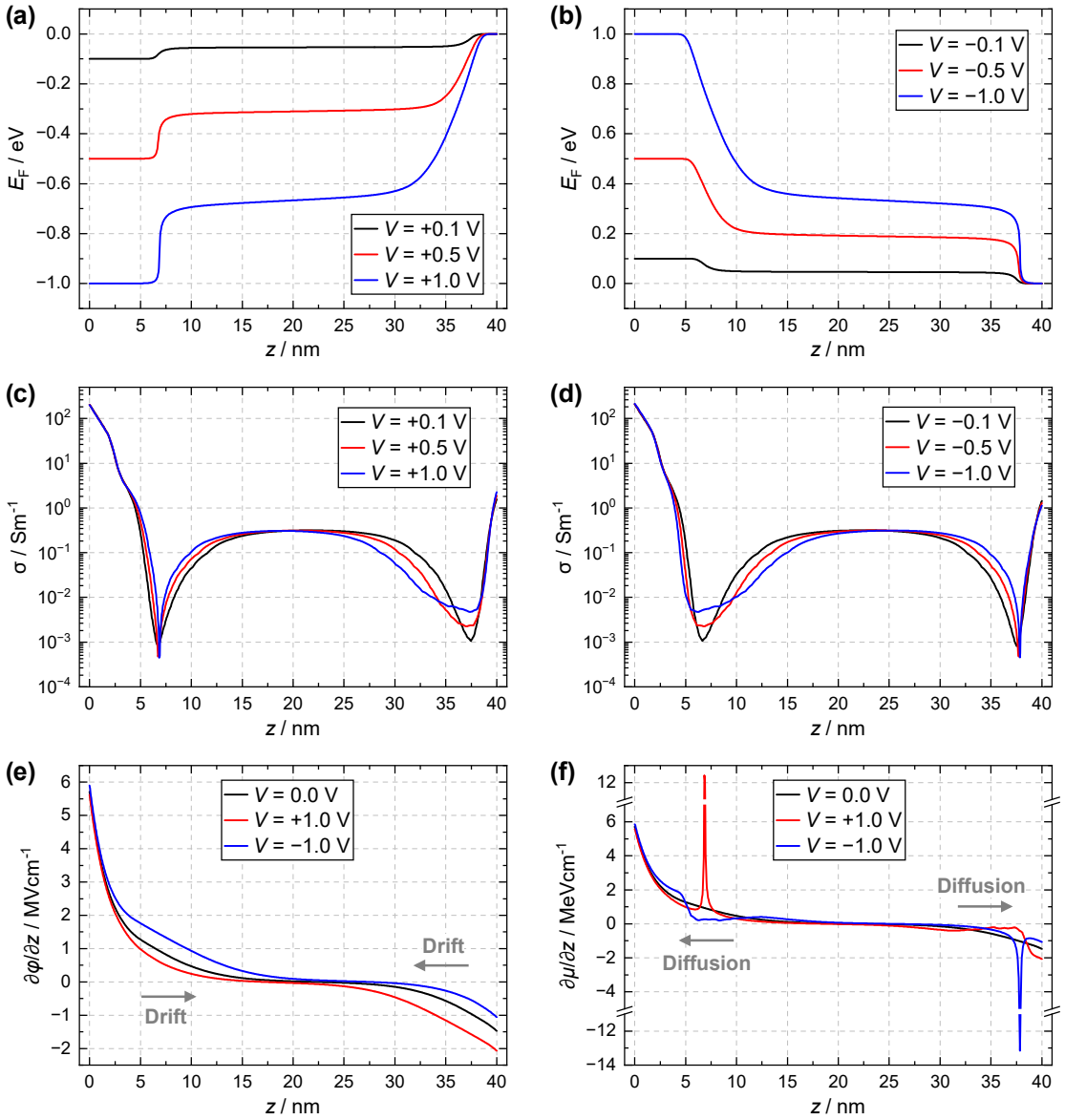


Figure 5.70: Voltage-dependence of the FEM results for Pt|a-GaO$_{1.0}$|ITO with $\epsilon_r = 1000$ (FEM-4 model with high-F effects). (a,b) Fermi level E_F (= $\tilde{\mu}$) for different positive and negative voltages. (c,d) VRH conductivity σ for different positive and negative voltages (cf. fig. B.31 for the corresponding FEM-3 results without high-F effects). (e) Electric potential gradient $\partial\varphi/\partial z$ (= $-\vec{F}$). (f) Chemical potential gradient $\partial\mu/\partial z$ of the electrons.

the F-dependency of the conductivity is neglected, then this broadening is no longer compensated and the device resistance increases, especially at low voltages. At higher voltages, the device resistance would stay approximately constant, because almost the whole voltage drop would occur in the growing regions with constant, low conductivity (fig. B.31). This shows that the exponential dependence between the VRH conductivity and the electric field strength is very important for the shape of the I-V curves.

Overall, the most interesting FEM result is the I-V curve for $\epsilon_r = 1000$ in fig. 5.69b, which has a

pronounced S-shape due to the F-dependency of the conductivity. The shape of this simulated curve resembles the I-V curves measured in the resistive switching experiments (cf. sec. 5.2). In contrast to classical back-to-back Schottky barriers with unequal electrodes, the curve is highly symmetric, which is another similarity to the measurements (e.g., fig. 5.12). Furthermore, the current density values are in a realistic range. The exact I-V relations are not reproduced, however. Contrary to the earlier expectation from $\ln(\sigma) \propto F^{1/2}$ (cf. sec. 5.5.4), the simulated curve cannot be described by $\ln(j/V) \propto V^{1/2}$ (fig. B.33a). This is actually no surprise, since the electric field is not constant in the dominating low-σ region and since the voltage drop does not occur exclusively in this region. Instead, the simulated curve approximately conforms to $\ln(j) \propto V^{1/2}$ (for $|V| > 0.5\,\mathrm{V}$, fig. B.33b). This slightly different relation has often been reported for a-GaO_x-based resistive switching devices, for example in [9, 26, 27, 177] (cf. sec. 2.4). Apparently, the devices in these works can be described by the idealized Pt|a-$GaO_{1.0}$|ITO model. This is yet another hint at the involvement of VRH in the resistive switching of a-GaO_x.

In general, it must be taken into account that a spatially constant composition is unlikely for real devices. For example, as already discussed in sec. 5.6.1, the O/Ga ratio could be increased in the low-σ regions due to the ion migration caused by the electric field of the Schottky barriers. Based on the KMC results for $x = 1.2$, one can expect that this would decrease the conductivity in these regions even further. The voltage drops would concentrate more to these regions, such that the I-V curve would probably converge towards the relation $\ln(j/V) \propto V^{1/2}$ that was measured in fig. 5.26. Another consequence would be that the narrow regions near the interfaces would dominate and produce an S-shaped curve even for lower permittivity values of $\epsilon_\mathrm{r} < 1000$. Further changes of the O/Ga ratio in the a-GaO_x layer, for example due to surface oxidation or during resistive switching, would of course affect the shape of the I-V curves as well. This would also explain the different degrees of curvature observed in the I-V experiments. Depending on the resulting conductivity and electric field strength in the dominating regions, the I-V curves could become more linear or more S-shaped.

Finally, it is time to pick up the thread of the resistive switching discussion in sec. 5.2.4. The experimental RS behavior could not be explained with the classical bulk or interface switching models due to three main reasons:

- The voltage- and stoichiometry-dependence of the electron transport through metal|a-GaO_x Schottky barriers was unclear.
- Bulk-dominated RS would fit to the symmetry and the occasional linearity of the measured I-V curves, but it would be inconsistent with the frequently observed S-shaped curvature of individual resistance states. Conversely, interface-type RS would provide an explanation for the curvature, but it would not lead to symmetric or linear I-V characteristics.
- Figure-8 switching would require that the major voltage drop occurs at or near the ITO-BE for negative voltage, but this seemed unlikely due to the previous HRS → LRS transition at positive voltage and due to the reverse-biased Pt|a-GaO_x interface.

The FEM simulations without ionic motion cannot fully reveal the resistive switching mechanism, of course, but they resolve the first and the second point on this list. According to the FEM results, the electron transport through a metal|a-GaO_x Schottky barrier is controlled by the low-σ region near the interface, where the Fermi level crosses the band gap. The conductivity measurements and the KMC simulations showed that this region would become more resistive when the local O/Ga ratio increases. The influence of an applied voltage on the electron transport through this region was clarified by the FEM calculations (e.g., fig. 5.70). Furthermore, the simulations demonstrated that the electric field strength in the devices may become high enough to affect the VRH conductivity, especially near the interfaces but probably also in the bulk. In combination with the exponential F-dependency of the conductivity, this leads to a more or less linear I-V curve, depending on the local electric field strength. Simultaneously, the I-V curve tends to be symmetric because the hopping conductivity does not depend on the direction of the electric field.

The third point concerning figure-8 RS cannot be resolved with certainty, because it can only be speculated about the ion transport in the devices. The most likely explanation is still a certain influx of oxygen ions from the ITO electrode into the highly oxygen-deficient a-GaO_x layer, due to the steep gradient of the chemical potential μ_O of oxygen at the interface (cf. sec. 5.2.4). Additionally, two plausible assumptions can be made about the ion transport:

- Oxygen ions are mobile, but only in regions with a strong electric field or a steep chemical potential gradient $\nabla\mu_O$, due to a generally low ionic conductivity (cf. sec. 2.3.4). Otherwise, the different resistance states of the devices would not be stable over time.
- For bipolar RS, the internal switching processes must occur in regions where the electric field reverses its direction depending on the voltage polarity. This essentially rules out switching directly at the interfaces, where the direction of the electric field is governed by the work function difference (cf. fig. 5.70e). In order to have sufficiently high $\nabla\varphi$ for at least some ion transport, these regions must be narrow with a high voltage drop (i.e., low conductivity). Furthermore, the local $\nabla\mu_O$ must be rather low, because otherwise the changes induced by the applied voltage would be reverted by diffusion when $V = 0$.

Together with the FEM results, this leads to two scenarios for the initial state of the Pt|a-GaO_x|ITO devices, depending on whether the surface of a-GaO_x was exposed to air before the TE deposition:

- **With surface oxidation:** Both the air exposure and the oxygen influx from ITO increase the O/Ga ratio at the respective interfaces, but the effect of the surface oxidation is stronger. The additional oxygen ions from both outer sources partially migrate away from the interfaces to the low-σ regions, due to the electric field of the Schottky barriers. Since both the initial increase of the O/Ga ratio and the electric field strength are higher at the TE than at the BE (cf. fig. 5.70e), this leads to a more resistive low-σ region near the TE than near the BE.
- **Without surface oxidation:** In this case, the only additional oxygen is provided by ITO. The electric field at the TE is still stronger, but without the oxygen from air, the increase of the O/Ga ratio is more pronounced for the low-σ region near the BE than near the TE.

Since the low-σ regions are near the inner end of the Schottky barriers and since the oxygen ion accumulation partially compensates the electron depletion, the electric field in these regions should be rather weak for $V = 0$. It is thus plausible that the electric field direction in these regions would switch with the voltage polarity. Furthermore, it explains why the I-V curves could still be symmetric when only one low-σ region dominates. It can also be expected that these regions are very narrow, especially when $\epsilon_r < 1000$ (cf. fig. 5.68a). The resulting high electric field strength for $V \neq 0$ then causes not only the switching transitions but also the observed S-shape of the I-V curves.

An applied voltage leads to oxygen ion migration to or from the dominating region and the resulting changes of the local VRH conductivity are visible as resistive switching. This voltage-induced exchange of oxygen with the adjacent material could occur either at the outer or the inner side of the low-σ region, because it is not located directly at the interface. This means that, in principle, the O/Ga ratio could increase or decrease for each voltage polarity, which would enable a large variety of different switching patterns. Since the electric field of the Schottky barrier likely prevents a reversible oxygen exchange in the direction of the interface, it is most probable that the switching transitions occur at the inner side of the low-σ region, in the direction of the bulk. This leads to the following idea for the resistive switching mechanism of Pt|a-GaO_x|ITO:

- **With surface oxidation → counter-8 bipolar RS:** The low-σ region near the Pt-TE dominates the device resistance.[123]
  - $V > 0$: Oxygen ions migrate from the bulk to the low-σ region near the TE, which leads to an LRS → HRS transition. This transition is not visible in the first cycle, because the device is initially in the HRS (O/Ga ratio already very high).
  - $V < 0$: Oxygen ions migrate from the low-σ region near the TE to the bulk, which leads to an HRS → LRS transition.
- **Without surface oxidation → figure-8 bipolar RS:** The low-σ region near the ITO-BE dominates the device resistance.
  - $V > 0$: Oxygen ions migrate from the low-σ region near the BE to the bulk, which leads to an HRS → LRS transition.
  - $V < 0$: Oxygen ions migrate from the bulk to the low-σ region near the BE, which leads to an LRS → HRS transition.

This is essentially a form of bulk switching in very confined regions near the interfaces. At sufficiently high V_{max} in the I-V experiment, both low-σ regions have substantial voltage drops and all four transitions are observed as abnormal bipolar RS (cf. fig. 5.25).

[123] The proposed mechanism also agrees with the XPS experiments in [10], where the changes of the O/Ga ratio in the first few nanometers below the Pt-TE were measured after counter-8 bipolar RS. In these experiments, the O/Ga ratio increased for $V > 0$ and it decreased for $V < 0$.

6

Conclusion

The primary aim of this work was the investigation of the resistive switching of Pt|a-GaO_x|ITO devices, with a focus on the electron transport. Previous studies showed promising results for these devices, but they also reported a wide variety of switching patterns and interpretations. Several variants of filament-, interface- and bulk-type switching have been proposed for the individual observations. However, a deeper understanding of the responsible processes or a unified RS mechanism have remained elusive. The results of this work, which are summarized in the following, contribute a lot of crucial pieces to this puzzle.

The initial characterization of the PLD-made a-GaO_x thin films with x in the range of 0.8–1.1 confirmed several expected trends, often for the first time:

- Stoichiometry: The higher the PLD fluence, the higher the O/Ga ratio x. No increase of x in the bulk during long-term storage in air.
- Growth rate: The higher the laser fluence, the higher the growth rate during PLD.
- Color: The lower the O/Ga ratio, the darker and less transparent are the thin films.
- Deposition in H_2/Ar instead of Ar: H_2 mainly reduces the growth rate, while it has only a weak effect on the O/Ga ratio.
- Surface oxidation: Exposure to air leads to a significant increase of the O/Ga ratio to approx. $x_{surf} \approx 1.3$. This effect is limited to the surface and the O/Ga ratio decreases gradually to its bulk value within the first 10–15 nm of the film (after > 1 year).
- Work function: The work function of a-GaO_x with $x_{surf} \approx 1.3$ is approx. 3.7–4.0 eV, i.e., much lower than for a-$GaO_{1.5}$ or β-Ga_2O_3.

The most influential of these findings is probably the surface oxidation, because of its relevance for any surface-sensitive measurement or application. It typically has not been taken into account in the a-GaO_x-related literature and most works not even mention whether the severely oxygen-deficient films have been exposed to air. Besides the consequences for the interpretation of experiments, the measured surface oxidation also confirms another important point: The ionic conductivity of oxygen in a-GaO_x must be very low, because otherwise the diffusion profile of the incorporated

oxygen would reach further into the sample. This is beneficial for resistive switching since it implies a good stability of the resistance states.

The subsequent I-V experiments with the prepared Pt|a-GaO_x|ITO devices reflected the confusing status quo in the literature. The devices exhibited counter-8, figure-8 and abnormal bipolar resistive switching, together with a wide range of curve shapes from linear to S-shaped. Nonetheless, it was possible to gain several insights:

- The measured I-V curves are often very symmetric and they obey different functional relationships (e.g., $j \propto V$, $j \propto V^2$, $\ln(j/V) \propto V^{1/2}$ or $\ln(j/V) \propto V^2$).
- No electroforming is necessary before resistive switching.
- The RS transitions are gradual, with an exponential time-dependence of the resistance changes and with rather low HRS/LRS resistance ratios.
- Counter-8 bipolar RS starts with an HRS → LRS transition at negative voltage.
- Figure-8 bipolar RS seems to occur when surface oxidation is prevented and it often has a much higher switching speed than counter-8 bipolar RS.
- Increasing V_{max} of the I-V sweeps leads to loops and abnormal bipolar RS.
- Two oppositely oriented switching regions seem to exist in the devices. One of them is responsible for counter-8 and the other for figure-8 bipolar RS. Together, they yield abnormal bipolar resistive switching.

In hindsight, the diversity of the RS results turned out as a strength, since it revealed new connections between counter-8, figure-8 and abnormal bipolar switching. The interpretation of the RS behavior in terms of mechanistic processes remained difficult, though, because one could only speculate about which regions within the devices might dominate the resistance.

The next step thus was a closer investigation of the electron transport in a-GaO_x. The conductivity measurements led to an important discovery: The temperature-dependence of the conductivity agrees very well with Mott's law for variable-range hopping (VRH), even at room temperature and above. This was a major surprise, since all previous publications about a-GaO_x have assumed mobility edge conduction (MEC) as the electron transport mechanism. For all measured samples, the fits according to the VRH law $\sigma \propto \exp\left[-(T_0/T)^{1/4}\right]$ were superior to the MEC law $\sigma \propto \exp\left[-\Delta E/(k_B T)\right]$. Additionally, the conductivity measurements yielded the following results:

- Thermal relaxation of the amorphous structure: Irreversible decay of the conductivity with an exponential time-dependence during the first heating ("healing" of electronic defects → higher T_0 → lower σ).
- Stoichiometry-dependence: The conductivity decreases non-linearly with increasing O/Ga ratio x, which is attributable to an increase of T_0 for higher x.

- Influence of magnetic fields: Very small, linear increase of σ with increasing magnetic field strength (negative magnetoresistance) and no measurable Hall effect, as expected for VRH.

This experimental evidence for electron hopping as the dominating charge transport process in a-GaO_x with $x \ll 1.5$ is certainly the most consequential finding of this work. Most effects that are relevant in electronic devices are different for VRH in comparison to band conduction – especially the way the electron transport is influenced by other factors like the material composition, electric fields or changes of the Fermi level. VRH thus must be taken into account in every electronics application of highly oxygen-deficient a-GaO_x. In the RS devices, two main aspects are affected: The stoichiometry-dependence of the conductivity, i.e., the influence of ion transport on the electron flux, and the electron transport through metal|a-GaO_x Schottky barriers. The exploration of these consequences became a main task of this work.

Since VRH normally just dominates at temperatures far below RT, it was suspected that a-GaO_x must have an exceptionally high density of localized electronic states near the Fermi level. This was confirmed by the HAXPES measurements and the DFT-DOS calculations:

- The HAXPES spectra for $x_{\mathrm{surf}} \approx 1.2$ have a broad defect band from the valence band up to the Fermi level, whereas a-$GaO_{1.5}$ only exhibits band tailing.
- The lower the O/Ga ratio, the more defect states emerge between the valence band and the conduction band in the DFT-DOS. The DOS for $x = 1.2$ already features some defect states near the Fermi level, while $x = 1.0$ and 0.8 lead to a high density of defect states at and around E_{F}.
- The defect states have approximately equal contributions from gallium and oxygen orbitals.
- The defect states and the tail states of CB and VB are localized states, as concluded from analyzing the charge density of the individual states. The distances from the Fermi level to the mobility edges E_{C} and E_{V} are approximately 0.5–1.0 eV and 1.5–3.5 eV, respectively.

These results demonstrate that the combination of two effects is responsible for the dominance of VRH over MEC: The severe non-stoichiometry creates a lot of defect states in the former band gap and the amorphicity prevents the formation of extended wave functions in the defect band.

The DFT-DOS also were the basis for the subsequent 3D-KMC simulations of the VRH conductivity. These simulations yielded the same exponential $T^{-1/4}$-dependence as the conductivity measurements and it was even possible to reproduce the experimental results quantitatively. This was the final confirmation that variable-range hopping is a proper model for the electron transport in a-GaO_x. The KMC simulations then were used to elucidate the influence of different physical parameters:

- Localization radius α: In accordance with the classical VRH theory by Mott, T_0 increased with decreasing α approximately via $T_0 \propto \alpha^{-3}$, which led to a strong decrease of σ for lower α.
- O/Ga ratio x: A non-linear decrease of σ and an increase of T_0 were observed for increasing x, just like in the conductivity measurements. In contrast to $x = 0.8$ and 1.0, the simulations for

x = 1.2 deviated from the $T^{-1/4}$-dependence of Mott's law because the Fermi level is in a small band gap.

- Electric field strength F: Strong electric fields caused an elevated electron temperature T_{eff} and an exponential increase of the conductivity, with $\ln(\sigma) \propto F^{1/2}$ for a-$GaO_{1.0}$ at high F. Both the increase of the charge carrier density n_{eff} through $T_{\text{eff}} > T$ and the increase of the mobility through the higher tunneling probability were responsible for the increasing conductivity.
- Chemical potential μ: The simulations with positively and negatively charged a-$GaO_{1.0}$ yielded the expected changes of the conductivity due to the shift of the Fermi level. The conductivity decreased by several orders of magnitude for $\mu \approx -0.5\,\text{eV}$, where the Fermi level is in a small band gap (still approx. $\ln(\sigma) \propto F^{1/2}$ at high F).

Besides these insights about a-GaO_x, the KMC simulations also led to several contributions to the VRH theory in general:

- Mott's VRH law is a valid approximation not only when the DOS is constant in the relevant energy range around E_{F} but also when the DOS varies moderately (i.e., as long as the correlation between distance and energy difference of the hops is rather weak).
- In general, the charge carrier density of VRH is given by the effective charge carrier density n_{eff} (as known from band conduction in metals). The concept of effective charge carriers is essential for a proper definition of the mobility and the diffusion coefficients for VRH.
- The Nernst-Einstein relation between the mobility and the Fickian diffusion coefficient is also valid for VRH. Remarkably, it retains its validity even at high F, as long as T is replaced by T_{eff}. This is a special consequence of VRH, where essentially no hop is completely perpendicular to the electric field. The Fickian diffusion coefficient – obtained from the mean squared displacements perpendicular to $\vec{F}$ – exhibits approximately the same T-, F- and μ-dependence as the conductivity.

The KMC simulations of this work are among the most comprehensive and sophisticated in the field of VRH research. To the best of the author's knowledge, they are the first VRH simulations with an ab-initio DOS of an amorphous material. This was made possible by the developed VHoppAS software, which enables efficient KMC-VRH simulations with an arbitrary DOS as distribution function for the localized states (free + open-source on GitHub or via [204]).

After the DFT calculations of the electronic structure and the KMC simulations of the electron movement, the computational multi-scale approach was completed by the FEM modeling of the electron transport on the device scale. The FEM calculations for Pt|a-$GaO_{1.0}$|ITO were likely the first ever simulations of an electronic device with variable-range hopping through Schottky barriers. Even though the FEM models did not contain ion movement, they led to significant advances with regard to resistive switching: The resistance of Pt|a-GaO_x|ITO devices is most likely dominated by narrow low-σ regions near each interface, where the Fermi level crosses a small band gap or DOS minimum. The existence of these low-σ regions in real devices is quite certain, because otherwise

the interfaces would have no influence on the I-V characteristics. This would contradict the experimental observations. The shape of the I-V curves is governed by the exponential F-dependence of the VRH conductivity in the low-σ regions, which explains the symmetry and the curvature of the I-V measurements. The variation of the local O/Ga ratio in these low-σ regions through ion transport thus seems to be responsible for the resistive switching of a-GaO_x. Based on these conclusions, a new unified RS mechanism was proposed for the counter-8, figure-8 and abnormal bipolar switching behavior. Its main principle is the reversible oxygen ion migration between the low-σ regions and the nearby bulk. Counter-8 bipolar RS results when the low-σ region near the Pt-TE dominates the device resistance, whereas figure-8 bipolar RS is observed when the other low-σ region near the ITO-BE is more resistive. This is controlled for example by the interface oxidation through air or ITO. If the applied voltage is rather high, then both regions have sufficient electric field strength for ion migration, which leads to abnormal bipolar RS. With the currently available evidence, this resistive switching mechanism is just a hypothesis. However, it is likely the first RS model for a-GaO_x that would explain the different switching patterns and curve shapes, and it is an important guide for the further experimental and theoretical investigation in the future.

In summary, the reported results represent a significant step forward for the understanding of a-GaO_x with $x \ll 1.5$ in general and for its resistive switching devices in particular. The insights even could be relevant for RS devices with $x \approx 1.5$, because VRH could dominate in the often postulated, oxygen-deficient filaments. Furthermore, the findings of this work should be transferable to the whole class of amorphous, highly non-stoichiometric materials like for example a-HfO_x or a-TaO_x. It is likely a general principle that VRH dominates even at room temperature once an amorphous material is prepared with a sufficient amount of compositional defects. The consequences then should be quite comparable to those observed for a-GaO_x.

