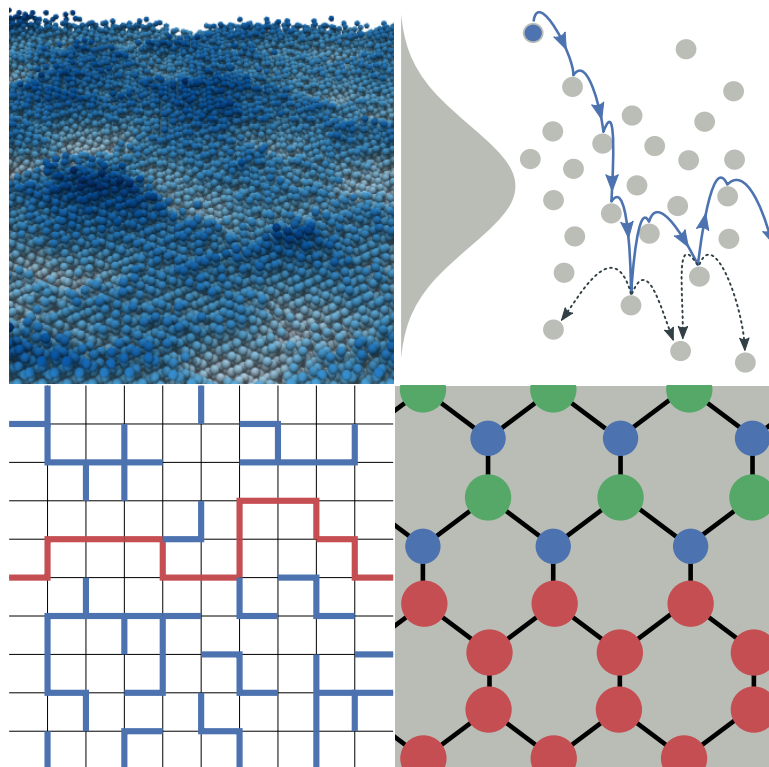# Theoretical Modeling of Kinetic Phenomena of Atoms and Charge Carriers in Disordered Materials

Jan Oliver Oelerich

Theoretical Modeling of Kinetic Phenomena of Atoms and Charge Carriers in Disordered Materials

D ISSERTATION

zur

Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem

Fachbereich Physik
der Philipps-Universität Marburg
vorgelegt von

Jan Oliver Oelerich
aus Köln

Marburg (Lahn), 2015

Vom Fachbereich Physik der Philipps-Universität Marburg
als Dissertation angenommen am 09.09.2015

Erstgutachter: **Prof. Dr. Sergei Baranovski**
Zweitgutachterin: **Prof. Dr. Kerstin Volz**

Tag der mündlichen Prüfung: 25.09.2015

Zusammenfassung

Die Entwicklung integrierter Schaltkreise Ende der 50er Jahre des vergangenen Jahrhunderts [1] hat wie keine andere Entdeckung die digitale Welt geprägt, in der wir heute leben. Die kurz danach vorgestellten Silizium (Si) basierten „complementary metal-oxide-semiconductors" (CMOS) Halbleiterbauelemente führten zu der Entwicklung des Intel 4004, der ersten „central processing unit" (CPU) und damit dem Vorfahren aller heutigen Computer Chips auf Basis von Si-Transistoren [2]. Seit ihrer Entdeckung wird die Leistung der Chips im wesentlichen durch eine immer größer werdende Transistorendichte erhöht. So erreichen die kleinsten Strukturen der im Jahre 2014 auf dem Markt erschienenen Intel Broadwell Architektur gerade einmal eine Größe von 14 nm [3]. In den ständig kleiner werdenden Strukturen spielen jedoch Quanteneffekte eine immer wichtigere Rolle und stellen damit eine natürliche Grenze für die Größe der Bauelemente dar. Zudem hat die digitale Revolution und der damit einhergehende wachsende Energiebedarf eine Fülle von neuen Anwendungen für Halbleiter hervorgebracht, wie etwa flexible Displays, transparente Elektronik oder günstige Solarzellen, deren Anforderungen herkömmliche Si-basierte Chips nicht gerecht werden können. Um die Leistung der Bauelemente weiter zu erhöhen und ihr Anwendungsspektrum zu erweitern, sind deshalb alternative Strategien nötig. Im Fokus der Materialwissenschaft stehen vor allem zwei Ansätze: Die Integration optisch aktiver, funktioneller Halbleiterheterostrukturen mit Si sowie nichtkristalline anorganische und organische Materialien mit geeigneten halbleitenden und optischen Eigenschaften. Die vorliegende Arbeit beschäftigt sich mit der theoretischen und computergestützten Modellierung unterschiedlicher kinetischer Aspekte von Ladungsträgern und Atomen in diesen Materialklassen.

Die Arbeit ist in zwei Teile und einen Anhang gegliedert, in welchem sich ein Literaturverzeichnis und Kopien der in der Arbeit beschriebenen Publikationen finden.

Teil I Der erste Teil der Arbeit behandelt den Ladungstransport in ungeordneten Halbleitern. Als ungeordnete oder *amorphe* Halbleiter bezeichnet man Materialien mit endlicher Bandlücke, deren atomare oder molekulare Struktur nicht durch ein Kristallgitter beschrieben werden kann sondern ungeordnet ist [4]. Die prominentesten Vertreter dieser Materialklasse sind Gläser aus Chalcogenen, amorphe anorganische Halbleiter, wie etwa amorphes Silizium (a-Si), und organische Halbleiter. Insbesondere die kostengünstigen und vergleichsweise einfachen Herstellungsprozesse sowie, im Falle der organischen Halbleiter, besonders flexible und kontrollierbare Materialeigenschaften sind verantwortlich für das große wissenschaftliche und industrielle Interesse an ungeordneten Halbleitern.

Während die elektronischen Wellenfunktionen in kristallinen Materialien durch Bloch-Wellen beschrieben werden können und räumlich ausgedehnt sind [5], gibt es in amorphen Halbleitern, bedingt durch ihre strukturelle Unordnung, eine hohe Dichte an lokalisierten elektronischen Zuständen [4, 5]. Diese lokalisierten Zustände wirken sich insbesondere auf die Mechanismen des Ladungstransports aus: Ladungsträger (Elektronen oder Löcher) bewegen sich entweder durch direktes inkohärentes Tunneln zwischen den Zuständen durch das System (*Hopping-Transport*) oder ihre Bewegung in den ausgedehnten Zuständen des Leitungsbands wird durch häufiges Einfangen in und Entkommen aus lokalisierten Zuständen unterbrochen (*Multiple-Trapping-Transport*) [4]. Diese Mechanismen können mit den für kristalline Halbleiter etablierten theoretischen Ansät-

zen nicht beschrieben werden, weshalb alternative Zugänge und Modelle zur Beschreibung von Ladungstransport in ungeordneten Materialien erforderlich sind.

Kapitel 1 befasst sich mit lokalisierten Zuständen in anorganischen und organischen ungeordneten Halbleitern. Die mathematische Beschreibung der Zustände und ihre räumliche und energetische Verteilung (Zustandsdichte) wird eingeführt. Es wird behandelt, wie sich Ladungsträger innerhalb dieser Zustandsdichte verteilen und die unterschiedlichen Mechanismen des Ladungstransports werden diskutiert. Weiterhin werden die Unterschiede und Gemeinsamkeiten zwischen anorganischen und organischen ungeordneten Halbleitern herausgestellt.

In Kapitel 2 bis 4 werden unterschiedliche theoretische Ansätze zur Beschreibung der Leitfähigkeit und anderer Eigenschaften des Ladungstransports in ungeordneten Systemen vorgestellt. Kapitel 2 behandelt das Transport Energie (TE)-Konzept: Die TE ist das für den Ladungstransport entscheidende Energieniveau innerhalb der lokalisierten Zustandsdichte [6]. Kenntnis der TE erlaubt nicht nur Aussagen über den energetischen Pfad von Ladungsträgern im ungeordneten System, sondern kann auch zur Vorhersage elektrischer Eigenschaften, etwa der elektrischen Leitfähigkeit, verwendet werden [6]. Vorgestellt werden in Kapitel 2 sowohl analytische als auch numerische Zugänge zur TE.

Kapitel 3 befasst sich mit der Beschreibung von Ladungstransport mittels der Perkolationstheorie. Leitende Ladungsträger in ungeordneten Systemen suchen sich die Wege des geringsten Widerstandes. Kennt man denjenigen Wegabschnitt mit dem größten Widerstand innerhalb dieser Wege, lassen sich daraus Transporteigenschaften ableiten [4, 7, 8]. Zur Bestimmung dieser Größen kann man das mathematische Konzept der Perkolationstheorie verwenden. Vorgestellt wird in Kapitel 3 eine Anwendung der Perkolationstheorie auf ein rein räumliches Netzwerkproblem sowie auf Hopping-Transport innerhalb der exponentiellen Zustandsdichte [8].

Erst kürzlich wurde ein drittes Konzept zur Beschreibung von Ladungstransport in ungeordneten Systemen vorgestellt [7]. Fasst man die einzelnen Verbindungen zwischen lokalisierten Zuständen als Netzwerk von Widerständen auf, so lassen sich durch Kenntnis der dissipierten Jouleschen Wärme an den einzelnen Widerständen unterschiedliche Aspekte des Ladungstransportes ermitteln, zum Beispiel das Energieniveau der in Kapitel 2 besprochenen TE. Dieser Ansatz wird in Kapitel 4 behandelt.

Kapitel 5 beschäftigt sich mit einem Modell zur Beschreibung von Multiple-Trapping-Transport in hohen elektrischen Feldern. Die Ansätze werden an Messungen zum Ladungstransport in Bleioxid überprüft und bestätigt. Ein Modell zur Beschreibung von Rekombination zwischen Elektronen und Löchern während des Multiple-Trapping-Transports wird in Kapitel 6 beschrieben.

Kapitel 7 dieses ersten Teils der Arbeit fasst kurz die möglichen Algorithmen zur Computersimulation von Hopping-Transport zusammen und zeigt ihre Vor- und Nachteile auf. Abschließend wird eine Zusammenfassung und ein Fazit zu Teil I in Kapitel 8 gegeben.

Teil II Im zweiten Teil der Arbeit werden Computersimulationen der Strukturbildung und -entwicklung von kristallinen III/V-Halbleitern behandelt. Diese Heterostrukturen, Mischungen aus zwei, drei oder sogar vier Elementen der dritten und fünften Hauptgruppe des Periodensystems, sind vielversprechende Kandidaten für die Funktionalisierung herkömmlicher Si Baugruppen [9]. Während Si ein indirekter Halbleiter ist und damit sehr schlechte optische Eigenschaften hat, weisen einige Vertreter der III/V-Heterostrukturen direkte Bandlücken auf und können daher als optisch aktive Schichten mit Si kombiniert werden. Zudem lassen sich die elektronischen und

strukturellen Eigenschaften dieser Heterostrukturen durch Variation der Mischverhältnisse und durch geeignete Herstellungsverfahren kontrolliert beeinflussen, womit die Materialien auf die gewünschte Anwendung angepasst werden können [9]. Die Herstellung dieser auf Si integrierten Heterostrukturen ist ein komplizierter chemischer Prozess, in welchem die einzelnen Atome mittels Trägermolekülen in der Gasphase angeboten werden, durch Gasphasen- oder Oberflächenreaktionen am Substrat von diesen dissoziieren und anschließend in der energetisch günstigsten Konfiguration in das Kristallgitter integriert werden [9]. Die Reinheit der so gewachsenen Strukturen sowie die Morphologie der Grenzflächen zwischen unterschiedlichen Materialschichten hängt entscheidend von den kinetischen Prozessen der Atome auf dem Substrat während des Wachstums ab.

Eine Einführung über metallorganische Gasphasenepitaxie, die populärste Technik zum Wachstum metastabiler Halbleiter, und die Materialeigenschaften von III/V-Halbleiterheterostrukturen finden sich in Kapitel 9 dieser Arbeit. Ferner wird in diesem einführenden Kapitel ein kurzer historischer Überblick über die theoretische Behandlung solcher Wachstumsprozesse mittels Computersimulationen gegeben.

Kapitel 10 befasst sich mit dem kinetischen Monte-Carlo (KMC) Algorithmus, der in dieser Arbeit zur Simulation von Halbleiterwachstum verwendet wird. Der Algorithmus besteht aus nacheinander ausgeführten Ereignissen der atomaren Kinetik, nämlich Adsorption auf und Desorption von der Substratoberfläche sowie Diffusionsprozessen zwischen zwei Gitterplätzen [10, 11]. Neben diesen elementaren Ereignissen werden das Simulationssystem und die sog. „Bond Counting" Methode zum Berechnen der Energiebarrieren eingeführt.

Die erste Anwendung des Simulationscodes wird in Kapitel 11 beschrieben. Beim experimentellen Wachstum der III/V-Verbindung Galliumphosphid (GaP) auf einem Si-(001) Substrat konnten, trotz der vorher atomar glatten Substratoberfläche, sehr starke Durchmischungseffekte zwischen GaP und Si an der Grenzfläche festgestellt werden. Diese Effekte konnten in der KMC Wachstumssimulation erfolgreich modelliert werden. Die für die Ergebnisse nötigen Energiebarrieren wurden in Dichtefunktionaltheorie (DFT) Rechnungen überprüft und bestätigt.

In Kapitel 12 wird gezeigt, dass die KMC Simulationen in der Lage sind, Tröpfchenbildung und Ätzeffekte metallischen Galliums (Ga) auf Si-(001) Substraten zu reproduzieren. Die in der Simulation in die Si Substratoberfläche geätzten Strukturen entsprechen in Form und Morphologie den experimentellen Beobachtungen. Somit können atomare kinetische Aspekte der Wachstumsprozesse als mögliche Verursacher für die beobachteten Ätzeffekte identifiziert werden.

Im abschließenden Kapitel 13 werden die Ergebnisse zusammengefasst und ein Ausblick auf zukünftige Arbeiten gegeben. Hier wird besonders auf die Integration von Methoden zur Behandlung von Gitterverzerrungen, Zwischengitterdiffusion und flüssig/fest Grenzflächen eingegangen.

Acknowledgements

I very much enjoyed the past three years of doctoral studies at the Physics Department of the Philipps-Universität in Marburg. My research and this thesis would not have been possible without the help and support of a number of people. In particular, I would like to thank…

… Prof. Sergei Baranovski for supervising my work in the past six years, teaching me so much about physics and the scientific community, and for always being available for scientific disucssions.

… Dr. Alexey Nenashev, who is probably the smartest person I know and tought me that there is a solution to almost every problem. Thank you for answering all my questions. It is fun doing research with you!

… Prof. Kerstin Volz for providing the second opinion on my thesis and for the successful collaborations. I enjoyed my participation in the graduate school (GRK 1782) very much and am grateful for your efforts organizing it.

… Prof. Wolfram Heimbrodt for chairing my thesis defense and Prof. Heinz Jänsch for taking part in the examination commission.

… Dr. Andreas Beyer for pleasant discussions, a lot of fun at conferences, and for introducing me to Indiaca.

… Ulrich Huttner for proofreading this work, Katharina Werner for teaching me about MOVPE, and the Theoretical Chemistry group for valuable discussions and for teaching me the basics of DFT.

… the Many Particle Physics group for being a pleasant crowd to be surrounded with at office times. Thank you for providing so much cake and the countless walks to the canteen.

… all my fellow PhD students for the great times in Rauischholzhausen, Rietzlern, San Sebastián, and Hofheim. Special thanks go to those of you who actively attended the Kindergarten seminar.

… the Fachschaft Φ for all the efforts to make our department a good place to study and the amazing SoFe.

… my friends for making Marburg a nice place to live, my family for all the support during my studies, and my dear Julia for putting up with my grumpiness.

… the German Research Foundation (DFG) and the graduate school "Functionalization of Semiconductors" (GRK 1782) for funding.

Author's Contributions

This thesis summarizes my work regarding charge transport in disordered semiconductors (Part I) and the simulation of atom kinetics during epitaxial growth of III/V semiconductors (Part II). Each part consists of an introduction into the topic, followed by summaries of the projects I worked on during my doctoral studies. The related research publications are appended to this work.

Publications in peer-reviewed journals

Article [I]   A. V. Nenashev, **J. O. Oelerich**, and S. D. Baranovskii, "Theoretical tools for the description of charge transport in disordered organic semiconductors", J. Phys. Condens. Matter **27**, 093201 (2015).

Article [II]   A. V. Nenashev, F. Jansson, **J. O. Oelerich**, D. Huemmer, A. V. Dvurechenskii, F. Gebhard, and S. D. Baranovskii, "Advanced percolation solution for hopping conductivity", Phys. Rev. B **87**, 235204 (2013).

Article [III]   **J. O. Oelerich**, F. Jansson, A. V. Nenashev, F. Gebhard, and S. D. Baranovskii, "Energy position of the transport path in disordered organic semiconductors", J. Phys. Condens. Matter **26**, 255801 (2014).

Article [IV]   A. Hofacker, **J. O. Oelerich**, A. V. Nenashev, F. Gebhard, and S. D. Baranovskii, "Theory to carrier recombination in organic disordered semiconductors", J. Appl. Phys. **115**, 223713 (2014).

Article [V]   K. Werner, A. Beyer, **J. O. Oelerich**, S. D. Baranovskii, W. Stolz, and K. Volz, "Structural characteristics of gallium metal deposited on si (001) by MOCVD", J. Cryst. Growth **405**, 102 (2014).

Publications submitted to peer-reviewed journals

Article [VI]   A. Beyer, A. Stegmüller, **J. O. Oelerich**, K. Jandieri, K. Werner, W. Stolz, S. D. Baranovskii, R. Tonner, and K. Volz, "Pyramidal Structure Formation at the Interface between III/V Semiconductors and Silicon", Manuscript submitted for publication in Chemistry of Materials.

Original contributions

During my doctoral studies, I contributed to two rather distinct fields of research, that are separately described in Part I and Part II of this thesis: "Charge transport in disordered semiconductors" and "Computer Simulation of Atom Kinetics in Semiconductor Heterostructures". In what follows, my original contributions to these fields of research are summarized.

I started working on charge transport in disordered semiconductors already in early 2010 as a bachelor student in the group[1]. In my bachelor's thesis, I confirmed the validity of the Einstein-Smoluchowski relation for hopping charge transport in disordered semiconductors [6]. During my master studies, I kept working in the group, focusing on the theoretical foundation of the transport energy (TE) concept, that is described in Chapter 2 of this work [6]. In my master's thesis, I focused on the distinct difference between the exponential and Gaussian density of states (DOS) regarding the dependence of transport coefficients on charge carrier concentration [12], that is described in Section 1.2.3, and developed a Monte Carlo (MC) computer code to simulate hopping charge transport [13]. I kept improving and extending the simulation code during the past years and it became a matured, easily applicable computer package, containing the numerical methods described in Chapters 2, 4, and 7.

I continued working on charge transport in disordered systems during my doctoral studies, starting in October 2012. The first project I was involved in regarded the description of variable range hopping (VRH) with percolation theory, outlined in Section 3.2. Together with Dr. Fredrik Jansson, I was responsible for the computer simulations in the work and for the comparison of the new approach with established theories. The project was published in Article [II]. I presented the results as a poster contribution at the 15th "Transport in Interacting Disordered Systems" (TIDS) conference in Sant Feliu de Guíxols, Spain (2013).

The second project I worked on in this field of research was the numerical determination of the transport energy (TE) level. The idea was to straightforwardly simulate hopping transport with my existing computer code, but distribute the energies of the localized states so that states within some energy range of the density of states (DOS) are absent in the system. The influence of this modification on hopping mobility is then tracked and the TE is found as that particular energy level, that turns out most decisive for transport coefficients. The TE values found in this way agree well with the established analytical optimization approach to the TE, that I worked on during my master studies [6]. Details of the method are described in Section 2.5. The results were published in Article [III]. The numerical calculations concerning the DOS modification approach presented therein have been performed with my simulation code for hopping transport, and the text of the corresponding chapter was mainly written by myself.

In the year 2013, I joined the work on charge carrier recombination during multiple trapping (MT) transport, that was started by Andreas Hofacker for his bachelor studies. In that project, a simple model for bimolecular recombination of excess charge carriers in the exponential and Gaussian DOS was developed, extending the known theory of Orenstein and Kastner [14]. The model was successfully compared to more sophisticated theories of recombination and reproduced known dynamics of certain limiting cases. Chapter 6 of this thesis summarizes the work in more detail and contains most of the calculations. The results were published in Article [IV]. All numerical calculations presented in the article have been done by myself and I was responsible for writing parts of the text.

In the first half of the year 2015, I worked on the dependence of multiple trapping (MT) transport in lead oxide (PbO) on strong external electric fields. I was able to show that a simple MT approach combined with an existing microscopic theory for the field dependent ionization of localized states can reproduce experimentally measured conductivities in PbO. A publication includeing these

---

[1]If not stated otherwise, by "group" I refer to the Disordered Many Particle Physics group at Philipps-Universität, Marburg, led by Prof. Sergei Baranovski.

results is currently in preparation.

Finally, I significantly contributed to the review Article [I], that contains many results of my abovementioned work and gives a comprehensive overview of existing theoretical tools for the description of charge transport in disordered semiconductors. In the article, I was partly responsible for the writing and contributed most of the figures. The structure of this thesis is loosely based on the contents of Article [I], which also contains the novel method of dissipated heat, that was developed by Dr. Alexey Nenashev and is described in Chapter 4 of this thesis.

My work on computer simulation of atom kinetics during epitaxial growth of semiconductor heterostructures began at the end of 2013, with the speculation that kinetic aspects during crystal growth are responsible for the experimentally observed strong intermixing effects at the buried interface between gallium phosphide (GaP) and silicon (Si). Until the beginning of 2014 I was busy developing a Kinetic Monte Carlo (KMC) computer simulation package for the computational modeling of atom kinetics during epitaxial growth of III/V semiconductors. Since then, the simulation code was significantly improved and extended, and matured into a user-friendly high-performance simulation package.

Using the simulation package, I was able to reproduce the intermixing effects of gallium phosphide (GaP) on silicon (Si)-(001) substrates as observed by the WZMW[2] under supervision of Prof. Kerstin Volz and Dr. Wolfgang Stolz, and to identify the driving forces responsible for the intermixing. The presumptions were later confirmed in ab-*initio* density functional theory (DFT) calculations carried out in the theoretical chemistry group of Prof. Ralf Tonner in Marburg. The results are submitted for publication. In the manuscript, I was responsible for the simulations and the writing of the corresponding parts of the text. I presented the results as a talk at the DPG spring meeting in Berlin (2015). Details of the work and a description of my simulation package are given in Chapters 10 and 11.

The simulation code was further applied to reproduce etching effects of gallium (Ga) droplets grown by MOVPE on the Si-(001) substrate. The experiments were also carried out in the WZMW. I was able to reproduce the droplet formation as well as the pyramidally shaped etching structures in the Si substrate with my computer simulation. The findings were published in Article [V] and I presented them as a poster on the 17th International Conference on Metalorganic Vapour Phase Epitaxy (ICMOVPE) in Lausanne, Switzerland (2014). In Article [V], all simulations were carried out by myself and I have written the corresponding parts of the text. The project is described in detail in Chapter 12 of this thesis.

---

[2]Wissenschaftliches Zentrum für Materialwissenschaften, Marburg

Contents

Part I

Charge Transport in Disordered Semiconductors

Chapter 1

Introduction

Although modern electronics are predominantly based on crystalline semiconductors of highest purity, there is a wide range of applications for disordered materials with semiconducting properties, such as chalcogenide glasses, amorphous inorganic semiconductors, or organic disordered semiconductors (ODSs) [4, 5]. These materials are used, for instance, in infrared detectors, inorganic and organic solar cells, optical storage devices, organic light emitting diodes (OLEDs), and organic field effect transistors (OFETs) [15]. Aside from their unique electronic structure, the main advantage of disordered semiconductors are low production costs, simple manufactoring processes, and, in case of ODSs, the possibility of tuning material properties to fit the required application.

The materials have in common, that the long-range order (LRO) of their atomic structure is broken and the atomic array is spatially disordered. This disorder completely changes the electronic structure of the solids, giving rise to the presence of localized electronic states that act as spatially confined traps for charge carriers [4, 5]. The high density of such traps dominates the charge transport mechanisms in disordered semiconductors: Conduction either consists of direct incoherent tunneling transitions (*hopping*) between the traps, or resembles trap-interrupted band transport, so-called *multiple trapping* (MT) [4, 7]. Consequently, the established theoretical framework for the description of electronic properties of crystalline semiconductors fails to describe disordered media, and completely new models and theories must be formulated.

In this Part I of this work, various concepts to describe charge transport in disordered semiconductors are introduced. In the following sections of this introductory chapter, we will get to know the physical framework of disordered organic and inorganic semiconductors. The localized electronic states, that determine the electronic properties, are introduced in Section 1.1. Their spatial and energetic distributions, along with the distribution of charge carriers in the systems, are discussed in Section 1.2. Section 1.3 focuses on the basic charge transport mechanisms, variable range hopping (VRH) and MT, that will be the foundation for the various theoretical concepts presented in the following chapters.

In Chapter 2, the concept of the transport energy (TE) is introduced. An analytical and numerical method to determine the TE is presented, along with a recipe on how to predict charge transport properties within the TE concept. Chapter 3 focuses on the description of hopping transport via percolation theory. The percolation theory is applied to hopping of charge carriers within an array of isoenergetic localized states, as well as to the description of variable range hopping (VRH) in an exponential density of states (DOS). A novel tool to study VRH transport based on the dissipated heat in a random resistor network is presented in Chapter 4. It is shown, how the TE level can be found within this approach and its values are compared to the traditional derivation of the TE presented in Chapter 2. In Chapter 5, we shown how the dependence of charge transport in lead oxide (PbO) on strong electric fields can be modeled using the MT concept, and Chapter 6 focuses on a model for charge carrier recombination during MT transport. Numerical methods to directly simulate hopping charge transport in a computer are briefly described in Chapter 7.

The contents and structures of this and the following chapters are loosely based on the review article by Nenashev et al. [7], that is appended in Article [I].

1.1 Localized Electronic States

For the conventional treatment of perfectly crystalline semiconductors, the translational invariance of the atomic array is an important assumption. It allows one to apply *Bloch's theorem*, stating that electronic wave functions can be expressed in a basis of the so-called *Bloch waves*, which can be factorized into a plane wave and some function $u(\vec{r})$ with the same periodicity as the underlying crystal lattice, $\Phi(\vec{r}) \propto \exp(i\vec{k}\vec{r})u(\vec{r})$ [5]. A solution of the Schrödinger equation with a perdiodic lattice potential in the Bloch wave basis can be obtained analytically and allows access to all kinds of electronic properties of ordered materials, e.g., the band structure or transport characteristics. Within this description, weak disorder can be accounted for by applying scattering theory based on the Boltzmann transport equation [4, 5]. This covers, for example, impurities or defects in crystalline semiconductors, given their concentration is small enough so that the mean free path of carriers between scattering events is much larger than their de Broglie wavelength [4].

Disordered materials, on the other hand, exhibit such a high degree of disorder that none of the abovementioned theories is applicable for the description of electronic phenomena. In non-crystalline disordered semiconductors, a significant amount of the electronic states are spatially localized. In this work, we often refer to such localized states as *sites*. A more detailed discussion of how the localized states arise in different material classes is given in Sections 1.1.1 and 1.1.2.

The wave function $\Psi(\vec{r})$ of a charge carrier trapped in one of these localized states is confined around the position $\vec{R}$ of the site and decays exponentially with distance $|\vec{R} - \vec{r}|$ [4],

$$\Psi(\vec{r}) \propto \exp\left(-\frac{|\vec{R} - \vec{r}|}{\alpha}\right) . \tag{1.1}$$

In Equation (1.1), α is the *localization length* of the electronic state. Although α is in general a non-trivial physical quantity, one usually assumes it equal for all states in the sample and independent on temperature, electric field, and other internal and external parameters. This assumption largely simplifies the theoretical treatment, but may not always be accurate. Typical values of α are in the range $1.0\,\text{Å}$ to $3.0\,\text{Å}$ for organic [4, 16] and inorganic disordered semiconductors [4, 17].

In general, the volume of the localized states is finite and varies between materials and material classes, and the carriers' wave functions are spread out over that volume. However, we will see later that transitions between the sites are so slow, that they completely dominate all charge transport properties. The radius of the sites is therefore unimportant for the theories presented in this work, while the concentration N and the localization length α are essential parameters. Therefore, one usually treats localized electronic states as point-shaped electron or hole traps and neglects the geometry of the sites, which greatly simplifies the calculations. A more in-depth discussion of that approximation can be found in Section 3 of Article [I].

In this work, we focus on two classes of disordered materials: Amorphous inorganic semiconductors and organic disordered semiconductors (ODSs). In the following two sections the chemical structure of these material classes is briefly described and it is discussed, how their localized states are formed and in which ways they differ from each other.
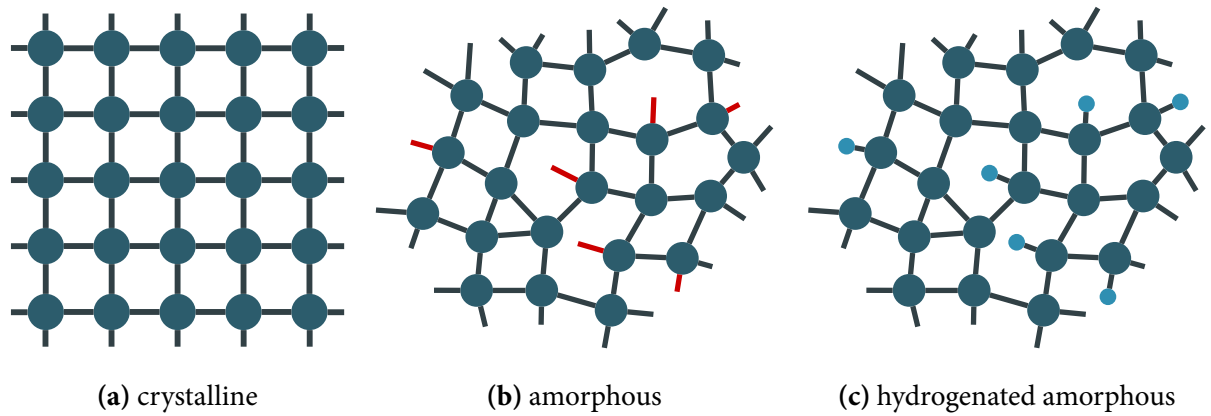
(a) crystalline   **(b)** amorphous   **(c)** hydrogenated amorphous

Figure 1.1: Schematic picture of a crystalline four-fold coordinated material (big circles) with (a) lattice structure and its amorphous counterpart, with (b) dangling bonds and (c) hydrogenated. Hydrogen atoms are marked as bright, small circles, dangling bonds are indicated as red lines.

1.1.1 Amorphous Inorganic Semiconductors

Amorphous inorganic semiconductors can be divided into two classes: There are chalcogenide glasses, which are compounds of chalcogen elements and one or more element serving as kations, and there are the amorphous counterparts of Si and of other conventional semiconductors. In what follows, a brief discussion of the structural characteristics of amorphous Si, the most prominent member of that class, is given, along with a discussion of its electronic density of states (DOS).

Crystalline silicon, c-Si, is an indirect-bandgap semiconductor that dominates today's semiconductor industry [9]. Si has four valence electrons in the outer shell (it is *four-fold coordinated*) and crystallizes in the diamond cubic lattice structure. Its structural and electronic properties are well understood and are easily accessible both theoretically and in experiments. Although a crystalline structure is the thermodynamically most stable state for bulk Si, it is possible to grow Si in a non-equilibrium, metastable state possessing an amorphous atomic structure. In that case, the material is called *amorphous silicon*, a-Si.

Examining the radial distribution function of a-Si (as obtained for example by x-ray diffraction or neutron scattering experiments [17]) reveals, that while the short-range order (SRO) of the Si atoms is similar to that of c-Si, the long-range order (LRO) of the constituents is broken and the momentum $\vec{k}$ is not a good quantum number anymore. This renders the momentum space useless for describing the electronic structure in a-Si. The absence of a LRO is induced by a broad distribution of bond lengths and angles in a-Si. In addition, three- and five-fold coordinated atoms are found along with the usual four-fold coordination, resulting in dangling bonds, which appear as structural defects of a-Si [17]. These dangling bonds can be saturated by adding hydrogen (H) atoms, which also releases some of the strain that is present in the disordered atomic array. In that case, the material is called *amorphous hydrogenated silicon*, a-Si:H [17]. The structure of a four-fold coordinated crystalline material (e.g., c-Si) as compared to its amorphous counterpart is shown schematically in Figure 1.1.

Breaking of the translational invariance and the existence of defects described above has essential implications on the electronic DOS in amorphous materials. The electronic structure of a-Si is composed of three parts: Extended bands, the band tails leaking into the gap, and the dangling
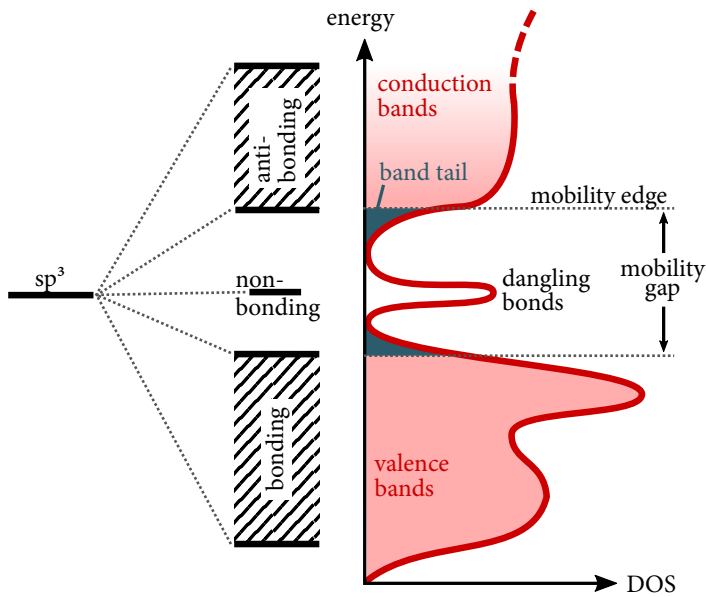
Figure 1.2: The density of states in amorphous semiconductors. The bands of extended states, formed by the sp^3 states of the atomic cluster, are broadened due to the disorder and exponential tails of localized states leak into the bandgap (or *mobility gap*). Structural defects appear as additional localized states in the center of the gap. Reproduced from Reference [17].

bonds' defect states [17]. The DOS of a-Si is depicted in Figure 1.2.

Because the SRO in a-Si is preserved and is similar to that of c-Si, valence and conduction bands still exist. They consist of the usual bonding (valence band) and anti-bonding (conduction band) states of the sp^2 hybridization of s and p orbitals. Electronic wave functions in these states are extended over large volumes of the sample and charge transport in such states is merely disturbed by scattering events at perturbations of the periodic potential landscape.

The second contribution to the DOS is due to the disordered nature of a-Si, i.e., the broad distribution of bond lengths and angles. It can be shown, that (three-dimensional) disordered materials possess a tail of localized electronic states that decreases exponentially from the conduction and valence bands into the bandgap, as shown in Figure 1.2 [18]. The width of the band tail depends on the amount of disorder in the atomic array. It is obvious that these localized states are tightly coupled to those positions in the bulk, where deviations from the equilibrium bond angle or length are the largest. The band tails are separated from the valence and conduction band by the so-called *mobility edges*, above which the conductivity in the extended states of the bands is by orders of magnitude higher than that in the tail states. The transport mechanisms in the band tail and conduction band are discussed in detail in Sections 1.3.1 and 1.3.3. A more in-depth discussion of the DOS of the band tails is given in Section 1.2.1.

Another contribution to the DOS stems from the dangling bonds of the wrongly coordinated Si atoms. Unless they are saturated with H (hydrogenated), these bonds can trap an electron (three-fold coordinated bonds) or a hole (five-fold coordinated bonds) and their electronic states are highly localized. These states are found in the center of the DOS, as depicted in Figure 1.2. Because detrapping from these states is very unlikely and their concentration is comparatively low, we will not take them into account in the theoretical considerations given in this work.

1.1.2 Organic Disordered Semiconductors

The second class of materials that is a subject of this work are organic disordered semiconductors (ODSs). ODSs usually consist of organic molecules that are weakly bound to each other via
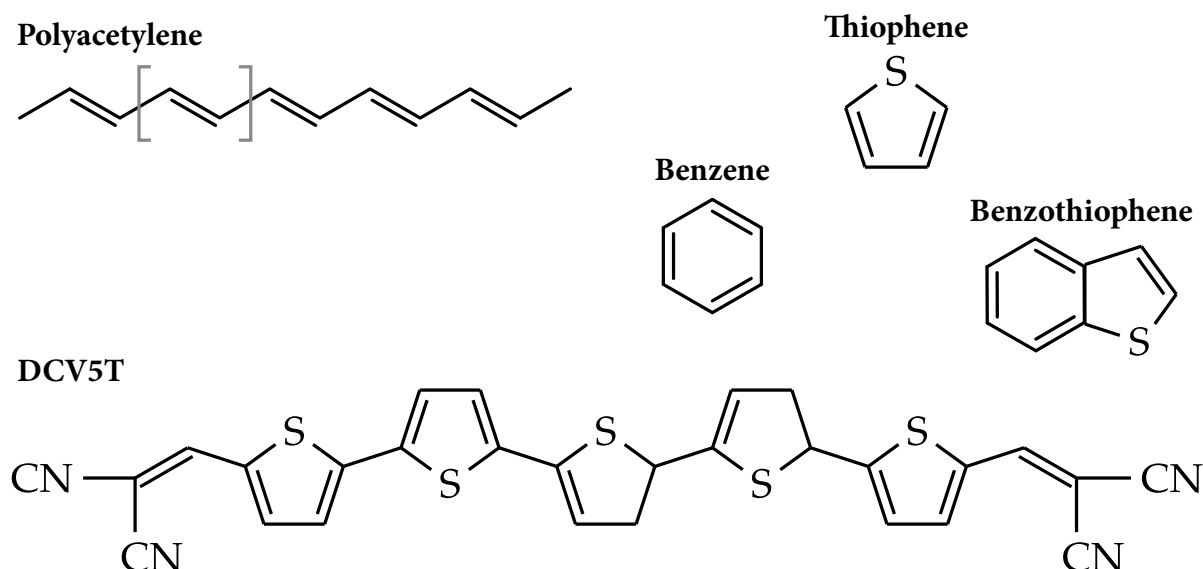
Polyacetylene

Thiophene

Benzene

Benzothiophene

DCV5T

Figure 1.3: The chemical structure of some organic molecules. All of the molecules have a conjugated
π-System with delocalized orbitals. The aromatic rings (Benzene, Thiophene, Benzothiophene) and
the polymer Polyacetylene can be integrated into or combined to larger organic polymers, as for
example the DCV5T.

van der Waals forces [19]. These molecules are based on chains of carbon (C) atoms with an
alternating sequence of single and double σ bonds (connected p orbitals), the so called *conjugated
systems*. Such chains can be either linear, forming a molecular backbone, or appear in closed
aromatic rings (*aromatics*), sometimes containing only C atoms and often with other elements
integrated into the ring (*heterocyclic aromatics* or *aromatic compounds*) [19]. These aromatics and
backbone structures are the basic building blocks for a large variety of organic molecules with
semiconducting properties. Some examples are shown in Figure 1.3. The endless possibilities of
extending and combining organic semiconductors to match the industry's requirements are the
reason for the growing interest of the scientific community in understanding the electrical and
chemical properties of such materials.

In the ground state, the electron configuration of an isolated C atom is $1s^2 2s^2 2p^2$, and that of
chemically bonded C is $1s^2 2s 2p^3$. In the bulk, three of the four valence electrons are found in the
three degenerate sp^2 orbitals that are in plane of the atomic system, and the fourth electron occupies
the p_z orbital perpendicular to that plane. Those sp^2 orbitals form strong σ-bonds in-plane between
the C atoms, while the overlapping p_z orbitals form π-systems, in which electrons are delocalized
over the backbone or atomic ring. An example is shown in Figure 1.4, where the electron density of
the π system of a benzene (C_6H_6) ring is displayed. Due to symmetry breaking, the π orbitals do
not have metallic character, but rather split into a bonding π and an anti-bonding π^* orbital. These
are the *highest occupied molecular orbital* (HOMO) and the *lowest unoccupied molecular orbital*
(LUMO), respectively. The π and π^* orbitals are localized on molecules or molecular segments.
When the molecules of an organic semiconductor are arranged in a periodic pattern, the π and
π^* can form delocalized bands. However, since these bands are so narrow, small perturbations
(due to defects or disorder) cause complete localization [20]. This is in contrast to the broad σ and
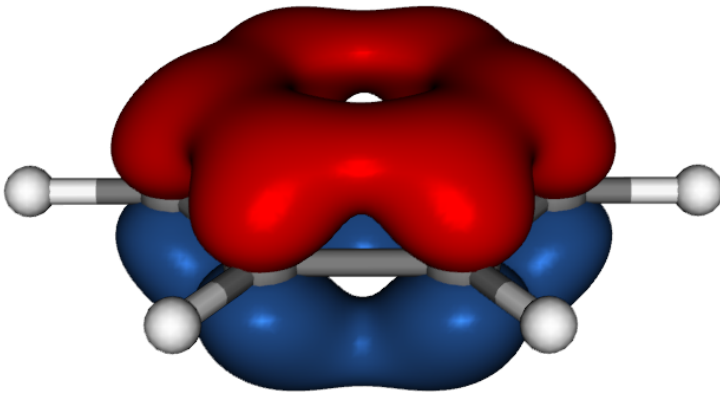
Figure 1.4: The conjugated π system of a benzene ring, an aromatic compound containing six C atoms (grey balls) and six hydrogen atoms (white balls). The π orbitals are delocalized over the ring. The electron densities were calculated with orca [21] using an LDA functional and visualized with Molekel [22].

σ^* bands in inorganic semiconductors, where a band structure is retained even when disorder is present and the localized states appear as band tails. (See Section 1.1.1.)

Although it is possible with some efforts to grow organic crystals, organic semiconductors are usually fabricated in a disordered structure. The electronic structure in such ODSs with spatially randomly distributed localized states is similar to the band tails in amorphous inorganic semiconductors, discussed in Section 1.1.1. However, in ODSs the DOS consists only of localized states without any underlying band structure. This is an important difference between inorganic and organic disordered semiconductors and forces researchers to often treat the two material systems separately. The shape of the DOS in ODSs is close to a Gaussian one [12] in contrast to the exponential DOS in the tails of amorphous inorganic semiconductors. The energy spectra of the two material classes are discussed in more detail in the following Section 1.2.

1.2 The Density of States and Equilibrium Charge Distribution

In the previous section we showed that electronic states in disordered semiconductors, be them organic or inorganic, are at least partly localized. In both material classes the existence of these localized states has large impact on charge transport. It is therefore crucial to understand the energy spectrum in disordered semiconductors for the appropriate theoretical description of transport effects.

In this section, we discuss the DOSs and the (equilibrium) charge carrier distributions in organic and inorganic disordered semiconductors. (The term DOS is used from now on for the spectrum of localized states only.) The DOS is the number of states $g(\varepsilon)\,\mathrm{d}\varepsilon\,\mathrm{d}V$ that are contained in the volume $\mathrm{d}V$ and have energies in the interval $[\varepsilon, \varepsilon + \mathrm{d}\varepsilon]$. Because one usually assumes the DOS homogeneous in space, which is correct up to statistical fluctuations and surface or interface effects, $g(\varepsilon)$ does not explicitly depend on the spatial position. Its energy dependence, however, requires careful investigation.

The normalized DOS $g(\varepsilon)$ is related to the spatial concentration N of sites via

$$N = \int\limits_{-\infty}^{+\infty} g(\varepsilon)\,\mathrm{d}\varepsilon\,. \tag{1.2}$$

While $g(\varepsilon)$ contains information about the energy distribution of localized states in the system, it makes no statement about the energy distribution of charge carriers. If a finite concentration

n of conducting electrons is created in the system, e.g., by injection or photo-excitation, and temperature is not too high, these carriers will relax towards energy states as low as possible (or, for holes, as high as possible) until thermal equilibrium is established. Then, they are distributed homogeneously in space and the probability for a site with energy ε to be occupied by a carrier becomes [5]:

$$f(\varepsilon, \varepsilon_F) = \left(\exp\left(\frac{\varepsilon - \varepsilon_F}{kT} \right) + 1 \right)^{-1} .$$

(1.3)

Equation (1.3) is the Fermi-Dirac distribution. The value of $f(\varepsilon, \varepsilon_F)$ is close to unity if $\varepsilon < \varepsilon_F$ and vanishes in the opposite case $\varepsilon > \varepsilon_F$. ε_F is the *Fermi energy* that (at $T = 0$) separates occupied from unoccupied localized states. If the system is poorly occupied with charge carriers, i.e., all relevant energies are much higher than the Fermi level, $\varepsilon \gg \varepsilon_F$, Equation (1.3) can be approximated by the distribution

$$f(\varepsilon, \varepsilon_F) \approx \exp\left(-\frac{\varepsilon - \varepsilon_F}{kT} \right) .$$

(1.4)

This form of the occupation probability is called the *Boltzmann limit* or *Boltzmann approximation* [15].

Because each energy state can only be occupied by a single charge carrier, the actual distribution of the carriers is the product $g(\varepsilon)f(\varepsilon, \varepsilon_F)$ and the Fermi energy ε_F is determined by the carrier concentration n via

$$n = \int\limits_{-\infty}^{+\infty} g(\varepsilon)f(\varepsilon, \varepsilon_F) \, \mathrm{d}\varepsilon .$$

(1.5)

Amorphous inorganic semiconductors, like a-Si, have been studied by the scientific community for a much longer period than the semiconducting polymers and related organic disordered materials. It is known both from theoretical considerations [18] and experiments (e.g., photo-spectroscopy) [4, 23], that the density of localized states in the tails of the conduction and valence bands of amorphous inorganic semiconductors decreases exponentially towards the center of the bandgap. It is often believed, that similar considerations can be applied to organic media, and an exponential DOS is frequently used to describe disordered organic semiconductors. However, the actual shape of the DOS in organic semiconductors was shown to be close to a Gaussian distribution [7, 12, 15, 20], as discussed in more detail in Section 1.2.3. Because there are some key differences between the exponential and Gaussian DOS that have significant impact on transport properties [12], we will separately discuss the energy spectrum of the two material classes in the following sections and in most parts of this work.

1.2.1 The Exponential DOS

As mentioned before, inorganic amorphous semiconductors possess band tails of localized states that decrease exponentially from the *mobility edge* towards the center of the bandgap, as shown in Figure 1.2. Mathematically, the DOS is described as follows:

$$g(\varepsilon) = \frac{N}{\varepsilon_0} \exp\left(\frac{\varepsilon - \varepsilon_c}{\varepsilon_0} \right) ,$$
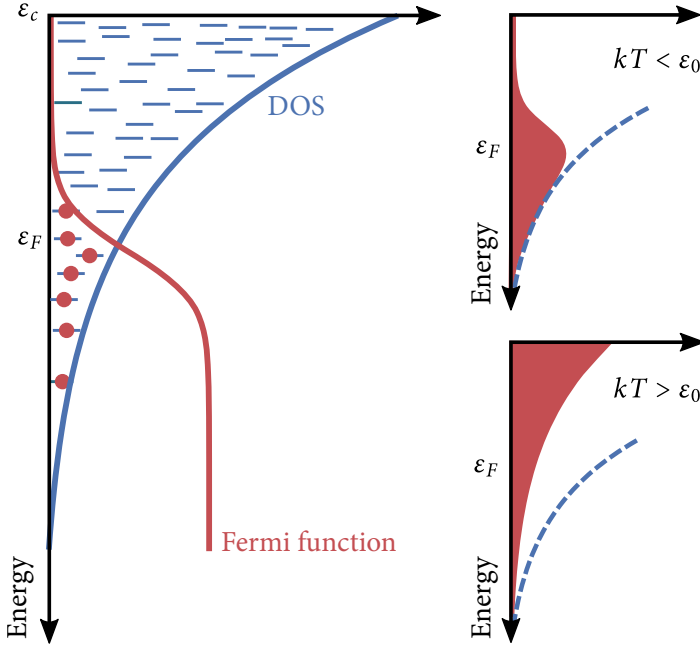
(1.6)

Figure 1.5: Left: The exponential DOS and the Fermi function versus energy. Localized states are indicated as short, blue lines. The states below the Fermi level ε_F are occupied by charge carriers (red circles). Right: Schematic display of the carrier distribution $g(\varepsilon)f(\varepsilon, \varepsilon_F)$ for the cases $kT < \varepsilon_0$ and $kT > \varepsilon_0$. The dashed, blue lines indicate the DOS.

where ε_0 is the energy scale of the system of localized states, that reflects the amount of disorder in the atomic structure, and ε_c is the mobility edge. We focus here on the electronic DOS below the conduction band. Similar considerations are valid for the distribution of hole states above the valence band. For simplicity, one usually counts energies relative to ε_c and therefore sets $\varepsilon_c = 0$. Equation (1.6) is normalized: it fulfills Equation (1.2) with the upper boundary of the integral replaced by $\varepsilon_c = 0$. Typical values of ε_0 for a-Si:H are 18 meV to 22 meV for the conduction band tail and 33 meV to 43 meV for the valence band [4, 17, 24].

If a finite concentration of thermally equilibrated charge carriers is created in a DOS, their energy distribution is described by the product of the DOS $g(\varepsilon)$ and the Fermi function $f(\varepsilon, \varepsilon_F)$ from Equation (1.3). The resulting distribution $g(\varepsilon)f(\varepsilon, \varepsilon_F)$ for the exponential DOS is drawn in Figure 1.5. Its properties depend decisively on the ratio between ε_0 and kT: At low temperatures $\varepsilon_0 > kT$, the exponential tail of the Fermi function decreases more steeply towards $\varepsilon = 0$ than the DOS increases. In this case, the states are occupied up to ε_F and occupation then decreases exponentially. That situation is shown in the upper inset of Figure 1.5. It is obvious, that in this case the majority of charge carriers occupy states around the Fermi level. In the opposite case, $\varepsilon_0 < kT$, the DOS $g(\varepsilon)$ wins over $f(\varepsilon, \varepsilon_F)$ and the occupation of carriers increases towards the mobility edge $\varepsilon = 0$. Then, the majority of charge carriers is situated right below the conduction band, as shown in the lower inset of Figure 1.5.

Typical values of ε_0 for many materials are of the order of kT at room temperature, $kT_{RT} \approx$ 25 meV [4]. Both of the abovementioned cases are therefore relevant for the interpretation of experimental results and must carefully be considered when dealing with the exponential DOS. In PbO, for instance, the scale of disorder ε_0 of hole states is different to that of electronic states, so that each of the two cases described above applies to one species of carriers in the same material. PbO is discussed in detail in Chapter 5.
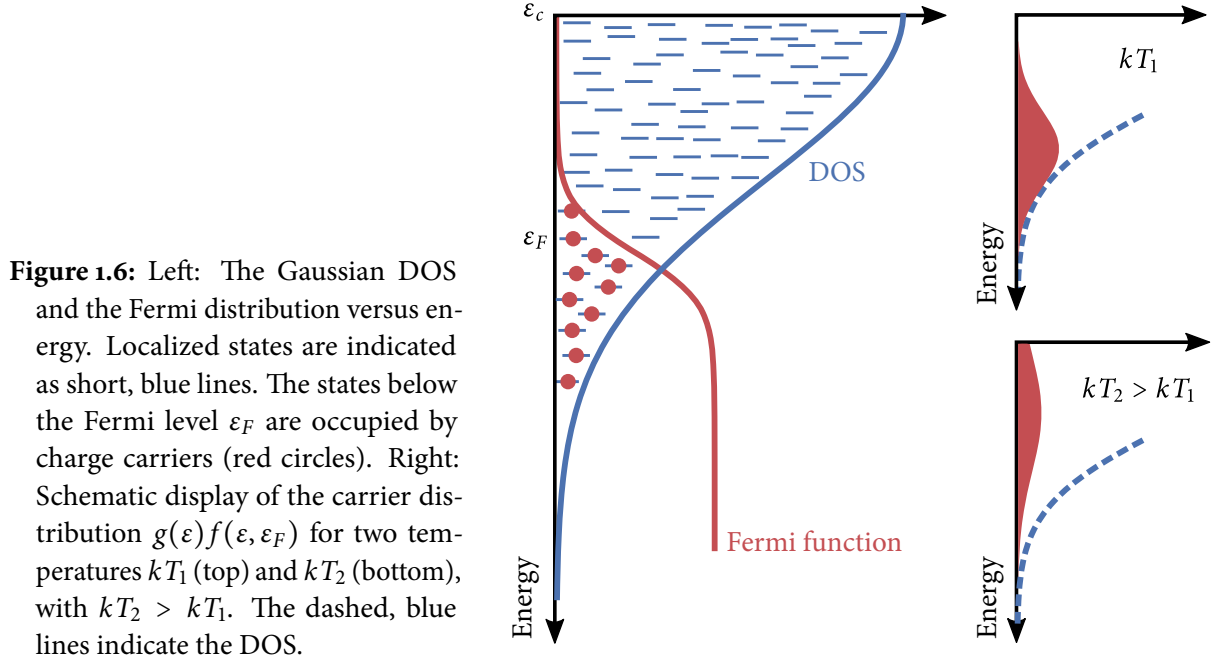
Figure 1.6: Left: The Gaussian DOS and the Fermi distribution versus energy. Localized states are indicated as short, blue lines. The states below the Fermi level ε_F are occupied by charge carriers (red circles). Right: Schematic display of the carrier distribution $g(\varepsilon)f(\varepsilon, \varepsilon_F)$ for two temperatures kT_1 (top) and kT_2 (bottom), with $kT_2 > kT_1$. The dashed, blue lines indicate the DOS.

1.2.2 The Gaussian Disorder Model

A Gaussian DOS in organic disordered semiconductors was first suggested by Gutmann and Lyons [25] and Silinsh [26] and later advertised in two review articles by Bässler [20, 27]. Bässler's argumentation relied on agreement of numerical simulations with experimental data, and even today there is no first-principles theory to determining the energy spectrum in ODSs. While the Gaussian DOS has since been studied extensively and numerous articles were published that point out the key differences between the exponential and Gaussian DOS, the shape of the density of states is still disputed controversially in the scientific community. In addition, it is often claimed that the exponential and Gaussian DOSs are so similar in structure that the actual shape is not important for calculations. Let us show here, that even the small structural differences between the Gaussian and exponential shapes of the DOS result in significantly different charge carrier distributions. Based on these results, we will formulate some strong arguments in favor of the Gaussian DOS in organic semiconductors in Chapter 2.

While the total energy spectrum of inorganic disordered semiconductors includes both extended and localized states (see Section 1.1.1), ODSs contain solely localized states [4, 20]. Their energy spectrum is described by a Gaussian density of states, as [7]

$$g(\varepsilon) = \frac{N}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right), \tag{1.7}$$

where σ is the energy scale of the system and the DOS is centered around $\varepsilon = 0$. (To emphasize the difference between the exponential and Gaussian shapes of the DOS, we use a different character than ε_0 in Equation (1.7).) ODSs typically have broader distributions of energy states than their inorganic counterparts: values for σ reported in the literature vary between 78 meV and 100 meV [20, 28]. The shape of the Gaussian DOS is shown schematically in Figure 1.6.

In ODSs, the energy distribution of charge carrier is significantly different from that in amorphous semiconductors, described in Section 1.2.1. In order to describe this distribution, we study

the product of the occupation probability $f(\varepsilon, \varepsilon_F)$ of sites with the DOS $g(\varepsilon)$. Let us first assume, that the Fermi level ε_F is deep in the tail of the DOS and that we can apply the Boltzmann approximation from Equation (1.4). Then

$$g(\varepsilon)f(\varepsilon, \varepsilon_F) \approx \exp\left(\frac{\varepsilon_F}{kT} - \frac{\varepsilon^2}{2\sigma^2} - \frac{\varepsilon}{kT}\right) = A(\varepsilon_F)\exp\left(-\left[\frac{\varepsilon}{\sqrt{2}\sigma} + \frac{\sigma}{\sqrt{2}kT}\right]^2\right), \tag{1.8}$$

which is a normal distribution with some prefactor $A(\varepsilon_F)$, centered at the energy $\varepsilon = -\sigma^2/kT$. We see from that result, that the carrier distribution has Gaussian shape and that its maximum position is independent of the Fermi level ε_F, as long as the Boltzmann approximation is applicable. The energy $-\sigma^2/kT$, the maximum position of the carrier distribution in Equation (1.8), is called *equilibration energy* ε_∞. As shown in the insets of Figure 1.6, the distribution broadens and shifts to higher energies with rising temperature.

When the Fermi level approaches ε_∞, the above derivation becomes invalid and the energy center of mass of the carrier distribution will become dependent on ε_F. In that case, as often found in the literature, the typical carrier energy is calculated as the energy expectation value of $g(\varepsilon)f(\varepsilon, \varepsilon_F)$ according to

$$\langle\varepsilon\rangle = \frac{\int_{-\infty}^{+\infty} \varepsilon\, g(\varepsilon)f(\varepsilon, \varepsilon_F)\, \mathrm{d}\varepsilon}{\int_{-\infty}^{+\infty} g(\varepsilon)f(\varepsilon, \varepsilon_F)\, \mathrm{d}\varepsilon}, \tag{1.9}$$

which also yields $\langle\varepsilon\rangle = \varepsilon_\infty = -\sigma^2/kT$ in the limit of low ε_F. This change in the dependence of the carrier distribution as a function of the Fermi level can be used to distinguish between the Gaussian and exponential shapes of the DOS in experiments. In the next Section 1.2.3, we discuss this important difference between the two DOS shapes in more detail.

1.2.3 Gaussian against exponential DOS

It is instructive to study the concentration dependence (i.e., dependence on ε_F) of the average carrier energy $\langle\varepsilon\rangle$ given by Equation (1.9) for the Gaussian DOS in comparison to the exponential DOS. The result is shown in Figure 1.7. It is clearly visible in the figure, that the majority of charge carriers in the exponential DOS occupy states slightly below ε_F at all concentrations, while the energy center of mass of the carrier distribution in the Gaussian DOS becomes concentration dependent only when $\varepsilon_F \gtrsim \varepsilon_\infty$.

As hinted above in Section 1.2.2, one expects $\langle\varepsilon\rangle$ to become dependent on the position of the Fermi level, when ε_F crosses the equilibration energy $\varepsilon_\infty = -\sigma^2/kT$, the maximum of Equation (1.8). Hence, the criterion

$$\varepsilon_F(n_c) \approx \varepsilon_\infty \tag{1.10}$$

defines a threshold concentration n_c (determined via Equation (1.5)) of charge carriers in the system, above which the transport properties of the Gaussian DOS are expected to become concentration dependent. This behaviour applies only, when an equilibration energy ε_∞ exists, which is not the case for the exponential DOS, where transport always depends on the carrier concentration n.

A schematic picture of the expected carrier mobility μ in the Gaussian and exponential DOS, reflecting this important difference, is shown in Figure 1.8. Measuring the dependence $\mu(n)$
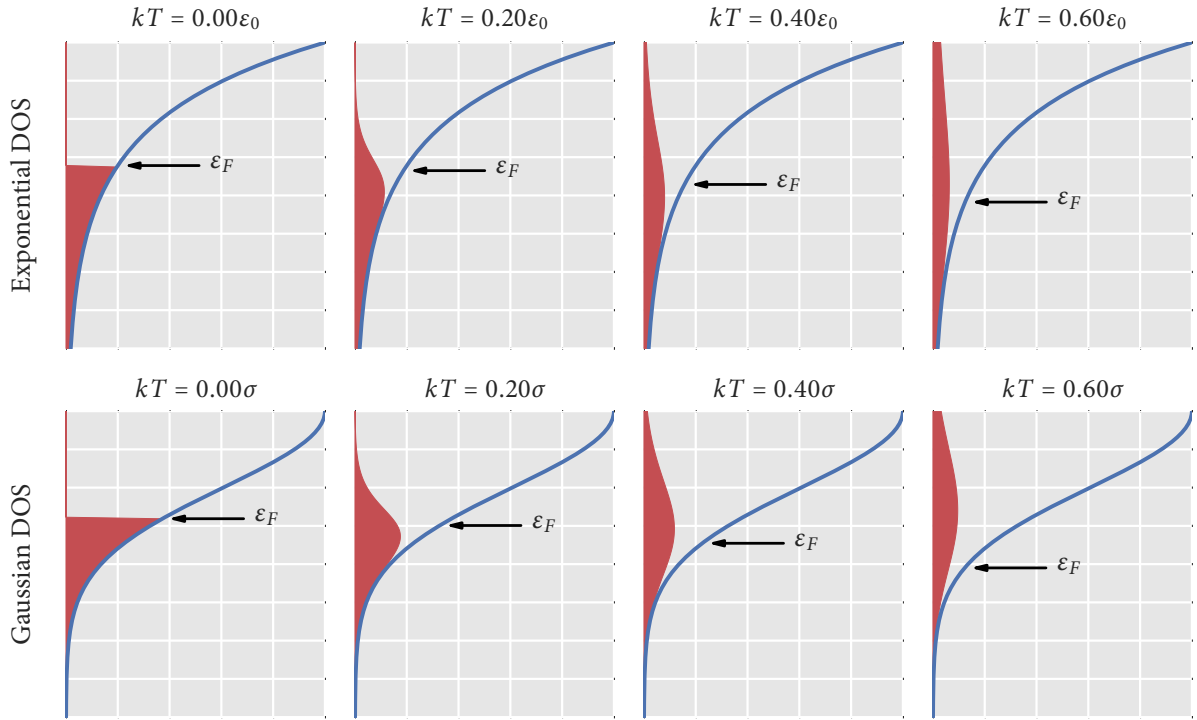
Figure 1.7: The carrier distribution (red shapes) in the exponential and Gaussian DOS (blue lines) at different temperatures. It is clearly visible, that the average carrier energy (given by Equation (1.9)) stays around the Fermi level in the exponential DOS but is shifted to higher energies in the Gaussian DOS with rising temperature. The slightly temperature dependent Fermi level ε_F, determined for fixed concentration n via Equation (1.5), is indicated by arrows.

over a suitable range of concentrations n is therefore a useful tool to gain information about the DOS shape of the system [13]. In fact, the characteristic concentration dependence of transport properties in the Gaussian DOS, as shown in Figure 1.8, is observed in various ODSs [29, 30], which is a strong argument for the existence of Gaussian disorder in such materials.

1.3 Transport Mechanisms in Disordered Semiconductors

In the previous section we studied the distribution of localized electronic states of inorganic (exponential DOS) and organic (Gaussian DOS) disordered semiconductors, along with their respective equilibrium charge carrier distributions. Due to their disordered atomic structure and the absence of translational invariance, the known concepts to describe charge transport in crystalline semiconductors are not applicable in disordered systems. In this section we introduce the basic transport modes in disordered materials.

There are two transport modes that take place in disordered media: *Hopping transport*, consisting of direct, phonon-assisted tunneling transitions between sites in the network, and *multiple trapping* (MT), which is trapping-interrupted transport in the extended states. Because a mobility edge with delocalized transport states is present only in inorganic semiconductors with exponential DOS, transport in ODSs happens exclusively by hopping transitions. The two distinct transport mechanisms are explained in the following sections. Chapters 2 to 4 will deal with transport

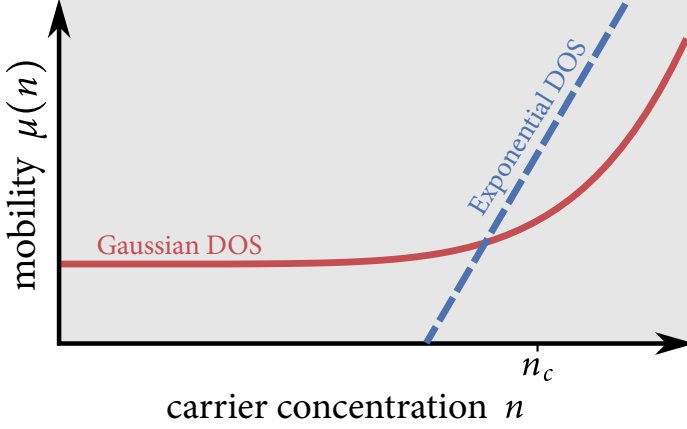**Figure 1.8:** Schematic dependence of the carrier mobility μ on the concentration n of charge carriers for the exponential and Gaussian DOS. The threshold concentration n_c is given by Equation (1.10).

phenomena for hopping transport, and in Chapters 5 and 6 we will study some properties of the MT transport mode.

1.3.1 Hopping Transport

Hopping transport is based on the incoherent tunneling transitions of a charge carrier between localized electronic states, by emission or absorption of phonons with appropriate energies. A hopping transition is depicted in Figure 1.9. The two most commonly used expressions for the hopping rate between an occupied and an unoccupied site are the Miller-Abrahams (MA) hopping rates [31]

$$\nu_{ij} = \nu_0 \exp\left(-\frac{2d_{ij}}{\alpha}\right) \exp\left(-\frac{\varepsilon_j - \varepsilon_i + |\varepsilon_j - \varepsilon_i|}{2kT}\right), \tag{1.11}$$

and the Marcus rates [32]

$$\nu_{ij} = \frac{J_0^2}{\hbar} \sqrt{\frac{\pi}{4E_a kT}} \exp\left(-\frac{2d_{ij}}{\alpha} - \frac{E_a}{kT}\right) \exp\left(-\frac{\varepsilon_j - \varepsilon_i}{2kT} - \frac{(\varepsilon_j - \varepsilon_i)^2}{16kTE_a}\right). \tag{1.12}$$

In both equations, α is the localization length introduced in Section 1.1, kT is thermal energy, d_{ij} is the distance between the two sites, and ε_i and ε_j are their energies. The prefactor ν_0 in Equation (1.11) is the *attempt-to-escape* frequency, which depends on the interaction mechanism inducing the hopping transition. Usually, electron-phonon interaction is assumed and ν_0 takes values of the order of the phonon frequency, 10^{12} s^{-1}. In the Marcus expression, Equation (1.12), J_0 is related to the transfer integral between the sites and E_a accounts for the reorganization energy of the system. Although the Marcus rates are used in the literature sometimes, the majority of studies neglect polaron effects and are based on the Miller-Abrahams (MA) expression, for it is much simpler to work with and proved sufficient in most cases [33]. A thorough derivation of the MA rates can be found in the original article of Miller and Abrahams [31] and in the monograph by Shklovskii and Efros [33, chapter 4.2]. In this work, we will use use the MA rates exclusively.

The form of the rates given in Equation (1.11) does not account for the carrier concentration present in the system, because it assumes that the initial site i is occupied and the target site j is

Figure 1.9: A single hopping transition of a charge carrier occupying some site i with energy ε_i at the position $\vec{r}_i$ to an unoccupied adjacent site j with energy ε_j at $\vec{r}_j$. Because $\varepsilon_j > \varepsilon_i$, the carrier must first absorb a high-energy phonon to be activated to site j. The transition rates ν_{ij} are given in Equation (1.11).

empty. A more general expression of the MA rates, including the electron or hole concentration, is given by [7, 33]

$$\Gamma_{ij} = v_0 f(\varepsilon_i, \varepsilon_F)(1 - f(\varepsilon_j, \varepsilon_F)) \exp\left(-\frac{2d_{ij}}{\alpha}\right) \exp\left(-\frac{\varepsilon_j - \varepsilon_i + |\varepsilon_j - \varepsilon_i|}{2kT}\right). \tag{1.13}$$
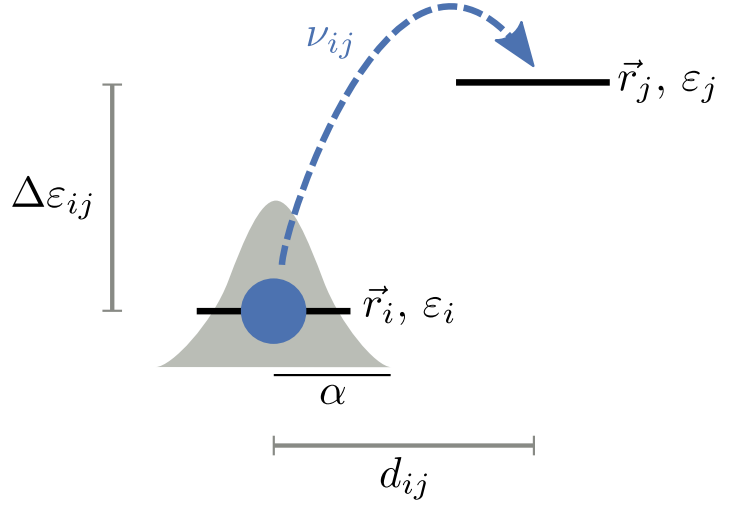
Here, the Fermi distributions include the probabilities for the initial site i to be occupied and the target site j to be empty. If the Fermi level is far below ε_i and ε_j, the Boltzmann approximation from Equation (1.4) can be applied and one arrives at the simple expression:

$$\Gamma_{ij} = v_0 \exp\left(-\frac{2d_{ij}}{\alpha}\right) \exp\left(\frac{\varepsilon_F - \max(\varepsilon_i, \varepsilon_j)}{kT}\right). \tag{1.14}$$

Considering a single charge carrier in an otherwise empty system, however, it is sufficient to work with Equation (1.11). We will do that in most of the Chapters 2 and 3. An important property of Equation (1.11) is that the rates of hops downwards in energy, i.e., where the target site j has lower energy than the initial site i ($\varepsilon_j < \varepsilon_i$), do not depend on site energies at all. The spatial distance between the two sites in that case completely determines the hopping rates and the transitions are by a factor of $\exp(\Delta\varepsilon/kT)$ faster than the reverse direction ($j \to i$), in which additional energy must be provided by a phonon.

1.3.2 Variable Range Hopping (VRH)

An important consequence of the MA rates, as given in Equation (1.11), is the interplay between the spatial contribution $\exp(-2d_{ij}/\alpha)$ and the energetic contribution $\exp(-\Delta\varepsilon/kT)$. A charge carrier occupying some site i in the system will, in the majority of cases, follow the path with the highest transition rate ν_{ij} to escape its original site. This corresponds to minimizing the hopping *difficulty*

$$u_{ij} = -\frac{2d_{ij}}{\alpha} - \frac{\varepsilon_j - \varepsilon_i + |\varepsilon_j - \varepsilon_i|}{2kT} \tag{1.15}$$

with respect to unoccupied adjacent sites. The optimal distance of carriers in hopping transitions therefore depends on the availability of unoccupied sites in the surrounding of the carrier, on the
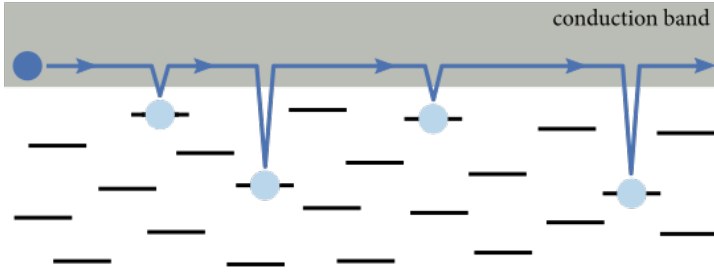
Figure 1.10: Schematic display of the MT transport mode. A charge carrier moves in the conduction band, but is frequently trapped into localized states. From there, detrapping is a thermally activated process.

energies of these sites, and on temperature. If, for example, an unoccupied site is very close to the electron but has a much higher energy, the contribution of the second term in Equation (1.15) may prevent the carrier from hopping to that site and it may be faster to tunnel to a more distant site with a lower energy difference. As obvious from Equation (1.15), the energetic contribution to the hopping difficulty scales with temperature: At high temperatures, it becomes less important and the carrier can overcome larger energy differences during hopping transitions. In that case it is more likely for the carrier to choose sites closer in space as hopping targets. This temperature dependence of the typical hopping distance is called *VRH*.

If kT becomes larger than the energy scale of the system, i.e., the width ε_0 of the exponential DOS or σ of the Gaussian DOS (see Equations (1.6) and (1.7)), hopping transitions become independent on energy and the VRH conduction merges into simple *nearest neighbor hopping (NNH)*, where carriers always tunnel to the sites closest in space.

Note, that VRH is present regardless of the localized states' sizes, as is shown in Section 3.1 of the attached Article [I] (Reference [8]). Even if the system is composed of large molecules, on which carriers are delocalized, the fastest hops are always determined by the interplay between spatial and energetic distance of the sites.

1.3.3 Multiple Trapping (MT)

The second important transport mode, multiple trapping (MT), is relevant for systems, that possess a density of localized states as well as a conduction band with extended states. The two regions are separated by the mobility edge (see Figure 1.2). Carrier mobilities above the mobility edge are of the order of $1\,\mathrm{V\,s\,cm^{-2}}$, which is several orders of magnitude larger than any hopping transport mobility (Section 1.3.1). During MT transport, the conduction process in the bands is interrupted by trapping of carriers into the localized states of the band tail. From there, carriers are thermally activated back into the bands and continue the transport process. Because these activation events are much slower than all other processes involved in long-range transport, they completely dominate the transport properties. The MT transport is schematically shown in Figure 1.10.

MT can only occur if temperature is high enough for carriers to be activated from localized states to the mobility edge ε_c. The typical release time for a carrier trapped in a localized state with energy ε_i is given by

$$\tau = \nu_0^{-1} \exp\left(\frac{\varepsilon_c - \varepsilon_i}{kT}\right), \tag{1.16}$$

where ν_0 is the attempt-to-escape frequency similar to that in the MA hopping rates (Equation (1.11)). At low temperatures, when tunneling transitions described by Equation (1.11) become faster than thermal detrapping (Equation (1.16)), the transport mode changes to VRH.

Chapter 2

The Transport Energy Concept

The concept of the transport energy (TE) is a powerful method for the description of hopping transport in disordered organic and inorganic semiconductors. The TE approach resembles the MT transport mode, described in Section 1.3.3, with the mobility edge replaced by a particular energy level, the so-called *transport energy*. Compared to the percolation theory discussed in Chapter 3, the TE concept is a more transparent approach, since it provides information about the energetic path that a carrier takes during hopping transport through spatially and energetically disordered localized states.

The transport energy was first mentioned in 1985 by Monroe [34]. Monroe managed to describe VRH transport in the exponential DOS using the known MT approach, but with the mobility edge replaced by a slightly temperature dependent energy level, the TE. The concept was later improved by Baranovskii et al. [35] and then extended for other energy spectra, for instance, the Gaussian DOS [36]. Since then, the theory was further developed to account for finite carrier concentrations [37–39] and for the percolation nature of hopping transport [6]. Note, that a variety of different interpretations of the term "transport energy" were proposed in the literature. This ambiguity requires one to carefully check the definition of a particular TE when comparing it to other TE values [7, 40].