This work naturally offers plenty of inspiration for further research. Two main points are definitely the experimental investigation of the proposed RS mechanism and the examination of comparable materials for VRH transport. Furthermore, it is for example necessary to clarify the role of hydrogen in a-GaO_x, the influence of Joule heating during resistive switching and whether MEC really begins to dominate for $x \to 1.5$. To conclude this work, several more specific ideas are listed in the following:

- Extension of the FEM models to include spatial variations of the O/Ga ratio x (or even dynamic ion movement): Instead of expensive DFT calculations on a fine-grained O/Ga ratio scale, it is likely more feasible to abstract the available DFT results into an x-dependent DOS model (incl. E_F vs. x), which is then fed to KMC in order to generate the $\sigma(x, \mu, F)$ data required for FEM. For a full description of resistive switching, the model for the electron transport could be paired with a calculation of the ion transport over time (from the continuity equation for $\partial x/\partial t$).
- In-situ HAXPES: The HAXPES results demonstrated that the O/Ga ratio can be determined through a thin metal layer. This could be used to monitor the O/Ga ratio beneath the TE during resistive switching.

- Photoelectron spectroscopy at the a-GaO_x|ITO interface: To investigate the proposed changes of the O/Ga ratio near the ITO electrode, the device setup could be reversed (ITO as TE, Pt as BE). The photoemission from gallium should also be visible through a thin ITO layer. Note that this approach would be limited to amorphous ITO. Alternatively, a water-soluble substrate like NaCl could be used for the Pt|a-GaO_x|ITO devices, which can be removed after the preparation of the sample.
- Improved contacts for resistive switching: The RS experiments could be optimized by using a crossbar sample architecture, where the individual devices are formed by intersections of line-shaped TEs and BEs. In this case, the electrodes can be contacted by conductive paste without restricting the device size. This would enable the combination of small devices and reliable contacts with negligible series resistance. Crosstalk between the devices should be no issue due to the rather low HRS/LRS ratios.
- Bubble formation at high voltage: In the *I-V* experiments, it was observed that very high positive voltage causes irreversible damage in the form of shallow bubbles. The respective devices had no surface oxidation and exhibited figure-8 bipolar RS. TEM measurements revealed that the "bubbles" correspond to voids at the a-GaO_x|ITO interface. Their round shape speaks against a simple film delamination. Instead, it seems like O_2 gas is created at the interface. This interpretation is supported by some first EELS results: At the border of the a-GaO_x layer to the void, the oxygen signal is significantly enhanced in comparison to the bulk (while the gallium signal is unchanged). This would fit to an enhanced oxygen transfer from ITO at high $V > 0$. Further investigation of this phenomenon could yield valuable information about the RS mechanism.
- Oxygen transfer at the a-GaO_x|ITO interface: The roughness of ITO can be optimized to produce a sharp interface. After ^{18}O-exchange of the ITO layer and subsequent a-GaO_x deposition, the oxygen diffusion from ITO to a-GaO_x could be analyzed by SIMS (for example after different waiting times). To avoid ion beam mixing during SIMS, it is probably advised to cut a FIB lamella from the sample.
- Surface oxidation with ^{18}O: Freshly prepared a-GaO_x films could be exposed to $^{18}O_2$ or $H_2^{18}O$, for example by injecting these compounds into the PLD chamber. The resulting diffusion profile of ^{18}O at the surface could be analyzed with SIMS after different waiting times or after applying a voltage.
- Hydrogen incorporation: The role of hydrogen could be investigated by supplying $^2H_2$ or ^2H_2O to the PLD chamber either during or after the deposition of a-GaO_x. The presence of deuterium in the thin films then could be analyzed by SIMS or solid-state NMR.

A

Calculations

A.1 Random Distances within a Sphere

Let $\vec{r}$ be a random point within a sphere with radius R. A spherical coordinate system with its origin in the center of the sphere is employed, such that the coordinates for $\vec{r}$ are the radius r (distance from the center), inclination θ (angle to z axis) and azimuth φ (angle to the x axis in the xy-plane). The volumetric average of a function corresponds to its integral over said volume, divided by the volume. The average of the distance r from the origin of the sphere is thus given by

$$\begin{aligned}\langle r\rangle &= \frac{1}{V_{\text{sph}}}\int_V r\,\mathrm{d}V = \frac{1}{V_{\text{sph}}}\int_0^{2\pi}\int_0^{\pi}\int_0^{R} r^3\sin(\theta)\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\varphi \\ &= \frac{1}{V_{\text{sph}}}\int_0^{2\pi}\mathrm{d}\varphi\int_0^{\pi}\sin(\theta)\,\mathrm{d}\theta\int_0^{R} r^3\,\mathrm{d}r = \frac{3}{4\pi R^3}2\pi 2\frac{1}{4}R^4 = \frac{3}{4}R \end{aligned} \tag{A.1}$$

where the definition of the volume element in spherical coordinates, $\mathrm{d}V = r^2\sin(\theta)\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\varphi$, and the sphere volume $V_{\text{sph}} = 4\pi R^3/3$ were used. In a similar way, the average distance in one direction can be obtained. Now, $\vec{r}$ shall be a random point within the semi-sphere with $z \geq 0$, i.e., θ is restricted to the range $[0, \pi/2]$. The distance of $\vec{r}$ to the xy-plane is $z = r\cos(\theta)$ and the average of this function is

$$\begin{aligned}\langle z\rangle &= \frac{1}{V_{\text{sph}}/2}\int_V r\cos(\theta)\,\mathrm{d}V = \frac{1}{V_{\text{sph}}/2}\int_0^{2\pi}\int_0^{\pi/2}\int_0^{R} r^3\cos(\theta)\sin(\theta)\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\varphi \\ &= \frac{1}{V_{\text{sph}}/2}\int_0^{2\pi}\mathrm{d}\varphi\int_0^{\pi/2}\cos(\theta)\sin(\theta)\,\mathrm{d}\theta\int_0^{R} r^3\,\mathrm{d}r = \frac{3}{2\pi R^3}2\pi\frac{1}{2}\frac{1}{4}R^4 = \frac{3}{8}R \end{aligned} \tag{A.2}$$

A.2 Effective Charge Carrier Density in Constant or Linear DOS

Let a DOS $g(\varepsilon)$ that is linear with regard to the state energy ε in the vicinity of the Fermi level (= μ on the ε-axis) be defined by

$$g(\varepsilon) = g(\mu) + \Delta g(\mu)(\varepsilon - \mu) \quad \text{with} \quad \Delta g(\mu) = \left(\frac{\partial g(\varepsilon)}{\partial \varepsilon} \right)_{\varepsilon=\mu} \tag{A.3}$$

This corresponds to a 1st order Taylor expansion of the DOS around the Fermi level. It includes also the case of constant DOS ($\Delta g(\mu) = 0$). With the definition of the effective charge carrier density n_{eff} (eq. 2.23) follows

$$\begin{aligned} n_{\text{eff}} &= \int_{-\infty}^{\infty} [g(\mu) + \Delta g(\mu)(\varepsilon - \mu)] f_{\text{FD}}(\varepsilon, \mu) [1 - f_{\text{FD}}(\varepsilon, \mu)] \, \mathrm{d}\varepsilon \\ &= g(\mu) \int_{-\infty}^{\infty} f_{\text{FD}}(\varepsilon, \mu) [1 - f_{\text{FD}}(\varepsilon, \mu)] \, \mathrm{d}\varepsilon + \Delta g(\mu) \int_{-\infty}^{\infty} (\varepsilon - \mu) f_{\text{FD}}(\varepsilon, \mu) [1 - f_{\text{FD}}(\varepsilon, \mu)] \, \mathrm{d}\varepsilon \end{aligned} \tag{A.4}$$

The second integral always equals zero, because $(\varepsilon - \mu)$ is antisymmetric and $f_{\text{FD}}[1 - f_{\text{FD}}]$ is symmetric around μ. The first integral is easily solved, since

$$\frac{\partial f_{\text{FD}}(\varepsilon, \mu)}{\partial \varepsilon} = -\frac{1}{k_{\text{B}} T} f_{\text{FD}}(\varepsilon, \mu) [1 - f_{\text{FD}}(\varepsilon, \mu)] \tag{A.5}$$

analog to the derivative with regard to μ that was encountered in the derivation of eq. 2.23. With $f_{\text{FD}} \to 0$ for $\varepsilon \to \infty$ and $f_{\text{FD}} \to 1$ for $\varepsilon \to -\infty$ then follows

$$n_{\text{eff}} = -k_{\text{B}} T g(\mu) \int_{-\infty}^{\infty} \frac{\partial f_{\text{FD}}(\varepsilon, \mu)}{\partial \varepsilon} \, \mathrm{d}\varepsilon = -k_{\text{B}} T g(\mu) [0 - 1] = k_{\text{B}} T g(\mu) \tag{A.6}$$

Both for constant and for linear DOS, n_{eff} therefore always equals $k_{\text{B}} T$ times the DOS at the Fermi level, regardless of the slope of the DOS. This result was also used by Mott in the derivation of the VRH law (eq. 2.54).[31, 72]

A.3 Weighted Mean, Inner Error and Outer Error

The weighted mean $\bar{x}$ of n different values $x_i \pm s_i$, each associated with an individual standard deviation s_i, is given by

$$\bar{x} = \frac{\sum_{i=1}^{n} \frac{x_i}{s_i^2}}{\sum_{i=1}^{n} \frac{1}{s_i^2}} \tag{A.7}$$

The standard error $u_{\bar{x}}$ of the weighted mean corresponds to the maximum of the inner error u_{in} and the outer error u_{out}, as defined by

$$u_{\mathrm{in}} = \sqrt{\frac{1}{\sum_{i=1}^{n} \frac{1}{s_i^2}}}, \quad u_{\mathrm{out}} = u_{\mathrm{in}} \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{s_i^2}} \quad \text{and} \quad u_{\bar{x}} = \max(u_{\mathrm{in}}, u_{\mathrm{out}}) \tag{A.8}$$

This calculation respects both the individual uncertainties and the deviations between the values. The equation for the inner error u_{in} originates from Gaussian error propagation of the weighted mean formula (eq. A.7), while the outer error u_{out} is a weighted standard deviation. Note that, when all s_i are similar (= s_0), these expressions correspond to the conventional "standard error of the arithmetic mean", as required:

$$u_{\bar{x}} = u_{\mathrm{in}} = u_{\mathrm{out}} = \frac{s_0}{\sqrt{n}} \quad \text{with} \quad s_0 = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{A.9}$$

The result $\bar{x} \pm u_{\bar{x}}$ of the general eqs. A.7 and A.8 gives a confidence interval that contains the "true" value of x with a probability of 68% (assuming normal distribution of x). However, this is only relevant when measuring a quantity x that is supposed to be a constant. When combining the measurements from different samples, for example, which are expected to have each a slightly different "true" value of x, one is interested in the spread of the value x among them, instead. Under the assumption that the different "true" values of x are identically distributed for the samples, a suitable error bar then can be obtained as

$$s_x = u_{\bar{x}} \sqrt{n} \tag{A.10}$$

which converts the standard error of the weighted mean back into a sample standard deviation s_x.[124] The final result $\bar{x} \pm s_x$ is a prediction interval for future, similarly determined values x_i, which lie within this interval around $\bar{x}$ with a certain probability (68% when x follows a normal distribution). In this work, for example the SKPM work functions from different, comparable a-GaO_x samples were aggregated in this way.

[124] The use of the sampling size n in the "$n - 1$" and "$\sqrt{n}$" terms of eqs. A.8 and A.10 is an approximation. For more exact unbiased estimators of the weighted standard deviations, n in these terms must be replaced by the so-called effective sampling size, which equals the square of the sum over s_i^{-2}, divided by the sum over s_i^{-4}. This has an influence especially when the s_i values have large scatter. The more equal the s_i, the closer is the effective sampling size to n. For the use cases in this work, there is virtually no difference between using n or the effective sampling size.

A.4 Tracer Statistics and Random Walk Theory for Electrons

In the KMC simulations, the movement of each electron k in the simulation cell is tracked individually. Its displacement vector

$$\vec{R}_k = \begin{pmatrix} X_k \\ Y_k \\ Z_k \end{pmatrix} \tag{A.11}$$

describes how far it moved in x-, y- and z-direction since the start of the simulation. However, this quantity is only meaningful if the electron experiences the same thermodynamic driving forces during the whole simulation. If the transport conditions would change, spatially or temporally, it would be impossible to retrieve material parameters just from the electron's overall displacement. Therefore, the simulation cell must be macroscopically homogeneous, with spatially and temporally constant chemical potential μ, electron density n, electric field $\vec{F}$ and temperature T, combined with periodic boundary conditions. This quite obvious requirement does not impede the simulation of a drift flux, since $\vec{F} \neq 0$ is still allowed, but it prevents the observation of Fickian diffusion because of $\nabla\mu = 0$ and $\nabla n = 0$. However, even though there is no net diffusion flux, there is still diffusive motion of the electrons in the homogeneous system, which is called tracer diffusion (or self-diffusion). The following derivations will show that by knowing the effective charge carrier density n_{eff}, which was introduced in sec. 2.1.1, it is possible to obtain the Fickian diffusion coefficient from the tracer diffusion observed in KMC.

In principle, the choice of which electrons are monitored as tracers is arbitrary. The common approach to track all electrons is just the most straightforward option. Generally, depending on which electrons are considered as tracers, there might be a tracer concentration gradient $\nabla n^* \neq 0$ and an associated chemical potential gradient $\nabla\mu^* \neq 0$ of the tracers. With starred quantities referring to tracers, the tracer flux thus can be written as

$$\vec{J}_{\mathrm{e}}^{\,*} = -\frac{\sigma^*}{e^2}\nabla\tilde{\mu}^* = -\frac{\sigma^*}{e^2}\nabla\mu^* + \frac{\sigma^*}{e}\nabla\varphi \tag{A.12}$$

There is a constraint, though, for the tracer choice, because individual electrons are not necessarily representative for others or for the whole collective. Electrons deep in the valence band, for example, are mostly immobile because the energy difference to the unoccupied states near and above the Fermi level is too large. They are essentially blocked from hopping and considering just them as tracers would lead to the false conclusion that there is almost no tracer diffusion. The tracer choice determines what fraction of the overall electron motion in the system is observed. In order to be representative for the material, the tracers n^* must be selected such that they together behave like a certain amount of effective charge carriers n_{eff}^*. By definition, each effective charge carrier – as a pseudo-particle – is representative for all other effective charge carriers and thus for the whole material (cf. sec. 2.1.1). The relation between n^* and n_{eff}^* is neither obvious nor unique. It is required,

though, that increasing the number of representative tracers for example by a factor 2 also doubles the observed tracer flux and the number of corresponding effective charge carriers. In other words, whatever the rule by which the tracers are selected, the ratio

$$\Phi^* := \frac{n^*}{n^*_{\mathrm{eff}}} \tag{A.13}$$

must be spatially and temporally constant. Note that this implies $\nabla n^* = \nabla n^*_{\mathrm{eff}}$. The actual value of Φ^* depends on the tracer choice. For example, if n^* is a uniformly random sample of all CB tail electrons (i.e., of n_{eff}) in an n-type, non-degenerate semiconductor, then is $\Phi^* = 1$. More generally, when n^* is a uniformly random sample of all electrons in the KMC cell (or when it corresponds to all electrons) then Φ^* must equal the thermodynamic factor Φ according to eq. 2.19. This would be an elegant tracer choice, since Φ is known – it is simply the ratio between n and n_{eff} for the KMC cell (cf. eqs. 4.16 and 4.18). The actual tracer choice used in the KMC simulations will be explained further below.

The effective charge carriers n^*_{eff} must react to an electric field like in an inhomogeneous system with similar local μ, which means that their mobility cannot depend on μ^*) and must correspond to u as defined in eq. 2.28. The drift flux has to be similar both in the n^*- and the n^*_{eff}-picture, which requires

$$\sigma^* = e n^* u^*_{\mathrm{N}} = e n^*_{\mathrm{eff}} u \quad \Rightarrow \quad u = u^*_{\mathrm{N}} \Phi^* \tag{A.14}$$

where u^*_{N} denotes the tracer mobility. From eq. A.14 follows that u^*_{N} must be spatially constant and independent of n^* and μ^*, though it depends on the way the tracers are selected. For the exemplary tracer choices mentioned above, it is $u^*_{\mathrm{N}} = u_{\mathrm{N}}$ when $\Phi^* = \Phi$ (cf. eq. 2.28) and $u^*_{\mathrm{N}} = u$ when $\Phi^* = 1$. Like u^*_{N}, also the tracer diffusivity must not vary in the homogeneous KMC cell. Therefore, the tracer flux can also be defined by

$$\vec{J}^*_{\mathrm{e}} = -D^*_{\mathrm{N}} \nabla n^* + \frac{\sigma^*}{e} \nabla \varphi \tag{A.15}$$

with a spatially constant tracer diffusion coefficient D^*_{N}, which is also called self-diffusion coefficient. This is just compatible to eq. A.12 if $\nabla \gamma^* = 0$ for the tracer activity coefficient γ^*, since otherwise the prefactor before ∇n^* would not be spatially constant:

$$\begin{aligned} -\frac{\sigma^*}{e^2} \nabla \mu^* &= -\frac{\sigma^*}{e^2} \frac{k_{\mathrm{B}} T}{\gamma^* n^*} \nabla (\gamma^* n^*) = -\frac{u^*_{\mathrm{N}}}{e} \frac{k_{\mathrm{B}} T}{\gamma^*} \nabla (\gamma^* n^*) = -\frac{u^*_{\mathrm{N}}}{e} \frac{k_{\mathrm{B}} T}{\gamma^*} \gamma^* \nabla n^* \\ &= -\frac{u^*_{\mathrm{N}} k_{\mathrm{B}} T}{e} \nabla n^* \end{aligned} \tag{A.16}$$

In combination with eqs. A.12 and A.15 results the Nernst-Einstein relation for the tracer diffusion

coefficient:

$$D_{\mathrm{N}}^{*} = \frac{u_{\mathrm{N}}^{*} k_{\mathrm{B}} T}{e} \tag{A.17}$$

Finally, from eqs. 2.29, A.14 and A.17 follows that the tracer diffusion coefficient D_{N}^{*} is related to the Fickian diffusion coefficient D by the factor Φ^{*}:[236]

$$D = D_{\mathrm{N}}^{*}\Phi^{*} = D_{\mathrm{N}}^{*}\frac{n^{*}}{n_{\mathrm{eff}}^{*}} \tag{A.18}$$

This relation is well-known from ion transport theory, where the only reasonable tracer choice leads to $\Phi^{*} = \Phi$.[237] Eqs. A.14 and A.18 allow the calculation of u and D – and thus also σ – from the respective tracer quantities u_{N}^{*} and D_{N}^{*}, provided that Φ^{*} is known.

It must again be emphasized that Φ^{*}, u_{N}^{*} and D_{N}^{*} are independent from any spatial variations of n^{*} or μ^{*}. Therefore, in contrast to the Fickian diffusion case with $\nabla n \neq 0$ (cf. eq. 2.14), there are no gradients of these quantities in the continuity equation for tracers in an homogeneous system:

$$\frac{\partial n^{*}}{\partial t} = -\nabla \vec{J}_{\mathrm{e}}^{*} = D_{\mathrm{N}}^{*}\nabla^{2} n^{*} - u_{\mathrm{N}}^{*}\nabla\varphi\nabla n^{*} \tag{A.19}$$

$$\text{while} \quad \frac{\partial n}{\partial t} = -\nabla \vec{J}_{\mathrm{e}} = \nabla(D\nabla n) - \nabla\varphi\nabla(n_{\mathrm{eff}} u) \tag{A.20}$$

This is essential for obtaining u_{N}^{*} and D_{N}^{*} – and ultimately u, D and σ – from the tracer displacements $\vec{R}_{k}$, as explained in the following adaptation of the classical random walk theory (e.g., [38]) to electrons.

In the KMC simulations of variable-range hopping, the tracers perform a random walk among the localized states. With an electric field in $-x$-direction ($\partial\varphi/\partial x > 0$), they tend to move in $+x$-direction, but they also exhibit isotropic diffusion in all directions. The tracer movement perpendicular to the electric field is examined first, using the y-direction as example. The tracer displacement in this direction shall be described by a normalized probability density function $p(Y, \tau)$, which means that $p(Y, \tau)\,\mathrm{d}Y$ is the probability that a tracer k has a displacement $Y_{k} \in [Y, Y + \mathrm{d}Y]$ after time τ. Accordingly, $n^{*}(y - Y, t)p(Y, \tau)\,\mathrm{d}Y$ is the amount of tracers that moves from the plane $y - Y$ to the plane y when the time advances from t to $t + \tau$. The tracer concentration in the plane y after time τ is then

$$n^{*}(y, t + \tau) = \int_{-\infty}^{\infty} n^{*}(y - Y, t)p(Y, \tau)\,\mathrm{d}Y \tag{A.21}$$

Note that $p(Y, \tau)$ does not depend on y or t, since the tracers always move under macroscopically

equal conditions. A Taylor expansion (or Kramers–Moyal expansion) of n^* around t and y yields

$$\begin{aligned} n^*(y,t) + \tau\frac{\partial n^*(y,t)}{\partial t} + \cdots = n^*(y,t)\int_{-\infty}^{\infty} p(Y,\tau)\,\mathrm{d}Y \\ - \frac{\partial n^*(y,t)}{\partial y}\int_{-\infty}^{\infty} Y p(Y,\tau)\,\mathrm{d}Y \\ + \frac{\partial^2 n^*(y,t)}{\partial y^2}\int_{-\infty}^{\infty}\frac{Y^2}{2} p(Y,\tau)\,\mathrm{d}Y + \cdots \end{aligned} \tag{A.22}$$

where the first terms cancel each other because $p(Y,\tau)$ is normalized. For sufficiently short τ, the higher order terms on the left-hand side may be omitted and, since $p(Y,\tau)$ must be narrowly localized around $Y = 0$ in this case, the higher order terms on the right-hand side become negligible too. This leads to

$$\frac{\partial n^*(y,t)}{\partial t} = -\frac{\langle Y\rangle_\mathrm{N}}{\tau}\frac{\partial n^*(y,t)}{\partial y} + \frac{\langle Y^2\rangle_\mathrm{N}}{2\tau}\frac{\partial^2 n^*(y,t)}{\partial y^2} \tag{A.23}$$

$$\text{with}\quad \langle Y\rangle_\mathrm{N} = \int_{-\infty}^{\infty} Y p(Y,\tau)\,\mathrm{d}Y \quad\text{and}\quad \langle Y^2\rangle_\mathrm{N} = \int_{-\infty}^{\infty} Y^2 p(Y,\tau)\,\mathrm{d}Y \tag{A.24}$$

where, analog to the previous convention, the subscript "N" symbolizes that p refers to tracers and not effective charge carriers. Since the tracer movement in $+y$- and $-y$-direction has equal probability, $p(Y,\tau)$ must be symmetric around $Y = 0$ and its 1st moment $\langle Y\rangle_\mathrm{N}$ must be zero. Comparison of the remaining

$$\frac{\partial n^*(y,t)}{\partial t} = \frac{\langle Y^2\rangle_\mathrm{N}}{2\tau}\frac{\partial^2 n^*(y,t)}{\partial y^2} \tag{A.25}$$

with the continuity equation for the tracer flux in y-direction (cf. eq. A.19)

$$\frac{\partial n^*(y,t)}{\partial t} = -\frac{\partial J_\mathrm{e}^*(y,t)}{\partial y} = D_{y,\mathrm{N}}^*\frac{\partial^2 n^*(y,t)}{\partial y^2} \tag{A.26}$$

shows that the tracer diffusion coefficient $D_{y,\mathrm{N}}^*$ is connected to the 2nd moment $\langle Y^2\rangle_\mathrm{N}$ by the so-called Einstein-Smoluchowski relation

$$D_{y,\mathrm{N}}^* = \frac{\langle Y^2\rangle_\mathrm{N}}{2\tau} \tag{A.27}$$

The probability density function $p(Y,\tau)$ can be approximated by the distribution of the tracer displacements Y_k, which means that – for a sufficiently high number N^* of representative tracers – $\langle Y^2\rangle_\mathrm{N}$ is given by the mean squared displacement of the tracers (with $N^* = n^* V_\mathrm{cell}$):

$$\langle Y^2\rangle_\mathrm{N} = \frac{1}{N^*}\sum_{k=1}^{N^*} Y_k^2 \tag{A.28}$$

Combining eqs. A.13, A.18, A.27 and A.28 finally yields the Fickian diffusion coefficient (with $N^*_{\mathrm{eff}} = n^*_{\mathrm{eff}} V_{\mathrm{cell}}$):

$$D_y = D^*_{y,\mathrm{N}} \Phi^* = \frac{\langle Y^2 \rangle_{\mathrm{N}} \Phi^*}{2\tau} = \frac{\langle Y^2 \rangle}{2\tau} \tag{A.29}$$

$$\text{with} \quad \langle Y^2 \rangle = \langle Y^2 \rangle_{\mathrm{N}} \Phi^* = \frac{\Phi^*}{N^*} \sum_{k=1}^{N^*} Y_k^2 = \frac{1}{N^*_{\mathrm{eff}}} \sum_{k=1}^{N^*} Y_k^2 \tag{A.30}$$

Dividing the sum of squared tracer displacements not by the number of tracers N^* but by the corresponding number of effective charge carriers N^*_{eff} in the KMC cell allows to obtain the Fickian diffusion coefficient from the simulations.

The same obviously results for the equivalent z-direction and since D_y and D_z are formally equal an average transverse diffusion coefficient D_{yz} can be defined:

$$D_z = \frac{\langle Z^2 \rangle}{2\tau} \quad \text{with} \quad \langle Z^2 \rangle = \frac{1}{N^*_{\mathrm{eff}}} \sum_{k=1}^{N^*} Z_k^2 \quad \Rightarrow \quad D_{yz} = \frac{D_y + D_z}{2} = \frac{\langle Y^2 \rangle + \langle Z^2 \rangle}{4\tau} \tag{A.31}$$

In x-direction, the analog $p(X, \tau)$ would be the probability density function for a certain displacement X through drift and diffusion combined. However, since diffusion has no preferential direction in an isotropic material, it does not affect the distribution's 1st moment

$$\langle X \rangle_{\mathrm{N}} = \int_{-\infty}^{\infty} X p(X, \tau) \, \mathrm{d}X \tag{A.32}$$

With the X_k-distribution of the tracers after time τ as estimate for $p(X, \tau)$, the average drift velocity of the tracers is thus given by (cf. eq. 2.27)

$$\langle v_{\mathrm{d}} \rangle^*_{\mathrm{N}} = \frac{\langle X \rangle_{\mathrm{N}}}{\tau} \quad \text{with} \quad \langle X \rangle_{\mathrm{N}} = \frac{1}{N^*} \sum_{k=1}^{N^*} X_k \tag{A.33}$$

Correspondingly, the average drift velocity of effective charge carriers is (cf. eq. 2.28)

$$\langle v_{\mathrm{d}} \rangle = \langle v_{\mathrm{d}} \rangle^*_{\mathrm{N}} \Phi^* = \frac{\langle X \rangle_{\mathrm{N}} \Phi^*}{\tau} = \frac{\langle X \rangle}{\tau} \tag{A.34}$$

$$\text{with} \quad \langle X \rangle = \langle X \rangle_{\mathrm{N}} \Phi^* = \frac{\Phi^*}{N^*} \sum_{k=1}^{N^*} X_k = \frac{1}{N^*_{\mathrm{eff}}} \sum_{k=1}^{N^*} X_k \tag{A.35}$$

The characterization of the diffusion in x-direction is not as simple, though, because the value of the 2nd moment

$$\langle X^2 \rangle_{\mathrm{N}} = \int_{-\infty}^{\infty} X^2 p(X, \tau) \, \mathrm{d}X \tag{A.36}$$

depends on both drift and diffusion. Moreover, the analog Taylor expansion around x becomes inaccurate when $p(X, \tau)$ is narrowly localized around $X = \langle X \rangle_\mathrm{N}$ instead of $X = 0$. These issues can be solved when $p(X, \tau)$ is symmetric around $\langle X \rangle_\mathrm{N}$. In this case, drift and diffusion may be considered independent: Each tracer moves by $\langle X \rangle_\mathrm{N} = \langle v_\mathrm{d} \rangle_\mathrm{N}^* \tau$ in the time τ due to the electric field and every deviation of X_k from $\langle X \rangle_\mathrm{N}$ is interpreted as diffusion. However, this approximation is only valid for electrons that actually exhibit a drift movement. For electrons far below E_F, for example, the difference between X_k and $\langle X \rangle_\mathrm{N}$ does not originate from diffusion but from their inability to reach an unoccupied state. Their movement is blocked by the other electrons (Pauli principle). Such immobilized electrons are not eligible as representative tracers. This led to the following tracer choice in the KMC simulations: N^* corresponds to the so-called "mobile" electrons N_mob in the KMC cell, which includes all electrons that are neither stationary (= zero hops) nor purely oscillating (= only oscillating hops). This constraint does not affect the corresponding number of effective charge carriers N_eff^*, because the excluded electrons all have no contribution to any charge transport. The appropriate N_eff^* for this tracer choice therefore is the total number of effective charge carriers N_eff in the KMC cell, which is easily obtained from eq. 2.23 for n_eff (or rather eq. 4.18).[125]

With the above approximation, the diffusive motion of the tracers can be described on a moving spatial axis $\hat{x} = x - \langle v_\mathrm{d} \rangle_\mathrm{N}^* t$. This corresponds to a Galilean transformation from an Eulerian to a Lagrangian frame of reference, which is inertial in this case because $\langle v_\mathrm{d} \rangle_\mathrm{N}^*$ is constant. The displacement $\hat{X} = X - \langle v_\mathrm{d} \rangle_\mathrm{N}^* \tau = X - \langle X \rangle_\mathrm{N}$ on the $\hat{x}$-axis in a time interval τ is then purely due to diffusion. Analog to the approach for the y- and z-axes before, it shall be described by a normalized probability density function $p(\hat{X}, \tau)$, which means that $p(\hat{X}, \tau)\,\mathrm{d}\hat{X}$ is the probability that a tracer deviates from the average displacement by $\hat{X}_k \in [\hat{X}, \hat{X} + \mathrm{d}\hat{X}]$ after time τ. For a moving observer (= stationary w.r.t. $\hat{x}$-axis), the tracer concentration in a plane $\hat{x}$ after time τ is thus

$$n^*(\hat{x}, t + \tau) = \int_{-\infty}^{\infty} n^*(\hat{x} - \hat{X}, t)\, p(\hat{X}, \tau)\, \mathrm{d}\hat{X} \tag{A.37}$$

With an analog reasoning as for Y and Z, a Taylor expansion around $\hat{x}$ and t then yields (with $\langle \hat{X} \rangle_\mathrm{N} = 0$)

$$\frac{\partial n^*(\hat{x}, t)}{\partial t} = \frac{\langle \hat{X}^2 \rangle_\mathrm{N}}{2\tau} \frac{\partial^2 n^*(\hat{x}, t)}{\partial \hat{x}^2} \quad \text{with} \quad \langle \hat{X}^2 \rangle_\mathrm{N} = \int_{-\infty}^{\infty} \hat{X}^2 p(\hat{X}, \tau)\, \mathrm{d}\hat{X} \tag{A.38}$$

Upon the transformation back into the static frame of reference, it must be taken into account that, with regard to the x-axis, $\hat{x}$ depends on t. The change of n^* with time is then given by its total derivative, which is also called material derivative in this case. The spatial derivatives can be transformed by $\partial/\partial\hat{x} = \partial/\partial x$, since they are related just by a shift without deformation ($\partial\hat{x}/\partial x = 1$). For a stationary observer (= stationary w.r.t. x-axis), the multivariate chain rule together with the

[125] Note that this specific tracer choice is none of the aforementioned examples with $\Phi^* = \Phi = N/N_\mathrm{eff}$ or $\Phi^* = 1$. Instead, it is $\Phi^* = N_\mathrm{mob}/N_\mathrm{eff}$.

above partial derivative yields

$$\begin{aligned}\frac{\mathrm{d}n^*(\hat{x},t)}{\mathrm{d}t} &= \frac{\mathrm{d}n^*(x-\langle v_\mathrm{d}\rangle_\mathrm{N}^* t,t)}{\mathrm{d}t} = \frac{\partial(x-\langle v_\mathrm{d}\rangle_\mathrm{N}^* t)}{\partial t}\frac{\partial n^*(\hat{x},t)}{\partial x} + \frac{\partial n^*(\hat{x},t)}{\partial t} \\ &= -\langle v_\mathrm{d}\rangle_\mathrm{N}^* \frac{\partial n^*(\hat{x},t)}{\partial x} + \frac{\langle \hat{X}^2\rangle_\mathrm{N}}{2\tau}\frac{\partial^2 n^*(\hat{x},t)}{\partial x^2}\end{aligned} \tag{A.39}$$

Since $\hat{X}$ is the deviation from the average x-displacement, $\langle \hat{X}^2\rangle_\mathrm{N}$ is the variance of X:

$$\mathrm{Var}(X)_\mathrm{N} := \langle \hat{X}^2\rangle_\mathrm{N} = \langle (X-\langle X\rangle_\mathrm{N})^2\rangle_\mathrm{N} \tag{A.40}$$

Eq. A.39 describes the change of the tracer concentration at a location $\hat{x}$, but since the same equation applies everywhere in the material – with the same, x- and t-independent $\langle v_\mathrm{d}\rangle_\mathrm{N}^*$ and $\mathrm{Var}(X)_\mathrm{N}$ – it can also be written with $n^*(\hat{x},t)$ replaced by $n^*(x,t)$:

$$\begin{aligned}\frac{\mathrm{d}n^*(x,t)}{\mathrm{d}t} &= \frac{\partial n^*(x,t)}{\partial t} = -\langle v_\mathrm{d}\rangle_\mathrm{N}^* \frac{\partial n^*(x,t)}{\partial x} + \frac{\langle \hat{X}^2\rangle_\mathrm{N}}{2\tau}\frac{\partial^2 n^*(x,t)}{\partial x^2} \\ &= -\frac{\langle X\rangle_\mathrm{N}}{\tau}\frac{\partial n^*(x,t)}{\partial x} + \frac{\mathrm{Var}(X)_\mathrm{N}}{2\tau}\frac{\partial^2 n^*(x,t)}{\partial x^2}\end{aligned} \tag{A.41}$$

Comparison with the continuity equation for the tracer flux in x-direction (cf. eq. A.19)

$$\frac{\partial n^*(x,t)}{\partial t} = -\frac{\partial J_\mathrm{e}^*(x,t)}{\partial x} = -u_\mathrm{N}^* F\frac{\partial n^*(x,t)}{\partial x} + D_{x,\mathrm{N}}^*\frac{\partial^2 n^*(x,t)}{\partial x^2} \tag{A.42}$$

and taking the $X-k$-distribution as estimate for the respective probability density function yields the tracer mobility and the tracer diffusion coefficient according to

$$u_\mathrm{N}^* = \frac{\langle X\rangle_\mathrm{N}}{F\tau} \quad \text{and} \quad D_{x,\mathrm{N}}^* = \frac{\mathrm{Var}(X)_\mathrm{N}}{2\tau} \tag{A.43}$$

$$\text{with} \quad \langle X\rangle_\mathrm{N} = \frac{1}{N^*}\sum_{k=1}^{N^*} X_k \quad \text{and} \quad \mathrm{Var}(X)_\mathrm{N} = \frac{1}{N^*}\sum_{k=1}^{N^*}(X_k-\langle X\rangle_\mathrm{N})^2 \tag{A.44}$$

The same well-known result can also be derived from the Langevin equation or the Fokker-Planck equation.[238] Finally, the mobility u and the Fickian diffusion coefficient D_x can be obtained by combining eqs. A.14, A.18, A.43 and A.44:

$$u = u_\mathrm{N}^*\Phi^* = \frac{\langle X\rangle_\mathrm{N}\Phi^*}{F\tau} = \frac{\langle X\rangle}{F\tau} \quad \text{with} \quad \langle X\rangle = \langle X\rangle_\mathrm{N}\Phi^* = \frac{\Phi^*}{N^*}\sum_{k=1}^{N^*} X_k = \frac{1}{N_\mathrm{eff}^*}\sum_{k=1}^{N^*} X_k \tag{A.45}$$

$$D_x = D_{x,\mathrm{N}}^*\Phi^* = \frac{\mathrm{Var}(X)_\mathrm{N}\Phi^*}{2\tau} = \frac{\mathrm{Var}(X)}{2\tau} \tag{A.46}$$

$$\begin{aligned}\text{with} \quad \mathrm{Var}(X) &= \mathrm{Var}(X)_\mathrm{N}\Phi^* = \frac{\Phi^*}{N^*}\sum_{k=1}^{N^*}(X_k-\langle X\rangle_\mathrm{N})^2 = \frac{1}{N_\mathrm{eff}^*}\sum_{k=1}^{N^*}(X_k-\langle X\rangle_\mathrm{N})^2 \\ &= \langle X^2\rangle - \frac{N_\mathrm{eff}^*}{N^*}\langle X\rangle^2\end{aligned} \tag{A.47}$$

$$\text{where} \quad \langle X^2 \rangle = \frac{1}{N^*_{\text{eff}}} \sum_{k=1}^{N^*} X_k^2 \quad (= \langle X^2 \rangle_{\text{N}} \Phi^*) \tag{A.48}$$

At first sight, it seems rather strange that eqs. A.46 and A.47 for the parallel Fickian diffusion coefficient D_x still contain an explicit dependence on N^*, which is not the case for u, D_y or D_z. However, the term $X_k - \langle X \rangle_{\text{N}}$ describes each tracer's deviation from its expectation value $\langle X \rangle_{\text{N}}$ and only the collective variance of all representative tracers can be reinterpreted as diffusion of effective charge carriers in x-direction. The validity of the artificial separation of drift and diffusion in this approach can be checked by comparing D_x and D_{yz}.

For the aforementioned, specific tracer choice that is used in this work, it is $n^*_{\text{eff}} = n_{\text{eff}}$ and thus $\sigma^* = \sigma$ according to eq. A.14. From the combination of eqs. A.14 and A.45 therefore follows that the VRH conductivity σ is given by

$$\sigma = \sigma^* = e n^*_{\text{eff}} \frac{\langle X \rangle}{F\tau} = \frac{e}{F\tau} \frac{n^*_{\text{eff}}}{N^*_{\text{eff}}} \sum_{k=1}^{N^*} X_k = \frac{e}{F\tau V_{\text{cell}}} \sum_{k=1}^{N^*} X_k \tag{A.49}$$

Note that this equation for σ has no longer an explicit dependence on N^*_{eff} or N_{eff}. Therefore, as long as one is only concerned with the conductivity, a proper definition of the charge carrier density for VRH is not necessary. This is probably the major reason why the effective charge carrier model or any likewise treatment was not found in previous KMC-VRH studies.

A.5 Variational Formulation of the FEM Models

This section contains the model equations that were supplied to the FEniCS software for the FEM simulations. The common notation for FEM models is adopted, where u is a spatially varying quantity that should be determined and where the test functions are denoted by v.

The variational formulation of the models was derived by multiplying the respective partial differential equation(s) of sec. 4.3 with the test function(s) and integrating over the model volume V. For the one-dimensional simulations carried out in this work, it is $\mathrm{d}V = \mathrm{d}z$ and $\nabla = \partial/\partial z$. Green's first identity, which here corresponds to integration by parts, was used to reduce the order of the differentials (resulting for example in $\nabla u \nabla v$ terms instead of $v\nabla^2 u$). The surface integrals in Green's first identity are zero because of the Dirichlet boundary conditions (since the test functions must have $v = 0$ at boundaries where u is known).

The Jacobians required for the iterative Newton solver were derived as Gateaux derivatives with $u = \bar{u} + \gamma \delta u$ in the limit of $\gamma \to 0$, where $\bar{u}$ is the result of the previous iteration and where the trial function δu represents the change of u in the current iteration step. The derivatives of the charge density and the conductivity with regard to γ that occurred in this context were transformed into derivatives with regard to μ and $|\nabla \varphi|$ by using the chain rule.

Model "FEM-1": $V = 0$ without high-field effects

The searched function is $u = \mu\ (= e\varphi)$. The respective trial function in the Jacobian has the symbol δu. The test function for the differential equation is denoted by v. The data set $\rho(u)$ was supplied as input from DFT. The required derivative $\partial\rho/\partial u$ as a function of u was calculated from the input data through numerical differentiation.

Variational formulation of FEM-1:

$$\begin{aligned} &\int v\nabla^2 u\,\mathrm{d}V + \int v\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\rho(u)\,\mathrm{d}V = 0 \\ \Rightarrow\quad &\int \nabla v\nabla u\,\mathrm{d}V - \int v\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\rho(u)\,\mathrm{d}V = 0 \end{aligned} \tag{A.50}$$

Jacobian of FEM-1:

$$\begin{aligned} &\lim_{\gamma\to 0}\frac{\partial}{\partial\gamma}\left[\int \nabla v\nabla(\bar{u}+\gamma\delta u)\,\mathrm{d}V - \int v\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\rho(\bar{u}+\gamma\delta u)\,\mathrm{d}V\right] \\ &= \int \nabla v\nabla\delta u\,\mathrm{d}V - \int v\delta u\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\left.\frac{\partial\rho(u)}{\partial u}\right|_{u=\bar{u}}\mathrm{d}V \end{aligned} \tag{A.51}$$

Model "FEM-2": $V = 0$ with high-field effects

The searched function is $u = \mu\ (= e\varphi)$. The respective trial function in the Jacobian has the symbol δu. The test function for the differential equation is denoted by v. The data set $\rho(u, |\nabla u|/e)$ was supplied as input from DFT and KMC. The required derivatives $\partial\rho/\partial u$ and $\partial\rho/\partial|\nabla u|$ as functions of u and $|\nabla u|/e$ were calculated from the input data through numerical differentiation.

Variational formulation of FEM-2:

$$\begin{aligned} &\int v\nabla^2 u\,\mathrm{d}V + \int v\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\rho\left(u, \frac{|\nabla u|}{e}\right)\mathrm{d}V = 0 \\ \Rightarrow\quad &\int \nabla v\nabla u\,\mathrm{d}V - \int v\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\rho\left(u, \frac{|\nabla u|}{e}\right)\mathrm{d}V = 0 \end{aligned} \tag{A.52}$$

Jacobian of FEM-2:

$$\begin{aligned} &\lim_{\gamma\to 0}\frac{\partial}{\partial\gamma}\left[\int \nabla v\nabla(\bar{u}+\gamma\delta u)\,\mathrm{d}V - \int v\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\rho\left(\bar{u}+\gamma\delta u, \frac{|\nabla(\bar{u}+\gamma\delta u)|}{e}\right)\mathrm{d}V\right] \\ &= \int \nabla v\nabla\delta u\,\mathrm{d}V - \int v\delta u\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\left.\frac{\partial\rho\left(u, \frac{|\nabla\bar{u}|}{e}\right)}{\partial u}\right|_{u=\bar{u}}\mathrm{d}V \\ &\qquad - \int v\nabla\delta u\frac{e}{\epsilon_0\epsilon_\mathrm{r}}\left.\frac{\partial\rho\left(\bar{u}, \frac{|\nabla u|}{e}\right)}{\partial|\nabla u|}\right|_{\nabla u=\nabla\bar{u}}\mathrm{d}V \end{aligned} \tag{A.53}$$

Model "FEM-3": $V \neq 0$ without high-field effects

The searched functions are $u_1 = \mu$ and $u_2 = \varphi$. The respective trial functions in the Jacobian have the symbols δu_1 and δu_2. The test functions for the two differential equations are denoted by ν_1 and ν_2. The data sets $\rho(u_1)$ and $\sigma(u_1)$ were supplied as input from DFT and KMC. The required derivatives $\partial\rho/\partial u_1$ and $\partial\sigma/\partial u_1$ as functions of u_1 were calculated from the input data through numerical differentiation.

Variational formulation of FEM-3:

$$\begin{aligned}
&\int \nu_1 \nabla \left[\frac{\sigma(u_1)}{e} \nabla u_1 \right] \mathrm{d}V - \int \nu_1 \nabla \left[\sigma(u_1) \nabla u_2 \right] \mathrm{d}V + \int \nu_2 \nabla^2 u_2 \, \mathrm{d}V \\
&\quad + \int \nu_2 \frac{1}{\epsilon_0 \epsilon_\mathrm{r}} \rho(u_1) \, \mathrm{d}V = 0 \\
\Rightarrow \quad &\int \frac{\sigma(u_1)}{e} \nabla \nu_1 \nabla u_1 \, \mathrm{d}V - \int \sigma(u_1) \nabla \nu_1 \nabla u_2 \, \mathrm{d}V + \int \nabla \nu_2 \nabla u_2 \, \mathrm{d}V \\
&\qquad - \int \nu_2 \frac{1}{\epsilon_0 \epsilon_\mathrm{r}} \rho(u_1) \, \mathrm{d}V = 0
\end{aligned} \tag{A.54}$$

Jacobian of FEM-3:

$$\begin{aligned}
&\lim_{\gamma_1 \to 0} \frac{\partial}{\partial \gamma_1} \left[\int \frac{\sigma(\bar{u}_1 + \gamma_1 \delta u_1)}{e} \nabla \nu_1 \nabla (\bar{u}_1 + \gamma_1 \delta u_1) \, \mathrm{d}V - \int \sigma(\bar{u}_1 + \gamma_1 \delta u_1) \nabla \nu_1 \nabla \bar{u}_2 \, \mathrm{d}V \right. \\
&\quad \left. - \int \nu_2 \frac{1}{\epsilon_0 \epsilon_\mathrm{r}} \rho(\bar{u}_1 + \gamma_1 \delta u_1) \, \mathrm{d}V \right] + \lim_{\gamma_2 \to 0} \frac{\partial}{\partial \gamma_2} \left[\int \nabla \nu_2 \nabla (\bar{u}_2 + \gamma_2 \delta u_2) \, \mathrm{d}V \right. \\
&\quad \left. - \int \sigma(\bar{u}_1) \nabla \nu_1 \nabla (\bar{u}_2 + \gamma_2 \delta u_2) \, \mathrm{d}V \right] \\
&= \int \frac{1}{e} \left. \frac{\partial \sigma(u_1)}{\partial u_1} \right|_{u_1 = \bar{u}_1} \delta u_1 \nabla \nu_1 \nabla \bar{u}_1 \, \mathrm{d}V + \int \frac{\sigma(\bar{u}_1)}{e} \nabla \nu_1 \nabla \delta u_1 \, \mathrm{d}V \\
&\quad - \int \left. \frac{\partial \sigma(u_1)}{\partial u_1} \right|_{u_1 = \bar{u}_1} \delta u_1 \nabla \nu_1 \nabla \bar{u}_2 \, \mathrm{d}V - \int \nu_2 \frac{1}{\epsilon_0 \epsilon_\mathrm{r}} \left. \frac{\partial \rho(u_1)}{\partial u_1} \right|_{u_1 = \bar{u}_1} \delta u_1 \, \mathrm{d}V \\
&\quad + \int \nabla \nu_2 \nabla \delta u_2 \, \mathrm{d}V - \int \sigma(\bar{u}_1) \nabla \nu_1 \nabla \delta u_2 \, \mathrm{d}V
\end{aligned} \tag{A.55}$$

Model "FEM-4": $V \neq 0$ with high-field effects

The searched functions are $u_1 = \mu$ and $u_2 = \varphi$. The respective trial functions in the Jacobian have the symbols δu_1 and δu_2. The test functions for the two differential equations are denoted by ν_1 and ν_2. The data sets $\rho(u_1, |\nabla u_2|)$ and $\sigma(u_1, |\nabla u_2|)$ were supplied as input from DFT and KMC. The required derivatives $\partial\rho/\partial u_1$, $\partial\rho/\partial|\nabla u_2|$, $\partial\sigma/\partial u_1$ and $\partial\sigma/\partial|\nabla u_2|$ as functions of u_1 and $|\nabla u_2|$ were calculated from the input data through numerical differentiation.