We will show in Section 2.1 that charge carriers in a steeply energy dependent DOS do not conduct around the Fermi level, but that transport is dominated by activated hops to an energy level situated much higher in the DOS, namely the transport energy. An analytical description of VRH using the TE concept is given in Sections 2.3 and 2.4, and a numerical method for finding the transport energy in straightforward hopping computer simulations is described in Section 2.5.

2.1 Carrier Activation During Hopping Transport

Consider a charge carrier occupying a site i in a steeply energy dependent DOS $g(\varepsilon)$, e.g., with Gaussian or exponential shape. According to the Miller-Abrahams hopping rates in Equation (1.11), after some time the carrier will perform a tunneling transition to an unoccupied site j in its surrounding. The most probable destination j is the one, that minimizes the hopping difficulty $u(d_{ij}, \varepsilon_j - \varepsilon_i)$:

$$u(d_{ij}, \varepsilon_j - \varepsilon_i) = \frac{2d_{ij}}{\alpha} + \frac{\varepsilon_j - \varepsilon_i + |\varepsilon_j - \varepsilon_i|}{2kT}. \tag{2.1}$$

Equation (2.1) is the negative exponent of the MA hopping rates as defined in Equation (1.11), where d_{ij} is the distance between the sites, ε_i and ε_j are their energies, α is the localization radius, and kT is thermal energy. The minimum of that difficulty corresponds to a maximal hopping rate and hence to the most probable hopping transition.

The second term in the r.h.s. of Equation (2.1), the energy contribution, vanishes when $\varepsilon_i > \varepsilon_j$, i.e, when the carrier relaxes towards lower energies. This makes hops downwards in energy much
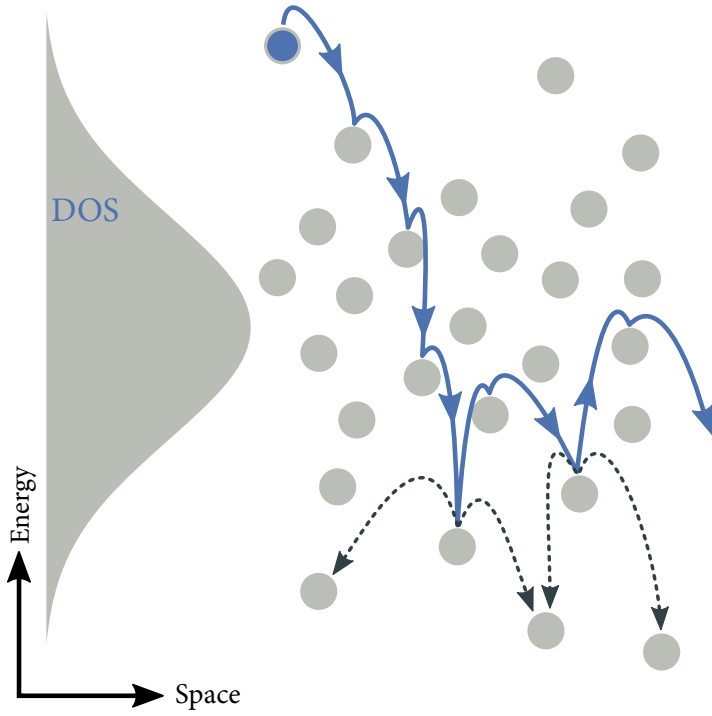
Figure 2.1: A carrier (blue circle) in a steeply energy dependent DOS can not dive indefinitely in energy. At some point, the next available target site for a hopping transition is so far away (dashed arrows), that it is faster for the carrier to absorb a phonon and be activated to a site higher in energy but closer in space.

faster than phonon-assisted upward hops, given that a suitable target site j is present close to the carrier's initial site i. If, however, the distance d_{ij} to the next site with $\varepsilon_i > \varepsilon_j$ is large, it may be more efficient for the carrier to pay the price of the energy difference in order to minimize the necessary tunneling distance.

There are two factors that increase the necessary hopping distance d_{ij} at deeper energies: First, in thermal equilibrium the occupation of sites by other charge carriers increases towards lower energies, as described by the Fermi-Dirac distribution in Equation (1.3). Hence, the deeper the carrier dives in energy, the lower is the probability of finding unoccupied sites suitable as hopping targets. Secondly, the DOS decreases steeply with energy, further lowering the amount of target states available for hopping transitions. At some point during the carrier's journey, the distance d_{ij} from its current site i to the next site j with lower energy will therefore become so large, that it is easier to be activated to another spatially close target site k with higher energy. This happens, when there is at least one site k with $\varepsilon_k > \varepsilon_i$ that fulfills the following criterion for all sites j with $\varepsilon_j < \varepsilon_i$:

$$\frac{2d_{ij}}{\alpha} > \frac{2d_{ik}}{\alpha} + \frac{\varepsilon_k - \varepsilon_i}{kT} \, . \tag{2.2}$$

The relaxation and subsequent activation process is depicted in Figure 2.1. In each DOS that decreases towards deeper energies, there exists some energy level below which the average distance to the next site is so large that carriers will instead be activated to higher energies. This threshold energy is exactly the TE level. Surprisingly, as we will show in Section 2.3, this critical energy level does not depend on the energy of the carrier's initial site ε_i.

2.2 Charge Transport and the TE

In the previous section, we found that hopping transport in a steeply energy dependent DOS consists of hops downwards in energy, if carriers occupy states above the TE ε_t, and activated hops upwards in energy, if the initial carrier energy is lower than ε_t. One might think that carriers therefore stay in the vicinity of this transport energy level during the whole conduction process: When relaxing downwards from some initial energy level ε_i, the most probable hopping target energy is dictated solely by the DOS. In DOS functions that decrease steeply with energy, it is likely that carriers jump to energies not far away from their starting energy ε_i. This relaxation happens until they reach energies below the TE ε_t. Then, the carriers perform hops upwards in energy again, with the hopping destination also situated in the vicinity of the TE. VRH therefore seems to be resembled by some kind of zigzag motion in the vicinity of the transport energy level. This assumption is, however, wrong for the following reason.

The probability $p_t(\varepsilon)$ for a carrier placed at some energy ε_i to be trapped into a site with energy $\varepsilon < \varepsilon_i$ is proportional to the DOS,

$$p_t(\varepsilon) \propto g(\varepsilon),\tag{2.3}$$

and the time $\tau_t(\varepsilon)$, that the carrier must wait until it is released and activated back into states around ε_t becomes exponentially longer for deeper energies,

$$\tau_t(\varepsilon) \propto \exp\left(\frac{\varepsilon_t - \varepsilon}{kT}\right).\tag{2.4}$$

Although trapping into deep states is less probable in steeply decaying DOS functions than trapping into shallow states, the carriers spend an exponentially long time trapped in those deep states. The equilibrium between the trapping probability in Equation (2.3) and the trapping time in Equation (2.4) therefore determines the equilibrium carrier distribution during VRH transport. It results in the carrier distributions introduced in Section 1.2: The majority of carriers in the exponential DOS occupy states around the Fermi level ε_F and most carriers in the Gaussian DOS are found in states close to the equilibration energy, ε_∞.

During charge transport, the slowest hopping transitions are therefore the rare hops from energies deep in the DOS, i.e., from the equilibrium carrier distribution, to the TE level. Because these hops are much slower than all other transitions, they almost completely determine charge transport coefficients, for instance the mobility. In addition, they account for the strongest temperature dependence of the transport properties, as is obvious from Equation (2.4). Based on these simple arguments, we can already here predict a dependency of the carrier mobility on the TE, the equilibrium carrier distribution, and temperature in the form

$$\mu \propto \tau^{-1} \propto \exp\left(-\frac{\varepsilon_t - \varepsilon_\infty}{kT}\right) \qquad \text{or} \qquad \mu \propto \tau^{-1} \propto \exp\left(-\frac{\varepsilon_t - \varepsilon_F}{kT}\right)\tag{2.5}$$

for the Gaussian and exponential DOS, respectively. (Note, that this is only true for the Gaussian DOS as long as $\varepsilon_F < \varepsilon_\infty$.) Knowledge of the position and properties of the transport energy is therefore crucial for the understanding of VRH and for the calculation of transport coefficients in steeply energy dependent DOSs.

2.3 Optimization Approach to the TE

Mathematically, the transport energy can be calculated by optimizing the MA hopping rates for a transition upwards in energy with respect to the target energy ε_j. In what follows, we will follow the derivation of the TE by Oelerich [6], taking into account both the percolation nature of hopping transport as well as the dependence on carrier concentration.

As we know from Equation (1.11), the rate $\nu_\uparrow(\varepsilon_i, \varepsilon_j)$ for a transition from some site i to another site j with higher energy can be written as

$$\nu_\uparrow(\varepsilon_i, \varepsilon_j) = \nu_0 \exp\left(-\frac{2d(\varepsilon_i, \varepsilon_j)}{\alpha} - \frac{\varepsilon_j - \varepsilon_i}{kT}\right), \tag{2.6}$$

with $d(\varepsilon_i, \varepsilon_j)$ being the typical hopping distance for transitions between energies ε_i and ε_j. The optimal target energy ε_j for a hop starting at ε_i is obtained from Equation (2.6) by optimizing the rates with respect to ε_j:

$$\frac{d\nu_\uparrow(\varepsilon_i, \varepsilon_j)}{d\varepsilon_j} = 0. \tag{2.7}$$

An expression for the energy dependent hopping distance $d(\varepsilon_i, \varepsilon_j)$ can be obtained from the site concentration N and the percolation threshold $B_c = 2.7$ (an introduction in percolation theory is given in Chapter 3):

$$d(\varepsilon_i, \varepsilon_j) \approx d(\varepsilon_j) = \left[\frac{4\pi}{3B_c}N(\varepsilon_j)\right]^{-1/3} = \left[\frac{4\pi}{3B_c}\int_{-\infty}^{\varepsilon_j} g(\varepsilon)\left[1 - f(\varepsilon, \varepsilon_F)\right]d\varepsilon\right]^{-1/3}. \tag{2.8}$$

Note, that we approximate the site concentration $N(\varepsilon_j)$ *at* the energy ε_j with the concentration of sites *up to* the energy level ε_j. This approximation is valid for steeply decreasing DOS shapes. Also note, that we account for the percolation nature of hopping transport by including the percolation threshold B_c, and for finite carrier concentrations by counting only unoccupied sites, i.e., adding the factor $[1 - f(\varepsilon, \varepsilon_F)]$ to the integrand.

Collecting Equations (2.6) to (2.8) and identifying the optimal target energy ε_j as the TE ε_t, we can write down the following defining equation for the TE:

$$\frac{2kT}{3\alpha}\left(\frac{4\pi}{3B_c}\right)^{-1/3}\left[\int_{-\infty}^{\varepsilon_t} g(\varepsilon)[1 - f(\varepsilon, \varepsilon_F)]d\varepsilon\right]^{-4/3}\left[1 - f(\varepsilon_t, \varepsilon_F)\right]g(\varepsilon_t) = 1 \tag{2.9}$$

Equation (2.9) can be solved numerically to obtain ε_t. In Figure 2.2, some of the results are shown. It is clear from the data, that the TE is situated close to the center of the DOS (or, in case of an exponential spectrum, the mobility edge). With rising temperature, the TE is pushed to higher energies and it is independent on carrier concentration for most reasonable fillings. The results show, that it is often a good approximation to assume $\varepsilon_t = 0$, especially in the Gaussian DOS [6].
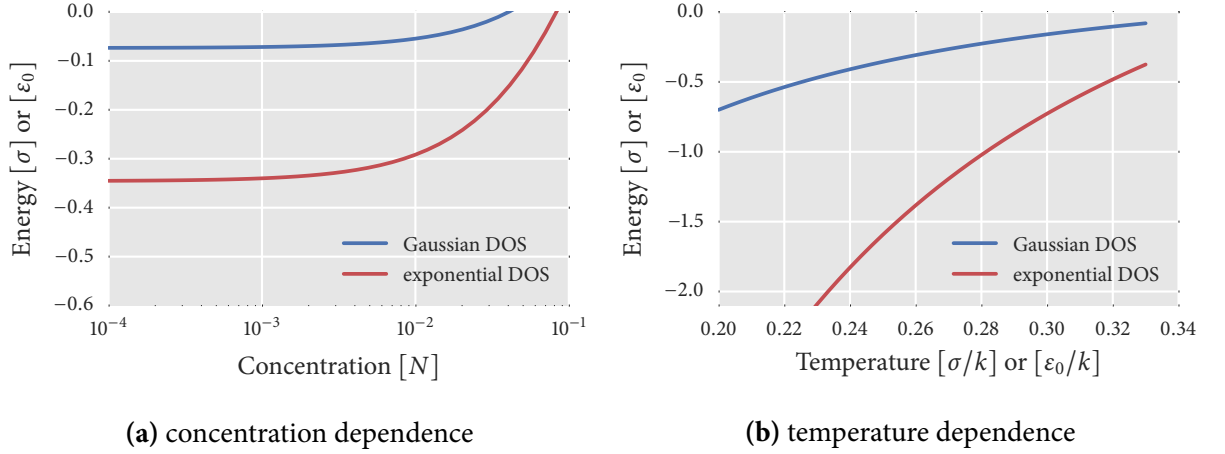
(a) concentration dependence **(b)** temperature dependence

Figure 2.2: The position of the TE ε_t for the exponential and Gaussian DOS versus charge carrier concentration and temperature. Energies and temperatures are given in units of the width of the DOS σ (or, in case of the exponential DOS, ε_0). In (a), temperature was fixed at $kT = 0.33\sigma$ and in (b) concentration was fixed at $n = 10^{-4}N$.

2.4 From the TE to Charge Carrier Mobility

In the previous sections we learned, that long-range charge transport is dominated by hops from deep states to the TE and how to calculate that energy level. Although this alone is valuable information for the understanding of VRH transport, we can use the knowledge of the TE to derive analytical expressions for the charge carrier mobility and related quantities of the conduction process.

In disordered semiconductors, when charge carriers are in thermal equilibrium and the carrier concentration is not too high, the *Einstein relation* holds, relating the diffusion coefficient D to the mobility μ of charge carriers [41, 42]:

$$\mu = \frac{e}{kT}D\,. \tag{2.10}$$

The diffusion coefficient D in one dimension (in this case the direction of charge transport) is defined as

$$D = \frac{d(\varepsilon_t)^2}{2\langle t\rangle}\,, \tag{2.11}$$

with $d(\varepsilon_t)$ being the typical distance of single hopping transitions and $\langle t\rangle$ being the average time of the hops. For the former, we use the hopping distance calculated in Equation (2.8) with the TE ε_t evaluated from Equation (2.9). The time $\langle t\rangle$ is calculated as an average over the corresponding hopping rates $\nu_\uparrow(\varepsilon_i, \varepsilon_t)$ from some initial site i to the TE ε_t, weighted with the density $g(\varepsilon)$ and availability $[1 - f(\varepsilon, \varepsilon_F)]$ of the sites:

$$\langle t\rangle = \langle \nu_\uparrow^{-1}(\varepsilon_t)\rangle = \frac{\int\limits_{-\infty}^{\varepsilon_t} \nu_\uparrow^{-1}(\varepsilon, \varepsilon_t)g(\varepsilon)[1 - f(\varepsilon, \varepsilon_F)]\,\mathrm{d}\varepsilon}{\int\limits_{-\infty}^{\varepsilon_t} g(\varepsilon)[1 - f(\varepsilon, \varepsilon_F)]\,\mathrm{d}\varepsilon}\,. \tag{2.12}$$

Collecting Equations (2.8) to (2.12), the following expression can be found for the mobility μ:

$$\mu = \frac{v_0}{2}\frac{e}{kT}\frac{3B_c}{4\pi d(\varepsilon_t)n_t}\exp\left(-\frac{2}{\alpha}d(\varepsilon_t) - \frac{\varepsilon_t - \varepsilon_F}{kT}\right), \tag{2.13}$$

with the concentration n_t of carriers occupying states below the TE:

$$n_t = \int_{-\infty}^{\varepsilon_t} f(\varepsilon, \varepsilon_F)g(\varepsilon)\,\mathrm{d}\varepsilon. \tag{2.14}$$

It is tempting to interpret Equation (2.13) as activation of charge carriers from the Fermi level ε_F to the transport energy, because the expression has a form similar to that of the MA hopping rates in Equation (1.11). However, the prefactor n_t carries an additional strong dependence on the Fermi level, and the TE (and, consequently, the hopping distance $d(\varepsilon_t)$) weakly depend on carrier concentration as well.

In an empty system, where the Boltzmann approximation in Equation (1.4) can be applied, we can write n_t as

$$n_t \approx \exp\left(\frac{\varepsilon_F}{kT}\right)\int_{-\infty}^{\varepsilon_t}\exp\left(-\frac{\varepsilon}{kT}\right)g(\varepsilon)\,\mathrm{d}\varepsilon. \tag{2.15}$$

Inserting this expression into Equation (2.13) it is ovious that the strong, exponential dependence on the value of the Fermi level ε_F cancels. In addition, we can assume that all relevant sites in the DOS are unoccupied, so that the terms $[1 - f(\varepsilon, \varepsilon_F)]$ drop out of the expressions for ε_t and $d(\varepsilon_t)$ in Equations (2.8) and (2.9), respectively. The mobility in this case becomes completely independent of the Fermi level and thereby of the charge carrier concentration n of the system.

This case applies only as long as the Boltzmann approximation is applicable. This is not the case for high concentrations of charge carriers and for certain shapes of the DOS. In the exponential DOS, for instance, mobility depends on n at all concentrations n, which is therefore to be expected in all amorphous inorganic semiconductors (see Section 1.2.1). The Gaussian DOS, present in ODSs, on the other hand, allows the application of the Boltzmann approximation and therefore exhibits concentration independent mobilities up to high carrier concentrations n. The transition from an independence to a dependence of μ on n can be used to determine the shape of the DOS experimentally, as shown by Oelerich et al. [12].

Example solutions of Equation (2.13) are shown in Figure 2.3 for the exponential and Gaussian DOS. The behaviour described above, with the mobility μ being independent on n at low n for the Gaussian DOS but not for the exponential one, is clearly visible from the data. When the carrier concentration n becomes high enough to fulfill the criterion Equation (1.10), the Fermi level surpasses ε_∞ and the carrier distribution is pushed to higher energies. Consequently, the energy distance between the carriers and the TE becomes smaller and mobility is increased. The results shown in Figure 2.3 confirm the schematic picture introduced in Figure 1.8.

2.5 Numerical Approach to the TE

Complementary to the analytical methods for calculating the transport energy, there are also attempts to determine ε_t in hopping computer simulations [40, 43–45]. Here, we will introduce the
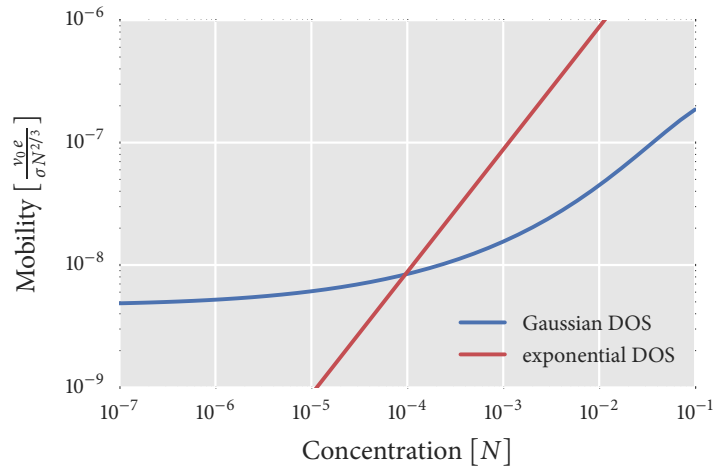
Figure 2.3: Concentration dependence of the carrier mobility $\mu(n)$ as calculated via Equation (2.13) for exponential and Gaussian DOS at $kT = 0.25\sigma$ for the Gaussian and $kT = 0.5\varepsilon_0$ for the exponential spectrum.

approach by Oelerich et al. [40], that is based on observing the influences of DOS modifications on simulation results. Another way to determine the TE computationally is by the heat method shown in Chapter 4.

Computer simulations of hopping transport in disordered media are typically rather straightforward: One assumes some particular DOS and generates a set of localized states with energies distributed accordingly. Depending on the simulation details, the states are placed randomly in space or on some lattice structure. Charge transport is then simulated either directly using Monte Carlo methods or via a balance equation approach, based on steady state characteristics of the current flow through the sample. Descriptions of these schemes are given in Chapter 7. In either of those algorithms, the carrier mobility of long-range charge transport can be extracted from the simulation results.

The value of the mobility determined in the simulations depends on the choice of the DOS and will therefore also be sensitive to variations of the DOS, even if its general shape (e.g., a Gaussian one) is fixed. This can be used in the following way to determine the energy range that is decisive for charge transport and therefore, by definition, includes the TE.

Hopping transport is simulated not in the complete DOS, but in a DOS in which sites of a certain energy range $[\varepsilon_{\text{cut}} - w, \varepsilon_{cut}]$ are cut out abd unavailable for hopping transport. An example of such a modification is schematically shown in Figure 2.4. The dependence of the mobility μ on the width w and the upper boundary ε_{cut} of the cutout interval is then determined via repeated hopping computer simulation of the system while sweeping the parameters ε_{cut} and w. The resulting $\mu(\varepsilon_{\text{cut}}, w)$ is expected to possess a minimum value for that particular cutout interval $[\varepsilon_{\text{cut}} - w, \varepsilon_{cut}]$, that is most important for charge transport. By definition of the transport energy, the TE is then expected to lie within the interval that was responsible for the most severe drop of the mobility values. The precision of the so-determined value of the TE depends on the width w of the interval: The fewer sites are cut out of the DOS, the more precise will be the found value of the TE. However, taking out fewer sites also results in a less pronounced dependence of the $\mu(\varepsilon_{\text{cut}}, w)$ curve on the modification.

Results of the cutout method for the Gaussian DOS are plotted in Figure 2.5, along with a comparison with the analytical optimization approach described in Section 2.3. The values obtained from the optimization approach nicely match the minima of the mobility curves. Hence, the
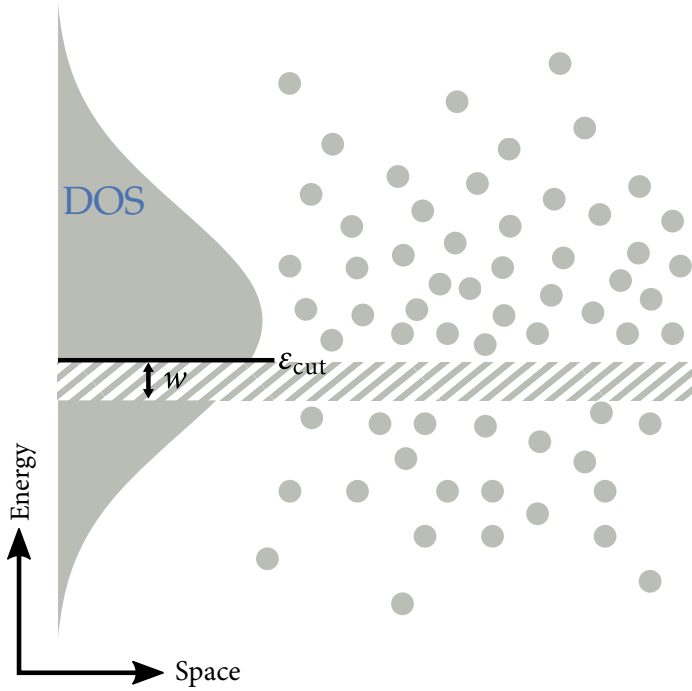
Figure 2.4: Schematic picture of a Gaussian DOS with sites (grey circles) in the interval $\left[\varepsilon_{\mathrm{cut}} - w, \varepsilon_{cut}\right]$ cut out from the system. The hatched region is therefore forbidden for charge transport.

method of optimizing the hopping rates evidently yields values for the TE that indeed are crucial for long-range charge transport. As expected, the absolute drop in the mobility depends on the width w of the cutout interval. The fewer sites are cut out, the less severely suffers the mobility from the modification.
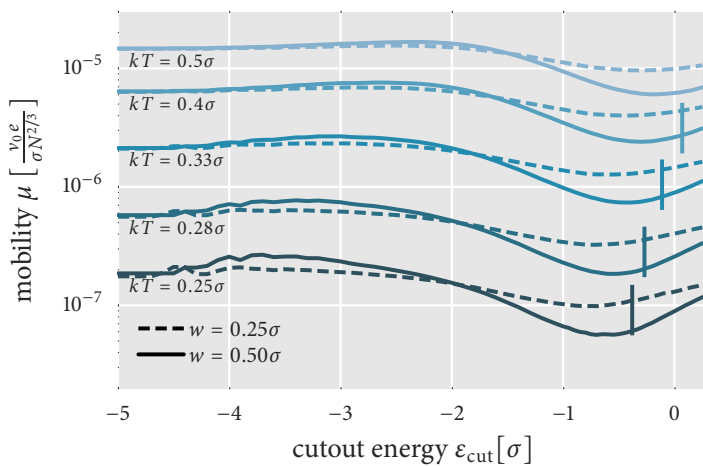
Figure 2.5: Mobility obtained from straightforward computer simulations of VRH with the Gaussian DOS modified according to the method described in Section 2.5. The drop in the mobility for certain cutout energies $\varepsilon_{\mathrm{cut}}$ is clearly visible. The straight lines correspond to the TE as obtained from the optimization approach in Section 2.3. Parameters: $\alpha = 0.215 N^{-1/3}$, $n \approx 10^{-6} N$.

Chapter 3

Percolation Theory

Percolation theory is the most popular theoretical approach to describe hopping charge transport in disordered systems. It can be applied to VRH and the simpler NNH [7]. Percolation theory is the theory of connectivity in random networks and a simple way to describe phase transitions. The basic concept can easily be explained by the following example.

Imagine a lattice of initially disconnected sites. Now, subsequently add connections between nearest neighboring sites until a connected path from one side of the system to the opposite side is established, as shown in Figure 3.1 (a). The ratio $x_c = N_{conn}/N_{tot}$ of connected to disconnected bonds is called the *percolation threshold*.

One distinguishes between two types of percolation: Bond and site percolation. The example described above and displayed in Figure 3.1 (a), is called *bond percolation* because the decisive entities for percolation are the connections between the sites. If the sites themselves are "enabled" during the search for the percolation threshold, the problem is called *site percolation*. Site percolation is schematically shown in Figure 3.1 (b).

Percolation theory can be applied to many kinds of network problems. In most cases, the interesting property one tries to find is the percolation threshold x_c. In Section 3.1 the percolation theory is applied to hopping in a random network of isoenergetic sites. In Section 3.2 we then develop an exact percolation theory approach for hopping conduction in disordered systems with exponential DOS. This exact solution is described in more detail in Article [II] — here, we will only point out the most important steps of the derivation.

3.1 Percolation Theory in a Resistor Network

Let us study the simple model of a spatially random network of localized states with equal energies. This model is applicable to disordered materials when temperature kT is much larger than the energy scale ε_0 or σ of the DOS (see Equations (1.6) and (1.7)). In that case, the contribution of the energetic term to the hopping rates in Equation (1.11) is negligible and the disorder can be treated as purely spatial. Consequently, as indicated in Section 1.3.2, the transport mode will be nearest neighbor hopping (NNH). We assume that the concentration of sites N is low enough, so that the average intersite distance $N^{-1/3}$ is much larger than the localization radius, $N^{-1/3} \gg \alpha$.

Following Miller and Abrahams [31] and Shklovskii and Efros [33], we will interpret the hopping problem as a network of random resistances: Each site i in our system represents a node in an electrical circuit and is connected via resistors (i, j) to adjacent sites. Each of these connections is assigned a resistance

$$R_{ij} = R_0 \exp\left(\frac{2d_{ij}}{\alpha}\right) \tag{3.1}$$

where d_{ij} is the distance between the two nodes and α is the localization radius. The exponential function corresponds to the spatial part of the hopping rates in Equation (1.11), which is appropriate
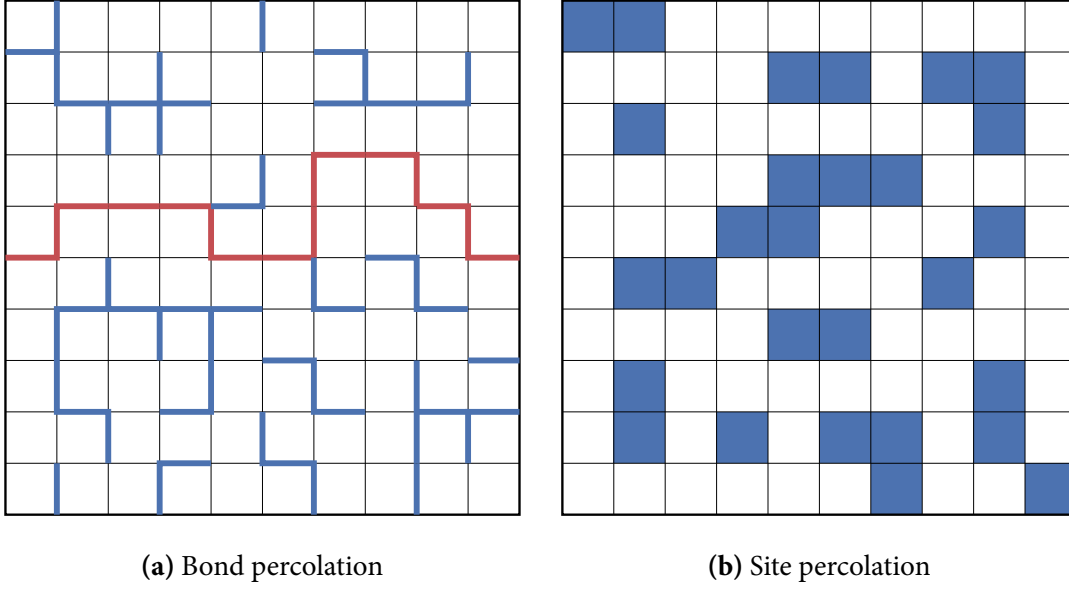
(a) Bond percolation   **(b)** Site percolation

Figure 3.1: The two percolation schemes, (a) bond percolation and (b) site percolation. "Enabled" entities are shown in color. In (a), a percolating path connecting two sides of the system is found and highlighted in red color.

for the isoenergetic model. The prefactor R_0 is equal for all bonds in the system.

Current in the network will always prefer the path of minimal resistance. We can therefore find the critical resistance $R_{ij} \equiv R_{\text{crit}}$ that electrons still need to pass through in order to provide long-range conduction. Because resistances increase monotonously with distance d_{ij} according to Equation (3.1), the problem is identical to finding the longest hopping distance d_{crit} required for the carriers to conduct. The critical resistance is then given by Equation (3.1) as

$$R_{\text{crit}} = R_0 \exp\left(\frac{2d_{\text{crit}}}{\alpha}\right). \tag{3.2}$$

The percolation problem is visualized in 2D in Figure 3.2. At fixed concentration N of sites one enables all connections with $d_{ij} < \tilde{d}$, where $\tilde{d}$ is a small number. Now, $\tilde{d}$ is increased and the additional connections fulfilling $d_{ij} < \tilde{d}$ are enabled. At some size of $\tilde{d} \equiv d_{\text{crit}}$, two opposite sides of the system will become connected and the critical distance is found. The same method can be applied for finding the critical concentration N_{crit} at fixed distance $\tilde{d}$. In fact, it is convenient to represent the percolation solution with a dimensionless quantity B_c, that is defined by

$$B_c = N V_{\text{crit}} = N \frac{4\pi}{3} d_{\text{crit}}^3. \tag{3.3}$$

B_c is the average number of sites in a spherical volume with radius d_{crit} at a site concentration N. At the percolation threshold, its numerical value for 3D systems is $B_c = 2.735$ [7, 33]. Because $d_{\text{crit}} \propto N^{-1/3}$, B_c does not depend on the concentration N and is a universal constant.

Knowing the value of B_c and the relation given by Equation (3.3), we can calculate the network's dc conductivity $\sigma_{\text{dc}} \propto R_{\text{crit}}^{-1}$, as [7]

$$\sigma_{\text{dc}}(N) \propto \exp\left(-\frac{\gamma N^{-1/3}}{\alpha}\right), \qquad \gamma = 2\left(\frac{3B_c}{4\pi}\right)^{1/3} \approx 1.7351. \tag{3.4}$$

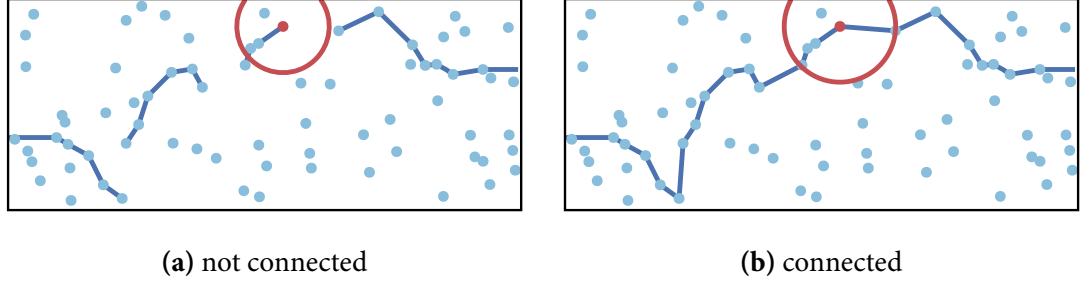**(a)** not connected   **(b)** connected

Figure 3.2: The percolation problem of finding the critical radius (or critical concentration) of sites in order to provide an infinite connected path. (a) shows the disconnected network, and the percolating network is shown in (b).

The actual percolation problem that was solved to arrive at Equation (3.4) is finding the value of B_c, which can only be done numerically. However, the above derivation is a first example on how to apply percolation theory to physical problems.

Equation (3.4) provides the strong exponential dependence of the conductivity on the concentration of sites. However, the prefactor of σ_{dc} also carries an additonal weak dependence on N that is yet to be determined. The reason for the dependence of the prefactor on site concentration is the following. If current would flow only through resistances with $R_{ij} \leq R_{crit}$, those resistors with $R_{ij} \approx R_{crit}$ will represent bottle necks in the connected cluster, since they are very sparse and by definition all of the current has to pass through them. If one adds some more resistors with $R_{crit} < R_{ij} \leq R_{crit} \exp(\delta)$ to the network, the conductivity will increase significantly at $\delta \approx 1$ [7, 8, 33]. Larger values of δ will have negligible effects on the conductivity [33]. The infinite network with resistances $R_{ij} \leq R_{crit} \exp(\delta)$ is called the *critical subnetwork*, because it completely determines the conductivity. For this critical subnetwork we can define a *correlation length L_{corr}* in the following way. If one cuts out of the critical subnetwork a cube of length L and measures its resistance $R(L)$, one will find two limiting cases. In case L is smaller than the correlation length L_{corr} of the critical subnetwork, the resistance $R(L)$ of the cube is determined by its *local* critical resistance $\tilde{R}_{crit}$. In the opposite case it reflects the macroscopic resistivity ρ of the whole system:

$$R(L) = \begin{cases} \tilde{R}_{crit} & \text{if } L < L_{corr}\,, \\ \rho/L & \text{if } L > L_{corr}\,. \end{cases} \tag{3.5}$$

At $L = L_{corr}$, the two cases meet and one observes $\rho \approx L_{corr}\tilde{R}_{crit}$. Moreover, at $L = L_{corr}$, the local critical resistance of the cube $\tilde{R}_{crit}$ is close to the critical resistance of the whole network from Equation (3.2), $\tilde{R}_{crit} \approx R_{crit}$. We can therefore conclude, that

$$\sigma_{dc} = \rho^{-1} = L_{corr}^{-1} R_{crit}^{-1}\,, \tag{3.6}$$

and the remaining task is to find the correlation length L_{corr}.

It is known [33], that the correlation length L_{corr} of the infinite cluster scales as

$$L_{corr} = N^{-1/3} \left(\frac{\xi - \xi_c}{\xi_c} \right)^{-\nu}\,, \tag{3.7}$$

with ξ_c being the exponent of the critical resistance Equation (3.2), and ξ being the exponent of the critical resistance of the critical subnetwork, i.e., $R_{\text{crit}} \exp(\delta)$. The power v depends on dimensionality. In 3D, $v = 0.875 \pm 0.008$ [7]. Inserting the resistances' exponents into Equation (3.7) yields the dependence of the correlation length on site concentration N:

$$L_{\text{corr}} \propto N^{-1/3} \left(\alpha N^{1/3} \right)^{-v} . \tag{3.8}$$

Combining Equations (3.2), (3.4), and (3.8) then gives the full expression for the conductivity of our problem of randomly distributed sites:

$$\sigma_{\text{dc}}(N) = A R_0^{-1} N^{1/3} \left(\alpha N^{1/3} \right)^v \exp \left(-\frac{\gamma N^{-1/3}}{\alpha} \right) . \tag{3.9}$$

Equation (3.9) contains the weak pre-exponential dependence of the conductivity on concentration N as well as the more pronounced exponential dependence. It is exact up to a dimensionless numerical coefficient A.

3.2 Percolation Theory for VRH in the Exponential DOS

In the past four decades, various attempts were made to describe VRH using methods from percolation theory [8, 46–50], with Grünewald and Thomas [50] being the first to apply it to transport in the exponential DOS. The way how percolation theory is typically incorporated into studies of VRH transport is similar to the purely spatial percolation problem discussed in the previous Section 3.1, but with an additional energy dependent term in the resistances R_{ij}. The value of the percolation threshold $B_c = 2.735$ from Equation (3.3), although obtained for problems without energy disorder, is usually applied also to the energy dependent problems. Until recently, it was not known whether this is a good approximation and how to appropriately average over connected sites in the network [7]. Nenashev et al. [8] finally showed in 2013 how to exactly solve the VRH problem in the exponential DOS by mapping the 3+1D space-energy VRH onto a 3D percolation problem of spheres with exponentially distributed radii, which can be solved exactly. The study is appended in Article [II]. In this section, the approach is briefly outlined.

The general procedure will be similar to the purely spatial problem discussed in the previous Section 3.1: Values for the critical resistance R_{crit} of the resistor network are found first, which includes the strong dependence of the volume resistivity ρ on the model parameters. For a derivation of the weakly varying pre-exponential factors is we refer to Article [II] in the appendix.

Let us start again by formulating the resistance of two sites i and j in the network of resistors, similar to Equation (3.1) but with an additional energy dependence to account for the energy distribution in our problem. We use the form of the MA rates given in Equation (1.14):

$$R_{ij} = \frac{kT}{e^2 v_0} \exp \left[\frac{2d_{ij}}{\alpha} + \frac{\max\left(\varepsilon_i, \varepsilon_j \right) - \varepsilon_F}{kT} \right] . \tag{3.10}$$

Equation (3.10) holds as long as the Fermi level ε_F is far below all energies decisive for the conduction process, which we will assume from now on. Similar to the procedure in the previous section, we search for the critical resistance $R_{ij} \leq R_{\text{crit}}$ that is needed for the infinite percolating cluster.

Because R_{crit} is expected to have a similar form as Equation (3.10), it is convenient to write it in the form

$$R_{\mathrm{crit}} = \frac{kT}{e^2 v_0} \exp\left[\frac{\varepsilon^* - \varepsilon_F}{kT}\right],\tag{3.11}$$

introducing a *critical energy* ε^* which is to be determined. By comparing Equations (3.10) and (3.11), the percolation criterion $R_{ij} \leq R_{\mathrm{crit}}$ can then be translated into the following constraints:

$$\frac{2d_{ij}}{\alpha} + \frac{\varepsilon_i}{kT} \leq \frac{\varepsilon^*}{kT} \qquad \text{and} \qquad \frac{2d_{ij}}{\alpha} + \frac{\varepsilon_j}{kT} \leq \frac{\varepsilon^*}{kT}.\tag{3.12}$$

Since $d_{ij} > 0$, the criteria in Equation (3.12) can only be fulfilled by sites with $\varepsilon < \varepsilon^*$. The concentration N^* of such sites is given by

$$N^* = \int_{-\infty}^{\varepsilon^*} g(\varepsilon)\, \mathrm{d}\varepsilon = N \exp\left(\frac{\varepsilon^*}{\varepsilon_0}\right),\tag{3.13}$$

using the exponential DOS from Equation (1.6) (with $\varepsilon_c = 0$). Because N^* cannot be larger than the total site concentration N, Equation (3.13) implies that $\varepsilon^* < 0$. For each of the sites i that fulfill $\varepsilon_i < \varepsilon^*$, let us define a "radius" r_i according to

$$r_i = \frac{\alpha}{2kT}\left(\varepsilon^* - \varepsilon_i\right),\tag{3.14}$$

which can be used to once again reformulate the bonding criteria from Equation (3.12) to read

$$d_{ij} \leq r_i \qquad \text{and} \qquad d_{ij} \leq r_j.\tag{3.15}$$

Because the radii r are linear functions of the site energies ε, their distribution $g(r)$ is given by the DOS $g(\varepsilon)$ via

$$g(r) = \frac{g(\varepsilon)}{|\,\mathrm{d}r/\,\mathrm{d}\varepsilon|} = \frac{N}{\varepsilon_0}\frac{2kT}{\alpha} \exp\left(\frac{\varepsilon}{\varepsilon_0}\right).\tag{3.16}$$

Using Equations (3.13) and (3.14), we can rewrite $g(r)$ as

$$g(r) = \frac{L^*}{N^*} \exp\left(-\frac{r}{L^*}\right), \qquad \text{with} \qquad L^* \equiv \frac{\alpha \varepsilon_0}{2kT}.\tag{3.17}$$

L^* is the natural length scale of the system of radii, allowing us to express the distances d_{ij} and radii r_i as dimensionless quantities:

$$\tilde{d}_{ij} = d_{ij}/L^* \qquad \text{and} \qquad \tilde{r}_i = r_i/L^*,\tag{3.18}$$

along with the dimensionless bonding criteria

$$\tilde{d}_{ij} \leq \tilde{r}_i \qquad \text{and} \qquad \tilde{d}_{ij} \leq \tilde{r}_j.\tag{3.19}$$

The distribution of dimensionless radii $g(\tilde{r})$ is given by

$$g(\tilde{r}) = n \exp\left(-\tilde{r}\right), \qquad \text{with} \qquad n \equiv N^*(L^*)^3.\tag{3.20}$$
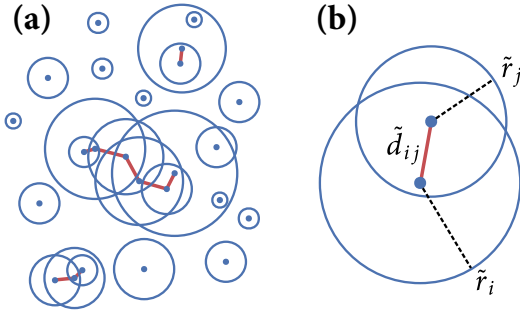
Figure 3.3: The percolation problem of the system of randomly distributed spheres with exponentially distributed radii. (a) shows an example distribution of spheres and some connections and (b) illustrates the connection criterion. Figure taken from Article [II].
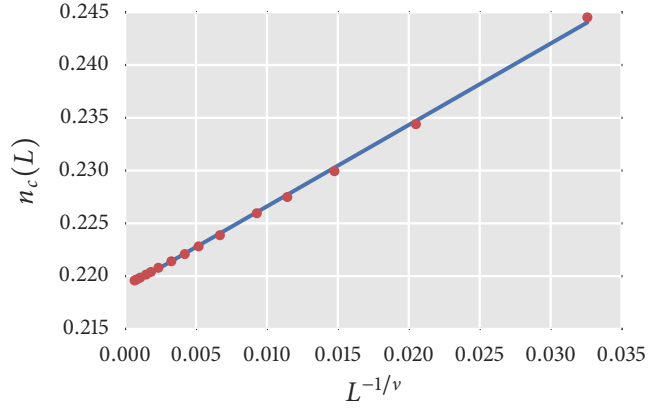
Figure 3.4: Finite size scaling of the percolation threshold n_c for the spheres problem in Figure 3.3 with the scaling exponent $v = 0.875$. Figure taken from Article [II].

Equations (3.19) and (3.20) correspond exactly to the following purely spatial percolation problem. Consider a system of randomly placed spheres with concentration n and radii $\tilde{r}$. The radii are distributed exponentially according to the distribution $g(\tilde{r})$ in Equation (3.20). Two spheres are considered as *connected*, if the criterion formulated in Equation (3.19) holds, i.e., the center of sphere i lies within sphere j and vice versa. See Figure 3.3 for an illustration of the problem. One can then find the necessary concentration $n = n_{\mathrm{crit}}$, that is needed for an infinite cluster of connected spheres in the system. This percolation problem carries no dependencies on system parameters whatsoever and can be solved numerically up to high precisions.

The critical concentration n_c of spheres is obtained by simulating the percolation problem for systems with different finite lengths L and then finding $n_c = n_c(L = \infty)$ with the finite size scaling law

$$n_c(L) - n_c(\infty) \propto L^{-1/v} . \tag{3.21}$$

The scaling exponent v is known precisely [51] as $v = 0.875$ in the 3D case. Such an evaluation is shown in Figure 3.4. The critical concentration n_c is then evaluated to

$$n = n_c \approx 0.219 . \tag{3.22}$$

Knowing the threshold concentration n_c necessary for an infinite connected network, we can now map the problem back onto VRH in space and energy. By combining Equations (3.13), (3.17), (3.20), and (3.22) one can formulate the following equation for the critical energy ε^*:

$$n_c = N \exp\left(\frac{\varepsilon^*}{\varepsilon_0}\right)\left(\frac{\alpha \varepsilon_0}{2kT}\right)^3 . \tag{3.23}$$

29

Figure 3.5: The hopping mobility μ for the exponential DOS as obtained from the TE concept and the exact solution found in Section 3.2 versus temperature. The absolute value of the exact solution is an unknown parameter.
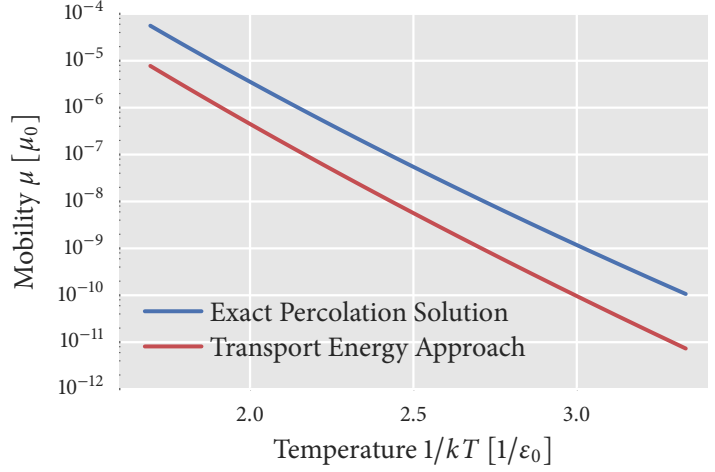
Assuming low temperatures, the Fermi level ε_F can be obtained from the carrier concentration n_e via [8]

$$n_e \equiv \int_{-\infty}^{\varepsilon_F} g(\varepsilon)\, \mathrm{d}\varepsilon = N \exp\left(\frac{\varepsilon_F}{\varepsilon_0}\right). \tag{3.24}$$

Equations (3.23) and (3.24) finally allow us to determine the critical resistance R_{crit} from Equation (3.11) and we obtain

$$R_{\mathrm{crit}} = \frac{kT}{e^2 \nu_0} \left[\frac{n_c}{n_e} \left(\frac{2kT}{\alpha \varepsilon_0}\right)^3 \right]^{\varepsilon_0/kT}. \tag{3.25}$$

The result for R_{crit} is exact up to small errors in the numerical evaluation of the critical concentration n_c, given that temperature is low enough that Equation (3.24) holds and the Fermi level is not too high. Equation (3.25) constitutes the strong dependence of the system's resistivity ρ on model parameters. The weaker, pre-exponential dependences can be found by taking the critical subnetwork into account, as was done in Section 3.1 for the isoenergetic case. The calculation is discussed in detail in Article [II] and yields the following final result for percolation within the VRH problem in the exponential DOS:

$$\sigma_{\mathrm{dc}} = A \frac{e^2 \nu_0}{\alpha \varepsilon_0} \left(\frac{kT}{\varepsilon_0}\right)^\nu \left[\frac{8 n_c}{n_e} \left(\frac{kT}{\alpha \varepsilon_0}\right)^3 \right]^{-\varepsilon_0/kT}, \tag{3.26}$$

where $\nu = 0.875$ is the known scaling exponent for 3D problems [51].

3.2.1 Comparison with the TE Concept

The conductivity σ_{dc} obtained in the previous section can be used to benchmark other approaches for the description of hopping charge transport. A comparison between the exact solution given in Equation (3.26) and the mobility found with the TE approach in Section 2.4, given by Equation (2.13), is shown in Figure 3.5. The mobility μ obtained with the TE concept is found by

evaluating Equation (2.13) for the exponential DOS given in Section 1.2.1 and assuming small carrier concentrations n. The following expression is obtained:

$$\mu_{\mathrm{TE}} = \frac{v_0 e}{n_e \varepsilon_0 \alpha} \frac{B_c}{2\pi} \left[\frac{4\pi}{3B_c} \frac{27 n_e}{8 \exp(3)} \left(\frac{\alpha \varepsilon_0}{kT} \right)^3 \right]^{\varepsilon_0/kT} . \tag{3.27}$$

Figure 3.5 shows, that the temperature dependences of the two approaches agree excellently. The small deviation in the slope stems from the missing temperature dependence of the prefactor of Equation (3.27) as compared to the exact expression obtained via percolation theory, Equation (3.26). The offset between the curves is chosen for easier comparison. Since the percolation approach, Equation (3.26), contains the unknown prefactor A, only the dependence on temperature but not the absolute numbers can be compared.

Chapter 4

The Method of Dissipated Heat

Very recently, a novel method to study VRH in disordered materials was suggested by Nenashev et al. [7], based on the calculation of dissipated heat in a classical resistor network. In this chapter we will repeat the derivation of that approach, discuss the numerical procedures suitable for solving it, and briefly compare some of its results to the TE approach described in Chapter 2.

For the method of dissipated heat, the system of localized states is again interpreted as a system of electrical resistors, as described in Chapter 3, with an applied external electric field $\vec{F}$. The general idea of the approach is to identify those resistors in the network, that produce the highest power during electrical conduction. Then, the dissipated heat of the network can be evaluated as a function of site energies or site distances, giving access to certain quantities of the VRH problem.

4.1 The Resistor Network Model

To find out hopping transport properties from the resistor network, let us start with the power $\mathbb{P}$ of the Joule heat that is dissipated by current flowing through our network. The total power $\mathbb{P}$ can be related to the conductance G and voltage U of the whole network, as well as to the individual resistors:

$$\mathbb{P} = U^2 G = \sum_{(i,j)} U_{ij}^2 G_{ij}, \tag{4.1}$$

where $\sum_{(i,j)}$ is a sum over all pairs of sites in the system. Let us now introduce a variation $\delta\mathbb{P}$ of the power and study its influence on both global and local quantities, while keeping the electric field $\vec{F}$ (and therefore the voltage U) constant. The differential of the local part of Equation (4.1) is

$$\delta\mathbb{P} = 2 \sum_{(i,j)} U_{ij} G_{ij} \delta U_{ij} + \sum_{(i,j)} U_{ij}^2 \delta G_{ij}. \tag{4.2}$$

Inserting the current $I_{ij} = G_{ij} U_{ij}$ and the local potential difference $U_{ij} = \varphi_j - \varphi_i + \vec{F}\vec{d}_{ij}$ ($\vec{d}_{ij} \equiv \vec{r}_j - \vec{r}_i$ is the vector connecting the sites) into Equation (4.2), the first sum vanishes due to Kirchhoff's junction rule $\sum_i I_{ij} = 0$:

$$2 \sum_{(i,j)} U_{ij} G_{ij} \delta U_{ij} = \sum_{i,j} \delta(\varphi_j - \varphi_i + \vec{F}\vec{d}_{ij}) I_{ij}$$

$$= \sum_j \delta\varphi_j \left(\sum_i I_{ij} \right) - \sum_i \delta\varphi_i \left(\sum_j I_{ij} \right) = 0. \tag{4.3}$$

In Equation (4.3) the sum over pairs of sites was transformed into sums over individual sites, losing the factor 2. In addition, $\vec{F}\vec{d}_{ij}$ is constant and therefore not sensitive to variations of $\mathbb{P}$.

We can now return to Equation (4.2) and further evaluate the second sum:

$$\delta\mathbb{P} = \sum_{(i,j)} U_{ij}^2 \delta G_{ij} = \sum_{(i,j)} P_{ij} \frac{\delta G_{ij}}{G_{ij}} = \sum_{(i,j)} P_{ij} \delta(\ln G_{ij}). \tag{4.4}$$

Together with Equation (4.1) we end up with the following result, relating variations of the global conductance G to variations of local conductances G_{ij}:

$$\delta\mathbb{P} = U^2\delta G = \mathbb{P}\frac{\delta G}{G} = \mathbb{P}\delta(\ln G) = \sum_{(i,j)} P_{ij}\delta(\ln G_{ij}) . \tag{4.5}$$

In the following sections, we will use Equation (4.5) to gain access to some physical quantities important for the description of the VRH.

4.2 Finding the Transport Path

Let us first try to relate the resistor network properties to some characteristic energies of hopping conduction. It will turn out, that we can perfectly describe activation energies with the heat method and thereby find the value of the transport energy (TE) (see Chapter 2). If the network is large enough, the charge carrier mobility μ of the VRH problem is proportional to the network's conductivity G. Since activated hopping transition rates $\nu_\uparrow$ are proportional to the Arrhenius equation (see Equation (1.11))

$$\nu_\uparrow \propto \mu \propto \exp\left(-\frac{\varepsilon_a}{kT}\right) \tag{4.6}$$

with the thermal energy kT and the activation energy ε_a, we can write ε_a together with Equation (4.5) as

$$\varepsilon_a = -\frac{\mathrm{d}(\ln\mu)}{\mathrm{d}(1/kT)} = -\frac{\mathrm{d}(\ln G)}{\mathrm{d}(1/kT)} = -\mathbb{P}^{-1}\sum_{(i,j)} P_{ij}\frac{\mathrm{d}(\ln G_{ij})}{\mathrm{d}(1/kT)} . \tag{4.7}$$

In order to evaluate the derivative of G_{ij} we now need to relate properties of our network of resistors to electronic transitions between two sites. This is derived in detail in the monograph of Shklovskii and Efros [33, chapter 4.2], and gives the following general expression for the local conductances G_{ij}:

$$G_{ij} = \frac{e^2}{kT}\nu_0\exp\left(-\frac{2d_{ij}}{\alpha}\right)\frac{f_i(1-f_j)}{\exp\left(\frac{\varepsilon_j-\varepsilon_i}{kT}\right)+1} , \tag{4.8}$$

with the localization length α (assumed equal for all sites), the site energies ε_i and ε_j, and the filling factors f_i and f_j. Assuming thermal equilibrium and small electron concentrations, we can use Fermi-Dirac distributions (Equation (1.3)) in the Boltzmann approximation (Equation (1.4)) for f_i and f_j. Furthermore, we assume

$$\frac{1}{\exp\left(\frac{\varepsilon_j-\varepsilon_i}{kT}\right)+1} \approx \begin{cases} \exp\left(-\frac{\varepsilon_j-\varepsilon_i}{kT}\right) & \text{if } \varepsilon_j > \varepsilon_i , \\ 1 & \text{if } \varepsilon_j < \varepsilon_i . \end{cases} \tag{4.9}$$

Inserting f_i, f_j, and Equation (4.9) into Equation (4.8) yields the following expression for G_{ij}, closely related to the expression of the MA hopping rates given in Equation (1.14):

$$G_{ij} = \frac{e^2}{kT}\nu_0\exp\left(-\frac{2d_{ij}}{\alpha}\right)\exp\left(\frac{\varepsilon_F - \max(\varepsilon_i,\varepsilon_j)}{kT}\right) , \tag{4.10}$$

of which the derivative in Equation (4.7) can easily be calculated. We can now finally write the activation energy as

$$\varepsilon_a = \varepsilon_2 - \varepsilon_1 - kT \tag{4.11}$$

with

$$\varepsilon_1 = \frac{1}{kT} \frac{\mathrm{d}(\varepsilon_F)}{\mathrm{d}(1/kT)} + \varepsilon_F \qquad \text{and} \qquad \varepsilon_2 = \mathbb{P}^{-1} \sum_{(i,j)} P_{ij} \max(\varepsilon_i, \varepsilon_j). \tag{4.12}$$

Let us discuss Equations (4.11) and (4.12) separately for the exponential and Gaussian DOS.

It can easily be verified via Equation (1.5), that the Fermi energy level in the purely exponential DOS depends only weakly on temperature. Therefore, the derivative in Equation (4.12) vanishes and $\varepsilon_1 \approx \varepsilon_F$. According to Equation (4.11), the activation energy ε_a now resembles the energy difference between the Fermi level ε_F and the concentration independent energy ε_2, up to corrections of the order of kT. Because VRH in the exponential DOS (at not too high temperatures) is dominated by activated hops from the Fermi level ε_f to the TE ε_t (see the discussion in Section 2.2), we quickly come to the conclusion that the target energy of the activation, ε_2, can be identified as the TE, $\varepsilon_2 \equiv \varepsilon_t$.

In DOS functions with steeper slope than purely exponential, for example in Gaussian DOSs, we are dealing with a slightly more complicated situation. As already discussed in Section 1.2.3, two cases must be dealt with separately: When the charge carrier concentration is large enough, so that $\varepsilon_F > \varepsilon_\infty$, the majority of carriers occupy states around the Fermi energy ε_F. In that case, the same arguments as for the exponential DOS above apply and we can again identify ε_1 with ε_F. In the opposite case, when $\varepsilon_F < \varepsilon_\infty$, the Fermi level can approximately be written as [7, 52]

$$\varepsilon_F(n, T) \approx -\frac{1}{2} \frac{\sigma^2}{kT} - \frac{1}{kT} \ln\left(\frac{N}{n}\right). \tag{4.13}$$

The derivative in Equation (4.12) is easy to calculate with Equation (4.13) and yields

$$\frac{1}{kT} \frac{\mathrm{d}(\varepsilon_F)}{\mathrm{d}(1/kT)} = -\frac{1}{2} \frac{\sigma^2}{kT} + \frac{1}{kT} \ln\left(\frac{N}{n}\right). \tag{4.14}$$

The result for ε_1 in the Gaussian DOS with small carrier concentration therefore becomes

$$\varepsilon_1 \approx -\frac{\sigma^2}{kT} = \varepsilon_\infty. \tag{4.15}$$

In conclusion, activation in the Gaussian DOS happens from the Fermi level ε_F, if the system is filled with carriers, or from the equilibration energy ε_∞, in case the carrier concentration is small. The target energy ε_2 can in both cases be identified with the transport energy ε_t.

With $\varepsilon_2 \equiv \varepsilon_t$ and the relation given in Equation (4.11), the calculation of the individual powers P_{ij} between each pair of sites in the system therefore allows us to directly calculate the transport energy ε_t for our hopping network as

$$\varepsilon_t = \mathbb{P}^{-1} \sum_{(i,j)} P_{ij} \max(\varepsilon_i, \varepsilon_j). \tag{4.16}$$

The values of the P_{ij} can be found by straightforward computer simulations or by numerically solving the resistor network described above. We will briefly discuss suitable numerical methods in Section 4.4. However, let us first reformulate the above results using distribution functions in order to visualize the findings.

4.3 The Heat Distribution Functions

Let us define the functions $H_\varepsilon(\varepsilon)$ and $H_r(r)$, which are sums of powers P_{ij} with an upper boundary ε for the site energies and r for the distance of the sites, respectively:

$$H_\varepsilon(\varepsilon) \equiv \mathbb{P}^{-1} \sum_{\substack{(i,j) \\ \varepsilon_i < \varepsilon \\ \varepsilon_j < \varepsilon}} P_{ij} \qquad\qquad H_r(r) \equiv \mathbb{P}^{-1} \sum_{\substack{(i,j) \\ |\vec{d}_{ij}| < r}} P_{ij}. \qquad (4.17)$$

It is easy to see, that $H_\varepsilon(-\infty) = H_r(0) = 0$ and $H_\varepsilon(\infty) = H_r(\infty) = 1$. Let us further define the following two distribution functions $h_\varepsilon(\varepsilon)$ and $h_r(r)$:

$$h_\varepsilon(\varepsilon) \equiv \frac{\mathrm{d}H_\varepsilon(\varepsilon)}{\mathrm{d}\varepsilon} \qquad\qquad h_r(r) \equiv \frac{\mathrm{d}H_r(r)}{\mathrm{d}r}. \qquad (4.18)$$

The distributions $h_\varepsilon(\varepsilon)$ and $h_r(r)$ give information about the contribution of hops to a certain energy ε, or with a certain distance r, to the total heat power $\mathbb{P}$ dissipated by the network. We can show, that the transport energy ε_t can be found as the expectation value of the function $h_\varepsilon(\varepsilon)$. Recall Equation (4.16) and calculate the contribution of those pairs of sites (i, j) to ε_t in that equation, that fulfill the condition $\max(\varepsilon_i, \varepsilon_j) \in [\varepsilon, \varepsilon + \mathrm{d}\varepsilon]$. By definition of $H_\varepsilon(\varepsilon)$, that contribution can be expressed via

$$\varepsilon \left[H_\varepsilon(\varepsilon + \mathrm{d}\varepsilon) - H_\varepsilon(\varepsilon) \right] = \varepsilon \frac{\mathrm{d}H_\varepsilon(\varepsilon)}{\mathrm{d}\varepsilon} \, \mathrm{d}\varepsilon = \varepsilon h_\varepsilon(\varepsilon) \, \mathrm{d}\varepsilon. \qquad (4.19)$$

Integration of Equation (4.19) and comparison with Equation (4.16) then leads to the final result

$$\varepsilon_t = \int_{-\infty}^{\infty} \varepsilon h(\varepsilon) \, \mathrm{d}\varepsilon. \qquad (4.20)$$

Hence, the expectation value of the distribution $h(\varepsilon)$ gives us the TE as defined in Equation (4.16). Similarly, we can find the typical hopping distance $\bar{r}$ of hops that are decisive for conduction via the spatial distribution function $h_r(r)$ as

$$\bar{r} = \int_{0}^{\infty} r h_r(r) \, \mathrm{d}r. \qquad (4.21)$$

Plotting the distributions $h_\varepsilon(\varepsilon)$ and $h_r(r)$ therefore shows the importance of different energies and hopping lengths for the long-range conductivity. Some results are presented in the following section.

4.4 Numerical Solutions of the Resistor Network

In order to gain quantitative information about our system using the heat method as described in the previous sections, the resistor network must be solved numerically to obtain the dissipated powers P_{ij} of the individual resistances. If those are known, we can use Equation (4.16) to calculate the TE or study the distribution functions $h_\varepsilon(\varepsilon)$ and $h_r(r)$ by following the derivation given in Section 4.3.

Numerical Procedures   The simplest way to obtain those powers is by numerical solution of the system of equations given by Kirchhoff's junction rule

$$
\begin{aligned}
\sum_j I_{ij} = 0 &= \sum_j U_{ij} G_{ij} \\
&= \sum_j \left(\varphi_j - \varphi_i + \vec{F}\vec{d}_{ij} \right) G_{ij}
\end{aligned}
\tag{4.22}
$$

for the unknown potentials φ_i. The conductances G_{ij} are given by Equation (4.10), assuming that all relevant energies are far above the Fermi level. The system can be brought into the form $\underline{M}\vec{\varphi} = \vec{b}$ by rewriting Equation (4.22) into

$$
\sum_j \underbrace{\left(G_{ij} - \delta_{ij} \sum_k G_{ik} \right)}_{M_{ij}} \varphi_j = \underbrace{- \sum_j \vec{F}\vec{d}_{ij} G_{ij}}_{b_i} \, ,
\tag{4.23}
$$

and can be solved for the potentials $\vec{\varphi}$ by one's favourite linear algebra library. With φ_i and given external electric field $\vec{F}$, the individual voltages U_{ij} can be calculated and the powers $P_{ij} = U_{ij}^2 G_{ij}$ can be obtained using the conductance from Equation (4.10).

Another way of obtaining P_{ij} is by numerically finding the occupation probabilities p_i of the sites during steady-state conduction, e.g., by the balance equation (BE) approach described in Section 7.1. In that case, the currents I_{ij} are given by

$$
I_{ij} = e p_i \left(1 - p_j \right) \Gamma_{ij} - e p_j \left(1 - p_i \right) \Gamma_{ji} \, ,
\tag{4.24}
$$

where Γ_{ij} is the hopping rate between sites i and j. The individual dissipated powers can then be calculated via $P_{ij} = I_{ij}^2/G_{ij}$, with G_{ij} given by Equation (4.10) in case of low Fermi level or, more generally, by

$$
G_{ij} = \frac{e^2}{kT} f(\varepsilon_i, \varepsilon_F) \left(1 - f(\varepsilon_j, \varepsilon_F) \right) \Gamma_{ij} \, ,
\tag{4.25}
$$

with the Fermi distribution $f(\varepsilon, \varepsilon_F)$. This second method may be preferable if one has already implemented the BE method and quickly wants to add the capability of solving the heat method to the program.

Note, that in both algorithms one should include periodic boundary conditions so that a steady state current flow is possible. In addition, one should carefully check for finite size effects in the calculation results. The requirements for large system sizes is probably the main culprit for numerically inaccurate results and, based on those, misinterpretation of the findings. The performance of the calculation can be greatly improved by only considering pairs of sites within a certain (sufficiently large) cut-off radius. The resulting matrix $\underline{M}$ then becomes sparse and the system can be solved much more efficiently.

Results   Let us finally present some example calculations of the heat distributions $h_\varepsilon(\varepsilon)$ and $h_r(r)$. The results are shown in Figures 4.1 and 4.2. The simulated system had $30 \times 30 \times 30$ sites (for $kT = 0.2\sigma$, the system size was chosen as $60 \times 60 \times 60$) randomly distributed in space. The site
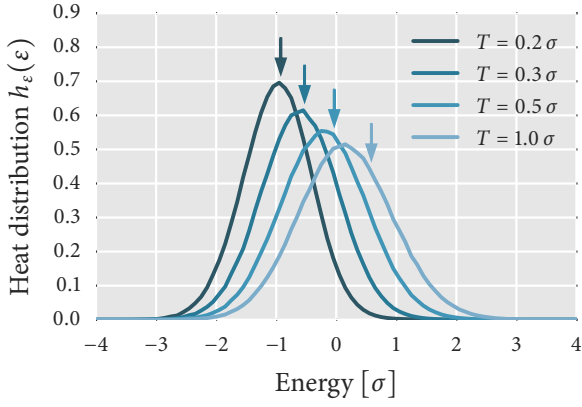
Figure 4.1: The heat distribution function $h_e(\varepsilon)$ from Equation (4.18) versus energy at different temperatures kT. The localization length was $\alpha = 0.3N^{-1/3}$. Arrows mark the position of the TE ε_t as calculated via Equation (2.9). Figure reproduced from Article [I].
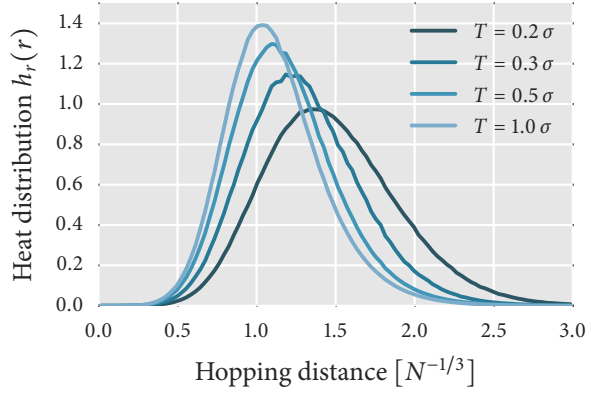

Figure 4.2: The spatial heat distribution $h_r(r)$ showing the most important hopping distances $\bar{r}$ for long-range conduction, obtained via Equation (4.18) at different values of the temperature kT. The localization length was $\alpha = 0.3N^{-1/3}$. Figure reproduced from Article [I].

energies were drawn from the Gaussian DOS as given in Equation (1.7). Different temperatures are plotted in the figures.

Figure 4.1 clearly shows, that sites with low energies do not participate in resistances that dissipate much power. Hence, they are not important for charge transport. The same is true for sites with very high energies. The main contribution to conduction comes from sites within about 1σ below the center of the DOS, $\varepsilon = 0$. The centers of mass of the curves (approximately corresponding to their maxima) resemble the TE ε_t as given by Equation (4.20). To compare this approach with the optimization method for the transport energy derived in Section 2.3, the arrows indicate the position of the TE as calculated via Equation (2.9). For lower temperatures, the results agree very well and deviations occur only at very high temperatures. (Note, that σ in ODSs is much larger than kT at room temperature.) This agreement confirms both the heat method as well as the optimization approach for the TE.

In Figure 4.2 a similar behaviour is found for the hopping distances. Very short hops are so sparse that they don't contribute much to charge transport (i.e., don't dissipate much heat). The same holds for very long hops, which are unlikely to happen due to the low tunneling probability. The centers of mass $\bar{r}$ of the curves mark the hopping distance decisive for transport. In agreement with the VRH model, the decisive hopping distances become shorter for higher temperatures T: The higher T, the less important is the energy difference between the sites participating in the hopping transitions and, consequently, carriers will choose spatially closer sites. $\bar{r}$ saturates for high temperatures to the typical intersite distance $N^{-1/3}$, indicating NNH.

Further results can be found in Chapter 6 of Article [I]. The heat method can also be exploited to calculate the influence of DOS modifications on the conductivity G of the sample, as shown in Chapter 6.4 of Article [I].

Chapter 5

Multiple Trapping in High Electric Fields

In this chapter, we show how the multiple trapping (MT) model can be used to explain the dependence of charge transport in disordered inorganic semiconductors on high electric fields.

The basis for the study is experimental data on field-dependent charge transport in polycrystalline lead oxide (PbO) measured by the group of Prof. A. Reznik at Lakehead University in Thunder Bay, Canada. In the following sections, the experimental data for electron and hole transport is reviewed separately. We show, that the transport mode for both carrier species is most likely MT. The hole transport is almost independent of the applied electric field and can be described by simple expressions for high-temperature MT. To explain the field dependence of electron transport, a simple phenomenological and a more sophisticated quantum mechanical model are considered.

5.1 Hole Transport in Lead Oxide

Figure 5.1 shows the hole mobility μ_h as a function of the applied electric field F measured via photo generated charge carrier extraction with linearly increasing voltage (photo-CELIV) [53]. It is clear from the data, that the hole mobility is almost independent of the electric field and changes only slightly with temperature. The weak temperature dependence indicates, that the energy scale of the DOS ε_0 is smaller than the thermal energy kT, so that carriers assemble right below the mobility edge ε_c. Transport is therefore solely determined by the trapping and detrapping coefficients and does not strongly depend on temperature and field. This situation in the exponential DOS is described in Section 1.2.1 and shown in Figure 1.5 for $kT > \varepsilon_0$.

An analytical expression for the temperature dependence of the MT carrier mobility for $kT > \varepsilon_0$ in a disordered system with exponential DOS can be quickly derived as follows. The drift mobility μ of charge carriers during multiple trapping is given by the free carrier mobility μ_0 multiplied by the ratio of free carriers in the conduction band to the total carrier concentration:

$$\mu = \mu_0 \frac{n_{\text{free}}}{n_{\text{total}}}, \tag{5.1}$$

where n_{free} and n_{total} are the free and total carrier concentrations. The latter concentration can be calculated by evaluating Equation (1.5) for the exponential DOS in the Boltzmann limit, which is applicable in case $kT > \varepsilon_0$:

$$
\begin{aligned}
n_{\text{total}} &= \int_{-\infty}^{0} g(\varepsilon) f(\varepsilon, \varepsilon_F)\, \mathrm{d}\varepsilon \\
&\approx N \exp\left(\frac{\varepsilon_F}{kT}\right) \frac{1}{1 - \frac{\varepsilon_0}{kT}}.
\end{aligned}
\tag{5.2}
$$

The concentration of carriers above the mobility edge, n_{free}, is found by multiplying the concentration N of traps at $\varepsilon = \varepsilon_c \equiv 0$ with the occupation probability at that energy (again in the Boltzmann
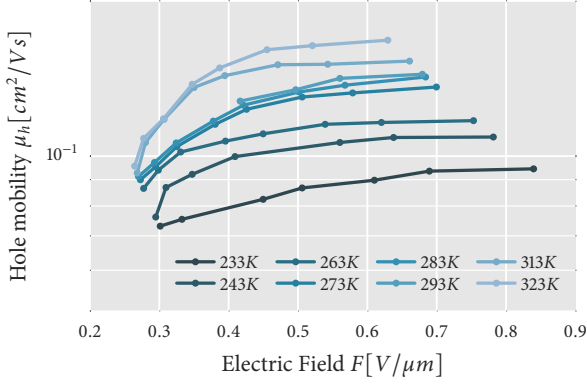
Figure 5.1: Experimental data for hole mobility versus applied electric field at different temperatures.
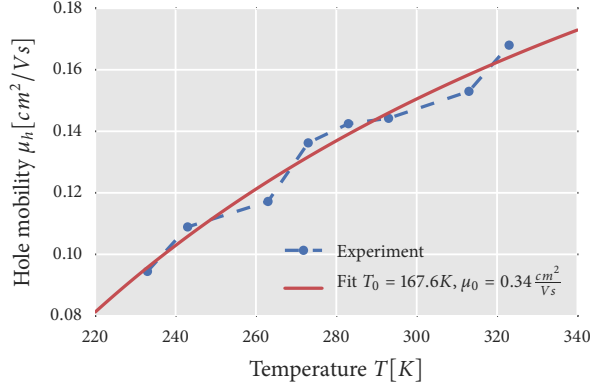

Figure 5.2: Fit by Equation (5.4) to the temperature dependence of the hole mobility extracted from Figure 5.1.

limit):

$$n_{\text{free}} \approx N \exp\left(\frac{\varepsilon_F}{kT}\right). \tag{5.3}$$

Inserting Equations (5.2) and (5.3) into Equation (5.1) gives the following simple expression for the carrier mobility during MT for the case $kT > \varepsilon_0$ [54]:

$$\mu_h = \mu_0 \left(1 - \frac{\varepsilon_0}{kT}\right). \tag{5.4}$$

Equation (5.4) can be fitted to the temperature dependence of the hole mobility from Figure 5.1 to obtain the values of ε_0 and μ_0. The fit yields $\mu_0 \approx 0.34 \, \text{cm}^2 \, \text{V}^{-1} \, \text{s}^{-1}$ and $\varepsilon_0/k \approx 167.6 \, \text{K}$, which are reasonable values for the exponential DOS and hole mobility in inorganic disordered materials. Clearly, $\varepsilon_0/k < T$ for all measured temperatures.