Variational formulation of FEM-4:

$$\begin{aligned}
&\int \nu_1 \nabla \left[\frac{\sigma(u_1, |\nabla u_2|)}{e} \nabla u_1 \right] \mathrm{d}V - \int \nu_1 \nabla \left[\sigma(u_1, |\nabla u_2|) \nabla u_2 \right] \mathrm{d}V \\
&\quad + \int \nu_2 \nabla^2 u_2 \, \mathrm{d}V + \int \nu_2 \frac{1}{\epsilon_0 \epsilon_\mathrm{r}} \rho(u_1, |\nabla u_2|) \, \mathrm{d}V = 0 \\
\Rightarrow \quad &\int \frac{\sigma(u_1, |\nabla u_2|)}{e} \nabla \nu_1 \nabla u_1 \, \mathrm{d}V - \int \sigma(u_1, |\nabla u_2|) \nabla \nu_1 \nabla u_2 \, \mathrm{d}V \\
&\quad + \int \nabla \nu_2 \nabla u_2 \, \mathrm{d}V - \int \nu_2 \frac{1}{\epsilon_0 \epsilon_\mathrm{r}} \rho(u_1, |\nabla u_2|) \, \mathrm{d}V = 0
\end{aligned} \tag{A.56}$$

Jacobian of FEM-4:

$$\begin{aligned}
&\lim_{\gamma_1 \to 0} \frac{\partial}{\partial \gamma_1} \left[\int \frac{\sigma(\bar{u}_1 + \gamma_1 \delta u_1, |\nabla \bar{u}_2|)}{e} \nabla \nu_1 \nabla (\bar{u}_1 + \gamma_1 \delta u_1) \, \mathrm{d}V \right. \\
&\quad - \int \sigma(\bar{u}_1 + \gamma_1 \delta u_1, |\nabla \bar{u}_2|) \nabla \nu_1 \nabla \bar{u}_2 \, \mathrm{d}V - \left. \int \nu_2 \frac{1}{\epsilon_0 \epsilon_\mathrm{r}} \rho(\bar{u}_1 + \gamma_1 \delta u_1, |\nabla \bar{u}_2|) \, \mathrm{d}V \right] \\
&\quad + \lim_{\gamma_2 \to 0} \frac{\partial}{\partial \gamma_2} \left[\int \frac{\sigma(\bar{u}_1, |\nabla (\bar{u}_2 + \gamma_2 \delta u_2)|)}{e} \nabla \nu_1 \nabla \bar{u}_1 \, \mathrm{d}V \right. \\
&\quad - \int \sigma(\bar{u}_1, |\nabla (\bar{u}_2 + \gamma_2 \delta u_2)|) \nabla \nu_1 \nabla (\bar{u}_2 + \gamma_2 \delta u_2) \, \mathrm{d}V \\
&\quad + \left. \int \nabla \nu_2 \nabla (\bar{u}_2 + \gamma_2 \delta u_2) \, \mathrm{d}V - \int \nu_2 \frac{1}{\epsilon_0 \epsilon_\mathrm{r}} \rho(\bar{u}_1, |\nabla (\bar{u}_2 + \gamma_2 \delta u_2)|) \, \mathrm{d}V \right] \\
&= \int \frac{1}{e} \left. \frac{\partial \sigma(u_1, |\nabla \bar{u}_2|)}{\partial u_1} \right|_{u_1 = \bar{u}_1} \delta u_1 \nabla \nu_1 \nabla \bar{u}_1 \, \mathrm{d}V + \int \frac{\sigma(\bar{u}_1, |\nabla \bar{u}_2|)}{e} \nabla \nu_1 \nabla \delta u_1 \, \mathrm{d}V \\
&\quad - \int \left. \frac{\partial \sigma(u_1, |\nabla \bar{u}_2|)}{\partial u_1} \right|_{u_1 = \bar{u}_1} \delta u_1 \nabla \nu_1 \nabla \bar{u}_2 \, \mathrm{d}V - \int \nu_2 \frac{1}{\epsilon_0 \epsilon_\mathrm{r}} \left. \frac{\partial \rho(u_1, |\nabla \bar{u}_2|)}{\partial u_1} \right|_{u_1 = \bar{u}_1} \delta u_1 \, \mathrm{d}V \\
&\quad + \int \frac{1}{e} \left. \frac{\partial \sigma(\bar{u}_1, |\nabla u_2|)}{\partial |\nabla u_2|} \right|_{\nabla u_2 = \nabla \bar{u}_2} |\nabla \delta u_2| \nabla \nu_1 \nabla \bar{u}_1 \, \mathrm{d}V \\
&\quad - \int \left. \frac{\partial \sigma(\bar{u}_1, |\nabla u_2|)}{\partial |\nabla u_2|} \right|_{\nabla u_2 = \nabla \bar{u}_2} |\nabla \delta u_2| \nabla \nu_1 \nabla \bar{u}_2 \, \mathrm{d}V - \int \sigma(\bar{u}_1, |\nabla \bar{u}_2|) \nabla \nu_1 \nabla \delta u_2 \, \mathrm{d}V \\
&\quad + \int \nabla \nu_2 \nabla \delta u_2 \, \mathrm{d}V - \int \nu_2 \frac{1}{\epsilon_0 \epsilon_\mathrm{r}} \left. \frac{\partial \rho(\bar{u}_1, |\nabla u_2|)}{\partial |\nabla u_2|} \right|_{\nabla u_2 = \nabla \bar{u}_2} |\nabla \delta u_2| \, \mathrm{d}V
\end{aligned} \tag{A.57}$$

B

Figures

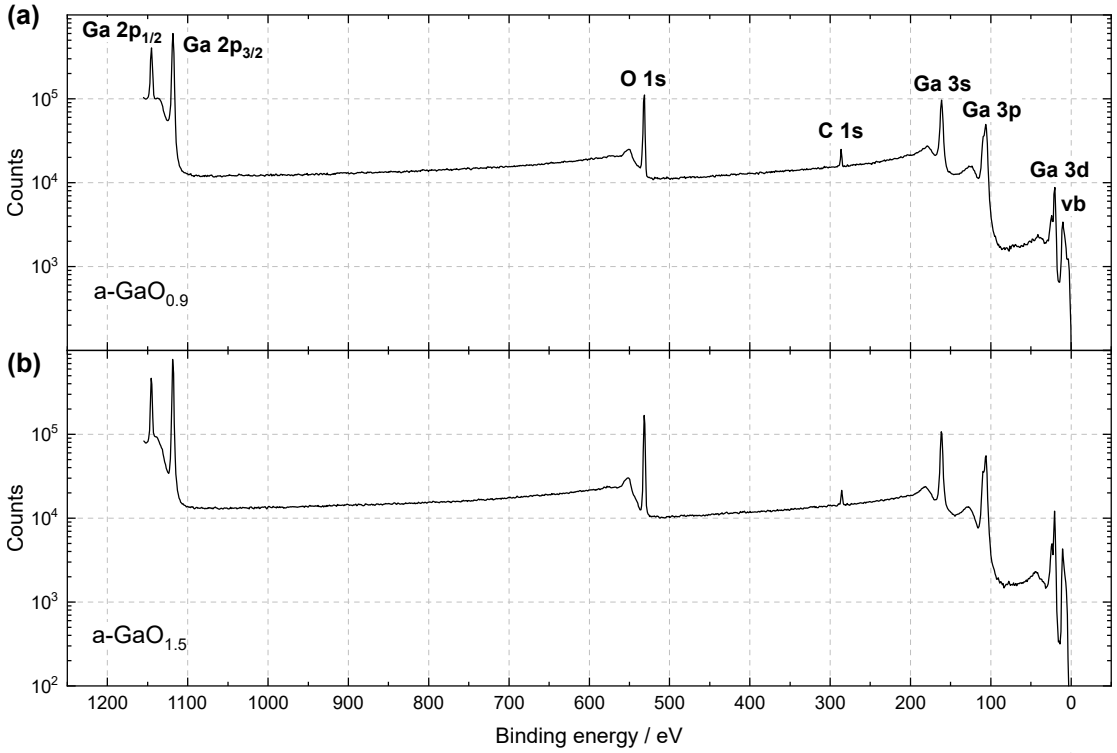


Figure B.1: Exemplary full HAXPES spectra, with peak assignment according to previous studies (e.g., [185]). (a) a-GaO_x with $x \approx 0.9$, deposited with 0.79 J/cm^2 in Ar (same sample as in figs. 5.5a, B.2 and B.3). (b) a-$GaO_{1.5}$, deposited in O_2. There is no visible difference between the two spectra except for the valence band region, which is analyzed more closely in fig. 5.34. The unlabeled, shallow peaks behind the Ga and O peaks are the associated plasmon loss peaks.

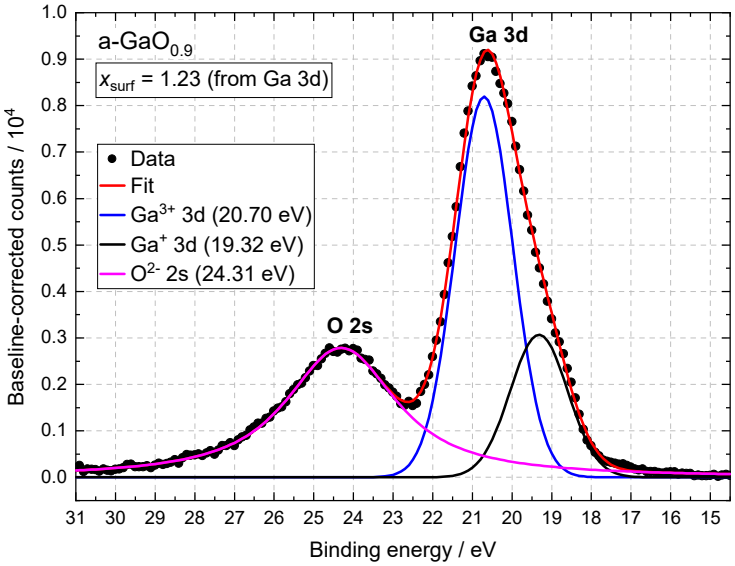


Figure B.2: Exemplary HAXPES Ga 3d signal for a-GaO_x with $x \approx 0.9$ (PLD with 0.79 J/cm^2 in Ar; same sample as in figs. 5.5a, B.1a and B.3). The peak positions for the two Ga^{3+} and Ga^+ sub-peaks and for the third Voigt peak of the overlapping O 2s signal are in excellent agreement with previous works.[10, 91, 115, 184] A surface stoichiometry of $x_{surf} = 1.23$ was calculated from the Ga 3d sub-peak area ratio, which is lower than the value $x_{surf} = 1.33$ that was obtained from the Ga $2p_{3/2}$ peak of the same sample (fig. 5.5a).

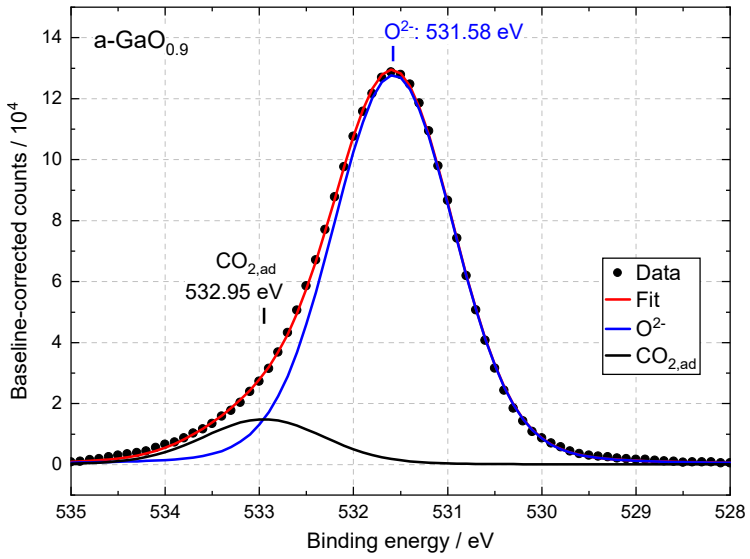


Figure B.3: Exemplary HAXPES O 1s signal for a-GaO_x with $x \approx 0.9$ (PLD with 0.79 J/cm^2 in Ar; same sample as in figs. 5.5a, B.1a and B.2). The main peak corresponds to oxygen bound to gallium,[104] while the overlapping sub-peak was attributed to carbonyl oxygen in adsorbed CO_2 (observed at 532.1 eV or 532.9 eV by XPS in [90, 92]).

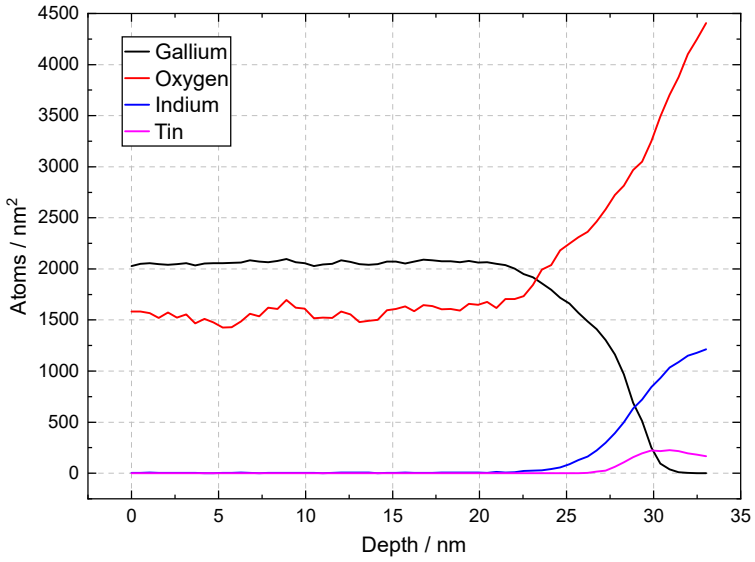


Figure B.4: EELS analysis of an a-GaO_x|ITO interface. The a-GaO_x layer was deposited with 0.49 J/cm² in Ar and has a bulk O/Ga ratio of 0.82–0.95 according to EPMA and EELS (same sample as in fig. 5.6). The left side of the graph belongs to bulk gallium oxide, while the ITO substrate is on the right side.

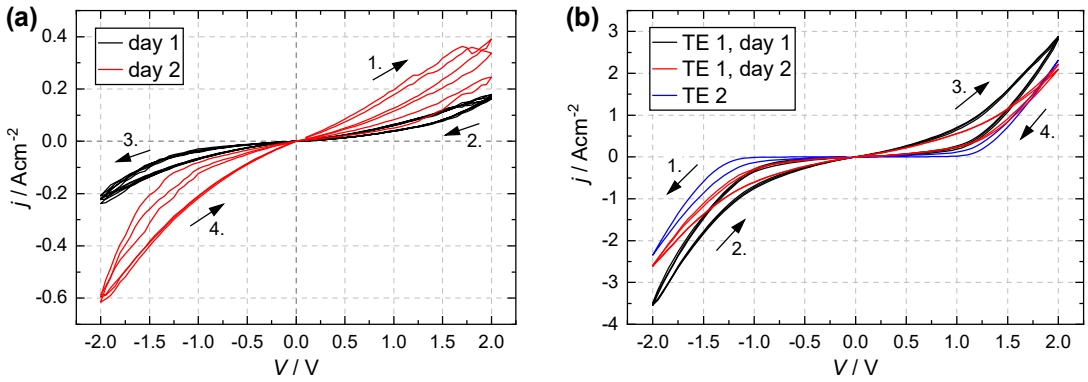


Figure B.5: Influence of the contact quality on the shape of the RS curves of different Pt|a-GaO_x|ITO devices. (a) Comparison between *I*-*V* curves measured on day 1 (v_{sw} = 0.1 V/s) with those measured on day 2 (v_{sw} = 0.2 V/s) after re-contacting of an RS device ($x \approx 0.9$, $d \approx 90$ nm, PLD with 0.78 J/cm² in Ar, chamber #1; sp-Pt-TE-∅: 300 µm, CT3, V_{max} = ±2.0 V). (b) *I*-*V* curves for two devices next to each other on one sample (TEs 1 and 2), together with measurements after re-contacting on the next day ($x \approx 0.9$, $d \approx 60$ nm, PLD with 0.80 J/cm² in Ar, chamber #1; sp-Pt-TE-∅: 400 µm, CT2, reversed sweep direction: 0 V → −2 V → 0 V → +2 V → 0 V, v_{sw} = 0.05 V/s).

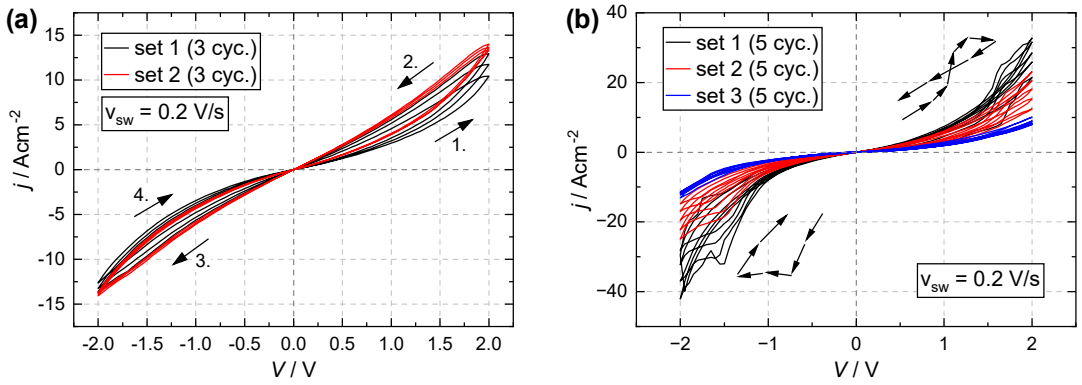


Figure B.6: Examples for the gradual variation of the *I*-*V* sweeps from measurements of two different Pt|a-GaO_x|ITO devices on the same sample (x: 0.98–1.07 by EELS, $d \approx 60$ nm, PLD with 0.79 J/cm^2 in Ar; lithographic Pt-TE, CT4, $V_{max} = \pm 2.0$ V). (a) Current increases from cycle to cycle (TE: 50 × 50 µm^2). (b) Current decreases from cycle to cycle and there is a loop at positive voltages (TE: 100 × 100 µm^2).

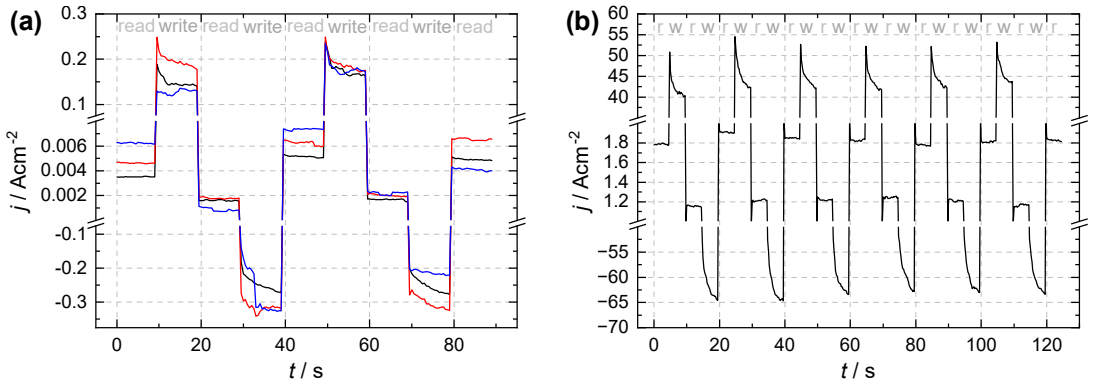


Figure B.7: Pulsed counter-8 bipolar RS of Pt|a-GaO_x|ITO devices on two different samples (contacted by W or Au probe tips): (a) $x \approx 0.9$, $d \approx 90$ nm (PLD with 0.78 J/cm^2 in Ar, chamber #1), with sp-Pt-TE (same device as in fig. 5.7b, TE-∅: 300 µm, CT3, $V_{write} = \pm 2.0$ V, $V_{read} = +0.1$ V, $\Delta t = 10$ s). The graph shows three consecutive pulse sequences (black = 1st, red = 2nd, blue = 3rd). (b) $x \approx 1.0$, $d \approx 120$ nm (PLD with 0.79 J/cm^2 in Ar), with lithographic Pt-TE (same device as in fig. 5.7d, TE: 20×20 µm^2, CT4, $V_{write} = \pm 2.25$ V, $V_{read} = +0.1$ V, $\Delta t = 5$ s).

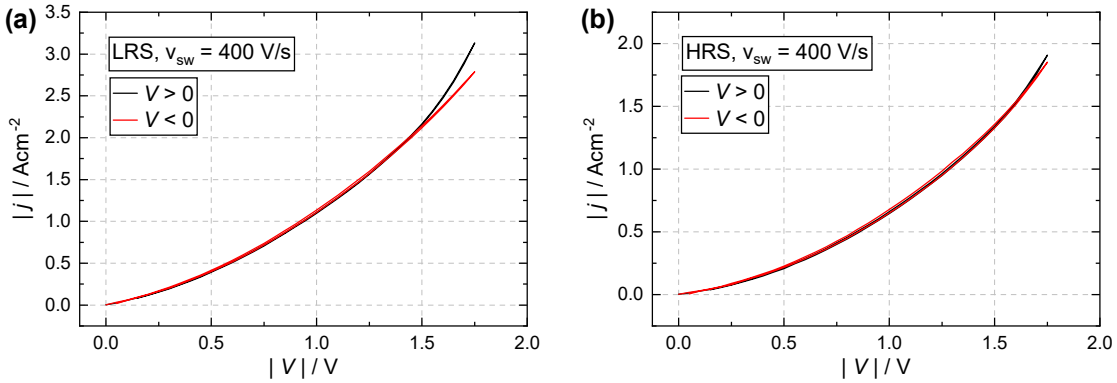


Figure B.8: Visualization of the symmetry of the fast *I*-*V* sweeps from fig. 5.12b, which represent the electron transport in an LRS and an HRS of Pt|a-GaO_x|ITO without further counter-8 resistive switching ($x \approx 0.9$, $d \approx 90$ nm, PLD with 0.78 J/cm^2 in Ar, chamber #1; CT5, TE area: 0.465 mm^2, $V_{max} = \pm 1.75$ V). These cycles were measured 3 min after the respective state was created by the constant ± − 1.75 V shown in fig. 5.12a.

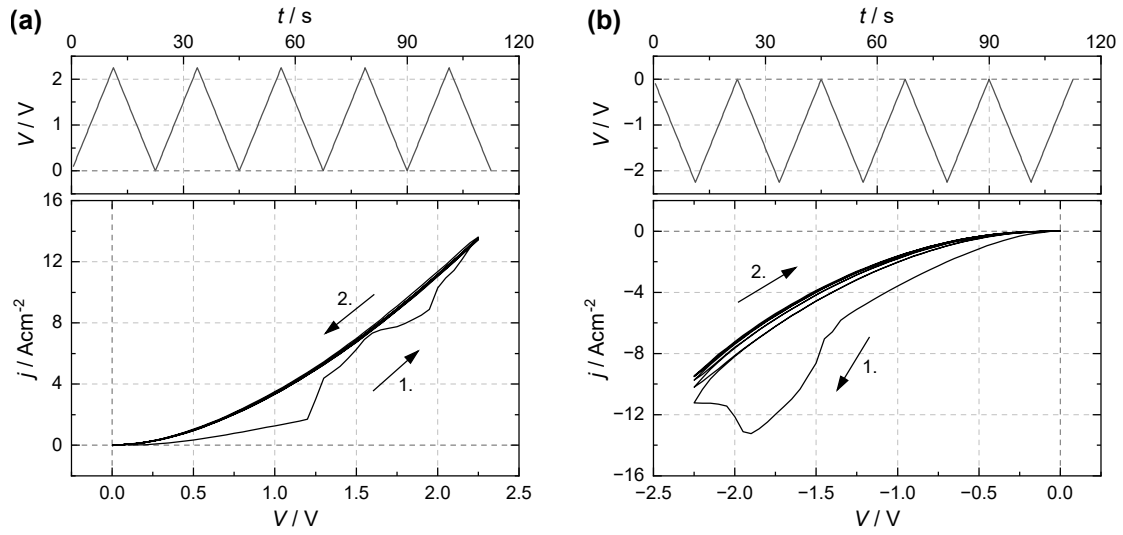


Figure B.9: Bipolarity check for figure-8 RS of a Pt|a-GaO_x|ITO device with $x \lesssim 0.8$ and $d \approx 50$ nm (same as in fig. 5.15a; PLD with less than 0.20 J/cm^2 in Ar, chamber #1, eb-Pt-TE-∅: 100 µm, CT3). The top graphs show the applied voltage $V(t)$, while the bottom graphs display the resulting I-V curve. (a) Triangular I-V sweep with only positive voltage ($V_{max} = +2.25$ V, $v_{sw} = 0.2$ V/s). (b) Triangular I-V sweep with only negative voltage ($V_{max} = -2.25$ V, $v_{sw} = 0.2$ V/s). In each graph, the numbered arrows refer to the first 0 V → V_{max} → 0 V cycle, which is the only one with a notable switching effect.

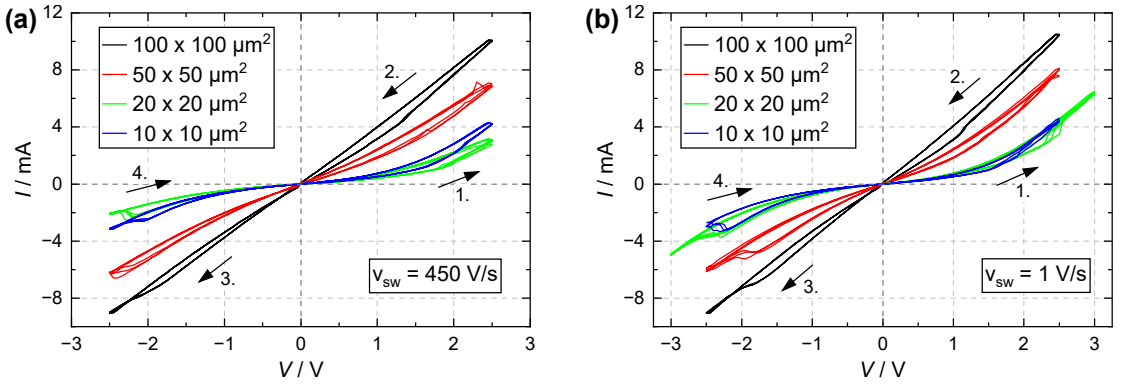


Figure B.10: Comparison of the figure-8 bipolar RS of four Pt|a-GaO_x|ITO devices with different TE sizes on the same sample (same devices as in fig. 5.19; $x \approx 1.1$, $d \approx 130$ nm, PLD with 4.90 J/cm^2 in Ar; lithographic Pt-TEs, CT2, $V_{max} = \pm 2.5$ V). (a) Resistive switching at high sweep rate (same measurements as in fig. 5.19 but here the current is shown instead of the current density). (b) Resistive switching of the same devices at low sweep rate. The device with a TE size of 50 × 50 µm^2 was not measured at low sweep rate for $V_{max} = \pm 2.5$ V, which is why the slow sweeps for $V_{max} = \pm 3.0$ V are shown instead.

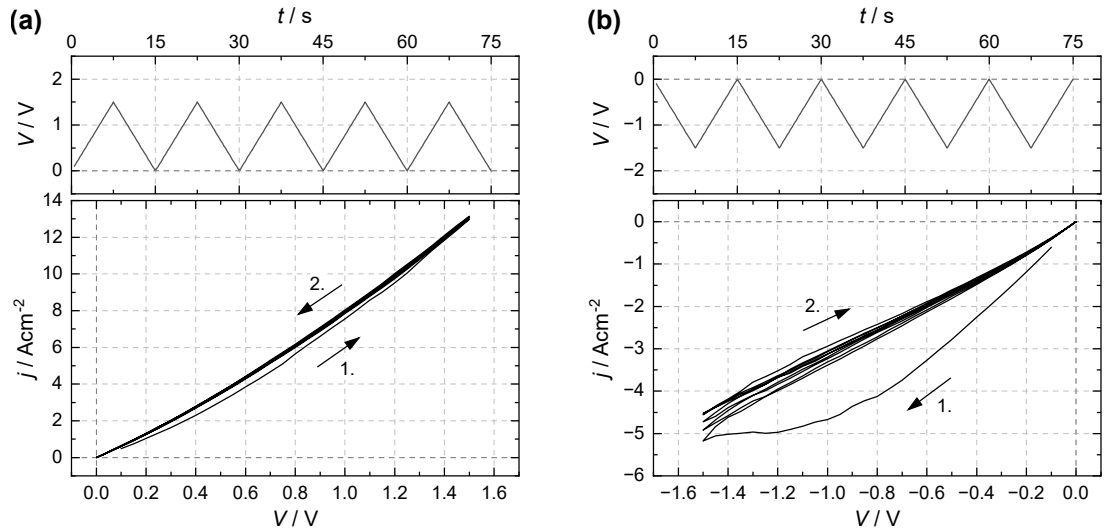


Figure B.11: Bipolarity check for figure-8 RS of a Pt|a-$GaO_{1.5}$|a-GaO_x|ITO device with $d \approx 15$ nm for the a-$GaO_{1.5}$ layer and $d \approx 55$ nm for the a-GaO_x layer (same device as in fig. 5.20; $x \approx 1.0$, PLD with 0.79 J/cm^2 in Ar, lithographic Pt-TE: 100×100 µm^2, CT4). The top graphs show the applied voltage $V(t)$, while the bottom graphs display the resulting I-V curve. (a) Triangular I-V sweep with only positive voltage (V_{max} = +1.5 V, v_{sw} = 0.2 V/s). (b) Triangular I-V sweep with only negative voltage (V_{max} = −1.5 V, v_{sw} = 0.2 V/s). In each graph, the numbered arrows refer to the first 0 V → V_{max} → 0 V cycle, which is the only one with a notable switching effect.

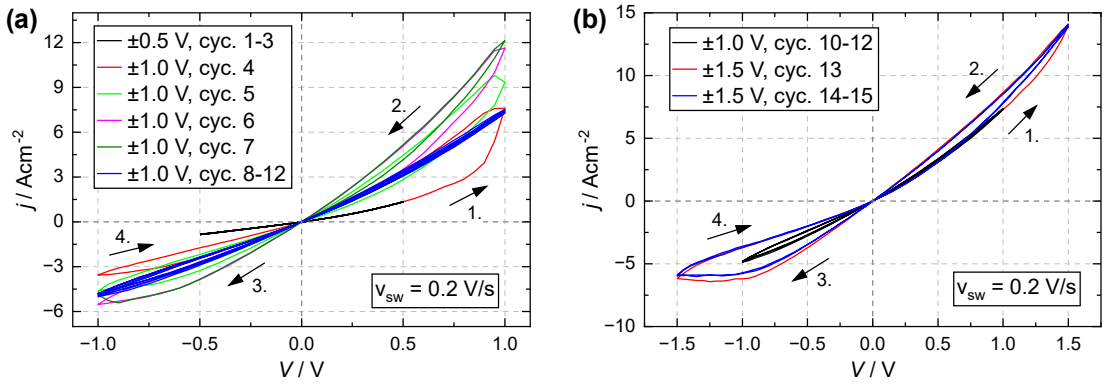


Figure B.12: Development of figure-8 bipolar RS from the as-prepared state of a Pt|a-$GaO_{1.5}$|a-GaO_x|ITO device with $d \approx 15$ nm for the a-$GaO_{1.5}$ layer and $d \approx 55$ nm for the a-GaO_x layer (same device as in fig. 5.20; $x \approx 1.0$, PLD with 0.79 J/cm^2 in Ar, lithographic Pt-TE: 100 × 100 µm^2, CT4). (a) I-V sweeps 1 to 12 after contacting of the device, with V_{max} = ±0.5 V and ±1.0 V. (b) I-V sweeps 10 to 15 with V_{max} = ±1.0 V and ±1.5 V, which led to the curve shape displayed in fig. 5.20a.

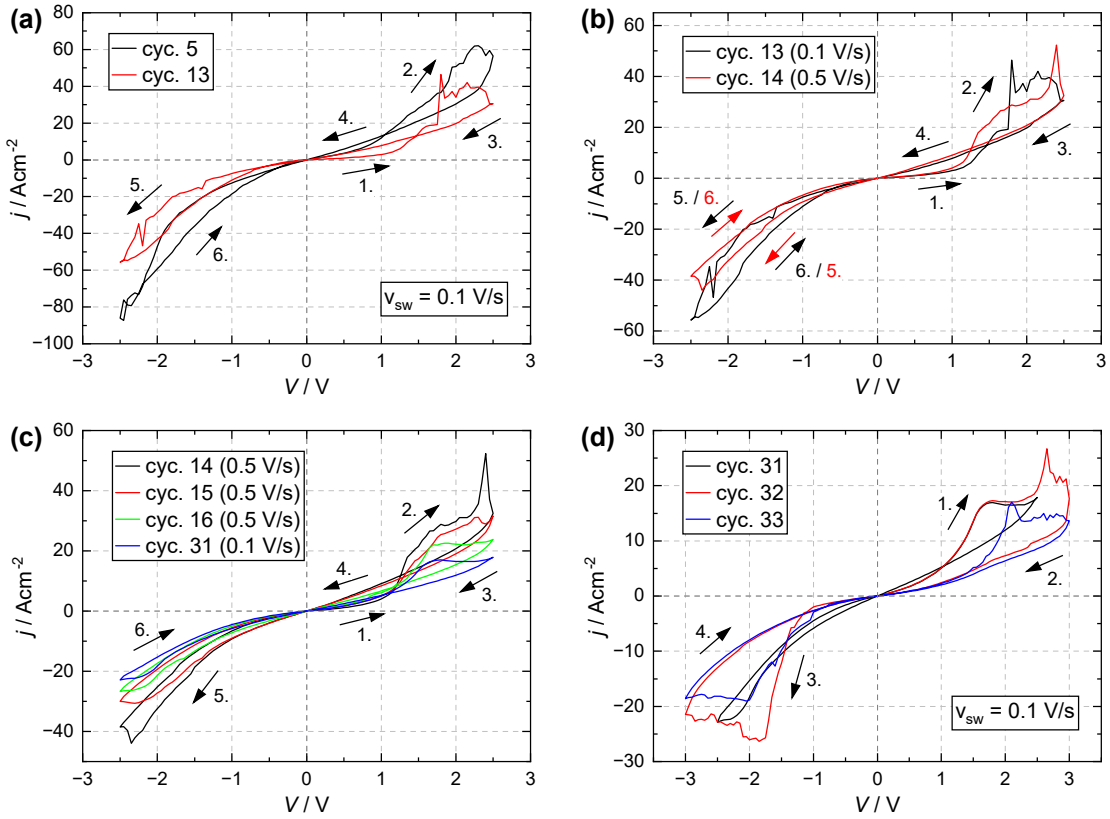


Figure B.13: Development of abnormal bipolar RS from the as-prepared state of a Pt|a-GaO_x|ITO device with $x \approx 0.9$ and $d \approx 90$ nm (same device as fig. 5.21b; PLD with 0.86 J/cm^2 in Ar, chamber #1; sp-Pt-TE-∅: 200 µm, CT1). (a) Initial counter-8 bipolar RS at $V_{max} = \pm 2.5$ V, with a crossing point at low positive voltage that gradually moves to higher voltages (cyc. 6–12 have the same shape as 5 and 13). (b) Reversal of the transition at negative voltages led to abnormal bipolar RS (HRS → LRS in cyc. 13, LRS → HRS in cyc. 14). (c) Gradual decrease of the overall current density and of the crossing point voltage from cycle to cycle (cyc. 17–30 have the same shape as 16 and 31). (d) Increase of V_{max} from ±2.5 V (cyc. 31) to ±3.0 V (cyc. 32–33) led to the symmetric abnormal bipolar RS displayed in fig. 5.21b. Similar plateaus in the abnormal bipolar I-V curves have been reported in [27] and [26].

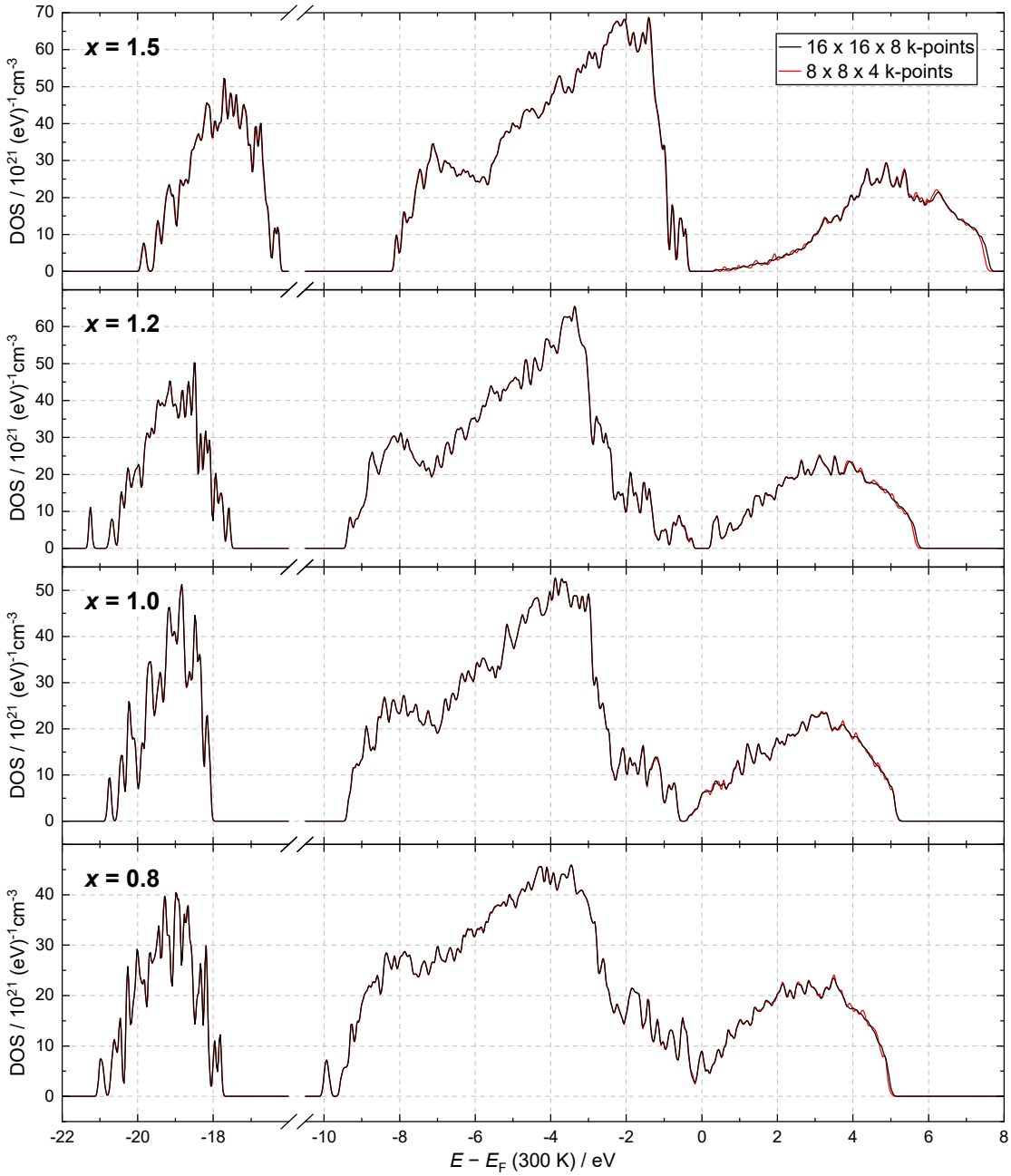


Figure B.14: Convergence check for the k-point density of the DFT-DOS calculations for a-$Ga_{64}O_{96}$ (x = 1.5), a-$Ga_{64}O_{78}$ (x = 1.2), a-$Ga_{64}O_{64}$ (x = 1.0) and a-$Ga_{72}O_{58}$ (x = 0.8).

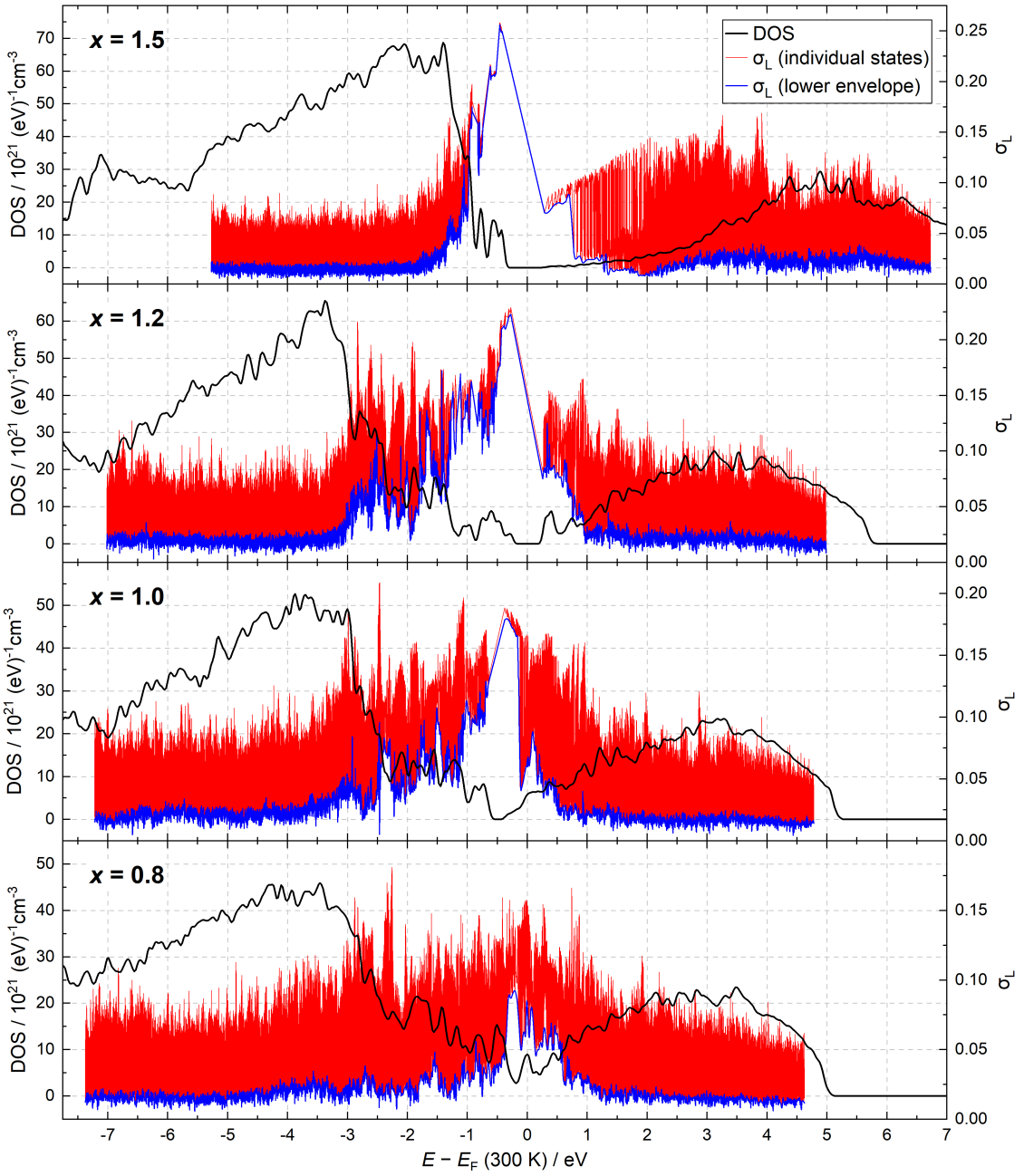


Figure B.15: Complete results of the LOS analysis for a-$Ga_{64}O_{96}$ (x = 1.5), a-$Ga_{64}O_{78}$ (x = 1.2), a-$Ga_{64}O_{64}$ (x = 1.0) and a-$Ga_{72}O_{58}$ (x = 0.8). For each of the stoichiometries, the DFT-DOS is shown together with the localization standard deviation $\sigma_L(E)$ for all electronic states in the selected energy range (in red), and with the lower envelope of $\sigma_L(E)$ in blue.

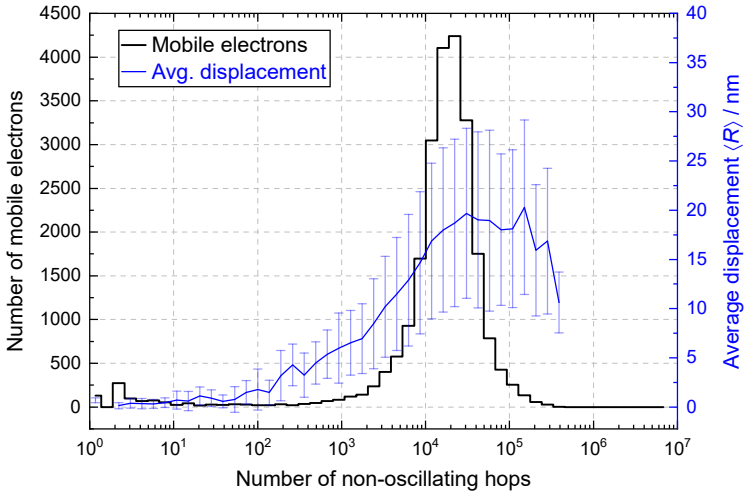


Figure B.16: Distribution of the non-oscillating hops per mobile electron during a KMC-VRH simulation for a-$GaO_{1.0}$ (same simulation as fig. 5.38). The data refers to equilibration and main simulation together ($N_{h,eq} = 3{\times}10^8$, $N_{h,sim} = 2{\times}10^{10}$), with a total number of non-oscillating hops of $5.4{\times}10^8$ (2.7%). The average number of non-osc. hops per mobile electron is $2.3{\times}10^4$ and the highest number of non-osc. hops reached by an individual electron is $4.0{\times}10^5$. For comparison, the blue graph shows the mean displacement $\langle R \rangle$ of the mobile electrons as a function of the number of non-osc. hops, with $\langle R \rangle$ as arithmetic average of the individual displacements $R_k = (X_k^2 + Y_k^2 + Z_k^2)^{1/2}$ and the error bars as the respective sample standard deviation. It reveals that the electrons with the highest displacement are not necessarily those with the most non-osc. hops. The reason for this is that the employed electric field strength of 0.0635 MV/cm is rather weak and only 50.045% of all non-osc. hops took place in $+x$-direction ($x_{ij} > 0$).

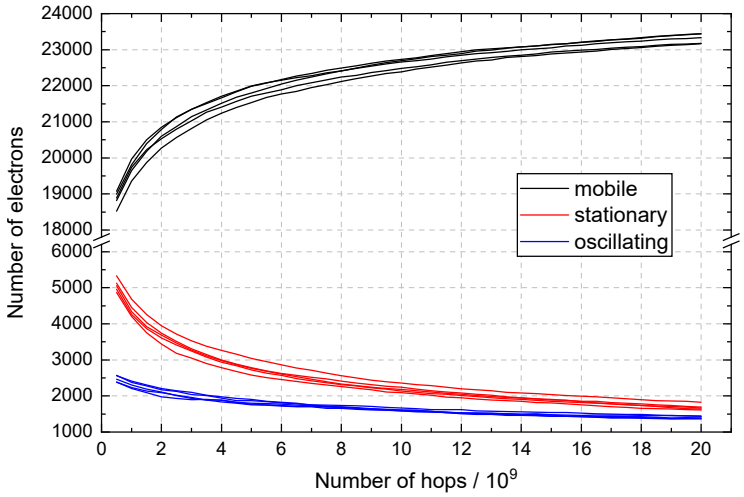


Figure B.17: Development of the different electron categories during the main simulation stage of 5 KMC repetitions for a-$GaO_{1.0}$, with an especially large number of $N_{h,sim} = 2{\times}10^{10}$ hops to investigate the long-term trend (same simulation as fig. 5.38). Each repetition has different localized states due to a varied seed for the random number generator. In the limit of $\tau \to \infty$, each electron in the KMC cell would make an non-oscillating hop and thus become mobile in the formal sense, while the number of completely stationary or oscillating electrons would converge to zero. This does not happen on physically reasonable timescales, though, due to the $\exp[-\varepsilon_{ij}/(k_B T)]$ term in the Miller-Abrahams equation. Reaching this limiting case is not necessary either, because electrons are indistinguishable and it is irrelevant which of them contribute to charge transport at a certain time.

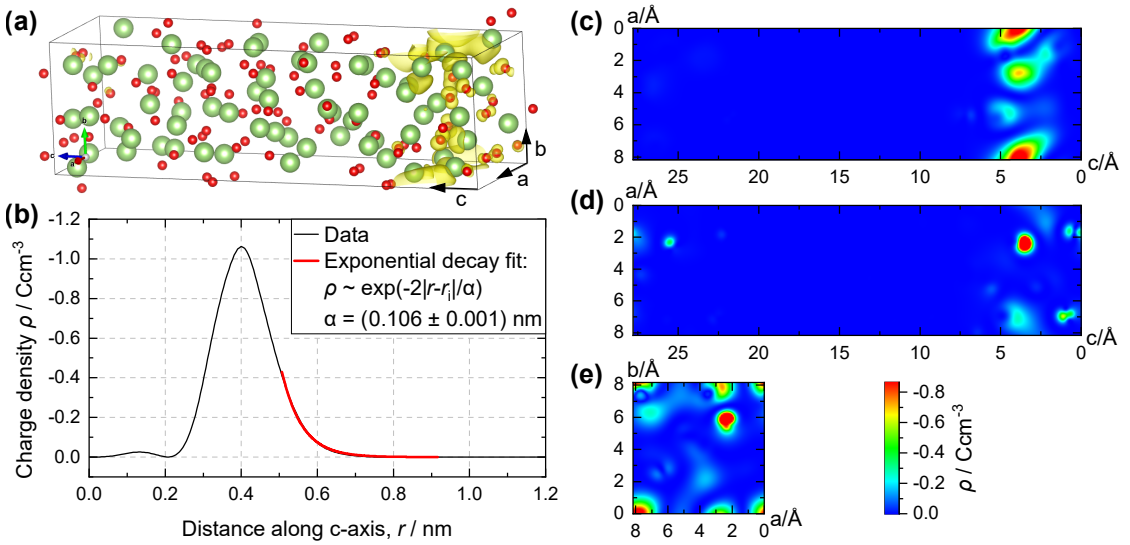


Figure B.18: Investigation of a localized electronic state at the Fermi level, as obtained from the DFT calculations for a-$Ga_{64}O_{64}$ (x = 1.0; same state as in fig. 5.36d). (a) Charge density $\rho = -e|\phi_i|^2$ of the localized state, displayed as yellow isosurface for $\rho = -0.17\,\mathrm{C/cm^3}$ (Ga atoms in green, O atoms in red). (b) Line plot of the charge density from the origin of the DFT cell in c-direction, i.e., through the largest bubble visible in (a). The exponential tail is fitted by $\rho \propto \exp(-2|\vec{r} - \vec{r}_i|/\alpha)$, which yields α = 0.106 nm. Note that this localized state also has 10 other regions with higher ρ, but this one is the most extended and thus probably the most influential for VRH. (c) Charge density map of the ac-plane at b = 0 nm, which contains the region fitted in (b) as red peak. (d) Charge density map of the ac-plane at b = 5.87 nm, which contains the regions with the overall highest ρ, for comparison. (e) Charge density map of the ab-plane at c = 3.65 nm, which contains both the region with highest ρ and the region fitted in (b) in the corners.

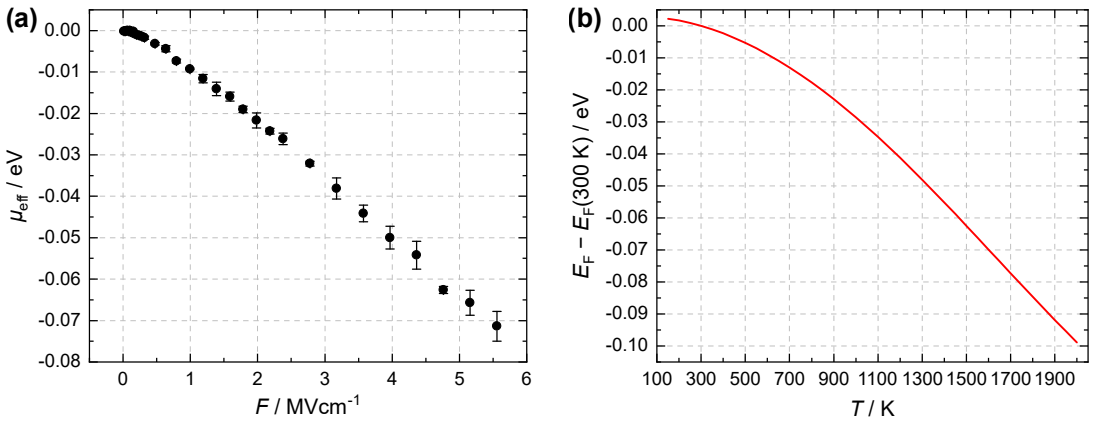


Figure B.19: Variation of the Fermi level in a-$GaO_{1.0}$. (a) Electric field-dependence of the effective chemical potential μ_{eff} at T = 300 K, as obtained from the KMC simulations beyond the weak-field region (cf. eq. 4.17). (b) Temperature-dependence of the Fermi level relative to T = 300 K, as obtained from the DFT-DOS by the adjustment procedure described in sec. 4.1.2. The result is similar to the relation between μ_{eff} and T_{eff} in the KMC simulations.