The free hole mobility μ_0 may also carry a weak dependence on the electric field, as shown for instance by Juška et al. [55]. Hence, the small field dependence observed experimentally in Figure 5.1 can be attributed to the free hole mobility μ_0.

With the $\mu_h(T)$ curve corresponding nicely to the simple model derived above for reasonable fitting parameters, we can conclude that the model of MT with $\varepsilon_0/k < T$ and holes distributed as shown in Figure 1.5 appropriately describes the electronic structure of the hole DOS in PbO.

5.2 Electron Transport in Lead Oxide

Electron transport properties were studied in the samples using the same technique (photo-CELIV) as applied for the hole transport described above. The results are shown in Figures 5.3 and 5.4. Clearly, the transport mode and dependences on temperature and field are significantly different to those for holes.

As shown in Figure 5.4, the dependence of the electron mobility μ_e on temperature T, both in the zero field case and for high electric field, is described by an Arrhenius equation and therefore
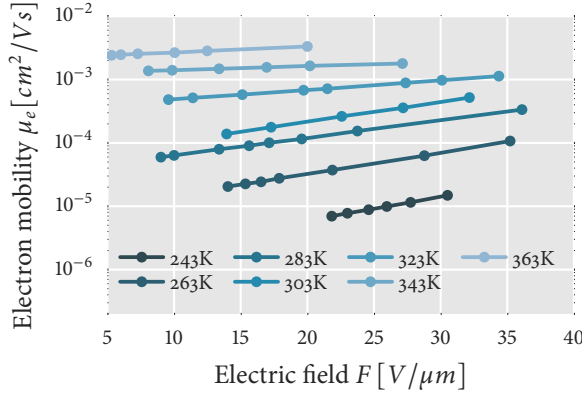
Figure 5.3: The electron mobility μ_e versus electric field F for different temperatures as measured by photo-CELIV.
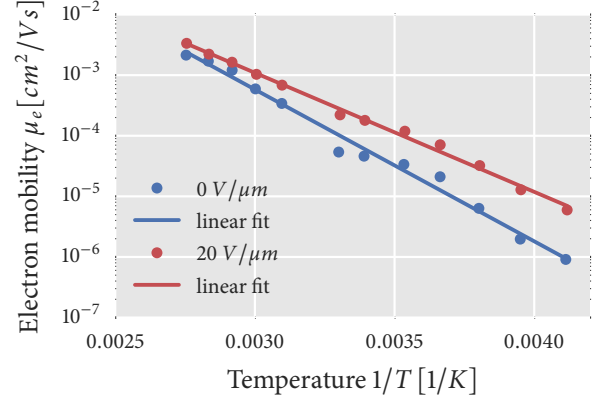
Figure 5.4: Inverse temperature $(1/T)$ dependence of the measured electron mobility μ_e for two different electric fields F. Arrhenius type lines are drawn as a guide to the eye.

$\log(\mu_e) \propto 1/T$. This observation strongly suggests activated MT like transport for electrons in PbO. In that case, one can calculate an effective activation energy ε_a via

$$\varepsilon_a = -\frac{\mathrm{d}\log(\mu_e)}{\mathrm{d}\frac{1}{kT}} . \tag{5.5}$$

From Figure 5.4 one can then find the field dependence of the activation energy $\varepsilon_a = \varepsilon_a(F)$, which is plotted in Figure 5.5 for the PbO samples studied here along with the data for electron transport in a-Si:H and electron and hole transport in a-Se measured in Reference [56]. As visible from the decreasing ε_a with increasing field F, the electric field supports activation towards the mobility edge during the MT process by decreasing the activation energy.

In the following sections, we present two possible theoretical models to explain the field dependences shown in Figures 5.3 and 5.5.

5.2.1 Phenomenological Model of the Field Dependency

A simple approach to explaining the presented experimental data is to follow the assumptions of Juška et al. [55] and Antoniadis and Schiff [57]. In both studies it is assumed, that the activation energy ε_a is simply lowered by the energy gain due to the electric field F within some distance x in field direction:

$$\mu_e(F) = \mu_0 \exp\left(-\frac{\varepsilon_a - eFx}{kT}\right) , \tag{5.6}$$

where the value of x in both studies serves as a fitting parameter. Obviously, Equation (5.6) perfectly reproduces the linear dependence of $\log(\mu_e)$ on electric field F, that is visible in Figure 5.3. Furthermore, the linear decrease of $\varepsilon_a(F)$ for PbO with respect to the field strength, shown in Figure 5.5, is trivially fulfilled. The value of x for the different experimental data shown in Figure 5.5 can be found by fitting the function $\varepsilon_a(F) = \varepsilon_a^0 - eFx$ to the measurements, and varies between approximately 30 Å for PbO and 15 Å for the a-Se sample measured in Reference [56].
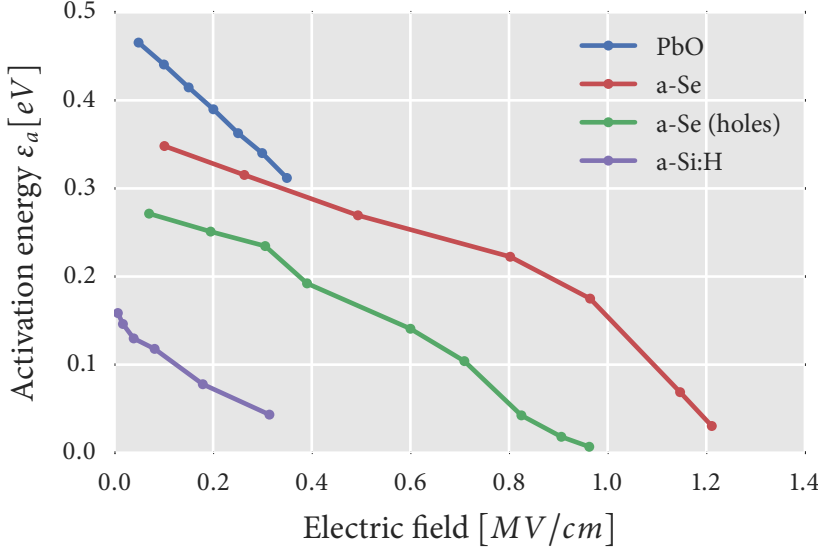
Figure 5.5: Activation energy ε_a versus applied electric field F for various materials. All measured data is for electrons, except for the curve marked as "a-Se (hole)". The data for a-Si:H and a-Se is taken from Reference [56]. Lines are drawn as a guide to the eye.

This raises the important question of the physical interpretation of the model. Electrons trapped in localized states below the mobility edge in PbO, a-Si:H, and a-Se should travel over distances of more than 10 Å in order to gain as much activation energy as shown in Figure 5.5. Neither Juška et al. [55] nor Antoniadis and Schiff [57] give any physical arguments for the value of x appearing in Equation (5.6). It is therefore an open question, whether this model is applicable to the measurements in PbO and other disordered inorganic semiconductors, and how to interpret the experimental observations within this approach.

5.2.2 Exact Quantum Mechanical Model of Trap Ionization

A radically different approach as compared to the phenomenological model described in Section 5.2.1 is presented in this section. Let's assume for simplicity, that carriers can be captured into a set of isoenergetic traps with energy ε_a^0 below the mobility edge ε_c. This assumption is a rough approximation for an exponential DOS, where most electrons gather around the Fermi level and becomes better for steeper DOSs due to the presence of the equilibration energy ε_∞, introduced in Section 1.1.2. Let us further assume, that ionization from the trap states, i.e., activation towards the mobility edge, depends on the electric field in the way calculated by Karpus and Perel [58]. The mobility then becomes

$$
\mu_e(F) = \mu_e^0 \exp\left[\frac{\varepsilon_a^0}{\hbar\omega} \left(2\omega\tau - 1 + C^{-1} \right) + \frac{e^2 F^2 \tau^3}{3\hbar m_{\text{eff}}} \right],
$$
$$
\tau \equiv \frac{1}{2\omega} \log\left[C \exp\left(\frac{\hbar\omega}{kT} \right) - C \right],
$$

(5.7)

with a (weakly) field dependent free electron mobility $\mu_e^0(F)$ [55], the atomic vibrational frequency ω, a parameter C related to the electron-phonon coupling which we will treat as a free parameter, and the effective electron mass m_{eff}. In Equation (5.7), the mobility edge is as usual set to zero, $\varepsilon_c = 0$.

Assuming $C = 1$ and $\hbar\omega \gg kT$ in Equation (5.7) results in the well-known Keldysh effect [58, 59]. Below we show, that this is sufficient to roughly reproduce the experimental findings. Under

Figure 5.6: Fit of Equation (5.9) to the field dependent activation energy of various inorganic disordered semiconductors. The solid lines are the fits, the points are experimental measurements. The data for a-Se and a-Si:H is taken from Reference [56].
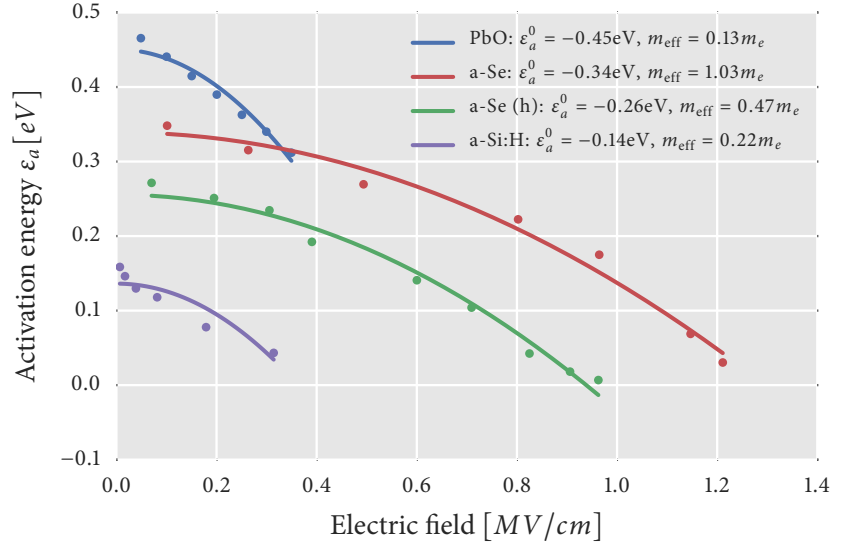
these assumptions, the expression for the mobility reads

$$
\mu_e(F) \approx \mu_e^0 \exp\left[\frac{\varepsilon_a^0}{kT} + \frac{e^2 F^2 \hbar^2}{24(kT)^3 m_{\text{eff}}} \right].
\tag{5.8}
$$

With the activation energy defined as in Equation (5.5), $\varepsilon_a(F)$ can be determined as

$$
\varepsilon_a \approx -\varepsilon_a^0 - \frac{e^2 \hbar^2}{8 m_{\text{eff}}} \frac{F^2}{(kT)^2}.
\tag{5.9}
$$

One can see from Equations (5.8) and (5.9), that the mobility is not Arrhenius-like and the activation energy decreases not linearly with the field. At first glance, both of these observations contradict the experimental measurements shown in Figures 5.4 and 5.5. Moreover, the linear dependence of $\log(\mu_e)$ on field F, shown in Figure 5.3, is not recovered. Let us nevertheless try to describe the measurements with the approach by Karpus and Perel [58].

Fitting Equation (5.9) to the curves shown in Figure 5.5 yields effective electron masses m_{eff} and trap energy levels ε_a^0 for the respective materials. The fit results are shown in Figure 5.6. The resulting values of the effective carrier masses, as well as the trap levels, are within reasonable ranges for amorphous inorganic materials. The parabolic decrease of the calculated activation energy resembles the expected behaviour from Equation (5.9).

Using the fit parameters of PbO, $\varepsilon_a^0 = -0.45\,\text{eV}$ and $m_{\text{eff}} = 0.13 m_e$ from Figure 5.6, we can study the field and temperature dependence of the mobility given in Equation (5.8) in comparison with the experimental measurements shown in Figures 5.3 and 5.4. The prefactor of the mobility expression, the free carrier mobility μ_e^0, is chosen so that the absolute values agree. The results for the temperature dependence of μ_e are plotted in Figure 5.7. The Arrhenius-like behaviour is almost recovered and the data agree sufficiently well with the experimental measurements. The field dependence of the mobility is shown in Figure 5.8. Also here, Equation (5.8) recovers the experimental data and gives a good approximation to the increasing mobility with increasing external electric field.
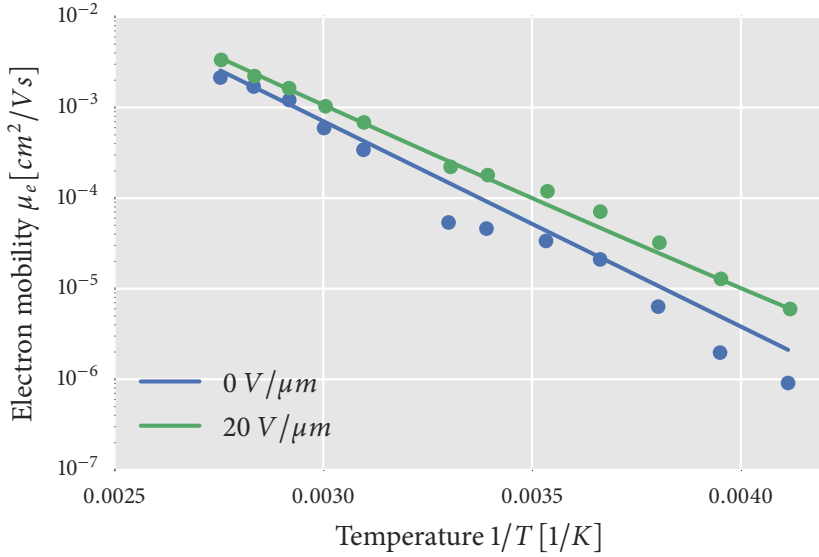
Figure 5.7: Equation (5.8) and the experimental values of the electron mobility μ_e of PbO versus temperature for two different values of the applied electric field.

In conclusion, the assumption of a single trap level ε_a^0 with release rates given by the microscopic theory of Karpus and Perel [58] in the limit of the Keldysh effect [59] seem to be a good approximation to the experimentally observed dependences of electron transport in PbO on temperature and electric field. By assuming not a single trap level, but a distributed DOS function, e.g., the exponential DOS as given in Equation (1.6), one may be able to find more accurate agreement between theory and experiment. In addition, using the full approach by Karpus and Perel [58] Equation (5.7), with the additional parameters C and ω, the analytic theory could be fitted to experimental data even better. One should also keep in mind that the free electron mobility μ_e^0, which we assumed constant in our derivation, also weakly depends on the electric field [55].
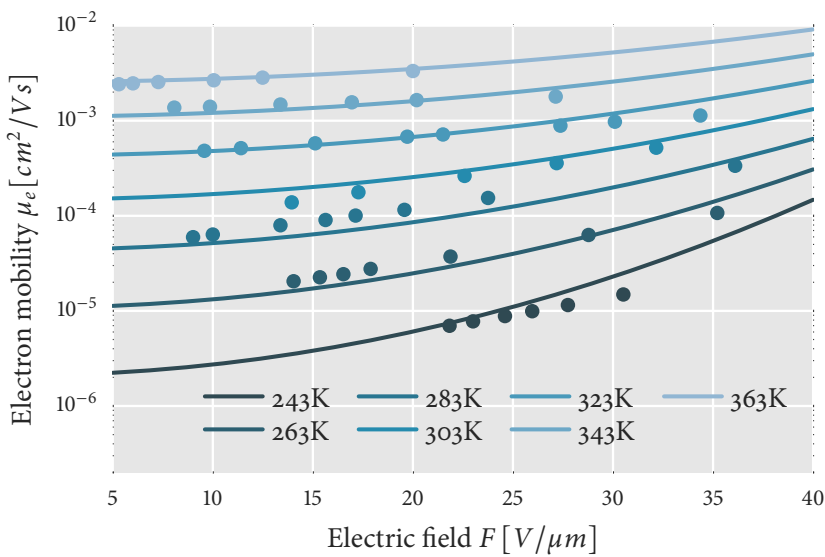

Figure 5.8: Equation (5.8) and the experimental values of the electron mobility μ_e of PbO versus electric field for different temperatures.

Chapter 6

Carrier Recombination in the MT model

Along with the charge transport characteristics, charge carrier recombination decisively influences the performance of optoelectronic devices, particularly those based on disordered semiconductors. In this chapter, we study bimolecular carrier recombination in the framework of the MT transport mode. The traditional MT description applies to amorphous inorganic semiconductors, with bands of extended states and exponentially decreasing localized band tails in the mobility gap (see Sections 1.1.1 and 1.2.1). However, as shown in Chapter 2, the TE concept allows us to exploit the MT formalism also for ODSs by replacing the mobility edge with the TE level ε_t, at which carriers conduct. Furthermore, as shown in Section 2.3, the TE ε_t is situated close to the center of the Gaussian DOS for reasonable parameter values [6, 15]. We will therefore use the MT formalism both for organic and inorganic disordered semiconductors, with the mobility edge or TE as the reference energy, $\varepsilon_t = \varepsilon_c = 0$.

A popular method to treat recombination dynamics in disordered systems was suggested by Orenstein and Kastner [14], who successfully treated recombination within the MT framework in the exponential DOS, under the assumption that recombination processes are more efficient than trapping events. Adriaenssens et al. [60] later applied the model of Orenstein and Kastner to describe the transient photoconductivity in bulk homogeneous systems. The problem of recombination in disordered systems was also tackled recently by Tachiya and Seki [61], who presented an exact analytic, but rather elaborate theory. We will use their results for testing our new approach in Section 6.2. Recombination during MT was also studied by means of computer simulations [62], the results of which, however, suffered from finite size effects and are therefore not reliable [63].

In Article [IV], a new method to describe recombination during MT is described that is free of any assumptions regarding recombination efficiency and can easily be applied to different DOS shapes. In the following sections, the method is summarized and some results for the carrier concentration dynamics in exponential and Gaussian DOSs are given.

6.1 Theoretical Model

We are interested in recombination kinetics after a pulsed photoexcitation of charge carriers from the valence into the conduction band of a disordered semiconductor. We assume the carriers to be excited into the extended states above the mobility edge (or, in case of a Gaussian DOS, above the TE) at time $t = 0$. The carriers then have two options:

1. They can be trapped into localized states below the mobility edge and continue the MT process. The efficiency of this process is described by the capture coefficient b_t.

2. They can recombine with carriers of the opposite kind with a recombination coefficient b_r.

After each trapping and subsequent detrapping event, carriers can again either recobine or be trapped into a localized state.
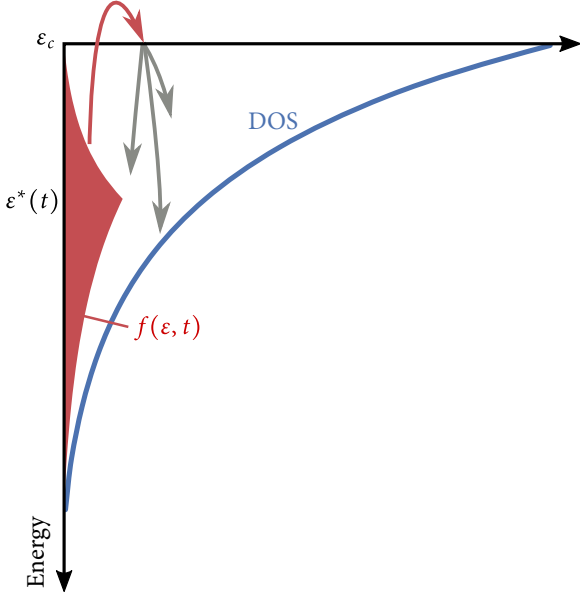
Figure 6.1: The threshold energy $\varepsilon^*(t)$ and the charge carrier distribution above and below $\varepsilon^*(t)$. The carriers with $\varepsilon > \varepsilon^*(t)$ are distributed according to Boltzmann statistics, and the carrier distribution at $\varepsilon < \varepsilon^*(t)$ parallels the shape of the DOS.

The time evolutions of two quantities are of interest: The total concentration of excited carriers $N(t)$, and the concentration of free carriers $n(t)$. The latter are those carriers that occupy states above the mobility edge at time t, i.e., contribute to the conduction process. The concentration $N(t)$ is equal for holes and electrons after an excitation from the valence into the conduction band, and we assume that only one species must be mobile for successful recombination. Therefore, recombination happens between free carriers of one species $n(t)$ and all carriers of the opposite species, and has bimolecular character. The dynamics for $N(t)$ is determined by [14]

$$\frac{dN(t)}{dt} = -b_r N(t) n(t).$$
(6.1)

In case a carrier is trapped into a localized state with energy $\varepsilon < 0$, on average it is detrapped and activated back to the mobility edge after the time

$$\tau(\varepsilon) = v_0^{-1} \exp\left(-\frac{\varepsilon}{kT}\right),$$
(6.2)

where v_0 is the *attempt-to-escape* frequency known from the Miller-Abrahams rates in Equation (1.11). After some time t has passed, carriers trapped in shallow states with $\tau(\varepsilon) < t$ have had enough time to be detrapped, while carriers in deep states with $\tau(\varepsilon) > t$ are still stuck. Let us introduce the time dependent threshold energy $\varepsilon^*(t)$, that is defined as follows: Carriers trapped in states with $\varepsilon > \varepsilon^*(t)$ had enough time for at least one detrapping event within time t and can be considered in thermal equilibrium. Consequently, their distribution follows Boltzmann statistics (Equation (1.4)) and decreases exponentially towards the mobility edge. On the contrary, carriers in deeper states, $\varepsilon < \varepsilon^*(t)$, have not had enough time to thermalize within time t and are assumed to be frozen into their initial form, parallel to the DOS. Because the probability of trapping from the mobility edge is not energy dependent, the distribution of carriers from $\varepsilon^*(t)$ downwards follows the shape of the DOS function. The resulting total carrier distribution $f(\varepsilon, t)$ in an exponential DOS is depicted in Figure 6.1.

Knowing $f(\varepsilon, t)$, we can now find equations for both the total density of charge carriers $N(t)$ and the density of free carriers $n(t)$. The former is given by the integral over the carrier distribution weighted with the DOS, similar to Equation (1.5) in thermal equilibrium. Because the carrier concentration cuts off almost abruptly above $\varepsilon^*(t)$ (see Figure 6.1), the following approximation can be used:

$$
\begin{aligned}
N(t) &= \int_{-\infty}^{0} f(\varepsilon, t) g(\varepsilon) \, d\varepsilon \\
&\approx F(t) \int_{-\infty}^{\varepsilon^*(t)} g(\varepsilon) \, d\varepsilon,
\end{aligned}
\tag{6.3}
$$

where $F(t)$ is the mean occupation number for sites below $\varepsilon^*(t)$. The dynamics of $F(t)$ is related to the concentration $n(t)$ of free carriers through [14]

$$
\frac{dF(t)}{dt} = b_t n(t) \left[1 - F(t) \right].
\tag{6.4}
$$

Equation (6.4) describes the situation where $n(t)$ carriers are available for trapping into the DOS of localized states, with only a fraction $[1 - F(t)]$ of the sites being unoccupied and therefore available as trapping targets. The factor b_t is the trapping coefficient that determines the efficiency of trapping.

For the concentration of free carriers $n(t)$, we can write down the relation [14]

$$
\frac{n(t)}{N(t)} = \frac{g(0)}{g(\varepsilon^*(t))} \exp\left(\frac{\varepsilon^*(t)}{kT} \right),
\tag{6.5}
$$

which arises from the ratio of the carrier distributions $g(\varepsilon) f(\varepsilon, \varepsilon_F)$ for the energies $\varepsilon = 0$ and $\varepsilon = \varepsilon^*(t)$ in the Boltzmann limit.

Equations (6.2) to (6.5) constitute a system of ordinary differential equations (ODEs) for the quantities $N(t)$, $n(t)$, $\varepsilon^*(t)$, and $F(t)$ that can be solved self-consistently. The parameters for the calculation are the ratio of recombination and trapping coefficients b_r/b_t, temperature kT, the shape of the DOS $g(\varepsilon)$, and the initial carrier concentration $N(0)$.

6.2 Recombination Dynamics in the Exponential DOS

Inserting an exponential DOS shape, as defined in Equation (1.6), into Equations (6.3) and (6.5) and solving the system of ODEs derived in the previous section numerically allows us to compare the method with the exact solution by Tachiya and Seki [61] and to study the time dependence of the new threshold energy $\varepsilon^*(t)$.

Figure 6.2 (a) shows the dynamics of the Fermi level $\varepsilon_F(t)$, the demarcation energy $\varepsilon_d(t)$, and the energy $\varepsilon^*(t)$ for two different recombination rates b_r. The demarcation energy is defined similar to $\varepsilon^*(t)$: Carriers above $\varepsilon_d(t)$ had enough time to thermalize while carriers below $\varepsilon_d(t)$ are still trapped. The energy is therefore given by the criterion

$$
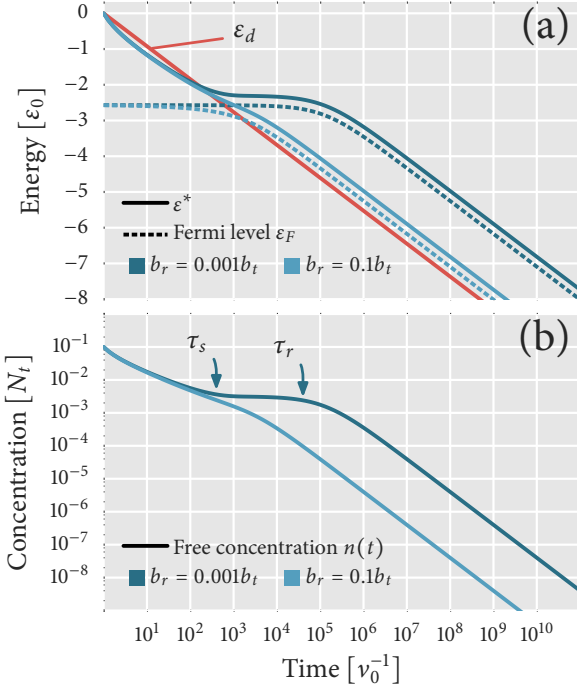\tau(t) t \approx 1 \quad \Rightarrow \quad \varepsilon_d(t) = -kT \ln(\nu_0 t).
\tag{6.6}
$$

Figure 6.2: The dynamics of (a) the energies $\varepsilon^*(t)$, ε_F, ε_d and (b) the free carrier concentration $n(t)$ versus time. Temperature was fixed at $kT = 0.4\varepsilon_0$. Figure taken from Article [IV].
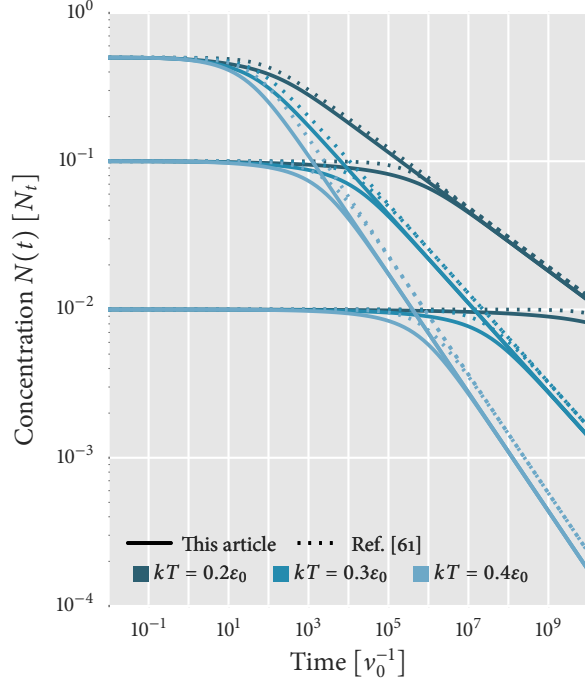
Figure 6.3: Comparison of the dynamics of the total concentration $N(t)$ obtained by the method introduced in Section 6.1 to the results of Tachiya and Seki [61]. b_r/b_t was fixed at 0.1. Figure taken from Article [IV].

However, the demarcation energy, as used for instance by Orenstein and Kastner [14] in place of our $\varepsilon^*(t)$, neglects saturation effects when recombination is not efficient enough to quickly deplete the DOS. In that case, $\varepsilon_d(t)$ crosses the Fermi level at some point and becomes meaningless for the description of the recombination dynamics. In Figure 6.2 (a) this behaviour is reproduced: $\varepsilon^*(t)$ follows approximately the demarcation energy $\varepsilon_d(t)$ until the Fermi level $\varepsilon_F(t)$ comes close. At that point, the dynamics of $\varepsilon_F(t)$ is recovered by $\varepsilon^*(t)$. When recombination is efficient enough so that no saturation can happen, the energy $\varepsilon^*(t)$ decreases with the demarcation energy $\varepsilon_d(t)$ at all times, as plotted for the case $b_r = 0.1b_t$.

The free carrier concentration $n(t)$ is shown in Figure 6.2 (b). Two features are clearly visible for slow recombination $b_r = 0.001b_t$: After the initial excitation, the density $n(t)$ first decreases but then shows a plateau after some time τ_s. This is the point, where the total distribution of charge carriers reaches thermal equilibrium and its maximum is given by the Fermi level $\varepsilon_F(t)$. After the time τ_r, recombination sets in, the DOS depletes, and $n(t)$ decreases again.

The results for the total carrier concentration $N(t)$ at different temperatures T and initial concentrations $N(0)$ is plotted in Figure 6.3, along with the solutions obtained by the method of Tachiya and Seki [61]. It is clearly visible, that the two approaches agree well with only minor deviations. For long times t, the known power law dependency $N(t) \propto t^{-kT/\varepsilon_0}$ is observed over several orders of magnitude of $N(t)$ [14, 61].
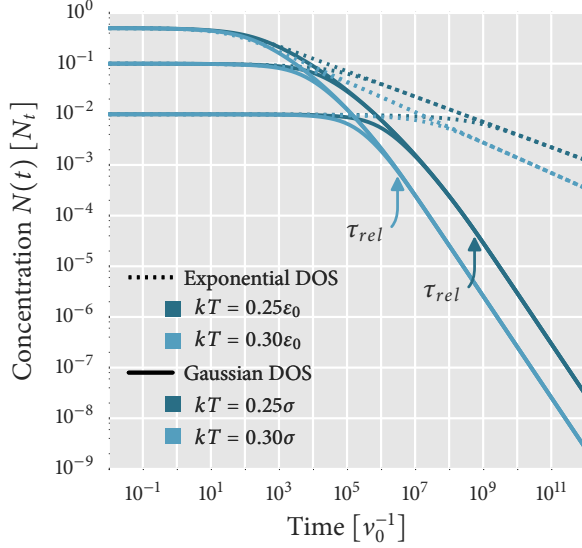
Figure 6.4: Total carrier concentration $N(t)$ for two temperatures and initial concentrations $N(0)$ in the exponential and Gaussian DOS. b_r/b_t was fixed at 0.1. Figure taken from Article [IV].
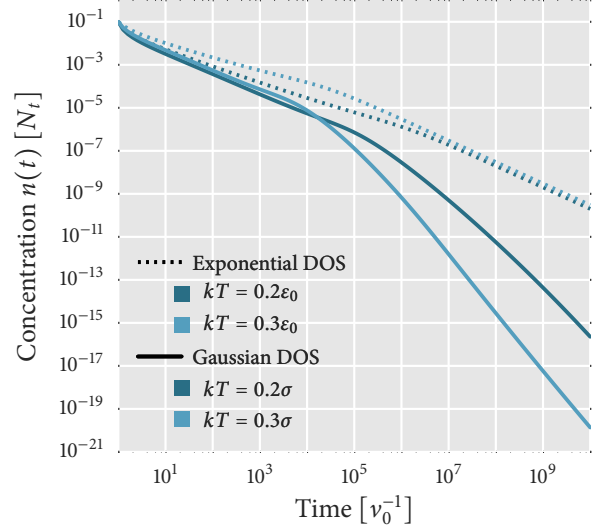
Figure 6.5: Free carrier concentration $n(t)$ for two temperatures and initial concentrations $N(0)$ in the exponential and Gaussian DOS. b_r/b_t was fixed at 0.1. Figure taken from Article [IV].

6.3 Recombination Dynamics in the Gaussian DOS

The Gaussian DOS, typical for ODSs, is a bit more complicated to treat due to its distinct features regarding the electron distribution. In Section 1.2.2 we found, that in thermal equilibrium the majority of charge carriers in the Gaussian DOS occupy states around the so-called *equilibration energy* ε_∞, which is independent on the concentration of carriers as long as $\varepsilon_F < \varepsilon_\infty$. This behaviour is not accounted for in the recombination treatment introduced in Section 6.1, where we assume the majority of carriers to sit around the threshold energy $\varepsilon^*(t)$ at all times. It is the approximation of Equation (6.3) that fails for $\varepsilon^*(t) \leq \varepsilon_\infty$, which, however, is needed to solve the dynamics of the carrier concentrations.

To take the equilibration energy into account, we need to artificially modify the dynamics of $\varepsilon^*(t)$ in the following way. During the relaxation process, when $\varepsilon^*(t) > \varepsilon_\infty$, we stick to the dynamics as obtained from the system of ODEs defined in Section 6.1. At time $\tau_{\rm rel}$, when carriers have relaxed to ε_∞, i.e., $\varepsilon^*(\tau_{\rm rel}) \approx \varepsilon_\infty$, we replace the threshold energy $\varepsilon^*(t)$ with the time independent equilibration energy ε_∞. In that regime, by combining Equations (6.1) and (6.5) and replacing $\varepsilon^*(t)$ with ε_∞, the following ODE for the carrier concentration $N(t)$ is obtained:

$$\frac{dN(t)}{dt} = -b_r N(t)^2 \exp\left[-\frac{1}{2}\left(\frac{\sigma}{kT} \right)^2 \right],\tag{6.7}$$

having the analytic solution

$$N(t) = \frac{N(\tau_{\rm rel})}{1 + (t - \tau_{\rm rel})N(\tau_{\rm rel})b_r \exp\left[-\frac{1}{2}\left(\frac{\sigma}{kT} \right)^2 \right]}.\tag{6.8}$$

The free carrier dynamics $n(t)$ in the regime $t > \tau_{\text{rel}}$ can easily be obtained from Equation (6.5) and evaluates to

$$n(t) = N(t) \exp\left[-\frac{1}{2}\left(\frac{\sigma}{kT}\right)^2\right]. \tag{6.9}$$

The solutions for carrier recombination in the Gaussian DOS, compared to those in the exponential DOS, are plotted in Figures 6.4 and 6.5. For simplicity, we assumed the energy scales ε_0 and σ of the spectra to be equal. The recombination rate was fixed at $b_r = 0.1b_t$. It is clearly visible from both the total carrier concentration $N(t)$ and the free carriers $n(t)$, that recombination in the Gaussian DOS is much more efficient than in the exponential DOS. This different behaviour is easy to understand: Carriers in the Gaussian DOS dive in energy only until they reach ε_∞, from where activation and subsequent recombination is faster than the corresponding processes in the exponential DOS, where the relaxation of carriers is limited only by the Fermi level.

The time evolution of the excited carrier concentration $N(t)$ can be determined experimentally by measuring the intensity of the transient photoabsorption, which is proportional to $N(t)$ [63]. Within seconds, the carrier concentration $N(t)$ in the Gaussian DOS can drop by more than eight orders of magnitude, which is impossible to achieve in a system with exponential DOS. Complementary to the concentration dependence of the carrier mobility as described in Section 6.3, the recombination dynamics therefore constitute another way to experimentally distinguish the shape of the DOS in disordered materials.

Chapter 7

Numerical Treatment of Charge Transport in Disordered Materials

For the sake of completeness, let us here briefly outline two different possibilities for studying hopping transport in disordered materials by means of computer simulations. The model framework introduced in Chapter 1, with the sites distributed according to the choice of the DOS $g(\varepsilon)$ and the transition rates given by the Miller-Abrahams hopping rates is simple enough to allow for the implementation of efficient and uncomplex computer algorithms, which is very useful to check the validity of theoretical approaches.

The vast majority of computer simulations of hopping transport are based on either Monte Carlo (MC) methods or on solving the master equation of the steady state charge flow through the system. In both cases, current is driven by an external electric field and periodic or helical boundary conditions are applied. The system is generated by randomly placing sites in space and energy, according to the shape of the DOS. Both algorithms therefore give results only for a specific configuration of localized states and an averaging procedure is advisable. The two methods are outlined in the following Sections 7.1 and 7.2.

Typical problems of all simulation approaches of hopping conduction are finite size effects. Available memory and computer time limit the amount of sites M in the simulated systems, which can influence the results for certain parameter choices. For instance, consider a system with $M = 10^6$ sites, a Gaussian DOS as in Equation (1.7), and a temperature of $kT = 0.2\sigma$. The equilibration energy ε_∞, where most carriers are situated in thermal equilibrium, then has a value of $\varepsilon_\infty = -\sigma^2/kT = -5\sigma$. One can easily check, that (on average) not a single one of the 10^6 sites would have an energy of ε_∞ or below, since ε_∞ is too deep in the tail of the DOS. Hence, finite size effects in this case will not allow any reasonable results to be obtained from the simulation. In the literature, this is unfortunately often not taken into account and one can find various studies, where results cannot be trusted for that reason [62, 64].

7.1 The Balance Equation (BE) Approach

The simplest and most efficient approach to simulate movement of a single charge carrier in a system of randomly placed sites with arbitrary energetic disorder $g(\varepsilon)$ is the balance equation (BE) approach. In this approach, the system is assumed to be in a steady state and transport properties are derived from the stationary occupation probabilities p_i of the sites i. The occupation probabilities p_i are found by solving the master equation, consisting of a set of BEs that relate the values of p_i and p_j of two sites i and j to the charge flow between those sites. Assuming, that in a steady state the total current into each site equals the total current out of that site, and that the occupation probabilities are independent from each other (corresponding to a mean-field

approximation), we can write down the following master equation [65, 66]:

$$\sum_{i \neq j} p_i \Gamma_{ij} \left(1 - p_j\right) = \sum_{i \neq j} p_j \Gamma_{ji} \left(1 - p_i\right) , \tag{7.1}$$

where $\sum_{i \neq j}$ runs over all pairs of sites. Each term in the sum corresponds to the directed flow of electrons between a pair of sites, weighted with the probability p_i for the initial site to be occupied and the target site to be empty, $(1 - p_j)$. Γ_{ij} is the MA hopping rate between sites i and j, including the electric field:

$$\Gamma_{ij} = \nu_0 \exp\left(-\frac{2d_{ij}}{\alpha}\right) \begin{cases} 1 & \text{if } \varepsilon_i > \varepsilon_j , \\ \left[\varepsilon_i - \varepsilon_j - F_z(z_i - z_j)\right]/kT & \text{otherwise} . \end{cases} \tag{7.2}$$

In Equation (7.2), F_z is the external electric field (in z-direction) and z_i are the z-coordinates of the sites. Equations (7.1) and (7.2) constitute a system of M equations for the occupation numbers p_i of the M sites. Once their values are found, the carrier mobility can be obtained via

$$\mu = \sum_{i \neq j} p_i \Gamma_{ij} \left(1 - p_j\right) \left(z_j - z_i\right) \Big/ \sum_i p_i . \tag{7.3}$$

Because of its non-linearity, the Equation System (7.1) is very difficult to solve. In the limit of low carrier concentration, one therefore assumes the target site of each hop to be always unoccupied, so that Equation (7.1) reduces to

$$\sum_{i \neq j} p_i \Gamma_{ij} = \sum_{i \neq j} p_j \Gamma_{ji} . \tag{7.4}$$

Equation (7.4) describes transport of a single charge carrier in the system of M sites. To model carrier concentrations that are *not* very low, the approximation is therefore invalid and one needs to solve the full non-linear Equation (7.1). Although that is possible and methods exist to reduce the complexity of that non-linear problem [65, 66], those methods are out of the scope of this work and we will here stick to the dilute case.

Equation (7.4) can be written as the following matrix equation of the form $\underline{M}\vec{p} = \vec{b}$, that is left to be solved for the occupation probabilities $\vec{p}$ [67]:

$$\begin{pmatrix} 1 & 1 & 1 & \cdots \\ \Gamma_{12} & -\Gamma_2 & \Gamma_{13} & \cdots \\ \Gamma_{13} & \Gamma_{23} & -\Gamma_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} , \tag{7.5}$$

where $\Gamma_i = \sum_j \Gamma_{ij}$ is the total outgoing rate of site i. Since most of the entries of the rates matrix in Equation (7.5) are very small (if pairs of sites are far apart or have large energy differences) such small elements can be approximated by 0 and the matrix becomes sparse. Solving Equation (7.5) is the computationally most demanding task of the BE approach. To do that, one can use any of the many linear algebra packages that are available for solving such simple (sparse) matrix equations, which makes the implementation of the BE method particularly simple.

Besides the availability of highly optimized software packages to do the heavy lifting of the BE approach, it has some more advantages over the MC algorithm: The assumption of the system being

in a steady state inherently ensures, that the resulting transport properties do not contain non-equilibrium effects. Furthermore, the algorithms for solving the matrix equation in Equation (7.5) are usually nicely parallelizable, which is why the BE method is typically faster than MC simulations.

Disadvantages of the BE method are difficulties in handling systems with finite charge carrier concentration and a worse scaling behaviour than that of MC simulations with system size. The memory footprint of the BE algorithm is typically larger than that of an efficiently implemented MC simulation.

7.2 MC Simulation of Hopping Transport

The MC simulation approach is probably the most straightforward and transparent numerical technique for studying hopping charge transport. Movement of carriers through the system of localized states is simulated as a sequence of elementary events, which are the single carrier transitions between a pair of sites. Application of MC algorithms to hopping problems started in the 1980's, and the first review article on that matter was published by Bässler in 1993 [20]. The principal method works as follows.

After preparation of the system, the probabilities p_{ij} for a hop from site i to site j are calculated for each pair of sites i and j as the hopping rate Γ_{ij} normalized by the total outgoing flow of site i:

$$p_{ij} = \tau_i \Gamma_{ij} \qquad \text{with} \qquad \tau_i^{-1} \equiv \sum_{k \neq i} \Gamma_{ik} \,. \tag{7.6}$$

Equation (7.6) fulfills the criterion $\sum_j p_{ij} = 1$. In order to reduce the required memory of the simulation, one can enable hopping only between sites that are separated by a distance lower than some cut-off radius r_c, that is typically chosen of the order of several intersite distances $N^{-1/3}$, e.g., $r_c = 3N^{-1/3}$, or by setting a lower boundary for the hopping rates Γ_{ij} to exclude hops that are very improbable.

One or more charge carriers are placed on random sites in the disordered system. Waiting times τ_i according to Equation (7.6) are then assigned to each carrier's site. The carrier with the shortest waiting time τ_i is then chosen for the next transition. A random number $p \in \,]0, 1]$ is drawn from a uniform probability distribution and the hopping target j is found via:

$$\sum_{k=1}^{j-1} p_{ik} < p \leq \sum_{k=1}^{j} p_{ik} \,. \tag{7.7}$$

After the target site j is determined (requiring that it is empty), the hop is executed, the carrier is assigned a new waiting time τ_j, and the system time is advanced by τ_i. The waiting time τ_i of the executed hop is then substracted from the waiting times of all other carriers, so that they are more likely to be chosen for the next transition. The process is repeated for the next hop.

Because charge carrier motion is modeled as a sequence of single hopping processes, one can find transport coefficients by simply collecting statistics of the carriers' trajectories through space and energy. Usually, the system is relaxed for some time prior to collecting such statistics, so that carriers can thermalize and the obtained results are those of the equilibrated system. Carrier mobility μ and diffusion coefficients parallel $D_\parallel$ and perpendicular $D_\perp$ to the field can be obtained via

$$\mu = \frac{\langle z \rangle}{F_z t_s} \,, \qquad D_\perp = \frac{\langle x^2 \rangle + \langle y^2 \rangle}{4 t_s} \,, \qquad D_\parallel = \frac{\langle z^2 \rangle - \langle z \rangle^2}{2 t_s} \,, \tag{7.8}$$

where t_s is the total time during which statistics were collected and x, y, z are the displacements in the respective directions of each charge carrier.

The MC algorithm described above can be optimized in many ways, which is, however, out of the scope of this work. Since the single steps of the simulation require very little work for the processor, the performance of the simulation can be pushed to millions of hops per second on a single processing core. It is also possible to parallelize certain parts of the algorithm, which can increase the performance noticeably.

Advantages of the MC method over the BE approach introduced in Section 7.1 are the ability to observe non-equilibrium effects, trace particle trajectories, and to easily simulate more than one charge carrier in the system. Furthermore, the memory footprint is usually lower and the scaling with system size is better, so that finite size effects can more easily be prevented.

Chapter 8

Conclusions

In Part I of this thesis, a number of theoretical tools for the description of charge transport in disordered media were described. Disorder in the atomic (or molecular) array of organic and inorganic semiconductors breaks the symmetry required for the description of electronic phenomena via reciprocal space. Charge carriers in such materials are not found in extended states, with their wave functions spread over large volumes of the samples, but predominantly occupy spatially localized states. Charge transport is therefore determined by tunneling transitions between localized states or by trapping and detrapping from bands of extended states into the localized density of states (DOS).

In the introduction of this Part I, the electronic structure of amorphous organic and inorganic semiconductors was described. It was discussed, where the localized states originate from and how they are described in the model framework used throughout this work. In inorganic disordered semiconductors, the conduction and valence bands have tails of localized states leaking into the bandgap, separated from the extended states by the mobility edges. The electronic structure of organic disordered semiconductors (ODSs), on the other hand, consists solely of the density of localized states, with no bands of extended states. Their spectra, the exponentially decreasing DOS in inorganic and the Gaussian DOS valid for organic disordered semiconductors, were introduced along with the corresponding characteristic equilibrium charge carrier distributions. The decisive differences between the exponential and Gaussian DOS were pointed out, with the transport properties in the latter DOS being independent of carrier concentration n (i.e., of the position of the Fermi level ε_F) while transport in the former DOS decisively depends on n. Finally, the different transport modes that occur in disordered media were introduced. It was discussed that charge carriers in inorganic disordered semiconductors move in the extended states, interrupted by trapping into the localized DOS, which is the multiple trapping (MT) transport model. In ODSs, extended states are absent and carriers move via incoherent tunneling transitions between the localized states. This transport mode is called variable range hopping (VRH) due to the strongly temperature dependent tunneling distances.

Within this model framework, various techniques were developed to theoretically describe long-range conduction in disordered semiconductors. The physically most transparent approach is the transport energy (TE) concept, introduced in Chapter 2. The concept is based on the observation, that hopping transport in exponential and Gaussian DOSs consists of activated hops to some particular energy level, the TE, that is independent of the Fermi level ε_F and is situated at much higher energies than the equilibrium position of the carriers. An analytical derivation of the TE was presented, along with a technique to determine the TE level in computer simulations. Furthermore, a method was introduced how knowledge of the TE can be used to predict the conductivity of hopping charge carriers in disordered media. In addition to the TE concept being a predictive theory on its own, the universality of the TE and the fact, that it usually lies far above the typical carrier energies, make it possible to treat the TE as a mobility edge and to apply the MT model even in systems without extended states.

The second important approach to describe conduction in disordered media is the percolation theory, which was studied in Chapter 3. Because the rates of hopping transitions between localized states depend exponentially on tunneling distances and energy differences, carriers use only the fastest path in the random network. Each connected path therefore contains critical hops, that are the slowest connections and therefore have the largest influence on transport properties. Percolation theory approaches aim to identify such critical hops and thereby infer long-range characteristics of VRH. To introduce the concept of percolation theory within hopping conduction, first the purely spatial problem of transport through randomly distributed localized states with equal energies was described. Then, it was shown that percolation theory can be used to exactly describe VRH in systems with exponential DOS, by mapping the 3+1D (space + energy) problem onto a purely spatial 3D problem of spheres with randomly distributed radii, which is solvable numerically with high precision. Finally, the results were compared to the TE approach and it is shown, that both methods agree well.

Very recently, a novel method for the description of charge transport in disordered media had been suggested, which was introduced in Chapter 4. Interpreting the system of localized states as a resistor network, one can find those connections between pairs of states, that transport the highest electrical power during conduction, i.e., dissipate the largest amount of Joule heat. By studying the dependence of the dissipated heat in the network on energy or distance of the connections, one can find the TE and the most decisive hopping radii in the framework of this approach. It was shown, that the results obtained by the method of dissipated heat agree well with the TE concept.

In Chapter 5, it is shown how the MT transport model can be exploited to describe electron and hole transport in disordered inorganic media at high external electric fields. The theories were developed on the basis of experimental data of the dependences of electron and hole mobility on temperature and field measured in lead oxide (PbO). It was shown, that hole transport can be explained by a simple MT model, assuming that temperature is higher than the width of the localized DOS. For the strong dependence of electron transport coefficients on field and temperature, a simple phenomenological approach and a sophisticated quantum mechanical approach are suggested. Both approaches are capable of reproducing the experimentally measured phenomena.

So far, the introduced techniques dealt with transport of excess charge carriers in the array of randomly placed localized states. In Chapter 6, it was shown how bimolecular recombination during MT transport can be described theoretically. Free carriers of one species (holes or electrons) in the extended states either recombine with trapped carriers of the opposite species or are trapped into the localized DOS. These competing mechanisms, trapping and recombination, result in characteristic dynamics of the carrier concentrations $N(t)$ that are significantly different for exponential and Gaussian DOS functions. The results were compared to existing results from the literature and it was shown, how the recombination model can be used to distinguish between exponential and Gaussian DOS in experiments.

In Chapter 7, numerical recipes for the treatment of hopping charge transport were briefly discussed. The two most popular algorithms, straight forward Monte Carlo (MC) simulations and the balance equation (BE) approach were described and their advantages and disadvantages were pointed out. The efficiency of the numerical methods available to simulate VRH makes them an important tool for the verification of analytical models.

Part II

Computer Simulation of Atom Kinetics in Semiconductor Heterostructures

Chapter 9

Introduction

The development of integrated circuits based on semiconductors was probably the most important technical advance of the past 60 years, laying the foundation for the digitalized world we live in today [1]. Shortly after the first silicon (Si) based integrated circuits appeared, the complementary metal–oxide–semiconductor (CMOS) technology led to the development of the Intel 4004, the first central processing unit (CPU) and the ancestor of all of today's Si based computer chips [2]. Until today, performance of microelectronic devices is increased simply by decreasing the size of the transistors, the basic active component of computer chips: The famous Moore's law [68] states, that the semiconductor industry manages to double the density of transistors of microprocessors each year, which is, as of today, still accomplished. Quantum effects, however, pose a fundamental lower limit for the size of the transistors, and will soon make it impossible to further improve the chips' performance by simple downscaling of device dimensions. Research on semiconductors in the past decade therefore focused on finding alternative strategies to increase device performance, for instance, by integration of new high-κ materials to reduce tunneling leakage currents, or by functionalizing Si based devices via integration of materials with interesting optoelectronic properties. Both of these approaches necessitate the development and understanding of novel compound semiconductors, that are suitable for monolithic integration with Si.

III/V semiconductors, combinations of elements from the third and fifth group of the periodic table of elements, are promising candidates for the integration of optoelectronic properties with Si. While Si itself is an indirect semiconductor, and therefore not optically excitable, many III/V semiconductor compounds have direct bandgaps that can often be tuned by changing the elementary composition. Figure 9.1 shows a map of the bandgap energy versus the lattice constant of various binary III/V alloys. The lines between the materials correspond to the gap energy and lattice constant of the mixture with different composition. For instance, the line between gallium arsenide (GaAs) and indium arsenide (InAs) shows the properties of $Ga_xIn_{1-x}As$. Due to such mixing behaviour, the electronic and structural properties of III/V alloys can be taylored by combining multiple elements with the right composition. Since the late 1990's, the complexity of III/V materials therefore grew with intensive research on ternary and quaternary alloys (mixtures of three and four different elements), and on metastable materials, e.g., dilute nitrides.

The fabrication of III/V semiconductors integrated monolithically with Si is a complex process for which various growth techniques were developed, for instance, atomic layer deposition (ALD), molecular beam epitaxy (MBE), and liquid phase epitaxy (LPE). The industry standard for growing such materials, however, is metalorganic vapour phase epitaxy (MOVPE), which allows non-equilibrium growth needed for the deposition of metastable materials. The increasing complexity of the III/V materials requires more and more detailed understanding of the chemical and physical processes both during growth and during post-processing of the devices, e.g., by high temperature annealing. Theoretical modeling of such procedures, especially that involving non-equilibrium techniques, is, however, still in its infancy.

In this Part II of the work, we show how Kinetic Monte Carlo (KMC) computer simulations can
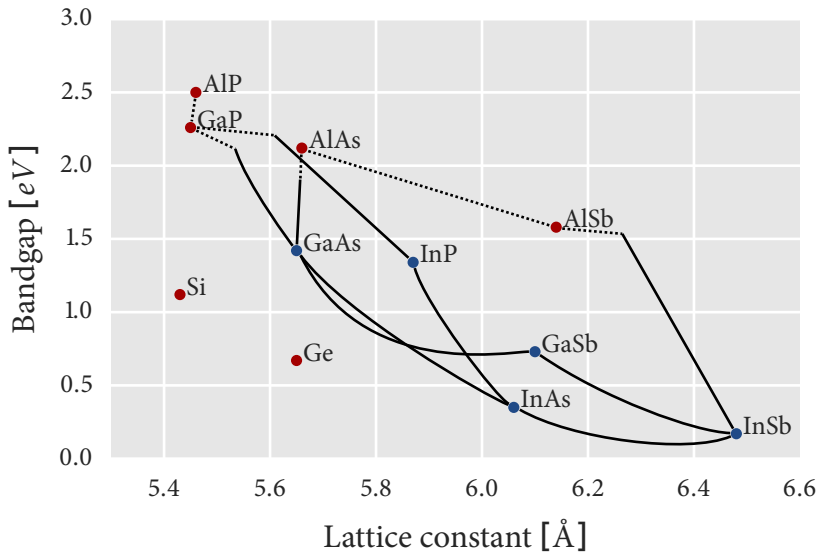
Figure 9.1: The bandgap energy versus lattice constant of various III/V semiconductors. Red points indicate indirect bandgaps, blue points are direct bandgaps. The lines between the materials represent the properties of mixtures of different composition, with dotted lines having indirect and solid lines having direct bandgaps. Data from Reference [69].

be used to model non-equilibrium epitaxial growth of semiconductor compounds. In the following Section 9.1, the lattice structure and material properties of III/V semiconductors is discussed and Section 9.2 contains a brief overview of the MOVPE growth technique with the different kinetic processes involved. A brief historical overview of computer simulations of semiconductor growth is given in Section 9.3.

In Chapter 10, the KMC method is introduced and discussed, which is then used to model intermixing effects of GaP growth on Si in Chapter 11 and etching effects of Ga droplets on Si substrates in Chapter 12. A conclusion and an outlook for Part II is given Chapter 13.

9.1 III/V Semiconductors on Silicon

III/V semiconductors are alloys which consist of elements from the third and fifth main group of the periodic table. From the former, the most prominently used elements are aluminum (Al), Ga, and indium (In), and from the latter, most materials are based on either phosphorus (P), arsenic (As), or antimony (Sb). All binary III/V combinations of those elements can be grown under suitable conditions, and they form the basis for ternary and quaternary III/V alloys. Their direct or indirect bandgaps and lattice constants are shown in Figure 9.1. The alloys crystallize in the zincblende crystal strcuture, have semiconducting properties (i.e., a bandgap of 0.1 eV to 2.5 eV), and are solid at room temperature. In the past 20 years, ternary or even quaternary III/V alloys were grown, containing mixtures of the abovementioned elements and often additional dilute concentrations of nitrogen (N), boron (B), or bismuth (Bi), for instance, Ga(N,As,P) / (B,Ga)P quantum well structures [70].

The perfect zincblende lattice structure, shown in Figure 9.2, consists of two interpenetrating face-centered cubic lattices, with the group III and group V elements each occupying their own sublattice. The unit cell contains eight atoms in total. The zincblende structure is tetrahedrally coordinated, i.e., each atom has four nearest neighbors of the opposite element, and is therefore similar to the diamond cubic lattice structure of Si and germanium (Ge). This similarity in lattice structure makes integration of thin layers of III/V materials with Si possible. However, usually
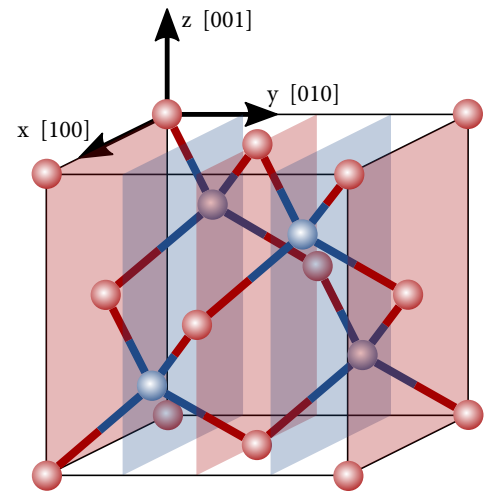
Figure 9.2: The unit cell of the zincblende lattice structure. The two atomic species are drawn as red and blue spheres. The two fcc sublattices are indicated by the colored planes. Some atoms of the red species of neighboring unit cells are also shown.

there is a mismatch between the lattice constants of Si and most III/V alloys, that significantly influences the growth procedures. The smallest lattice mismatch of the binary alloys to Si is found in GaP, amounting to less than 0.5 %, what makes it the ideal material for buffer layers between Si and more complex III/V alloys [9].

An important aspect of III/V semiconductors on Si is the polarity of the interface. As a *polar* III/V alloy, for instance GaP, consists of elements with different numbers of valence electrons, the abrupt interface between this material and a *non-polar* substrate, e.g., Si, will be charged postively or negatively, depending on the polarity of the alloy. This can be understood by studying the (110) projection of the crystal, as shown in Figure 9.3. The two possible polarities of the GaP-Si interface are shown, the positively charged Ga-polar (Figure 9.3 (a)) and the negatively charged P-polar (Figure 9.3 (b)) interface. Such charged interfaces are energetically unfavorable, which results in a driving force for interface intermixing within one or two layers in order to reduce the electrostatic forces [71].

(a) Ga polar interface.

(b) P polar interface.

Figure 9.3: The different polarities of a III/V semiconductor with an abrupt interface to a non-polar material, projected onto the (110) plane. Here, as an example, GaP on Si is shown, with Si atoms in red color, Ga atoms in blue color, and P atoms colored green. If the first layer above the Si substrate consists of Ga atoms, the interface is called Ga-polar (a), in the opposite case it is P-polar (b).

Figure 9.4: Schematic drawing of the MOVPE growth processes. The carrier gas picks up precursor molecules in the bubblers, which are then t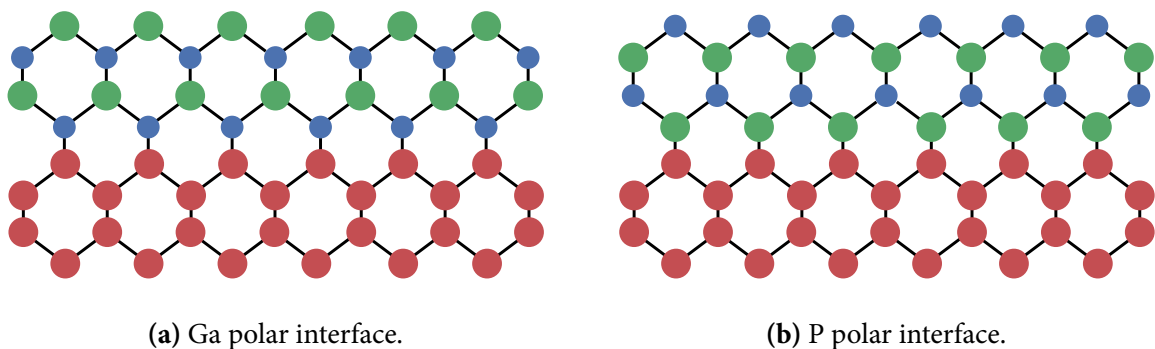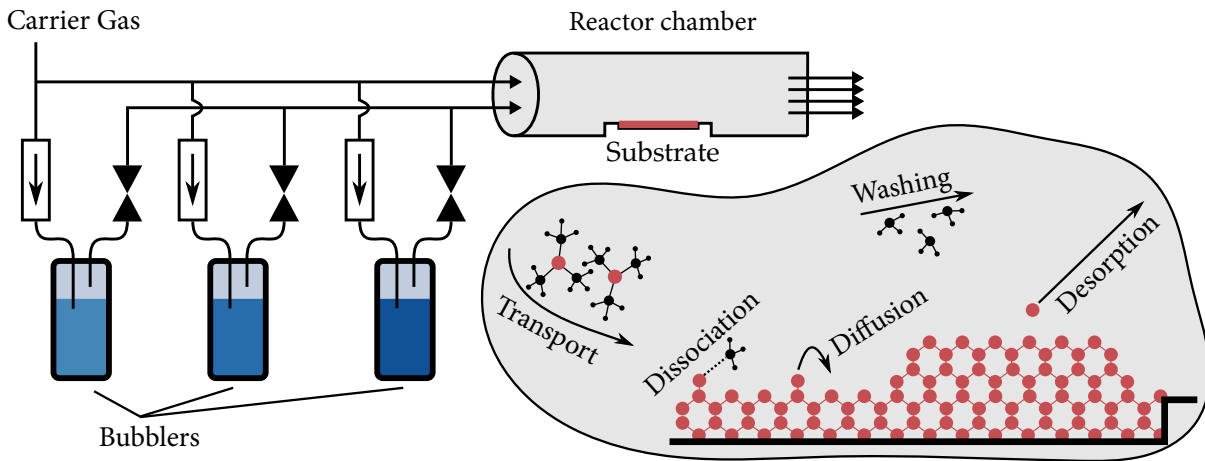ransported to the heated substrate. There, they dissociate and the cargo element binds to the surface. Via atomic kinetics on the surface the crystal arranges itself. The remains of the precursor molecules are washed away from the surface.

9.2 Metalorganic Vapour Phase Epitaxy (MOVPE)

Metalorganic vapour phase epitaxy (MOVPE), sometimes also referred to as organometallic vapour phase epitaxy (OMVPE) or metalorganic chemical vapour deposition (MOCVD), is a non-equilibrium method to grow crystalline layers by chemical reactions of the organometallic precursors with the substrate. The MOVPE process is schematically shown in Figure 9.4. A carrier gas, typically hydrogen, flows through the bubblers, which contain the metalorganic precursors in (usually) liquid form. The carrier gas is saturated with the precursor molecules, which are then transported to the reaction chamber. By adjusting the parameters of this process, the partial pressure of the different precursors can be controled. The gases then impinge on the heated substrate, for example, a Si wafer, where surface and gas-phase reactions lead to dissociation of the precursor molecules. During these reactions the desired atom to be incorporated into the crystal is detached from its carrier molecule and adsorbed on the substrate. Finally, in the ideal case the remaining molecules are purged from the surface. The adsorbed atoms left on the substrate then diffuse on the surface until a stable configuration is found, which, under suitable growth conditions, leads to single crystalline layer growth.

MOVPE is a highly complex process with many adjustable parameters, e.g., the (relative) partial pressures of the precursors, growth temperature and time, precursor choice, and many more. The atomistic processes on the substrate also decisively influence the quality of the grown structures: Many defect types are direct consequences of the adsorbate kinetics, for example, antiphase boundaries, vacancies, step or island formation, antisites, or phase separations. The ability to theoretically model the kinetic aspects of epitaxial growth of semiconductors is therefore important to understand and avoid such effects, and improve device quality.

9.3 Computational Modeling of Semiconductor Growth

There are certain basic requirements for the theoretical modeling of epitaxial growth, that already rule out many of the established computational methods in the field of condensed matter. Growth procedures typically take long times (up to several minutes) whereas the fastest involved atomistic processes happen at the timescale of picoseconds. These scales differ by many orders of magnitude, which makes it impossible to exactly model the particles' dynamics and rules out most molecular dynamics (MD) techniques as candidate methods. Furthermore, the interesting structural aspects of semiconductor growth typically involve many atoms and can easily become as large as tens of nanometres in diameter. The required supercell sizes of the simulated systems are therefore by far too large for traditional ab-*initio* density functional theory (DFT) calculations. On top of the requirements regarding size and time, the epitaxy of metastable materials, to which most of the III/V semiconductors belong, is a non-equilibrium procedure, which renders the application of mesoscopic continuum methods inappropriate. Currently, the only simulation technique to fulfill all of those requirements is Kinetic Monte Carlo (KMC). In KMC, the physical processes are reduced to a limited set of well-defined discrete events, that are described only by their rate constants and are executed one by one. The details of the KMC method are given in Chapter 10.

The first attempts to simulate epitaxial growth by means of KMC simulations were carried out in the early 1970's, by Abraham and White [72] and Gilmer and Bennema [73]. In both studies, the authors simulated vapour phase epitaxy (VPE). A few years later, Bortz et al. [74] suggested a new MC mechanism, in which all tried events are actually carried out. This so-called BKL algorithm allowed for much more efficient and thereby larger and longer simulations. As of today, the BKL algorithm is the standard method for KMC simulations and is almost exclusively used [10, 11].

During the first years, researchers were mainly interested in the surface structures and in growth modes, i.e., 2D Frank-van-der-Merwe or 3D Volmer-Weber growth [75–77]. The first studies to simulate growth of III/V materials, for instance MBE growth of GaAs [78], appeared in the early 1990's. From the interest in general growth properties, the focus shifted towards growth of nanostructures in the past decades, e.g., quantum dots or nanowires [79, 80], which nowadays is the main interest of KMC structure simulations.

Today, many epitaxial techniques utilize the lattice mismatch between the grown material and the substrate for the formation of nanostructures. A lot of effort has therefore been made to accurately account for strained crystal lattices in the KMC codes, starting with the basic ball-and-spring model by Orr et al. [77]. The last advance in this respect was the application of an effective bond order potential, the Keating potential [81], in KMC simulations of strained structures [80].

Further improvements of the traditional KMC techniques are the parallel simulation of solid and liquid phases in a lattice-KMC model [82], as well as the combination of KMC with more accurate simulation methods, e.g., DFT [83] or MD [84]. Especially in the field of surface chemistry, KMC simulations with rate constants precalculated by DFT are often used [85, 86].

The basic KMC algorithm for the computer simulation of heteroepitaxy is described in detail in the next Chapter 10.

Chapter 10

KMC Computer Simulation of Atom Kinetics

In this chapter, the algorithm for the simulation of atom kinetics of semiconductor growth is described. It is based on a KMC method, where discrete elementary atomistic events are simulated sequentially. The rate constants of the different kinetic processes are determined phenomenologically by the *bond counting* [10, 87] method, that previously proved successful for accurate KMC simulation of epitaxial growth [82, 88, 89]. An overview of the history of growth simulations with KMC methods is given in Section 9.3.

10.1 Preparation of the System

The first step of the simulation procedure is to prepare the simulated system, also called *supercell*. All material systems we are interested in crystallize in the zincblende lattice structure. In addition, as explained in more detail in the following Section 10.4, the elementary events we take into account depend only on the lattice sites of the simulated crystal. Hence, we choose our system geometry to correspond to the zincblende lattice structure, with the lattice sites positioned at the equilibrium atom positions. While this choice significantly increases the performance of the simulation and, thereby, allows us to simulate the required time ranges and large supercells, it is a rough approximation of the atom kinetics that may be insufficient in certain situations.

Let the cartesian coordinates x, y, and z correspond to the crystallographic directions [100], [010], and [001], respectively. Throughout this work, periodic boundary conditions are applied in x and y direction. The bottom two monolayers (MLs) of the system, i.e., those with lowest z-coordinate, are at all times occupied with atoms that are fixed during the simulation. This ensures a substrate, where adatoms (atoms adsorbed on the substrate) have all nearest and second-nearest neighboring positions below occupied. The atomic species of the two substrate MLs is an input parameter of the simulation and depends on the simulated problem; however, in this work we always start with Si substrates. The fixed substrate and the crystallographic directions are shown in Figure 10.1, with the higher atoms (those with larger z coordinate) colored brighter.

In principal, the system can be prepopulated with the desired starting configuration of the material. If, for example, one wishes to simulate vacancy diffusion in a grown material, one could populate the system in advance with only some few lattice positions left unoccupied, and then start the simulation. This allows the efficient simulation of dynamics and/or stability of surface or interface structures.

10.2 The Elementary Events

There are three types of events present in the simulation, that can each be enabled or disabled prior to and during the simulation.
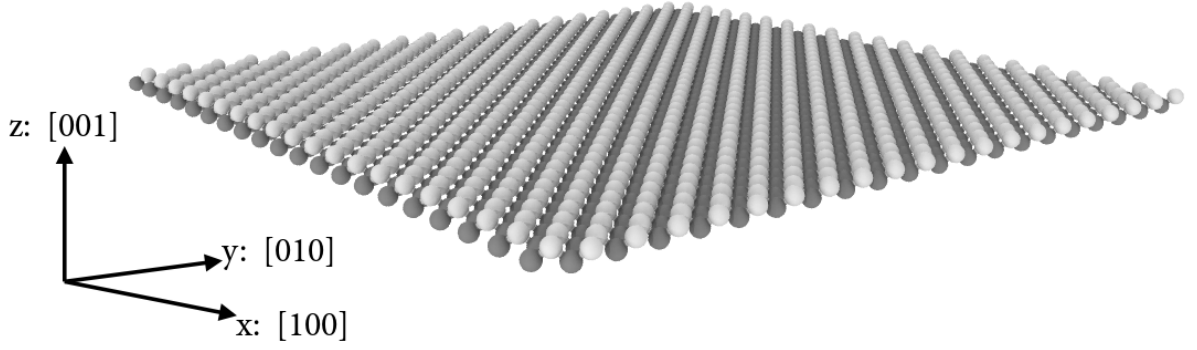
Figure 10.1: The two fixed substrate layers of the zincblende simulation lattice with the choice of the coordinate system. The atoms of the upper layer are displayed in brighter color. All atoms are of the same species. Figure created with Ovito [90].

Adsorption   In an adsorption event, a new atom arrives at a random position on the substrate's surface. This is implemented in the following way: A random lattice site i of the topmost layer of the system, i.e., that with the highest z coordinate, is chosen randomly. If none of the two nearest neighbors below the site i is occupied with an atom, one of them is randomly chosen as the new site i. This process is repeated until a site is found that has at least one occupied nearest neighbor below. This implementation ensures, that new adsorbates are always connected to the substrate. The rate v^A of the adsorption events is chosen such that the desired growth rate is reproduced. It therefore depends on system size and is calculated as

$$v^A = \Gamma n_{\mathrm{ML}},\tag{10.1}$$

where n_{ML} is the number of lattice sites in a single monolayer and Γ is the desired growth rate per unit time. If, for instance, we are simulating a system of $N_x \times N_y \times N_z$ unit cells (the cells as shown in Figure 9.2) and want to grow at a rate of $\Gamma = 2\,\mathrm{ML}/s$, the total rate for the adsorption event is $v^A = 4N_x N_y/s$, since there are two lattice positions per unit cell in each ML.

Desorption   Desorption from the surface of the simulated supercell works simply by deleting the corresponding atom from its occupied site, given it is not the only connection of any other atom to the substrate. Desorption is a thermally activated process and happens with a rate

$$v_i^D = v_0 \exp\left(-\frac{\varepsilon_i(X_i) + \varepsilon_d}{kT}\right),\tag{10.2}$$

where v_0 is the attempt-to-escape frequency, usually estimated as the atomic vibrational frequency (or Debye frequency) of the crystal, $v_0 \approx 10^{13}\,\mathrm{s}^{-1}$, and kT is thermal energy. $\varepsilon_i(X_i)$ is the energy barrier for diffusive events from site i of the atomic species X_i (see next Section 10.3) and ε_d is an additional desorption energy barrier, which is an input parameter of the simulation. Because desorption is a rare event as compared to diffusion, the barrier ε_d is usually chosen high enough to prevent rapid evaporation of the material.

Diffusion   In a diffusive event, an atom hops from one lattice site i to a site j within the second coordination sphere, i.e., to a nearest or second-nearest neighbor. This is allowed only if the target

site j is not already occupied and, as for the desorption event, if this process leaves no other atoms disconnected from the substrate. Such events are thermally activated and their rates are given by

$$v^H_{i \to j} = v_0 \exp\left(-\frac{\varepsilon_i(X_i) + \Delta\varepsilon_{ij}}{kT}\right), \tag{10.3}$$

where $\varepsilon_i(X_i)$ is the same energy barrier for diffusion out of site i as in the desorption rates, and $\Delta\varepsilon_{ij}$ contains additional energy offsets to account for anisotropic effects, e.g., Ehrlich-Schwoebel barriers [91, 92] or dimerization on surfaces. Throughout our simulations, such effects are neglected and $\Delta\varepsilon_{ij} = 0$.

10.3 The Bond Counting Method

In the previous section, the microscopic events, that are possible in the simulation, were introduced. So far, we know that their rates (Equations (10.2) and (10.3)) are simple Arrhenius equations, that depend on the energy barriers $\varepsilon_i(X_i)$. These energy barriers are generally difficult to calculate: Depending on the required accuracy, one might need to employ ab-*initio* methods, for instance DFT calculations, or bond order potentials, e.g., the Tersoff or second-moment tight-binding potential. For KMC simulations of crystal growth, however, the large number of events in the simulation and the requirement for big supercells render such accurate methods unfeasible. Instead, we assume that the energy barrier for an atom to escape its lattice position is proportional to the number of atoms in its surrounding. This approximation is sometimes called *bond counting KMC* [87].

In bond counting KMC, the barriers $\varepsilon_i(X_i)$ are calculated as

$$\varepsilon_i(X_i) = \sum_{j \in \text{NN}} \Delta\varepsilon_{NN}(X_i, X_j) + \sum_{j \in 2\text{NN}} \Delta\varepsilon_{2NN}(X_i, X_j), \tag{10.4}$$

where the sums run over the occupied nearest (NN) and second-nearest ($2NN$) neighbors, respectively, and X_i and X_j are the atomic species of the particles at the sites i and j. Empty lattice positions do not contribute to the sum. The energies $\Delta\varepsilon_{NN}(X_i, X_j)$ and $\Delta\varepsilon_{2NN}(X_i, X_j)$ are the energy contributions of NN or $2NN$ bonds between atoms of species X_i and X_j. Hence, the necessary input of the bond counting method is a table that contains the contributions of all possible atomic combinations of nearest and second-nearest neighbor bonds to the total energy barrier.

Bond counting KMC is a rough approximation of the binding energies in solids, that will surely not give accurate absolute numbers and may fail completely in some situations. However, it is extremely efficient to compute and often sufficiently accurate to reproduce kinetic aspects of epitaxy. In Chapters 11 and 12 we present two example studies, where bond counting KMC was successfully applied to model experimental results.

10.4 The Simulation Algorithm

The simulation is generally split into different growth phases, that are executed sequentially. A growth phase is defined by a set of atomic species to grow, the growth rates of these elements, and the total duration of the phase. Depending on the growth phase, adsorption events with
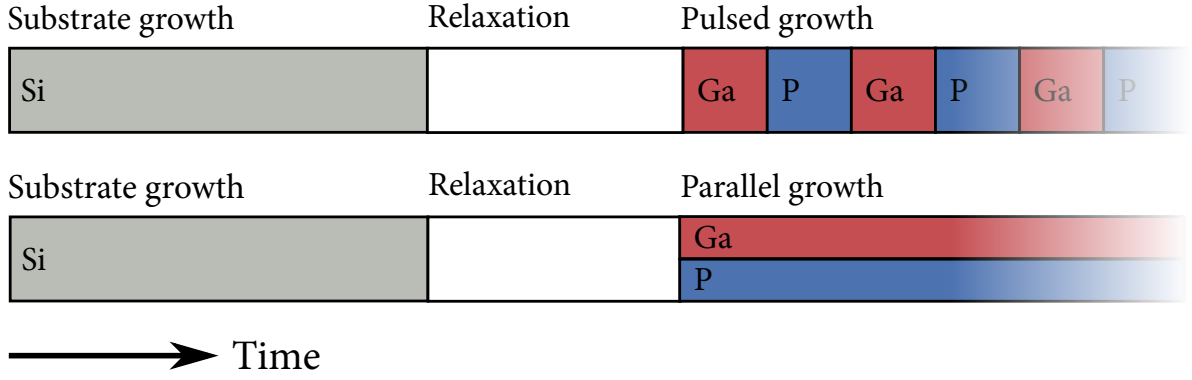
Figure 10.2: Two typical growth phase configurations of a Si substrate and growth of GaP, once in pulsed mode and once in parallel.

corresponding growth rates for the different atomic species are enabled or disabled. Two examples are shown in Figure 10.2. Initially, Si is grown to generate the substrate. Then, an empty relaxation phase is executed in which no adsorption events are enabled and the system relaxes by diffusion and desorption. Afterwards, Ga and P are grown layer by layer in pulsed mode or continuously in parallel mode.

The KMC algorithm during the growth phases works as follows.

1. A list of all possible microscopic events is generated. Each event i is associated with the probability p_i for it to occur. When v_i is the physical rate of event i, the probability p_i is calculated via

$$p_i = \frac{v_i}{\sum_j v_j},$$
(10.5)

   where the sum $\sum_j$ runs over all possible events. The rates for diffusion and desorption, as described in Section 10.2, are calculated by iterating over the nearest and second-nearest neighbors of the respective atoms, and the adsorption rates are included depending on the currently active growth phase.

2. A random number $r \in]0,1]$ is drawn from the continuous uniform distribution $\mathcal{U}(0,1)$. An event k is then determined via

$$\sum_{j=1}^{k-1} p_j < r \le \sum_{j=1}^{k} p_j$$
(10.6)

   and is executed. Hence, the probability for an event to be chosen is proportional to its physical rate constant, as calculated via Equation (10.5).

3. The total simulation time is updated by

$$\Delta t = -\frac{\ln(s)}{\sum_j v_j},$$
(10.7)

   where $\sum_j v_j$ is the cumulative rate of all possible events and s is a random number drawn from $\mathcal{U}(0,1)$.

4. The simulation is stopped if the termination criterion is reached (i.e., the total simulation time exceeds the configured value) or continues by going to step 1.

The most time consuming steps are the updating of the list of events along with their rates in step 1, and the choice of the next event to be executed, step 2. In what follows, we briefly discuss possible optimizations of these two steps, making them independent on system size and therefore largely reducing the required computer time.

Updating the list of events   After one of the events in Section 10.2 is executed, the list of events and event rates becomes invalid and needs updating. Rather than recreating the whole list, it is beneficial to update only those events that change, or need to be added or removed. Because all of the events depend at most on atoms within the second coordination sphere, i.e., on nearest and second-nearest neighbors, this affects only a small number of events. For instance, if a diffusive hop from lattice site i to lattice site j is executed, the following events change. There are new hops possible *to* the now available site i, while hops *to* site j are now forbidden, because it is occupied. Similarly, hops *from* site i are removed and hops *from* site j have to be added. In addition, the diffusion and desorption rates of most of the nearest and second-nearest neighbors of sites i and j change due to the configuration change. Updating of the events and event rates in this way is a local process and therefore independent on the size of the system.

Sampling from the list of events   The choice of the next event, in the form given by Equation (10.6), requires iterating over the list of all possible events, the number of which increases with system size. This iteration can be improved by imposing a two-dimensional storage scheme for the possible events. In bond counting KMC, the event rates given in Equations (10.1) to (10.3) can only take a finite number of discrete values, since the number of possible configurations of the neighbors of an atom is limited. Hence, many events that are possible at the same time will share the same event rate. Let $\mathcal{N}_i$ be the number of events with the rate v_i. The probability for an event with rate v_i to happen is then given by

$$p_i = \frac{v_i \mathcal{N}_i}{\sum\limits_j v_j \mathcal{N}_j} \, , \tag{10.8}$$

with the number of different v_i (and, consequently, p_i) independent on system size. Each p_i corresponds to a list of events that share the same event rates and therefore are equally probable to be executed. Hence, one can choose an event *rate v_k* to happen according to Equation (10.6) and then simply execute a random one of the $\mathcal{N}_k$ events that share the rate v_k. This second random sampling is constant in time and therefore independent on the number of possible events. This storage scheme, albeit a bit more complex than the trivial one, makes the random sampling of the next microscopic event completely independent on system size.

With handling the list of possible events in the way described above, the only quantity in the simulation that depends on the size of the system is the time step Δt given in Equation (10.7). The larger the system, the more atoms (and lattice sites) are present and the more events are possible at any time during the simulation. Hence, the time step Δt shrinks linearly with the number of lattice sites and the simulation takes proportionally longer to finish. In total, the complexity of the

simulation regarding system size is therefore $\mathcal{O}(N_x \times N_y \times N_z)$, when N_x, N_y, N_z are the number of unit cells in direction x, y, z, respectively. This flat scaling allows efficient simulation of millions of lattice sites and atoms within hours on a common desktop computer.

Chapter 11

Interface Intermixing of GaP on Si

In this chapter, we present our findings concerning unusually strong intermixing effects at the interface between GaP and Si.

Due to the similarity in lattice constants between Si and GaP, the latter is a perfect candidate for buffer layers between the Si substrate and optically active alloys, e.g., Ga(N,As,P) quantum well structures [70]. As device dimensions are decreased to increase the density of components integrated on chips, the interface morphology between the different materials becomes increasingly important. Recently, it was observed that GaP grown by MOVPE on atomically flat Si-(001) substrates induces strong intermixing of the materials at the GaP/Si interface. The intermixing structures appear as pyramidally shaped Si clusters within a layer thickness of more than seven ML.

We were able to reproduce these effects with KMC computer simulations of the growth procedure and identify kinetic aspects of the growth as the main driving force of intermixing. The energy barriers necessary for the interdiffusion and formation of the structured interface have been confirmed in DFT calculations. Our findings suggest that such intermixing behaviour is a general effect of the growth of heterostructures on Si.

In Section 11.1, the experimental results are summarized. The KMC computer simulation results are presented in Section 11.2 and the DFT results for the hopping energy barriers are discussed in Section 11.3.

11.1 Experimental Results

The GaP layers were grown onto the Si-(001) substrate by MOVPE in an Aixtron AIX 200 GFR reactor. Prior to the growth, the substrate was atomically flat, as verified by scanning tunneling microscopy (STM) measurements. Ga and P were grown in pulsed mode at an overall growth rate of $1\,\mathrm{ML\,s^{-1}}$. For Ga, the precursor triethylgallium (TEGa) was used, and P was grown using tertiarybutylphosphine (TBP). The growth temperature was $T = 723\,\mathrm{K}$.

After growth, the samples were treated with high energy argon ion milling to create thin foils, that are transparent for electrons and can therefore be studied in the scanning transmission electron microscope (STEM). With that technique, foils with a thickness of 8 nm and 20 nm (from the same sample) were prepared, with the viewing direction along the [110] crystal axis. The samples were analyzed in a JEOL 2200 FS microscope, using the high angle annular dark field (HAADF)-STEM technique. In STEM, a beam of electrons is focused onto the sample and the scattered intensity of the transmitted electrons is detected. The focus point is then rastered over the sample surface to obtain a spatially resolved picture. When viewing along a crystal axis, the detected intensity is an average of the atoms in the atomic column along that axis. The intensity of the signal is roughly proportional to the square of the proton number Z^2 of the atoms in that column, which is why elements with higher proton numbers appear brighter in the measurements.

Table 11.1: The energy barrier contributions for the bond counting KMC method of the GaP/Si interface simulation. Each pair of atomic species is assigned a nearest (NN) and second-nearest (2NN) value.

	Si		Ga		P	
	NN	2NN	NN	2NN	NN	2NN
Si	0.3 eV	0.15 eV	0.3 eV	0.15 eV	0.6 eV	0.2 eV
Ga	0.3 eV	0.15 eV	0.3 eV	0.15 eV	0.3 eV	0.15 eV
P	0.6 eV	0.2 eV	0.3 eV	0.15 eV	0.3 eV	0.15 eV

The obtained measurements are shown in Figure 11.1: Subfigure (a) is a picture measured in a 20 nm thick region and (b) was measured at a thin point with only 8 nm thickness. The darker spots in the lower halfs of the pictures are the Si substrate, and the dumbbells with one brighter and one darker spot in the upper halfs of the pictures are Ga (bright spots) and P (dark spots). Clearly, the interface is not abrupt but exhibits strong intermixing between the substrate and the GaP layers. In both regions, the intermixing is not uniform but appears as pyramidally shaped structures that seem to be limited by the {111} crystal planes. This effect is more clearly visible in the thin sample, subfigure (b), since less material was averaged over.

The raw STEM data can be back-simulated using an absorptive potential approximation [93] method. With that procedure, the Si content of each atomic column can be determined. The resulting density maps, corresponding to the raw experimental pictures, are shown in Figure 11.1 (c) and (d), where the red circles correspond to 0 % Si content and the blue circles indicate 100 % Si content. It is clearly visible, that intermixing of about 7 ML to 8 ML is present at the interface in both the thick (c) and thin (d) sample region.

11.2 KMC Computer Simulation

The growth procedure was simulated using the KMC technique described in Chapter 10. The simulated system contained $55 \times 55 \times 60$ unit cells, i.e., in total 1 452 000 lattice sites. The desorption event described in Section 10.2 was disabled at all times during the simulation. The growth rates of all three atomic species were fixed at 1 ML/s. First, 25 ML of Si were grown onto the two fixed substrate layers. The resulting Si substrate was atomically flat, with only some adatoms and small islands on the surface. Then, P and Ga were provided layer by layer in pulsed growth mode (see Figure 10.2), with a relaxation pause of 1 s after each phase, until a total of 20 ML of GaP were grown. The growth temperature was chosen as $T = 775$ K, which is slightly higher than the experimental value. However, since temperature simply scales the energies in Equation (10.3), the difference has no influence on the results. The energy barrier contributions for the bond counting KMC scheme are listed in Table 11.1. All nearest neighbor bonds contribute with 0.3 eV and all second-nearest neighbor bonds add 0.15 eV to the total barrier. The only exception are the Si-P bonds, which contribute with 0.6 eV and 0.2 eV. P therefore is much more tightly bound to Si than all other atomic species.

The simulated Si substrate prior to the arrival of any Ga or P atoms is shown in Figure 11.2 (a). Clearly, similar to the substrate used in the experiments, the surface is atomically flat with only single atoms and atomic islands in the first ML on the surface. After deposition of Ga and P, the surface morphology of the Si substrate significantly changed. The substrate after the simulation is shown in Figure 11.2 (b), with all Ga and P atoms being hidden. The figure shows, that during

(a) STEM picture thick sample

(b) STEM picture thin sample

(c) Si content thick sample

(d) Si content thin sample

(e) Simulation results thick sample

(f) Simulation results thin sample

Relative Si content: 0.1 0.5 0.9
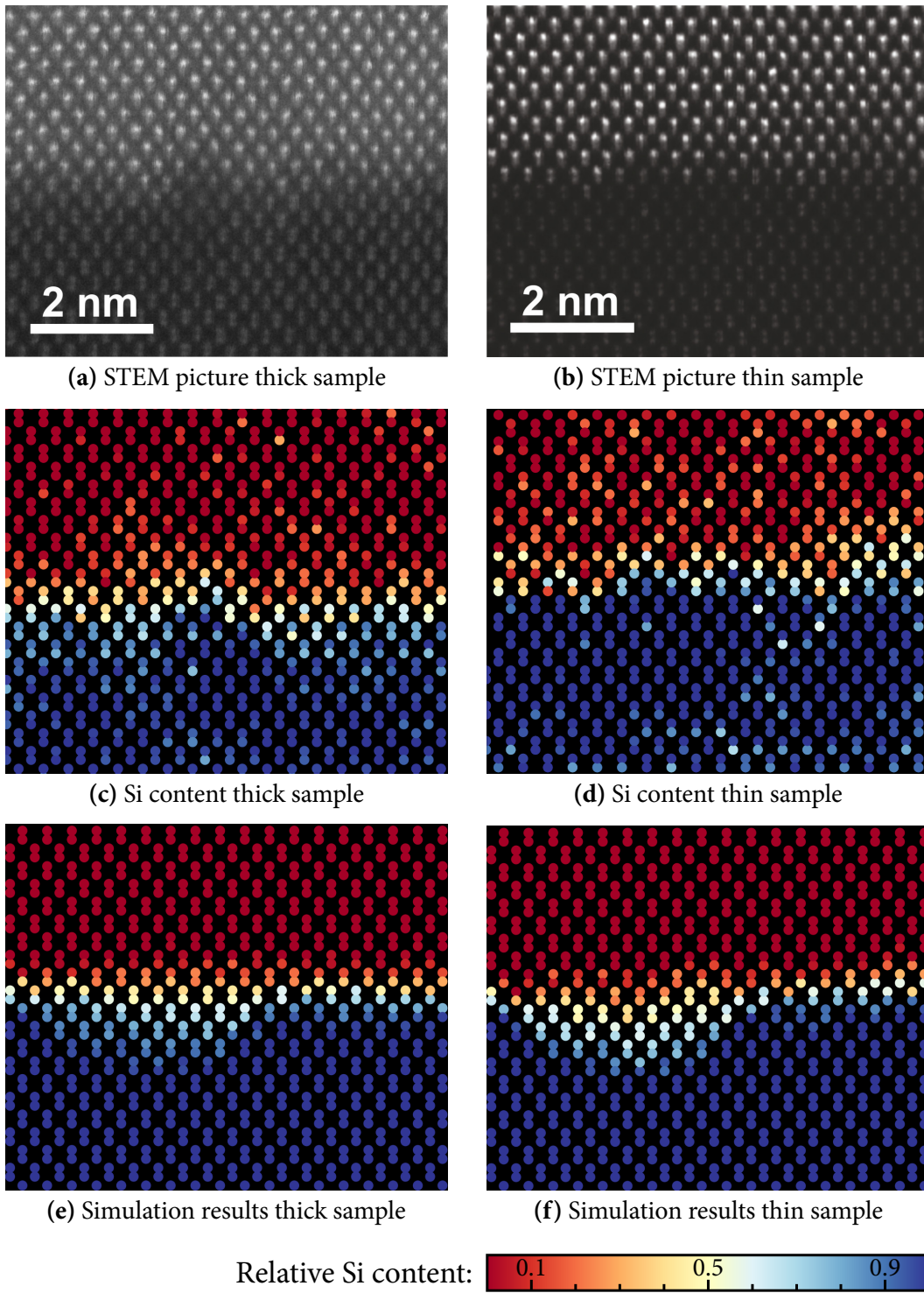
Figure 11.1: The GaP/Si interface viewed along the [110] crystal axis in STEM. Figure (a) and (b) show the raw HAADF-STEM data obtained at thick (20 nm) and thin (8 nm) regions of the sample, respectively. (c) and (d) are maps of the corresponding Si content as simulated by the absorptive potential approximation [93]. Figure (e) and (f) show the Si content from the KMC simulation results for thick and thin sample cutouts, respectively. Figure (a) and (b) by WZMW.
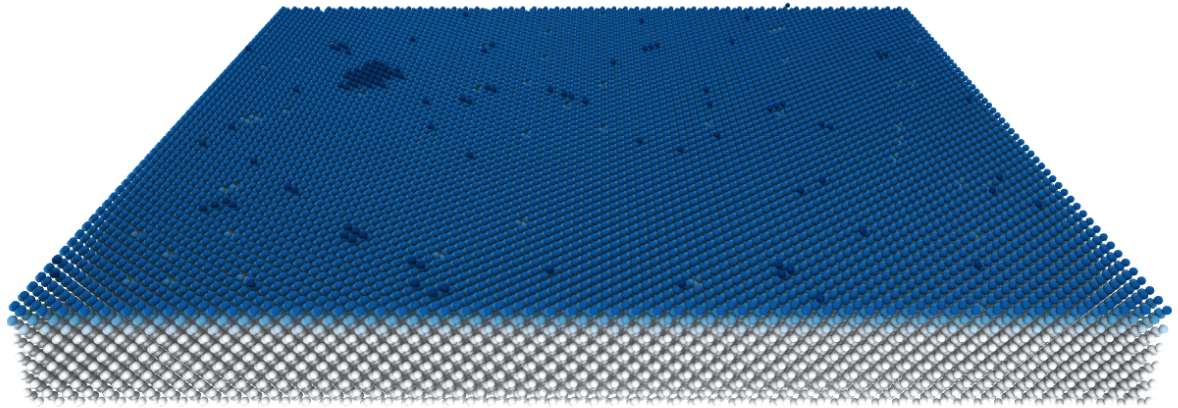
(a) Si substrate before GaP deposition
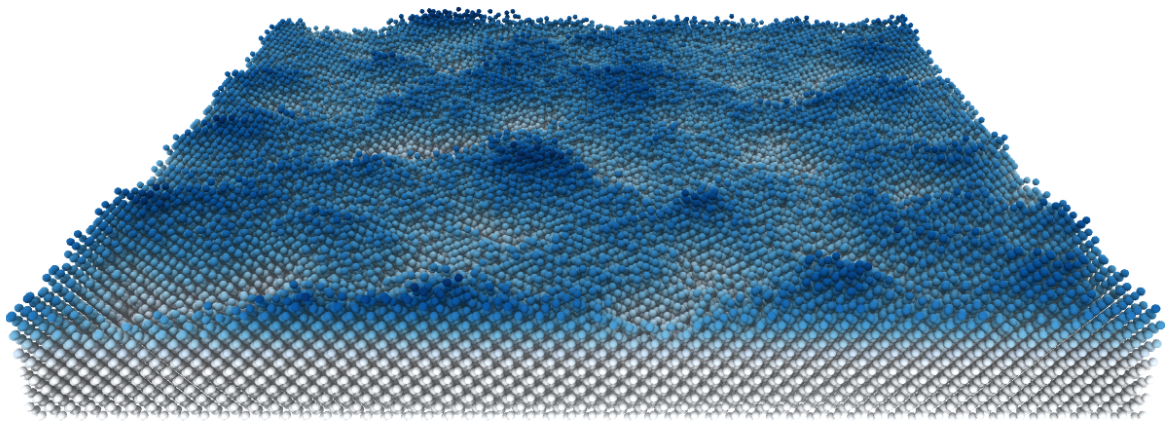
(b) Si substrate after GaP deposition

Figure 11.2: Simulation results of the Si substrate (a) prior to and (b) after the growth of GaP (Ga and P atoms are hidden). The colors of the atoms are chosen according to their positions in growth direction, with higher atoms appearing darker. The simulated system size is $55 \times 55 \times 60$ unit cells. Figure created with Ovito [90].

deposition of GaP the substrate became rough and significant intermixing was induced at the buried interface.

For a comparison with the experimental observations, the 3D results in Figure 11.2 were analyzed similar to the functionality of the STEM microscope. The sample was cut into slices of 20 nm and 8 nm and the Si content of the atomic columns was determined in viewing direction [110] and plotted as a 2D map, similar to Figure 11.1 (c) and (d). The results are shown in Figure 11.1 (e) and (f). Clearly, both the thick and thin cutting show intermixing of about 7 ML to 8 ML. The clustering of the Si structures is more noticeable in the thin sample, both the experimental one Figure 11.1 (d) and the simulated one (f), and is nicely visible also in the 3D picture Figure 11.2. The averaging procedure to obtain the 2D maps of the Si content smears out the Si clustering in the thick samples, (c) and (e), so that the clusters at the interface are less pronounced.

To quantify both the experimental and theoretical results, one can plot the Si content of each

(a) Thick sample   **(b)** Thin sample

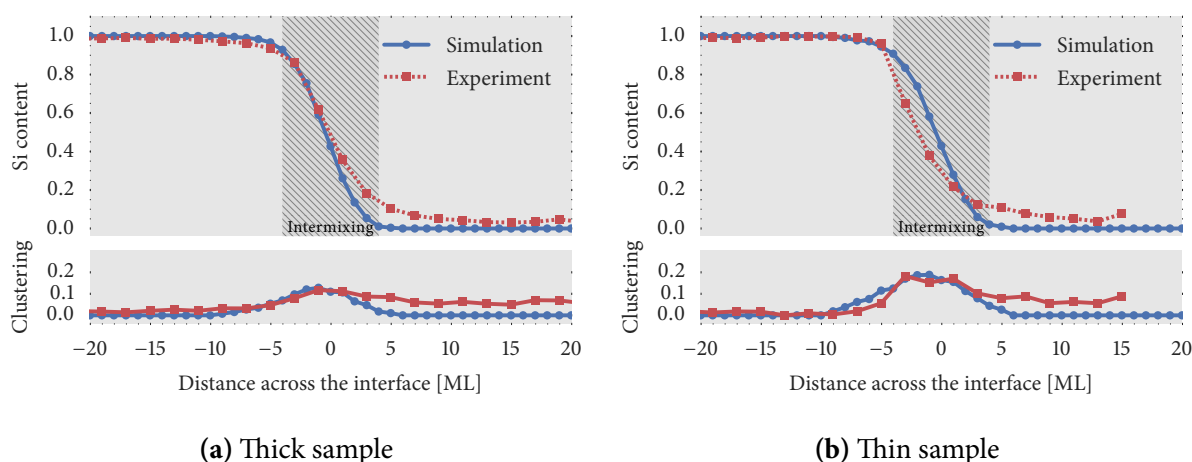**Figure 11.3:** The Si content versus distance across the interface in units of single ML. The variance of the averaging procedure is a measure of the clustering at the interface and is plotted at the bottom of the figure. Figure (a) shows the thick and (b) the thin region of the sample. The hatched area in both pictures is the intermixing region.

ML against the distance across the interface, i.e., the number of the ML in growth direction (z coordinate of Figure 10.1). To obtain the Si content of each ML, the 2D pictures from Figure 11.1 were used and all values of the Si content corresponding to the same ML (z coordinate) were averaged. This procedure yields not only the total Si content of the whole ML, but the variance of the averaging is also a measure for the amount of clustering visible in the picture. If, for instance, the intermixing was completely homogeneous, each atomic column would contain the same amount of Si and the variance of the averaging procedure would be zero. If, on the other hand, the interface is structured, columns in the same ML will contain different amounts of Si and the variance will be large. This measure is expected to depend on the thickness of the sample, since the intermixing is more homogeneous when the atomic columns contain more atoms in total.

The 1D results are shown in Figure 11.3. Both the thick (Figure 11.3 (a)) and thin (Figure 11.3 (b)) region show excellent agreement between the KMC simulations and experiment. Furthermore, the variance of the averaging procedure agrees nicely and, as expected, shrinks with increasing thickness of the sample. Because P and Si have very similar proton numbers, they appear almost equally bright in the STEM pictures and are hardly distinguishable. Therefore, only data points from those MLs are shown, that belong to the Ga sublattice.

The strong intermixing in the simulation results is induced by the existence of one slow atomic species on the surface. In our simulations, this role is taken by the P atoms, who are much more strongly bound to the Si substrate than the Si-Si and Si-Ga bonds. (See Table 11.1.) Diffusion for P atoms therefore is much slower and they act as immobile centers, around which Si clusters can form. This effect significantly supports the interface intermixing and gives rise to the observed results.

11.3 Ab-*initio* Verification of the Barriers

The premise of the high diffusion barriers for P atoms on Si has been checked in ab-*initio* DFT calculations by the theoretical chemistry group of Prof. R. Tonner at the Philipps-University in

Figure 11.4: Setup of DFT calculations for the energy barriers of different atomic species hopping on different substrates. The adatom species are Si, Ga, and P, and the substrate materials were (a) Si, (b) P-terminated GaP, and (c) Ga-terminated GaP. Blue spheres are Si atoms, the orange spheres are P, and Ga is shown as green spheres. Three MLs are shown, indicated by decreasing atomic radii with increasing distance from the surface. The arrows indicate the direction of movement. Figure by Andreas Stegmüller.
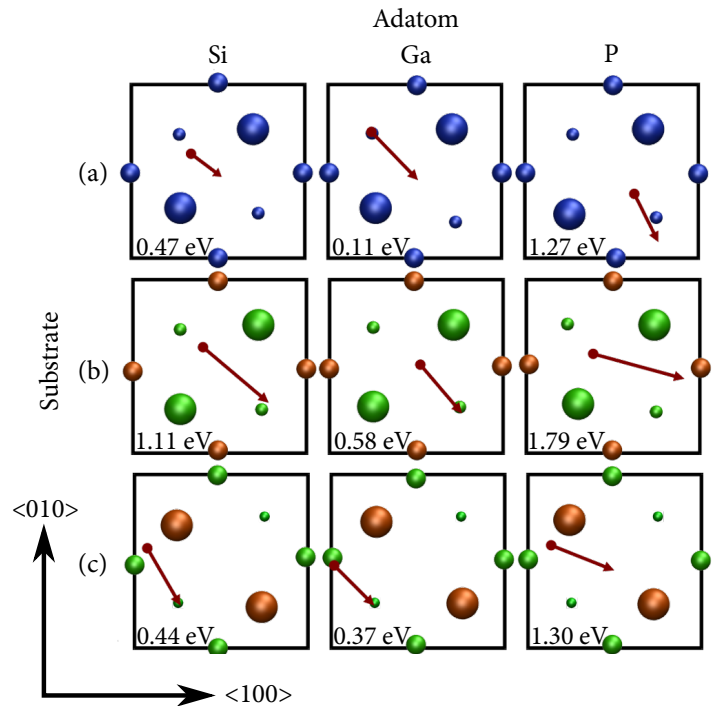
Marburg. Hops of single Si, Ga, and P atoms from one equilibrium lattice position to a second-nearest neighboring position within the same ML were calculated on perfectly flat Si, P-terminated GaP, and Ga-terminated GaP substrates. In the P terminated GaP substrate, the upmost atomic ML is filled with P atoms and vice versa for Ga termination. Since only relative barrier heights are of interest for our considerations, the lattices were not relaxed but a perfect zincblende (or diamond) lattice structure was assumed.

The matrix of calculations and the energy barriers are shown in Figure 11.4. Each column represents the calculations of an adatom of different species (Si, Ga, and P) and each row is a different substrate lattice, Si, P-terminated GaP, and Ga-terminated GaP. It is clearly visible from the results, that the barriers for P atoms on either surface are much higher than those of the other species. Similarly, the barriers for a Si atom hopping on a P-terminated substrate are significantly higher than for the other substrates. This confirms the assumption of P being more tightly bound to Si than all the other bonds in the system, as shown in Table 11.1. It therefore represents the slow atomic species that is responsible for the strong interface intermixing.

11.4 Other Growth Parameters and Pure Diffusion

To rule out the possibility of the experimental results being an artefact, the measurements were repeated with slightly different growth parameters or substrates. In addition to the "exact" sample which was used for the analysis in Section 11.1, three samples were grown and investigated: The "high-T" sample was grown at $T = 948\,\mathrm{K}$, the "2° off" sample had a substrate off orientated by 2°, and in the "reversed polarity" sample, the element binding to Si, i.e., the element forming the first monolayer on the substrate, was changed. The resulting Si content versus distance across the interface is shown in Figure 11.5. The intermixing in all samples except the "high-T" one is comparatively strong, ruling out the possibility for the intermixing effects being an experimental
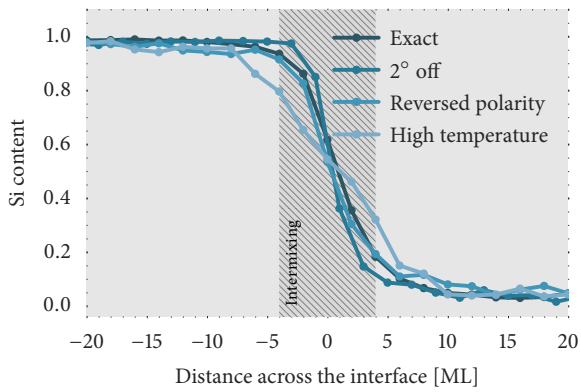
Figure 11.5: Si content versus distance across the interface for the "exact" sample, which was analyzed in Section 11.1, and three other samples: The "high-T" sample, the "2° off" sample and a sample with "reversed polarity".
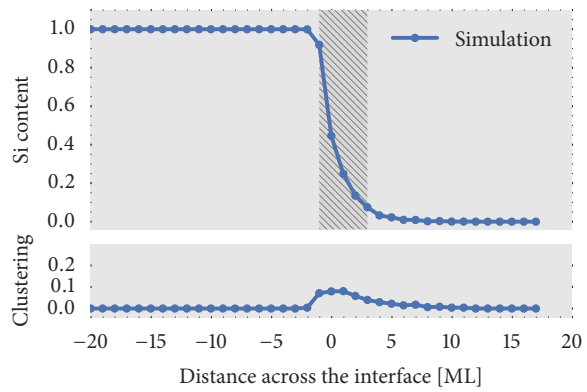
Figure 11.6: The simulation results for the case of equal energy barriers for all atomic species. This corresponds to the pure interdiffusion with indistinguishable particle types. The Si content versus distance across the interface in units of ML is plotted.

artefact. In the case of high growth temperature $T = 948\,\text{K}$, the "high-T" sample, the intermixing is dominated by temperature driven diffusion and appears even stronger than in the other samples.

To verify that the strong Si-P bonds, corresponding to the presence of a slow atomic species at the surface, induce the strong intermixing, the simulation was repeated with the energy barriers of all atomic species being equal, i.e., set according to Table 11.1 but with the Si-P contribution also set to 0.3 eV for the nearest and 0.15 eV for the second-nearest neighbors. The resulting intermixing is plotted in Figure 11.6 and amounts to approximately four ML with small amounts of clustering. This result proves that interdiffusion and the clustering at the GaP-Si interface is driven by deviating bond strengths between Si and P.

Chapter 12

Structural Characteristics of Ga Metal Deposited on Si

Ga containing III/V semiconductors on Si substrates offer a wide variety of applications. Although the electronic and structural properties of the grown materials are well accessible and studied, little is known about the buried interfaces and the atomisticc processes at the substrate surfaces during the growth procedure. Recently, it was observed, that Ga deposition by MOVPE on Si-(001) substrates leads to the formation of metallic Ga droplets, which significantly etch into the Si substrates [94], resulting in large, pyramidally shaped cavities in the formerly flat substrates.

In this chapter, we show that KMC computer simulations of MOVPE growth of Ga on Si-(001) are able to reproduce such etching effects and therewith identify the atomic kinetics during epitaxial growth as a possible cause of such observations. The experimental measurements and our theoretical findings are summarized in the following Sections 12.1 and 12.2, respectively. The published work, Article [V], is attached.

12.1 Experimental Observations

Ga was grown by MOVPE in a commercial Aixtron AIX 200 GFR reactor onto atomically flat Si-(001) substrates. Two precursors were used in comparison, TEGa and trimethylgallium (TMGa), with growth temperatures varied from 648 K to 723 K for TEGa and 723 K to 823 K for TMGa. Some of the samples were later annealed at temperatures between 673 K to 1073 K.

The sample surface morphology was analyzed by atomic force microscopy (AFM). For some samples, the grown Ga was first removed by selective etching using hydrochloric acid (HCl), which leaves the Si surface intact. In addition, STEM pictures were taken in a JEOL 2200 FS microscope, with the viewing direction along the [110] crystal axis. For the STEM measurements, electron transparent foils were prepared using high energy argon milling.

The measured results are shown in Figure 12.1. The Ga droplets are clearly visible as bright spots in Figures 12.1 (a) and 12.1 (b), which show the surface structure as measured by AFM for the two different precursors. The density and size of the droplets depend on the used precursor. Analyzing the solid/liquid interface between the droplets and the substrate by cross-sectional HAADF-STEM in Figures 12.1 (c) and 12.1 (d) reveals significant etching of the Ga droplets into the substrate, independent of the choice of the precursor. The etched structures are pyramidally shaped and limited by the {111} family of planes of the diamond cubic Si lattice. The depth of the pyramids stretches over many MLs, reaching more than 10 nm. In case of the TMGa precursor, Figure 12.1 (d), recrystallization of the Si material dissolved in the Ga droplet is visible at the boundaries of the cavity, which is shown as an inset of the figure.

12.2 KMC Computer Simulation

Ga deposition on the Si-(001) substrate was simulated by means of the bond counting KMC method described in Chapter 10. The chosen system size was $40 \times 40 \times 80$ unit cells, which corresponds
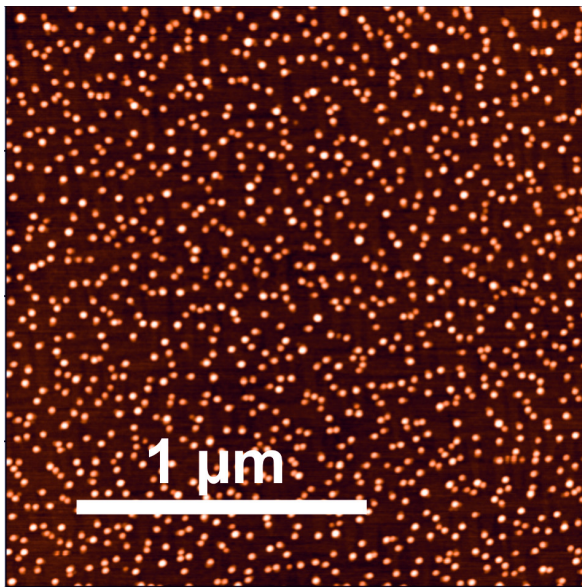
(a) AFM measurement TEGa precursor
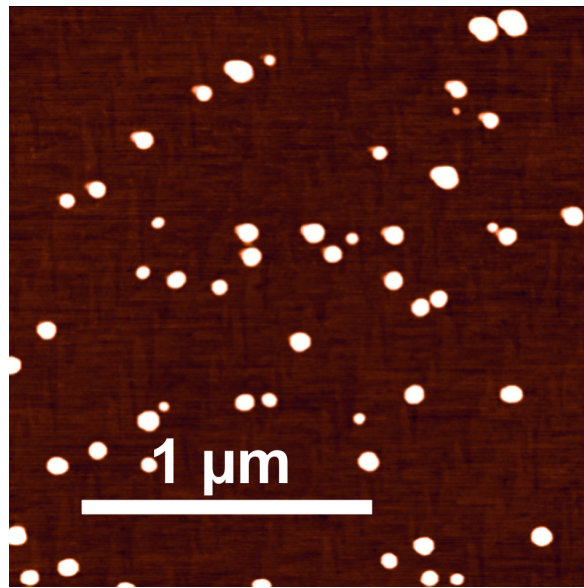
(b) AFM measurement TMGa precursor
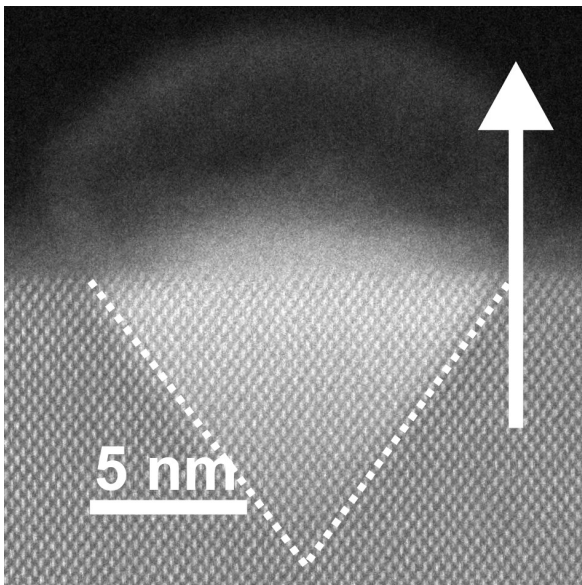
(c) HAADF-STEM measurement TEGa precursor
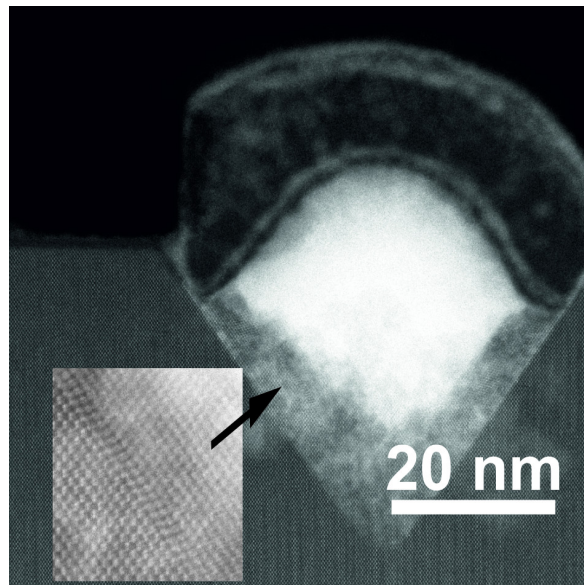
(d) HAADF-STEM measurement TMGa precursor

Figure 12.1: Experimental results of the surface morphology of Ga grown on Si-(001) substrates. (a) and (b) show AFM pictures of the surface for the different precursors (a) TEGa and (b) TMGa. The bright spots are metallic Ga droplets. Cross-sectional HAADF-STEM images of the samples are shown in (c) and (d), for the different precursors respectively. As clearly visible, in both cases the Ga droplets etched large, pyramidal structures framed by the {111} lattice planes into the Si substrate. Figures (a)–(d) are taken from Article [V].

Table 12.1: The energy barrier contributions for the bond counting KMC method of the Ga on Si-(001) growth simulation. Each pair of atomic species is assigned a nearest (NN) and second-nearest (2NN) value.

	Si NN	Si 2NN	Ga NN	Ga 2NN
Si	0.3 eV	0.15 eV	0.3 eV	0.15 eV
Ga	0.3 eV	0.15 eV	0.1 eV	0.25 eV

to 1 024 000 simulated lattice sites. Enabled events were the adsorption and the diffusion event described in Section 10.2. Similar to the procedure of the GaP on Si growth simulation, described in Section 11.2, a 25 ML thick Si substrate was first grown, resulting in an atomically flat Si-(001) surface. Afterwards, a total of 10 ML of Ga was grown during a time of 30 s (hence, with a growth rate of $\Gamma = 0.33\,\text{ML/s}$) and the system was relaxed for another 5 s, during which the adsorption event was disabled. Growth was simulated at a temperature of $T = 750\,\text{K}$, which is close to the experimentally used temperatures. The energy barrier contributions of the Si-Si, Si-Ga, and Ga-Ga nearest and second-nearest neighbor bonds are listed in Table 12.1.

The chosen barriers are the same as in the GaP-Si interface simulation described in Chapter 11, except for the Ga-Ga interaction. Small nearest neighbor and the high second-nearest neighbor contribution of Ga-Ga are chosen to approximate the liquid character of the metallic Ga.
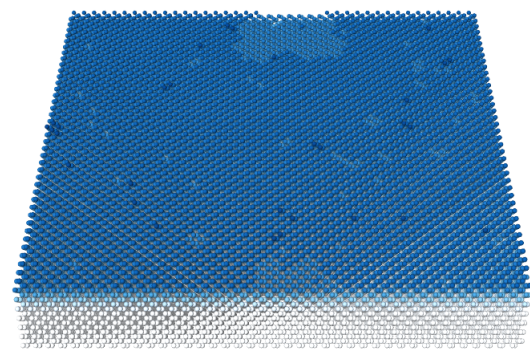
(a) Si substrate before Ga deposition
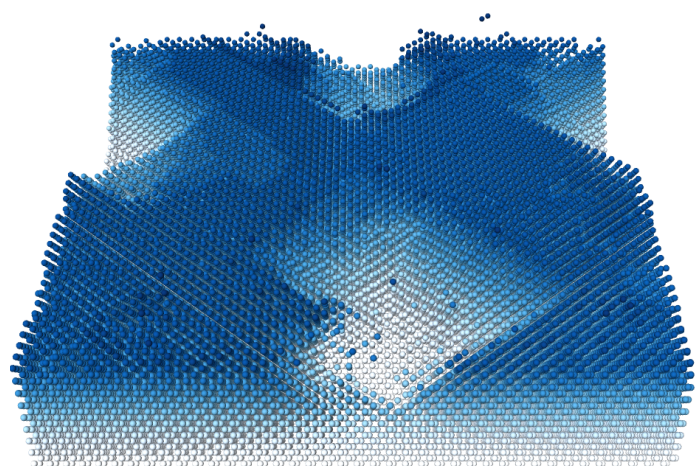
Figure 12.2: The simulated Si substrate (a) before and (b) after the deposition of Ga. The atoms are colored according to the position in growth direction, i.e., higher atoms are darker and lower atoms are colored brighter. Figure created with Ovito [90].

(b) Si substrate after Ga deposition
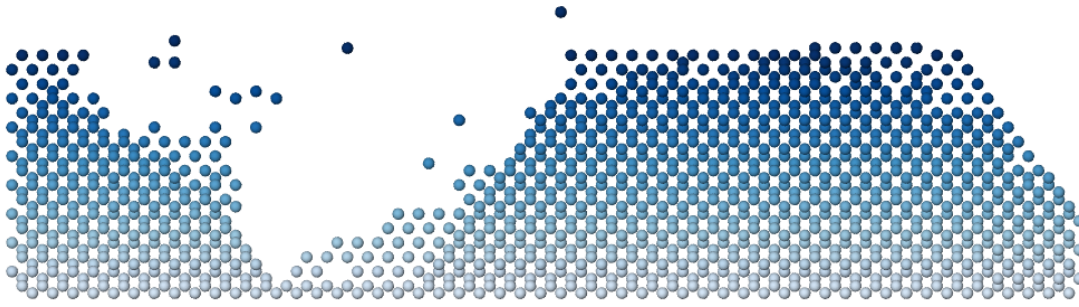
Figure 12.3: A sliced view along the [110] direction of the substrate after Ga deposition. Ga atoms are hidden. The pyramidal, {111} limited etching structure is clearly visible. Figure created with Ovito [90].

Figure 12.2 shows the Si substrate before (a) and after (b) deposition of Ga, with Ga atoms hidden. Clearly, the substrate is initially atomically flat and became structured only after Ga arrived at the surface. Figure 12.2 (b), clearly shows the large pyramidally shaped etching structures in the Si sustrate, as deep as the simulated supercell. A sliced view of the substrate, parallel to the (110) crystal plane, is shown in Figure 12.3. The pyramidal structures are clearly visible in the figure and the etching structures are framed by the {111} planes of the substrate lattice, as observed experimentally. Since the pyramid is too large for the 25 ML thick substrate in the simulation, it is cut off at the bottom and the tip is not visible.

The excellent agreement between the simulation results and experimental findings suggests kinetic aspects during the deposition of Ga on Si-(001) as a possible reason for the observed substrate etching effects. The stability of the {111} family of lattice planes is nicely reproduced in the simulation and gives rise to similar pyramidal shapes of the etching structures in the simulations and in experiments.

Chapter 13

Conclusions and Outlook

The increasing complexity of III/V semiconductors integrated on Si based devices requires a detailed theoretical understanding of the growth and post-processing procedures. Non-equilibrium epitaxial growth of semiconductors, for instance via MOVPE, is a complex procedure that involves many chemical and physical processes on long time and length scales. In this Part II of the thesis, a method to theoretically model such growth procedures on the basis of KMC computer simulations was introduced and applied to study two different phenomena of III/V semiconductor growth on Si.

We showed, that the KMC method is an efficient algorithm to simulate epitaxy with large supercells on long time scales by discretizing the kinetics of single atoms and employing effective event rates. Rather than determining these event rates using ab-*initio* methods or bond order potentials, we used the bond counting KMC scheme, which is a simple and extremely efficient approximation for the calculation of energy barriers of the atomistic processes. It was shown, how the single steps of the KMC algorithm can be optimized so that the overall complexity of the simulations grows linearly with the size of the supercell. Altogether, the method allows for the direct simulation of systems of several tens of nanometres side lengths on time scales of minutes.

The KMC code was first applied to simulate GaP growth on Si-(001) substrates. It was previously observed in samples grown by MOVPE, that the buried interface between the Si substrate and the GaP crystal exhibits significant intermixing of the two materials, albeit the Si substrate was formerly atomically flat. Such intermixing behaviour was shown to be reproducible in KMC simulations, identifying the atomic kinetics on the substrate during growth as a possible cause. The required simulation parameters for the intermixing were then confirmed in DFT studies: P adatoms are much more tightly bound to the Si substrates than Ga and Si adatoms. P adatoms therefore diffuse very slowly on the substrate as compared to the other atomic species, which strongly assists the intermixing. The excellent agreement between experimental observations, KMC simulations, and DFT calculations demonstrates that KMC is a useful tool to study atomic kinetics during epitaxial growth and can be used to predict certain properties of the material's structure formation.

The second application of the KMC code to experimental findings was the modeling of etching effects of Ga droplets on the Si-(001) substrate. It was observed experimentally, that Ga layers grown by MOVPE formed metallic droplets on the substrate. HAADF-STEM measurements show, that these droplets dissolved the Si material beneath and therewith etched large, pyramidally shaped holes into the substrate. The experiments could be reproduced with KMC simulations of Ga growth on Si. In the simulations, the Ga atoms gathered in droplets as well and the etching of large, pyramidally shaped structures into the Si substrate could be reproduced. The agreement between experiment and theory in this second study once more proves the usefulness of KMC for the investigation of atomic kinetics during epitaxial growth.

Herewith, it was shown that KMC simulations are viable candidates for the theoretical modeling of non-equilibrium growth procedures of metastable III/V semiconductors.

Although the basic KMC method introduced in this work proved sufficient to reproduce some

experimental observations, it needs to be significantly extended to describe other material combinations and more complex growth mechanisms. Here we give a brief perspective on the possible ways, in which the basic KMC code can be extended to account for additional physical and chemical phenomena.

Most III/V semiconductors have lattice constants different to that of Si. Growth of such materials on Si substrates therefore results in strained layers and the elastic energy becomes a nonneglibile driving force for the structure formation. The forces introduced by the strained lattices are non-local, and therefore difficult to integrate into KMC without completely destroying the advantages of the KMC technique. There are, however, methods to include lattice mismatch in the calculation of the rate constants of the atomistic events, in order to approximate the effects of elastic strain. The most popular approach is the "ball-and-spring" model, that was introduced by Orr et al. [77] and was since enhanced and applied several times [88, 89, 95, 96]. Rudin et al. [80] recently suggested another technique for the inclusion of lattice strain in KMC, based on the interatomic Keating potential [81]. With that technique, the authors successfully modeled Stranski-Krastanov growth of Ge islands on Si substrates [80]. The implementation of such algorithms on top of KMC is necessary in order to broaden the number of material combinations that can be modeled by the simulation code.

With the KMC procedure introduced in this work it is possible to simulate vacancy diffusion in grown crystals. When a lattice position is unoccupied, one of the atoms on neighboring sites can hop into the vacancy by the usual diffusion event. A sequence of such processes allows, for instance, to model annealing effects caused by vacancy diffusion. However, atoms can not only diffuse via vacancies through the bulk, but also by interstitial diffusion, which can be the dominant mechanism especially for small atoms. During interstitial diffusion, the atoms diffuse off-lattice between the atoms of the host material. In order to accurately model the diffusion kinetics during high-temperature annealing of grown materials, such interstitial diffusion mechanisms should therefore be integrated into the KMC code. One way of doing this is to overlay the crystal lattice with a virtual lattice, allowing transitions between the lattices and employing special diffusion kinetics on the virtual lattice.

A third interesting extension of the basic KMC code is the possibility to simulate liquid and solid phases in parallel, for instance, in order to model droplet epitaxy. A corresponding KMC algorithm was recently suggested by Reyes et al. [82]. The authors assumed constant diffusion rates in those parts of the sample, that are labeled as "liquid", and employed additional energy barriers for the attachment and detachment of atoms at the solid/liquid interface. With that method, formation of nanostructures in droplet epitaxy could successfully be modeled and experiments were nicely reproduced [82].

Finally, it would be desirable to check the assumptions of the bond counting KMC scheme by comparison with more accurate bond order potentials, e.g., the Tersoff potential, or ab-*initio* DFT methods. To do that, one could, for example, study the energy barriers of single diffusion steps by means of DFT calculations as a function of the number of atoms in the surrounding. In bond counting KMC, a linear dependence of the barriers on number of neighbors is presumed.

Appendices

Bibliography

[1] T. Instruments, *The chip that Jack built*, (1997) `http://www.ti.com/corp/docs/kilbyctr/jackbuilt.shtml` (visited on 07/08/2015).

[2] F. Faggin, Byte **17**, 145 (1992).

[3] Wikipedia, *14 nanometer - Wikipedia, The Free Encyclopedia*, (2015) `https://en.wikipedia.org/wiki/14%5C_nanometer` (visited on 07/03/2015).

[4] S. D. Baranovskii, ed., *Charge transport in disordered solids with applications in electronics* (John Wiley & Sons, Ltd., Chichester, 2006).

[5] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, 1976).

[6] J. O. Oelerich, "On the validity of the Einstein relation for disordered systems", Bachelor Thesis (Philipps-Universität Marburg, 2010).

[7] A. V. Nenashev, J. O. Oelerich, and S. D. Baranovskii, Journal of Physics: Condensed Matter **27**, 093201 (2015).

[8] A. V. Nenashev, F. Jansson, J. O. Oelerich, D. Huemmer, A. V. Dvurechenskii, F. Gebhard, and S. D. Baranovskii, Physical Review B **87**, 235204 (2013).

[9] K. Volz, W. Stolz, A. Dadgar, and A. Krost, "Growth of III/Vs on Silicon: Nitrides, Phosphides, Arsenides and Antimonides", in *Handbook of crystal growth*, edited by T. Nishinga and T. F. Kuech, 2nd ed. (Elsevier, Amsterdam, 2015) Chap. 31, pp. 1249–1300.

[10] A. C. Levi and M. Kotrla, Journal of Physics: Condensed Matter **9**, 299 (1997).

[11] A. Groß, *Theoretical Surface Science* (Springer, Ulm, 2009).

[12] J. O. Oelerich, D. Huemmer, and S. D. Baranovskii, Physical Review Letters **108**, 226403 (2012).

[13] J. O. Oelerich, "Theoretical Description of Charge Transport in Organic Semiconductors", Master Thesis (Philipps-Universität Marburg, 2012).

[14] J. Orenstein and M. Kastner, Solid State Communications **40**, 85 (1981).

[15] S. D. Baranovskii, physica status solidi (b) **251**, 1 (2014).

[16] S. J. S. Lemus and J. Hirsch, Philosophical Magazine B **53**, 25 (1986).

[17] R. A. Street, *Hydrogenated Amorphous Silicon* (Cambridge University Press, Cambridge, 1991).

[18] C. M. Soukoulis, M. H. Cohen, and E. N. Economou, Physical Review Letters **53**, 616 (1984).

[19] M. Schwoerer and H. C. Wolf, *Organic Molecular Solids* (Wiley-VCH, Weinheim, 2007).

[20] H. Bässler, physica status solidi (b) **175**, 15 (1993).

[21] F. Neese, Wiley Interdisciplinary Reviews: Computational Molecular Science **2**, 73 (2012).

[22] U. Varetto, *Molekel*.

[23] L. Ley, "Photoemission and optical properties", in *The physics of hydrogenated amorphous silicon ii*, edited by J. D. Joannopoulos and G. Lucovsky (Springer, Berlin, 1984) Chap. 3, pp. 61–168.

[24] W. E. Spear, Journal of Non-Crystalline Solids **59–60**, 1 (1983).

[25] F. Gutmann and L. E. Lyons, *Organic Semiconductors* (Wiley, New York, 1967).

[26] E. A. Silinsh, physica status solidi (a) **3**, 817 (1970).

[27] H. Bässler, physica status solidi (b) **107**, 9 (1981).

[28] F. Maddalena, C. D. Falco, M. Caironi, and D. Natali, Organic Electronics **17**, 304 (2015).

[29] C. Tanase, E. J. Meijer, P. W. M. Blom, and D. M. de Leeuw, Physical Review Letters **91**, 216601 (2003).

[30] C. Tanase, P. W. M. Blom, D. M. de Leeuw, and E. J. Meijer, physica status solidi (a) **201**, 1236 (2004).

[31] A. Miller and E. Abrahams, Physical Review **120**, 745 (1960).

[32] R. A. Marcus, Annual Review of Physical Chemistry **15**, 155 (1964).

[33] B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer, Berlin, 1984).

[34] D. Monroe, Physical Review Letters **54**, 146 (1985).

[35] S. D. Baranovskii, P. Thomas, and G. J. Adriaenssens, Journal of Non-Crystalline Solids **190**, 283 (1995).

[36] S. D. Baranovskii, T. Faber, F. Hensel, and P. Thomas, Journal of Physics: Condensed Matter **9**, 2699 (1997).

[37] V. I. Arkhipov, P. Heremans, E. V. Emelianova, G. J. Adriaenssens, and H. Bässler, Journal of Physics: Condensed Matter **14**, 9899 (2002).

[38] R. Schmechel, Physical Review B **66**, 235206 (2002).

[39] R. Schmechel, Journal of Applied Physics **93**, 4653 (2003).

[40] J. O. Oelerich, F. Jansson, A. V. Nenashev, F. Gebhard, and S. D. Baranovskii, Journal of Physics: Condensed Matter **26**, 255801 (2014).

[41] S. D. Baranovskii, H. Cordes, F. Hensel, and G. Leising, Physical Review B **62**, 7934 (2000).

[42] O. Rubel, S. D. Baranovskii, P. Thomas, and S. Yamasaki, Physical Review B **69**, 014206 (2004).

[43] J. P. Gonzalez-Vazquez, J. A. Anta, and J. Bisquert, Physical Chemistry Chemical Physics **11**, 10359 (2009).

[44] B. Hartenstein and H. Bässler, Journal of Non-Crystalline Solids **190**, 112 (1995).

[45] S. V. Novikov and G. G. Malliaras, physica status solidi (b) **243**, 387 (2006).

[46] B. I. Shklovskii and A. L. Efros, Soviet Physics JETP **33** (1971).

[47] V. Ambegaokar, B. Halperin, and J. Langer, Physical Review B **4**, 2613 (1971).

[48] M. Pollak, Journal of Non-Crystalline Solids **11**, 1 (1972).

[49] M. C. J. M. Vissenberg and M. Matters, Physical Review B **57**, 12964 (1998).

[50] M. Grünewald and P. Thomas, physica status solidi (b) **125**, 125 (1979).

[51] C. D. Lorenz and R. M. Ziff, Journal of Physics A: Mathematical and General **31**, 8147 (1998).

[52] S. D. Baranovskii, I. P. Zvyagin, H. Cordes, S. Yamasaki, and P. Thomas, physica status solidi (b) **230**, 281 (2002).

[53] G. Juška, K. Arlauskas, M. Viliunas, and J. Kočka, Physical Review Letters **84**, 4946 (2000).

[54] T. Tiedje and A. Rose, Solid State Communications **37**, 49 (1981).

[55] G. Juška, K. Arlauskas, J. Kočka, M. Hoheisel, and P. Chabloz, Physical Review Letters **75**, 2984 (1995).

[56] G. Juška, Journal of Non-Crystalline Solids **137&138**, 401 (1991).

[57] H. Antoniadis and E. A. Schiff, Physical Review B **43**, 13957 (1991).

[58] V. Karpus and V. I. Perel, Soviet Physics JETP **64**, 1376 (1986).

[59] L. V. Keldysh, Soviet Physics JETP **7**, 665 (1958).

[60] G. J. Adriaenssens, S. D. Baranovskii, W. Fuhs, J. Jansen, and Ö. Öktü, Physical Review B **51**, 9661 (1995).

[61] M. Tachiya and K. Seki, Physical Review B **82**, 085201 (2010).

[62] J. Nelson, Physical Review B **67**, 155209 (2003).

[63] A. Hofacker, J. O. Oelerich, A. V. Nenashev, F. Gebhard, and S. D. Baranovskii, Journal of Applied Physics **115**, 223713 (2014).

[64] G. Schönherr, H. Bässler, and M. Silver, Philosophical Magazine Part B **44**, 369 (1981).

[65] Z. G. Yu, D. L. Smith, A. Saxena, R. L. Martin, and a. R. Bishop, Physical Review Letters **84**, 721 (2000).

[66] W. F. Pasveer, J. Cottaar, C. Tanase, R. Coehoorn, P. A. Bobbert, P. W. M. Blom, D. M. de Leeuw, and M. A. J. Michels, Physical Review Letters **94**, 206601 (2005).

[67] F. Jansson, "Charge transport in disordered materials", PhD Thesis (Åbo Akademi University, 2011).

[68] G. E. Moore, Proceedings of the IEEE **86**, 82 (1998).

[69] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, Journal of Applied Physics **89**, 5815 (2001).

[70] S. Gies, M. Zimprich, T. Wegele, C. Kruska, A. Beyer, W. Stolz, K. Volz, and W. Heimbrodt, Journal of Crystal Growth **402**, 169 (2014).

[71] W. Harrison, E. Kraut, J. Waldrop, and R. Grant, Physical Review B **18**, 4402 (1978).

[72] F. F. Abraham and G. M. White, Journal of Applied Physics **41**, 1841 (1970).

[73] G. H. Gilmer and P. Bennema, Journal of Crystal Growth **13-14**, 148 (1972).

[74] A. Bortz, M. Kalos, and J. Lebowitz, Journal of Computational Physics **17**, 10 (1975).

[75] P. Smilauer and D. D. Vvedensky, Physical Review B **52**, 14263 (1995).

[76] D. G. Vlachos, L. D. Schmidt, and R. Aris, Physical Review B **47**, 4896 (1993).

[77] B. G. Orr, D. Kessler, C. W. Snyder, and L. Sander, Europhysics Letters **19**, 33 (1992).

[78] B. Meng and W. H. Weinberg, Surface Science **364**, 151 (1996).

[79] B. Liedke, K.-H. Heinig, A. Muücklich, and B. Schmidt, Applied Physics Letters **103**, 133106 (2013).

[80] S. A. Rudin, V. A. Zinov'ev, A. V. Nenashev, A. Y. Polyakov, Z. V. Smagina, and A. V. Dvurechenskii, Optoelectronics, Instrumentation and Data Processing **49**, 461 (2013).

[81] P. N. Keating, Physical Review **145**, 637 (1966).

[82] K. Reyes, P. Smereka, D. Nothern, J. Millunchick, S. Bietti, C. Somaschini, S. Sanguinetti, and C. Frigeri, Physical Review B **87**, 165406 (2013).

[83] F. Grosse, W. Barvosa-Carter, J. J. Zinck, and M. F. Gyure, in Mrs proceedings, Vol. 701 (2001).

[84] A. A. Schmidt, V. S. Kharlamov, K. L. Safonov, Y. V. Trushin, E. E. Zhurkin, V. Cimalla, O. Ambacher, and J. Pezoldt, in Computational materials science, Vol. 33, 1-3 (2005), pp. 375–381.

[85] P. Kratzer, E. Penev, and M. Scheffler, Applied Physics A **75**, 79 (2002).

[86] K. Reuter, "First-Principles Kinetic Monte Carlo Simulations for Heterogeneous Catalysis: Concepts, Status and Frontiers", in *Modeling heterogeneous catalytic reactions: from the molecular process to the technical system*, edited by O. Deutschmann (Wiley-VCH, Weinberg, 2009) Chap. 3, 72ff.

[87] P. Zhang, X. Zheng, S. Wu, J. Liu, and D. He, Vacuum **72**, 405 (2004).

[88] K. E. Khor and S. Das Sarma, Physical Review B **62**, 16657 (2000).

[89] T. P. Schulze and P. Smereka, Journal of the Mechanics and Physics of Solids **57**, 521 (2009).

[90] A. Stukowski, Modelling and Simulation in Materials Science and Engineering **18**, 015012 (2010).

[91] G. Ehrlich and F. G. Hudda, The Journal of Chemical Physics **44**, 1039 (1966).

[92] R. L. Schwoebel and E. J. Shipsey, Journal of Applied Physics **37**, 3682 (1966).

[93] K. Ishizuka, Ultramicroscopy **90**, 71 (2002).

[94] K. Werner, A. Beyer, J. O. Oelerich, S. D. Baranovskii, W. Stolz, and K. Volz, Journal of Crystal Growth **405**, 102 (2014).

[95] A.-L. Barabási, Applied Physics Letters **70**, 2565 (1997).

[96] C.-H. Lam, C.-K. Lee, and L. M. Sander, Physical Review Letters **89**, 216102 (2002).

Contributed Articles

This appendix contains a compilation of the publications described in this thesis. An overview of the articles with their abstracts is given first.

Article I
A. V. Nenashev, J. O. Oelerich, and S. D. Baranovskii

Theoretical tools for the description of charge transport in disordered organic semiconductors

Hopping conduction is widely considered as the dominant charge transport mechanism in disordered organic semiconductors. Although theories of hopping transport have been developed in detail for applications to inorganic amorphous materials, these theories are often out of scope for the community working with organic amorphous systems. Theoretical research on charge transport in organic systems is overwhelmed by phenomenological fittings of numerical results by equations, which often make little physical sense. The aim of the current review is to bring analytical theoretical methods to the attention of the community working with disordered organic semiconductors.

Article II
A. V. Nenashev, F. Jansson, J. O. Oelerich, D. Huemmer, A. V. Dvurechenskii, F. Gebhard, and S. D. Baranovskii

Advanced percolation solution for hopping conductivity

Hopping of carriers between localized states dominates charge transport in amorphous organic and inorganic semiconductors. We suggest a comprehensive description of this transport regime based on the percolation approach that allows one to determine not only very pronounced exponential dependencies of the hopping conductivity on material parameters, but also the more weakly dependent preexponential factors. The problem of the variable-range hopping (VRH) via sites with exponential energy distribution is mapped onto a universal geometrical problem of percolation via spheres with distributed sizes. An exact solution of the latter problem provides accurate results for the VRH in systems with exponential density-of-states (DOS). Our analytical results are confirmed by straightforward computer simulations and compared to former results present in the literature. We also discuss the case of nearest-neighbor hopping on a lattice, where the preexponential factors are provided by the percolation approach for any shape of the DOS.

Article III
J. O. Oelerich, F. Jansson, A. V. Nenashev, F. Gebhard, and S. D. Baranovskii

Energy position of the transport path in disordered organic semiconductors

The concept of *transport energy* is the most transparent theoretical approach to describe hopping transport in disordered systems with steeply energy dependent density of states (DOS), in particular

in organic semiconductors with Gaussian DOS. This concept allows one to treat hopping transport in the framework of a simple multiple-trapping model, replacing the mobility edge by a particular energy level called the transport energy. However, there is no consensus among researchers on the position of this transport level. In this article, we suggest a numerical procedure to find out the energy level most significantly contributing to charge transport in organic semiconductors. The procedure is based on studying the effects of DOS modifications on the charge carrier mobility in straightforward computer simulations. We also show why the *most frequently visited energy*, computed in several numerical studies to determine the transport energy, is not representative for charge transport.

Article IV

A. Hofacker, J. O. Oelerich, A. V. Nenashev, F. Gebhard, and S. D. Baranovskii

Theory to carrier recombination in organic disordered semiconductors
A theoretical description for recombination kinetics of charge carriers in a disordered system with a broad energy distribution of localized states (DOS) is suggested. This kinetics is governed by the exchange of carriers between transport states and traps. Concentration transients in systems with Gaussian DOS, typical for organic semiconductors, appear much steeper than those obtained for systems with exponential DOS. This difference in recombination kinetics is caused by the difference in thermalization kinetics for these two types of the DOS functions. The comparison of the recombination transients for mobile and trapped carriers in exponential and Gaussian DOS might help to distinguish between these two possible shapes of the DOS using experimental data for transient photoconductivity and photoabsorption.

Article V

K. Werner, A. Beyer, J. O. Oelerich, S. D. Baranovskii, W. Stolz, and K. Volz

Structural characteristics of gallium metal deposited on Si (001) by MOCVD
The epitaxial growth of high quality gallium containing III/V material on silicon offers a wide field of possible applications. While the growth of III/V material has already been studied extensively, little is known about the atomistic processes at the interface between silicon and the III/V material. Here we study the deposition of gallium on silicon (0 0 1) by metal organic chemical vapour deposition for the two different precursors triethyl gallium and trimethyl gallium, the influence of varying growth conditions as well as the influence of an annealing after growth. We use atomic force microscopy and high angle annular dark field scanning transmission electron microscopy to analyse the detailed structures occurring for this deposition. For all conditions studied, pyramidal gallium containing structures form in the silicon with boundaries on the silicon {1 1 1} lattice planes. The pyramidal etching structures can be reproduced in straightforward kinetic Monte Carlo simulations of the growth process. When triethyl gallium is used, those structures are covered by carbon containing passivating mounds. After an additional annealing at 8000°C no gallium can be detected anymore for any sample, while an annealing at considerably lower temperatures leads to an alteration of the silicon gallium alloy.

Article I

A. V. Nenashev, **J. O. Oelerich**, and S. D. Baranovskii
Theoretical tools for the description of charge transport in disordered organic semiconductors
J. Phys. Condens. Matter **27**, 093201 (2015)

Topical Review

Theoretical tools for the description of charge transport in disordered organic semiconductors

A V Nenashev[1,2], J O Oelerich[3] and S D Baranovskii[3]

[1] Institute of Semiconductor Physics, 630090 Novosibirsk, Russia
[2] Novosibirsk State University, 630090 Novosibirsk, Russia
[3] Faculty of Physics and Material Sciences Center, Philipps-Universität, D-35032 Marburg, Germany

E-mail: baranovs@staff.uni-marburg.de

CrossMark

Abstract

Hopping conduction is widely considered the dominant charge transport mechanism in disordered organic semiconductors. Although theories of hopping transport have been developed in detail for applications to inorganic amorphous materials, these theories are often out of scope for the community working with organic amorphous systems. Theoretical research on charge transport in organic systems is overwhelmed by phenomenological fittings of numerical results by equations, which often make little physical sense. The aim of the current review is to bring analytical theoretical methods to the attention of the community working with disordered organic semiconductors.

Keywords: charge transport, disorder, organic semiconductors

(Some figures may appear in colour only in the online journal)

1. Introduction

Electronic properties of organic disordered semiconductors (ODSs) are currently the focus of intensive experimental and theoretical research. ODSs already dominate the industrial electrophotographic image recording and their applications in light-emitting diodes (OLEDs) [1, 2], in field-effect transistors (OFETs) [3] and in organic solar cells (OSCs) [4] are intensively worked on. ODSs are a broad class of materials, which can essentially differ from each other in morphology and chemical composition. Most prominent representatives of the ODSs are conjugated polymers [5, 6], molecularly doped polymers [7–9] and low-molecular-weight organic glasses [10, 11]. Understanding the electronic properties of the ODSs is challenging with respect to the fundamental research and device development. In the current report, we focus on the theoretical tools suitable for description of charge transport properties of the ODSs, bearing in mind that charge transport is decisive for all device applications of such materials.

In spite of the broad variety in the materials' chemistry and morphology, it is widely accepted that charge transport in most ODSs is due to incoherent tunneling (hopping) of electrons or holes between strongly localized states, which are distributed randomly in space. This was shown by Bässler in two detailed review papers [9, 12] and confirmed later in numerous review articles, monographs and edited books about charge transport in ODSs [13–24]. The frequency of a hopping transition from an occupied site i to an empty site j is typically [9, 12, 24] described either by the Miller–Abrahams (MA) expression [25]

$$\nu_{ij} = \nu_0 \exp\left(-\frac{2r_{ij}}{\alpha} - \frac{\varepsilon_j - \varepsilon_i + |\varepsilon_i - \varepsilon_j|}{2kT} \right), \qquad (1)$$

or [26] by the Marcus expression [27]

$$\nu_{ij} = \frac{J_o^2}{\hbar} \sqrt{\frac{\pi}{4E_a kT}} \exp\left(-\frac{2r_{ij}}{\alpha} - \frac{E_a}{kT} \right) \qquad (2)$$

$$\times \exp\left(-\frac{\varepsilon_j - \varepsilon_i}{2kT} - \frac{(\varepsilon_j - \varepsilon_i)^2}{16kT E_a} \right).$$
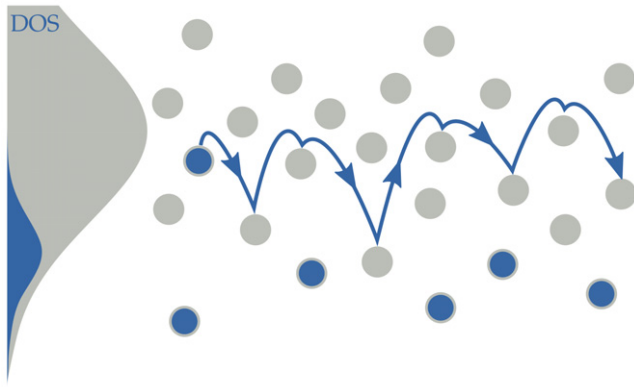
Figure 1. Charge carrier motion via hopping in the Gaussian DOS. Occupied states are marked blue.

In equations (1) and (2), α is the localization length of charge carriers in the localized states, which takes values of the order of 10^{-8} cm [7, 28], ε_i and ε_j are the energies of the states i and j, respectively, k is the Boltzmann constant and T is temperature. The prefactor ν_0 in equation (1) depends on the interaction mechanism responsible for the hopping transitions and is called the *attempt-to-escape frequency*. Usually, electron-phonon interaction is assumed, with ν_0 being close to the phonon frequency 10^{12} s^{-1} [12, 29]. The prefactor J_o^2 in equation (2) is related to the transfer integral connecting the sites involved in the transition and E_a is the polaron activation energy related to the reorganization energy of the matrix.

Often, a Gaussian density of states (DOS) is assumed for organic disordered materials [9, 12],

$$g(\varepsilon) = \frac{N}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right). \qquad (3)$$

In equation (3), σ is the characteristic energy scale of the disorder, which in ODSs is usually of the order of $\sigma \approx 0.1$ eV [9] and N is the concentration of localized states, typically between $N \simeq 10^{20}$ cm^{-3} and $N \simeq 10^{21}$ cm^{-3} [30–37]. It was recently proven that the organic semiconductors P3HT, OC$_1$C$_{10}$-PPV and quinquethiophenes (5T) possess a Gaussian DOS [24, 38], confirming the claim of Bässler [9, 12]. The corresponding model is called the *Gaussian Disorder Model* (GDM) when spatial positions and energies can be considered as independent from each other. If, in the opposite case, site positions and energies are correlated, one usually speaks of the *Correlated Disorder Model* (CDM) [39, 40]. The DOS in the CDM can be Gaussian as well. Hopping motion in the Gaussian DOS is schematically shown in figure 1.

Sometimes a purely exponential DOS is assumed in the literature [32, 41–59]:

$$g(\varepsilon) = \frac{N}{\varepsilon_0} \exp\left(\frac{\varepsilon}{\varepsilon_0}\right). \qquad (4)$$

We denote the energy scale in equation (4) as ε_0 in order to distinguish it in the following discussion from the energy scale σ of the Gaussian DOS given by equation (3).

In spite of the clearly formulated model, the development of theoretical research related to charge transport in ODSs

cannot be considered satisfactory. It is particularly unsatisfactory when we take into account the fact that comprehensive theories for the description of hopping transport in inorganic disordered systems, such as doped crystalline and amorphous inorganic semiconductors, have existed for decades and been successfully applied to the description of similar processes. Most prominent representatives of the latter systems are chalcogenide glasses, like amorphous selenium (a-Se) and amorphous semiconductors, like hydrogenated amorphous silicon (a-Si:H). The DOS function in such materials is believed to be purely exponential, as given by equation (4) [19]. The theory of hopping charge transport in the exponential DOS was developed in detail in the 1970s. This theory is, however, only rarely applied to ODSs. Most researchers working with ODSs are instead focused on equations obtained by phenomenological fitting of numerical results. Two almost non-overlapping worlds seem to coexist: The world of theory (WoT), in which researchers use well approved analytical methods developed to describe hopping transport and the world of simulations (WoS), in which computers are made responsible for developing analytical equations on the basis of numerical simulations. The aim of the current report is to discuss the state of affairs in the WoS and WoT and to highlight theoretical tools appropriate for the treatment of ODSs in the WoT.

It will be shown in section 2 that equations provided in the WoS by phenomenological fitting of numerical results, albeit very popular, can seldomly be considered as reliable. In the succeeding sections, analytical theories to describe hopping transport in application to ODSs will be discussed. This report is complementary to the recently published review articles by Tessler *et al* [22] and by Baranovskii [24]. The paper by Tessler *et al* [22] focused mainly on applications of general transport concepts to organic thin-film devices, while the review by Baranovskii [24] focused on the applications of the general theoretical concepts to account for particular dependences of the carrier mobility on the concentration of charge carriers, on temperature, on the applied electric field, on the concentration of hopping sites etc, without going into detail on the analytical theories. In the current report, we pay much attention to the details of the theoretical concepts, so that one could use the text as a guide for calculations. In particular, the *percolation theory* and the concept of the *transport energy* and their application to the *variable-range hopping* (VRH) problem are highlighted in detail. In addition, a new analytical approach to VRH based on calculations of the dissipated heat in a resistor network is described.

2. WoS: phenomenological fitting of numerical results

As indicated in the introduction, there exists a broad community of researchers who numerically simulate transport properties in ODSs for some given sets of parameters and appoint a computer to provide analytical equations suitable to describe the numerical data. The corresponding equations mostly lack any theoretical or physical basis. We shall call this community the *World of Simulations* (WoS). Let us highlight the dangers of the phenomenological fitting in the case of the

two most popular equations in the field of ODSs obtained in the framework of the GDM.