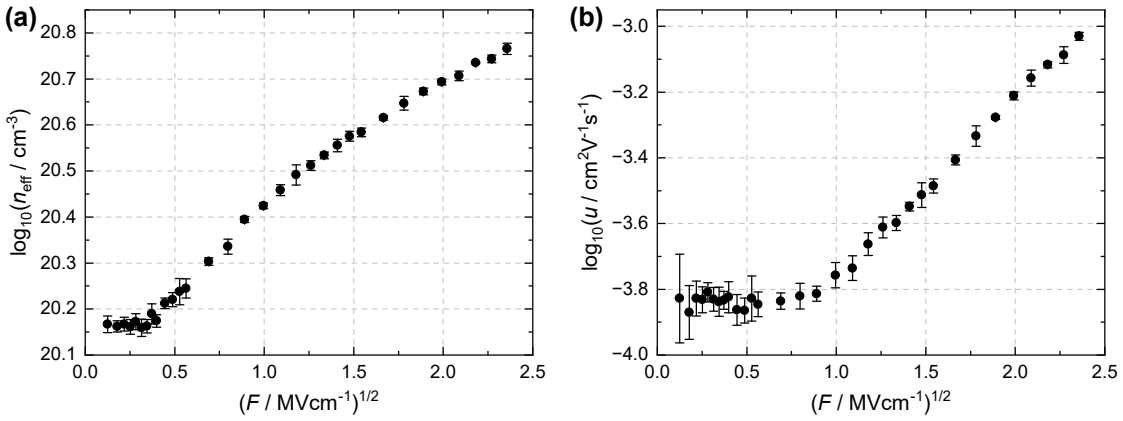


Figure B.20: Electric field-dependence of the effective charge carrier density n_{eff} and the mobility u of a-$GaO_{1.0}$ at T = 300 K, as obtained through the KMC simulations with varied electric field strength F (same simulations as in fig. 5.56).

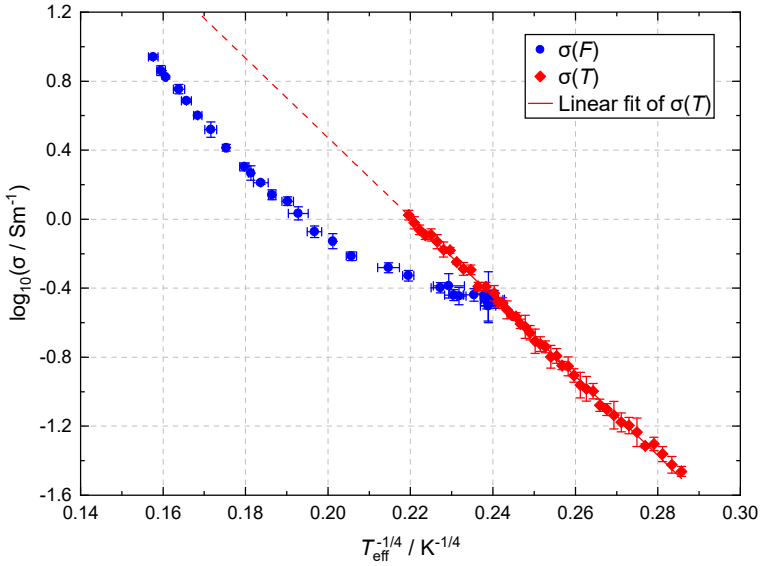


Figure B.21: Assessment of the hypothesis that the electric field-dependence of the VRH conductivity can be described by the classical VRH law with T replaced by the electrons' effective temperature T_{eff}. The weak-field KMC simulations for a-$GaO_{1.0}$ at different temperatures ("$\sigma(T)$"; cf. fig. 5.49) are compared to the simulations at high electric field strength F ("$\sigma(F)$"; T = 300 K; cf. fig. 5.56).

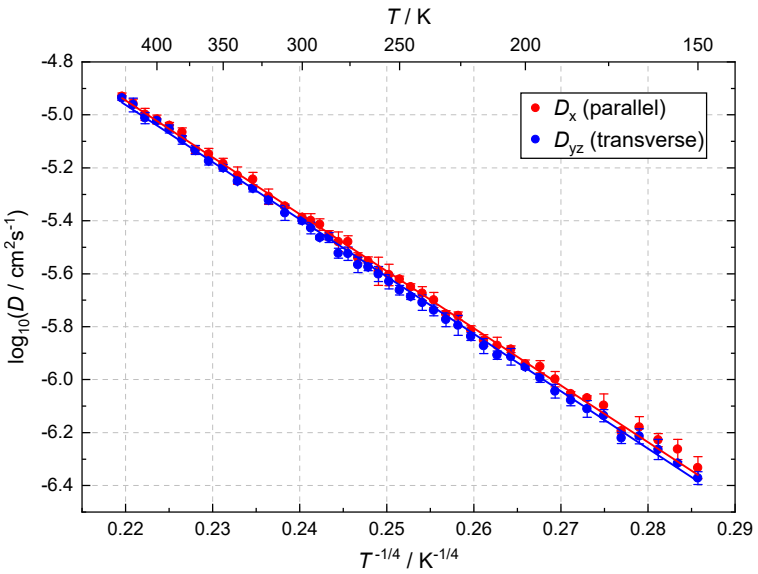


Figure B.22: Comparison of the Fickian diffusion coefficients parallel (D_x) and perpendicular ($D_{yz} = D$) to the electric field, as obtained in the VRH simulations for a-$GaO_{1.0}$ with varied temperature (same simulations as in fig. 5.49). The lines are the respective linear fits.

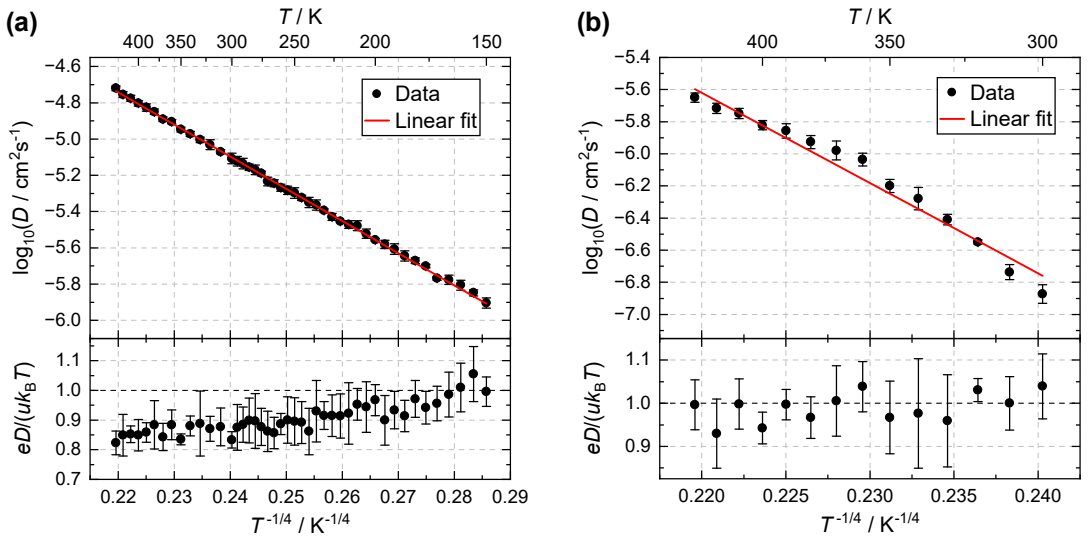


Figure B.23: Temperature-dependence of the Fickian diffusion coefficient D (top) and the corresponding Nernst-Einstein ratio eD/uk_BT (bottom) for VRH in a-GaO_x with (a) $x = 0.8$ and (b) $x = 1.2$ (same simulations as in fig. 5.52).

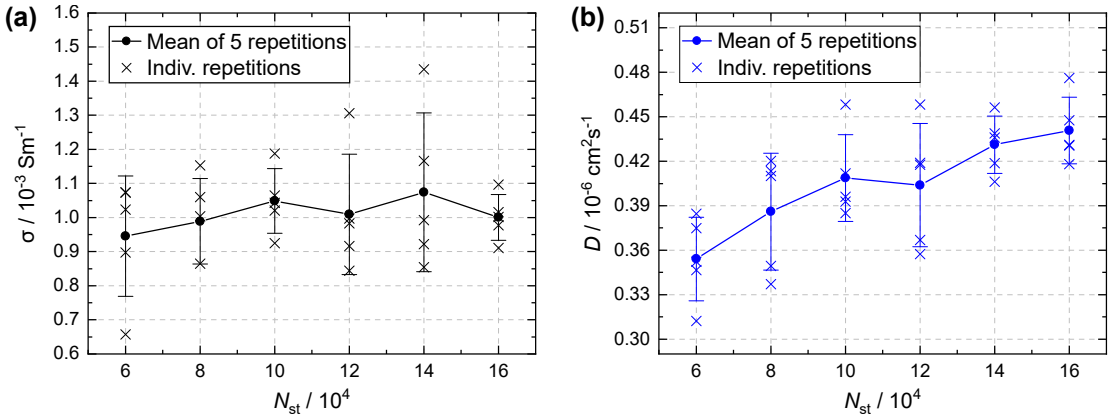


Figure B.24: KMC-VRH simulations with different number of states N_{st}, for positively charged a-$GaO_{1.0}$ with $\mu = -0.5$ eV at $T = 300$ K. (a) The conductivity σ is independent of N_{st} (apart from statistical scatter), as required. (b) The simultaneously obtained Fickian diffusion coefficient D exhibits a slight systematic increase for increasing N_{st}, which hints at incomplete convergence (too low $N_{h,sim}$) or – more likely – insufficient equilibration (too low $N_{h,eq}$).

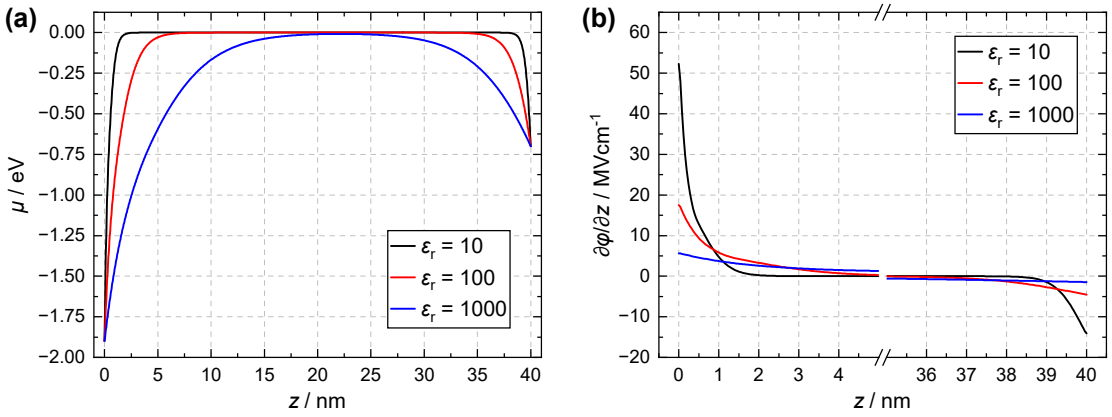


Figure B.25: Electron equilibrium in Pt|a-$GaO_{1.0}$|ITO with $V = 0$, for different relative permittivities ϵ_r of the a-$GaO_{1.0}$ layer (FEM-1 model without high-F effects). (a) Chemical potential μ of the electrons, which corresponds to the electric potential according to $\mu = e\varphi$. (b) Electric potential gradient $\partial\varphi/\partial z$ ($= -\vec{F}$).

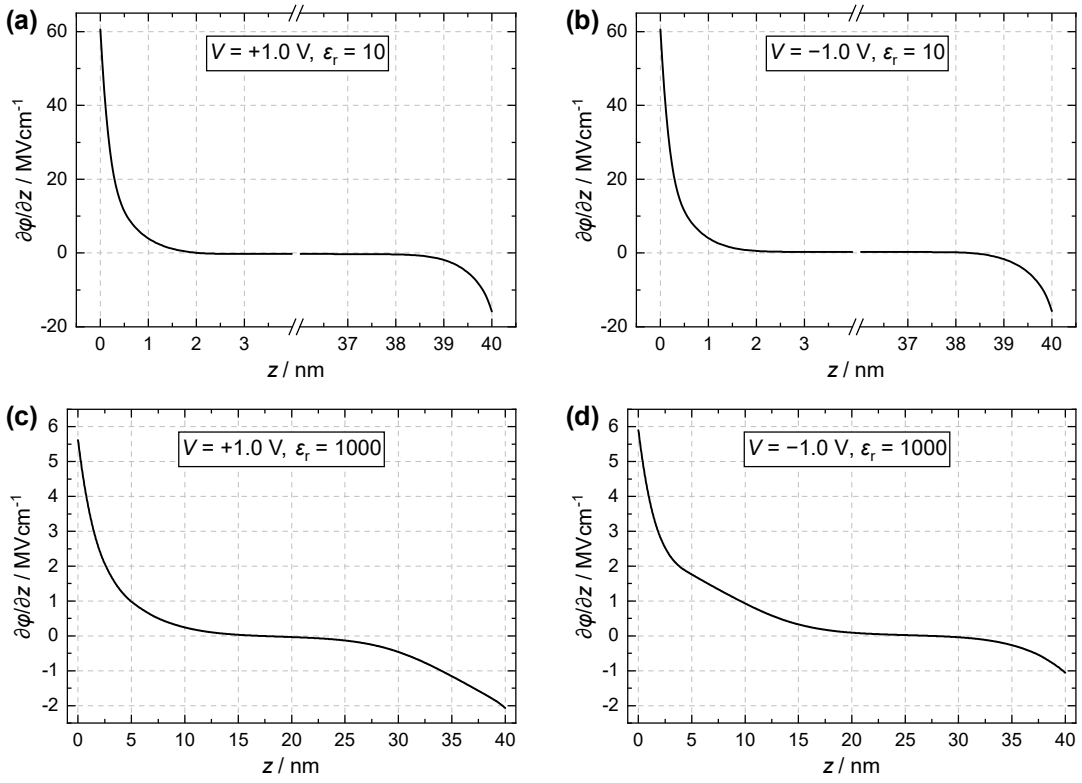


Figure B.26: Gradient of the electric potential $\partial\varphi/\partial z$ $(= -\vec{F})$ in Pt|a-$GaO_{1.0}$|ITO with an applied voltage of (a,c) $V = +1\,\mathrm{V}$ or (b,d) $V = -1\,\mathrm{V}$, for different permittivities of either (a,b) $\epsilon_r = 10$ or (c,d) $\epsilon_r = 1000$ (FEM-4 model with high-F effects). In (a,b), it is $\partial\varphi/\partial z = -0.27\,\mathrm{MV/cm}$ in the whole bulk of the a-$GaO_{1.0}$ layer (approx. $3\,\mathrm{nm} < z < 37\,\mathrm{nm}$).

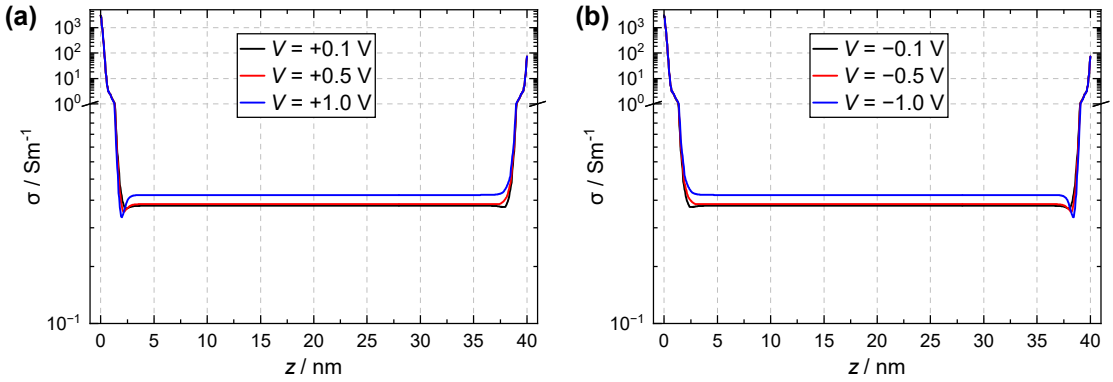


Figure B.27: Voltage-dependence of the hopping conductivity σ in Pt|a-$GaO_{1.0}$|ITO with $\epsilon_r = 10$ (FEM-4 model with high-F effects).

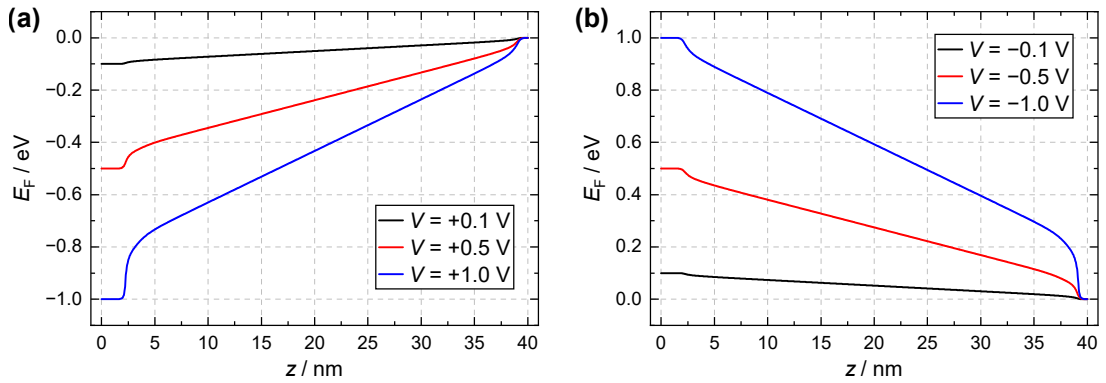


Figure B.28: Voltage-dependence of the Fermi level E_F (= $\tilde{\mu}$) in Pt|a-$GaO_{1.0}$|ITO with ϵ_r = 100 (FEM-4 model with high-F effects). (a) For $V > 0$, the bulk and the TE interface dominate the resistance. (b) For $V < 0$, the bulk and the BE interface have relevant voltage drops.

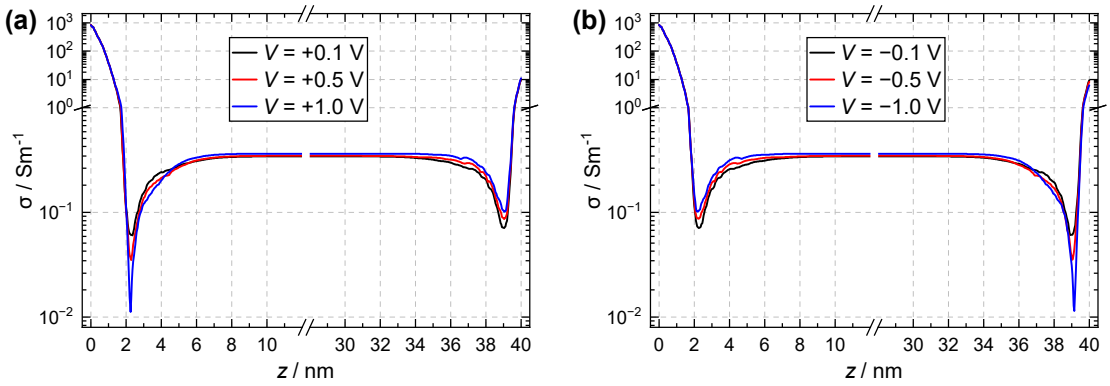


Figure B.29: Voltage-dependence of the hopping conductivity σ in Pt|a-$GaO_{1.0}$|ITO with ϵ_r = 100 (FEM-4 model with high-F effects). (a) For $V > 0$, the conductivity increases in the bulk and near the BE, but it decreases near the TE. (b) For $V < 0$, it increases in the bulk and near the TE, while it decreases near the BE.

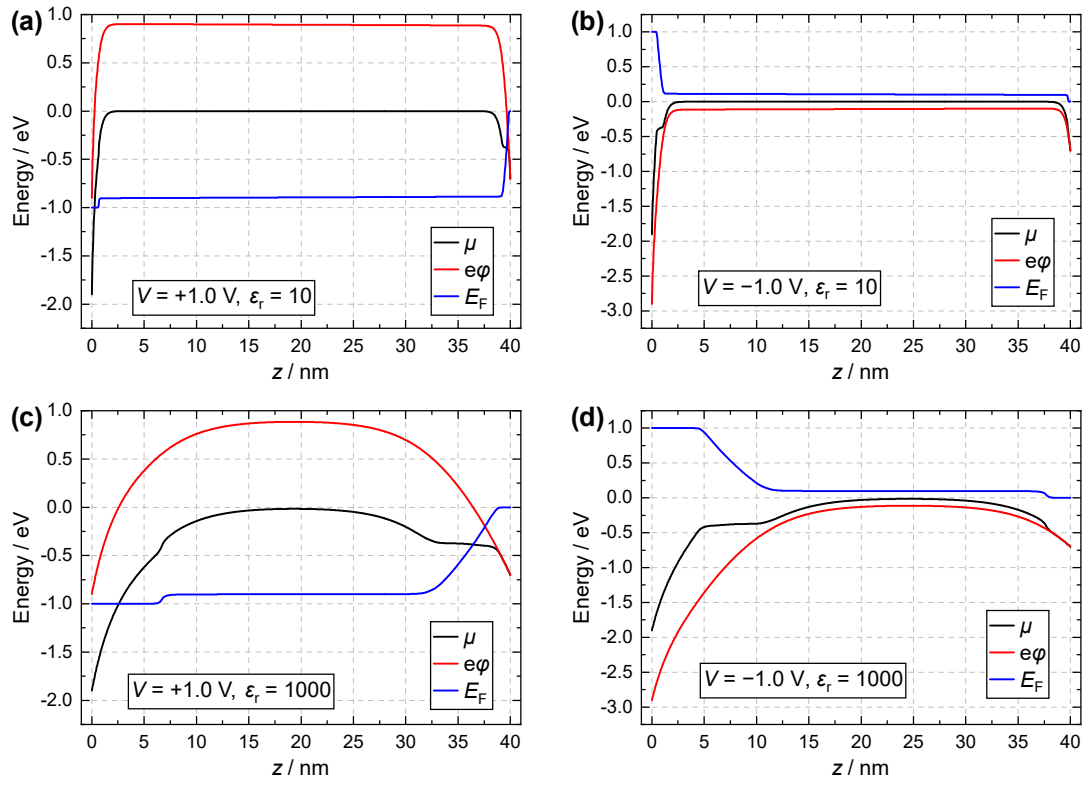


Figure B.30: Chemical potential μ, electric potential φ and Fermi level E_F (= $\tilde{\mu}$) in Pt|a-$GaO_{1.0}$|ITO with an applied voltage of (a,c) V = +1 V or (b,d) V = −1 V, for different permittivities of either (a,b) ϵ_r = 10 or (c,d) ϵ_r = 1000 (FEM-3 model without high-F effects).

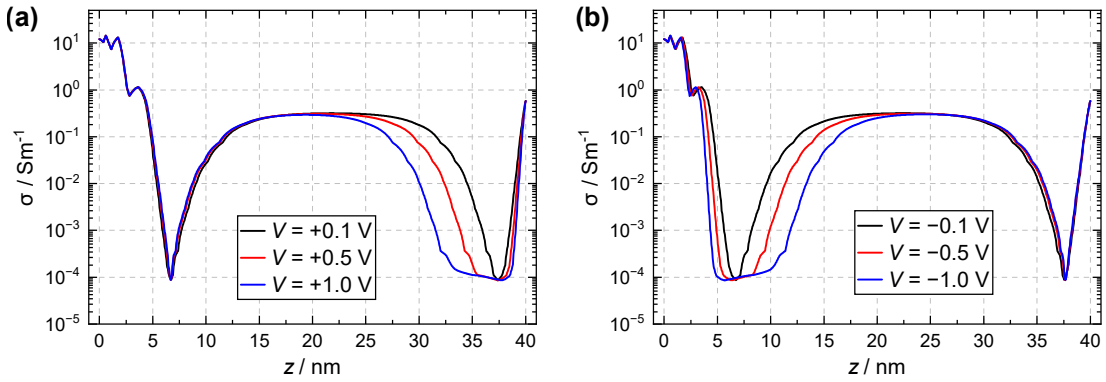


Figure B.31: Voltage-dependence of the hopping conductivity σ in Pt|a-$GaO_{1.0}$|ITO with ϵ_r = 1000 (FEM-3 model without high-F effects).