In the seminal review article by Bässler [9], the dependence of the carrier mobility μ on the strength of electric field F in the GDM with MA rates was given in the form

$$\mu(F) = \mu_0 \exp\left\{-\left(\frac{2}{3}\frac{\sigma}{kT}\right)^2\right\}$$

$$\times \exp\left\{\widetilde{C}\left[\left(\frac{\sigma}{kT}\right)^2 - \widetilde{B}\right]\sqrt{F}\right\}. \tag{5}$$

Here μ_0 is a field-independent prefactor and $\widetilde{B} = 2.25$ for $\Sigma < 1.5$ and $\widetilde{B} = \Sigma^2$ for $\Sigma \geqslant 1.5$, with Σ being the numerical parameter of the so-called *non-diagonal disorder* responsible for the variation of the intersite coupling on a lattice [9]. For a spatially random distribution of sites, this extra parameter Σ would correspond to the additional random distribution of localization parameters α. This equation has been noted as the most widely used equation in the context of organic semiconductors [22]. Usually, experimentally measured dependences $\mu(F)$ are compared to equation (5) in an attempt to determine material parameters [60–62].

A rather similar approach to determine $\mu(F)$ was used by Pasveer *et al* [33], who also numerically determined $\mu(F)$ in the GDM with MA rates on a lattice and fitted the results by the analytical formula

$$\mu(T, n, F) \approx \mu(T, n)\phi(T, F), \tag{6}$$

with $\phi(T, F)$ in the form:

$$\phi(T, F) = \exp\left\{0.44\left[\left(\frac{\sigma}{kT}\right)^{3/2} - 2.2\right]\right.$$

$$\left.\times\left[\sqrt{1 + 0.8\left(\frac{Feb}{\sigma}\right)^2} - 1\right]\right\}, \tag{7}$$

where b is the lattice constant.

It is remarkable that equations (5)–(7) are very popular in the field of ODSs in spite of the fact that these equations can hardly be correct. At small energy disorder, $\sigma < 1.5kT$, for equation (5) and $\sigma < 1.7kT$, for equation (7), when the exponents in these equations become negative, both equations predict exponential decrease of the carrier mobility with rising electric field at all field values. This result can hardly be rationalized if we take into account that the increase of the mobility at low fields was predicted in numerous transparent theoretical calculations of hopping conductivity [63–65].

In the opposite case of high energy disorder, $\sigma > 1.5kT$ for equation (5) and $\sigma > 1.7kT$ for equation (7), when the exponents become positive, both equations predict exponential increase of the mobility with rising electric field at any field. Let us consider equation (7) and show that this equation is erroneous at high electric fields. At $eFb > \sigma$ we can assume $\sqrt{1 + 0.8(eFb/\sigma)^2} \propto F$ and $\phi(T, F)$ of equation (6) therefore becomes exponentially increasing as a function of the field strength F. One should note, however, that if the inequality $Feb > \sigma$ is valid, the energy gain of a charge carrier due
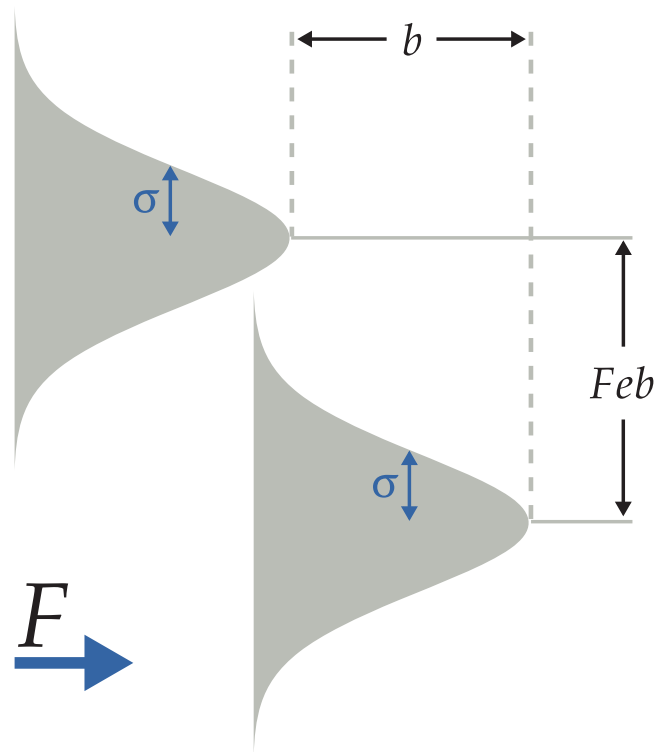
Figure 2. By hopping over a distance b along the field direction, the (positively charged) carrier gains energy Feb, which is to be compared with the energy scale of disorder σ.

to hopping transitions over the shortest possible distance b along the field direction is larger than the energy scale σ of the DOS, as illustrated in figure 2. Under these circumstances, disorder does not provide any obstacle for hopping motion with MA rates and the drift velocity $v_{\rm d}$ is field-*in*dependent. Concomitantly, the drift mobility $\mu = v_{\rm d}/F$ should decrease with rising field as $\mu \propto 1/F$ at $Feb > \sigma$. This clearly conflicts with the prediction of equation (7). Apparently, equation (7) cannot be valid at $\sigma > 1.7kT$ and $Feb > \sigma$. Notably, this invalidity is not only quantitative but also qualitative: While equation (7) predicts an exponential increase of μ with rising F, in reality μ decreases with F. The same deficiency can also be found in equation (5).

The lack of awareness of the WoS with respect to the main results obtained in the WoT for description of hopping conduction is striking. For instance, simulating the carrier mobility at different carrier concentrations n in the framework of the GDM in 2005, Pasveer *et al* [33] recognized a concentration dependence of the mobility $\mu(n)$. Although the dependence of the mobility μ on carrier concentration n in the hopping regime is a trivial effect, well-known to the scientific community for decades, it somehow appeared remarkable for researchers working with ODSs [33, 43, 46] and the GDM was renamed as the *Extended Gaussian Disorder Model* (EGDM) [66].

Let us therefore briefly reason why the hopping mobility should be inherently concentration-dependent. The transition rates of a charge carrier hopping between a pair of sites i and j given in equation (1) are valid only if site i is occupied by a charge carrier and site j is empty. In the general case, the

contribution of the pair of sites i and j to charge transport depends on the position of the Fermi level ε_F [25]:

$$\nu_{ij} = \nu_0 \exp\left(-\frac{2r_{ij}}{\alpha} - \frac{|\varepsilon_i - \varepsilon_j| + |\varepsilon_i - \varepsilon_F| + |\varepsilon_j - \varepsilon_F|}{2kT} \right).$$

(8)

As a consequence of equation (8), pairs of sites i and j play little role for charge transport when both sites are likely to be unoccupied, i.e. energies ε_i and ε_j lie far above ε_F, or both sites are blocked by charge carriers when the energies ε_i and ε_j are far below ε_F. Since the transition rates therefore depend decisively on the carrier concentration n via $\varepsilon_F(n)$, the carrier mobility μ should depend on n as well. In many experimental studies of ODSs, however, the dependence $\mu(n)$ is *not* observed at low n, which is a remarkable and unexpected result. The issue of the concentration-independent mobility is discussed in section 3.3.

With respect to ODSs, the dependence of μ on n has been recognized as an inherent property of Bässler's GDM in numerous theoretical studies [31, 67–72]. For instance, Shaked *et al* [71] clearly stated '*that within the disorder-model framework developed by Bässler and co-workers [14] it is possible to account for strong density dependence of the mobility*'. Furthermore, Schmechel [69, 70] demonstrated the ability of the GDM to provide the concentration-dependent mobilities in ODSs and applied his analysis to account for existing experimental data. Apparently being unaware of this development achieved in the WoT, researchers in the WoS promoted their numerical result as the new EGDM [33, 66].

The above analysis of equations (5)–(7), which belong to the most cited of theoretical results in the field of ODSs, shows that acting in the framework of the WoS can hardly lead to rational conclusions and that naively following its results can be dangerous. Empirical fitting of numerical data not supported by solid theoretical concepts can be valid in some restricted range of parameters, but not in general. Some dependences of transport coefficients on material parameters were considered in the WoS as universal simply because other parameters were not changed in course of simulations. Therefore we dedicate the rest of this report to theoretical concepts developed in the WoT for description of charge transport in the hopping regime.

Application of equations (6) and (7) can have another, not just academic, consequence. As announced in the recent feature article by Coehoorn and Bobbert [66], such phenomenological equations are the basis for the commercially available OLED simulation software tools (Simulation software SETFOS3.2, product of Fluxim (www.fluxim.com); Simulation software SimOLED3.x, product of Sim4tec (www.sim4tec.com)). The accuracy of the commercially available software is, however, out of the scope of our report.

3. WoT: general remarks on transport regimes

Before diving into details of the analytical tools developed to study hopping charge transport in disordered materials, let us first discuss some characteristic regimes of the long-range hopping transport and address the decisive properties of the energy spectrum of the ODSs. In particular, we will discuss in the following sections when hopping transport can be treated as nearest-neighbour hopping (NNH) and when the more general regime of the variable-range hopping (VRH) should be considered. Furthermore, under which circumstances one can model charge transport on regular lattice grids and what influence the shape of the DOS has on the transport properties, will be discussed.

3.1. Nearest-neighbour or variable-range hopping?

Due to the dependences of the hopping rates given by equations (1) and (2) on the intersite distances r_{ij} and on site energies ε_i, ε_j, hopping conduction is in general a variable-range hopping (VRH) process. At low temperatures, as compared to the energy scale of disorder, energy factors in the exponents of equations (1) and (2) play an essential role and carriers mostly tunnel to spatially remote sites in order to optimize the activation in energy. The higher T is, the less important the energy-dependent factors become and the closer in space the carrier transitions are, as prescribed by the tunneling probability $\exp(-2r_{ij}/\alpha)$ in equations (1) and (2). In the limit of sufficiently high temperatures, when $kT > \sigma$ in equation (3) or $kT > \varepsilon_0$ in equation (4), only spatial factors determine the hopping rates and the VRH regime turns into the nearest-neighbour hopping (NNH) regime, in which carriers mostly tunnel to the spatially nearest sites [73].

Sometimes it is argued that the VRH cannot serve as the appropriate transport mechanism in ODSs consisting of large molecules, as illustrated in figure 3(*a*). At first glance, it seems that only the NNH is possible in such media, because the nearest neighbours block the electron's way to more distant molecules, preventing the long-distance jumps. If so, the localization sites are often placed on a regular lattice grid and it is assumed that hopping transitions are possible only between the nearest neighbours even at low temperatures, $kT \ll \sigma$ [74, 75].

Below, we show that these naive arguments against the VRH are irrelevant. The nearest neighbours do not block distant jumps. Indeed, they even enhance the probability of such jumps due to the broader spread of the electron wavefunctions. As a result of such spreading, the effective localization length in a dense array of molecules turns out to be larger than that in a single-molecule case.

Let us prove this effect for the simplest model: the one-dimensional model of Anderson localization with equally spaced potential wells, whose energy levels are randomly distributed (figure 3(*b*)). Let D be the distance between the centers of the wells, σ—the characteristic scatter of the energy levels and τ—the hopping integral (tunneling integral) between the adjacent wells. We consider only the case of a strong Anderson localization, where $\tau \ll \sigma$. Wells are numbered from left to right, so that the coordinate of the nth well is $x_n = Dn$.

According to the tight-binding approach, the electron wavefunction is characterized by the set $\{\ldots, a_{-1}, a_0, a_1, a_2, \ldots\}$ of amplitudes (expansion coefficients) related to the corresponding wells (figure 3(*c*)). Let us consider the localized
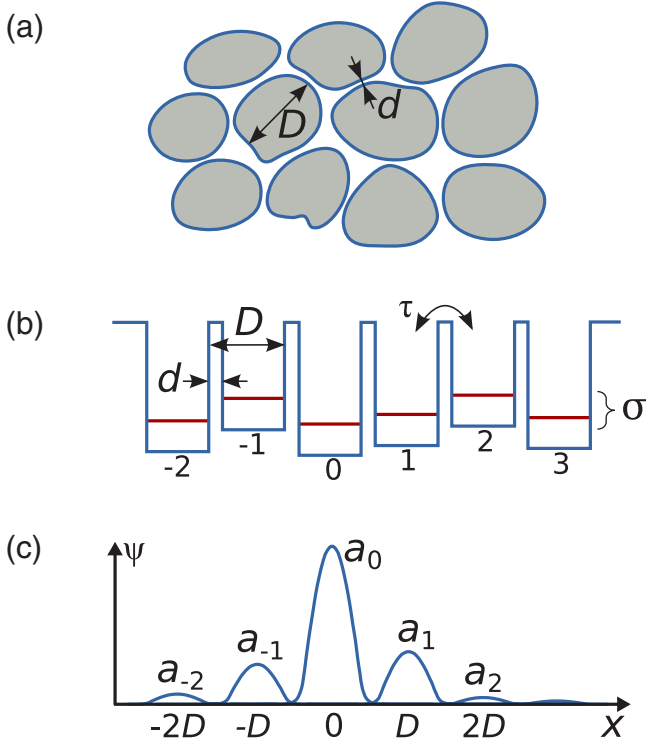
Figure 3. Dense array of organic molecules: (*a*) schematic representation; the molecule diameter D is large compared with the intermolecular distance d; (*b*) one-dimensional model of potential wells with energy disorder (Anderson model); (*c*) typical shape of the electron wavefunction in a strongly localized state; amplitudes a_n are related to probabilities p_n of finding the electron in corresponding wells as $|a_n|^2 = p_n$.

state centered at the 0th well, so that the amplitude a_0 is the largest one. It is easy to see (e.g., from the first-order perturbation theory) that

$$a_{\pm 1} \approx a_0 \frac{\tau}{\Delta\varepsilon_{0,\pm 1}} , \qquad (9)$$

where $\Delta\varepsilon_{mn}$ is the energy difference between the mth and the nth sites [73]. The typical values of this difference are of the order of σ and therefore

$$\frac{|a_{\pm 1}|}{|a_0|} \simeq \frac{\tau}{\sigma} . \qquad (10)$$

Analogous consideration shows that this estimate holds also for more distant wells:

$$\frac{|a_{\pm(n+1)}|}{|a_{\pm n}|} \simeq \frac{\tau}{\sigma} \quad \text{for} \quad n \geqslant 0 . \qquad (11)$$

Applying equation (11) repeatedly, one can see that

$$|a_n| \simeq |a_0| \left(\frac{\tau}{\sigma}\right)^{|n|} , \qquad (12)$$

i.e. amplitudes a_n decay exponentially with distance from the central (0th) well. This decay can be expressed as

$$|a_n| \simeq |a_0| \exp\left(-\frac{|x_n|}{\alpha_{\text{eff}}}\right) , \qquad (13)$$

where $x_n = Dn$ is the position of the nth well (with respect to the well at which the wavefunction is centered) and α_{eff} is the effective localization radius. The value of α_{eff} can be obtained by comparison between equations (12) and (13):

$$\alpha_{\text{eff}} = \frac{D}{\ln(\sigma/\tau)} . \qquad (14)$$

Now let us return to the dense three-dimensional array of large molecules. One can expect that equation (14) is valid in this case at least qualitatively, since there is nothing special to the one-dimensional Anderson model in the considerations above. (In the three-dimensional case, one should consider D as the diameter of the molecules.) One can conclude from equation (14) that α_{eff} and D have the same order of magnitude (unless τ is smaller than σ by many orders of magnitude), which makes variable-range hopping possible. In a localized state, the wavefunction penetrates not only into the adjacent molecules, but also into the more remote ones. Therefore, hops between the non-nearest-neighbouring molecules contribute to the hopping transport.

Let us express the effective localization length α_{eff} via the bare localization length in the barrier, α_0. It is obvious that the tunneling integral τ depends on the barrier thickness d as $\propto \exp(-d/\alpha_0)$. The preexponential factor can be roughly estimated quasiclassically [76, section 50], yielding

$$\tau \simeq \frac{\Delta E}{2\pi} \exp\left(-\frac{d}{\alpha_0}\right) , \qquad (15)$$

where ΔE is the typical distance between the energy levels in a single molecule. Substituting equation (15) into equation (14), one obtains the relation between the bare and the 'effective' localization radii, α_0 and α_{eff}:

$$\alpha_{\text{eff}} \approx \frac{D}{d/\alpha_0 - \ln(\Delta E/2\pi\sigma)} . \qquad (16)$$

Here, both d (spacing between the molecules) and α_0 are of the order of several Ångströms, hence d/α_0 is not a large number. The same is true for $\ln(\Delta E/2\pi\sigma)$. Therefore, one can see again, that α_{eff} has the order of D (molecule diameter), rather than of α_0.

The arguments above can be viewed as a justification of the dominating role of the VRH in organic semiconductors, both for small and large molecules as compared to the intermolecular distance. It is therefore important to account for the VRH regime when dealing with ODSs in theoretical studies.

3.2. When are lattice models appropriate?

Very often, studies of charge transport, particularly in the WoS, are performed on regular lattice grids, although the characteristic property of disordered materials is not only the energetic, but also the *spatial* disorder. Such studies can well mimic the VRH regime as long as the decisive temperature-dependent length of the hops in the VRH essentially exceeds the lattice constant of the grid and hops to remote lattice sites are included [9]. However, studies on grids make little sense
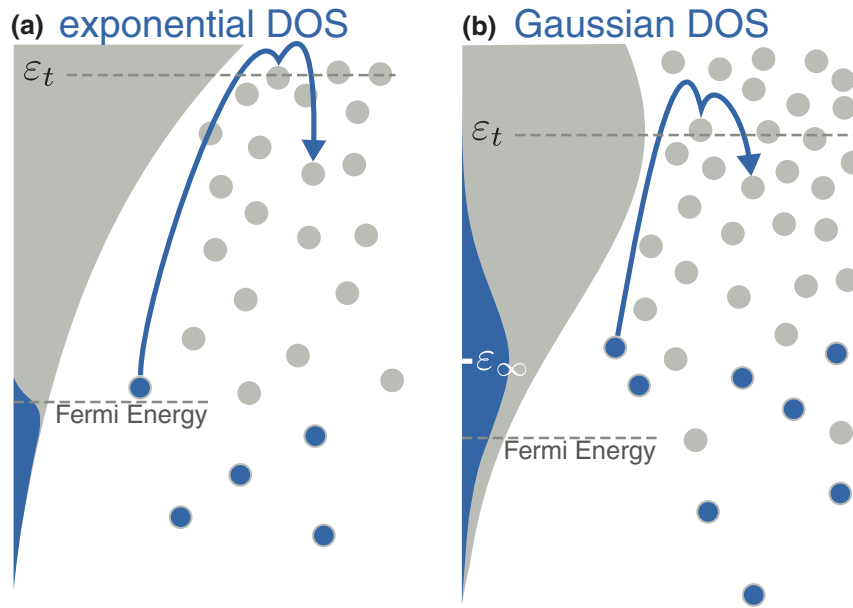
Figure 4. Schematic picture of the equilibrium carrier distribution in the exponential (*a*) and Gaussian (*b*) DOS. Occupied states are marked blue.

if carriers are artificially forced only to hop to the nearest neighbours in simulation algorithms or in analytical studies. It is particularly erroneous at low temperatures, when the characteristic hopping length in reality exceeds the nearest-neighbour spacing (the VRH regime), while the carriers in the modeling are forced to move only via short hops to the nearest neighbours on the grid [74, 75]. Constraining the model to the nearest-neighbour transitions on a lattice is also not correct in the case of high temperatures, when charge transport occurs in the NNH regime, since the results of the modeling depend essentially on the choice of the particular lattice structure [74, 75], while the real disordered media can hardly recall a regular lattice grid. Therefore, even in the case of the NNH when the typical hopping length is of the order of the inter-site distance, one should always consider spatially random systems.

3.3. Gaussian DOS against exponential DOS

If the thermal energy kT is smaller than the energy scale of disorder σ, which is the case for ODSs, where at room temperature $kT \approx 0.3\sigma$ [9, 24], charge transport occurs in the VRH regime, i.e. the interplay between the spatial and energy factors determines the transport path. Under these circumstances, the shape of the energy spectrum plays the crucial role for charge transport and the decisive task is to determine the DOS for the underlying material. Exponential DOS given by equation (4) or Gaussian DOS given by equation (3) are most often used in the literature for ODSs and it is claimed sometimes [41] that there is no principle difference between the steeply energy-dependent DOS functions in their ability to describe charge transport, particularly to account for the dependence of carrier mobility μ on the concentration of carriers n. Below we show that, on the contrary, the shape of the dependence $\mu(n)$ for the exponential DOS given by equation (4) is not only quantitatively, but also

qualitatively different to that for the Gaussian DOS given by equation (3). Herewith we claim that studying the dependence $\mu(n)$ experimentally one can distinguish between these two possible shapes of the DOS.

In order to understand this message, one should clarify what is different and what is similar for charge transport in the case of Gaussian DOS as compared to the case of exponential DOS. As will be shown in sections 5 and 6, transport in both cases is dominated by thermally activated transitions of charge carriers towards the energy level ε_t situated far above the energy region occupied by carriers. This feature is qualitatively similar for all steeply energy-dependent DOS functions. The position of the level ε_t, traditionally called the *transport energy*, depends slightly on the shape of the DOS. What is, however, qualitatively different in the case of Gaussian DOS, as compared to the case of the exponential DOS, is the kind of thermal distribution of carriers along the energy spectrum of the system.

The distribution of carriers in thermal equilibrium is described by the product $g(\varepsilon)f(\varepsilon)$, where $g(\varepsilon)$ is the DOS and $f(\varepsilon)$ is the Fermi distribution,

$$f(\varepsilon) = \left[1 + \exp\frac{(\varepsilon - \varepsilon_{\mathrm{F}})}{kT} \right]^{-1}. \qquad (17)$$

In the exponential DOS given by equation (4), the maximum of the carrier distribution $g(\varepsilon)f(\varepsilon)$ at $kT < \varepsilon_0$ is in close vicinity of the concentration-dependent Fermi energy $\varepsilon_{\mathrm{F}}(n)$, as illustrated schematically in figure 4(*a*). Mobility, determined by carrier activation from $\varepsilon_{\mathrm{F}}(n)$ towards ε_t is therefore always a function of n, as schematically illustrated by the dashed line in figure 6.

In the Gaussian DOS, on the contrary, most carriers at low n are distributed not around the Fermi level, but are instead situated around the so-called *equilibration energy* ε_∞. This

energy is usually identified with the average carrier energy determined as [9, 77]

$$\varepsilon_\infty = \frac{\int_{-\infty}^{\infty} \varepsilon g(\varepsilon) \exp(-\varepsilon/kT) \mathrm{d}\varepsilon}{\int_{-\infty}^{\infty} g(\varepsilon) \exp(-\varepsilon/kT) \mathrm{d}\varepsilon} = -\frac{\sigma^2}{kT}. \quad (18)$$

Here we follow the traditional definition of ε_∞ via equation (18) [9], since for the Gaussian DOS given by equation (3) the maximum of the product $g(\varepsilon)f(\varepsilon)$ corresponds to ε_∞ given by equation (18). This situation, with most carriers occupying in thermal equilibrium energy levels far above the Fermi level, is counter-intuitive. It occurs because the number of states in the Gaussian DOS increases above the Fermi level at a much steeper rate than the decrease of the occupation probability described by the Fermi function. The distribution of carriers in the Gaussian DOS at low n is illustrated schematically in figure 4(b). In such a case, transport in the Gaussian DOS is due to carrier transitions via energy levels in the range between ε_∞ and ε_t. Neither of these energies depend on the concentration of carriers n. Furthermore, the occupation numbers of states in the range between ε_∞ and ε_t is very low, $f(\varepsilon) \ll 1$ and the carriers can be considered as independent from each other (filling of states in the range between ε_∞ and ε_t can be neglected). Therefore, the carrier mobility in the Gaussian DOS at low n does not depend on n, which is in striking contrast to the situation in the exponential DOS, where carrier mobility depends on n at all n values.

This picture with most carriers occupying energy levels around ε_∞ in the Gaussian DOS is valid at low n, when the Fermi energy $\varepsilon_F(n)$ is far below the equilibration energy ε_∞. With rising n, the Fermi energy $\varepsilon_F(n)$ moves upwards and it crosses ε_∞ at a carrier concentration n_c determined by the equation [67, 69, 72]

$$\varepsilon_F(n_c) = \varepsilon_\infty, \quad (19)$$

where the Fermi energy $\varepsilon_F(n)$ is determined via the relation

$$\int_{-\infty}^{\infty} g(\varepsilon) f(\varepsilon, \varepsilon_F) \mathrm{d}\varepsilon = n. \quad (20)$$

In figure 5, the crossing between ε_F and ε_∞ with varying n is illustrated schematically [78].

Since both ε_∞ and ε_F depend on temperature T, the criterion given by equation (19) can also be treated as a function of T rather than of the carrier concentration:

$$\varepsilon_F(T_c, n) = \varepsilon_\infty(T_c). \quad (21)$$

In fact, this temperature-dependent form was suggested first [67] and only later reformulated in the form of equation (19) [38, 69, 72].

The dependence $\mu(n)$ is shown schematically in figure 6 for the exponential and Gaussian DOS. It is clearly visible that the mobility in the Gaussian DOS is concentration-*in*dependent at low concentrations $n < n_c$ and becomes dependent on n at $n > n_c$, whereas the carrier mobility in the exponential DOS depends on n at all concentrations. A concentration-independent mobility at low n was observed experimentally for
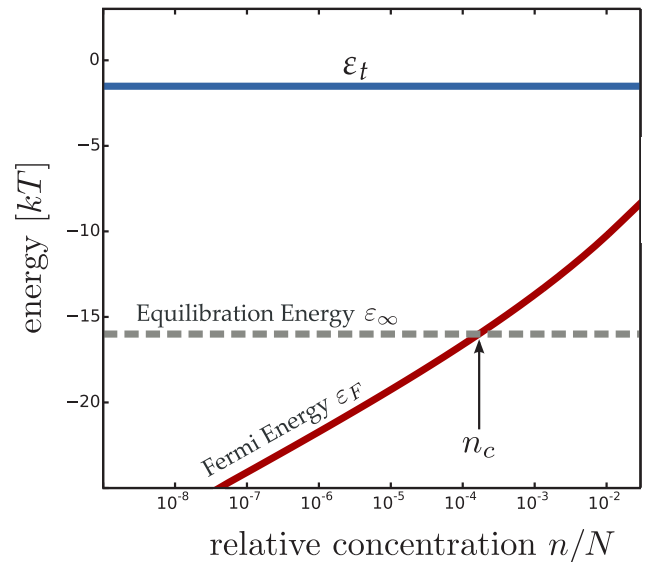
Figure 5. Positions of the concentration-dependent Fermi energy ε_F given by equation (20) and of the equilibration energy ε_∞ given by equation (18) at $\sigma/kT = 4$. Position of the transport energy ε_t is calculated via equation (60).
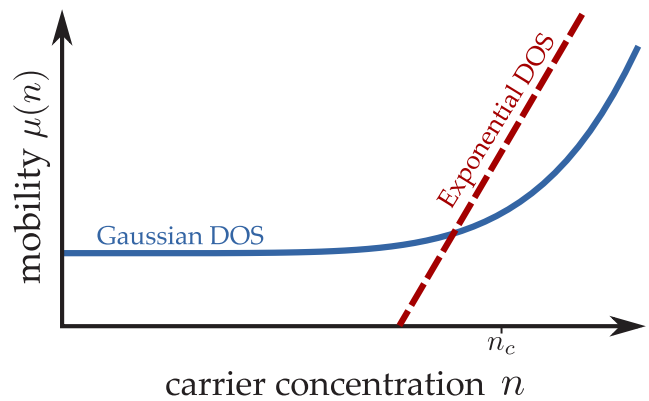
Figure 6. Schematic dependences of the carrier mobility on concentration of carriers n in a Gaussian DOS (solid line) and in an exponential DOS (dashed line). Concentration n_c corresponds to the condition given by equation (19).

the widely studied ODSs PPV and P3HT [43, 79], which is a strong indicator for these materials possessing a DOS steeper than a purely exponential one. Mobility independent of the concentration n at low n was also found for three derivatives of quinquethiophenes (5T) [80, 81]. It is worth emphasizing once more that the independence of the carrier mobility on n at low carrier concentrations n clearly rules out the exponential DOS as a candidate for such materials, whereas the measured data fit perfectly into the GDM picture [19, 38, 67, 69, 72, 78].

Note, that ε_∞ can be found in any DOS decaying steeper than purely exponentially and the threshold concentration n_c can be calculated via equation (19). Concomitantly, the plateau of $\mu(n)$ at $n < n_c$ is to be found for any of these steep DOS functions. Moreover, the value of n_c drastically depends on the shape of the DOS, as shown in the detailed study by Oelerich *et al* [38]. Comparison between n_c from equation (19) and the experimental data for PPV and P3HT from [43] indicates that

these materials possess a DOS very close to the Gaussian one. Note also, that equation (19) is not only valid in the GDM, but should hold for systems with space-energy correlations, i.e. for the CDM.

In the following sections, we present several theoretical tools to describe hopping charge transport in ODSs.

4. WoT: percolation theory

The most successful theoretical approach among those developed so far to describe the VRH and NNH in disordered systems is the *percolation theory*. Therefore we start exploring the WoT by discussing this particular approach.

The basis for applying percolation theory to analyze hopping transport is the concept of the *resistance network* suggested by Miller and Abrahams [25]. In this concept, each site is modeled by a node of an electrical circuit and for each pair of sites (i, j), between which carrier jumps are possible, a resistance R_{ij} connecting the corresponding nodes is introduced. Then, the resistance of the whole sample can be found as the resistance of the equivalent network constructed from the elements R_{ij}. The values of R_{ij} are defined as [73, 82]

$$R_{ij} = \frac{kT}{e^2 \Gamma_{ij}}, \qquad (22)$$

where $\Gamma_{ij} = \nu_{ij} f(\varepsilon_i)[1 - f(\varepsilon_j)]$ is the equilibrium rate of carrier transitions from site i to site j. According to the detailed balance principle, $\Gamma_{ij} = \Gamma_{ji}$. For the case of MA hopping rates, the values of Γ_{ij} are given by equations (8) and (17).

It is crucial for evaluating the network resistance that the magnitudes of R_{ij} are distributed in an exponentially broad range. (This can be easily seen from equation (8), keeping in mind that the scatter of hop lengths r_{ij} is, typically, large as compared to the localization length α and the distribution of site energies has a width much larger than kT.) In such conditions, very large resistances do not contribute to transport because they can be bypassed through paths with much lower resistances. On the other hand, resistances of very small values serve simply as short-cuts between larger resistances, hence their exact values are not important. These simple arguments give a hint that there is such a 'critical' value R_{crit} that pairs of sites with $R_{ij} \gg R_{crit}$ and with $R_{ij} \ll R_{crit}$ do not play a decisive role for the overall conductivity of the network.

Percolation theory provides a precise meaning for R_{crit} via the following arguments [73]. Let us sequentially delete resistors from the network, starting from the largest one and proceeding in the order of descending resistances. At some 'critical' step in this procedure, deleting the next resistor disconnects the opposite sides of the sample. Then, R_{crit} is, by definition, the resistance of the element deleted at this 'critical' step, taken in the limit of an infinitely large sample. The corresponding value R_{crit} determines the main (exponential) factor in the macroscopic resistivity [73]. As we will see below, in the most important cases the evaluation of R_{crit} can be reduced to purely geometrical problems of the percolation theory. Also the slowly varying pre-exponential factor, which

appears in resistivity along with R_{crit}, can be estimated on the basis of the percolation theory [73].

In section 4.1 we will apply the percolation approach to the simplest model of isoenergetic sites. Then, in sections 4.2–4.4 we will consider the low-temperature conductivity described by the Mott law and the transport in systems with exponential and Gaussian DOS. If not specified otherwise, we will assume the MA transition rates given by equation (8).

4.1. Percolation theory for the NNH with isoenergetic sites

Let us consider hopping transport via localized states, which are randomly distributed in 3D space and possess equal carrier energies. We assume strong localization, implying that the localization length α is small as compared to the typical distance between sites $N^{-1/3}$, where N is the concentration of sites. In real systems, this model can be relevant at high enough temperatures, when kT is larger than the characteristic energy scale of the DOS. In this case, elementary resistances R_{ij} differ from each other only due to the different distances r_{ij} between sites: $R_{ij} = R_0 \exp(2r_{ij}/\alpha)$, where the prefactor R_0 is the same for all pairs of sites. We will consider R_0 as constant while varying the site concentration, which means keeping the fraction of occupied sites unchanged.

Finding the critical resistance R_{crit} is reduced to the following purely geometrical problem. Let us consider a pair of sites as connected if the distance between the sites is less than some fixed value r. When r is large enough, connected pairs form an *infinite cluster*, i.e., there exists a path over the whole structure passing only through connected pairs of sites. At small values of r, only finite clusters of connected sites can exist. Let r_c be the smallest value of r at which the infinite cluster still exists. Then,

$$R_{crit} = R_0 \exp(2r_c/\alpha). \qquad (23)$$

The quantity r_c depends on the concentration of sites N. It is convenient to express r_c through a dimensionless parameter B_c—an average number of sites within the sphere of radius r_c around a given site [73, 83, 84]:

$$\frac{4\pi}{3} N r_c^3 = B_c. \qquad (24)$$

As a dimensionless quantity, B_c does not depend on N; thus, its value is universal: $B_c \cong 2.735$ [85]. Since r_c does not depend on temperature, the NNH regime is valid for this model of conduction.

Let us consider the dependence of the dc conductivity σ_{dc} through isoenergetic sites on their concentration N. This dependence is exponentially strong, because the concentration-dependent value r_c appears in the exponent in equation (23). From equations (23)–(24) one can express σ_{dc} as follows [73]:

$$\sigma_{dc}(N) \propto R_{crit}^{-1} \propto \exp\left(-\frac{2r_c}{\alpha}\right) \equiv \exp\left(-\frac{\gamma N^{-\frac{1}{3}}}{\alpha}\right), \quad (25)$$

where $\gamma = 2(3B_c/4\pi)^{1/3} \cong 1.7351$ according to equation (24).

Equation (25) provides a solid basis for the understanding and quantitative analysis of the concentration dependence of

σ_{dc} in systems with low concentration of localization sites ($N^{-1/3} \gg \alpha$) at high temperatures. Notably, equation (25) shows that the conductivity monotonously decreases with decreasing N. Hence, there is no such threshold concentration N_c (often assumed in ODSs) that σ_{dc} strictly vanishes at $N < N_c$.

It is possible to go beyond equation (25) and to estimate also the non-exponential factors in the expression for σ_{dc}. The key point is that adding pairs with $R_{\text{crit}} < R_{ij} \leqslant R_{\text{crit}} \exp(\delta)$, where $\delta \simeq 1$, to the percolation cluster gives the major contribution to the conduction [73]. The infinite cluster formed by this set of pairs, the so-called 'critical subnetwork' [73], or 'fat percolation cluster' [86], determines the conductivity. An important characteristic of the critical subnetwork is its correlation length L_{corr}, which possesses the following feature: on length scales larger than L_{corr} the subnetwork can be considered homogeneous, whereas on smaller scales it is highly inhomogeneous [73]. As explained in section 5.6 of the monograph by Efros and Shklovskii [73], one can estimate the conductivity σ_{dc} of the resistor network in the 3D case via the relation

$$\sigma_{\text{dc}} \simeq R_{\text{crit}}^{-1} L_{\text{corr}}^{-1}. \qquad (26)$$

The correlation length, in turn, depends on the fraction ϵ of pairs with resistances above R_{crit}, in the manifold of the resistors added in order to form the critical subnetwork. When ϵ is small, L_{corr} is proportional to $\epsilon^{-\nu}$, where ν is the universal critical index of the correlation length of the percolation cluster. In 3D, $\nu = 0.875 \pm 0.008$ [87]. Since in the system of isoenergetic sites $\epsilon \simeq \alpha/r_c \simeq \alpha N^{1/3}$, the correlation length of the critical subnetwork depends on α as

$$L_{\text{corr}} \propto (\alpha N^{1/3})^{-\nu}. \qquad (27)$$

Finally, collecting equations (23)–(27) one can obtain the following expression for the conductivity σ_{dc} in 3D case:

$$\sigma_{\text{dc}}(N) = A R_0^{-1} N^{\frac{1}{3}} (\alpha N^{\frac{1}{3}})^{\nu} \exp\left(-\frac{\gamma N^{-\frac{1}{3}}}{\alpha}\right), \qquad (28)$$

where $\gamma \approx 1.735$, $\nu \approx 0.875$ and A is a numerical coefficient of order unity. The dependence $\sigma_{\text{dc}}(N)$ over several orders of magnitude has been observed in numerous experimental studies of hopping transport via impurity atoms in doped crystalline semiconductors, where the values of the localization length α for randomly placed impurities are well known [73]. The prefactor R_0 may depend on hopping distance. If so, one should use the value of R_0 at r_c.

4.2. Percolation theory for the VRH, Mott law

Let us now turn to considering the VRH transport regime, which, as explained in section 3.1, comes in to play when kT is smaller than the characteristic width of the DOS. Analysis of the VRH conduction in the framework of the percolation theory is more sophisticated than in the NNH case, because the dependence of transition rates on site energies also turns out to be essential, along with the dependence of transition

rates on the intersite distances. One can consider the charge transport in these conditions as percolation in 4D space, whose coordinates are three spatial axes and the energy. In this context, an important question arises: is it still possible to analyze the charge transport in the VRH regime in terms of the average number of bonds B_c given by equation (24), or is it necessary to go beyond this simple approach?

In the current subsection we focus on the temperature dependence of the conductivity at very low temperatures, when transport occurs only in close vicinity of the Fermi level ε_F. As explained in [73], this dependence can be found as a solution of a 4D percolation problem yielding the prominent Mott law. The corresponding 4D percolation problem looks as follows. Consider a set of points randomly placed in 3D space with some given concentration. Each point is characterized by its dimensionless position vector $\tilde{\mathbf{r}}_i$ and a dimensionless 'energy' $\tilde{\varepsilon}_i$. The latter quantity is uniformly distributed in the range $[-1, 1]$. Two points i and j are considered as connected if

$$\frac{|\tilde{\varepsilon}_i| + |\tilde{\varepsilon}_j| + |\tilde{\varepsilon}_i - \tilde{\varepsilon}_j|}{2} + |\tilde{\mathbf{r}}_i - \tilde{\mathbf{r}}_j| \leqslant 1. \qquad (29)$$

The problem is to find the minimal concentration $\tilde{n}$, at which the connected system of points forms an infinite percolation cluster. Below we will assume that the concentration of points is fixed at this threshold value $\tilde{n}$.

Let us further introduce a spatial scale D and an energy scale W, connected to each other as

$$\frac{W}{kT} = \frac{2D}{\alpha} \qquad (30)$$

(the exact values of D and W will be determined later). One should keep in mind that each point i is associated with a site in a real sample, located at

$$\mathbf{r}_i = D \tilde{\mathbf{r}}_i \qquad (31)$$

and having the energy

$$\varepsilon_i = W \tilde{\varepsilon}_i + \varepsilon_F. \qquad (32)$$

These sites are uniformly distributed within the energy range $[\varepsilon_F - W, \varepsilon_F + W]$ and their concentration is equal to $\tilde{n}/D^3$. The density of states around the Fermi level is therefore

$$g(\varepsilon_F) = \frac{\tilde{n}}{2W D^3}. \qquad (33)$$

The connection criterion (29) corresponds to the following relation between sites i and j:

$$\frac{|\varepsilon_i - \varepsilon_F| + |\varepsilon_j - \varepsilon_F| + |\varepsilon_i - \varepsilon_j|}{2kT} + \frac{2r_{ij}}{\alpha} \leqslant \frac{W}{kT}. \qquad (34)$$

Using equations (8) and (22), one can rewrite this criterion as

$$R_{ij} \leqslant \frac{kT}{e^2 \nu_0} \exp\left(\frac{W}{kT}\right). \qquad (35)$$

Since the criterion given by equation (29) provides the percolation threshold by construction, the same is true for equation (35), i.e. pairs of sites obeying equation (35) form

the least dense infinite cluster. Hence, the critical value of the resistance R_{crit} is equal to the r. h. s. of equation (35). (Note that, for finding R_{crit}, it is sufficient to take into account only sites with energies in the range $[\varepsilon_F - W, \varepsilon_F + W]$, because all other sites do not satisfy the inequality (34).) Hence, the conductivity σ_{dc} scales as

$$\sigma_{dc} \propto R_{crit}^{-1} \propto \exp(-W/kT), \tag{36}$$

where W is the half-width of the energy strip around ε_F, in which conduction takes place. The value of W can be found by resolving the system of equations (30) and (33) with respect to W and D:

$$W = kT \left[(T_0/T)^{1/4} \right], \tag{37}$$

where $T_0 = \beta/[kg(\varepsilon_F)\alpha^3]$ and $\beta \equiv 4\tilde{n}$ is a numerical factor that can be obtained from the solution of the dimensionless percolation problem described above. The energy scale for the most important hops decreases with decreasing temperature ($W \propto T^{3/4}$) and simultaneously the spatial scale increases ($D \propto T^{-1/4}$). This indicates the VRH character of the conduction: at smaller temperatures, the carrier chooses more distant hops in order to reduce the difficulty of overcoming the energy barrier.

Equations (36) and (37) give the prominent Mott law [88] for the temperature dependence of hopping conductivity:

$$\sigma_{dc} = \sigma_{dc}^o \exp[-(T_0/T)^{1/4}], \tag{38}$$

where the prefactor σ_{dc}^o only weakly (non-exponentially) depends on temperature. It is worth emphasizing that the percolation approach is capable of evaluating this prefactor, up to a constant factor of the order unity. One can find the details in section 9.4 of [73]. The above consideration shows that the Mott law (38) is valid when the DOS can be considered as constant in the energy range $\varepsilon_F \pm W$.

4.3. Percolation theory for the VRH in the case of exponential DOS

As mentioned in section 1, as early as the 1970s and 1980s, a comprehensive theory of the VRH was developed for systems with the purely exponential DOS given by equation (4) and we first focus on this shape of the DOS. For the exponential DOS, the Mott law is valid at temperatures satisfying the strong inequality $kT(T_0/T)^{1/4} \ll \varepsilon_0$. Using the estimate $g(\varepsilon_F) \approx n/\varepsilon_0$ for concentration of carriers n, one obtains the condition $kT/\varepsilon_0 \ll (n\alpha^3)^{1/3}$ for the validity of the Mott's law.

In 1979, Grünewald and Thomas [89] used percolation theory to study the VRH in the exponential DOS at temperatures above the validity range of the Mott law and discovered a new regime, in which the carrier mobility is described as

$$\mu = \frac{\sigma_{dc}}{en} = \frac{\sigma_{dc}^0}{en} \exp\left[-\frac{\varepsilon^* - \varepsilon_F(n, T)}{kT} \right], \tag{39}$$

where σ_{dc}^0 is the preexponential factor in the expression for the conductivity and ε^* is the characteristic highest energy of sites

contributing to charge transport [89]. Unlike the Mott law, in this new regime the energy ε^* is fully detached from the Fermi level and does not depend on n.

Two decades later, the very same problem was addressed by the community working with ODSs [41]. While Grünewald and Thomas [89] in 1979 used the classical percolation approach [82, 90, 91], Vissenberg and Matters in 1998 used a 'somewhat different approach' [41] to that by Grünewald and Thomas. It is instructive to find out what was different and which of the two approaches gives more accurate results.

According to the percolation approach one first rewrites the rate given by equation (8) in the form

$$\nu_{ij} = \nu_0 \exp\left(-\xi_{ij} \right) \tag{40}$$

and determines $B(\varepsilon_i, \xi)$, the average number of sites accessible from a site with energy ε_i using hopping rates with the exponent ξ_{ij} smaller than some value ξ. The aim is to find out the magnitude of the exponent ξ_c that provides the infinite percolation path for carriers through the system of sites using only rates with $\xi \leqslant \xi_c$. In order to determine ξ_c, it is necessary to formulate the percolation criterion that relates the quantity $B(\varepsilon_i, \xi)$, appropriately averaged over site energies, to some numerical factor given by the percolation theory [82, 90, 91].

In a first approximation one might perform the averaging over site energies ε_i ascribing equal weights for all sites in the range $\varepsilon_F - kT\xi_c \leqslant \varepsilon_i \leqslant \varepsilon_F + kT\xi_c$ leading to the equation

$$B_c = \frac{\displaystyle\int_{\varepsilon_F - kT\xi_c}^{\varepsilon_F + kT\xi_c} d\varepsilon\, g(\varepsilon) B(\varepsilon, \xi_c)}{\displaystyle\int_{\varepsilon_F - kT\xi_c}^{\varepsilon_F + kT\xi_c} d\varepsilon\, g(\varepsilon)}, \tag{41}$$

where B_c is equal to that in equation (24). It is this way of averaging that was used by Vissenberg and Matters [41] for the exponential DOS (and later by Coehoorn *et al* [92] for the Gaussian DOS).

Close inspection performed by Pollak [91] and by Overhof [93] shows, however, that it is necessary for the averaging to use a weighting factor that is proportional to the average number of bonds for a given site energy ε, which leads to the percolation criterion [91, 93, 94]

$$B_c = \frac{\displaystyle\int_{\varepsilon_F - kT\xi_c}^{\varepsilon_F + kT\xi_c} d\varepsilon\, g(\varepsilon) B^2(\varepsilon, \xi_c)}{\displaystyle\int_{\varepsilon_F - kT\xi_c}^{\varepsilon_F + kT\xi_c} d\varepsilon\, g(\varepsilon) B(\varepsilon, \xi_c)}. \tag{42}$$

It is this way of averaging that was applied by Grünewald and Thomas [89] for the exponential DOS (and later by Baranovskii *et al* [67] for Gaussian DOS).

Remarkably, one can check which of the two percolation criteria is more accurate for the case of the exponential DOS. The exponential DOS, although probably not relevant for all ODSs [24, 38], is still useful as a test field for analytical and numerical methods, because its VRH problem can be solved exactly. Nenashev *et al* [95] recently showed that the four-dimensional (three spatial dimensions and energy) VRH problem in the exponential DOS can be mapped onto
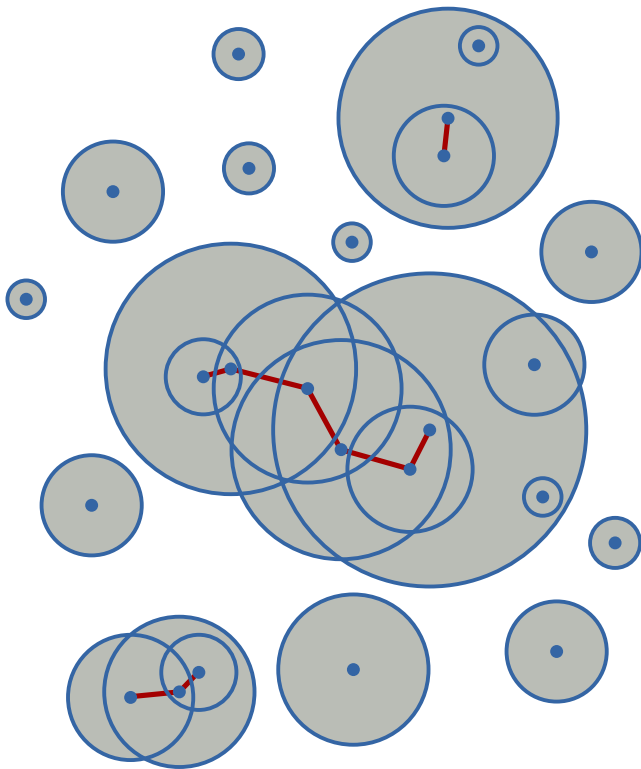
Figure 7. The system of spheres with exponentially distributed radii. Two spheres are considered connected, if the center of each of them lies within the other sphere. Paths between connected spheres are marked as red lines.

a purely geometrical problem of percolation via spheres with exponentially distributed radii. The latter problem is solvable exactly and provides not only the exponential dependence of the conductivity on system parameters, but even the pre-exponential factors with much weaker dependences [95]. This mapping is possible at the reasonable condition $\varepsilon_0 > kT$. In this case, one can show that the percolation problem in a 3D system of randomly placed spheres with radii distributed via $g(r) = N \exp(-\tilde{r})$ ($\tilde{r}$ being a dimensionless radius) represents the 4D VRH problem. This percolation problem is schematically shown in figure 7. The critical concentration $\widetilde{N}_c$ of spheres in this 3D problem is accessible numerically with high precision, yielding the result $\widetilde{N}_c \approx 0.219$ [95]. This result can now be converted back into an accurate solution of the initial VRH problem.

The comparison between the results of Grünewald and Thomas [89], those obtained by Vissenberg and Matters [41] and the exact solution by Nenashev *et al* [95] is shown in figure 8. In addition, the figure contains a solution based on the transport energy (TE) concept, as suggested by Oelerich *et al* [38] (this approach is described in detail in section 5). All the analytical solutions shown in figure 8 can be presented in the form

$$\mu = \mathbf{A} \frac{v_0 e}{n \varepsilon_0 \alpha} \left[\mathbf{B} \frac{n}{8} \left(\frac{\alpha \varepsilon_0}{kT} \right)^3 \right]^{\varepsilon_0/kT}, \qquad (43)$$

with **A** and **B** being slowly varying or constant factors that differ between the solutions. The values of **A** and **B** are listed in table 1. The exponential dependences on the system
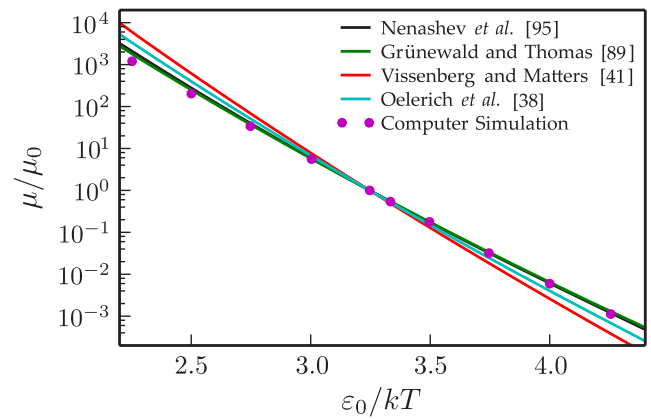

Figure 8. Normalized mobility values obtained from computer simulation and from equation (43) for the different analytical solutions, at $N^{1/3}\alpha = 0.3$ and $n/N = 0.001$. Reproduced with permission from [95]. Copyright 2013 by the American Physical Society.

Table 1. Values of **A** and **B** in equation (43) for the different compared expressions for the mobility μ. C is an unknown numerical coefficient, $\nu \approx 0.875$ and $\widetilde{N}_c \approx 0.219$ are known from percolation theory [87, 95] and $B_c = 2.7$ is a percolation threshold taken from [73]. Reproduced with permission from [24].

Article	A	B
Nenashev *et al* [95]	$0.36 \left(\frac{kT}{\varepsilon_0} \right)^\nu$	$\frac{1}{\widetilde{N}_c}$
Oelerich *et al* [38]	$\frac{B_c}{2\pi}$	$\frac{4\pi}{3B_c} \frac{27}{\exp(3)}$
Vissenberg and Matters [41]	C	$\frac{\pi}{B_c}$
Grünewald and Thomas [89]	$\frac{1}{3}$	$\frac{4\pi}{3B_c} \frac{68}{27}$

parameters in equation (43) are the same in all approaches. Note, that equation (43) is not valid for the low temperatures, when criterion $kT/\varepsilon_0 > (n\alpha^3/8N_c)^{1/3}$ is violated [95]. All solutions are therefore restricted to temperatures above those, at which the well-known Mott law, given by equation (38), is expected.

In figure 8 the data for each of the four approaches is normalized to its value at $kT = 0.3\varepsilon_0$. The deviations between the solutions are therefore only due to different temperature dependences. It is clearly visible from the figure that the result by Grünewald and Thomas [89] and the exact one by Nenashev *et al* [95] agree almost perfectly and match the simulation results very well. This is remarkable since the solution by Grünewald and Thomas was not only the first theory for the VRH in the exponential DOS based on the percolation theory, but it also uses a very simplistic approach with a percolation threshold $B_c = 2.7$ for overlapping spheres with equal sizes, which so far is proven valid only in the case of the NNH. The TE concept, which is introduced in section 5, performs slightly worse, but it still provides a good approximation for the temperature dependence. The solution by Vissenberg and Matters [41] shows the least agreement with the exact result and the computer simulations, although the deviations are still not large.

Regarding the different averaging procedures in equations (42) and (41), one can conclude that the classical averaging, as suggested by Pollak [91], is supported by the perfect

agreement between the theory of Grünewald and Thomas [89] and the exact solution by Nenashev *et al* This averaging was also used by Baranovskii *et al* [67] for the Gaussian DOS. It is worth emphasizing once again that when dealing with percolation approaches, which rely on the critical number of bonds per site as a percolation criterion, it is important to carefully select the averaging procedure.

4.4. Percolation theory for the VRH in the framework of the Gaussian Disorder Model (GDM)

Percolation theory for the VRH in systems with Gaussian DOS has not yet been developed to the extent achieved in the case of exponential DOS, for which the physical problem has been converted into an exactly solvable geometrical one, as discussed in the previous section 4.3. Nevertheless, it is certainly possible to use percolation arguments to successfully describe the VRH in systems with a Gaussian DOS given by equation (3). Below we sketch this approach, closely following the arguments presented in [67, 68, 96].

Similar to the case of the exponential DOS, the Mott law is valid also for the Gaussian DOS at temperatures satisfying the strong inequality $kT(T_0/T)^{1/4} \ll \sigma$, where T_0 in equation (38) contains the same combination of parameters as in the case of the exponential DOS.

Let us show that the VRH regime described by equation (39) is also valid for the Gaussian DOS. In accordance with equation (8), the average number of sites $B(\varepsilon_i, \xi)$ accessible from a site with energy ε_i using hopping rates with the exponent ξ_{ij} smaller than some value ξ is equal to

$$B(\varepsilon_i, \xi) = \int d\varepsilon_j g(\varepsilon_j) \qquad (44)$$
$$\times \int d^3 r_{ij} \theta \left(\xi - \frac{2r_{ij}}{\alpha} - \frac{\varepsilon_{ij}}{kT} \right),$$

with

$$\varepsilon_{ij} = \left[|\varepsilon_i - \varepsilon_j| + |\varepsilon_i - \varepsilon_F| + |\varepsilon_j - \varepsilon_F| \right]/2. \qquad (45)$$

In order to determine the percolation path, these expressions should be inserted into equation (42) and the Gaussian DOS given by equation (3) should be used. The concentration of carriers n in any experimental situation is much smaller than the concentration of localized states N so that the Fermi level is situated in the low-energy part of the DOS and the lower half of the Gaussian DOS is relevant for charge transport. Due to the sharp increase of the Gaussian DOS above the Fermi level, the main contribution to the conductivity comes from the states with energies above the Fermi level. These energies are close to the upper edge of the conducting energy layer, $\varepsilon_F + \xi_c kT$ [96]. In this region (for $|\varepsilon_F + \xi_c kT|, \xi_c kT \gg \sigma$) the average number of bonds can be calculated exactly and from equation (42) one obtains the percolation criterion [96]

$$B_c = \frac{s^2}{6\sqrt{\pi}} \exp(-s^2), \qquad (46)$$

where

$$s = \frac{\varepsilon_F + \xi_c kT}{\sqrt{2}\sigma}. \qquad (47)$$

These equations lead to the percolation criterion for the Gaussian DOS:

$$\xi_c = \frac{\varepsilon_G^* - \varepsilon_F}{kT} \qquad (48)$$

with

$$\varepsilon_G^* = -\sqrt{2}\sigma \ln^{1/2} Q, \qquad (49)$$

$$Q = \frac{\gamma N \alpha^3}{B_c s_c^2} \left(\frac{\sigma}{kT} \right)^3, \qquad (50)$$

where $s_c = -\ln[Q(s_c)]^{1/2}$ [67, 96].

Using equation (48) one obtains an equation similar to equation (39) for carrier mobility in the case of a Gaussian DOS:

$$\mu = \frac{\sigma_{dc}}{en} = \frac{\widetilde{\sigma}_{dc}^0}{en} \exp \left[-\frac{\varepsilon_G^* - \varepsilon_F(n, T)}{kT} \right], \qquad (51)$$

where $\widetilde{\sigma}_{dc}^0$ is the pre-exponential factor in the expression for the conductivity independent from the carrier concentration n. Note, that the notation of ε_G^* in equation (51) was chosen to be different to that in equation (39), ε^*, in order to reflect the difference in DOS shapes used in their derivations. In the framework of the percolation theory, these energies appear as the highest energy levels which carriers have to achieve in order to perform long-range transport. More details regarding these energy levels will be discussed in section 5.

Although the functional forms of equations (39) and (51) look very similar to each other, the dependences of the mobility μ on the carrier concentration n and on temperature in the GDM are essentially different from those in systems with the exponential DOS, as discussed already in section 3.3. At $T > T_c$, or $n < n_c$, where n_c and T_c are given by equations (19) and (21), respectively, the Fermi level is situated below ε_∞. In this regime, the Fermi energy can be approximated [67, 68, 96] as

$$\varepsilon_F(n, T) \simeq -\frac{1}{2} \frac{\sigma^2}{kT} - kT \ln \left(\frac{N}{n} \right). \qquad (52)$$

Inserting this expression into equation (51) leads to the linear dependence of the exponential term in equation (51) on n and concomitantly to the remarkable result that carrier mobility does not depend on the concentration of carriers n at $n < n_c$ due to the presence of another n in the denominator in equation (51). This regime is in drastic contrast to the case of the exponential DOS, where the carrier mobility depends on the concentration of carriers n at all n values. Furthermore, inserting equation (52) into equation (51) shows that at $T > T_c$, or $n < n_c$, the temperature dependence of the mobility in the GDM has the form $\ln(\mu/\mu_0) \propto -(1/2)(\sigma/kT)^2$ [19, 24, 67, 97]. Often the dependence $\ln(\mu/\mu_0) \propto T^{-2}$ is considered as evidence for a Gaussian DOS, while the Arrhenius temperature dependence $\ln(\mu/\mu') \propto T^{-1}$ is claimed to indicate a purely exponential DOS. The important conclusion from the above consideration is the possibility of accounting for both kinds of temperature dependences of hopping conductivity in the framework of the

universal theoretical model based on the Gaussian DOS. The temperature dependence of the mobility is sensitive to the concentration of charge carriers n [19, 24, 67].

It was mentioned above that the VRH problem cannot be solved for a Gaussian DOS with the same precision as has been achieved for the exponential DOS, in the sense that in the latter case not only the exponential terms, but also pre-exponential factors in the expression for the conductivity can be obtained using the percolation theory [24, 95]. If one, however, simplifies the problem and considers not the real system of the randomly placed sites, as in the case of the VRH, but assumes that sites are placed on a regular grid and allows only transitions between the nearest neighbours, one can solve the percolation problem including the derivation of the pre-exponential terms in the expression for hopping conductivity for any shape of the DOS, including the Gaussian one [24, 73, 95].

The problem of transport on lattices was in fact solved several decades ago; the solution can be found in the monograph by Shklovskii and Efros [73]. By applying the method given in this monograph, one can show [95] that the mobility of charge carriers on a grid with lattice constant b is given by

$$\mu(T, n) \simeq \frac{e\omega_0}{kTbn} \left(\frac{kT}{\sigma}\right)^\nu \exp\left(\frac{\varepsilon_F(T, n) - \varepsilon^*}{kT}\right), \quad (53)$$

with the tunneling frequency $\omega_0 = \nu_0 \exp(-2b/\alpha)$. The energy level ε^* in this case can be obtained from the solution of the site percolation problem on a lattice. The number ν in equation (53) is the critical index of the correlation length of the percolation cluster already discussed in section 4.1. This number is universal in the sense that it does not depend on model parameters at scales shorter than the correlation length of the percolation cluster [73, 95]. Cottaar *et al* [74] recently found a similar solution for the GDM on a lattice. However, in that study, the numerical constant ν appears as a fit parameter depending on the lattice structure (SC or FCC) and on the choice of the hopping rates (Miller–Abrahams or Marcus equations). This example from the WoS once again shows that fitting of numerical results does not necessarily lead to correct conclusions.

5. WoT: the concept of the transport energy (TE)

The simple form of equation (51) might lead to the conclusion that the VRH in systems with Gaussian DOS takes place by activation of charge carriers towards the energy level ε_G^*. However, the r-dependent term in equation (8) plays an equally important role for the VRH as the thermal activation. Therefore, one can hardly assume that the description of the VRH as simple activation of carriers towards the energy level ε_G^* would mean that carriers actually conduct moving around that energy level. One should instead interpret the energy level ε_G^* in equation (51) as an 'effective' energy level serving formally to mimic the combined effects of the spatial tunneling and of the activation in energy. The role of ε_G^* is discussed in detail in section 6.3. Only in the case of the NNH, can the level
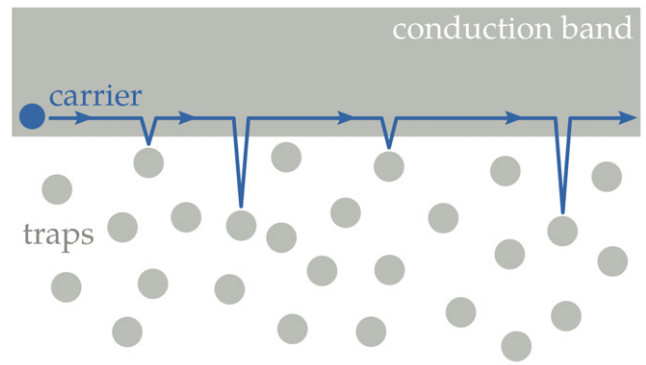

Figure 9. A sketch of the multiple-trapping process.

ε_G^* be treated as decisive for hopping conduction. Therefore, percolation theory is not a suitable tool to determine the 'real' energy path that charge carriers have to visit in order to provide transport in the VRH regime.

A powerful theoretical method for the description of hopping transport in systems with steeply energy-dependent DOS, complementary to the percolation theory, is the concept of the TE. Initially, this approach was developed for systems with the exponential DOS [98, 99] and was later extended for systems with other energy spectra, for instance, with a Gaussian DOS [100]. The essence of the TE approach is the ability to describe hopping transport via sites distributed in space and energy in full analogy with the multiple-trapping (MT) model by replacing the mobility edge in the MT model with a particular energy level ε_t [98–100].

The MT model was initially developed to describe charge transport in inorganic disordered materials, such as chalcogenide glasses and amorphous silicon [101–103], which possess the so-called mobility edge, i.e. the energy level that separates extended states with rather high carrier mobility from the localized states, which can be considered as traps. In the MT model, a charge carrier moves only via delocalized states with energies above the mobility edge. This motion is interrupted by trapping into localized states with subsequent activation of carriers back into the conducting states above the mobility edge. Transport in the framework of the MT model is shown schematically in figure 9. The advantage of the MT charge transport mode as compared to the VRH, is the possibility of exact theoretical treatment by analytical equations under equilibrium and non-equilibrium conditions for various shapes of the DOS [104–106].

There is some confusion in the literature with respect to the term 'transport energy'. Mixing up the effective TE ε_G^* in equation (51) with the real TE ε_t is probably the main reason for that confusion. While the effective TE ε_G^* serves as a formal tool to describe the VRH as simple activation, the real TE ε_t characterizes the transport level at which carriers actually conduct. The difference between these effective and real TEs becomes more transparent if one compares the expressions for the mobility of charge carriers μ derived in various approaches. After deriving ε_t for the GDM [100] in full analogy with the derivations for the exponential DOS, Baranovskii *et al* [97] suggested the following expression for the carrier mobility at

$n \ll n_c$:

$$\mu = \mu_0 \exp\left[-\frac{2r(\varepsilon_t)}{\alpha} - \frac{\varepsilon_t}{kT} - \frac{1}{2}\left(\frac{\sigma}{kT}\right)^2 \right], \quad (54)$$

where $r(\varepsilon_t)$ is the typical distance between localized states with energies below ε_t. Working with the effective TE ε_{eff}, Arkhipov *et al* [107] and Nikitenko *et al* [108] suggested a slightly different expression for the carrier mobility:

$$\mu = \widetilde{\mu_0} \exp\left[-\frac{\varepsilon_{eff}}{kT} - \frac{1}{2}\left(\frac{\sigma}{kT}\right)^2 \right]. \quad (55)$$

Taking into account that values of μ_0 in equation (54) and $\widetilde{\mu_0}$ in equation (55) are very close to each other [97, 107, 108], it is apparent that the meaning of the effective TE ε_{eff} is different to that of the TE ε_t. ε_{eff} in equation (55) is denoted differently to the energy ε_G^* in equation (51) particularly because the percolation nature of the VRH has not been taken into account in the derivation of ε_{eff} by Arkhipov *et al* [107] and by Nikitenko *et al* [108]. The same is true for the initial derivation of the ε_t in the GDM by Baranovskii *et al* [100]. Later Oelerich *et al* [78] derived the equation for ε_t in the GDM, taking into account the percolation criterion and extending this concept for finite carriers concentrations n. The dependence of the 'effective' transport energy, ε_{eff} on n has been derived by Arkhipov *et al* [109].

Another interesting definition of the TE was suggested by Schmechel [69], who, in contrast to all other derivations of the TE, did not use this quantity to calculate the carrier mobility, but instead derived the TE *from* the previously calculated dc conductivity by introducing a differential energy-dependent dc conductivity. The percolation nature of the cluster of sites responsible for the VRH has not been taken into account in that derivation.

Martens *et al* [110] essentially repeated the derivation by Baranovskii *et al* [100], also not taking into account the percolation nature of the VRH transport.

Studying different derivations of the transport energy, one should keep in mind that, not only is the value ε_t important, but so also is the way that it is used to calculate transport coefficients. One popular and transparent approach to obtain the TE is the optimization of hopping rates [78, 99, 100, 111]. Below we briefly describe this approach and point out how the carrier mobility can be obtained from the calculated ε_t.

5.1. Optimization approach to the TE

Consider a carrier in a state with energy ε_i. According to equation (1), the typical rate of a hop of this carrier to a site deeper in energy is

$$\nu_\downarrow = \nu_0 \exp\left[-\frac{2r(\varepsilon_i)}{\alpha} \right], \quad (56)$$

where

$$r(\varepsilon) = \left[\frac{4\pi}{3B_c} \int_{-\infty}^{\varepsilon} g(\varepsilon')d\varepsilon' \right]^{-1/3}. \quad (57)$$
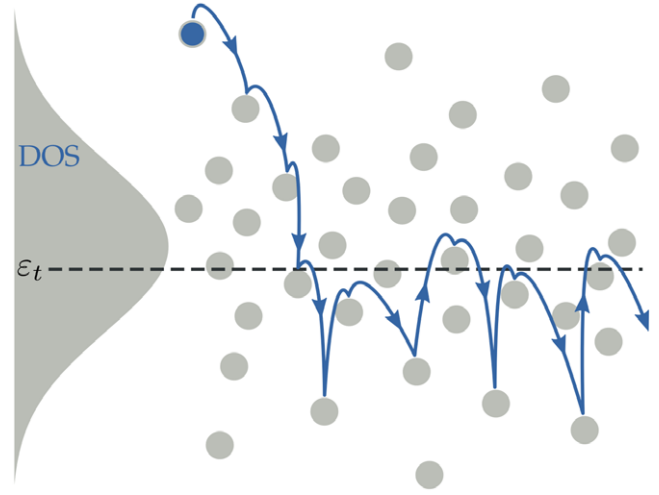

Figure 10. Schematic picture of carrier energy relaxation via the transport energy in an empty system with Gaussian DOS.

The factor B_c accounts for the percolation nature of hopping transport, i.e. the necessity for the hopping carrier to find a continuous path through the system, as discussed in detail in section 4. The rate of an activated jump of the carrier to a site with higher energy, $\varepsilon_x > \varepsilon_i$, is

$$\nu_\uparrow(\varepsilon_i, \varepsilon_x) = \nu_0 \exp\left[-\frac{2r(\varepsilon_x)}{\alpha} - \frac{\varepsilon_x - \varepsilon_i}{kT} \right]. \quad (58)$$

The transport energy ε_t is obtained from equation (58) by finding the most likely targeted energy in a hop from the level ε_i upwards in energy:

$$\frac{\partial \nu_\uparrow(\varepsilon_i, \varepsilon_t)}{\partial \varepsilon_t} = 0. \quad (59)$$

Inserting equations (57) and (58) into equation (59), one arrives at the following expression for the transport energy:

$$\exp\left(\frac{\varepsilon_t^2}{2\sigma^2} \right) \left[\int_{-\infty}^{\varepsilon_t/\sigma\sqrt{2}} e^{-t^2} dt \right]^{4/3}$$
$$= (9\sqrt{2\pi} B_c^{-1} N\alpha^3)^{-1/3} \frac{kT}{\sigma}. \quad (60)$$

It can be easily shown [99, 100], that ε_t does not depend on the initial energy ε_i and is therefore a universal quantity for a given set of system parameters $N\alpha^3$ and σ/kT. Therefore, charge carriers situated at energies below ε_t will most probably be activated to energies in the vicinity of ε_t and carriers above ε_t will most probably fall down in energy. As soon as they reach energies below ε_t, the carriers will perform MT-like motion [112] with the mobility edge replaced by the TE ε_t [98–100]. Figure 10 schematically shows hopping transport within the GDM [97, 100].

A solution of equation (60) for reasonable values of parameters $\sigma/kT = 4$ and $N\alpha^3 = 0.01$ is shown in figure 5. The TE is situated in close vicinity of the maximum of the DOS, an assumption that can be found in previous studies [9, 113], yet without rigorous justification. Note, that the dependence of the TE on the charge carrier concentration was not taken into

account in the derivation of equation (60). This dependence will be introduced below in equation (64). It turns out that the TE is essentially independent of the concentration of charge carriers n up to rather high concentrations $n \approx 0.05\,\mathrm{N}$ [78].

It is important to find out how broad the distribution of the hopping rates is, or equivalently, how sharp the maximum calculated in equation (59) appears for reasonable parameter values. The energy width of the transport path around ε_t appears rather narrow [100]. As long as the distance between ε_t and the carrier distribution at ε_F or ε_∞ is larger than the energy scale of the DOS, σ, the TE model is an accurate tool for description of charge transport [38].

5.2. Numerical calculation of the TE

Along with many analytical approaches, the position of the transport energy was also studied using computer simulation techniques. Since the TE is, by most definitions, the energy path of conducting charge carriers, the first method that comes to mind is to simply trace the most frequently visited energies in straightforward computer simulations. This algorithm was suggested by Cleve *et al* [64] and used in several numerical studies [53, 114, 115]. Hartenstein and Bässler [114] recognized, however, that fast oscillations of carriers between localized states close in space and energy dominate the statistics of the most frequently targeted energies, though do not contribute to the long-range transport. This effect was addressed in analytical studies [116] and in computer simulations [53]. Comprehensive studies of such oscillations were provided by Mendels and Tessler [37] and by Oelerich *et al* [111], whose results clearly show that the algorithm suggested by Cleve *et al* can hardly be used to determine the position of the TE. More details can be found in [37, 111].

Oelerich *et al* [111] recently suggested a computer algorithm that allows one to determine the energy range that dominates the charge transport. The idea is to cut out sites with energies in an interval $[\varepsilon_\mathrm{cut} - w, \varepsilon_\mathrm{cut}]$ and to check whether the resulting mobility, determined in straightforward computer simulations, is affected by such a modification of the DOS. In figure 11 the corresponding modification is depicted. Charge carriers must avoid the sites with energies in this region and find other transport paths. Note, that the DOS is *not* renormalized after the modification. Depending on how much the withdrawn sites contribute to charge transport, the cutting will affect the mobility of the system more or less severely. The highest drop in the carrier mobility is expected when the cut-out sites are the most important ones for the long-range transport. This algorithm has been proven valid for revealing the position of the TE [111] by comparison with the exactly solvable hopping model, the NNH on a lattice.

In order to study the effect of the DOS modification for the case of the VRH, Oelerich *et al* [111] calculated the charge carrier mobilities using the balance equation approach [117, 118]. By varying the width w and the upper boundary ε_cut of the cutout interval (see figure 11), one can find at which w and ε_cut the maximal effect in the decrease of the carrier mobility is achieved and thereby reveal the energy range most important for charge transport. The data are presented in
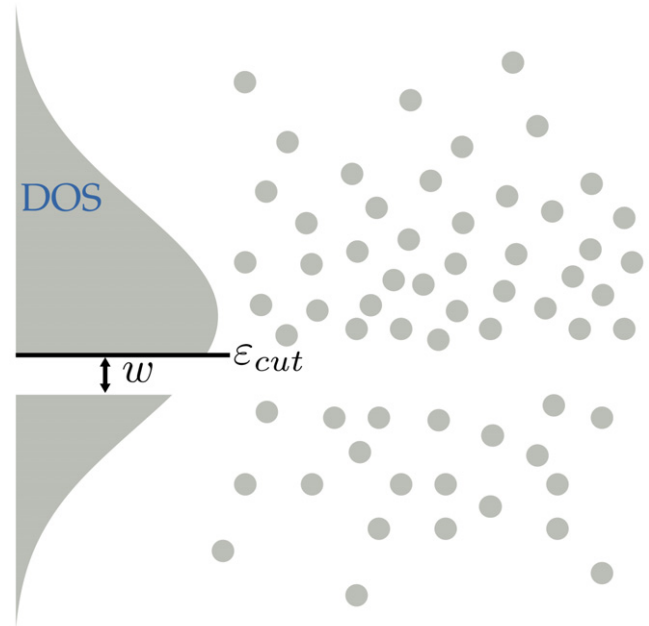

Figure 11. Schematic picture of the modified DOS $g(\varepsilon)$. Sites with energies in the interval $[\varepsilon_\mathrm{cut} - w, \varepsilon_\mathrm{cut}]$ are removed from the system.

figure 12. Results for two different widths, $w = 0.25\sigma$ and $w = 0.5\sigma$ at different temperatures kT are compared. In all cases, motion of a single charge carrier was simulated. The localization length was chosen as $\alpha = 0.215\,\mathrm{N}^{-1/3}$ and the number of sites in the simulated system was equal to 90^3.

It is clearly visible in the figure that a significant decrease in the resulting mobility appears in each of the curves for a certain energy range of withdrawn sites. The effect is larger for larger energy intervals w, yet more accurate for lower values of w, since in that case the system is affected less severely. The choice of the interval width w should therefore be a compromise between the visibility of the effect and the accuracy in determining the position of the most efficient ε_cut. We interpret the minima on the curves plotted in figure 12 as pointing at the position of the real TE responsible for the long-range charge transport. The position of the TE, as determined from figure 12, shifts upwards in energy with rising temperature. This agrees with all previous analytical and numerical studies. The vertical lines close to the minima of the mobility curves in figure 12 show the positions of the TE calculated from equation (60). Apparently, the optimization approach leading to equation (60) is supported by the results of the straightforward computer simulations of the carrier mobility presented in figure 12.

Figure 12 also shows that when sites in the vicinity of the average carrier energy $\varepsilon_\infty = -\sigma^2/kT$ are cut out, the mobility is slightly increased. This happens for the following reason. In the Gaussian DOS in thermal equilibrium most carriers occupy sites around the equilibrium energy ε_∞. Removing sites around this energy pushes carriers to higher energies, decreasing the activation energy necessary for activation to the transport path, leading concomitantly to the increase of the carrier mobility.

The minima on the $\mu(\varepsilon_\mathrm{cut})$ curves in figure 12, which we interpret as indicative for the real transport energy related to
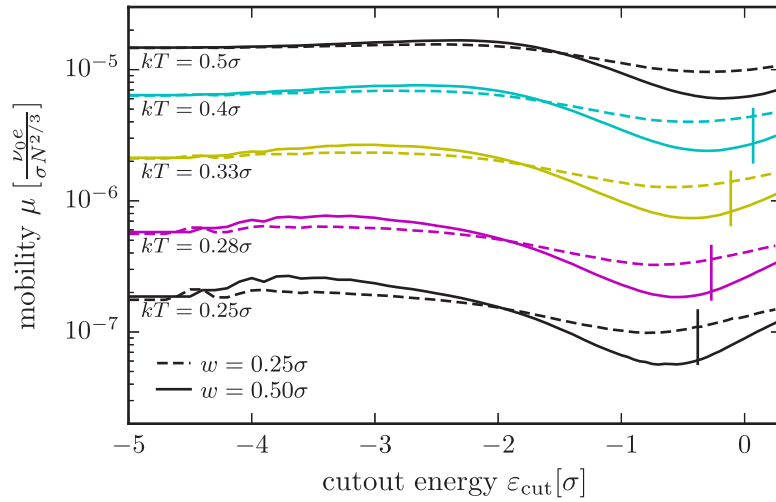
Figure 12. The mobility in a system with absent sites in the interval $[\varepsilon_{cut} - w, \varepsilon_{cut}]$. The vertical lines next to the minimum of the mobility are the analytical results for the transport energy from equation (60).

the long-range charge transport, are situated, for the given temperature range, close to the center of the DOS. It seems therefore correct to assume, in a simplified approach, that the TE coincides with the center of the Gaussian DOS, as has previously been suggested [9, 113]. It means that although charge carriers spend most time in states with energies around the equilibration energy $\varepsilon_\infty = -\sigma^2/kT$ transport takes place via sites with much higher energies than ε_∞.

The data in figure 12 show another remarkable feature, namely that cutting sites with energies in the vicinity of the level $-\sigma^2/2kT$ does not affect the carrier mobility significantly. This shows that it is wrong to interpret the VRH charge transport in the Gaussian DOS as activation of carriers from some particular energy level ε_{start} to the transport energy ε_t. If the latter interpretation were correct, the observation of the mobility temperature dependence in the form of $\ln(\mu/\mu_0) \propto -1/2(\sigma/kT)^2$ [9] would yield the value of $\varepsilon_{start} \approx -\sigma^2/2kT$ for $\varepsilon_t \approx 0$. The fact that cutting sites with energies in the vicinity of the level $-\sigma^2/2kT$ does not affect the mobility means that this energy level does not play any significant role for the charge transport. The temperature dependence of carrier mobility in the form of $\ln(\mu/\mu_0) \propto -1/2(\sigma/kT)^2$ is instead the result of the time averaging over the upward hops to the TE from deeper energy levels [24, 38, 97].

From the above consideration one can conclude that for steeply energy-dependent DOS, such as the Gaussian or exponential DOS, the concept of the TE helps a lot to get a clear insight into the underlying physical processes responsible for charge transport. Using the concept of the TE, Oelerich et al [38] derived the equation for the VRH mobility in a system with a steeply energy-dependent DOS $g(\varepsilon)$ in the form

$$\mu \simeq v_0 \frac{e}{kT} \frac{3B_c F_{ER}}{4\pi r(\varepsilon_t) n_t}$$

$$\times \exp\left(-\frac{2B_c^{1/3}}{\alpha} r(\varepsilon_t) - \frac{\varepsilon_t - \varepsilon_F}{kT}\right) \qquad (61)$$

with n_t determined as

$$n_t = \int_{-\infty}^{\varepsilon_t} f(\varepsilon, \varepsilon_F) g(\varepsilon) d\varepsilon \qquad (62)$$

and the function F_{ER} derived by Roichman and Tessler [119] in the form

$$F_{ER} = \frac{\int_{-\infty}^{\infty} d\varepsilon g(\varepsilon) \frac{\exp[(\varepsilon - \varepsilon_F)/kT]}{(1 + \exp[(\varepsilon - \varepsilon_F)/kT])^2}}{\int_{-\infty}^{\infty} d\varepsilon g(\varepsilon) \frac{1}{1 + \exp[(\varepsilon - \varepsilon_F)/kT]}} . \qquad (63)$$

The value of the TE ε_t is, in the general case, determined via the equation [24, 38]

$$\frac{2}{3}\left(\frac{4\pi}{3B_c}\right)^{-\frac{1}{3}} \frac{kT}{\alpha} \left[\int_{-\infty}^{\varepsilon_t} [1 - f(\varepsilon, \varepsilon_F)] g(\varepsilon) d\varepsilon\right]^{-\frac{4}{3}}$$

$$\times [1 - f(\varepsilon_t, \varepsilon_F)] g(\varepsilon_t) = 1. \qquad (64)$$

The function $r(\varepsilon)$ in the above equations is given by equation (57) and the factor $B_c = 2.735$ [85] accounts for the percolation nature of the hopping transport. Equation (64) is an extension of equation (60) that accounts for finite charge carrier concentrations.

Using equation (61) and (64), one obtains the data shown in figure 8 marked as Oelerich et al for the VRH mobility in the case of the exponential DOS given by equation (4) [38]. Comparison with the exact solution in figure 8 shows that equation (61) based on the TE concept is as accurate as the results based on the percolation approach and computer simulations.

6. WoT: mathematical solution for the VRH using the method of dissipated heat

As explained above, the appropriate transport regime at temperatures smaller than the energy scale of disorder, $kT \ll \sigma$, is the VRH, in which the interplay between the spatial- and energy-dependent terms in the hopping probabilities determines the transport path. Appropriate theoretical methods to account for this transport mechanism

are the percolation theory, described in section 4 and the approach based on the TE concept, described in section 5. In this section, we introduce another analytical method to study charge transport in the VRH regime based on calculating the distribution of dissipated heat. Following the classical recipe, the system will be considered as a resistor network [25, 73, 82, 90, 91].

The new approach is outlined in section 6.1. In section 6.2, it is used to derive information about hopping activation energies. The distribution of lengths of decisive hops is studied in section 6.3, providing a clear picture of the VRH nature of hopping conduction. In section 6.4 we show that the method of cutting out parts of the DOS, as outlined in section 5, indeed yields reliable information about the transport energy.

6.1. Response of a resistor network's conductance on variations of individual resistors' conductances

Let us consider a resistor network under some voltage U. The conductance of the resistor connecting sites i and j is denoted by G_{ij}. The relation of this quantity with the hopping transition rates and the resistor method in general is described in the monograph by Shklovskii and Efros [73]. Voltages U_{ij} and currents $I_{ij} = U_{ij}G_{ij}$ are determined by a system of Kirchhoff's equations including balance of currents at each site i,

$$\sum_j I_{ij} = 0, \qquad (65)$$

and connections between voltages U_{ij}, potentials φ_i, φ_j and electromotive forces $\mathcal{E}_{ij}$,

$$U_{ij} = \varphi_i - \varphi_j + \mathcal{E}_{ij}. \qquad (66)$$

(A proper expression for the electromotive forces reads $\mathcal{E}_{ij} = \mathbf{F}\,\mathbf{d}_{ij}$, where $\mathbf{F}$ is the external electric field, $\mathbf{d}_{ij}$ is the vector joining sites i and j. Periodic boundary conditions in the field direction are assumed.)

In order to relate the conductance G of the whole network to local quantities G_{ij}, U_{ij}, let us consider the power $\mathbb{P}$ of Joule heat generating by the network. On the one hand, this power consists of powers P_{ij} produced by individual resistors (summation is over pairs of sites, each pair is counted only once),

$$\mathbb{P} = \sum_{(i,j)} P_{ij} \equiv \sum_{(i,j)} U_{ij}^2 G_{ij}. \qquad (67)$$

On the other hand, $\mathbb{P}$ can be expressed through the network's conductance G and the voltage U applied to the network (which is the sum of electromotive forces along a line connecting opposite edges of the network),

$$\mathbb{P} = U^2 G. \qquad (68)$$

These two equations provide the relation between quantities G and G_{ij}.

Let us consider infinitely small variations of conductances G_{ij},

$$G_{ij} \to G_{ij} + \delta G_{ij}, \qquad (69)$$

while electromotive forces $\mathcal{E}_{ij}$ and concomitantly the voltage U, are fixed. We are interested in the relative change of the network's conductance $\delta G/G \equiv \delta(\ln G)$ due to these variations. In order to calculate this quantity, we will consider the variation $\delta\mathbb{P}$ of the total power $\mathbb{P}$. First, we will obtain this variation from equation (67):

$$\delta\mathbb{P} = \sum_{(i,j)} \delta(U_{ij}^2 G_{ij}) = 2\sum_{(i,j)} U_{ij}G_{ij}\,\delta U_{ij} + \sum_{(i,j)} U_{ij}^2\,\delta G_{ij}. \qquad (70)$$

Let us show that the first sum in the right-hand side of equation (70) vanishes (This sum will be denoted below as S_1.). Note that each term in this sum remains unchanged under swap of indices $i \leftrightarrow j$. This gives the possibility of representing the sum over non-repeating pairs of sites via double sum over all sites (where each pair is counted twice): $2\sum_{(i,j)} \to \sum_i \sum_j$. So,

$$S_1 = \sum_i \sum_j U_{ij}G_{ij}\,\delta U_{ij} \equiv \sum_i \sum_j I_{ij}\,\delta U_{ij}. \qquad (71)$$

According to equation (66),

$$\delta U_{ij} = \delta\varphi_i - \delta\varphi_j. \qquad (72)$$

Substituting this equation into equation (71), one obtains the relation:

$$S_1 = \sum_i \sum_j I_{ij}(\delta\varphi_i - \delta\varphi_j)$$
$$= \sum_i \delta\varphi_i \left(\sum_j I_{ij}\right) - \sum_j \delta\varphi_j \left(\sum_i I_{ij}\right). \qquad (73)$$

The sums in the brackets are equal to zero, due to the Kirchhoff's law, equation (65). Therefore $S_1 = 0$ and in the right-hand side of equation (70) only the second sum survives:

$$\delta\mathbb{P} = \sum_{(i,j)} U_{ij}^2\,\delta G_{ij}. \qquad (74)$$

The latter equation can be written equivalently as

$$\delta\mathbb{P} = \sum_{(i,j)} P_{ij}\frac{\delta G_{ij}}{G_{ij}} \equiv \sum_{(i,j)} P_{ij}\,\delta(\ln G_{ij}). \qquad (75)$$

Evaluating $\delta\mathbb{P}$ from equation (68),

$$\delta\mathbb{P} = U^2\delta G = U^2 G\,\frac{\delta G}{G} \equiv \mathbb{P}\,\delta(\ln G), \qquad (76)$$

and comparing equation (75) with equation (76), we obtain

$$\delta(\ln G) = \mathbb{P}^{-1}\sum_{(i,j)} P_{ij}\,\delta(\ln G_{ij}). \qquad (77)$$

Equation (77) makes it possible to evaluate the response of the network's conductance to variations G_{ij} of individual resistors. We will use this result for calculating the activation energy of hopping mobility and for calculating the change of the mobility when a small part of the DOS is cut out.

6.2. Activation energy of hopping conductivity and the transport energy

Let us consider mobility μ of some hopping systems as a function of temperature T. Electron concentration is assumed to be constant. One can define the activation energy ε_a as

$$\varepsilon_a = -\frac{d(\ln \mu)}{d(1/kT)}, \tag{78}$$

i.e. as a slope of the temperature dependence of mobility in the Arrhenius plot. Below we denote $1/kT$ as β.

Taking a sufficiently large sample of the conducting medium, one can assume that the conductance G of the sample is proportional to μ, as a function of temperature. Therefore

$$\varepsilon_a = -\frac{d(\ln G)}{d\beta}, \tag{79}$$

and, according to equation (77),

$$\varepsilon_a = -\mathbb{P}^{-1} \sum_{(i,j)} P_{ij} \frac{d(\ln G_{ij})}{d\beta}, \tag{80}$$

where $\mathbb{P}$ is the sum of all P_{ij} as defined by equation (67).

Assuming Miller–Abrahams hopping rates given by equation (8), one can evaluate the conductance G_{ij} corresponding to the transition between sites i and j as

$$G_{ij} = \frac{e^2}{kT} \Gamma_{ij}^0 \exp\left(\frac{\varepsilon_F - \max(\varepsilon_i, \varepsilon_j)}{kT}\right), \tag{81}$$

where Γ_{ij}^0 is a temperature-independent prefactor that depends on the distance between the sites. Equation (81) is valid only when both ε_i and ε_j lie far above the Fermi level. This condition is fulfilled for all *decisive* transitions, for which P_{ij} is non-vanishing. We therefore will use equation (81) for *all* transitions.

Taking the derivative of equation (81), one obtains

$$\frac{d(\ln G_{ij})}{d\beta} = kT + \beta \frac{d\varepsilon_F}{d\beta} + \varepsilon_F - \max(\varepsilon_i, \varepsilon_j). \tag{82}$$

Substituting this equation into equation (80) provides the activation energy in the form:

$$\varepsilon_a = \varepsilon_2 - \varepsilon_1 - kT, \tag{83}$$

where

$$\varepsilon_1 = \beta \frac{d\varepsilon_F}{d\beta} + \varepsilon_F, \tag{84}$$

$$\varepsilon_2 = \frac{\sum_{(i,j)} P_{ij} \max(\varepsilon_i, \varepsilon_j)}{\sum_{(i,j)} P_{ij}}. \tag{85}$$

Note that ε_2 does not depend on the carrier concentration n. Indeed, the only concentration-dependent factor in equation (81) is $\exp(\varepsilon_F/kT)$, the same for all conductances G_{ij}. Therefore all ratios between conductances remain unchanged with varying n and consequently the same is true for the ratios between powers P_{ij} contributing to equation (85).

In the case of a slowly varying DOS, such as the exponential DOS with a characteristic energy scale $\varepsilon_0 \gg kT$, the Fermi energy is almost temperature-independent. Hence, one can neglect the first term of the r.h.s. in equation (84) and get $\varepsilon_1 \approx \varepsilon_F$. According to equation (83) the activation energy ε_a is approximately equal to the distance from the Fermi level to the concentration-independent energy level ε_2. This result resembles the schematic picture shown in figure 4(*a*) for the exponential DOS. In this picture, long-range transport can be understood as activation of carriers from the Fermi level to the transport energy ε_t. Thus, it is reasonable to identify the transport energy with the quantity ε_2.

The case of a Gaussian DOS given by equation (3) is somewhat more complicated. Let us calculate ε_1 for this case. The temperature dependence of ε_F is sensitive to its relative position with regards to the equilibration energy ε_∞ determined by equation (18). When $\varepsilon_F > \varepsilon_\infty$, the Fermi level is almost independent of temperature [96]. Therefore $\varepsilon_1 \approx \varepsilon_F$ in this case. In the opposite case, $\varepsilon_F < \varepsilon_\infty$, the position of the Fermi level can be approximated by equation (52). Taking the derivative, one obtains

$$\beta \frac{d\varepsilon_F}{d\beta} \simeq -\frac{\sigma^2 \beta}{2} - \frac{1}{\beta} \ln \frac{n}{N}, \tag{86}$$

which together with equation (52) leads to the result:

$$\varepsilon_1 \simeq -\sigma^2 \beta = \varepsilon_\infty. \tag{87}$$

Thus, the quantity ε_1 for Gaussian DOS is found to be

$$\varepsilon_1 \simeq \begin{cases} \varepsilon_F, & \text{if } \varepsilon_F > \varepsilon_\infty, \\ \varepsilon_\infty, & \text{if } \varepsilon_F < \varepsilon_\infty. \end{cases} \tag{88}$$

This result shows that, in the Gaussian DOS, the quantity ε_2 determined by equation (85) plays the role of the TE. At high carrier concentrations ($\varepsilon_F > \varepsilon_\infty$), the activation energy $\varepsilon_a \approx \varepsilon_2 - \varepsilon_1$ corresponds to the transitions from the Fermi level to ε_2, as in the case of the exponential DOS discussed above. At low carrier concentrations ($\varepsilon_F < \varepsilon_\infty$), the equilibration energy ε_∞ appears as a starting point for carrier activation instead of the Fermi level, in full accordance with the schematic picture shown in figure 4(*b*).

We have seen that, for both exponential DOS and Gaussian DOS, the quantity ε_2 acts as the target energy of carrier activation, i.e., as the TE. Thus, we *define* the TE ε_t as being equal to ε_2:

$$\varepsilon_t \equiv \varepsilon_2. \tag{89}$$

According to this definition, equation (85) expresses the TE through the values of dissipated powers P_{ij}, which can be found by computer simulations.

Finally, let us offer a more convenient representation for the TE $\varepsilon_t \equiv \varepsilon_2$ defined by equation (85). Let $H(\varepsilon)$ be the sum of P_{ij} over such pairs (i, j) that $\max(\varepsilon_i, \varepsilon_j) < \varepsilon$.

$$H(\varepsilon) = \sum_{\substack{(i,j) \\ \varepsilon_i < \varepsilon \\ \varepsilon_j < \varepsilon}} P_{ij}. \tag{90}$$

Obviously, $H(\varepsilon)$ is a monotonously increasing function with limits

$$H(-\infty) = 0, \quad H(+\infty) = \sum_{(i,j)} P_{ij} \equiv \mathbb{P}. \quad (91)$$

If the sample is sufficiently large, $H(\varepsilon)$ can be considered as a smooth function and it is possible to define the following function $h(\varepsilon)$:

$$h(\varepsilon) = \mathbb{P}^{-1} \frac{\mathrm{d}H(\varepsilon)}{\mathrm{d}\varepsilon}. \quad (92)$$

By definition, $h(\varepsilon)$ is normalized:

$$\int_{-\infty}^{+\infty} h(\varepsilon)\mathrm{d}\varepsilon = 1. \quad (93)$$

One can consider $h(\varepsilon)$ as the site distribution of the generated Joule heat.

The energy ε_t can be expressed via the function $h(\varepsilon)$. To show this, consider the contribution into the sum

$$\sum_{(i,j)} P_{ij} \max(\varepsilon_i, \varepsilon_j) \quad (94)$$

of such pairs (i, j) that $\max(\varepsilon_i, \varepsilon_j) \in [\varepsilon, \varepsilon + \mathrm{d}\varepsilon]$. This contribution is equal to

$$\varepsilon\left[H(\varepsilon + \mathrm{d}\varepsilon) - H(\varepsilon)\right] = \varepsilon \frac{\mathrm{d}H(\varepsilon)}{\mathrm{d}\varepsilon}\mathrm{d}\varepsilon = \varepsilon\mathbb{P}h(\varepsilon)\mathrm{d}\varepsilon. \quad (95)$$

Consequently, the whole sum equation (94) can be obtained by integration of equation (95):

$$\sum_{(i,j)} P_{ij} \max(\varepsilon_i, \varepsilon_j) = \int_{-\infty}^{+\infty} \varepsilon\mathbb{P}h(\varepsilon)\mathrm{d}\varepsilon. \quad (96)$$

Dividing both parts of the latter equation by $\mathbb{P}$, one obtains, according to equations (85) and (89), that

$$\varepsilon_\mathrm{t} = \int_{-\infty}^{+\infty} \varepsilon h(\varepsilon)\mathrm{d}\varepsilon. \quad (97)$$

It means that the TE ε_t is equal to the center of mass of the heat distribution curve $h(\varepsilon)$. This is the central result of this section. It provides an accurate and easy-to-understand meaning of the TE and gives a method of evaluation ε_t from computer simulation data.

An example of heat distributions $h(\varepsilon)$ at different temperatures in the GDM is shown in figure 13. These curves were calculated numerically by the following procedure. The dissipated powers P_{ij} were obtained by numerical solution of a system of Kirchhoff's equations (see the first paragraph of section 6.1) in a cube containing $30 \times 30 \times 30$ sites with periodic boundary conditions.[4] Conductances G_{ij} were taken in the form of equation (81) for all pairs (i, j), assuming infinitely small carrier concentration. Then, the function $h(\varepsilon)$ was obtained from the values of P_{ij} and averaged over 50 realizations.

The curves $h(\varepsilon)$ indicate the energy ranges, at which the sites responsible for long-range transport are located.

---

[4] For $kT/\sigma = 0.2$, a $60 \times 60 \times 60$ cube was used.

Sites with energies *above* this region do not participate in conductivity at all. On the other hand, some sites *below* this region are parts of current-carrying paths, but their resistances are too low to contribute to the whole network's resistance. (See [74, 75] for a spatial picture of the current paths simulated for the lattice model.) In this sense, the exact values of conductances G_{ij} for connections between such sites are not important. This justifies the argumentation based on using the 'non-degenerate' expression (81) for all pairs (i, j) even when some sites lie below the Fermi level, as long as the Fermi level is below the range of non-vanishing $h(\varepsilon)$. Under these circumstances, the function $h(\varepsilon)$ does not depend on the carrier concentration. (The reasoning above, as to why the transport energy ε_2 is concentration-independent, can be applied to the function $h(\varepsilon)$ as well).

Figure 13 clearly demonstrates the VRH nature of conduction in the GDM. The very essence of VRH is the temperature dependence of the optimal hops. With decreasing temperature, a carrier is forced to choose upward hops to sites lower in energy but more distant in space. Consequently, the smaller the temperature is, the lower is the energy range of sites decisive in the long-range transport. This tendency is seen in figure 13: the curve $h(\varepsilon)$ and, concomitantly, the TE, shifts downwards with decreasing temperature.

Let us briefly summarize the main results from this section:

1. The heat distribution function $h(\varepsilon)$ indicates the range of energies of the most important sites for the long-range transport. The TE ε_t is located in the middle of this energy range and is related to the function $h(\varepsilon)$ by equation (97).

2. The activation energy ε_a of hopping conduction is equal to $\varepsilon_\mathrm{t} - \varepsilon_1 - kT$, where the quantity ε_1 is defined by equation (84). This can be interpreted as follows: in order to participate in the long-range conduction, carriers should be thermally activated from the level ε_1 to the transport level ε_t (the additional term $-kT$ is related to a pre-exponential factor). For exponential DOS, $\varepsilon_1 \approx \varepsilon_\mathrm{F}$. The same is true for Gaussian DOS at high enough carrier concentrations, $n > n_\mathrm{c}$. If carrier concentration is low ($n < n_\mathrm{c}$) in the case of Gaussian DOS, ε_1 approaches the equilibration energy $\varepsilon_\infty = -\sigma^2/kT$.

3. The transport energy ε_t can be easily calculated on the basis of computer simulations of hopping transport, either from dissipated powers P_{ij} by equation (85), or from the heat distribution $h(\varepsilon)$ by equation (97). Both these formulas are exact.

4. In the VRH regime, the heat distribution $h(\varepsilon)$ and the TE ε_t move downwards with decreasing temperature. This reflects the interplay between the spatial and energetical terms in the typical rate of an upward hop, given by equation (58) and is the essence of the VRH.

5. The concept of the transport energy is applicable when the Fermi level ε_F lies well below the curve $h(\varepsilon)$ (i.e. when $h(\varepsilon_\mathrm{F})$ is negligible). In this regime, the function $h(\varepsilon)$ and concomitantly the transport energy ε_t do not depend on the carrier concentration.
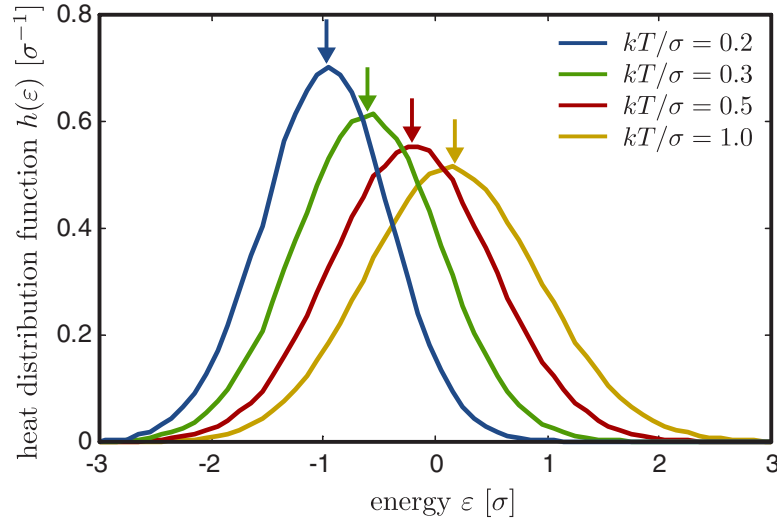
Figure 13. Heat distribution functions $h(\varepsilon)$ at four different temperatures in the system with Gaussian DOS and randomly placed sites (GDM). Arrows indicate positions of the transport energy, calculated by equation (97), at corresponding temperatures. Localization length $\alpha = 0.3\,N^{-1/3}$.

6.3. Distribution of dissipated heat over hopping distances

In the full analogy with the distribution of dissipated heat over energies, $h(\varepsilon)$, one can introduce the heat distribution function over hopping distances, $h_r(r)$:

$$h_r(r) = \mathbb{P}^{-1}\frac{\mathrm{d}H_r(r)}{\mathrm{d}r}, \tag{98}$$

where

$$H_r(r) = \sum_{\substack{(i,j) \\ r_{ij} < r}} P_{ij}. \tag{99}$$

Let us also introduce the notation $\bar{r}$ for the 'mean hopping distance', where the mean value is taken using the dissipated powers P_{ij} as weights:

$$\bar{r} = \frac{\sum_{(i,j)} P_{ij}\, r_{ij}}{\sum_{(i,j)} P_{ij}}. \tag{100}$$

The relation between the function $h_r(r)$ and the quantity $\bar{r}$ is analogous to the relation between $h(\varepsilon)$ and ε_t:

$$\bar{r} = \int_{-\infty}^{+\infty} r\, h_r(r)\, \mathrm{d}r, \tag{101}$$

i.e. $\bar{r}$ is the center of mass of the distribution $h_r(r)$.

In addition, the physical meaning of $h_r(r)$ and $\bar{r}$ are analogous to that of $h(\varepsilon)$ and ε_t. The value of $\bar{r}$ reflects the typical length of hops decisive for conductivity and the curve $h_r(r)$ shows the distribution of such hops. The distributions $h_r(r)$ are given in figure 14 for the same model and parameters used in the previous section for figure 13. One can see that the distributions are restricted from above and from below. With decreasing temperature, the curve $h_r(r)$ shifts towards larger hopping lengths. This is again a direct evidence of the variable-range hopping in the GDM.

Just as the TE determines the conductivity response to a small change of temperature (via the activation energy), the 'mean hopping distance' $\bar{r}$ determines the response to small changes of the localization length α and of the site concentration N. To see this, recall that each elementary conductance G_{ij} is proportional to $\exp(-2r_{ij}/\alpha)$ and, consequently,

$$\frac{\partial(\ln G_{ij})}{\partial(\alpha^{-1})} = -2r_{ij}. \tag{102}$$

From equations (77) and (102) one can express the response of the logarithm of the sample's conductance $\ln G$ to small variation of α^{-1}:

$$\frac{\partial(\ln G)}{\partial(\alpha^{-1})} = \mathbb{P}^{-1}\sum_{(i,j)} P_{ij}(-2r_{ij}) = -2\bar{r}. \tag{103}$$

Note by comparison of equations (102) and (103) that the whole sample reacts to small changes of α as if all hops had the length $\bar{r}$.

It is hardly possible to control α in a real experiment. More experimentally relevant is the system's response to small changes of the site concentration N. These two kinds of responses (to α and to N) are related to each other due to the fact that α and N can appear in the conductivity σ_{dc} not independently, but only in the combination $N\alpha^3$ (apart from the factor $N^{1/3}$ that is needed by dimensionality considerations):

$$\sigma_{dc}(N,\alpha) = N^{1/3}\chi(N\alpha^3), \tag{104}$$

where χ is some unknown function. Taking logarithmic derivatives of σ_{dc},

$$-2\bar{r} = \frac{\partial(\ln\sigma_{dc})}{\partial(\alpha^{-1})} = -3N\alpha^4\frac{\chi'}{\chi}, \tag{105}$$

$$\frac{\partial(\ln\sigma_{dc})}{\partial(\ln N)} = \frac{1}{3} + N\alpha^3\frac{\chi'}{\chi}, \tag{106}$$
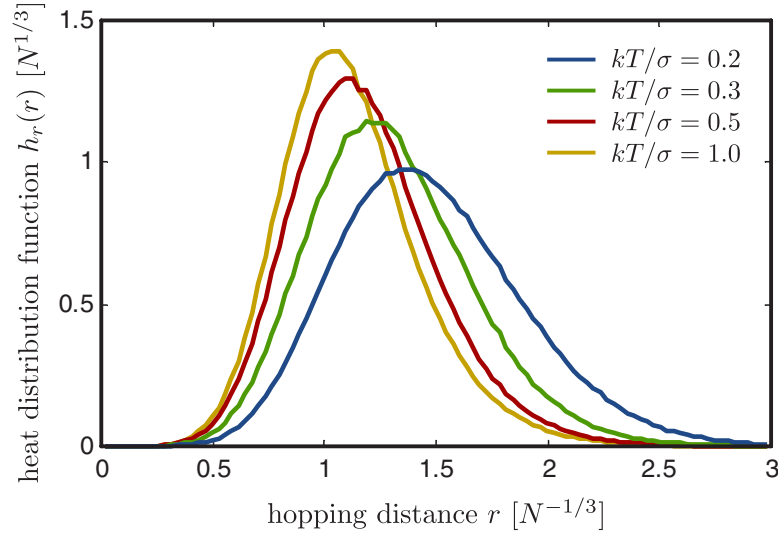
Figure 14. Heat distributions over hopping distances, $h_r(r)$, at four different temperatures in the system with Gaussian DOS and randomly placed sites (GDM). Localization length $\alpha = 0.3\,N^{-1/3}$.

and excluding the function χ from equations (105) and (106), one can obtain the response of σ_{dc} to small variations of the site concentration N:

$$\frac{\partial(\ln \sigma_{\mathrm{dc}})}{\partial(\ln N)} = \frac{1}{3} + \frac{2}{3}\frac{\bar{r}}{\alpha}. \tag{107}$$

(Note that we assumed in equations (104)–(107) that the fraction of occupied sites n/N remains constant during variation of N.)

Let us test equations (103) and (107) on the simplest case of the NNH with isoenergetic sites, as discussed in section 4.1. Applying either equation (103) or equation (107) to the solution of the NNH problem, equation (28), one can find the value of $\bar{r}$,

$$\bar{r} = \frac{\gamma}{2}N^{-1/3} + \frac{\nu}{2}\alpha \equiv r_{\mathrm{c}} + \frac{\nu}{2}\alpha, \tag{108}$$

where r_{c} is the percolation threshold determined by equation (24). In the limit $N\alpha^3 \to 0$ (strong percolation) $\bar{r}$ tends to r_{c}, supporting the interpretation of $\bar{r}$ as of the 'length of the most important hops'.

It is also possible to study not only one-dimensional distributions of the dissipated heat, $h(\varepsilon)$ and $h_r(r)$, but also its two-dimensional distribution over the energy ε and the hopping distance r. An example of such distribution in the GDM is shown in figure 15. One can see that the major part of the heat is generated by hops placed in the (r, ε)-plane near the green line defined by the equation

$$\frac{\varepsilon}{kT} + \frac{2r}{\alpha} = \mathrm{const} \equiv \frac{\varepsilon_G^*}{kT}. \tag{109}$$

Here ε_G^* has the meaning of the 'effective transport energy', which is discussed in sections 4.4 and 5. Figure 15 clearly demonstrates the difference between the 'real' transport energy ε_{t} and the 'effective' one, ε_G^*. Namely, sites with energies near ε_{t} do participate in conductivity, unlike sites with energies close to ε_G^*. The quantity ε_G^* is simply a measure of hopping difficulty including both energetical and spatial barriers.
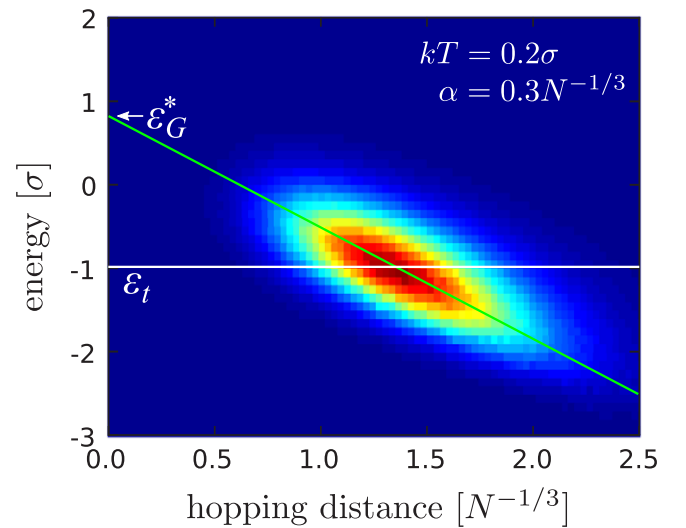

Figure 15. Distribution of the dissipated heat over hopping distances (horizontal axis) and site energies (vertical axis) in the GDM: the maximal dissipation is colored red, zero dissipation is colored blue. The white horizontal line indicates the transport energy ε_{t}. The green inclined line connects equally hard hops: $\varepsilon/kT + 2r/\alpha = \mathrm{const}$. The position of the 'effective transport energy' ε_G^*, discussed in sections 4.4 and 5, is also shown.

6.4. Effect of cutting out small part of DOS

Having found the precise way to determine the TE via equation (97), let us check the validity of the procedure based on the method of cutting out parts of the DOS used in section 5 for the VRH. Let us calculate the change δG of the conductance G of a sample due to cutting out sites within the energy range $[\varepsilon_{\mathrm{cut}}, \varepsilon_{\mathrm{cut}} + w]$, while carrier concentration remains unchanged.[5] The width w of the cut range is considered small as compared to kT. Therefore the initial DOS $g(\varepsilon)$ turns into

---

[5] Unlike section 5, where $\varepsilon_{\mathrm{cut}}$ denotes the *upper* edge of the cut-out energy band, in this subsection we use the symbol $\varepsilon_{\mathrm{cut}}$ for the *lower* edge of the band. This small correction is motivated by our effort to keep the formulas as short as possible.

the following DOS $\tilde{g}(\varepsilon)$:

$$\tilde{g}(\varepsilon) = \begin{cases} 0, & \text{if } \varepsilon \in [\varepsilon_{\text{cut}}, \varepsilon_{\text{cut}} + w], \\ g(\varepsilon), & \text{otherwise}, \end{cases} \quad (110)$$

as illustrated in figure 11. Assuming that changes of all conductances caused by the DOS modification are small, one can calculate δG using equation (77). The small change of local conductances can be achieved by the following change of site energies ε_i:

$$\varepsilon_i \to \varepsilon_i + \Delta(\varepsilon_i), \quad (111)$$

where

$$\Delta(\varepsilon) = \begin{cases} 0, & \text{if } \varepsilon < \varepsilon_{\text{cut}}, \\ w\, g(\varepsilon_{\text{cut}})/g(\varepsilon), & \text{if } \varepsilon \geqslant \varepsilon_{\text{cut}}. \end{cases} \quad (112)$$

The modification of site energies given by equation (111) converts the DOS $g(\varepsilon)$ into $\tilde{g}(\varepsilon)$, while the change of all site energies is small (except very high energies that do not contribute to the charge transport). In order to keep the fixed carrier concentration n, the Fermi level ε_F must be changed as well.

Relative variations $\delta(\ln G_{ij})$ of the local conductances G_{ij} can then be found from equation (81):

$$\delta(\ln G_{ij}) = \frac{\delta\varepsilon_F}{kT} - \frac{\Delta(\max(\varepsilon_i, \varepsilon_j))}{kT}, \quad (113)$$

where $\delta\varepsilon_F$ is the change of the Fermi level. Substitution of equation (113) into equation (77) gives the relative change $\delta(\ln G)$ of the sample's conductance G:

$$\delta(\ln G) = \frac{\delta\varepsilon_F}{kT} - \frac{1}{\mathbb{P}kT} \sum_{(i,j)} P_{ij}\, \Delta(\max(\varepsilon_i, \varepsilon_j)). \quad (114)$$

In equation (114) one can apply the same reduction of the summation $\sum_{(i,j)} P_{ij}\ldots$ to an integration $\int \mathrm{d}\varepsilon\, h(\varepsilon)\ldots$ that led from equation (85) to equation (97). The only difference is that $\Delta(\max(\varepsilon_i, \varepsilon_j))$ is contributing instead of $\max(\varepsilon_i, \varepsilon_j)$. The result is

$$\delta(\ln G) = \frac{\delta\varepsilon_F}{kT} - \frac{1}{kT} \int_{-\infty}^{+\infty} h(\varepsilon)\, \Delta(\varepsilon)\, \mathrm{d}\varepsilon \quad (115)$$

$$= \frac{\delta\varepsilon_F}{kT} - \frac{w}{kT} \int_{\varepsilon_{\text{cut}}}^{+\infty} h(\varepsilon)\, \frac{g(\varepsilon_{\text{cut}})}{g(\varepsilon)}\, \mathrm{d}\varepsilon.$$

Finally, let us find $\delta\varepsilon_F$ from the following condition of constant carrier concentration n:

$$n = \int_{-\infty}^{\infty} g(\varepsilon)\, f(\varepsilon, \varepsilon_F)\, \mathrm{d}\varepsilon \quad (116)$$

$$= \int_{-\infty}^{\infty} \tilde{g}(\varepsilon)\, f(\varepsilon, \varepsilon_F + \delta\varepsilon_F)\, \mathrm{d}\varepsilon,$$

where $f(\varepsilon, \varepsilon_F)$ is the Fermi function. The right-hand side can be rewritten (using smallness of $\delta\varepsilon_F$ and w) as follows:

$$n = n + \frac{\mathrm{d}n}{\mathrm{d}\varepsilon_F} \delta\varepsilon_F - g(\varepsilon_{\text{cut}})\, f(\varepsilon_{\text{cut}}, \varepsilon_F)\, w, \quad (117)$$

from where one can find

$$\delta\varepsilon_F = w\, \frac{g(\varepsilon_{\text{cut}})\, f(\varepsilon_{\text{cut}}, \varepsilon_F)}{\mathrm{d}n/\mathrm{d}\varepsilon_F}. \quad (118)$$

Substitution of $\delta\varepsilon_F$ given by equation (118) into equation (115) provides the final result:

$$\frac{\delta G}{G} = \frac{w}{kT}\, \frac{g(\varepsilon_{\text{cut}})\, f(\varepsilon_{\text{cut}}, \varepsilon_F)}{\mathrm{d}n/\mathrm{d}\varepsilon_F} \\ - \frac{w}{kT} \int_{\varepsilon_{\text{cut}}}^{\infty} h(\varepsilon)\, \frac{g(\varepsilon_{\text{cut}})}{g(\varepsilon)}\, \mathrm{d}\varepsilon. \quad (119)$$

The first (positive) term is responsible for moving the Fermi energy upward due to cutting off part of the DOS, while the second (negative) one reflects a moving upward of the TE.

Figure 16 provides an illustration to equation (119). The upper plot reproduces the heat distribution function $h(\varepsilon)$ taken from figure 13 at $kT = 0.3\sigma$. The lower plot shows the corresponding dependence $\delta G(\varepsilon_{\text{cut}})$. It is clearly visible that figure 16 reproduces all features of figure 12, where the effect of the cut-off was simulated straightforwardly: The region of positive δG near the equilibration energy ε_∞ is described by the first term in equation (119), while the decrease of δG at higher energies reflects the impact of cutting out sites around the transport path (second term in equation (119)).

The quantity δG takes negative values in figure 16 only within the range of non-vanishing $h(\varepsilon)$, i.e. near the TE ε_t. This can be easily understood on the basis of equation (119). When ε_{cut} is too high, the function $h(\varepsilon)$ vanishes at $\varepsilon > \varepsilon_{\text{cut}}$ and so does the integral in equation (119). If ε_{cut} is too low, this integral is proportional to $g(\varepsilon_{\text{cut}})$, therefore strongly decaying with lowering ε_{cut}.

With the above results one can conclude that the procedure based on cutting out parts of the DOS used in section 5 is valid. The TE $\varepsilon_t^{\text{cut}}$ obtained by this procedure (the minimum position of the curve $\delta G(\varepsilon_{\text{cut}})$) is located in the region of energies decisive for the long-range transport. The same is true for the TE ε_t derived from the heat distribution as explained in section 6.2. However, one can see in figures 12 and 16 that $\varepsilon_t^{\text{cut}}$ appears slightly lower than TEs determined by other methods. The reason for that lies in the fact that cutting-out sites *below* the transport energy causes a larger effect in the conductivity than cutting-out sites *above* the TE.

One can conclude that the energy distribution of the dissipated heat, $h(\varepsilon)$, provides enough information to predict the response of conductivity on small changes of the DOS. Together with the observations of sections 6.2 and 6.3, it makes the heat method presented here a powerful tool for analyzing the hopping conductivity in organic materials.

The heat method developed in this section was illustrated in the example of GDM. It should be noted, however, that the method can be applied to other models as well, including the exponential DOS model and the Correlated Disorder Model (CDM). It is especially notable that in the case of CDM, where existence of a transport energy was in doubt, the heat method also provides a meaningful definition of the transport energy.
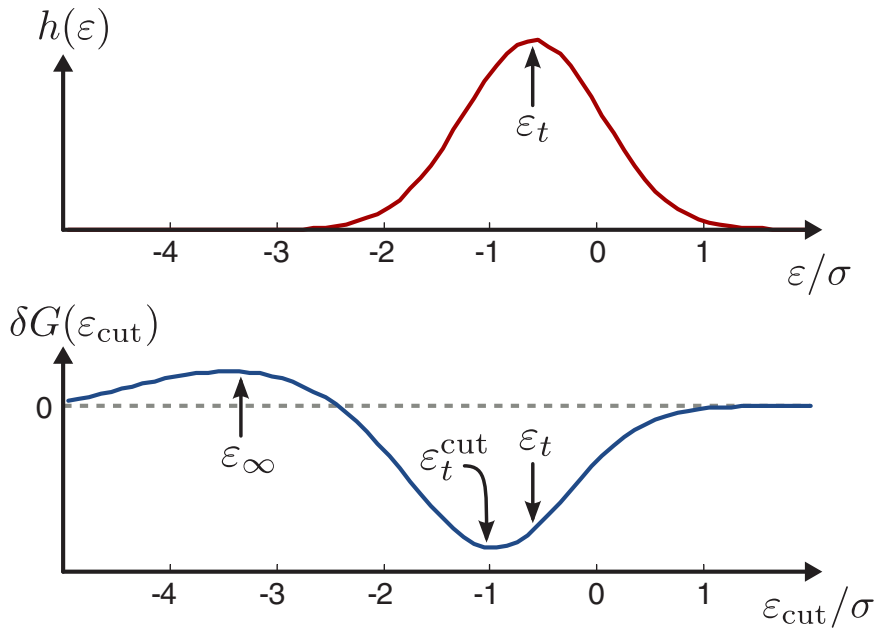
Figure 16. Heat distribution function $h(\varepsilon)$ (upper plot) and change in conductivity $\delta G(\varepsilon_{\text{cut}})$ when a small part of the DOS near the energy ε_{cut} is cut out (lower plot) for GDM. Parameters: $\alpha = 0.3\,N^{-1/3}$, $kT = 0.3\sigma$, $\varepsilon_F = -5\sigma$. Indicated energies: the equilibration energy ε_∞, the transport energy ε_t defined from the heat distribution and the transport energy $\varepsilon_t^{\text{cut}}$ defined by the cutting method introduced in section 5.

7. WoT: On the effective medium approximation

Effective medium approximation (EMA) belongs to the theoretical tools often used to study charge transport in disordered systems [120–122] including ODSs [26, 123]. The EMA is based on the averaging of various characteristics of the inhomogeneous media [73]. Starting from the method of a random resistor network considered in section 4, the EMA leads to the following self-consistency equation for the effective conductance G between hopping sites,

$$\left\langle \frac{G_{ij} - G}{G_{ij} + \chi G} \right\rangle = 0, \tag{120}$$

where G_{ij} is the random quantity describing the conductance between sites i and j in the random resistor network. The symbol $\langle\rangle$ denotes the configurational averaging and $\chi = d-1$, where d is the spatial dimension of the considered problem [120, 122].

The EMA has been applied to study both the NNH transport regime [26, 121] and the VRH regime [122, 124]. Remarkably, many results obtained in the framework of the EMA [122, 124] differ essentially from those obtained in the framework of the percolation theory [73, 94]. For instance, at low temperatures, when hopping conduction occurs in the narrow energy range in close proximity to the Fermi level, the temperature dependence of the hopping conductivity in the VRH regime is predicted in the framework of the EMA for $d = 3$ in the form [122]

$$\sigma_{\text{dc}} = \widetilde{\sigma}_{\text{dc}}^o \exp[-(\widetilde{T}_0/T)^{2/5}], \tag{121}$$

where $\widetilde{T}_0$ is the characteristic temperature and the prefactor $\widetilde{\sigma}_{\text{dc}}^o$ depends on temperature weakly.

Apparently, this result conflicts with the prominent Mott law [88] given by equation (38). Overhof and Thomas [94] analyzed the disagreement between equation (121) and equation (38) and came to the conclusion that equation (121) provided by the EMA is not sufficiently accurate, whereas equation (38) can be viewed as valid. A similar result with respect to the accuracy of the EMA can be found in the monograph by Shklovskii and Efros [73], who concluded that '*the effective medium theory gives accurate results only in the case of a weak relative inhomogeneity. This approach, by its very essence, cannot lay claim to good results for a strongly (exponentially) inhomogeneous medium.*' For weakly inhomogeneous systems, the EMA might be applicable, but such systems are not within the scope of the current review article.

8. Summary

Though claiming in the title of this review that the paper is related to theoretical treatment of charge transport in organic disordered semiconductors (ODSs), we have not focused strongly on specific details of ODSs. Instead, the main content of the article is devoted to general theoretical tools suitable to describe hopping transport in strongly inhomogeneous media. These general tools are applicable to the ODSs since it is widely concluded in the community that the transport mechanism in these materials is hopping transport [13–24].

What is, however, specific to the state of research related to charge transport in ODSs is the unawareness of large parts of the community with respect to the general theoretical tools suitable to successfully describe hopping transport. Therefore, much emphasis in the research on ODSs has so far been put on the phenomenological fitting of results obtained in numerical calculations, by equations containing numerous

adjustable parameters. The most popular examples of such phenomenological fittings are provided by equations (5)–(7), which are shown in section 2 as hardly acceptable. Furthermore, it is shown in section 3 that forcing charge carriers to perform only the nearest-neighbour transitions on a regular lattice grid, performed in several recent theoretical studies, can hardly bring meaningful results if the thermal energy kT is smaller than the energy scale of disorder, as is typical for ODSs at room temperature. The disproportion in the theoretical research on ODSs between the world of simulations (WoS) and the world of theory (WoT) is still remarkable.

In sections 4 and 5 we highlighted the percolation theory and the concept of transport energy (TE) as suitable tools to describe the variable-range hopping transport in strongly inhomogeneous disordered systems, including the ODSs. These well-approved theoretical tools, developed initially for applications in the field of inorganic materials, are perfectly suitable to describe charge transport in ODSs. Among other results, it is shown that transport features are sensitive to the energy spectrum of the system. For instance, in systems with the density of states steeper than the exponential function, carrier mobility μ does not depend on the concentration of carriers n at low n values, while in systems with the DOS described by an exponential or by a weaker energy-dependent function, μ, on the contrary, depends on n at all n values.

In section 6 a novel analytical approach based on calculations of the heat power production in the random resistor network is described as an accurate tool to study hopping transport. This accurate theoretical tool confirms the validity and accuracy of the TE concept.

Among the not yet clarified issues, one could highlight the question as to whether the GDM or the CDM (correlated disorder model) is more realistic for ODSs and the question on the role of polaron effects for charge transport in ODSs. Furthermore, one should go beyond the simplest set of phenomenological assumptions, which is the basis for the GDM and CDM. Some existing theoretical attempts in this direction were however beyond the scope of our review.

Acknowledgments

References

[1] Tang C W and Van Slyke S A 1987 *Appl. Phys. Lett.* **51** 913
[2] Burroughes J H, Bradley D D C, Brown A R, Marks R N, Mackay K, Friend R H, Burns P L and Holmes A B 1990 *Nature* **347** 539
[3] Brown A R, Jarrett C P, de Leeuw D M and Matters M 1997 *Synth. Met.* **88** 37
[4] Yu G, Gao J, Hummelen J C, Wudl F and Heeger A J 1995 *Science* **270** 1789
[5] Abkowitz M 1992 *Phil. Mag. B* **65** 817
[6] Kryukov A Y, Saidov A C and Vannikov A V 1992 *Thin Solid Films* **209** 84
[7] Gill W 1972 *J. Appl. Phys.* **43** 5033
[8] Schein L A, Peled A and Glatz D 1989 *J. Appl. Phys.* **66** 686

[9] Bässler H 1993 *Phys. Status Solidi* B **175** 15
[10] Borsenberger P M and Shi J 1995 *Phys. Status Solidi* B **191** 461
[11] Borsenberger P M, Gruenbaum W T and Magin E H 1996 *Physica* B **228** 226
[12] Bässler H 1981 *Phys. Status Solidi* B **107** 9
[13] Borsenberger P M, Magin E H, van der Auweraer M and de Schryver F C 1993 *Phys. Status Solidi* A **140** 9
[14] van der Auweraer M, de Schryver F C, Borsenberger P M and Bässler H 1994 *Adv. Mater.* **6** 199
[15] Pope M and Swenberg C E 1999 *Electronic Processes in Organic Crystals and Polymers* (Oxford: Oxford University Press)
[16] Hadziioannou G and van Hutten P F (ed) 2000 *Semiconducting Polymers* (New York: Wiley)
[17] Brabec C, Dyakonov V, Parisi J and Sariciftci N (ed) 2003 *Organic Photovoltaics: Concepts and Realization* (Berlin: Springer)
[18] Bruetting W (ed) 2005 *Physics of Organic Semiconductors* (New York: Wiley)
[19] Baranovski S (ed) 2006 *Charge Transport in Disordered Solids with Applications in Electronics* (Chichester: Wiley)
[20] Schwoerer M and Wolf H C 2007 *Organic Molecular Solids* (New York: Wiley)
[21] Sun S S and Dalton L (ed) 2008 *Organic Electronic and Optoelectronic Materials and Devices* (Boca Raton, FL: CRC Press)
[22] Tessler N, Preezant Y, Rappaport N and Roichman Y 2009 *Adv. Mater.* **21** 2741
[23] Meller G and Grasser T (ed) 2010 *Organic Electronics* (Berlin: Springer)
[24] Baranovskii S D 2014 *Phys. Status Solidi* B **251** 487
[25] Miller A and Abrahams E 1960 *Phys. Rev.* **120** 745
[26] Fishchuk I I, Kadashchuk A, Hoffmann S T, Athanasopoulos S, Genoe J, Bässler H and Köhler A 2013 *Phys. Rev.* B **88** 125202
[27] Marcus R A 1964 *Ann. Rev. Phys. Chem.* **15** 155
[28] Rubel O, Baranovskii S. D, Thomas P and Yamasaki S 2004 *Phys. Rev.* B **69** 014206
[29] Bredas J L, Beljonne D, Coropceanu V and Cornil J 2004 *Chem. Rev.* **104** 4971
[30] Martens H C F, Blom P W M and Schoo H F M 2000 *Phys. Rev.* B **61** 7489
[31] Arkhipov V I, Heremans P, Emelianova E V, Adriaenssens G J and Bässler H 2002 *J. Phys.: Condens. Matter* **14** 9899
[32] Tanase C, Blom P W M, de Leeuw D M and Meijer E J 2004 *Phys. Status Solidi* A **201** 1236
[33] Pasveer W F, Cottaar J, Tanase C, Coehoorn R, Bobbert P A, Blom P W M, de Leeuw D M and Michels M A J 2005 *Phys. Rev. Lett.* **94** 206601
[34] Arkhipov V I, Heremans P, Emelianova E V and Bässler H 2005 *Phys. Rev.* B **71** 045214
[35] Craciun N I, Wildeman J and Blom P W M 2008 *Phys. Rev. Lett.* **100** 056601
[36] Germs W C, van der Holst J J M, van Mensfoort S L M, Bobbert P A and Coehoorn R 2011 *Phys. Rev.* B **84** 165210
[37] Mendels D and Tessler N 2013 *J. Phys. Chem.* C **117** 24740
[38] Oelerich J O, Huemmer D and Baranovskii S D 2012 *Phys. Rev. Lett.* **108** 226403
[39] Gartstein Y N and Conwell E M 1995 *Chem. Phys. Lett.* **245** 351
[40] Novikov S V, Dunlap D H, Kenkre V M, Parris P E and Vannikov A V 1998 *Phys. Rev. Lett.* **81** 4472
[41] Vissenberg M C J M and Matters M 1998 *Phys. Rev.* B **57** 12964

[42] Meijer E J, Tanase C, Blom P W M, van Veenendaal E, Huisman B H, de Leeuw D M and Klapwijk T M 2002 *Appl. Phys. Lett.* **80** 3838

[43] Tanase C, Meijer E, Blom P and de Leeuw D 2003 *Phys. Rev. Lett.* **91** 216601

[44] Tanase C, Meijer E, Blom P and de Leeuw D 2003 *Org. Electron.* **4** 33

[45] Nelson J 2003 *Phys. Rev.* B **67** 155209

[46] Tanase C, Blom P and de Leeuw D 2004 *Phys. Rev.* B **70** 193202

[47] Salleo A, Chen T W, Völkel A R, Wu Y, Liu P, Ong B S and Street R A 2004 *Phys. Rev.* B **70** 115311

[48] Nelson J, Choulis S and Durrant J 2004 *Thin Solid Films* **451–2** 508

[49] Anthopoulos T, Tanase C, Setayesh S, Meijer E J, Hummelen J C, Blom P W M and de Leeuv D M 2004 *Adv. Mater.* **16** 2174

[50] Blom P W M, Tanase C, de Leeuw D M and Coehoorn R 2005 *Appl. Phys. Lett.* **86** 092105

[51] Sedghi N, Donaghy D, Raja M, Badriya S, Higgins S and Eccleston W 2006 *J. Non-Cryst. Solids* **352** 1641

[52] Estrada M, Meja I, Cerdeira A, Pallares J, Marsal L and Iiguez B 2008 *Solid-State Electron.* **52** 787

[53] Gonzalez-Vazquez J, Anta J and Bisquert J 2009 *Phys. Chem. Chem. Phys.* **11** 10359

[54] Tachyia M and Seki K 2010 *Phys. Rev.* B **82** 085201

[55] Montero J and Bisquert J 2011 *J. Appl. Phys.* **110** 043705

[56] Street R A, Song K W, Northrup J E and Cowan S 2011 *Phys. Rev.* B **83** 165207

[57] Germs W C, Guo K, Janssen R A J and Kemerink M 2012 *Phys. Rev. Lett.* **109** 016601

[58] Brondijk J J, Roelofs W S C, Mathijssen S G J, Shehu A, Cramer T, Biscarini F, Blom P W M and de Leeuw D M 2012 *Phys. Rev. Lett.* **109** 056601

[59] Schubert M, Preis E, Blakesley J C, Pingel P, Scherf U and Neher D 2013 *Phys. Rev.* B **87** 024203

[60] Borsenberger P M and Fitzgerald J J 1993 *J. Phys. Chem.* **97** 4815

[61] Kageyama H, Ohnishi K, Nomura S and Shirota Y 1997 *Chem. Phys. Lett.* **277** 137

[62] Laquai F, Wegner G and Bässler H 2007 *Phil. Trans. R. Soc.* A **365** 1473

[63] Shklovskii B I 1973 *Sov. Phys. Semicond.* **6** 1964

[64] Cleve B, Hartenstein B, Baranovskii S. D, Scheidler M, Thomas P and Bässler H 1995 *Phys. Rev.* B **51** 16705

[65] Jansson F, Baranovskii S D, Sliaužys G, Österbacka R and Thomas P 2008 *Phys. Status Solidi* C **5** 722–4

[66] Coehoorn R and Bobbert P A 2012 *Phys. Status Solidi* A **209** 2354

[67] Baranovskii S D, Zvyagin I P, Cordes H, Yamasaki S and Thomas P 2002 *Phys. Status Solidi* B **230** 281

[68] Baranovskii S D, Zvyagin I P, Cordes H, Yamasaki S and Thomas P 2002 *J. Non-Cryst. Solids* **299** 416

[69] Schmechel R 2002 *Phys. Rev.* B **66** 235206

[70] Schmechel R 2003 *J. Appl. Phys.* **93** 4653

[71] Shaked S, Tal S, Roichman Y, Razin A, Xiao S, Eichen Y and Tessler N 2003 *Adv. Mater.* **15** 913

[72] Roichman Y and Tessler N 2003 *Synth. Met.* **135** 443

[73] Shklovskii B I and Efros A L 1984 *Electronic Properties of Doped Semiconductors* (Berlin: Springer)

[74] Cottaar J, Koster L J A, Coehoorn R and Bobbert P A 2011 *Phys. Rev. Lett.* **107** 136601

[75] Cottaar J, Coehoorn R and Bobbert P A 2012 *Phys. Rev.* B **85** 245205

[76] Landau L D and Lifshitz E M 1965 *Quantum Mechanics: Non-Relativistic Theory* (Oxford: Pergamon)

[77] Ries B, Bässler H, Grünewald M and Movaghar B 1988 *Phys. Rev.* B **37** 5508

[78] Oelerich J O, Huemmer D, Weseloh M and Baranovskii S D 2010 *Appl. Phys. Lett.* **97** 143302

[79] Campbell I H, Smith D L, Neef C J and Ferraris J P 1999 *Appl. Phys. Lett.* **75** 841

[80] Schulze K, Riede M, Brier E, Reinold E, Bäuerle P and Leo K 2008 *J. Appl. Phys.* **104** 074511

[81] Alesi S, Brancolini G, Viola I, Capobianco M L, Venturini A, Camaioni N, Gigli G, Melucci M and Barbarella G 2009 *Chem. Eur. J.* **15** 1876

[82] Ambegaokar V, Halperin B I and Langer J S 1971 *Phys. Rev.* B **4** 2612

[83] Dalton N W, Domb C and Sykes M F 1964 *Proc. Phys. Soc.* **83** 496

[84] Domb C and Dalton N W 1966 *Proc. Phys. Soc.* **89** 859

[85] Lorenz C D and Ziff R M 2001 *J. Chem. Phys.* **114** 3659

[86] Dyre J C and Schrøder T B 2000 *Rev. Mod. Phys.* **72** 873

[87] Lorenz C D and Ziff R M 1998 *Phys. Rev.* E **57** 230

[88] Mott N F 1969 *Phil. Mag.* **19** 835

[89] Grünewald M and Thomas P 1979 *Phys. Status Solidi* B **94** 125

[90] Shklovskii B I and Efros A L 1971 *Sov. Phys.—JETP* **33** 468

[91] Pollak M 1972 *J. Non-Cryst. Solids* **11** 1

[92] Coehoorn R, Pasveer W F, Bobbert P A and Michels M A J 2005 *Phys. Rev.* B **72** 155206

[93] Overhof H 1975 *Phys. Status Solidi* B **67** 709

[94] Overhof H and Thomas P 1996 *Phys. Rev.* B **53** 13187

[95] Nenashev A V, Jansson F, Oelerich J O, Huemmer D, Dvurechenskii A V, Gebhard F and Baranovskii S D 2013 *Phys. Rev.* B **87** 235204

[96] Zvyagin I P 2008 *Phys. Status Solidi* C **5** 725

[97] Baranovskii S D, Cordes H, Hensel F and Leising G 2000 *Phys. Rev.* B **62** 7934

[98] Monroe D 1985 *Phys. Rev. Lett.* **54** 146

[99] Baranovskii S D, Thomas P and Adriaenssens G J 1995 *J. Non-Cryst. Solids* **190** 283

[100] Baranovskii S D, Faber T, Hensel F and Thomas P 1997 *J. Phys.* C **9** 2699

[101] Noolandi J 1977 *Phys. Rev.* B **16** 4466

[102] Schmidlin F W 1977 *Phys. Rev.* B **16** 2362

[103] Silver M and Cohen L 1977 *Phys. Rev.* B **15** 3276

[104] Rudenko A I and Arkhipov V I 1982 *Phil. Mag.* B **45** 177

[105] Arkhipov V I and Rudenko A I 1982 *Phil. Mag.* B **45** 189

[106] Nenashev A V, Jansson F, Baranovskii S D, Österbacka R, Dvurechenskii A V and Gebhard F 2010 *Phys. Rev.* B **81** 115204

[107] Arkhipov V I, Emelianova E V and Bässler H 2001 *Phil. Mag.* B **81** 985

[108] Nikitenko V R, von Seggern H and Bässler H 2007 *J. Phys.: Condens. Matter* **19** 136210

[109] Arkhipov V I, Heremans P and Bässler H 2003 *Appl. Phys. Lett.* **82** 4605

[110] Martens H C F, Hulea I N, Romijn I, Brom H B, Pasveer W F and Michels M A J 2003 *Phys. Rev.* B **67** 121203

[111] Oelerich J O, Jansson F, Nenashev A V, Gebhard F and Baranovskii S D 2014 *J. Phys.: Condens. Matter* **26** 255801

[112] Orenstein J and Kastner M 1981 *Solid State Commun.* **40** 85

[113] Neumann F, Genenko Y A and von Seggern H 2006 *J. Appl. Phys.* **99** 013704

[114] Hartenstein B and Bässler H 1995 *J. Non-Cryst. Solids* **190** 112

[115] Novikov S V and Malliaras G G 2006 *Phys. Status Solidi* B **243** 387

[116] Arkhipov V I, Emelianova E V and Adriaenssens G J 2001 *Phys. Rev.* B **64** 125125

[117] Yu Z G, Smith D L, Saxena A, Martin R L and Bishop A R 2000 *Phys. Rev. Lett.* **84** 721

[118] Cottaar J and Bobbert P A 2006 *Phys. Rev.* B **74** 115204

[119] Roichman Y and Tessler N 2002 *Appl. Phys. Lett.* **80** 1948

[120] Kirkpatrick S 1973 *Rev. Mod. Phys.* **45** 574

[121] Böttger H, Bryksin V V and Schulz F 1994 *Phys. Rev. B* **49** 2447

[122] Bleibaum O, Böttger H, Bryksin V V and Schulz F 1995 *Phys. Rev. B* **51** 14020

[123] Fishchuk I I, Arkhipov V I, Kadashchuk A, Heremans P and Bässler H 2007 *Phys. Rev. B* **76** 045210

[124] Mott N F and Twose W 1961 *Adv. Phys.* **10** 707

Article II

A. V. Nenashev, F. Jansson, **J. O. Oelerich**, D. Huemmer, A. V. Dvurechenskii, F. Gebhard, and S. D. Baranovskii
Advanced percolation solution for hopping conductivity
Phys. Rev. B **87**, 235204 (2013)

Advanced percolation solution for hopping conductivity

A. V. Nenashev,[1,2] F. Jansson,[3] J. O. Oelerich,[3,*] D. Huemmer,[3] A. V. Dvurechenskii,[1,2]
F. Gebhard,[3] and S. D. Baranovskii[3]

[1]*Institute of Semiconductor Physics, 630090 Novosibirsk, Russia*
[2]*Novosibirsk State University, 630090 Novosibirsk, Russia*
[3]*Department of Physics and Material Sciences Center, Philipps-University, D-35032 Marburg, Germany*

Hopping of carriers between localized states dominates charge transport in amorphous organic and inorganic semiconductors. We suggest a comprehensive description of this transport regime based on the percolation approach that allows one to determine not only very pronounced exponential dependencies of the hopping conductivity on material parameters, but also the more weakly dependent pre-exponential factors. The problem of the variable-range hopping (VRH) via sites with exponential energy distribution is mapped onto a universal geometrical problem of percolation via spheres with distributed sizes. An exact solution of the latter problem provides accurate results for the VRH in systems with exponential density-of-states (DOS). Our analytical results are confirmed by straightforward computer simulations and compared to former results present in the literature. We also discuss the case of nearest-neighbor hopping on a lattice, where the pre-exponential factors are provided by the percolation approach for any shape of the DOS.

I. INTRODUCTION

Interest in optoelectronic properties of amorphous organic and inorganic semiconductors has been steeply growing in the last decades. This is caused by successful applications of such materials in various devices and by their promises for future applications. Manufacturability and low production costs of amorphous materials, along with their specific charge transport properties, make such materials favorable and in some cases unique for various applications, particularly for large-area devices, where demands to the mobilities of charge carriers are not very high. Already at an early stage of the research on disordered semiconductors it was recognized that within wide ranges of parameters charge transport is dominated by incoherent tunneling transitions (hopping) of carriers between spatially localized states with a broad energy distribution. For the energy spectrum, also called the density of states (DOS), two shapes are usually discussed: (i) an exponential DOS,

$$g(\varepsilon) = \begin{cases} 0 & \text{when } \varepsilon > 0, \\ \dfrac{N}{\sigma} \exp(\varepsilon/\sigma) & \text{when } \varepsilon \leqslant 0, \end{cases} \quad (1)$$

is often assumed for inorganic amorphous materials,[1,2] and (ii) a Gaussian DOS,

$$g(\varepsilon) = \frac{N}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right), \quad (2)$$

is assumed for organic disordered semiconductors,[3,4] although the latter materials are sometimes claimed to possess an exponential DOS.[5,6]

The parameter σ in Eqs. (1) and (2) characterizes the width of the DOS varying from $\sigma \simeq 0.025$ eV for inorganic amorphous materials[2] to $\sigma \simeq 0.1$ eV for organic systems.[3] In this report the hopping of carriers through randomly distributed sites with concentration N will be considered. For the sake of simplicity, no correlations between positions and energies of sites will be assumed.

The rate Γ_{ij} of carrier transitions from an occupied site i to an empty site j, separated by the distance d_{ij}, is assumed to follow the Miller-Abrahams expression,

$$\Gamma_{ij} = \nu_0 \exp\left[-\frac{2d_{ij}}{a} - \frac{\max(\Delta\varepsilon_{ij}, 0)}{kT}\right], \quad (3)$$

where ν_0 is the attempt-to-escape frequency, a is the localization length, k is the Boltzmann constant, T is temperature, and $\Delta\varepsilon_{ij} = \varepsilon_j - \varepsilon_i$ is the difference between carrier energies ε_j and ε_i on sites j and i, respectively. The latter quantities depend on the applied electric field.

Theoretical studies of hopping transport in disordered semiconductors are often performed by empirical fitting of numerical results, which leads to analytical equations that can hardly be valid in a broad range of parameters. This is particularly true for organic materials, for which analytical studies are believed to be very difficult.[3] The aim of the current report is to present an analytical theoretical approach that allows one to determine not only the very pronounced exponential dependencies of carrier mobility on various material parameters, but also the more weakly dependent pre-exponential factors. This approach is based on classical percolation theory.

The percolation approach for the description of variable-range-hopping (VRH) has been known for about 40 years.[7–9] Grünewald and Thomas[10] used this approach for studying hopping transport in the exponential DOS given by Eq. (1). Almost two decades later, the same problem in the framework of the percolation approach was addressed by Vissenberg and Matters,[5] who used a somewhat different percolation criterion than Grünewald and Thomas.[10] Since no comparison with the results of Grünewald and Thomas has been performed by Vissenberg and Matters, it is not clear which percolation approach is correct or more accurate.

In Sec. II it is shown that the VRH problem in the exponential DOS can be mapped onto a purely geometrical problem that allows a numerical solution described in Sec. III.

This numerical solution provides results for the conductivity of the system and for the carrier mobility, accurate up to a dimensionless numerical coefficient. The value of this coefficient is determined via straightforward simulations described in Sec. IV. In Sec. V the new analytical approach is compared to previous results from the literature. It is shown that the result of Grünewald and Thomas[10] is more accurate than that of Vissenberg and Matters.[5] Furthermore, the obtained exact solution for the exponential DOS is used to check in Sec. V the validity of the recently suggested theoretical approach for calculating hopping mobilities.[4]

The percolation approach to hopping transport in the Gaussian DOS [Eq. (2)][11,12] appears less transparent than in the case of an exponential DOS. For the Gaussian DOS the exact mapping on a purely geometrical model is not possible. However, in the case of carrier hopping via neighboring sites on a regular lattice, suggested recently in the frame of empirical fitting for disordered organic semiconductors,[13] a theoretical approach similar to that described in Sec. II is feasible also for a Gaussian DOS. This approach is presented in Sec. VI. Concluding remarks are given in Sec. VII.

II. VRH IN THE EXPONENTIAL DOS

The standard method for studying the problem of hopping transport in small electric fields (within the Ohmic regime) is the method of a resistance network.[7–9,14,15] In this method, a resistance R_{ij} is assigned to each pair of sites with nonvanishing hopping rate Γ_{ij} given by Eq. (3). The expression for R_{ij} becomes especially simple when the Fermi level ε_F is far below the site energies ε_i, ε_j. In this case,

$$R_{ij} = \frac{kT}{e^2 \nu_0} \exp\left[\frac{2d_{ij}}{a} + \frac{\max(\varepsilon_i, \varepsilon_j) - \varepsilon_F}{kT}\right]. \quad (4)$$

We will assume that the carrier concentration is low enough, so that Eq. (4) holds for all pairs of sites significant for the conductivity. The quantitative meaning of this "low enough concentration" constraint will be described below. Also, low temperatures will be considered, $kT \ll \sigma$.

Since the resistances R_{ij} are distributed in an exponentially broad region, the percolation approach can be used for determining the net resistivity. The main part of this approach consists of seeking the *critical* value R_{crit} of the resistance, i.e., the value at which the opposite sides of a sample become connected, when resistances are "switched on" in ascending order. The system of connected site with bond resistances R_{ij} defined by the bonding criterion

$$R_{ij} \leqslant R_{crit} \quad (5)$$

is at the percolation threshold, by definition of R_{crit}. It is convenient to write R_{crit} in a way similar to Eq. (4),

$$R_{crit} = \frac{kT}{e^2 \nu_0} \exp\left(\frac{\varepsilon^* - \varepsilon_F}{kT}\right), \quad (6)$$

introducing a "critical energy" ε^*. The bonding criterion (5) can be expressed via ε^* as

$$\frac{2d_{ij}}{a} + \frac{\varepsilon_i}{kT} \leqslant \frac{\varepsilon^*}{kT}, \quad \frac{2d_{ij}}{a} + \frac{\varepsilon_j}{kT} \leqslant \frac{\varepsilon^*}{kT}. \quad (7)$$

Since $d_{ij} > 0$, only sites with energies $\varepsilon < \varepsilon^*$ can have connections satisfying the criterion (7). The concentration N^* of such sites is

$$N^* = \int_{-\infty}^{\varepsilon^*} g(\varepsilon) d\varepsilon = N \exp\left(\frac{\varepsilon^*}{\sigma}\right). \quad (8)$$

This implies that $\varepsilon^* < 0$, which puts a constraint on the density N discussed below. Introducing for each site i with $\varepsilon_i < \varepsilon^*$ a "radius" r_i according to

$$r_i = \frac{a}{2kT}(\varepsilon^* - \varepsilon_i), \quad (9)$$

one can express the bonding criterion in a purely geometrical form:

$$d_{ij} \leqslant r_i, \quad d_{ij} \leqslant r_j. \quad (10)$$

One can find the distribution function $g(r)$ of the "radii" of sites, using the relation $|g(r) dr| = |g(\varepsilon) d\varepsilon|$:

$$g(r) = \frac{g(\varepsilon)}{|dr/d\varepsilon|} = \frac{N}{\sigma} \exp(\varepsilon/\sigma) \frac{2kT}{a}. \quad (11)$$

Expressing ε via r in accordance with Eq. (9) and using Eq. (8) one obtains

$$g(r) = \frac{N^*}{L^*} \exp(-r/L^*), \quad (12)$$

where the quantity

$$L^* = \frac{a\sigma}{2kT} \quad (13)$$

has the meaning of the average "radius."

Since L^* represents a natural length scale for the "radii," it is convenient to measure d_{ij} and r_i in units of L^*, introducing dimensionless quantities

$$\tilde{d}_{ij} = d_{ij}/L^*, \quad \tilde{r}_i = r_i/L^*. \quad (14)$$

Then, the bonding criterion becomes

$$\tilde{d}_{ij} \leqslant \tilde{r}_i, \quad \tilde{d}_{ij} \leqslant \tilde{r}_j, \quad (15)$$

and the distribution of "dimensionless radii" becomes exponential with the mean value 1:

$$g(\tilde{r}) = n \exp(-\tilde{r}), \quad (16)$$

where n is a dimensionless concentration of sites having positive "radii,"

$$n = N^*(L^*)^3. \quad (17)$$

Now, it will be shown that n does not depend on physical parameters (such as σ, T, a, etc.), and has a universal value. For this purpose, a purely geometrical percolation problem is formulated that will provide the desired quantity. Consider a set of spheres of different radii in an infinite three-dimensional space [Fig. 1(a)]. The distribution of their radii is exponential with the mean value 1. Centers of spheres are randomly placed, uncorrelated with their radii. The average number of spheres per unit cube is n. If the center of sphere i lies inside the sphere j, and simultaneously the center of sphere j lies inside the sphere i, then we call these spheres connected to each other [Fig. 1(b)]. If there is a path from one sphere to another one via connected pairs of spheres, we call these spheres belonging to the same cluster. The percolation problem is to find the
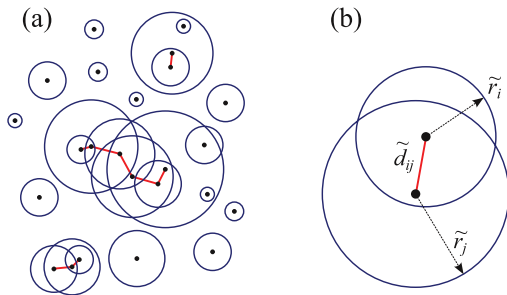
FIG. 1. (Color online) The system of spheres of different radii (a) and the connection criterion (b). Lines join centers of connected spheres.

critical concentration n_c such that there exists no infinite cluster when $n < n_c$ and there is an infinite cluster when $n > n_c$. The numerical solution of this problem with the result $n_c \approx 0.219$ is described in Sec. III.