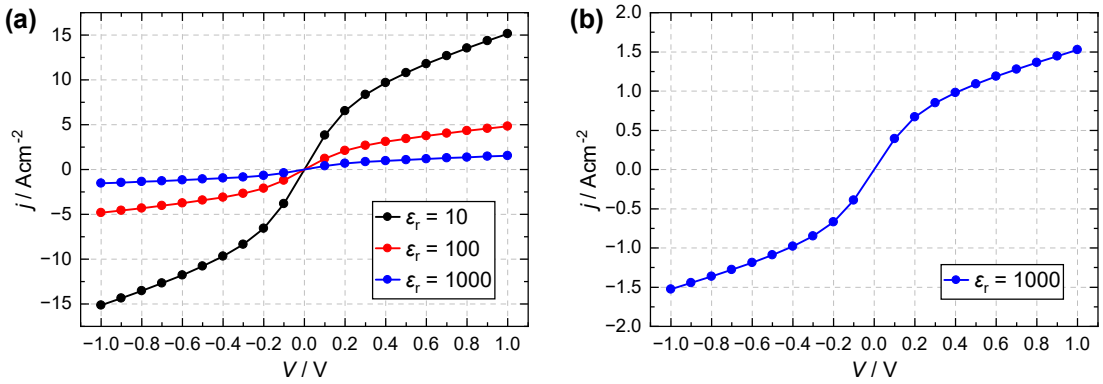


Figure B.32: Current density j for different voltages V applied to Pt|a-$GaO_{1.0}$|ITO (FEM-3 model without high-F effects). The points represent the individual FEM simulations, which are connected by straight lines as guide for the eye. In (a), the results for the different relative permittivities ϵ_r = 10, 100 and 1000 are displayed together, while (b) shows a close-up plot of the data for ϵ_r = 1000.

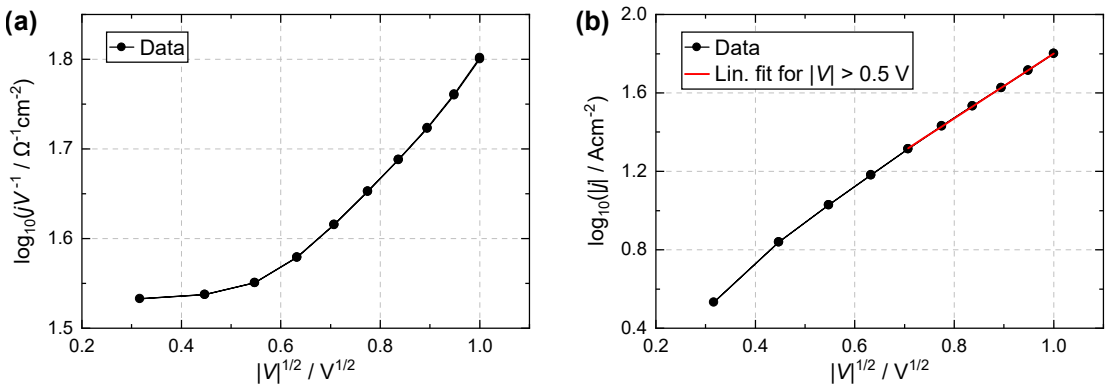


Figure B.33: Logarithmic plots of the simulated j-V curve for Pt|a-$GaO_{1.0}$|ITO with ϵ_r = 1000 (same data as in fig. 5.69b; FEM-4 model with high-F effects).

C

Tables

Table C.1: Binding energies (E_{bin} in eV) of the fitted HAXPES peaks of all measured samples. Shown is the average peak position ± max. deviation.

Ga $2p_{3/2}$		Ga 3d			
Ga^{3+}	Ga^{+}	Ga^{3+}	Ga^{+}	O 1s	O 2s
1118.50 ± 0.15	1117.30 ± 0.22	20.66 ± 0.12	19.34 ± 0.09	531.44 ± 0.14	24.13 ± 0.25

Table C.2: Definition of the different contact types – abbreviated by "CT" – for the *I-V* measurements.

Type	Measuring instrument	TE contact method	BE contact method
CT1	simple *I-V* station	W tip (cat-whisker)	Ag paste + Au wire
CT2	MPS150 probe station	W tip (cat-whisker)	W tip (needle)
CT3	MPS150 probe station	W tip (needle)	W tip (needle)
CT4	MPS150 probe station	etched Au wire	W tip (needle)
CT5	MPS150 probe station	Pt paste + Au wire	Pt paste + Au wire

References

[1] M. Kryder, Chang Soo Kim, *IEEE Transactions on Magnetics* **2009**, *45*, 3406–3413.

[2] G. E. Moore in International Electron Devices Meeting, IEEE, **1975**, pp. 11–13.

[3] D. S. H. Rosenthal, D. C. Rosenthal, E. L. Miller, I. F. Adams, M. W. Storer, E. Zadok in The Memory of the World in the Digital Age: Digitization and Preservation, Vancouver, Canada, **2012**.

[4] T. N. Theis, H.-S. P. Wong, *Computing in Science & Engineering* **2017**, *19*, 41–50.

[5] T. Shi, R. Wang, Z. Wu, Y. Sun, J. An, Q. Liu, *Small Structures* **2021**, *2*, 2000109.

[6] S. H. Lee, X. Zhu, W. D. Lu, *Nano Research* **2020**, *13*, 1228–1243.

[7] Y. Xi, B. Gao, J. Tang, A. Chen, M.-F. Chang, X. S. Hu, J. V. D. Spiegel, H. Qian, H. Wu, *Proceedings of the IEEE* **2021**, *109*, 14–42.

[8] S. Slesazeck, T. Mikolajick, *Nanotechnology* **2019**, *30*, 352003.

[9] X. Gao, Y. Xia, J. Ji, H. Xu, Y. Su, H. Li, C. Yang, H. Guo, J. Yin, Z. Liu, *Applied Physics Letters* **2010**, *97*, 193501.

[10] Y. Aoki, C. Wiemann, V. Feyer, H.-S. Kim, C. M. Schneider, H. Ill-Yoo, M. Martin, *Nature Communications* **2014**, *5*, 3473.

[11] S. I. Stepanov, V. I. Nikolaev, V. E. Bougrov, A. E. Romanov, *Reviews on Advanced Materials Science* **2016**, *44*, 63–86.

[12] M. Higashiwaki, G. H. Jessen, *Applied Physics Letters* **2018**, *112*, 060401.

[13] *Gallium Oxide*, (Eds.: S. Pearton, F. Ren, M. Mastro), Elsevier, Amsterdam, **2019**.

[14] J. Zhang, J. Shi, D.-C. Qi, L. Chen, K. H. L. Zhang, *APL Materials* **2020**, *8*, 020906.

[15] H. Sheoran, V. Kumar, R. Singh, *ACS Applied Electronic Materials* **2022**, *4*, 2589–2628.

[16] Shivani, D. Kaur, A. Ghosh, M. Kumar, *Materials Today Communications* **2022**, *33*, 104244.

[17] J. Tsai, C.-H. Hsu, C.-Y. Hsu, C.-S. Yang, *Electronics Letters* **2013**, *49*, 554–555.

[18] C.-C. Yang, J.-Q. Huang, K.-Y. Chen, P.-H. Chiu, H.-T. Vu, Y.-K. Su, *IEEE Access* **2019**, *7*, 175186–175191.

[19] K.-J. Gan, P.-T. Liu, T.-C. Chien, D.-B. Ruan, S. M. Sze, *Scientific Reports* **2019**, *9*, 14141.

[20] C.-Y. Wang, S.-P. Chang, W.-L. Huang, *ECS Advances* **2022**, *1*, 023501.

[21] S. Wang, C. He, J. Tang, R. Yang, D. Shi, G. Zhang, *Chinese Physics B* **2019**, *28*, 017304.

[22] N. H. Al-Hardan, M. A. Abdul Hamid, A. Jalar, M. Firdaus-Raih, *Materials Today Physics* **2023**, *38*, 101279.

[23] Y. Xiao, S. Ke, Y. Jin, T. Zhang, B. Jiang, J. Meng, S. Chen, R. Li, Z. Zhang, C. Ye, *Ceramics International* **2024**, *50*, 8981–8986.

[24] K.-J. Gan, P.-T. Liu, D.-B. Ruan, C.-C. Hsu, Y.-C. Chiu, S. M. Sze, *Journal of Electronic Materials* **2020**, *49*, 6817–6822.

[25] M. N. Almadhoun, M. Speckbacher, B. C. Olsen, E. J. Luber, S. Y. Sayed, M. Tornow, J. M. Buriak, *Nano Letters* **2021**, *21*, 2666–2674.

[26] D. Y. Guo, Y. P. Qian, Y. L. Su, H. Z. Shi, P. G. Li, J. T. Wu, S. L. Wang, C. Cui, W. H. Tang, *AIP Advances* **2017**, *7*, 065312.

[27] D. Y. Guo, Z. P. Wu, L. J. Zhang, T. Yang, Q. R. Hu, M. Lei, P. G. Li, L. H. Li, W. H. Tang, *Applied Physics Letters* **2015**, *107*, 032104.

[28] K. Sato, Y. Hayashi, N. Masaoka, T. Tohei, A. Sakai, *Scientific Reports* **2023**, *13*, 1261.

[29] C. Kura, Y. Aoki, E. Tsuji, H. Habazaki, M. Martin, *RSC Advances* **2016**, *6*, 8964–8970.

[30] C. Ader, A. Falkenstein, M. Martin, *Scientific Reports* **2021**, *11*, 14384.

[31] N. F. Mott, E. A. Davis, *Electronic Processes in Non-Crystalline Materials*, 2nd ed., Oxford University Press, Oxford, **1979**.

[32] L. Nagarajan, R. A. De Souza, D. Samuelis, I. Valov, A. Börger, J. Janek, K.-D. Becker, P. C. Schmidt, M. Martin, *Nature Materials* **2008**, *7*, 391–398.

[33] A. von der Heiden, M. Bornhöfft, J. Mayer, M. Martin, *Physical Chemistry Chemical Physics* **2019**, *21*, 4268–4275.

[34] S. R. de Groot, P. Mazur, *Non-Equilibrium Thermodynamics*, Dover, Dover Publications, New York, **1984**, 510 pp.

[35] N. W. Ashcroft, N. D. Mermin, *Solid State Physics*, Holt, Rinehart and Winston, New York, **1976**, 826 pp.

[36] T. Yamamoto, M. Ogata, H. Fukuyama, *Journal of the Physical Society of Japan* **2022**, *91*, 044704.

[37] A. R. Allnatt, A. B. Lidiard, *Atomic Transport in Solids*, Cambridge University Press, Cambridge, England, **1993**, 572 pp.

[38] H. Mehrer, *Diffusion in Solids: Fundamentals, Methods, Materials, Diffusion-Controlled Processes*, Springer, Berlin ; New York, **2007**, 651 pp.

[39] C. Korte, J. Janek, *J. Phys. Chem Solids* **1997**, *58*, 623–637.

[40] P. Hein, T. Romstadt, F. Draber, J. Ryu, T. Böger, A. Falkenstein, M. Kim, M. Martin, *Advanced Electronic Materials* **2025**, *11*, 2400407.

[41] S. M. Sze, K. N. Kwok, *Physics of Semiconductor Devices*, 3rd, John Wiley & Sons, Inc., **2006**.

[42] V. Palenskis, *AIP Advances* **2014**, *4*, 047119.

[43] V. Palenskis, E. Žitkevičius, *Crystals* **2021**, *11*, 622.

[44] U. Mizutani, *Introduction to the Electron Theory of Metals*, 1st ed., Cambridge University Press, **2001**.

[45] N. Zettili, *Quantum Mechanics: Concepts and Applications*, Wiley, Chichester ; New York, **2001**, 649 pp.

[46] R. F. Pierret, *Advanced Semiconductor Fundamentals*, 2nd ed, Prentice Hall, Upper Saddle River, N.J, **2003**, 221 pp.

[47] A. Moliton, *Solid-State Physics for Electronics*, ISTE, London, UK, **2009**.

[48] D. K. Ferry, *Quantum Mechanics: An Introduction for Device Physicists and Electrical Engineers*, 2nd ed, Taylor & Francis, New York, **2001**, 345 pp.

[49] S. Goßner, *Grundlagen der Elektronik: Halbleiter, Bauelemente und Schaltungen*, 7. ergänzte Auflage, Shaker Verlag, Aachen, **2008**, 527 pp.

[50] C. Hu, *Modern Semiconductor Devices for Integrated Circuits*, Pearson Education (US), **2009**.

[51] M. Ogata, H. Fukuyama, *Journal of the Physical Society of Japan* **2019**, *88*, 074703.

[52] P. Landsberg, S. Hope, *Solid-State Electronics* **1977**, *20*, 421–429.

[53] J. Singh, K. Shimakawa, *Advances in Amorphous Semiconductors*, Taylor & Francis, London ; New York, **2003**, 329 pp.

[54] *Charge Transport in Disordered Solids with Applications in Electronics*, (Ed.: S. Baranovski), John Wiley & Sons Ltd, **2006**.

[55] P. W. Anderson, *Physical Review* **1958**, *109*, 1492–1505.

[56] N. Bacalis, E. N. Economou, M. H. Cohen, *Physical Review B* **1988**, *37*, 2714–2717.

[57] A. Miller, E. Abrahams, *Physical Review* **1960**, *120*, 745–755.

[58] N. F. Mott, *Journal of Non-Crystalline Solids* **1968**, *1*, 1–17.

[59] V. Ambegaokar, B. I. Halperin, J. S. Langer, *Physical Review B* **1971**, *4*, 2612–2620.

[60] B. I. Shklovskiĭ, A. L. Éfros, *Soviet Physics Uspekhi* **1975**, *18*, 845–862.

[61] B. I. Shklovskii, A. L. Efros, *Electronic Properties of Doped Semiconductors*, Springer, Berlin, **1984**, 386 pp.

[62] A. Voje, MSc thesis, University of Oslo, Oslo, Norway, **2009**.

[63] S. Marianer, B. I. Shklovskii, *Physical Review B* **1992**, *46*, 13100–13103.

[64] B. Cleve, B. Hartenstein, S. D. Baranovskii, M. Scheidler, P. Thomas, H. Baessler, *Physical Review B* **1995**, *51*, 16705–16713.

[65] J. Kurchan, *Nature* **2005**, *433*, 222–225.

[66] A. M. Somoza, M. Ortuño, M. Caravaca, M. Pollak, *Physical Review Letters* **2008**, *101*, 056601.

[67] T. Kluge, T. Cowan, A. Debus, U. Schramm, K. Zeil, M. Bussmann, *Physical Review Letters* **2011**, *107*, 205003.

[68] H. Mizuno, H. Shiba, A. Ikeda, *Proceedings of the National Academy of Sciences* **2017**, *114*, E9767–E9774.

[69] K. Prasai, P. Biswas, D. A. Drabold, *Semiconductor Science and Technology* **2016**, *31*.

[70] W.-X. Zhou, Y. Cheng, K.-Q. Chen, G. Xie, T. Wang, G. Zhang, *Advanced Functional Materials* **2020**, *30*, 1903829.

[71] P. B. Allen, J. L. Feldman, J. Fabian, F. Wooten, *Philosophical Magazine B* **1999**, *79*, 1715–1731.

[72] N. F. Mott, *The Philosophical Magazine* **1969**, *19*, 835–852.

[73] M. Pollak, *Journal of Non-Crystalline Solids* **1972**, *11*, 1–24.

[74] J. M. Marshall in *Functional Properties of Nanostructured Materials*, *Vol. 223*, (Eds.: R. Kassing, P. Petkov, W. Kulisch, C. Popov), Springer Netherlands, Dordrecht, **2006**, pp. 45–56.

[75] E. M. Hamilton, *Philosophical Magazine* **1972**, *26*, 1043–1045.

[76] A. L. Efros, B. I. Shklovskii, *Journal of Physics C: Solid State Physics* **1975**, *8*, L49.

[77] I. I. Fishchuk, A. Kadashchuk, A. Bhoolokam, A. de Jamblinne de Meux, G. Pourtois, M. M. Gavrilyuk, A. Köhler, H. Bässler, P. Heremans, J. Genoe, *Physical Review B* **2016**, *93*, 195204.

[78] E. H. Rhoderick, R. H. Williams, *Metal-Semiconductor Contacts*, 2nd ed., Clarendon Press ; Oxford University Press, Oxford [England]; New York, **1988**.

[79] M. Higashiwaki, K. Konishi, K. Sasaki, K. Goto, K. Nomura, Q. T. Thieu, R. Togashi, H. Murakami, Y. Kumagai, B. Monemar, A. Koukitu, A. Kuramata, S. Yamakoshi, *Applied Physics Letters* **2016**, *108*, 133503.

[80] F.-C. Chiu, *Advances in Materials Science and Engineering* **2014**, *2014*, 1–18.

[81] C. R. Wronski, D. E. Carlson, R. E. Daniel, *Applied Physics Letters* **1976**, *29*, 602–605.

[82] W. B. Jackson, R. J. Nemanich, M. J. Thompson, B. Wacker, *Physical Review B* **1986**, *33*, 6936–6945.

[83] P. Schlupp, H. Von Wenckstern, M. Grundmann, *physica status solidi (a)* **2017**, *214*, 1700210.

[84] J. Åhman, G. Svensson, J. Albertsson, *Acta Crystallographica Section C Crystal Structure Communications* **1996**, *52*, 1336–1338.

[85] S. S. Kumar, E. Rubio, M. Noor-A-Alam, G. Martinez, S. Manandhar, V. Shutthanandan, S. Thevuthasan, C. Ramana, *The Journal of Physical Chemistry C* **2013**, *117*, 4194–4200.

[86] Y. Qin, S. Long, Q. He, H. Dong, G. Jian, Y. Zhang, X. Hou, P. Tan, Z. Zhang, Y. Lu, C. Shan, J. Wang, W. Hu, H. Lv, Q. Liu, M. Liu, *Advanced Electronic Materials* **2019**, *5*, 1900389.

[87] M. I. Pintor-Monroy, M. G. Reyes-Banda, C. Avila-Avendano, M. A. Quevedo-Lopez, *IEEE Sensors Journal* **2021**, *21*, 14807–14814.

[88] K. Zheng, J. L. Zhao, X. W. Sun, V. Q. Vinh, K. S. Leck, R. Zhao, Y. G. Yeo, L. T. Law, K. L. Teo, *Applied Physics Letters* **2012**, *101*, 143110.

[89] H. Liang, S. Cui, R. Su, P. Guan, Y. He, L. Yang, L. Chen, Y. Zhang, Z. Mei, X. Du, *ACS Photonics* **2019**, *6*, 351–359.

[90] L.-X. Qian, Z.-H. Wu, Y.-Y. Zhang, P. T. Lai, X.-Z. Liu, Y.-R. Li, *ACS Photonics* **2017**, *4*, 2203–2211.

[91] A. Petitmangin, B. Gallas, C. Hebert, J. Perrière, L. Binet, P. Barboux, X. Portier, *Applied Surface Science* **2013**, *278*, 153–157.

[92] X. Li, H.-L. Lu, H.-P. Ma, J.-G. Yang, J.-X. Chen, W. Huang, Q. Guo, J.-J. Feng, D. W. Zhang, *Current Applied Physics* **2019**, *19*, 72–81.

[93] W. Li, J. Wan, Z. Tu, H. Li, H. Wu, C. Liu, *Ceramics International* **2022**, *48*, 3185–3191.

[94] A. Mahmoodinezhad, C. Janowitz, F. Naumann, P. Plate, H. Gargouri, K. Henkel, D. Schmeißer, J. I. Flege, *Journal of Vacuum Science & Technology A* **2020**, *38*, 022404.

[95] H. Liang, Z. Han, Z. Mei, *physica status solidi (a)* **2021**, *218*, 2000339.

[96] J. Brendt, PhD thesis, RWTH Aachen University, Aachen, Germany, **2011**.

[97] A. von der Heiden, PhD thesis, RWTH Aachen University, Aachen, Germany, **2017**.

[98] D. Y. Guo, Z. P. Wu, Y. H. An, P. G. Li, P. C. Wang, X. L. Chu, X. C. Guo, Y. S. Zhi, M. Lei, L. H. Li, W. H. Tang, *Applied Physics Letters* **2015**, *106*, 042105.

[99] P. C. Wang, P. G. Li, Y. S. Zhi, D. Y. Guo, A. Q. Pan, J. M. Zhan, H. Liu, J. Q. Shen, W. H. Tang, *Applied Physics Letters* **2015**, *107*, 262110.

[100] P. G. Li, Y. S. Zhi, P. C. Wang, Z. B. Sun, L. H. Li, Y. H. An, D. Y. Guo, W. H. Tang, J. H. Xiao, *Applied Physics A* **2016**, *122*, 663.

[101] J. Gonzalo, C. Afonso, J. Perrie‘re, R. Go´mez San Roman, *Applied Surface Science* **1996**, *96–98*, 693–696.

[102] A. Petitmangin, C. Hébert, J. Perrière, B. Gallas, L. Binet, P. Barboux, P. Vermaut, *Journal of Applied Physics* **2011**, *109*, 013711.

[103] M. Martin, R. Dronskowski, J. Janek, K.-D. Becker, D. Roehrens, J. Brendt, M. W. Lumey, L. Nagarajan, I. Valov, A. Börger, *Progress in Solid State Chemistry*, Ionic Solid Solutions – Mixing in the Solid State **2009**, *37*, 132–152.

[104] F. Zhang, H. Li, Y.-T. Cui, G.-L. Li, Q. Guo, *AIP Advances* **2018**, *8*, 045112.

[105] M. D. Heinemann, J. Berry, G. Teeter, T. Unold, D. Ginley, *Applied Physics Letters* **2016**, *108*, 022107.

[106] Y. Chen, O. Sakata, H. Morita, A. Matsuda, F. Jia, O. Seo, L. S. R. Kumara, T. Ina, E. Kobayashi, J. Kim, C. Song, S. Hiroi, N. Palina, Y. Lou, W. Ren, M. Yoshimoto, *Applied Surface Science* **2022**, *578*, 151943.

[107] M. Passlack, E. F. Schubert, W. S. Hobson, M. Hong, N. Moriya, S. N. G. Chu, K. Konstadinidis, J. P. Mannaerts, M. L. Schnoes, G. J. Zydzik, *Journal of Applied Physics* **1995**, *77*, 686–693.

[108] Z. Yu, C. D. Overgaard, R. Droopad, M. Passlack, J. K. Abrokwah, *Applied Physics Letters* **2003**, *82*, 2978–2980.

[109] R. O'Donoghue, J. Rechmann, M. Aghaee, D. Rogalla, H.-W. Becker, M. Creatore, A. D. Wieck, A. Devi, *Dalton Transactions* **2017**, *46*, 16551–16561.

[110] D. J. Comstock, J. W. Elam, *Chemistry of Materials* **2012**, *24*, 4011–4018.

[111] M. Nieminen, L. Niinistö, E. Rauhala, *J. Mater. Chem.* **1996**, *6*, 27–31.

[112] J. B. Varley, J. R. Weber, A. Janotti, C. G. Van De Walle, *Applied Physics Letters* **2010**, *97*, 142106.

[113] T. Zacherle, PhD thesis, RWTH Aachen University, Aachen, Germany, **2012**.

[114] T. Zacherle, P. C. Schmidt, M. Martin, *Physical Review B* **2013**, *87*, 235206.

[115] S. Bhowmick, R. Saha, M. Mishra, A. Sengupta, S. Chattopadhyay, S. Chakrabarti, *Materials Today Communications* **2022**, *33*, 104766.

[116] M. Orita, H. Ohta, M. Hirano, H. Hosono, *Applied Physics Letters* **2000**, *77*, 4166–4168.

[117] D. Shiojiri, D. Fukuda, R. Yamauchi, N. Tsuchimine, K. Koyama, S. Kaneko, A. Matsuda, M. Yoshimoto, *Applied Physics Express* **2016**, *9*, 105502.

[118] A. Ortiz, J. C. Alonso, E. Andrade, C. Urbiola, *Journal of The Electrochemical Society* **2001**, *148*, F26.

[119] J. Hao, M. Cocivera, *Journal of Physics D: Applied Physics* **2002**, *35*, 433–438.

[120] M. F. Al-Kuhaili, S. M. A. Durrani, E. E. Khawaja, *Applied Physics Letters* **2003**, *83*, 4533–4535.

[121] F. K. Shan, G. X. Liu, W. J. Lee, G. H. Lee, I. S. Kim, B. C. Shin, *Journal of Applied Physics* **2005**, *98*, 023504.

[122] H. Kröncke, F. Maudet, S. Banerjee, J. Albert, S. Wiesner, V. Deshpande, C. Dubourdieu, *Journal of Vacuum Science & Technology A* **2021**, *39*, 052408.

[123] J. Kim, T. Sekiya, N. Miyokawa, N. Watanabe, K. Kimoto, K. Ide, Y. Toda, S. Ueda, N. Ohashi, H. Hiramatsu, H. Hosono, T. Kamiya, *NPG Asia Materials* **2017**, *9*, e359.