It is easy to see that the dimensionless concentration n defined by Eq. (17) is equal to n_c. Indeed, consider sites with energies $\varepsilon_i < \varepsilon^*$, using L^* as a unit of length, and for each of such sites i draw a sphere with radius $\tilde{r}_i$ centered at this site. According to Eq. (16), the distribution of the spheres' radii is exponential with the mean value of unity. The concentration of spheres is n. Since the distance between spheres i and j is $\tilde{d}_{ij}$, the bonding criterion given by Eq. (15) is equivalent to these spheres being connected as shown in Fig. 1(b). Therefore, the system of connected resistors with bonding criterion Eq. (5) perfectly maps onto the system of connected dimensionless spheres described above. Moreover, by definition of R_{crit}, the system of resistors is exactly at the percolation threshold, hence the same holds true for the corresponding system of spheres, which means that the concentration of spheres n is equal to the critical one:

$$n = n_c \approx 0.219 \,. \tag{18}$$

Using Eqs. (8), (13), (17), and (18), one obtains the following equation for the (yet unknown) "critical energy" ε^*:

$$N \exp\left(\frac{\varepsilon^*}{\sigma}\right)\left(\frac{a\sigma}{2kT}\right)^3 = n_c. \tag{19}$$

For the case of low temperatures, the Fermi energy ε_{F} is related to the carrier concentration n_e via

$$\int_{-\infty}^{\varepsilon_{\text{F}}} g(\varepsilon)d\varepsilon = N \exp\left(\frac{\varepsilon_{\text{F}}}{\sigma}\right) = n_e. \tag{20}$$

Expressing ε^* and ε_{F} from the last two equations, and substituting their values into Eq. (6), one obtains the following result for the critical resistance R_{crit}:

$$R_{\text{crit}} = \frac{kT}{e^2 \nu_0}\left[n_c\left(\frac{2kT}{a\sigma}\right)^3 n_e^{-1}\right]^{\sigma/kT}. \tag{21}$$

Due to its exponential form given by Eq. (6), the critical resistance R_{crit} represents the most strongly varying factor in the volume resistivity ρ as a function of the model parameters. The other (pre-exponential) factors in ρ are more slowly varying. Nevertheless, the latter are also important,

primarily because they provide the correct dimensionality of ρ. Therefore, the resistivity ρ must be calculated, taking into account also the pre-exponential factors.

As explained in Section 5.6 of the monograph by Shklovskii and Efros,[16] the conductivity of a strongly inhomogeneous medium is defined almost completely by the so-called *critical subnetwork*. This subnetwork is an infinite cluster formed by resistances R_{ij} that are smaller than *or comparable with* the critical value R_{crit}:

$$R_{ij} \leqslant R_{\text{crit}} \exp(\delta), \tag{22}$$

where δ is a quantity of order unity which we can consider as a constant.

In order to obtain the resistivity of the resistor network, cut out from the resistor network a cube of some given size L and consider the resistance $R(L)$ between opposite faces of this cube. There are two limiting cases depending on how L compares with L_{corr}, the correlation length of the critical subnetwork. If $L \gg L_{\text{corr}}$, then the cube content can be treated as a continuous medium, yielding

$$R(L) = \rho/L, \quad \text{if } L \gg L_{\text{corr}},$$

where ρ is the macroscopic resistivity. In the opposite case, $L \ll L_{\text{corr}}$, the cube resistance is determined by the "local" critical resistance $\tilde{R}_{\text{crit}}$, i.e., a resistance at which the opposite sides of the cube become connected when resistors are "switching on" in ascending order:

$$R(L) = \tilde{R}_{\text{crit}}, \quad \text{if } L \ll L_{\text{corr}}.$$

The two cases meet each other at $L \simeq L_{\text{corr}}$. Moreover, the values of $\tilde{R}_{\text{crit}}$ are of the same order of magnitude as $R_{\text{crit}} \equiv R_0 \exp(\xi_c)$ for nearly all cubes of the size L_{corr}. Therefore

$$R(L_{\text{corr}}) \simeq \rho/L_{\text{corr}} \simeq R_{\text{crit}},$$

which gives

$$\rho \simeq L_{\text{corr}} R_{\text{crit}}. \tag{23}$$

Here and below we use the sign $\simeq$ to indicate equality up to some dimensionless factor of order unity.

In order to find L_{corr}, it is useful to start from the bonding criterion for the critical subnetwork given by Eq. (22). Expressed via ε^* in Eq. (6), this criterion reads

$$\frac{2d_{ij}}{a} + \frac{\varepsilon_i}{kT} \leqslant \frac{\varepsilon^*}{kT} + \delta, \quad \frac{2d_{ij}}{a} + \frac{\varepsilon_j}{kT} \leqslant \frac{\varepsilon^*}{kT} + \delta. \tag{24}$$

The difference between Eq. (7) and Eq. (24) can be eliminated by the following increase of ε^*:

$$\varepsilon^* \to \varepsilon^* + \delta kT. \tag{25}$$

Therefore the renormalization (25) allows us to turn the consideration of the percolation threshold into a consideration of the critical network, while the derivation of Eqs. (8)–(17) remains unchanged. According to Eq. (13), the value of L^* for the critical subnetwork does not depend on ε^*, and therefore stays the same as for the percolation threshold. However, the values of N^* and n are increased by a factor $\exp(\delta kT/\sigma)$, because they depend on ε^* due to Eqs. (8) and (17). Hence, the dimensionless concentration $n_{\text{subnetwork}}$ of sites necessary to form the critical subnetwork is

$$n_{\text{subnetwork}} = n_c \exp(\delta kT/\sigma), \tag{26}$$

where n_c is the corresponding concentration at the percolation threshold.

It is known[16] that the correlation length of the infinite cluster scales with the site concentration n as $(n - n_c)^{-\nu}$, where ν is the known critical index of the correlation length. This allows estimating the dimensionless correlation length of the critical subnetwork, L_{corr}/L^*, as

$$L_{\text{corr}}/L^* \simeq (n_{\text{subnetwork}} - n_c)^{-\nu}. \tag{27}$$

The value of ν depends only on the dimensionality of space. In 3D, $\nu = 0.875 \pm 0.008$.[17] Since

$$n_{\text{subnetwork}} - n_c = n_c[\exp(\delta kT/\sigma) - 1] \simeq \frac{kT}{\sigma}, \tag{28}$$

it follows that

$$L_{\text{corr}} \simeq L^* \left(\frac{\sigma}{kT}\right)^{\nu}. \tag{29}$$

Note that the factor n_c in Eq. (28) is dropped in the Taylor expansion since all constant prefactors are described via a single fitting coefficient. Combining Eqs. (13), (23), and (29), one obtains the prefactor of the resistivity ρ:

$$\rho \simeq \frac{a\sigma}{kT} \left(\frac{\sigma}{kT}\right)^{\nu} R_{\text{crit}}. \tag{30}$$

Finally, substituting Eq. (21) for R_{crit} one obtains for the resistivity ρ:

$$\rho = A \frac{a\sigma}{e^2 \nu_0} \left(\frac{\sigma}{kT}\right)^{\nu} \left[8n_c \left(\frac{kT}{a\sigma}\right)^3 n_e^{-1}\right]^{\sigma/kT}, \tag{31}$$

where A is a numerical constant. This equation is the central result of the current report for the resistivity of a system with exponential DOS in the frame of the VRH that contains only one unknown dimensionless constant A. The latter is determined via fitting to computer simulation results described in Sec. IV.

The above results are applicable under the following conditions.

(i) The temperature T must be low enough:

$$kT \ll \sigma. \tag{32}$$

This requirement is used in Eqs. (27) and (28), where it is important that $(n_{\text{subnetwork}} - n_c) \ll 1$. It is also used for the evaluation of the Fermi energy ε_{F} via Eq. (20).

(ii) The electron concentration n_e should be low enough to guarantee that the Fermi level ε_{F} is far below the "critical energy" ε^*:

$$\left[8n_c \left(\frac{kT}{a\sigma}\right)^3 n_e^{-1}\right]^{\sigma/kT} \gg 1. \tag{33}$$

This condition allows one to use Eq. (4) for the resistances R_{ij}. One can treat condition (33) also as a restriction on temperature T from below. The condition is satisfied if the bracket in the l.h.s. is larger than unity, i.e., if $kT/\sigma > (n_e a^3/8n_c)^{1/3}$. Furthermore, one can show that condition (33) determines temperatures far above those, at which the well-known Mott law for VRH is expected: The Mott law is valid at temperatures satisfying the strong inequality $kT(T_0/T)^{1/4} \ll \sigma$, where $T_0 = 1/[kg(\varepsilon_{\text{F}})a^3]$. Using the estimate $g(\varepsilon_{\text{F}}) \approx n_e/\sigma$, one obtains the condition $kT/\sigma \ll (n_e a^3)^{1/3}$ for the validity of Mott law.

Apparently, the latter condition is fulfilled at much lower temperatures than those determined by condition (33). In other words, the range of applicability for our theory determined by condition (33) does not overlap with the range of applicability for Mott VRH.

(iii) There should be enough sites for variable-range hopping:

$$N > 8n_c \left(\frac{kT}{a\sigma}\right)^3. \tag{34}$$

This requirement is used in Eq. (8), which implies that $\varepsilon^* < 0$. One can also consider it as a restriction on temperature, rewriting it as $kT/\sigma < (Na^3/8n_c)^{1/3}$. The latter condition does not contradict to condition (33) as long as the number of charge carriers is much smaller than the total number of localized states available for hopping transport: $n_e \ll N$.

Finally, a few remarks are in order for the two-dimensional case. There are a few differences between 2D and 3D VRH in a system with exponential DOS.

(i) The dimensionless concentration is $n = N^*(L^*)^2$ instead of $n = N^*(L^*)^3$ as in Eq. (17).

(ii) A different value $n_{c2} \approx 1.303$ of the critical dimensionless concentration applies, which is obtained numerically in Sec. III. (iii) $\rho \simeq R_{\text{crit}}$ must be used instead of $\rho \simeq L_{\text{corr}} R_{\text{crit}}$ in Eq. (23).[18,19]

As a result, one obtains [instead of Eq. (31), valid in 3D] the following expression in 2D:

$$\rho = A_2 \frac{kT}{e^2 \nu_0} \left[4n_{c2} \left(\frac{kT}{a\sigma}\right)^2 n_e^{-1}\right]^{\sigma/kT}, \tag{35}$$

where A_2 is a numerical constant.

III. PERCOLATION VIA SPHERES WITH DISTRIBUTED SIZES

The percolation threshold n_c was found using the algorithm described by Newman and Ziff.[20] According to this algorithm, spheres are generated randomly and added to the system one by one. For each sphere, one keeps track of the cluster to which it belongs. Initially, each sphere forms its own cluster. When a sphere added to the system is connected with already existing spheres, the clusters are merged. An efficient so-called union/find algorithm is used to keep track of the clusters.[20,21] Spheres are added until a connected cluster appears that spans the system in a given direction. Periodic boundary conditions are used, so that the system is considered to percolate when a cluster wraps around the system in a given direction and connects to itself. The linked cell method[22] is used to efficiently find the neighbors of each newly added sphere, in order to avoid iterating through a list of all spheres for every sphere that is added.

When determining the percolation threshold numerically, two kinds of errors affect the result. First, the concentration required for a spanning cluster will be different from one realization to the next. This error can be reduced by averaging over many realizations. The second error is a finite size effect: the critical concentration in a finite system depends[23] on the system size L,

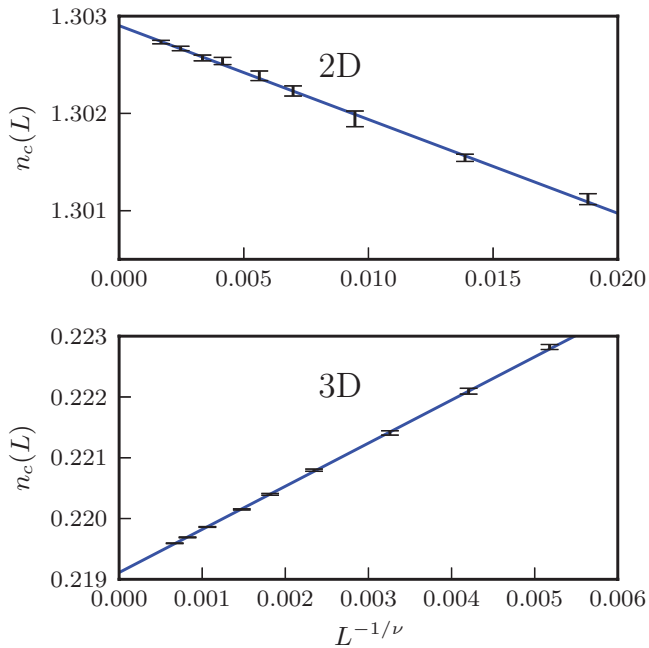$$n_c(L) - n_c \propto L^{-1/\nu}. \tag{36}$$

FIG. 2. (Color online) Finite size scaling of the percolation threshold n_c for disks (2D) and spheres (3D) with exponentially distributed radii. The scaling exponent ν is 0.875 in 3D,[17] and 4/3 in 2D.[23]

To obtain the limit of infinite system size one can plot $n_c(L)$ against $L^{-1/\nu}$, perform a linear fit, and extrapolate n_c for $1/L \to 0$. Figure 2 shows these plots for percolation in two and three dimensions, and demonstrates that the scaling law (36) applies in both cases. Between 20 000 and 40 000 realizations for each system size L were evaluated in 3D, and over 70 000 for each L in 2D. By extrapolating to $1/L \to 0$, the critical concentration of spheres in 3D,

$$n_c \approx 0.219 , \qquad (37)$$

was found. In two dimensions, the corresponding critical concentration of disks is $n_{c2} \approx 1.303$. To verify the program, it was applied to the case of equally sized spheres, where the percolation threshold is accurately known,[24] $n_c^{\mathrm{uni}} \approx 0.65296$. The known value was reproduced with an accuracy of four digits.

IV. STRAIGHTFORWARD SIMULATION OF THE VRH IN THE EXPONENTIAL DOS

In order to test the resistivity expression (31) derived in Sec. II, a numerical simulation of the charge transport was performed. The simulation also provides a value for the prefactor A in Eq. (31). The numerical results were obtained using the method of balance equations,[25–28] linearized for the limit of a small electric field.[29] Details of the procedure are given in the Appendix.

Simulation results for the temperature dependence of the mobility are shown in Fig. 3, and for the concentration dependence in Fig. 4. The data were obtained for a system with 125 000 sites with the localization length chosen as $a = 0.3N^{-1/3}$. The simulation was repeated for five realizations of the system, yielding error bars smaller than the symbol size.
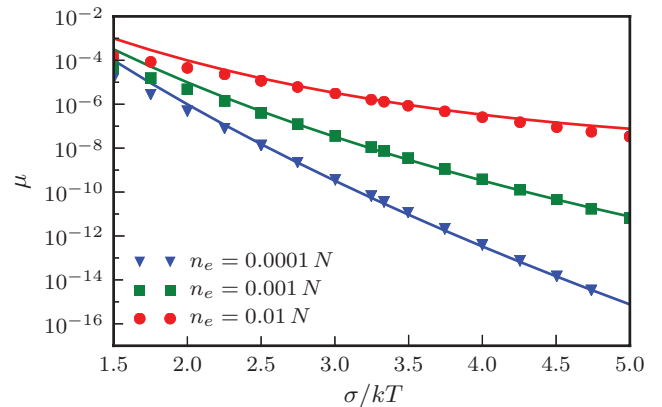

FIG. 3. (Color online) Temperature dependence of the mobility. The symbols show simulation results, while the curves are calculated with Eq. (31), for $N^{1/3}a = 0.3$. Deviations between the simulation results and the theory can be seen at high temperatures, when conditions (32) and (34) are not satisfied. For the highest concentration, a deviation is seen also at low temperatures where condition (33) is violated.

Fitting Eq. (31) against the simulation results, the value

$$A = 0.36 \qquad (38)$$

was obtained for the constant in the resistivity expression (31). Figures 3 and 4 show good agreement between expression (31) and the simulation results, within the limits given by the conditions (32)–(34). Note that A is the only fitted parameter.

V. COMPARISON WITH LITERATURE RESULTS FOR EXPONENTIAL DOS

The result in the form of Eq. (31) for the resistivity of a 3D system in the variable-range-hopping (VRH) regime and the computer simulations described in Sec. IV allows for checking the validity of theoretical approaches tried in the literature for the same problem or similar ones. After the basics for theoretical description of the VRH were clarified in
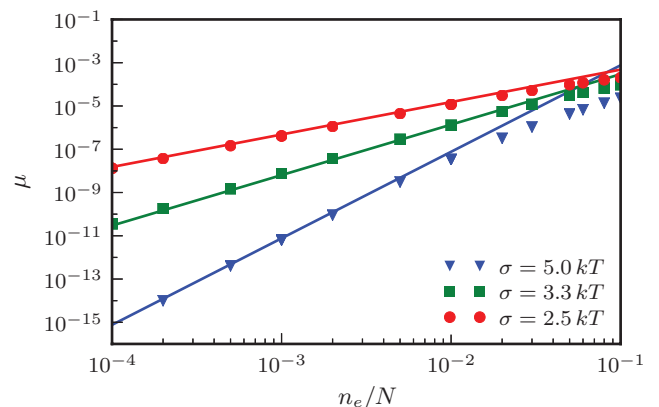

FIG. 4. (Color online) Concentration dependence of the mobility. The symbols show simulation results, while the curves are calculated with Eq. (31), for $N^{1/3}a = 0.3$. Deviations between the simulation results and the theory appears at high concentrations when condition (33) is violated.

TABLE I. Values for **A** and **B** in Eq. (40) for the different compared expressions of the mobility μ. C is an unknown numerical coefficient, $\nu \approx 0.875$ and $n_c \approx 0.219$ are known from percolation theory (see Sec. III), and $B_c = 2.7$ is a percolation threshold taken from Ref. 16.

Article	**A**	**B**
This paper, Eq. (31)	$0.36 \left(\frac{kT}{\sigma}\right)^{\nu}$	$\frac{1}{n_c}$
Oelerich *et al.* (Ref. 4)	$\frac{B_c}{2\pi}$	$\frac{4\pi}{3B_c} \frac{27}{\exp(3)}$
Vissenberg and Matters (Ref. 5)	C	$\frac{\pi}{B_c}$
Grünewald and Thomas (Ref. 10)	$\frac{1}{3}$	$\frac{4\pi}{3B_c} \frac{68}{27}$

the early 1970s[7–9] the main effort of researchers was directed to studying the effect of the DOS for the VRH charge transport. Grünewald and Thomas[10] were the first who studied hopping transport in the exponential DOS given by Eq. (1) within the percolation approach. Almost two decades later the same problem in the framework of the percolation approach was addressed by Vissenberg and Matters.[5] Vissenberg and Matters claimed to use a somewhat different percolation criterion as compared to that used by Grünewald and Thomas,[10] although no comparison with the result of Grünewald and Thomas has been performed by Vissenberg and Matters, and it is not clear which percolation approach is more accurate.

A percolation approach has been derived also for the VRH in systems with Gaussian DOS, valid for organic disordered semiconductors.[11,12] In this derivation, the percolation criterion from classical papers[9,10] and not that of Vissenberg and Matters[5] has been used. It is therefore highly desirable to check the validity of the classical recipe.[9,10] Furthermore, aiming at a description of the VRH in disordered organic materials with Gaussian DOS, a transparent approach based on the concept of transport energy has recently been put forward.[4] The latter approach should work also for the case of the exponential DOS, and herewith its validity can be checked by comparison with Eq. (31). Therefore, the results of Grünewald and Thomas,[10] the results of Vissenberg and Matters,[5] and the results of Oelerich *et al.*[4] for the VRH in the exponential DOS are compared with that in Eq. (31) based on the mapping of the problem onto a geometrical problem.

For the sake of clarity, the results obtained for the charge carrier mobility μ are compared, which is related to the resistivities ρ via

$$\mu^{-1} = e n_e \rho. \tag{39}$$

All analytical solutions mentioned above can be represented in the following form:

$$\mu = \mathbf{A} \frac{\nu_0 e}{n_e \sigma a} \left[\mathbf{B} \frac{n_e}{8} \left(\frac{a\sigma}{kT}\right)^3 \right]^{\sigma/kT}, \tag{40}$$

where **A** and **B** are different factors for different approaches to be checked. Table I lists the different approaches and the corresponding values of **A** and **B**. Note that in all cases, these values are either constants or slowly varying functions of σ/kT so that the strong dependencies on system parameters are equal in all solutions.

In Fig. 5, the results from Eq. (40) and Table I are plotted together with data points obtained by computer simulations.
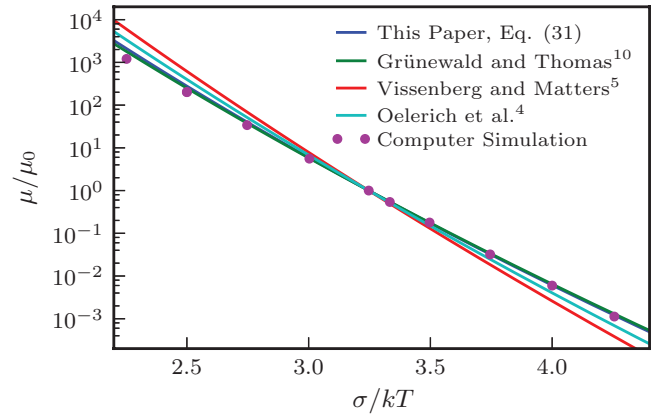

FIG. 5. (Color online) Normalized mobility values obtained from computer simulation and Eq. (40) for the different analytical solutions, for $N^{1/3}a = 0.3$ and $n_e/N = 0.001$.

Each data set is normalized to its value at $kT \approx 0.3\sigma$ so that only deviations in the temperature dependencies are shown.

It is clear from the figure that the result given by Eq. (31) and the one of Grünewald and Thomas[10] show very similar dependencies on temperature, and they match the simulation results. This is particularly remarkable not only because the result of Grünewald and Thomas[10] was the first obtained for the VRH in the exponential DOS, but mainly because it was based on a simplified approach with percolation threshold $B_c = 2.7$ for overlapping spheres with equal sizes. The result of Eq. (31) based on the geometrical problem with distributed sizes of spheres agrees with the simulation results even better. The comparison shown in Fig. 5 clearly confirms the validity of the approach suggested by Oelerich *et al.*[4] for description of the VRH based on the transport-energy concept.

The worst performance in comparison with the simulation data in Fig. 5 is demonstrated by the results of Vissenberg and Matters.[5] Therefore, it is worthwhile to discuss why the two percolation results of Grünewald and Thomas and of Vissenberg and Matters differ from each other. In both approaches, the average number of bonds per site is used as a percolation criterion. The system is assumed to percolate, when the sites *in the largest cluster* in average have $B_c \approx 2.7$ bonds each. However, in the two approaches the average is evaluated in different ways. Vissenberg and Matters evaluate the average number of bonds as the total number of bonds in the system divided by the number of sites with energies low enough to have bonds. Grünewald and Thomas[10,30] (following Pollak[9]) performed weighted averaging for these two quantities, where the weight of each site is the probability for that site to belong to the largest cluster. This probability is taken to be proportional to the number of bonds connected to the site.

This difference in the averaging procedure could be the reason for the different temperature dependencies obtained. While our result is very close to that of Grünewald and Thomas, our approach is different to theirs. We do not use the number of bonds per site as the percolation criterion, making the derivation more transparent. As the difference between the results of the two percolation approaches show, the number

of bonds is a nontrivial quantity to use when all sites are not equivalent.

One can view the good agreement between our numeric and analytic results with those of Grünewald and Thomas as supporting the averaging procedure used by Grünewald and Thomas.[10] This information is important for applying the percolation theory to systems with a DOS different from the exponential one considered in the current report.

VI. HOPPING TRANSPORT ON A LATTICE

So far, variable range hopping over sites randomly distributed in space was considered. With the percolation approach developed in Sec. II, the mobility was determined for the particular case of an exponential distribution of site energies. In another special case, that of nearest-neighbor hopping on a lattice, the mobility can be found for *any* shape of the DOS, using a similar percolation approach. For this reason, a discussion of the lattice case is included here. The only properties of the DOS needed are the characteristic width σ and the critical energy ε^* required for percolation.

Consider a lattice of sites with the lattice constant l_0, where resistances R_{ij} connect nearest-neighbor sites i and j:

$$ R_{ij} = \frac{kT}{e^2 \nu_0} \exp\left(\frac{2l_0}{a}\right) \exp\left[\frac{\max(\varepsilon_i, \varepsilon_j) - \varepsilon_F}{kT}\right]. \quad (41) $$

The energy-dependent exponent is denoted by ξ_{ij}, so that $R_{ij} = R_0 \exp(\xi_{ij})$. Let $\sigma_{dc}(\xi)$ be the dc conductivity of the lattice, in which resistances with $\xi_{ij} > \xi$ are cut off. There is a critical value ξ_c, such that $\sigma_{dc}(\xi) > 0$ only when $\xi > \xi_c$. The conductivity rapidly increases above ξ_c, and saturates at $\xi \approx \xi_c + O(1)$, determined by the critical subnetwork of resistances with $\xi_{ij} < \xi_c + O(1)$.[16] As in the VRH case, the resistivity of the system is determined using Eq. (23), where L_{corr} and R_{crit} are obtained for the lattice case below. The critical resistance in the lattice case can be written as

$$ R_{crit} = \frac{kT}{e^2 \nu_0} \exp\left(\frac{2l_0}{a}\right) \exp\left(\frac{\varepsilon^* - \varepsilon_F}{kT}\right), \quad (42) $$

where ε^* again denotes the critical energy. Because the resistance expression (41) depends only on the larger of the two energies ε_i and ε_j, the bond between sites i and j has $\xi_{ij} < \xi_c$ and is present in the percolating network if both inequalities $\varepsilon_i < \varepsilon^*$ and $\varepsilon_j < \varepsilon^*$ hold. This turns the problem of finding ε^* into a *site*-percolation problem, and allows for determining the critical energy using the site-percolation threshold p_s of the lattice,

$$ \int_{-\infty}^{\varepsilon^*} g(\varepsilon) d\varepsilon = p_s, \quad (43) $$

with the threshold p_s depending on the lattice type.

The correlation length L_{corr} of the percolating network above the percolation threshold p_c is determined by the scaling relation[16]

$$ L_{corr} \simeq l_0 (p - p_c)^{-\nu}, \quad (44) $$

where p is the fraction of bonds with $\xi_{ij} < \xi$ and l_0 is the lattice constant. This expression is similar to Eq. (29), but formulated in quantities appropriate for the lattice problem.

For a distribution of site energies characterized by the width σ, the distribution of ξ_{ij} has a width of order σ/kT. Then,

$$ p - p_c = (\xi - \xi_c)\frac{dp}{d\xi} \simeq (\xi - \xi_c)\left(\frac{\sigma}{kT}\right)^{-1} \simeq \frac{kT}{\sigma}, \quad (45) $$

and

$$ L_{corr} \simeq l_0 \left(\frac{\sigma}{kT}\right)^\nu. \quad (46) $$

Using Eqs. (23), (42), and (46), one obtains the resistivity ρ in the form

$$ \rho \simeq \frac{kT l_0}{e^2 \omega_0}\left(\frac{\sigma}{kT}\right)^\nu \exp\left(\frac{\varepsilon^* - \varepsilon_F(T, n_e)}{kT}\right), \quad (47) $$

with $\omega_0 = \nu_0 \exp(-2l_0/a)$ for consistency with the VRH expressions.

A scaling theory for charge transport on a regular lattice with Gaussian distribution of site energies was recently presented[13] by Cottaar *et al.* The main result is an expression for the charge carrier mobility, as a function of the fraction of occupied sites c and temperature T,

$$ \mu(T, c) = B\frac{e\omega_0}{N^{2/3}kTc}\left(\frac{kT}{\sigma}\right)^\lambda \exp\left[\frac{\varepsilon_F(T, c) - \varepsilon^*}{kT}\right]. \quad (48) $$

Here, B and λ are parameters assumed to be independent of T and c. These two parameters were determined by fitting Eq. (48) to mobility data obtained in numerical simulations.

Expressing the mobility μ as $(en_e\rho)^{-1}$ with ρ given by Eq. (47), the carrier concentration n_e as Nc, and the lattice constant l_0 as $N^{-1/3}$, the expression (48) for the mobility μ is obtained, except that ν replaces λ. Therefore the exponent λ introduced in Ref. 13 is equal to the critical exponent ν of the correlation length of the percolation cluster, with the magnitude $\nu = 0.875 \pm 0.008$ in the 3D case.[17] The exponent ν is independent of the details of the lattice and the hopping rate model, and depends only on the number of spatial dimensions of the problem. The values for λ obtained by fitting in Ref. 13 are very close to ν.

VII. SUMMARY

A percolation approach was developed to solve the variable range hopping (VRH) problem in the exponential DOS. By mapping the physical VRH problem onto a purely geometrical problem using randomly placed spheres with distributed sizes, an expression was derived for the hopping resistivity in the exponential DOS in two and three dimensions, exact up to a single numerical coefficient. The solution contains the pronounced exponential dependencies of the resistivity on material parameters as well as the more weakly dependent prefactors.

The purely geometrical percolation problem of spheres with distributed sizes was studied in detail, and the percolation thresholds were found with high accuracy in two and three dimensions using numerical simulations. The single fitting parameter in the analytical expression for the resistivity was determined via straightforward computer simulations. Comparison of the analytical theory with the simulation results shows a good agreement within the range of parameters, for which the theory is applicable.

The obtained expression for the VRH mobility was compared to the literature results. The result of Grünewald and Thomas[10] performs remarkably well in comparison to the analytical solution and to the computer simulations.

Finally, the problem of the nearest-neighbor hopping on a lattice was discussed. It is shown that the percolation approach provides a solution for this problem including pre-exponential factors also for DOS functions different from the exponential one.

ACKNOWLEDGMENTS

APPENDIX: BALANCE EQUATIONS FOR THE OHMIC REGIME

The numerical results for the charge carrier mobility presented in Sec. IV were obtained by solving balance equations for the charge carrier flow. In this method, a system of randomly placed sites with randomly chosen energies is generated. The steady state current in response to an electric field is then calculated for the system. First, the occupation probability for each site is determined by solving a set of balance equations. The occupation probabilities are then used to evaluate the current. The system is considered with periodic boundary conditions in all directions, so that when an electric field is applied, there will be a net current flowing through the system in steady state.

In the steady state, the electron flow into each site equals the electron flow out,

$$\sum_j n_i \Gamma_{ij}(1 - n_j) = \sum_j n_j \Gamma_{ji}(1 - n_i) , \quad \text{(A1)}$$

where n_i is the probability that site i is occupied by an electron.

Equation (A1) is a nonlinear system of equations, with one equation for each site in the system. When the charge carrier concentration can be considered small, the equations are typically linearized by omitting the factors $(1 - n_i)$, because the resulting linear equation system is much easier to solve than the full nonlinear one.[31,32]

For the present problem, the charge carrier concentration can not be considered small. However, the equations can be linearized for the limit of a small electric field.[29] Only the small-field limit is of interest in the current case, because the percolation theory is derived for the Ohmic regime.

Following Ref. 29, we define n_i^0 as the occupation probability of site i at zero electric field. These occupation probabilities follow the Fermi-Dirac distribution,

$$n_i^0 = f(\varepsilon_i, \varepsilon_F) = \frac{1}{1 + \exp\left[(\varepsilon_i - \varepsilon_F)/kT\right]}, \quad \text{(A2)}$$

since the hopping rates obey detailed balance. The problem will be solved in the variables

$$p_i = \frac{n_i - n_i^0}{n_i^0 (1 - n_i^0)} , \quad \text{(A3)}$$

which measure the electric-field-induced deviations in the probabilities n_i. Further, the hopping rate Γ_{ij} is split into two parts,

$$\Gamma_{ij} = \Gamma_{ij}^0 + \Delta\Gamma_{ij} , \quad \text{(A4)}$$

where Γ_{ij}^0 denotes the hopping rate from site i to j in the absence of an electric field and $\Delta\Gamma_{ij}$ is the field-induced change. The zero-field quantities obey the detailed balance relation

$$n_i^0 \Gamma_{ij}^0 \left(1 - n_j^0\right) = n_j^0 \Gamma_{ji}^0 \left(1 - n_i^0\right) . \quad \text{(A5)}$$

Next, n_i from Eq. (A3) and Γ_{ij} from Eq. (A4) are inserted into Eq. (A1). Terms quadratic in p are omitted, and Eq. (A5) is used to cancel terms without p:s. Finally, $\Delta\Gamma_{ij}$ is expressed by linearizing the hopping rate (3), yielding

$$-p_i \sum_j \gamma_{ij} + \sum_j p_j \gamma_{ij} = \frac{eE}{kT} \sum_j \gamma_{ij}(z_j - z_i) , \quad \text{(A6)}$$

where E is the magnitude of the electric field (the field is assumed to be directed in the z direction), and

$$\gamma_{ij} = n_i^0 \Gamma_{ij}^0 \left(1 - n_j^0\right) . \quad \text{(A7)}$$

Equation (A6) is a linear system of equations. The equations are not all independent, since the total number of charge carriers is free to change. One further equation is given by fixing the total number of charge carriers in the system,

$$\sum_j p_j n_j^0 \left(1 - n_j^0\right) = 0 . \quad \text{(A8)}$$

One can simply replace one of the equations in (A6) with Eq. (A8), however, for numerical stability it is preferable to spread the additional constraint over the whole system by adding it, with a suitably chosen scaling factor, to *each* equation in (A6).[33] The zero-field quantities n_i^0 can be determined either by fixing the total concentration n_e of charge carriers in the system or by specifying the Fermi level. To decrease the variations of the current between different realizations of the system, the latter approach was chosen.[33] The Fermi energy ε_F was found numerically by solving

$$\int_{-\infty}^{\infty} g(\varepsilon) f(\varepsilon, \varepsilon_F) d\varepsilon = n_e \quad \text{(A9)}$$

for the desired n_e.

Once the p_i are known from solving Eq. (A6), the current through the system in the z direction is given by

$$I = e \sum_{i>j} \frac{z_j - z_i}{L_z} \gamma_{ij} \left[p_i - p_j + \frac{eE}{kT}(z_j - z_i) \right], \quad \text{(A10)}$$

where L_z is the length of the system in the z direction. This expression was obtained by evaluating the net charge carrier flow, using the same linearizations as for Eq. (A6).

The expression for the current I shows the physical meaning of the quantities p_i and γ_{ij}. Performing the linearization corresponds to considering a resistor network picture, where γ_{ij} plays the role of conductivity between sites i and j, while p_i corresponds to the potential at site i. The resulting equations are then simply Kirchhoff's laws for the network.

*jan.oliver.oelerich@physik.uni-marburg.de

[1] J. Orenstein and M. Kastner, Solid State Commun. **40**, 85 (1981).

[2] R. A. Street, *Hydrogenated Amorphous Silicon* (Cambridge University Press, Cambridge, 1991).

[3] H. Bässler, Phys. Status Solidi B **175**, 15 (1993).

[4] J. O. Oelerich, D. Huemmer, and S. D. Baranovskii, Phys. Rev. Lett. **108**, 226403 (2012).

[5] M. C. J. M. Vissenberg and M. Matters, Phys. Rev. B **57**, 12964 (1998).

[6] J. Nelson, Phys. Rev. B **67**, 155209 (2003).

[7] B. I. Shklovskii and A. L. Efros, Sov. Phys. JETP **33**, 468 (1971).

[8] V. Ambegaokar, B. I. Halperin, and J. S. Langer, Phys. Rev. B **4**, 2612 (1971).

[9] M. Pollak, J. Non-Cryst. Solids **11**, 1 (1972).

[10] M. Grünewald and P. Thomas, Phys. Status Solidi B **94**, 125 (1979).

[11] S. D. Baranovskii, I. P. Zvyagin, H. Cordes, S. Yamasaki, and P. Thomas, Phys. Status Solidi B **230**, 281 (2002).

[12] S. D. Baranovskii, I. P. Zvyagin, H. Cordes, S. Yamasaki, and P. Thomas, J. Non-Cryst. Solids **299**, 416 (2002).

[13] J. Cottaar, L. J. A. Koster, R. Coehoorn, and P. A. Bobbert, Phys. Rev. Lett. **107**, 136601 (2011).

[14] A. Miller and E. Abrahams, Phys. Rev. **120**, 745 (1960).

[15] V. Ambegaokar, S. Cochran, and J. Kurkijärvi, Phys. Rev. B **8**, 3682 (1973).

[16] B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, Berlin, 1984).

[17] C. D. Lorenz and R. M. Ziff, Phys. Rev. E **57**, 230 (1998).

[18] S. Tyč and B. I. Halperin, Phys. Rev. B **39**, 877 (1989).

[19] P. Le Doussal, Phys. Rev. B **39**, 881 (1989).

[20] M. E. J. Newman and R. M. Ziff, Phys. Rev. E **64**, 016706 (2001).

[21] D. E. Knuth, *The Art of Computer Programming, volume 1 (3rd ed.): Fundamental Algorithms* (Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, 1997).

[22] J. Thijssen, *Computational Physics*, 2nd ed. (Cambridge University Press, Cambridge, 2007).

[23] D. Stauffer and A. Aharony, *Introduction to Percolation Theory, 2nd revised edition* (Taylor & Francis, London, 1994).

[24] C. D. Lorenz and R. M. Ziff, J. Chem. Phys. **114**, 3659 (2001).

[25] Z. G. Yu, D. L. Smith, A. Saxena, R. L. Martin, and A. R. Bishop, Phys. Rev. Lett. **84**, 721 (2000).

[26] Z. G. Yu, D. L. Smith, A. Saxena, R. L. Martin, and A. R. Bishop, Phys. Rev. B **63**, 085202 (2001).

[27] W. F. Pasveer, J. Cottaar, C. Tanase, R. Coehoorn, P. A. Bobbert, P. W. M. Blom, D. M. de Leeuw, and M. A. J. Michels, Phys. Rev. Lett. **94**, 206601 (2005).

[28] J. Cottaar and P. A. Bobbert, Phys. Rev. B **74**, 115204 (2006).

[29] W. F. Pasveer, P. A. Bobbert, H. P. Huinink, and M. A. J. Michels, Phys. Rev. B **72**, 174204 (2005).

[30] H. Overhof, Phys. Status Solidi B **67**, 709 (1975).

[31] S. Marianer and B. I. Shklovskii, Phys. Rev. B **46**, 13100 (1992).

[32] A. V. Nenashev, F. Jansson, S. D. Baranovskii, R. Österbacka, A. V. Dvurechenskii, and F. Gebhard, Phys. Rev. B **78**, 165207 (2008).

[33] J. Cottaar, Ph.D. thesis, Technische Universiteit Eindhoven, 2012.

Article III

J. O. Oelerich, F. Jansson, A. V. Nenashev, F. Gebhard, and S. D. Baranovskii

Energy position of the transport path in disordered organic semiconductors

J O Oelerich[1], **F Jansson**[1,2], **A V Nenashev**[3,4], **F Gebhard**[1] and **S D Baranovskii**[1]

[1] Department of Physics and Material Sciences Center, Philipps-University, D-35032 Marburg, Germany
[2] Section Computational Sciences, University of Amsterdam, Amsterdam, The Netherlands
[3] Laboratory of Non-Equilibrium Semiconductor Systems, Institute of Semiconductor Physics, 630090 Novosibirsk, Russia
[4] Physical Faculty Novosibirsk State University, 630090 Novosibirsk, Russia

E-mail: jan.oliver.oelerich@physik.uni-marburg.de

Abstract

The concept of *transport energy* is the most transparent theoretical approach to describe hopping transport in disordered systems with steeply energy dependent density of states (DOS), in particular in organic semiconductors with Gaussian DOS. This concept allows one to treat hopping transport in the framework of a simple multiple-trapping model, replacing the mobility edge by a particular energy level called the transport energy. However, there is no consensus among researchers on the position of this transport level. In this article, we suggest a numerical procedure to find out the energy level most significantly contributing to charge transport in organic semiconductors. The procedure is based on studying the effects of DOS modifications on the charge carrier mobility in straightforward computer simulations. We also show why the *most frequently visited energy*, computed in several numerical studies to determine the transport energy, is not representative for charge transport.

Keywords: organic semiconductors, transport energy, electronic transport

(Some figures may appear in colour only in the online journal)

1. Introduction

Although the description of charge transport in disordered organic semiconductors is of high importance for device applications and for the fundamental understanding of material properties, a comprehensive, transparent and self-consistent theory that can accurately predict transport properties is yet to be found. This is particularly unsatisfactory because a simple and clearly defined model framework existed for years: Charge carriers move via incoherent tunneling (*hopping*) between spatially and energetically distributed localized states with a Gaussian energy spectrum [1–3]:

$$g(\varepsilon) = \frac{N}{\sigma\sqrt{2\pi}}\exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right), \qquad (1)$$

where σ is the energy scale of the *density of states* (DOS), usually estimated as $\sigma \approx 0.1$ eV [1] and N is the concentration of randomly distributed localized states (sites) that charge

carriers can use for hopping transport. The rate for a carrier transition from an occupied site i to an empty site j, separated by a distance r_{ij}, is usually assumed [1, 4] to be described by the Miller–Abrahams expression [5]

$$\nu_{ij} = \nu_0 \exp\left(-\frac{2r_{ij}}{\alpha} - \frac{\varepsilon_j - \varepsilon_i + |\varepsilon_i - \varepsilon_j|}{2kT}\right). \qquad (2)$$

Here, α is the localization length of charge carriers in the localized states, usually of the order of $10^{(-8)}$ cm [6, 7], ε_i and ε_j are the carrier energies on sites i and j, respectively, k is the Boltzman constant and T is the temperature. The energy difference between states i and j must be compensated, for example, by absorption or emission of phonons. The prefactor ν_0, often called the *attempt-to-escape* frequency, depends on the interaction mechanism that causes transitions. In the case of interaction with phonons, its value is usually assumed close to the phonon frequency [4], $\nu_0 \approx 10^{12}$ s^{-1}. Due to the interplay between the spatial term in equation (2), dependent on
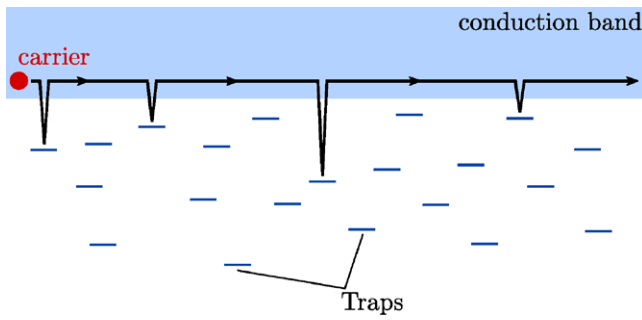
Figure 1. Schematic view of the MT transport mode. The carrier motion in the conduction band is interrupted by trapping into and subsequent activation from localized states (traps).

Figure 2. Schematic view of charge transport via the transport energy in a system with Gaussian DOS. The carrier distribution, with its maximum at the equilibration energy $-\sigma^2/kT$, is drawn with a red background.

parameters characterized by the product $N\alpha^3$ and the energy term, dependent on temperature and site energies characterized by the ratio σ/kT, charge transport at low temperatures is dominated by more distant transitions than at higher temperatures, i.e., transport takes place in the *variable-range-hopping* (VRH) mode [8]. If spatial positions and energies of localized states can be considered independent from each other, this model is traditionally called the *Gaussian Disorder Model* (GDM).

The most comprehensive theoretical tool to describe hopping transport in disordered materials is the percolation theory [8–11]. Particularly simple is its application to lattice models with transitions only between nearest neighboring sites [8]. Such a transport mode is called the *nearest-neighbor-hopping* (NNH) on a lattice. Percolation theory for this kind of transport was developed in the 1970s [8]. For the NNH on a cubic lattice with Gaussian DOS given by equation (1), it predicts that thermal activation of carriers to a particular energy level E_c dominates charge transport [12]. The energy E_c is determined by the condition [12]

$$\int_{-\infty}^{E_c} g(\varepsilon)\, d\varepsilon = x_c, \tag{3}$$

where x_c is the percolation threshold for the site percolation problem. Using the value $x_c = 0.31$, Schönherr *et al* [12] obtained $E_c \approx -0.6\sigma$. A very close result $E_c \approx -0.5\sigma$ for the same problem has recently been published by Cottaar *et al* [13]. Since the value of x_c is sensitive to the geometrical structure of the chosen regular lattice, the value of the energy E_c determined via equation (3) must also depend on the lattice structure. Therefore, the NNH on a lattice can hardly be suitable to study charge transport in a spatially disordered organic material, where transport was shown to take place via VRH [1]. While the application of the percolation theory to NNH on a lattice is a straightforward exercise [8], the derivation of the percolation approach for the VRH is a much more complicated task [8]. Aside from the uniform DOS [8], only in the case of an exponential DOS, $g(\varepsilon) = g_0 \exp(\varepsilon/\varepsilon_0)$, precise results not only for the exponentially strong dependences of transport coefficients on material parameters, but also for the slowly varying preexponential factors, could recently be obtained by percolation theory applied to VRH [14].

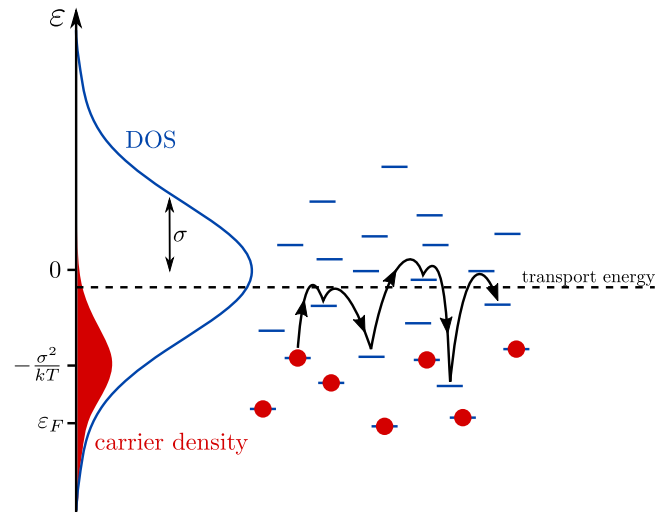An alternative powerful theoretical method to describe hopping transport in systems with steeply energy-dependent DOS, complementary to the percolation theory, is the concept of the *transport energy* (TE). The essence of the TE approach is the ability to describe hopping transport via sites distributed in space and energy in full analogy with the multiple-trapping (MT) model [15]. The MT model has been developed to describe charge transport in inorganic disordered materials, such as chalcogenide glasses and amorphous silicon [16–18], which possess a so-called *mobility edge*, i.e., the energy level that separates extended states with rather high carrier mobility from the localized states, which can be considered as traps. In the MT model a charge carrier moves only via delocalized states with energies above the mobility edge. This motion is interrupted by trapping into localized states with subsequent activation of the carrier back into the conducting states at the mobility edge. Transport in the framework of the MT model is shown schematically in figure 1. The advantage of the MT charge transport mode as compared to the VRH, is the possibility of exact theoretical treatment by analytical equations under equilibrium and non-equilibrium conditions and for various shapes of the DOS [19–21].

In the following section 2, various approaches and different definitions of the TE are summarized in order to point out the ambiguity of the term *transport energy*. In section 4 the popular numerical approach to the TE as the *most frequently visited energy* in Monte Carlo simulations is proven not to be related to the long-range transport. We suggest an alternative numerical procedure in section 5 and show that its results agree well with the analytic derivation of the transport energy by Oelerich *et al* [22], the important steps of which are briefly presented in section 3. In section 6 we address the issue of using the TE for description of non-equilibrium processes in the GDM. Concluding remarks are gathered in section 7.

2. The transport energy

In 1985 Monroe [23] has shown that in the case of the exponential DOS one can describe VRH in the framework of the

MT formalism by replacing the mobility edge in the MT model by some particular energy level ε_t, that he called the *transport energy* (TE). A slightly different derivation of ε_t was later suggested by Baranovskii *et al* [24]. The latter approach was soon extended [25] for systems with other than exponential energy spectra, in particular with a Gaussian DOS described by equation (1). Schematic movement of charge carriers via activation to the TE in a system with Gaussian DOS is shown in figure 2. It is worth emphasizing that the value ε_t of the TE is determined by the interplay between the term in equation (2) dependent on spatial parameters and the term dependent on temperature and site energies, i.e., the TE reflects the essence of the VRH transport mechanism [23–25].

Numerous derivations of the value of ε_t have been suggested in the literature and there is a lot of confusion in this respect. Studying VRH in systems with exponential DOS in the framework of the percolation theory, Grünewald and Thomas [26] have shown that in thermal equilibrium the dc conductivity σ_{dc} can be represented in the Arrhenius form,

$$\sigma_{dc} \propto \exp\left(-\frac{\varepsilon^* - \varepsilon_F}{kT}\right), \qquad (4)$$

where ε_F is the Fermi energy and ε^* is some energy level dependent on the parameters $N\alpha^3$ and kT/ε_0, where ε_0 denotes the energy scale of the exponential DOS. Later, the description of the VRH in the exponential DOS using percolation theory has been attempted by Vissenberg and Matters [27], though with a lower accuracy as compared to the initial derivation of Grünewald and Thomas [14]. The lower accuracy of the approach by Vissenberg and Matters is due to the different interpretation of the percolation criterion as compared to the classical one formulated by Pollak [11] and used by Grünewald and Thomas: In order to establish infinite percolation, the classical recipe demands an average number of $B_c \approx 2.7$ neighbors around each site, to which the transition rates are the highest and which *belong to the percolation cluster*. Vissenberg and Matters [27] instead averaged over *all* sites in the system, which leads to less accurate results as evidenced by comparison to the exact solution of the VRH problem for the exponential DOS obtained by Nenashev *et al* [14].

Percolation theory based on the classical recipe [11] has also been developed to describe the VRH in the GDM [28–30], i.e., in a random system of sites with a Gaussian energy distribution. It was shown that in thermal equilibrium at low temperatures or at high concentrations of carriers, cases in which the Fermi level ε_F is situated above the equilibration energy $\varepsilon_\infty = -\sigma^2/kT$, the temperature dependence of the dc conductivity σ_{dc} exhibits a simple activated behavior similar to equation (4) with ε^* determined by the relation

$$\varepsilon^* = -\sqrt{2\ln(Q)}\,\sigma, \quad Q = \beta\frac{N\alpha^3}{s^2}\left(\frac{\sigma}{kT}\right)^3, \qquad (5)$$

where β is a constant and $s \approx -\ln^{1/2}(Q)$ is a logarithmic function of $N\alpha^3$ and σ/kT.

The simple forms of equations (4) and (5) might lead to the conclusion that in systems with exponential or Gaussian DOS the VRH takes place by activation of charge carriers

towards the energy ε^*. However, the r-dependent term in equation (2) plays an equally important role for VRH as the thermal activation. Therefore, one can hardly assume that the description of the VRH as simple activation of carriers towards the energy level ε^* would mean that carriers actually conduct moving around that energy level. One should instead interpret the energy level ε^* in equation (4) as an 'effective' energy level serving formally to mimic the combined effects of the spatial tunneling and of the activation in energy. Therefore, percolation theory is not a suitable tool to determine the 'real' energy level that charge carriers have to visit in order to provide transport in the VRH regime. Mixing up the effective TE ε^* with the real TE ε_t is probably the main reason for the confusion with respect to the term 'transport energy'. While the effective TE ε^* serves as a formal tool to describe the VRH as a simple activation, the real TE ε_t characterizes the actual transport level at which carriers conduct via spatial tunneling.

The difference between these effective and real TEs becomes more transparent if one compares the expressions for the mobility of charge carriers μ derived in various approaches. After deriving ε_t for the GDM [25] in full analogy with the derivations for the exponential DOS, Baranovskii *et al* [31] suggested the following expression for the carrier mobility:

$$\mu = \mu_0 \exp\left[-\frac{2R(\varepsilon_t)}{\alpha} - \frac{\varepsilon_t}{kT} - \frac{1}{2}\left(\frac{\sigma}{kT}\right)^2\right], \qquad (6)$$

where $R(\varepsilon_t)$ is the typical distance between localized states with energies below ε_t. Working with the effective TE ε_{eff}, Arkhipov *et al* [32] and Nikitenko *et al* [33] suggested a slightly different expression for the carrier mobility:

$$\mu = \widetilde{\mu_0} \exp\left[-\frac{\varepsilon_{eff}}{kT} - \frac{1}{2}\left(\frac{\sigma}{kT}\right)^2\right]. \qquad (7)$$

Taking into account that the values of μ_0 in equation (6) and $\widetilde{\mu_0}$ in equation (7) are very close to each other [31–33], it is apparent that the meaning of the effective TE ε_{eff} is different to that of the TE ε_t. The effective TE ε_{eff} in equation (7) is denoted differently to the energy ε^* in equation (4) particularly because the percolation nature of the VRH has not been taken into account in the derivation of ε_{eff} by Arkhipov *et al* [32] and Nikitenko *et al* [33]. The same is true for the initial derivation of the ε_t in the GDM by Baranovskii *et al* [25]. Later Oelerich *et al* [22] derived the equation for ε_t in the GDM taking into account the percolation criterion and extending this concept for finite carrier concentrations n. This approach is outlined in section 3. The dependence of the 'effective' transport energy ε_{eff} on n has been derived by Arkhipov *et al* [34].

Another interesting definition of the TE was suggested by Schmechel [35], who, in contrast to all other derivations of the TE, did not use this quantity to calculate the carrier mobility, but instead derived the TE *from* the previously calculated dc conductivity by introducing a differential energy-dependent dc conductivity. The percolation nature of the cluster of sites responsible for the VRH has not been taken into account in Schmechel's derivation.

Martens *et al* [36] essentially repeated the derivation by Baranovskii *et al* [25] also not taking into account the percolation nature of the VRH transport.

Sometimes the following arguments are used to determine the position of the TE in the GDM [1, 37–39]. First one assumes that the temperature dependence of the carrier mobility $\mu(T)$ is completely decoupled from the dependence on the concentration of sites N. Then one uses the expression for $\mu(T)$ in the form obtained in computer simulations [1]

$$\mu \propto \exp\left[-\left(C\frac{\sigma}{kT}\right)^2\right], \tag{8}$$

with $C = \frac{2}{3}$. Assuming that transport is governed by activation of carriers from the equilibration energy $\varepsilon_\infty = -\sigma^2/kT$ towards some transport level ε_B, one obtains from equation (8) $\varepsilon_B \approx -\left(\frac{\sigma^2}{kT}\right) + 4/9\left(\frac{\sigma^2}{kT}\right) = -5/9\left(\frac{\sigma^2}{kT}\right)$. Oelerich *et al* [22] argued that the level ε_B can hardly be considered as a good candidate for the TE since in thermal equilibrium not all carriers occupy energies equal to $\varepsilon_\infty = -\sigma^2/kT$, but are instead distributed around ε_∞ in a broad energy range with the width [1] of the order σ. Carriers with energies in the upper part of that distribution contribute significantly more to transport than carriers from the lower part of the distribution, since activation to the transport level ε_t is exponentially easier for carriers with higher energies. The effective activation energy $\sigma^2/2kT$ is therefore essentially smaller than σ^2/kT.

Several algorithms have been developed and applied to determine the value of the TE by Monte Carlo computer simulations. Schönherr *et al* [12] assumed that $\varepsilon_\infty(T)$ must ultimately approach an asymptotic value as a function of temperature because of the Gaussian DOS profile. Their computer simulations suggest that for $\sigma/kT > 6$, the value $\varepsilon_\infty(T) \approx -1.8\,\sigma$ is attained. The activation energy of the mobility obtained in those simulations caused the assumption that the transport level is situated at an energy 0.70σ *above* the center of the Gaussian DOS [12]. The drastic discrepancy between this result and the value $E_c \approx -0.6\sigma$ obtained using equation (3) ultimately led to the conclusion that percolation theory is not suitable to account for hopping transport in the GDM [12]. In fact, this conclusion is due to the small size of the simulated array of sites. At small temperatures sites with energies around $\varepsilon_\infty = -\sigma^2/kT$ were statistically not present in the system and the simulated position of $\varepsilon_\infty(T)$ at low temperatures corresponded simply to the smallest energy still present in the simulated array of sites. Therefore this value did not depend on T in the simulations by Schönherr *et al* [12]. No saturation would have been recognized if the size of the simulated system had been increased.

Another computer algorithm to determine the position of the TE was used by Cleve *et al* [40], who suggested to trace the energies of sites targeted by carriers in hopping transitions during Monte Carlo simulations and to determine the position of the TE as the energy of the most frequently targeted sites. This algorithm has been used in several numerical studies [37, 41, 42]. Hartenstein and Bässler [37] recognized,

however, that due to oscillations of carriers within pairs of spatially and energetically close sites the most frequently targeted energy is not the energy essential for charge transport. The effect of such oscillations for the most frequently targeted sites has been addressed in analytical studies [43] and in computer simulations [42]. Very recently a comprehensive study of this effect has been provided by Mendels and Tessler [44], who suggested the analytical and numerical methods how to exclude the effect of frequent oscillations. The obtained distribution of targeted energies have not been related though to long-range charge transport [44]. In section 4 we address the issue of frequent carrier oscillations in detail, also showing that the algorithm suggested by Cleve *et al* is not suitable to find the position of the TE.

In section 5 we suggest a novel numerical algorithm as an alternative approach, which allows one to determine the position of the TE responsible for the long-range VRH transport in the GDM. The algorithm is checked by application to the exactly solvable model of the NNH on a lattice. Straightforward computer simulations of the VRH in the GDM carried out in the framework of this algorithm reveal the analytical estimate of the TE by Oelerich *et al* [22] as valid. In that analytical estimate one assumes that the TE corresponds to the target energy which maximizes the rate of carrier transitions upward in energy accompanied by tunneling over a distance typical for sites with energies below TE. This approach is outlined in section 3.

3. Analytic estimate of the transport energy

Let us briefly repeat the important steps in the analytical derivation of the TE following [22, 24, 25, 31, 45]. The obtained value for the TE will be compared with numerical results in section 5.

Consider a carrier in a state with energy ε_i. According to equation (2), the typical rate of a downward hop of such a carrier to a localized state deeper in energy is [22]

$$\nu_\downarrow = \nu_0 \exp\left[-\frac{2r(\varepsilon_i)}{\alpha}\right], \tag{9}$$

where

$$r(\varepsilon) = \left[\frac{4\pi}{3}\int_{-\infty}^{\varepsilon} g(\varepsilon')\left[1 - f(\varepsilon', \varepsilon_F)\right]d\varepsilon'\right]^{-1/3}. \tag{10}$$

The typical rate of an upward hop of such a carrier to a localized state with energy $\varepsilon_x > \varepsilon_i$ is

$$\nu_\uparrow(\varepsilon_i, \varepsilon_x) = \nu_0 \exp\left[-\frac{2r(\varepsilon_x)}{\alpha} - \frac{\delta}{kT}\right], \tag{11}$$

where $\delta = \varepsilon_x - \varepsilon_i > 0$. Note, that this expression is not exact: The considered distance $r(\varepsilon_x)$ is based on all empty states with energies deeper than ε_x. For the Gaussian DOS this is equivalent to considering a slice of energy with the width of the order σ. This works for a DOS that varies slowly compared to kT, but not in general. It is also assumed for simplicity that the localization length α does not depend on energy. The latter

assumption can be released easily on the cost of a somewhat more complicated form of the equations. We will analyze the hopping rates at a given temperature T and try to find out the energy level ε_x, which provides the fastest typical hopping rate for a charge carrier initially placed at a site with the energy ε_i. The corresponding energy ε_x is determined by the condition

$$\frac{\partial \nu_\uparrow(\varepsilon_i, \varepsilon_x)}{\partial \varepsilon_x} = 0. \tag{12}$$

It is easy to show [24, 25] that the target energy ε_x determined by equation (12) does not depend on the initial energy ε_i and hence is universal for given parameters $N\alpha^3$ and σ/kT and can be considered as the transport energy ε_t. It means that from states below ε_t carrier transitions essential for charge transport occur to the vicinity of ε_t, while transitions from states with energies higher than ε_t occur downwards in energy. This is the essence of the TE concept. After relaxation towards ε_t the behavior of carriers resembles the MT mode [15] with ε_t playing the role of the mobility edge [23–25].

Solving equation (12) for ε_x, the TE ε_t is then determined by [22]

$$\frac{2kT}{3\alpha} \left(\frac{4\pi}{3B_c} \right)^{-1/3} \left[\int_{\varepsilon_t}^{\infty} g(\varepsilon) \left[1 - f(\varepsilon, \varepsilon_F) \right] d\varepsilon \right]^{-4/3} \tag{13}$$
$$\times [1 - f(\varepsilon_t, \varepsilon_F)] \, g(\varepsilon_t) = 1.$$

The factor B_c in this expression accounts for the percolation nature of the hopping transport, i.e., for the necessity to provide a continuous path via sites separated by appropriate distances. Accurate estimates give [46] the value $B_c = 2.735$.

4. On the most frequently visited energy

In this Section we highlight why computer algorithms based on tracing the energies of most frequently visited sites are not suitable to determine the position of the TE. In such algorithms one counts the number of carrier transitions, which, in course of a standard Monte Carlo simulation, bring the carrier into a particular energy interval. The statistics of such events is then plotted as a histogram with number of hops per energy interval. The peak of this histogram, i.e., the *most frequently visited energy* is interpreted as pointing at the position of the TE [40–42].

Although it has been doubted that such an algorithm is accurate enough to determine the position of the TE [37, 42], the most frequently visited energy is often used not only to determine the TE, but even for conclusions on the existence or non-existence of the TE [41]. Therefore we find it instructive to analyze the approach based on the most frequently visited energy in more detail. Below, we derive an analytic expression for the most frequently visited energy and validate it via comparison with computer simulations. We will focus on the Gaussian DOS, although the treatment of the exponential DOS is very similar. The derivation shows which hops determine the energy in question and that these hops *cannot* be the ones decisive for mobility or other long-range transport coefficients.

We will determine the frequency of hops $f(\varepsilon_j)$ d ε_j with the target energy within the interval $[\varepsilon_j, \varepsilon_j + d \varepsilon_j]$. Let us assume that the electron concentration is low and that only a weak external

electric field is applied. The occupation probability for any site with energy ε is then given by a Boltzmann distribution

$$P_B(\varepsilon) \approx e^{\varepsilon_F/kT} e^{-\varepsilon/kT} \equiv \mathcal{N} e^{-\varepsilon/kT}, \tag{14}$$

where the normalization constant $\mathcal{N}$ is determined by the concentration of electrons. The expected frequency of hops to a particular site with the energy ε_j is determined by integrating the hopping rate $\nu_{ij} = \nu(\varepsilon_i, \varepsilon_j, r)$ (see equation (2)) over all possible initial energies ε_i and hopping distances r. The contributions from sites with the energy ε_i and distance r are weighted by the expected number of such sites $g(\varepsilon_i)$ and by the probability $P_B(\varepsilon_i)$ that the initial site is occupied. The resulting frequency of hops $f(\varepsilon_j)$ to one site with energy ε_j, weighted by the concentration of such sites $g(\varepsilon_j)$, gives the frequency of hops to the energy ε_j.

It is convenient to evaluate the hop frequency separately for hops upward and downward in energy, since the form of the hopping rates equation (2) is different for upward and downward hops. The frequency of upward hops to the energy interval $[\varepsilon_j, \varepsilon_j + d\varepsilon_j]$ per unit system volume and energy, is

$$f_{up}(\varepsilon_j) \, d\varepsilon_j = g(\varepsilon_j) \, d\varepsilon_j \times \int_{-\infty}^{\varepsilon_j} d\varepsilon_i \int_0^{\infty} d r \, 4\pi r^2 g(\varepsilon_i) \, P_B(\varepsilon_i) \, \nu(\varepsilon_i, \varepsilon_j, r). \tag{15}$$

The integral over r can be evaluated as

$$\int_0^{\infty} d r \, 4\pi r^2 \, e^{-2r/\alpha} = \pi\alpha^3. \tag{16}$$

By inserting the Gaussian DOS $g(\varepsilon)$ from equation (1), we obtain

$$f_{up}(\varepsilon_j) = \frac{\pi \nu_0 \mathcal{N} N\alpha^3}{2} \times g(\varepsilon_j) e^{-\varepsilon_j/kT} \left[1 + \text{erf}\left(\frac{\varepsilon_j}{\sqrt{2}\,\sigma} \right) \right]. \tag{17}$$

Similarly, by adjusting the limits of the energy integral in equation (15) and using the right form of the hopping rates $\nu(\varepsilon_i, \varepsilon_j, r)$ (see equation (2)), the frequency of downward hops arriving at ε_j can be evaluated as

$$f_{dn}(\varepsilon_j) = \frac{\pi \nu_0 \mathcal{N} N\alpha^3}{2} \times g(\varepsilon_j) e^{\sigma^2/2(kT)^2} \left[1 - \text{erf}\left(\frac{\varepsilon_j + \sigma^2/kT}{\sqrt{2}\,\sigma} \right) \right]. \tag{18}$$

The total frequency of hops to the energy ε_j is obtained by adding the frequencies of upward and downward hops

$$f(\varepsilon_j) = f_{up}(\varepsilon_j) + f_{dn}(\varepsilon_j). \tag{19}$$

The energy ε_{peak}, which maximizes $f(\varepsilon_j)$, is our desired quantity, the most frequently visited energy. It has previously been calculated in Monte Carlo simulations [37, 42]. In addition to the peak energy of $f(\varepsilon_j)$, we also calculate its average energy,

$$\langle \varepsilon_j \rangle = \int_{-\infty}^{\infty} \varepsilon_j f(\varepsilon_j) \, d\varepsilon_j \; / \int_{-\infty}^{\infty} f(\varepsilon_j) \, d\varepsilon_j. \tag{20}$$

Both $\langle \varepsilon_j \rangle$ and ε_{peak} are determined numerically.

The two energies ε_{peak} and $\langle \varepsilon_j \rangle$ turn out to be very close to each other, since the distribution of $f(\varepsilon_j)$ is almost symmetric. This will become clear from the results in figure 3. We prefer $\langle \varepsilon_j \rangle$ as compared to ε_{peak}, since the average is easier to determine in a simulation. Furthermore, a numeric evaluation of equation (20) shows that

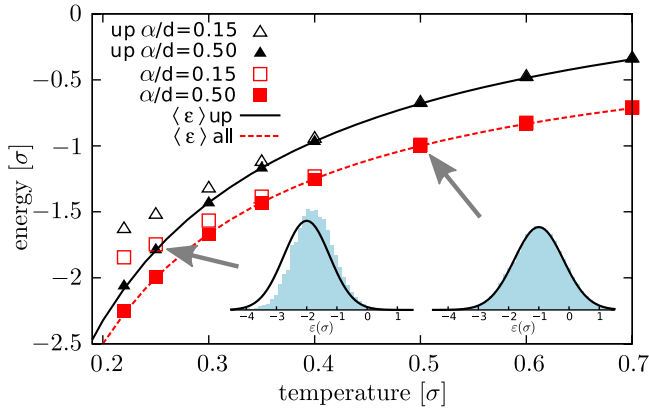**Figure 3.** The average visited energy (squares) as a function of temperature, obtained by Monte Carlo simulation in a system with 10^6 randomly placed sites at zero electric field. The triangles show the average energy of upward hops. The curves show the predicted average of the visited energies equation (20) and the corresponding prediction for upward hops. The insets show histograms of the number of visits as a function of energy, for $\alpha = 0.15\, d$ at two different temperatures (arrows). The curves show the function $f(\varepsilon_j)$, given by equation (19).

$$\langle \varepsilon_j \rangle \approx \frac{-\sigma^2}{2kT}, \tag{21}$$

to an accuracy of 10 digits. Note that $f(\varepsilon_j)$ depends on the localization length α only via the prefactor α^3. This means that both $\varepsilon_{\text{peak}}$ and $\langle \varepsilon_j \rangle$ are independent of α, which is a first hint to why this approach is insufficient for transport energy calculations.

The integral in equation (16) reveals the hops contributing mostly to the frequency of visits. The integrand has a maximum at $r = \alpha$, which means that hops of this length are the most important. Thus the most frequently visited energy is determined by pairs of spatially close sites, separated by the localization length α which is much smaller than the typical site separation. Such pairs are not important for the charge carrier mobility, [8, 47] since a charge carrier cannot move any significant distance using only such pairs.

In order to verify the results of the calculation above, we perform a Monte Carlo computer simulation of VRH and calculate the average of the visited energies. The algorithm is similar to that of Hartenstein and Bässler, [37] except that we consider sites randomly placed in space rather than distributed on a lattice. The average of the visited energies is shown in figure 3. The average over all upward hops is also shown since the transport energy is usually derived taking into account only upward hops in the hopping-rate optimization approach (see section 3). The insets in figure 3 show histograms of the number of visits to different energy intervals and a comparison with the function $f(\varepsilon_j)$. Good agreement is seen between the simulation results and equation (20), except for the small localization length at low temperatures. The latter is due to the finite-size effect, which comes in play at small localization lengths. In order to warrant representative results, sites in the low-energy tail of the DOS must not only be present in the system, but be present in pairs separated by the distance of

the order of α. This is apparently not the case in the left-hand histogram, where the number of visits in the low-energy part of the histogram is lower than predicted.

The data was obtained for a single electron hopping in an otherwise empty system. The electron was started from a random position and allowed to perform 10^8 hops; this procedure was repeated 1000 times. No electric field was applied. The simulation was repeated for five realizations of the disorder and the results were averaged. The hops in the simulation were restricted to $r_{\text{cut}} = 3d$, where $d = N^{-1/3}$ gives the typical site separation. Increasing the cut-off length r_{cut} did not alter the results.

In summary, the *most frequently visited energy* in the GDM can be determined analytically. The result agrees with those of a Monte Carlo simulation, yielding that the most frequent hops lead to an energy in the vicinity of $-\sigma^2/2kT$. However, these hops are much shorter than the typical distance between sites responsible for the charge carrier mobility. Therefore, the most frequently visited energy cannot be considered as relevant for charge transport and cannot serve as a transport energy.

Very recently, Mendels and Tessler [44] have attempted to overcome the problem by removing oscillatory hops between adjacent sites *before* the statistical analysis. Although the resulting distribution peaks are significantly shifted towards higher energies as compared to statistics including oscillations, it is not sufficient to eliminate cycles between only two sites in order to mimic the long-range hopping transport. In fact, the approach of tracing the most frequently visited energy becomes useful only when *all* closed cycles in the paths of charge carriers are removed before the statistical analysis.

However, removing all closed cycles from a directed graph is a very computationally demanding task, especially considering the large amount of data that is needed for a sufficiently accurate numerical study of the long-range hopping transport.

In the following Section, we suggest another approach which allows one to determine the transport energy using computer simulations. In that approach, the TE will be deduced from the results of carrier mobility and is therefore directly related to the long-range charge transport. It will be shown, that the transport energy is situated much higher in the DOS than $-\sigma^2/2kT$ given by equation (20).

5. How to find the transport energy in numerical simulations

We suggest to determine the TE by studying the effect of a DOS modifications on the charge carrier mobility. Thereby, it is ensured that the determined energy range is decisive for long-range transport. The idea is to cut out sites with energies in an interval $[\varepsilon_{\text{cut}} - w, \varepsilon_{\text{cut}}]$ from the system and to check, whether and how the resulting mobility, determined in straightforward computer simulations, is affected by such a modification of the DOS.

In figure 4, the corresponding modification is depicted. Sites in the interval $[\varepsilon_{\text{cut}} - w, \varepsilon_{\text{cut}}]$ are cut out from the DOS and are hence unaccessible for carriers during charge transport. Charge carriers must avoid the sites with energies in this
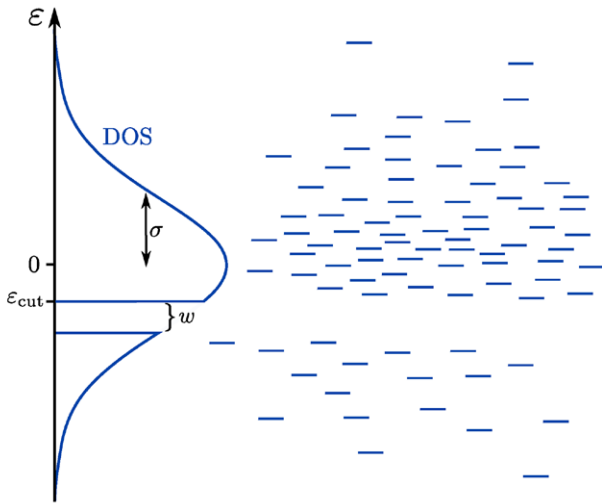
Figure 4. Schematic picture of the modified DOS $g(\varepsilon)$. Sites with energies in the interval $[\varepsilon_{cut}-w, \varepsilon_{cut}]$ are removed from the system and therefore not available for charge transport.

region and find other transport paths. Note, that the DOS is *not* renormalized after the modification. Depending on how important the withdrawn sites are for charge transport, the cutting will affect the mobility of the system more or less severely. We expect the highest drop in the carrier mobility when the cut-out sites have been those of highest importance for the long-range transport.

In order to justify the suggested method, let us check it by application to the exactly solvable hopping model, namely, the nearest-neighbor hopping (NNH) on a lattice. We assume a system of localized states distributed in an fcc or sc lattice structure. The site energies follow a Gaussian DOS. Charge carriers with finite concentration move via nearest neighbor transitions with Miller–Abrahams hopping rates given by equation (2). Percolation theory provides an exact analytical expression for the carrier mobility in this system [48]:

$$\mu(\varepsilon_{crit}, \varepsilon_F) = \mu_0 [\, g(\varepsilon_{crit}) \,]^\lambda \exp\!\left(\frac{\varepsilon_F - \varepsilon_{crit}}{kT} \right), \qquad (22)$$

where μ_0 and λ do not depend on the DOS. The critical energy ε_{crit} is obtained from the DOS via equation (3). Since equation (22) can be applied to any DOS function, it provides a perfect tool to test the approach of searching for the TE by DOS modifications.