[124] S. Cui, Z. Mei, Y. Zhang, H. Liang, X. Du, *Advanced Optical Materials* **2017**, *5*, 1700454.

[125] L. Xiong, L. Zhang, Q. Lv, T. Li, W. Song, J. Si, W. Zhu, L. Wang, *Semiconductor Science and Technology* **2021**, *36*, 045010.

[126] Y. Chen, Y. Lu, X. Yang, S. Li, K. Li, X. Chen, Z. Xu, J. Zang, C. Shan, *Materials Today Physics* **2021**, *18*, 100369.

[127] H. J. Lee, J.-H. Kim, J. Choi, Y. S. Kim, S.-N. Lee, *Heliyon* **2023**, *9*, e23157.

[128] J. E. N. Swallow, R. G. Palgrave, P. A. E. Murgatroyd, A. Regoutz, M. Lorenz, A. Hassa, M. Grundmann, H. von Wenckstern, J. B. Varley, T. D. Veal, *ACS Applied Materials & Interfaces* **2021**, *13*, 2807–2819.

[129] V. M. Pereira, F. Guinea, J. M. B. Lopes dos Santos, N. M. R. Peres, A. H. Castro Neto, *Physical Review Letters* **2006**, *96*, 036801.

[130] R. Atta-Fynn, P. Biswas, P. Ordejón, D. A. Drabold, *Physical Review B* **2004**, *69*, 085207.

[131] A. Usseinov, Z. Koishybayeva, A. Platonenko, V. Pankratov, Y. Suchikova, A. Akilbekov, M. Zdorovets, J. Purans, A. I. Popov, *Materials* **2021**, *14*, 7384.

[132] H. Zhou, *Journal of Materials Chemistry C* **2019**.

[133] N. Ueda, H. Hosono, R. Waseda, H. Kawazoe, *Applied Physics Letters* **1997**, *70*, 3561–3563.

[134] Y. Xiang, C. Zhou, W. Wang, *Journal of Alloys and Compounds* **2017**, *699*, 1192–1198.

[135] L. Yuan, S. Li, G. Song, X. W. Sun, X. Zhang, *Journal of Materials Science: Materials in Electronics* **2021**, *32*, 8347–8353.

[136] T.-H. Yang, H. Fu, H. Chen, X. Huang, J. Montes, I. Baranowski, K. Fu, Y. Zhao, *Journal of Semiconductors* **2019**, *40*, 012801.

[137] Z. Dela Cruz, C. Hou, R. F. Martinez-Gazoni, R. J. Reeves, M. W. Allen, *Applied Physics Letters* **2022**, *120*, 083503.

[138] A. Parisini, P. Mazzolini, O. Bierwagen, C. Borelli, K. Egbo, A. Sacchi, M. Bosi, L. Seravalli, A. Tahraoui, R. Fornari, *Journal of Vacuum Science & Technology A* **2022**, *40*, 042701.

[139] K. Matsuzaki, H. Hiramatsu, K. Nomura, H. Yanagi, T. Kamiya, M. Hirano, H. Hosono, *Thin Solid Films*, Proceedings of the Fourth International Symposium on Transparent Oxide Thin Films for Electronics and Optics (TOEO-4) **2006**, *496*, 37–41.

[140] K. D. Chabak, K. D. Leedy, A. J. Green, S. Mou, A. T. Neal, T. Asel, E. R. Heller, N. S. Hendricks, K. Liddy, A. Crespo, N. C. Miller, M. T. Lindquist, N. A. Moser, R. C. Fitch, D. E. Walker, D. L. Dorsey, G. H. Jessen, *Semiconductor Science and Technology* **2019**, *35*, 013002.

[141] N. Blumenschein, T. Paskova, J. F. Muth, *physica status solidi (a)* **2019**, *216*, 1900098.

[142] S. Jiao, H. Lu, X. Wang, Y. Nie, D. Wang, S. Gao, J. Wang, *ECS Journal of Solid State Science and Technology* **2019**, *8*, Q3086.

[143] J. Zhang, J. Willis, Z. Yang, X. Lian, W. Chen, L.-S. Wang, X. Xu, T.-L. Lee, L. Chen, D. O. Scanlon, K. H. L. Zhang, *Cell Reports Physical Science* **2022**, *3*, 100801.

[144] P. D. C. King, I. McKenzie, T. D. Veal, *Applied Physics Letters* **2010**, *96*, 062110.

[145] Y. Qin, M. Stavola, W. B. Fowler, P. Weiser, S. J. Pearton, *ECS Journal of Solid State Science and Technology* **2019**, *8*, Q3103.

[146] M. Stavola, W. B. Fowler, Y. Qin, P. Weiser, S. Pearton in *Gallium Oxide*, Metal Oxides Series, Elsevier, Amsterdam, **2019**, pp. 191–210.

[147] M. Fleischer, H. Meixner, *Sensors and Actuators B: Chemical* **1991**, *4*, 437–441.

[148] M. Lorenz, J. Woods, R. Gambino, *Journal of Physics and Chemistry of Solids* **1967**, *28*, 403–404.

[149] P. Rao, S. Kumar, *Superlattices and Microstructures* **2014**, *70*, 117–130.

[150] V. M. Kalygina, T. Z. Lygdenova, V. A. Novikov, Yu. S. Petrova, A. V. Tsymbalov, T. M. Yaskevich, *Semiconductors* **2019**, *53*, 388–394.

[151] H. J. von Bardeleben, J. L. Cantin, A. Parisini, A. Bosio, R. Fornari, *Physical Review Materials* **2019**, *3*, 084601.

[152] H. J. von Bardeleben, J. L. Cantin, *Journal of Applied Physics* **2020**, *128*, 125702.

[153] L. Friedman, M. Pollak, *Philosophical Magazine B* **1981**, *44*, 487–507.

[154] M. S. Reistad, MSc thesis, University of Oslo, Oslo, Norway, **2019**.

[155] P. Mandal, S. Roy, U. P. Singh, *Optical and Quantum Electronics* **2022**, *54*, 476.

[156] H. Ebrahimi-Darkhaneh, Z. Shamsi, M. G. R. Banda, M. Quevedo-Lopez, L. Colombo, S. K. Banerjee, *physica status solidi (a)* **2022**, *219*, 2200054.

[157] T. Yeh, Q. Zhu, D. Buchholz, A. Martinson, R. Chang, T. Mason, *Applied Surface Science* **2015**, *330*, 405–410.

[158] *CRC Handbook of Chemistry and Physics*, 95th ed., (Eds.: W. M. Haynes, D. R. Lide, T. J. Bruno), CRC Press, Taylor & Francis Group, Boca Raton, Florida, **2014**.

[159] R. Schlaf, H. Murata, Z. Kafafi, *Journal of Electron Spectroscopy and Related Phenomena* **2001**, *120*, 149–154.

[160] M. Helander, M. Greiner, Z. Wang, Z. Lu, *Applied Surface Science* **2010**, *256*, 2602–2605.

[161] Y. Park, V. Choong, Y. Gao, B. R. Hsieh, C. W. Tang, *Applied Physics Letters* **1996**, *68*, 2699–2701.

[162] F. Nüesch, L. J. Rothberg, E. W. Forsythe, Q. T. Le, Y. Gao, *Applied Physics Letters* **1999**, *74*, 880–882.

[163] J. Kim, B. Lägel, E. Moons, N. Johansson, I. Baikie, W. Salaneck, R. Friend, F. Cacialli, *Synthetic Metals* **2000**, *111–112*, 311–314.

[164] A. Fiedler, R. Schewski, Z. Galazka, K. Irmscher, *ECS Journal of Solid State Science and Technology* **2019**, *8*, Q3083.

[165] M. Passlack, N. E. J. Hunt, E. F. Schubert, G. J. Zydzik, M. Hong, J. P. Mannaerts, R. L. Opila, R. J. Fischer, *Applied Physics Letters* **1994**, *64*, 2715–2717.

[166] L. Zhang, Z.-J. Tang, *Physical Review B* **2004**, *70*, 174306.

[167] W. Hu, Y. Liu, R. L. Withers, T. J. Frankcombe, L. Norén, A. Snashall, M. Kitchin, P. Smith, B. Gong, H. Chen, J. Schiemer, F. Brink, J. Wong-Leung, *Nature Materials* **2013**, *12*, 821–826.

[168] K. Tsuji, H. Han, S. Guillemet-Fritsch, C. A. Randall, *Physical Chemistry Chemical Physics* **2017**, *19*, 8568–8574.

[169] L.-W. Wang, C.-W. Huang, K.-J. Lee, S.-Y. Chu, Y.-H. Wang, *Nanomaterials* **2023**, *13*, 1851.

[170] J. Uhlendorf, Z. Galazka, H. Schmidt, *Applied Physics Letters* **2021**, *119*, 242106.

[171] J. Uhlendorf, H. Schmidt, *Physical Review Materials* **2023**, *7*, 093402.

[172] S.-Y. Xu, M. Yu, D.-Y. Yuan, B. Peng, L. Yuan, Y.-M. Zhang, R.-X. Jia, *Chinese Physics B* **2024**, *33*, 017302.

[173] A. Azarov, V. Venkatachalapathy, E. V. Monakhov, A. Y. Kuznetsov, *Applied Physics Letters* **2021**, *118*, 232101.

[174] A. Kyrtsos, M. Matsubara, E. Bellotti, *Physical Review B* **2017**, *95*, 245202.

[175] F. Strauß, L. Dörrer, T. Geue, J. Stahn, A. Koutsioubas, S. Mattauch, H. Schmidt, *Physical Review Letters* **2016**, *116*, 025901.

[176] J.-H. Kim, H. J. Lee, H.-J. Kim, J. Choi, J.-H. Oh, D.-C. Choi, J. Byun, S.-E. Ahn, S.-N. Lee, *Advanced Electronic Materials* **2024**, 2300863.

[177] J.-J. Huang, T.-C. Chang, J.-B. Yang, S.-C. Chen, P.-C. Yang, Y.-T. Chen, H.-C. Tseng, S. M. Sze, A.-K. Chu, M.-J. Tsai, *IEEE Electron Device Letters* **2012**, *33*, 1387–1389.

[178] Z. Yang, J. Wu, P. Li, Y. Chen, Y. Yan, B. Zhu, C. S. Hwang, W. Mi, J. Zhao, K. Zhang, R. Guo, *Ceramics International* **2020**, *46*, 21141–21148.

[179] J. A. Röhr, D. Moia, S. A. Haque, T. Kirchartz, J. Nelson, *Journal of Physics: Condensed Matter* **2018**, *30*, 105901.

[180] R. Waser, M. Aono, *Nature Materials* **2007**, *6*, 833–840.

[181] *Hard X-ray Photoelectron Spectroscopy (HAXPES)*, (Ed.: J. Woicik), Springer International Publishing, Cham, **2016**.

[182] J. E. Rault, C.-T. Kuo, H. P. Martins, G. Conti, *Journal of Vacuum Science & Technology A* **2021**, *39*, 31.

[183] G.-L. Li, F. Zhang, Y.-T. Cui, H. Oji, J.-Y. Son, Q. Guo, *Applied Physics Letters* **2015**, *107*, 022109.

[184] C. Surdu-Bob, S. Saied, J. Sullivan, *Applied Surface Science* **2001**, *183*, 126–136.

[185] G. Schön, *Journal of Electron Spectroscopy and Related Phenomena* **1973**, *2*, 75–86.

[186] M. P. Seah, G. C. Smith, M. T. Anthony, *Surface and Interface Analysis* **1990**, *15*, 293–308.

[187] F. Zhang, A. Vollmer, J. Zhang, Z. Xu, J. Rabe, N. Koch, *Organic Electronics* **2007**, *8*, 606–614.

[188] L. J. van der Pauw, *Philips Res. Repts* **1958**, *13*, 1–9.

[189] E. Engel, R. M. Dreizler, *Density Functional Theory: An Advanced Course*, Springer Berlin Heidelberg, Berlin, Heidelberg, **2011**.

[190] D. S. Sholl, J. A. Steckel, *Density Functional Theory: A Practical Introduction*, 1st ed., Wiley, **2009**.

[191] G. Kresse, J. Hafner, *Physical Review B* **1993**, *47*, 558–561.

[192] G. Kresse, J. Hafner, *Physical Review B* **1994**, *49*, 14251–14269.

[193] G. Kresse, J. Furthmüller, *Computational Materials Science* **1996**, *6*, 15–50.

[194] G. Kresse, J. Furthmüller, *Physical Review B* **1996**, *54*, 11169–11186.

[195] P. E. Blöchl, *Physical Review B* **1994**, *50*, 17953–17979.

[196] G. Kresse, D. Joubert, *Physical Review B* **1999**, *59*, 1758–1775.

[197] J. P. Perdew, K. Burke, M. Ernzerhof, *Physical Review Letters* **1996**, *77*, 3865–3868.

[198] M. Ortuño, F. Estellés-Duart, A. M. Somoza, *physica status solidi (b)* **2021**, *259*, 2100340.

[199] D. N. Tsigankov, A. L. Efros, *Physical Review Letters* **2002**, *88*, 176602.

[200] D. M. Kaplan, V. A. Sverdlov, K. K. Likharev, *Physical Review B* **2003**, *68*, 045321.

[201] S. Ihnatsenka, X. Crispin, I. V. Zozoulenko, *Physical Review B* **2015**, *92*, 035201.

[202] G. Schönherr, H. Bässler, M. Silver, *Le Journal de Physique Colloques* **1981**, *42*, C4-111-C4–114.

[203] J. P. Gonzalez-Vazquez, J. A. Anta, J. Bisquert, *Physical Chemistry Chemical Physics* **2009**, *11*, 10359.

[204] P. Hein, *VHoppAS: Variable-range Hopping in Amorphous Solids*, version 2.0.0, Zenodo, **2024**, DOI 10.5281/zenodo.10974845.

[205] S. Baranovskii, B. Cleve, R. Hess, P. Thomas, *Journal of Non-Crystalline Solids* **1993**, *164*, 437–440.

[206] M. S. Alnaes, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. N. Richardson, J. Ring, M. E. Rognes, G. N. Wells, *Archive of Numerical Software* **2015**, *3*, 9–23.

[207] *Automated Solution of Differential Equations by the Finite Element Method*, (Eds.: A. Logg, K.-A. Mardal, G. N. Wells), Springer, **2012**.

[208] A. Logg, G. N. Wells, *ACM Transactions on Mathematical Software* **2010**, *37*, 1–28.

[209] A. Logg, G. N. Wells, J. Hake in *Automated Solution of Differential Equations by the Finite Element Method*, (Eds.: A. Logg, K.-A. Mardal, G. N. Wells), Lecture Notes in Computational Science and Engineering, Springer, **2012**.

[210] R. C. Kirby, A. Logg, *ACM Transactions on Mathematical Software* **2006**, *32*, 417–444.

[211] A. Logg, K. B. Ølgaard, M. E. Rognes, G. N. Wells in *Automated Solution of Differential Equations by the Finite Element Method*, (Eds.: A. Logg, K.-A. Mardal, G. N. Wells), Lecture Notes in Computational Science and Engineering, Springer, **2012**.

[212] K. B. Ølgaard, G. N. Wells, *ACM Transactions on Mathematical Software* **2010**, *37*, 1–23.

[213] R. C. Kirby, *ACM Transactions on Mathematical Software* **2004**, *30*, 502–516.

[214] R. C. Kirby in *Automated Solution of Differential Equations by the Finite Element Method*, (Eds.: A. Logg, K.-A. Mardal, G. N. Wells), Lecture Notes in Computational Science and Engineering, Springer, **2012**.

[215] M. S. Alnaes, A. Logg, K. B. Ølgaard, M. E. Rognes, G. N. Wells, *ACM Transactions on Mathematical Software* **2014**, *40*, 1–37.

[216] P. Vogt, O. Bierwagen, *Applied Physics Letters* **2016**, *108*, 072101.

[217] C. Petersen, S. Vogt, M. Kneiß, H. Von Wenckstern, M. Grundmann, *APL Materials* **2023**, *11*, 061122.

[218] A. Plech, U. Klemradt, H. Metzger, J. Peisl, *Journal of Physics: Condensed Matter* **1998**, *10*, 971–982.

[219] S.-Y. Tang, C. Tabor, K. Kalantar-Zadeh, M. D. Dickey, *Annual Review of Materials Research* **2021**, *51*, 381–408.

[220] V. B. Kumar, A. Gedanken, Z. Porat, *Ultrasonics Sonochemistry* **2015**, *26*, 340–344.

[221] N. Syed, A. Zavabeti, M. Mohiuddin, B. Zhang, Y. Wang, R. S. Datta, P. Atkin, B. J. Carey, C. Tan, J. van Embden, A. S. R. Chesman, J. Z. Ou, T. Daeneke, K. Kalantar-zadeh, *Advanced Functional Materials* **2017**, *27*, 1702295.

[222] W. Schirmacher, *Physical Review B* **1990**, *41*, 2461–2468.

[223] A. J. Scholten, J. I. Dijkhuis, *Physical Review B* **1996**, *53*, 3837–3840.

[224] T. Damker, V. Bryksin, H. Böttger, *Physica B: Condensed Matter* **1999**, *263*, 133–135.

[225] W. Lv, A. Henry, *Scientific Reports* **2016**, *6*, 35720.

[226] G. McArdle, I. V. Lerner, *Scientific Reports* **2021**, *11*, 24293.

[227] B. I. Shklovskii, *Soviet physics. Semiconductors* **1976**, *10*, 855–860.

[228] T. Hayashi, Y. Tokura, A. Fujiwara, *Physica B: Condensed Matter* **2018**, *541*, 19–23.

[229] N. Apsley, H. P. Hughes, *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics* **1975**, *31*, 1327–1339.

[230] M. Pollak, I. Riess, *Journal of Physics C: Solid State Physics* **1976**, *9*, 2339–2352.

[231] T. Miyatani, Y. Nishi, T. Kimoto, *Japanese Journal of Applied Physics* **2019**, *58*, 090914.

[232] H. J. H. Ma, J. F. Scott, *Physical Review Letters* **2020**, *124*, 146601.

[233] R. M. Hill, *Physica Status Solidi (a)* **1976**, *34*, 601–613.

[234] A. V. Nenashev, F. Jansson, S. D. Baranovskii, R. Österbacka, A. V. Dvurechenskii, F. Gebhard, *Physical Review B* **2010**, *81*, 115203.

[235] A. Awada, PhD thesis, Technische Universität Carolo-Wilhelmina zu Braunschweig, Braunschweig, Germany, **2012**.

[236] J. Van De Lagemaat, N. Kopidakis, N. R. Neale, A. J. Frank, *Physical Review B* **2005**, *71*, 035304.

[237] J. Crank, *The Mathematics of Diffusion*, 2d ed, Clarendon Press, Oxford, [Eng], **1975**, 414 pp.

[238] J. Spiechowicz, I. G. Marchenko, P. Hänggi, J. Łuczka, *Entropy* **2022**, *25*, 42.

Abbreviations

AC	Alternating current
ADF	Annular dark-field
AFM	Atomic force microscopy
a-GaO_x	Amorphous, non-stoichiometric gallium oxide (with O/Ga ratio $x < 1.5$)
AI	Artificial intelligence
AIMD	Ab-initio molecular dynamics
ALD	Atomic layer deposition
approx.	approximately
BE	Bottom electrode
C-AFM	Conductive atomic force microscopy
CB	Conduction band
cf.	compare (*conferatur*)
const.	constant
CT	Contact type
cyc.	Cycle
DC	Direct current
DFT	Density functional theory
DOS	Density of states
eb-Pt	Pt deposited by EBPVD
EBPVD	Electron-beam physical vapor deposition
EDX	Energy-dispersive X-ray spectroscopy
EELS	Electron energy loss spectroscopy
e.g.	for example (*exempli gratia*)
EPMA	Electron probe microanalysis
eq.	Equation
etc.	and other similar things (*et cetera*)
EXAFS	Extended X-ray absorption fine structure
FEM	Finite element method
FIB	Focused ion beam
fig.	Figure
GGA	Generalized gradient approximation
GW	"G" = Green's function, "W" = screened Coulomb interaction
HAXPES	Hard X-ray photoelectron spectroscopy
HI	Terminal of a source-measure unit (plus pole for positive voltage)
HRS	High resistive state

i.e.	that is (*id est*)
IM	Interference microscopy
incl.	including
IPR	Inverse participation ratio
ITO	Indium tin oxide
KMC	Kinetic Monte-Carlo
LO	Terminal of a source-measure unit (minus pole for positive voltage)
LOS	Localization of states
LRS	Low resistive state
MEC	Mobility edge conduction
osc.	oscillating
p.	Page
PAW	Projector augmented wave method
PBE	Perdew-Burke-Ernzerhof functional
PICS	Photoionization cross-section
PLD	Pulsed laser deposition
PPMS	Physical properties measurement system
RBS	Rutherford backscattering spectrometry
RDF	Radial distribution function
RF	Radio frequency
RRAM	Resistive random-access memory
RS	Resistive switching
RT	Room temperature
RTA	Rapid thermal annealing
SAED	Selected area electron diffraction
sec.	Section
SEM	Scanning electron microscopy
SIMS	Secondary-ion mass spectrometry
SKPM	Scanning Kelvin probe microscopy
sp-Pt	Pt deposited by DC magnetron sputtering
STEM	Scanning transmission electron microscopy
tbl.	Table
TE	Top electrode
TEM	Transmission electron microscopy
UPS	Ultraviolet photoelectron spectroscopy
UV-Vis	Ultraviolet and visible light
VB	Valence band
VBM	Valence band maximum
VRH	Variable-range hopping
vs.	versus

w.r.t.	with regard to
XPS	X-ray photoelectron spectroscopy
XRD	X-ray diffraction

Symbols

A	Area	cm^2
α	Localization radius	nm
C	Capacitance	F
D	Fickian diffusion coefficient	cm^2/s
d	Film thickness	nm
D^*	Tracer diffusion coefficient	cm^2/s
e	Elementary charge: 1.602×10^{-19}	C
E	Energy	eV
E_C	Conduction band mobility edge (in crystals: CB edge)	eV
E_V	Valence band mobility edge (in crystals: VB edge)	eV
E_F	Fermi level	eV
ε	Energy without electric potential energy ($E = \varepsilon - e\varphi$)	eV
ϵ_0	Vacuum permittivity: 8.854×10^{-12}	$C\,V^{-1}\,m^{-1}$
ϵ_r	Relative permittivity	
η	Ideality factor	
F	Electric field ($\vec{F} = -\nabla\varphi$, $F = \lvert\vec{F}\rvert$)	V/cm
f_{FD}	Fermi-Dirac distribution	
γ	Activity coefficient	
g	Density of states	cm^{-3}
G	Gibbs free energy	eV
H	Magnetic field strength	$A\,m^{-1}$
$\hat{h}$	Single-electron Hamilton operator	eV
h	Planck constant: 4.135×10^{-15}	eV s
$\hbar$	Reduced Planck constant: 6.582×10^{-16}	eV s
I	Electrical current	A
i	Imaginary unit ($i^2 = -1$)	
j	Electrical current density	A/cm^2
J	Particle flux	$cm^{-2}\,s^{-1}$
k_B	Boltzmann constant: 8.617×10^{-5}	$eV\,K^{-1}$
k	Wave number	cm^{-1}
λ	Wavelength	nm
ℓ_m	Mean free path	cm
m	Particle mass	kg
m^*	Effective electron mass	kg
μ	Chemical potential	eV

$\tilde{\mu}$	Electrochemical potential	eV
N	Number of electrons	
n	Electron density ($n = N/V$)	cm^{-3}
N_{eff}	Number of effective charge carriers	
n_{eff}	Effective charge carrier density ($n_{eff} = N_{eff}/V$)	cm^{-3}
N_h	Number of hops	
N_{st}	Number of localized states	
ν	Frequency	s^{-1}
ν_{sw}	Sweep rate for a voltage cycle	V/s
ω	Angular frequency ($\omega = 2\pi\nu$)	s^{-1}
φ_B	Schottky barrier height	V
φ_{bi}	Built-in potential	V
φ	Electric potential	V
ϕ	Kohn-Sham wave function	
Φ	Thermodynamic factor	
P	Random number, typically uniformly distributed in $[0,1)$	
p	Pressure	Pa
p_{O_2}	Oxygen partial pressure	Pa
ψ	Single-electron wave function	
q	Charge	C
ρ	Charge density	C/cm^3
R	Resistance	Ω
S_e	Seebeck coefficient	V/K
S	Entropy	eV/K
σ	Conductivity	S/cm
σ_L	Localization standard deviation (LOS analysis)	
S_z	Maximum area roughness (= lowest valley to highest peak)	nm
s	Sample standard deviation	
τ	Time interval	s
τ_m	Mean free time	s
T	Temperature	K
T_{eff}	Effective temperature of electrons	K
T_0	Characteristic "temperature" in Mott's VRH law	K
t	Time	s
u	Mobility	$cm^2\,V^{-1}\,s^{-1}$
v_d	Drift velocity (displacement per time)	cm/s
v_g	Group velocity of a wave packet	cm/s
v_p	Phase velocity of a wave packet	cm/s
V	Voltage	V
V	Volume	cm^3

V	Potential energy	eV
W	Work function	eV
x	O/Ga ratio in gallium oxide	