In order to perform such a test, let us cut an interval $[\varepsilon_{cut}-w, \varepsilon_{cut}]$ from the Gaussian DOS and check how it affects the carrier mobility. The modified DOS vanishes in the chosen energy interval:

$$g^*(\varepsilon) = \begin{cases} 0 & \text{if } \varepsilon \in [\varepsilon_{cut} - w, \varepsilon_{cut}], \\ g(\varepsilon) & \text{otherwise.} \end{cases} \qquad (23)$$

Both ε_{crit} and ε_F in equation (22) depend on the density of states and therefore need to be recalculated. Since the percolation threshold x_c and the carrier concentration n are independent of the shape of the DOS, we can use the following equations to determine the values ε_{crit}^* and ε_F^* for the modified
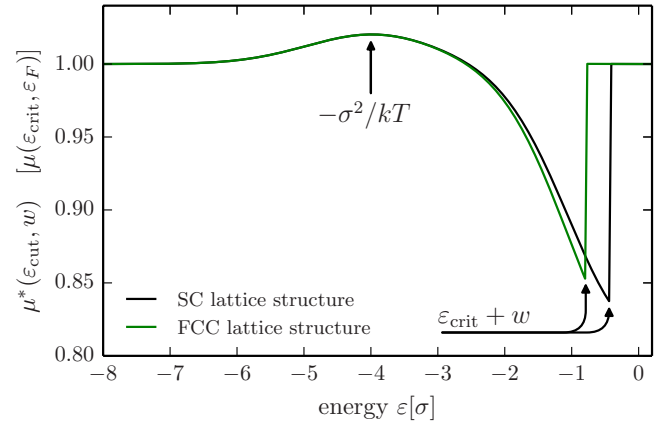

Figure 5. The mobility value of the modified DOS relative to the unmodified one as a function of the upper cutout interval boundary. Temperature $kT = 0.25\sigma$ and the cutout interval width $w = 0.05\sigma$. Evaluation is stopped when $\varepsilon_{crit} < \varepsilon_{cut} - w$, hence the discontinuity in the curves.

DOS from the values ε_{crit} and ε_F valid for the unmodified Gaussian DOS:

$$x_c = \int_{-\infty}^{\varepsilon_{crit}} g(\varepsilon)\,d\varepsilon = \int_{-\infty}^{\varepsilon_{cut}-w} g(\varepsilon)\,d\varepsilon + \int_{\varepsilon_{cut}}^{\varepsilon_{crit}^*} g(\varepsilon)\,d\varepsilon \qquad (24)$$

$$n = \int_{-\infty}^{\infty} g(\varepsilon) f(\varepsilon, \varepsilon_F)\,d\varepsilon \qquad (25)$$

$$= \int_{-\infty}^{\varepsilon_{cut}-w} g(\varepsilon) f(\varepsilon, \varepsilon_F^*)\,d\varepsilon + \int_{\varepsilon_{cut}}^{\infty} g(\varepsilon) f(\varepsilon, \varepsilon_F^*)\,d\varepsilon$$

In the above equations, the cut-out interval $[\varepsilon_{cut}-w, \varepsilon_{cut}]$ was simply removed from the integration ranges. Inserting the values for ε_{crit}^* and ε_F^* into equation (22) one obtains the mobility dependent on the cutout interval:

$$\mu^*(\varepsilon_{cut}, w) = \mu(\varepsilon_{crit}^*, \varepsilon_F^*). \qquad (26)$$

By comparing this result to that for the unmodified DOS one can study the effect of the DOS modification on the carrier mobility.

In figure 5, the ratio $\mu^*(\varepsilon_{cut}, w)/\mu(\varepsilon_{crit}, \varepsilon_F)$ is plotted as a function of the upper boundary of the cutout interval, ε_{cut}. The interval width w is fixed at $w = 0.05\sigma$ and results for two different lattice structures, *fcc* and *sc* are depicted.

The figure clearly demonstrates that the mobility drops significantly when cutting out sites in the vicinity of the critical energy ε_{crit}. This is the expected result.

There is another interesting peculiarity demonstrated by the curves in figure 5: when the cutout interval overlaps with the energy $\varepsilon_\infty = -\sigma^2/kT$ the carrier mobility increases. This happens for the following reason. In the Gaussian DOS and in thermal equilibrium most carriers occupy sites around the so-called equilibration energy ε_∞. Removing sites around this energy pushes carriers to higher energies and hence decreases the activation energy necessary for activation to the transport path, which leads concomitantly to the increase of the carrier mobility.
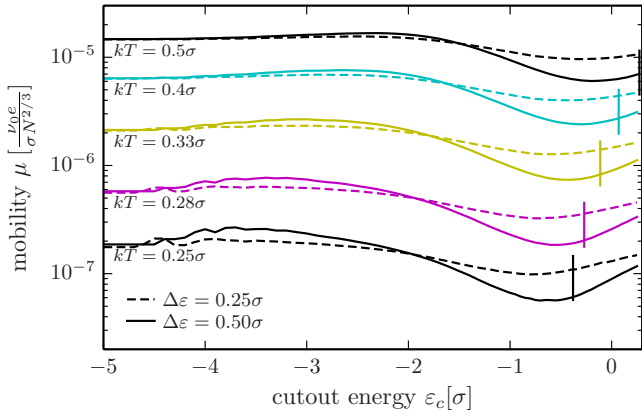
Figure 6. The mobility in a system with absent sites in the interval $[\varepsilon_{cut}-w, \varepsilon_{cut}]$. The vertical lines next to the minimum of the mobility are the analytical results from equation (13).

The above results, obtained by the exactly solvable model of nearest-neighbor hopping on a lattice, demonstrate the validity of the suggested approach for revealing the position of the TE: To cut out sites with particular energies and check the effect of such a modification on the calculated carrier mobility. Below we apply this approach to the problem relevant to the charge transport in real systems, namely to the VRH in the Gaussian DOS.

In order to study the effect of the DOS modification for the case of the VRH, we calculate the charge carrier mobilities numerically using the balance equation (BE) approach. [49, 50] By varying the width w and the upper boundary ε_{cut} of the cutout interval (see figure 4), one can find at which w and ε_{cut} the maximal effect in the decrease of the carrier mobility is achieved, thereby revealing the energy range most important for charge transport. The results are presented in figure 6. Results for two different widths, $w = 0.25\sigma$ and $w = 0.5\sigma$ at different temperatures kT are compared. In all cases, motion of a single charge carrier in an otherwise empty system was simulated. The localization length was chosen as $\alpha = 0.215 \, N^{-1/3}$ and the number of sites in the simulated system was equal to 90^3.

It is clearly visible in the figure, that a significant decrease in the resulting mobility appears in each of the curves for a certain energy range of withdrawn sites. The effect is larger for larger energy intervals w, yet more accurate for lower values of w, since in that case the system is affected less severely. The choice of the interval width w should therefore be a compromise between the visibility of the effect and the accuracy in determining the position of the most efficient ε_{cut}. We interpret the minima on the curves plotted in figure 6 as pointing at the position of the real transport energy (TE) responsible for the long-range charge transport. The position of the TE, as determined from figure 6, shifts upwards in energy with rising temperature. This agrees with all previous analytical and numerical studies. The vertical lines close to the minima of the mobility curves in figure 6 show the positions of the TE calculated from equation (13). Apparently, the optimization approach leading to equation (13) (outlined in section 3) is supported by the results of the straightforward computer simulations of the carrier mobility presented in figure 6.

Figure 6 also shows, that when sites in the vicinity of the average carrier energy $\varepsilon_{\infty} = -\sigma^2/kT$ are cut out, the mobility is

slightly increased. This is literally the effect discussed above with respect to the exactly solvable model of the nearest-neighbor hopping on a lattice; It is clearly visible in figure 5. When sites with energies in the vicinity of the equilibration level $\varepsilon_{\infty} = -\sigma^2/kT$ are removed from the system, the average carrier energy increases, which diminishes the activation energy to the TE and leads to an increase of the mobility.

For the given temperature range the minima of the $\mu(\varepsilon_{cut})$ curves in figure 6, which we interpret as the real *transport energy* related to long-range charge transport, are situated close to the center of the DOS. It seems therefore correct to assume in a simplified approach that the TE coincides with the center of the Gaussian DOS, as was suggested before [1, 51]. It means that although charge carriers spend most time in states with energies around the equilibration energy $\varepsilon_{\infty} = -\sigma^2/kT$, transport takes place via sites with much higher energies than ε_{∞}.

The data in figure 6 shows another remarkable feature, namely that the removal of sites with energies in the vicinity of the energy $-\sigma^2/2kT$ does not affect the carrier mobility significantly. This shows that it is wrong to interpret the VRH charge transport in the Gaussian DOS as activation of carriers from some particular energy level ε_{start} to the transport energy ε_t. If the latter interpretation was correct, the observation of the mobility temperature dependence in the form of equation (8) with $C^2 \approx 1/2$ would yield the value $\varepsilon_{start} \approx -\sigma^2/2kT$ for $\varepsilon_t \approx 0$. The fact that cutting sites with energies in the vicinity of the level $-\sigma^2/2kT$ does not affect the mobility means that this energy level does not play any significant role for charge transport. The temperature dependence of carrier mobility in the form of equation (8) is instead the result of the time averaging over the upward hops to the TE from deeper energy levels [2, 31].

6. Energy level responsible for non-equilibrium energy relaxation

While it is well established that the percolation nature of the VRH plays a decisive role for long-range conductivity [8], it is not clear whether the percolation features have anything to do with the description of the energy relaxation of charge carriers towards the equilibrium distribution. It is worth noting that the concept of the TE was initially invented to study the relaxation behavior of charge carriers, showing particularly that these dynamics can be described in the framework of the MT model by replacing the mobility edge by the TE [23]. The MT model, in its turn, was initially developed [16–18] also in order to describe the energy relaxation of the non-equilibrium carrier distribution in systems with a mobility edge. During the relaxation phase, before thermal equilibrium is achieved, the carrier mobility depends on time, which is the essence of the so-called *dispersive transport*. In the GDM, dispersive transport takes place before charge carriers have relaxed in energy either to the equilibration energy $\varepsilon_{\infty} = -\sigma^2/kT$, if $\varepsilon_{\infty} > \varepsilon_F$, or to the Fermi energy, if $\varepsilon_{\infty} < \varepsilon_F$.

Let us consider the case of small carrier concentrations n so that $\varepsilon_{\infty} > \varepsilon_F$. Baranovskii *et al* [31] estimated the time τ_{rel} necessary for carriers to achieve thermal equilibrium around

$\varepsilon_\infty = -\sigma^2/kT$ in the Gaussian DOS: They calculated τ_{rel} as the time needed for activation from ε_∞ towards the level ε_t^{rel}, calculated without taking percolation arguments into account [25, 31]. This led to a temperature dependence of τ_{rel} in the form [31]

$$\tau_{rel} \propto \exp\left[\left(B\frac{\sigma}{kT}\right)^2\right], \qquad (27)$$

with $B \approx 1$ resolving the puzzle about the difference between the values of the coefficients $C^2 \simeq 1/2$ and $B \simeq 1$ in equations (8) and (27), respectively, raised by the results of computer simulations [1, 12, 52, 53]. The quantity ε_t^{rel} was calculated [25, 31] via equation (13) with $B_c = 1$. Apparently, the level ε_t^{rel} for the energy relaxation of carriers does not coincide with the TE ε_t responsible for the long-range transport calculated with $B_c \simeq 2.7$.

A similar result has recently been obtained by Germs *et al* [54], who performed Monte Carlo (MC) simulations of the carrier energy relaxation in the GDM and found the position of the energy level E_c^{MC} that would describe numerical results in the framework of the MT model with replacing the mobility edge by E_c^{MC}. It was found that E_c^{MC} depends on the ratio σ/kT, i.e., depends on temperature T for a given value σ. Although it is not specified explicitly, one can guess that Germs *et al* [54] studied the nearest-neighbor hopping (NNH) process on a lattice. Otherwise the localization length α, not at all specified in the paper by Germs *et al* [54], would have come in play as an additional decisive parameter. On the other hand, equation (3), that determines the position of the TE E_c responsible for the long-range transport in the NNH on a lattice, depends only on the energy scale σ of the DOS and not on temperature T. This clearly illustrates the difference between E_c^{MC} and E_c. It would be desirable to extend the study from the oversimplified model of the NNH on a lattice, that yields the transport energy E_c dependent on the choice of the lattice structure [13], to a more realistic problem of the VRH in the GDM.

Assuming carrier relaxation in the VRH, the interesting question is whether or not to include the percolation nature of transport in such considerations, e.g., when calculating the transport energy ε_t to replace the mobility edge in the MT model. If the average spatial displacement of carriers during relaxation, i.e., before they reach the equilibration energy ε_∞, exceeds the correlation length $\mathcal{L}$ of the percolating cluster [8, 14], percolation criteria are expected to be necessary for the description of the relaxation dynamics; in case the displacement is smaller than $\mathcal{L}$, percolation should probably not play a significant role in the relaxation properties. This topic exceeds the scope of the current report and remains an open question.

7. Conclusions

A new numerical approach is suggested to determine the transport path of hopping charge carriers in a steeply decreasing DOS. The approach is based on simulating the effect of a DOS modification on charge transport properties. The method is applied to the Gaussian DOS relevant for organic disordered semiconductors. Our numerical calculations confirm the

validity of the analytical approach to calculate the transport energy via optimization of hopping rates [22].

It is also rigorously proven by analytical calculations and numerical simulations that algorithms based on tracing the energies of the most frequently targeted localized states, although widely applied, are not suitable to determine the position of the transport energy.

Acknowledgments

References

[1] Bässler H 1993 *Phys. Status Solidi* B **175** 15
[2] Oelerich J O, Huemmer D and Baranovskii S D 2012 *Phys. Rev. Lett.* **108** 226403
[3] Baranovskii S D 2014 *Phys. Status Solidi* B **251** 487–25
[4] Bässler H 1981 *Phys. Status Solidi* B **107** 9
[5] Miller A and Abrahams E 1960 *Phys. Rev.* **120** 745–55
[6] Gill W D 1972 *J. Appl. Phys.* **43** 5033
[7] Rubel O, Baranovskii S D, Thomas P and Yamasaki S 2004 *Phys. Rev.* B **69** 014206
[8] Shklovskii B I and Efros A L 1984 *Electronic Properties of Doped Semiconductors* (Heidelberg: Springer)
[9] Shklovskii B and Efros A 1971 *Sov. Phys.—JETP* **33** 468
[10] Ambegaokar V, Halperin B and Langer J 1971 *Phys. Rev.* B **4** 2612
[11] Pollak M 1972 *J. Non-Cryst. Solids* **11** 1–24
[12] Schönherr G, Bässler H and Silver M 1981 *Phil. Mag.* B **44** 369
[13] Cottaar J, Koster L J A, Coehoorn R and Bobbert P A 2011 *Phys. Rev. Lett.* **107** 136601
[14] Nenashev A V, Jansson F, Oelerich J O, Huemmer D, Dvurechenskii A V, Gebhard F and Baranovskii S D 2013 *Phys. Rev.* B **87** 235204
[15] Orenstein J and Kastner M A 1981 *Solid State Commun.* **40** 85–9
[16] Noolandi J 1977 *Phys. Rev.* B **16** 4466
[17] Schmidlin F W 1977 *Phys. Rev.* B **16** 2362
[18] Silver M and Cohen L 1977 *Phys. Rev.* B **15** 3276
[19] Rudenko A I and Arkhipov V I 1982 *Phil. Mag.* B **45** 177
[20] Arkhipov V I and Rudenko A I 1982 *Phil. Mag.* B **45** 189
[21] Nenashev A V, Jansson F, Baranovskii S D, Österbacka R, Dvurechenskii A V and Gebhard F 2009 *Phys. Rev.* B **81** 115204
[22] Oelerich J O, Huemmer D, Weseloh M and Baranovskii S D 2010 *Appl. Phys. Lett.* **97** 143302
[23] Monroe D 1985 *Phys. Rev. Lett.* **54** 146–9
[24] Baranovskii S D, Thomas P and Adriaenssens G J 1995 *J. Non-Cryst. Solids* **190** 283–87
[25] Baranovskii S D, Faber T, Hensel F and Thomas P 1997 *J. Phys. Condens. Matter* **9** 2699–706
[26] Grünewald M and Thomas P 1979 *Phys. Status Solidi* B **125** 125–33
[27] Vissenberg M C J M and Matters M 1998 *Phys. Rev.* B **57** 12964
[28] Baranovskii S D, Zvyagin I P, Cordes H, Yamasaki S and Thomas P 2002 *J. Non-Cryst. Solids* **299** 416
[29] Zvyagin I P 2008 *Phys. Status Solidi* C **5** 725

[30] Baranovskii S D, Zvyagin I P and Cordes H 2002 *Phys. Status Solidi* B **287** 281–7

[31] Baranovskii S D, Cordes H, Hensel F and Leising G 2000 *Phys. Rev.* B **62** 7934

[32] Arkhipov V I, Emelianova E V and Adriaenssens G J 2001 *Phys. Rev.* B **64** 125125

[33] Nikitenko V R, von Seggern H and Bässler H 2007 *J. Phys.: Condens. Matter* **19** 136210

[34] Arkhipov V I, Heremans P, Emelianova E V, Adriaenssens G J and Bässler H 2003 *Appl. Phys. Lett.* **82** 3245

[35] Schmechel R 2002 *Phys. Rev.* B **66** 235206

[36] Martens H C F, Hulea I N, Romijn I, Brom H B, Pasveer W F and Michels M A J 2003 *Phys. Rev.* B **67** 121203

[37] Hartenstein B and Bässler H 1995 *J. Non-Cryst. Solids* **190** 112–6

[38] Tanase C 2005 *PhD Thesis* University of Groningen

[39] Fumagalli L, Binda M, Natali D, Sampietro M, Salmoiraghi E, Gianvincenzo P D and Di Gianvincenzo P 2008 *J. Appl. Phys.* **104** 084513

[40] Cleve B, Hartenstein B, Baranovskii S D, Scheidler M, Thomas P and Bässler H 1995 *Phys. Rev.* B **51** 16705

[41] Novikov S V and Malliaras G G 2006 *Phys. Status Solidi* B **243** 387

[42] Gonzalez-Vazquez J P, Anta J A and Bisquert J 2009 *Phys. Chem. Chem. Phys.* **11** 10359–67

[43] Arkhipov V I, Emelianova E V and Adriaenssens G J 2001 *Phys. Rev.* B **64** 125125

[44] Mendels D and Tessler N 2013 *J. Phys. Chem.* C **117** 24740

[45] Baranovskii S D (ed) 2006 *Charge Transport in Disordered Solids with Applications in Electronics* (Chichester: Wiley)

[46] Lorenz C D and Ziff R M 2001 *J. Chem. Phys.* **114** 3659

[47] Baranovskii S D and Thomas P 1996 *J. Non-Cryst. Solids* **198** 140–45

[48] Cottaar J, Coehoorn R and Bobbert P A 2012 *Phys. Rev.* B **85** 245205

[49] Yu Z G, Smith D L, Saxena A, Martin R L and Bishop A R 2000 *Phys. Rev. Lett.* **84** 721–24

[50] Cottaar J and Bobbert P A 2006 *Phys. Rev.* B **74** 115204

[51] Neumann F, Genenko Y A and von Seggern H 2006 *J. Appl. Phys.* **99** 013704

[52] Ries B, Bässler H, Grünewald M and Movaghar B 1988 *Phys. Rev.* B **37** 5508

[53] Bässler H 1990 *Transport and Relaxation of Excitations in Random Organic Solids Advances in Disordered Semiconductors* (Singapore: World Scientific) 491–520 (Monte Carlo Simulation and Experiment)

[54] Germs W C, van der Holst J J M, van Mensfoort S L M, Bobbert P A and Coehoorn R 2011 *Phys. Rev.* B **84** 165210

Article IV

A. Hofacker, **J. O. Oelerich**, A. V. Nenashev, F. Gebhard, and S. D. Baranovskii
Theory to carrier recombination in organic disordered semiconductors
J. Appl. Phys. **115**, 223713 (2014)

Theory to carrier recombination in organic disordered semiconductors

A. Hofacker,[1] J. O. Oelerich,[1,a)] A. V. Nenashev,[2,3] F. Gebhard,[1] and S. D. Baranovskii[1]
[1]*Department of Physics and Material Sciences Center, Philipps-University, D-35032 Marburg, Germany*
[2]*Institute of Semiconductor Physics, 630090 Novosibirsk, Russia*
[3]*Department of Physics, Novosibirsk State University, 630090 Novosibirsk, Russia*

A theoretical description for recombination kinetics of charge carriers in a disordered system with a broad energy distribution of localized states (DOS) is suggested. This kinetics is governed by the exchange of carriers between transport states and traps. Concentration transients in systems with Gaussian DOS, typical for organic semiconductors, appear much steeper than those obtained for systems with exponential DOS. This difference in recombination kinetics is caused by the difference in thermalization kinetics for these two types of the DOS functions. The comparison of the recombination transients for mobile and trapped carriers in exponential and Gaussian DOS might help to distinguish between these two possible shapes of the DOS using experimental data for transient photoconductivity and photoabsorption. © *2014 AIP Publishing LLC*.
[http://dx.doi.org/10.1063/1.4883318]

I. INTRODUCTION

The problem of charge carrier recombination belongs to the central topics in research on semiconductors, particularly organic disordered semiconductors. Recombination efficiency determines the performance of all optoelectronic devices, and therefore, its understanding is of vital importance for device applications of semiconducting materials. Recombination processes in disordered semiconductors demonstrate several peculiar features, which do not appear in ordered crystalline materials and make the appropriate theoretical description of such processes a challenging issue.

The absence of any crystalline structure in disordered materials leads to an energetically broad distribution of localized states (DOS) that can act as traps for charge carriers if the energy scale of the DOS is larger than the thermal energy kT. In inorganic amorphous semiconductors, such as amorphous silicon and chalcogenide glasses, localized states form the so-called *band tails*, which are separated from the extended states by the mobility edge. At not extremely low temperatures, charge transport takes place via motion of carriers in the extended states, occasionally interrupted by trapping into the localized states with a subsequent release back into the transport states.[1,2] This so-called *multiple trapping* (MT) transport model is illustrated in Fig. 1. The MT model has been suggested in order to describe the time-dependent carrier mobility in the so-called *dispersive transport* regime in inorganic disordered materials, such as chalcogenide glasses and amorphous silicon.[3–5]

The DOS in inorganic amorphous materials is believed to have an exponential shape[1–7]

$$g(\varepsilon) = \frac{N_t}{\varepsilon_0} \exp\left(\frac{\varepsilon}{\varepsilon_0}\right), \quad (1)$$

where N_t is the total concentration of localized states and ε_0 is the energy scale of the DOS. The energy of a localized state ε is counted negative into the band gap from the mobility edge as a reference level ($\varepsilon = 0$).

Theoretical description of charge transport and recombination in organic amorphous semiconductors is not yet as well developed as that in inorganic disordered semiconductors. Some researchers suggest that a band structure with mobility edges separating extended and localized states is valid also for organic materials,[8–10] in which case mechanisms for charge transport and recombination should be very similar to those in inorganic disordered materials. Other researchers suggest that all electronic states in organic disordered systems are localized, and charge carriers move via incoherent tunneling (hopping) between spatially and energetically distributed localized states.[11] In any case, the energy spectrum in disordered organic materials has a Gaussian shape[11–14]

$$g(\varepsilon) = \frac{N_t}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right), \quad (2)$$

where σ denotes the energy scale of the DOS. It has been well established[14–18] that also in the case of hopping transport in a system with a steeply energy-dependent DOS, like
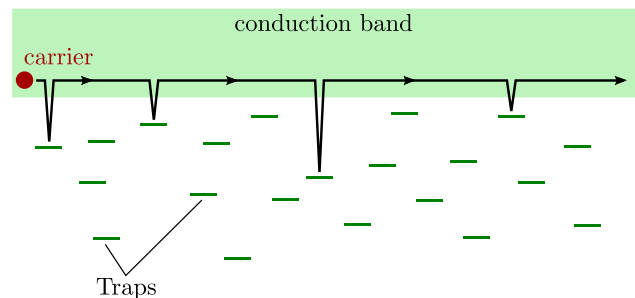

FIG. 1. Schematic view of the MT transport mode. Carrier motion in the conduction band is interrupted by trapping into and subsequent release from localized tail states.

a)Electronic mail: jan.oliver.oelerich@physik.uni-marburg.de

a Gaussian one, motion of charge carriers can be described in the framework of the MT formalism by replacing the mobility edge in the MT model by some particular energy level ε_t, called the *transport energy*. Furthermore, it has been shown that ε_t for the Gaussian DOS is situated close to the maximum of the DOS at conditions relevant for experimental studies.[14,18] For simplicity, we will therefore exploit the MT formalism below for inorganic and organic systems, assuming the position of the mobility edge (or the transport energy) at $\varepsilon = 0$ for both shapes of the DOS given by Eqs. (1) and (2).

A powerful model of charge carrier recombination accompanied by thermalization of carriers has been suggested by Orenstein and Kastner.[2] In this model, carriers perform the MT process, i.e., they are trapped into localized states with subsequent activation towards the mobility edge. The activation rate from a localized state with energy $\varepsilon < 0$ toward the mobility edge ($\varepsilon = 0$) is given by

$$\nu(\varepsilon) = \nu_0 \exp(\varepsilon/kT), \tag{3}$$

where ν_0 is the *attempt-to-escape* frequency usually taken of the order of the phonon frequency $\nu_0 \approx 10^{12}\,\text{s}^{-1}$, k is the Boltzmann constant, and T is temperature. As soon as a carrier appears at the mobility edge, there are two options for this carrier: (i) to be recaptured by a localized state with a capture coefficient b_t and herewith to continue the MT process, or (ii) to recombine via one of the recombination centers with a recombination coefficient b_r. The quantities of interest are the total time-dependent concentration of carriers $N(t)$ and the concentration of free carriers (those at the mobility edge) $n(t)$ after a pulsed excitation. The former concentration can be measured, for instance, by the transient photoabsorption and the latter one by the transient photoconductivity.[2,6]

Orenstein and Kastner obtained the dependencies $N(t)$ and $n(t)$ via an elegant and transparent analytical solution,[2] though valid only for the case $b_r \geq b_t$. The model suggested by Orenstein and Kastner[2] has later been exploited[19,20] to describe the transient photoconductivity in bulk homogeneous systems with exponential DOS given by Eq. (1). The assumption of quasiequilibrium allowed Adriaenssens *et al.*[19] to treat a realistic relation $b_r < b_t$ in contrast to the less realistic condition $b_r \geq b_t$ considered by Orenstein and Kastner.[2]

Literally, the same model has later been used by Nelson[21] to describe experimental data obtained for transient concentration of charge carriers in organic heterojunction solar cells. Nelson performed a straightforward computer simulation of the energy relaxation and recombination of carriers in a spatially homogeneous system and compared the results with experimental data obtained in heterojunction inhomogeneous materials. No comparison with analytical results obtained in the framework of the same model[2] has been attempted. The model has once more been studied in the framework of a detailed, though rather elaborate, analytical theory by Tachiya and Seki,[22] who were able to reproduce the data obtained numerically by Nelson.[21] In all these studies, the DOS was assumed purely exponential as given by

Eq. (1). In full analogy to Adriaenssens *et al.*,[19] Tachiya and Seki[22] assumed the system of recombining carriers to be in thermal equilibrium. In Sec. III, we will use the result by Tachiya and Seki[22] to check the validity of our approach in the case of thermal equilibrium for exponential DOS given by Eq. (1).

In the current report, we suggest an extension of the transparent approach by Orenstein and Kastner[2] for the realistic case $b_r < b_t$. Our approach is free from the assumption of thermal equilibrium necessary for the analytical treatments by Tachiya and Seki[22] and by Adriaenssens *et al.*,[19] i.e., it is valid for the case of thermal equilibrium as well as for the case of non-equilibrium carriers relaxing in energy. Furthermore, we apply the extended approach to describe recombination processes in systems with a Gaussian DOS inherent for organic disordered semiconductors and compare the results with those obtained for systems with exponential DOS. This comparison provides a recipe for how to distinguish between the Gaussian and exponential DOS by experimentally studying the transient phenomena.

In Sec. II, we briefly summarize the original theoretical approach by Orenstein and Kastner[2] valid in the case $b_r \geq b_t$. In Section III, we extend this approach for the realistic case $b_r < b_t$ and, in Sec. IV, check the results for the exponential DOS by comparison with the exact solution by Tachiya and Seki[22] valid in the case of thermal equilibrium. In Sec. V, we compare the results for a Gaussian DOS with those for exponential DOS and suggest a recipe for how to distinguish between these two shapes of the DOS by studying the transient concentrations of free and trapped carriers experimentally. Concluding remarks are gathered in Sec. VI.

II. DESCRIPTION OF RECOMBINATION VIA THE DEMARCATION ENERGY AT $b_r \geq b_t$

The key quantity determining recombination kinetics in the model by Orenstein and Kastner[2] is the so-called *demarcation energy*

$$\varepsilon_d(t) = -kT \ln(\nu_0 t), \tag{4}$$

determined by the condition[2,15]

$$\nu[\varepsilon_d(t)]t \approx 1. \tag{5}$$

In Eqs. (4) and (5), t is the time passed since the pulsed photoexcitation. For Eq. (5), one simply requires, according to Eq. (3), that within this time approximately one activation event to $\varepsilon = 0$ occurs. Hence, activation to the mobility edge $\varepsilon = 0$ is possible within time t from states with energies above $\varepsilon_d(t)$ and improbable from states below $\varepsilon_d(t)$. The states *above* $\varepsilon_d(t)$ can therefore be considered in thermal equilibrium, while states *below* $\varepsilon_d(t)$ are not able to thermalize within time t and are therefore occupied in accord with the DOS. Because the equilibrium distribution decays upwards above $\varepsilon_d(t)$ within scale kT and the DOS decays downwards within scale ε_0, the maximum of the carrier distribution corresponds to the vicinity of $\varepsilon_d(t)$.

Under these circumstances, the ratio between the concentration n of free carriers (those at the mobility edge) and

the concentration N of trapped carriers (those with energies around $\varepsilon_d(t)$) is

$$\frac{n}{N} = \frac{g(0)}{g(\varepsilon_d(t))}\exp\left[\frac{\varepsilon_d(t)}{kT}\right], \qquad (6)$$

assuming the equilibrium Boltzmann distribution for carriers above $\varepsilon_d(t)$. Because of the low concentration of intrinsic carriers, recombination in disordered materials is usually of a bimolecular character, i.e., photoinduced electrons recombine with photoinduced holes.[23] In this case, the recombination process is governed by the following dynamics:[2,23]

$$\frac{dN(t)}{dt} = -b_r N(t)n(t). \qquad (7)$$

This equation takes into account that mobile electrons (holes) with concentration n recombine with all holes (electrons) characterized by the concentration $N(t)$.

The concentration of trapped carriers, $N(t)$, is equal to the integral of the product of the density of states $g(\varepsilon)$ and the time-dependent energy distribution of carriers $f(\varepsilon, t)$. Because $f(\varepsilon, t)$ cuts off abruptly above $\varepsilon_d(t)$, the following assumption was used by Orenstein and Kastner:[2]

$$N(t) \simeq \int_{-\infty}^{0} f(\varepsilon,t)g(\varepsilon)d\varepsilon \simeq F(t)\int_{-\infty}^{\varepsilon_d} g(\varepsilon)d\varepsilon, \qquad (8)$$

where $F(t)$ is defined as the mean occupation number for $\varepsilon < \varepsilon_d$. Since $F(t)$ describes the occupation of states below ε_d, i.e., states with not yet released carriers, this quantity is determined by the equation[2]

$$\frac{dF(t)}{dt} \simeq b_t[1 - F(t)]n(t). \qquad (9)$$

Equations (7)–(9) determine the kinetics of both $N(t)$ and $n(t)$ self-consistently.[2] In this approach by Orenstein and Kastner, it is, however, assumed that $1 - F(t) \simeq 1$, i.e., that the average filling of states F is very low at all times t. This assumption is only valid when recombination is significantly more efficient than trapping ($b_r \gg b_t$) and no saturation of the carrier distribution can be established. Assuming the more realistic condition of $b_r \ll b_t$, at which carrier distribution saturates and $F(t)$ approaches unity, one comes to the clear difficulty of diverging concentrations using the above equations by Orenstein and Kastner.

In the next section, we introduce an extension of the model that does not depend on the assumption of low carrier occupancy and therefore covers more realistic parameter ranges than the original model.[2]

III. EXTENSION OF THE MODEL FOR THE CASE $b_r < b_t$

The key to the simplicity of the approach by Orenstein and Kastner is the approximation that carriers have a step-like distribution in energy, i.e., states below ε_d are uniformly occupied and states above ε_d are empty. We suggest to keep this approximation, which is reflected in Eq. (8), but to replace the demarcation energy ε_d (see Eq. (4)) by some effective energy ε^*, which is calculated in consistence with the carrier concentration dynamics (see below). Eq. (8) then reads

$$N(t) \simeq F(t)\int_{-\infty}^{\varepsilon^*(t)} g(\varepsilon)d\varepsilon, \qquad (10)$$

where $F(t)$ is the mean occupation number for $\varepsilon < \varepsilon^*$, in full analogy to the original approach by Orenstein and Kastner.[2] With ε_d replaced by ε^*, Eq. (9) means that detrapping is considered as inefficient for states with $\varepsilon \leq \varepsilon^*$.

In addition to Eqs. (7), (9), and (10), one more equation is needed in order to determine the dynamics of $F(t)$, $N(t)$, $n(t)$, and the new quantity $\varepsilon^*(t)$ self-consistently. For that purpose, we can use Eq. (6), which connects $n(t)$ and $N(t)$, replacing ε_d with ε^*,

$$\gamma(\varepsilon^*) = \frac{n}{N} = \frac{g(0)}{g[\varepsilon^*(t)]}\exp[\varepsilon^*(t)/kT]. \qquad (11)$$

By doing so, we assume that the concentration of free carriers is given by the Boltzmann tail of the trapped carrier concentration, which resembles thermal equilibrium above ε^*.

In order to obtain the carrier concentration dynamics, we now have to solve the system of ordinary differential equations: Eqs. (7) and (9) to (11). Straightforward calculations yield two coupled differential equations for $N(t)$ and $\varepsilon^*(t)$,

$$\frac{dN}{dt} = -b_r N^2(t)\gamma(\varepsilon^*), \qquad (12)$$

$$\frac{d\varepsilon^*}{dt} = -\frac{\gamma(\varepsilon^*)}{g(\varepsilon^*)}G(\varepsilon^*)N(t)\left[b_r - b_t + b_t\frac{G(\varepsilon^*)}{N(t)}\right], \qquad (13)$$

with

$$G(\varepsilon^*) = \int_{-\infty}^{\varepsilon^*} g(\varepsilon)d\varepsilon. \qquad (14)$$

To check the validity of the suggested approach, let us consider how $\varepsilon^*(t)$ behaves for $F(t) \ll 1$, where we would expect that $\varepsilon^* \simeq \varepsilon_d$. The condition $F(t) \ll 1$ is equal to the condition $N(t) \ll G(\varepsilon^*)$. If we assume the latter, Eq. (13) yields

$$\frac{d\varepsilon^*}{dt} = -b_t\frac{\gamma(\varepsilon^*)}{g(\varepsilon^*)}G^2(\varepsilon^*). \qquad (15)$$

Equation (15) can be separated and integrated over t and ε^*. Estimating $G(\varepsilon^*) \simeq Cg(\varepsilon^*)$ with some constant C, we arrive at

$$\varepsilon^*(t) = -kT\ln\left(\frac{b_t g(0)C^2}{kT}(t - t_0) + 1\right), \qquad (16)$$

where t_0 is the time, at which $\varepsilon^* = 0$. This time t_0 should be equal to ν_0^{-1} in the multiple-trapping description. If we

compare Eq. (16) with Eq. (4), we notice that the logarithmic time-dependence of $\varepsilon_d(t)$ is reproduced.

In the long-time limit, ε^* should approach the Fermi level ε_F, when $F(t)$ approaches unity. Indeed, setting $F(t) = 1$ in Eq. (10) makes ε^* equal to the Fermi level in the limit $T = 0$ K. At finite temperatures, the position of ε^* appears very close to ε_F, as clearly seen in Fig. 2(a). The difference between ε^* and ε_F is small as compared to the energy scale of the DOS.

IV. RESULTS FOR THE EXPONENTIAL DOS

Let us check the approach suggested in Sec. III by comparison with the exact, though elaborate solution by Tachiya and Seki[22] obtained for the exponential DOS. Inserting Eq. (1) into Eqs. (11)–(14), one obtains

$$\frac{dN}{dt} = -b_r N^2(t) \exp\left(\frac{\varepsilon^*}{kT} - \frac{\varepsilon^*}{\varepsilon_0}\right), \quad (17)$$

$$\frac{d\varepsilon^*}{dt} = N(t)\varepsilon_0 \exp\left(\frac{\varepsilon^*}{kT} - \frac{\varepsilon^*}{\varepsilon_0}\right)$$
$$\times \left[b_t - b_r - b_t \frac{N_t}{N(t)} \exp\left(\frac{\varepsilon^*}{\varepsilon_0}\right) \right]. \quad (18)$$

Equation (17) allows a direct comparison with the results in Ref. 22. The system of ordinary differential equations in Eq. (17) was solved numerically for different initial carrier concentrations N_0 and temperatures T. The value of the ratio
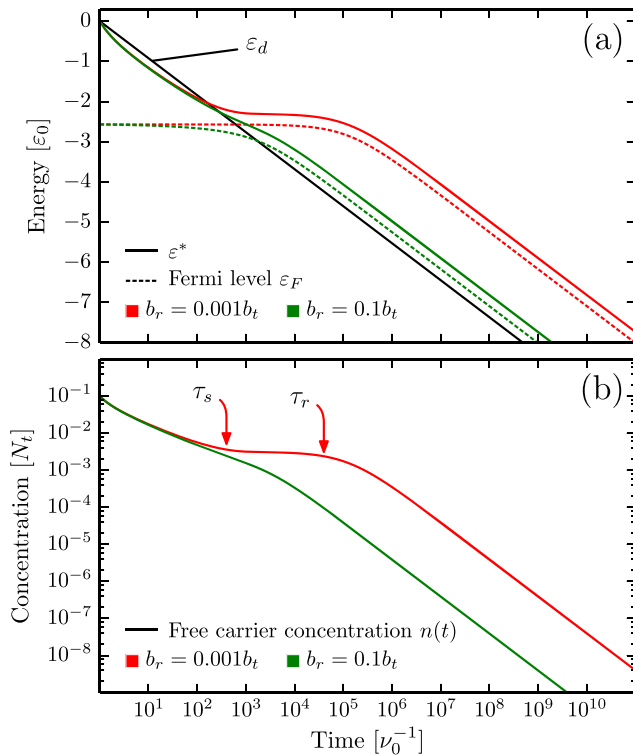
b_r/b_t was fixed at $b_r = 0.1 b_t$. The results are compared with the solution of Eq. (43) in Ref. 22. In order to perform the comparison, we assume that $\nu_0 = b_t N_t$. The results of Eq. (17) are plotted in Fig. 3 along with those from Ref. 22.

It is clearly visible from the figure that the approach from Eq. (17) agrees well with the results of the much more elaborate theory by Tachiya and Seki.[22] This good agreement confirms the validity of the approach suggested in Sec. III. Parameters for the comparison were chosen to justify that the thermal equilibrium is established before the launch of recombination, which is necessary for the approach in Ref. 22. For long times t, the typical power-law behavior[2,22] is observed, i.e., $N(t) \propto t^{-\alpha}$, where $\alpha = kT/\varepsilon_0$.

In Fig. 2(b), the time dependence of the free-carrier concentration $n(t)$ is plotted together with the dynamics of ε^*, showing a plateau for the case of very slow recombination $b_r = 0.001 b_t$. These results can be interpreted as follows. At some time, τ_s saturation is established, i.e., ε^* arrives at the Fermi level ε_F. At this point, $n(t)$ attains a stable value and does not decay further until recombination starts. When recombination comes in play at some time $\tau_r > \tau_s$, $n(t)$ starts to decay again because ε^* follows ε_F into the bandgap. Such behavior has been predicted by Orenstein and Kastner, though only qualitatively.[2] However, if recombination is fast enough, $b_r = 0.1 b_t$ in Fig. 2, this saturation behavior vanishes and the dynamics of $\varepsilon^* \approx \varepsilon_d$, as studied by Orenstein, and Kastner is recovered.

V. RESULTS FOR THE GAUSSIAN DOS

The extended approach suggested in Sec. III allows one to theoretically treat the charge carrier recombination

FIG. 2. Dynamics of ε^* and ε_F (a) and of the free carrier concentration (b) in the exponential DOS for two different recombination rates b_r. The beginnings of saturation at τ_s and of recombination at τ_r for the case $b_r = 0.001 b_t$ are marked by arrows. The demarcation energy ε_d from Eq. (4) is shown in (a) for comparison.
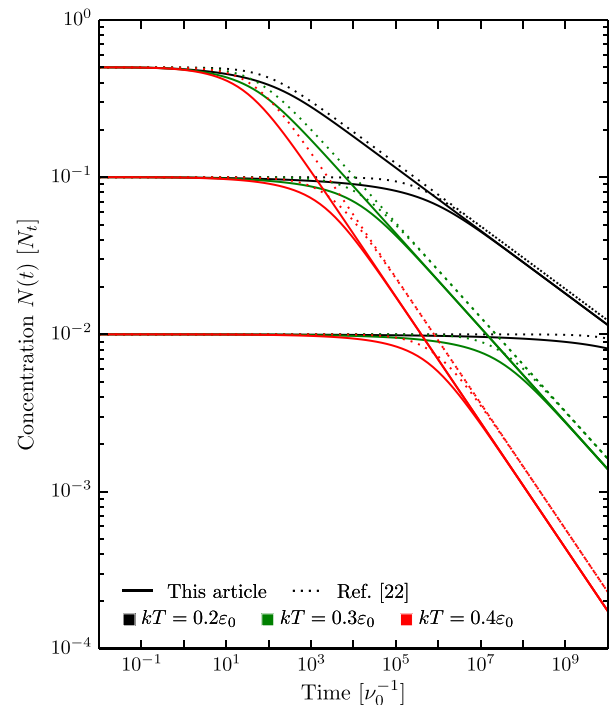

FIG. 3. Concentration of the carriers $N(t)$ given by Eq. (17) at $b_r = 0.1 b_t$ for different initial carrier concentrations and different temperatures compared to the results of Ref. 22.

accompanied by the energy relaxation of carriers not only in the case of exponential DOS, as described in Sec. IV, but also for other shapes of the energy spectrum, in particular, for the Gaussian DOS given by Eq. (2), inherent for disordered organic semiconductors.[11–14] Treating systems with the DOS steeper than the exponential one given by Eq. (1), one should, however, take into account the principal difference in the way of carrier thermalization in systems with the exponential DOS and those with steeper DOS functions.

In the exponential DOS, carriers dive in energy until no more unoccupied states with deeper energy are available, i.e., the system is saturated. In thermal equilibrium, the carrier distribution $g(\varepsilon)f(\varepsilon, \varepsilon_F)$ has a maximum at the Fermi level ε_F, decaying downwards in accord with the DOS and upwards in accord with the Boltzmann tail of the Fermi function. Hence, the equilibrium carrier distribution is determined by the position of the Fermi level, rendering all electronic processes in the exponential DOS concentration-dependent.

In the Gaussian DOS, the equilibration behavior appears essentially different to that in the exponential DOS. At low carrier concentrations and in thermal equilibrium, the carrier energy distribution $g(\varepsilon)f(\varepsilon, \varepsilon_F)$ has its maximum not at the Fermi level, but rather at some energy ε_∞ *independent* of the concentration of charge carriers.[11,14] This happens because the DOS decays downwards in energy more steeply than $f(\varepsilon, \varepsilon_F)$ increases. The energy ε_∞ is called *equilibration energy* and is calculated as[11]

$$\varepsilon_\infty = \frac{\int_{-\infty}^{\infty} \varepsilon g(\varepsilon)\exp(-\varepsilon/kT)d\varepsilon}{\int_{-\infty}^{\infty} g(\varepsilon)\exp(-\varepsilon/kT)d\varepsilon} = -\frac{\sigma^2}{kT}. \tag{19}$$

In the Gaussian DOS, a charge carrier relaxes from high-energy states downwards in energy only until it arrives at the equilibration energy ε_∞, even in an empty system without any interaction between relaxing carriers.[11,14] These relaxation dynamics have been studied in detail by analytical theories[24] and computer simulations.[11] In the simulations, noninteracting carriers were initially distributed uniformly over states in the Gaussian DOS, and their energy relaxation was traced. In course of time, the distribution moves downwards until its maximum arrives at the equilibration energy ε_∞. Besides other consequences, this observation means that our picture of carrier energy relaxation, described by the time-dependent energy $\varepsilon^*(t)$ in Sec. III, should be replaced by the time-independent equilibrium energy distribution of carriers (governed by ε_∞), as soon as $\varepsilon^*(t)$ arrives at ε_∞ at a time τ_{rel} determined by the condition $\varepsilon^*(\tau_{rel}) = \varepsilon_\infty$. This picture of carrier relaxation with the decisive role of ε_∞ is valid if the concentration of carriers is not too large and the condition $\varepsilon_\infty > \varepsilon_F$ holds, which we assume throughout the following analysis. Because of the role of the equilibration energy ε_∞, one should separately consider the time intervals $t < \tau_{rel}$ and $t \geq \tau_{rel}$.

At $t < \tau_{rel}$, one can evaluate Eqs. (11)–(14) using the DOS from Eq. (2), obtaining

$$\frac{dN}{dt} = -b_r N^2(t)\exp\left[\left(\frac{\varepsilon^*}{\sqrt{2}\sigma} + \frac{\sigma}{\sqrt{2}kT}\right)^2\right]\exp\left[-\frac{1}{2}\left(\frac{\sigma}{kT}\right)^2\right], \tag{20}$$

$$\frac{d\varepsilon^*}{dt} = N(t)\sigma\sqrt{\frac{\pi}{2}}\exp\left[\left(\frac{\varepsilon^*}{\sigma} + \frac{\sigma}{2kT}\right)^2\right]\exp\left[-\frac{1}{4}\left(\frac{\sigma}{kT}\right)^2\right]$$
$$\times \left[1 + \mathrm{erf}\left(\frac{\varepsilon^*}{\sqrt{2}\sigma}\right)\right]$$
$$\times \left[b_t - b_r - b_t\frac{N_t}{N(t)}\left(1 + \mathrm{erf}\left(\frac{\varepsilon^*}{\sqrt{2}\sigma}\right)\right)\right]. \tag{21}$$

Equations (20) and (21) describe the dynamics of ε^* and $N(t)$ in the case of a Gaussian DOS before the carrier distribution has relaxed to ε_∞.

At $t = \tau_{rel}$, i.e., as soon as $\varepsilon^*(t)$, obtained from Eq. (21) arrives at ε_∞, the recombination becomes controlled by the conventional equation for bimolecular recombination with the time-independent recombination coefficient

$$\frac{dN}{dt} = -b_r N^2(t)\exp\left[-\frac{1}{2}\left(\frac{\sigma}{kT}\right)^2\right]. \tag{22}$$

Equation (22) is the result of Eqs. (7) and (11) with $\varepsilon^*(t)$ replaced by ε_∞. It has the solution

$$N(t) = \frac{N(\tau_{rel})}{1 + (t - \tau_{rel})N(\tau_{rel})b_r\exp\left[-\frac{1}{2}\left(\frac{\sigma}{kT}\right)^2\right]}. \tag{23}$$

Results for the concentration of trapped carriers $N(t)$ obtained at $\varepsilon^*(t) > \varepsilon_\infty$ via solution of Eqs. (20) and (21) and at $\varepsilon^*(t) < \varepsilon_\infty$ via Eq. (23) are shown by solid lines in Fig. 4.
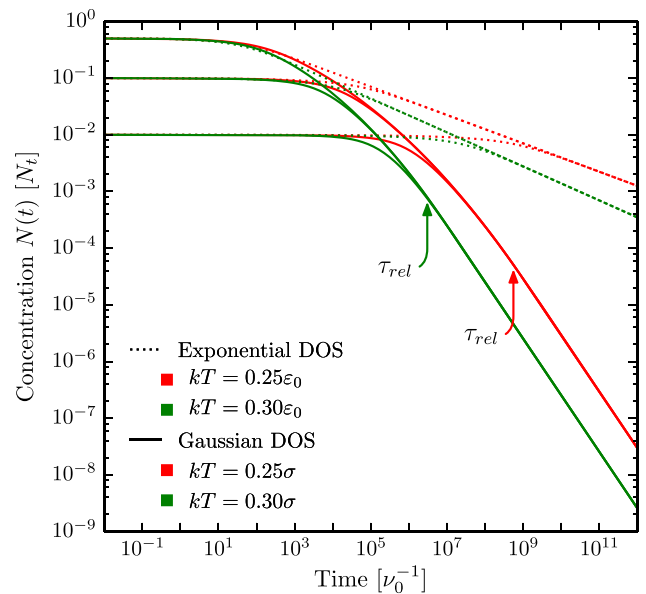
FIG. 4. Comparison of the dynamics of $N(t)$ for the Gaussian and exponential DOS, shown at $b_r = 0.1b_t$ for different temperatures T and different initial concentrations $N(0)$. The values of the relaxation time τ_{rel} determined by the condition $\varepsilon^*(\tau_{rel}) = \varepsilon_\infty$ are marked by arrows.

The time evolution of the concentration of trapped carriers $N(t)$ can be traced experimentally by measuring the intensity of the transient photoabsorption, which is linearly proportional to $N(t)$. A question arises herewith on whether it is possible to distinguish between systems with exponential DOS and those with Gaussian DOS by viewing experimental transients. In order to answer this question, in Fig. 4, we compare the results for the Gaussian DOS with those obtained for $N(t)$ in the case of the exponential DOS, discussed in Sec. IV, for the same set of parameters. In the calculation, the recombination coefficient was $b_r = 0.1b_t$ and, for simplicity, we chose the same values of ε_0 and σ for the exponential and Gaussian DOS, respectively. The difference in the magnitudes of the time-dependent decrease of $N(t)$ for the two different shapes of the DOS seen in Fig. 4 is very pronounced. While the exponential DOS leads to a change in $N(t)$ by just a couple of orders of magnitude for the given time interval, the decrease of $N(t)$ in the case of the Gaussian DOS corresponds to more than six orders of magnitude in the same time interval. In the case of the Gaussian DOS, Eq. (23) predicts the dependence $N(t) \propto 1/t$ at large t, whereas in the case of the exponential DOS the recombination appears much less efficient than in the case of the Gaussian DOS. This happens because the recombination rate $dN(t)/dt$, in accord with Eq. (7), is proportional to the concentration of free carriers $n(t)$, which decays in course of time for the exponential DOS not only because of the recombination process, but also due to the energy relaxation of carriers into deeper states. In the Gaussian DOS, this energy relaxation is stopped at $t \simeq \tau_{rel}$. Concomitantly, $n(t)$ becomes linearly proportional to $N(t)$ at $t \geq \tau_{rel}$, meaning that $n(t)/N(t) = \text{const}$ in the Gaussian DOS whereas $n(t)/N(t)$ decreases with time in the exponential DOS. Since a larger proportion of the carriers is present above the mobility edge, this effect enhances the recombination rate in systems with Gaussian DOS as compared to those with exponential DOS.

The concentration of mobile carriers $n(t)$ can be measured, for instance, by tracing the transients of the photoconductivity after terminating the generation of carriers. Theoretical transients of $n(t)$ for two different types of the DOS are plotted in Fig. 5, for equal values of ε_0 and σ. The decays of $n(t)$ for Gaussian DOS are apparently steeper than those for exponential DOS.

VI. CONCLUSIONS

A theoretical description for recombination of charge carriers accompanied by their energy relaxation through localized states is provided. For systems with exponential DOS, the description by Orenstein and Kastner,[2] initially suggested only for fast recombination, is extended to the realistic case of slow recombination as compared to trapping of carriers into localized states. The theory is further extended for applications to systems with Gaussian DOS inherent for disordered organic semiconductors. In the latter case, thermal equilibration of carriers at the equilibration energy leads to a much more pronounced recombination than in the case of an exponential DOS. This more pronounced recombination in the case of the Gaussian DOS is reflected in the time-dependent transients of the concentrations of trapped $(N(t))$ and mobile $(n(t))$ carriers that appear steeper by several orders of magnitude than those in the exponential DOS. Since the dependences $N(t)$ and $n(t)$ can be checked experimentally by measuring the transient photoabsorption and the transient photoconductivity, respectively, the theory may help to experimentally distinguish between different DOS functions.
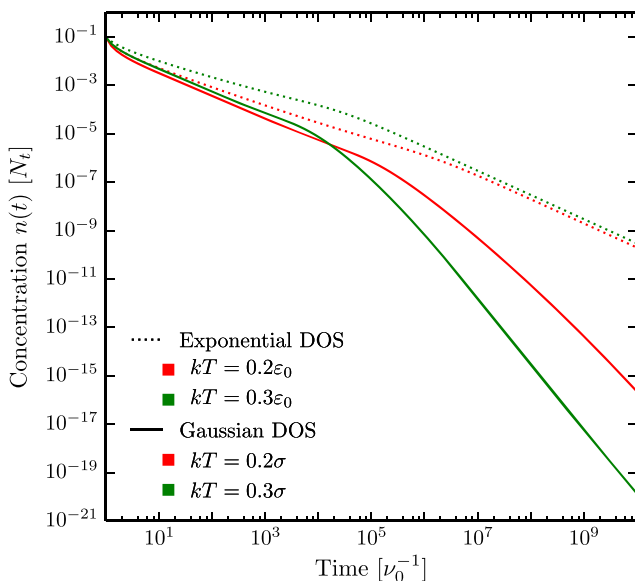
ACKNOWLEDGMENTS

FIG. 5. Dynamics of the free carrier concentration $n(t)$ at $b_r = 0.1b_t$ for both shapes of the DOS given by Eqs. (1) and (2) at $\varepsilon_0 = \sigma$.

[1]T. Tiedje and A. Rose, Solid State Commun. **37**, 49 (1981).
[2]J. Orenstein and M. Kastner, Solid State Commun. **40**, 85 (1981).
[3]J. Noolandi, Phys. Rev. B **16**, 4466 (1977).
[4]F. W. Schmidlin, Phys. Rev. B **16**, 2362 (1977).
[5]M. Silver and L. Cohen, Phys. Rev. B **15**, 3276 (1977).
[6]J. Orenstein and M. Kastner, Phys. Rev. Lett. **46**, 1421 (1981).
[7]*Charge Transport in Disordered Solids with Applications in Electronics*, edited by S. Baranovski (John Wiley & Sons, Ltd., Chichester, 2006).
[8]R. A. Street, J. E. Northrup, and A. Salleo, Phys. Rev. B **71**, 165202 (2005).
[9]R. A. Street, K. W. Song, J. E. Northrup, and S. Cowan, Phys. Rev. B **83**, 165207 (2011).
[10]P. Stallinga, Adv. Mater. **23**, 3356 (2011).
[11]H. Bässler, Phys. Status Solidi B **175**, 15 (1993).
[12]M. Pope and C. E. Swenberg, *Electronic Processes in Organic Crystals and Polymers* (Oxford University Press, Oxford, 1999).
[13]J. O. Oelerich, D. Huemmer, and S. D. Baranovskii, Phys. Rev. Lett. **108**, 226403 (2012).
[14]S. D. Baranovskii, Phys. Status Solidi B **251**, 487 (2014).
[15]D. Monroe, Phys. Rev. Lett. **54**, 146 (1985).
[16]S. D. Baranovskii, P. Thomas, and G. J. Adriaenssens, J. Non-Cryst. Solids **190**, 283 (1995).
[17]S. D. Baranovskii, T. Faber, F. Hensel, and P. Thomas, J. Phys.: Condens. Matter **9**, 2699 (1997).
[18]J. O. Oelerich, D. Huemmer, M. Weseloh, and S. D. Baranovskii, Appl. Phys. Lett. **97**, 143302 (2010).

[19]G. J. Adriaenssens, S. D. Baranovskii, W. Fuhs, J. Jansen, and O. Öktü, Phys. Rev. B **51**, 9661 (1995).

[20]H. Cordes, G. H. Bauer, and R. Brüggemann, Phys. Rev. B **58**, 16160 (1998).

[21]J. Nelson, Phys. Rev. B **67**, 155209 (2003).

[22]M. Tachiya and K. Seki, Phys. Rev. B **82**, 085201 (2010).

[23]B. I. Shklovskii, E. I. Levin, H. Fritzsche, and S. D. Baranovskii, "Hopping photoconductivity in amorphous semiconductors: Dependence on temperature, electric field and frequency," in *Advances in Disordered Semiconductors*, vol. 3 (World Scientific, Singapore, 1990), pp. 161–191.

[24]S. D. Baranovskii, H. Cordes, F. Hensel, and G. Leising, Phys. Rev. B **62**, 7934 (2000).

Article V

K. Werner, A. Beyer, **J. O. Oelerich**, S. D. Baranovskii, W. Stolz, and K. Volz
Structural characteristics of gallium metal deposited on Si (001) by MOCVD
J. Cryst. Growth **405**, 102 (2014)
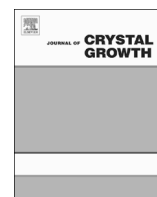
Structural characteristics of gallium metal deposited on Si (0 0 1) by MOCVD

K. Werner *, A. Beyer, J.O. Oelerich, S.D. Baranovskii, W. Stolz, K. Volz

Materials Science Center and Faculty of Physics, Philipps-Universität Marburg, 35032 Marburg, Germany

ABSTRACT

The epitaxial growth of high quality gallium containing III/V material on silicon offers a wide field of possible applications. While the growth of III/V material has already been studied extensively, little is known about the atomic processes at the interface between silicon and the III/V material. Here we study the deposition of gallium on silicon (0 0 1) by metal organic chemical vapour deposition for the two different precursors triethyl gallium and trimethyl gallium, the influence of varying growth conditions as well as the influence of an annealing after growth. We use atomic force microscopy and high angle annular dark field scanning transmission electron microscopy to analyse the detailed structures occurring for this deposition. For all conditions studied, pyramidal gallium containing structures form in the silicon with boundaries on the silicon {1 1 1} lattice planes. The pyramidal etching structures can be reproduced in straightforward kinetic Monte Carlo simulations of the growth process. When triethyl gallium is used, those structures are covered by carbon containing passivating mounds. After an additional annealing at 800 °C no gallium can be detected anymore for any sample, while an annealing at considerably lower temperatures leads to an alteration of the silicon gallium alloy.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Epitaxial growth of Ga-containing III/V semiconductors on silicon has attracted great attention due to the wide range of potential applications it offers [1–6]. Here GaP on Si is of special interest, due to the similarity in lattice constant of the two materials. Although the epitaxial deposition of polar III/V materials on non-polar Si substrates poses severe challenges [7,8], the usage of GaP as nucleation layer is a promising approach for the growth of high quality III/V alloys on Si substrate, as one is able to separate defects arising due to lattice mismatch from those arising due to the growth of a binary zinc-blende on a diamond structure material. A lot of work has already been done on the growth of GaP [9], GaN [10–12] and other Ga-containing III/V materials [7,13,14] on Si. The fundamentals of the Ga deposition on Si (0 0 1) in ultra high vacuum (UHV) by an elemental Ga beam have been investigated by STM, LEED and Auger electron spectroscopy for low surface coverage in a wide temperature regime by several groups [15–17]. It was found that, depending on the surface coverage, the Ga atoms form dimer rows or well ordered areas on Si. For those studies, though, the conditions differ drastically from those

used for metal organic chemical vapour deposition (MOCVD) of Ga containing material, since elemental Ga was deposited in an UHV environment at relatively low temperatures between 20 °C and 450 °C. The deposition of Ga and other metal droplets on GaAs by molecular beam epitaxy for growth conditions similar to those used in MOCVD has been investigated before experimentally and by kinetic Monte Carlo (KMC) simulations [18,19]. It was shown that the droplet formation strongly depends on structural characteristics such as orientation or surface termination of the substrate used as well as on the occurrence of surface defects. While our findings are in some parts in good agreement with the previous theoretical work [18], the atomic processes occurring for the droplet formation on Si seem to differ from those observed for III/V substrates. Also, the influence of phosphorus deposition on a Si surface has already been investigated using phosphine (PH_3) as precursor [20,21]. It was found that although an exposure of Si to PH_3 leads to a roughening of the surface, a short pre-flow of the P precursor PH_3 or tertiarybutyl phosphine (TBP) is necessary to obtain a sufficient quality of the GaP nucleation layer [9,22,23]. Meanwhile, not much is known yet about the atomic processes during the deposition of Ga on Si by MOCVD. However, a thorough understanding of those growth processes is essential to optimize the interfaces of Ga containing III/V material on Si. For GaN on Si grown at comparatively high temperatures the effect of melt-back etching is well known [24,25], however the atomistic processes associated with melt-back etching have not been extensively studied.

---

* Corresponding author. Tel.: +49 6421 28 25610.
E-mail address: Katharina.werner@physik.uni-marburg.de (K. Werner).

It was shown that this effect can also occur for GaP grown on Si (0 0 1) [9,26] and that it influences the characteristics of the III/V layers on Si decisively [27], evidencing that growth conditions have to be chosen with great care in order to achieve smooth interfaces and reproducible performance. For Ga deposition on Si by molecular beam epitaxy at temperatures between 440 °C and 550 °C, thus in the regime used for GaP nucleation by MOCVD, it was found that pits form in the Si substrate underneath the deposited Ga droplets [26]. Although this will supposedly have an influence on the interface of GaP and Si, to our knowledge no further investigation was carried out on this effect.

Aside from applications in III/V nucleation on Si, the deposition of metallic Ga on Si has gained significance in the fabrication of low dimensional structures in future devices [28,29]. Metallic Ga, having one of the largest liquid ranges (29.78–2400 °C) of any metal, combined with a low vapour pressure, is a suitable candidate as solvent and catalyst for the synthesis of several crystals. Its applicability as solvent for the low-temperature synthesis of several kinds of one-dimensional nanostructures like carbon nanotubes [30], Si [31] and SiO_2 [32] nanowires or GaAs whiskers [33] or for the growth of GaAs crystals [34] has already been shown. In order to also apply this to MOCVD growth, a method for the controlled deposition of metallic Ga droplets has to be established first and the structure of the deposited Ga has to be understood.

With the atomic resolution achieved by high-resolution scanning transmission electron microscopy (STEM) we are now able to thoroughly analyse the detailed structures occurring after the deposition of Ga on Si (0 0 1). In addition, using kinetic Monte Carlo (KMC) modelling, we are able to model the experimental findings theoretically and understand the kinetic driving forces of structure formation in the Ga–Si system.

This paper presents the results on the deposition of metallic Ga on Si (0 0 1) by MOCVD using the Ga precursors trimethyl gallium (TMGa) and triethyl gallium (TEGa). The influence of several growth parameters, such as the growth temperature, the Ga partial pressure and the growth time, on the structure and composition of the resulting Ga droplets has been investigated and will be discussed. The structures found will be compared to results from atomistic KMC modelling.

2. Experimental and theoretical methods

All investigated samples were grown via MOCVD in a commercial, horizontal MOCVD reactor system (AIX 200) using Pd-purified H_2 as carrier gas. To achieve oxygen free surfaces the nominal exact Si (0 0 1) substrates were first treated by a wet chemical etching procedure similar to the Shiraki process [35] followed by a bake out at 975 °C in the MOCVD reactor under H_2 atmosphere directly before growth. For all samples a homoepitaxial 1 μm thick Si buffer was grown firstly using silane (SiH_4) as precursor.

For the Ga growth TMGa and TEGa were used to study the influence of the precursor molecule on the grown structure. The reactor pressure was 50 mbar, while the growth temperature as well as the growth time and the Ga partial pressure were varied systematically to analyse the dependence of the Ga deposition on these parameters. Growth temperatures were chosen between 375 °C and 450 °C when TEGa was used while the Ga partial pressure for this precursor was varied in the range between 7.6×10^{-4} mbar and 1×10^{-2} mbar and the growth time was selected between 1 s and 15 s. For TMGa the growth temperature was varied between 450 °C and 550 °C, the Ga partial pressure was chosen between 1×10^{-2} mbar and 1 mbar and the growth time was varied between 1 s and 40 s. Selected samples were annealed in a separate experiment after the growth to study the influence of temperatures between 400 °C and 800 °C on the Ga containing

structures. The annealing was performed in the growth reactor under H_2 atmosphere at 50 mbar reactor pressure.

Atomic force microscopy (AFM) (Nanoscope Multimode IIIa), utilizing tapping mode under ambient conditions, was carried out to study the surface morphology of the samples after growth. For some samples AFM analysis was used in combination with selective etching with hydrochloric acid (HCl), which removes Ga while Si is unaffected by the etchant. For STEM imaging, electron transparent samples were prepared in cross-sectional geometry by conventional mechanical thinning and argon ion milling. The viewing direction was chosen along a ⟨1 1 0⟩ direction of Si. A JEOL JEM 2200FS, equipped with a probe aberration-corrector and operating at 200 kV, in combination with energy dispersive X-ray spectroscopy (EDX), was used to characterize the structural quality and the composition of the samples.

The deposition of Ga on Si substrate was simulated using the Bortz–Kalos–Lebowitz kinetic Monte Carlo algorithm [36] on a diamond cubic lattice structure. Only diffusion and adsorption events were enabled in the simulation. Diffusion barriers were calculated by simple counting of first- and second-nearest-neighbour bonds between the current atom and its surrounding [19]. Periodic boundary conditions of the system were assumed perpendicular to the growth direction [0 0 1]. Surface reconstruction and lattice strain were not taken into account. The system size was chosen as 22 nm × 22 nm × 44 nm, which corresponds to about 1 million simulated sites.

3. Results and discussion

3.1. TEGa deposition

For the Ga deposition using TEGa as precursor molecule the growth parameters were adopted from earlier results on the growth of GaP on Si (0 0 1) [9]. Judging from those, the amount of Ga deposited for different samples corresponds to 0.1 to 1.5 monolayers (ML). AFM and high angle annular dark field (HAADF) STEM images of those samples, grown at different growth temperatures, are shown in Fig. 1. After 15 s of Ga deposition at 450 °C, corresponding to 1.5 ML, the sample surface is densely covered with mounds, as can be seen in the AFM image (Fig. 1(a)). For all investigated samples, we found that the density and size of the mounds increases with increasing growth time. In addition, the Si surface steps between the mounds can only be clearly detected for samples with less than 0.5 ML of TEGa deposited, indicating that some material is also deposited across the entire surface. An AFM image of the sample after selective etching (Fig. 1(b)) shows that the mounds deposited with TEGa are only marginally affected in size and height by hydrochloric acid. This evidences that the deposited structures are not pure metallic Ga droplets. Results obtained by selective etching with organic solvents suggest that the mounds are carbon containing. This can be explained by a formation and deposition of polyethylene chains by the residual organic groups of the TEGa precursor molecule due to the catalytic effect of the deposited metallic Ga. A HAADF STEM measurement of a sample deposited under equivalent growth conditions is displayed in Fig. 1(c). It clearly reveals a bright pyramidal structure in Si after the Ga deposition, showing significant Ga diffusion into the originally smooth Si substrate. The Si substrate is dissolved in triangular structures by Ga. For better visibility, the triangular structure in Si is marked by a dashed white line. The position of this guide to the eye was determined in the low pass filtered image (Fig. 1(d)), which reveals the difference in contrast between the pure Si and the pyramidal structure more clearly. This was also done for all other presented STEM images, without showing the filtered images here. EDX measurements (data from line scan in Fig. 1(e)) evidence that the brighter contrast arises due to Ga incorporated in Si. In the marked section in Fig. 1(c) and (f), the
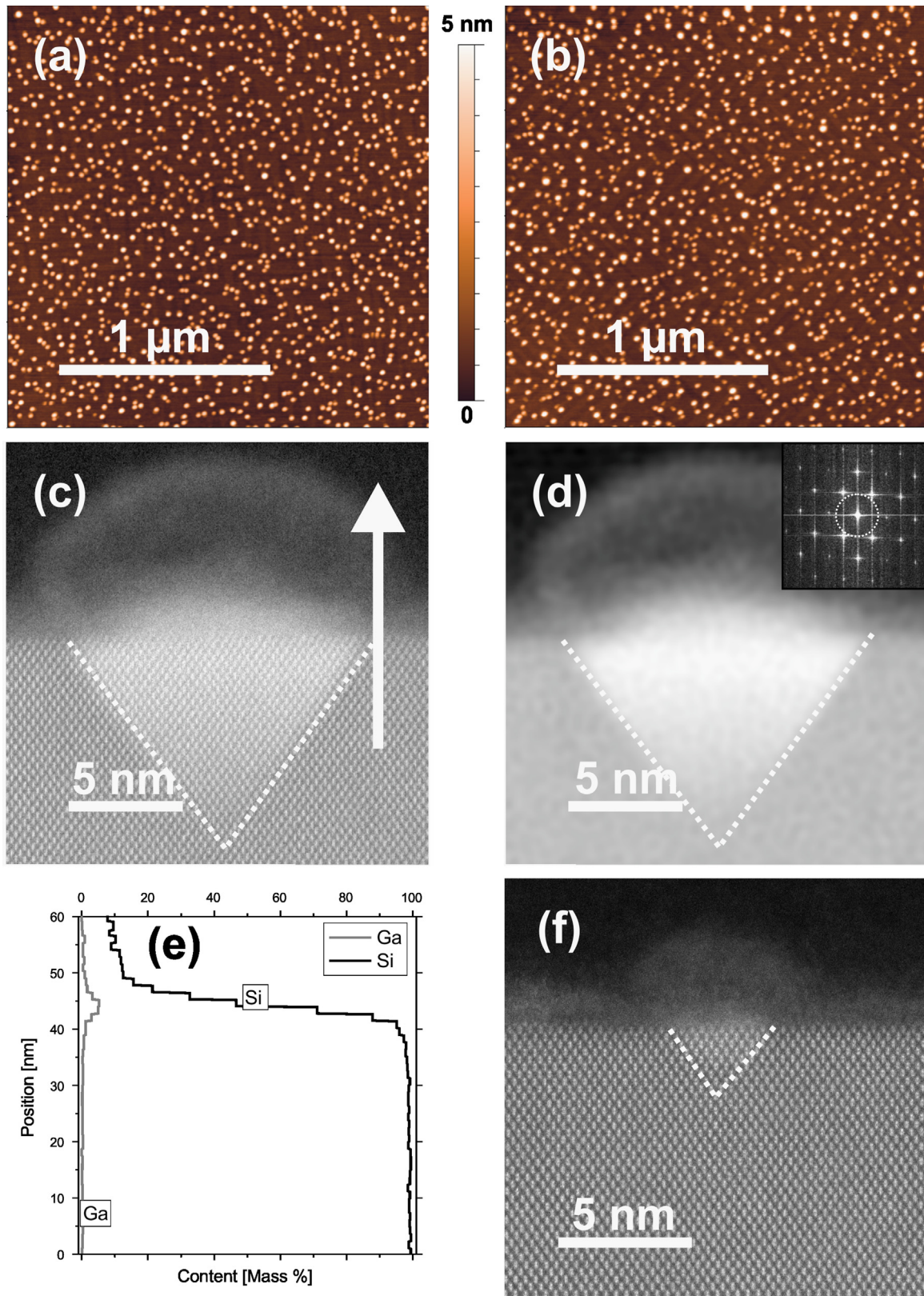
Fig. 1. AFM measurements of the sample surface after 15 s deposition of TEGa at 450 °C before (a) and after etching with HCl for 30 s (b). The cross-sectional HAADF STEM image of a sample after deposition using the same growth conditions (c) reveals bright pyramidal structures which are framed by boundaries on {1 1 1} Si lattice planes (marked by dashed lines). The contrast of the pyramid becomes more pronounced in the low pass filtered image (d). The inset shows the corresponding Fourier transformation and the size of the mask used for filtering. EDX line profiles acquired along the growth direction (e) for Si (black line) and Ga (grey line) indicate that the pyramids contain Ga. To obtain a sufficient EDX intensity a larger STEM spot size was chosen for the measurement. Additionally the signal was averaged over the width of the pyramid along the horizontal axis of the image. In the HAADF image of the structure grown at a lower temperature of 400 °C (f) the pyramidal structure is less distinctive due to the decreased size of the Ga containing structure in relation to the Si lattice in the direction of the electron beam.

Si crystal structure can clearly be seen in the STEM images. As the investigated TEM samples have a finite thickness, we attribute this to a fraction of intact Si crystal situated before and behind the pyramidal Ga containing structure with respect to the electron beam direction rather than to a mixed crystalline lattice of Si and Ga. This hypothesis is supported by the fact that the bright contrast in the pyramid mainly originates from the regions between, rather than on, the atomic columns, resulting in the better visibility of the structures in the low pass filtered image (Fig. 1(d)). The boundaries of the Ga pyramids are located on the Si {1 1 1} planes. We assume this to be based on the noticeably larger stability of the Si {1 1 1} planes in comparison to the other Si planes, which has already been shown in etching experiments [37]. Additionally, the STEM image reveals that the sample is capped by an amorphous layer. It stretches across the whole sample surface, but is clearly more pronounced above the pyramidal Ga containing structures and, as mentioned above, can be explained by the deposition of residual organic groups from the metal organic precursor molecule used.

Lowering the growth temperature from 450 °C to 400 °C, while at the same time increasing the TEGa partial pressure to compensate for the decreased decomposition of the precursor at this temperature, in order to achieve a similar amount of deposited material on the surface, leads to less pronounced pyramidal structures in the sample (Fig. 1(f)). This implies that less Ga etches into Si with decreasing growth temperature. The amorphous layer forming mounds above the Ga pyramids can still clearly be detected. But selective etching with hydrochloric acid, not shown here, suggests that the Ga content in those mounds is increased in comparison to samples grown with higher growth temperatures. This is in good agreement with less diffusion of Ga into Si, since the integral amount of Ga remains constant.

3.2. Theoretical modelling

To gain insight into the formation mechanisms of the etching structures, the Ga deposition on Si substrate was modelled using a standard KMC technique for epitaxial growth simulation, as described in the methods section above. The event rates of the diffusion events were calculated by a simple Arrhenius equation, with the constant prefactor of $2.5 \times 10^{13}\ \mathrm{s^{-1}}$. The simulated growth procedure was the following. 25 ML of Si were grown onto the two initial layers of Si atoms. The resulting Si surface was flat with only single atoms or unstable islands in the first ML above the surface. Then, 10 ML of Ga were deposited over a total time of 30 s, followed by a relaxation time of 5 s without any growth events. Note, that due to sample size restrictions in the simulation, more MLs of Ga are needed than in the experiment where Ga can be accumulated from larger areas. The temperature in the simulation was chosen as 477 °C, close to the experimental conditions. The bond strengths for the surface hopping barriers were Si-Si= (0.3 eV, 0.15 eV), Si–Ga=(0.3 eV, 0.15 eV), and Ga–Ga=(0.1 eV, 0.25 eV), for the first- and second-nearest-neighbour bonds, respectively. These bond strengths are in the same range as values used in comparable simulations in [19] and were adjusted to fit the experimental results. Of importance here are, not the absolute values, but the ratios of the values. The liquid character of the metallic Ga is represented by the small overall Ga–Ga bond strength but comparatively large bond strength for the second-nearest-neighbour bonds.

The adsorbed Ga atoms led to a significant etching of the Si substrate, as shown in Fig. 2. The figure shows only the Si content of the sample after the simulation, projected onto the {1 1 0} plane (a) and as a 3D image (b). It is clearly visible that the etched
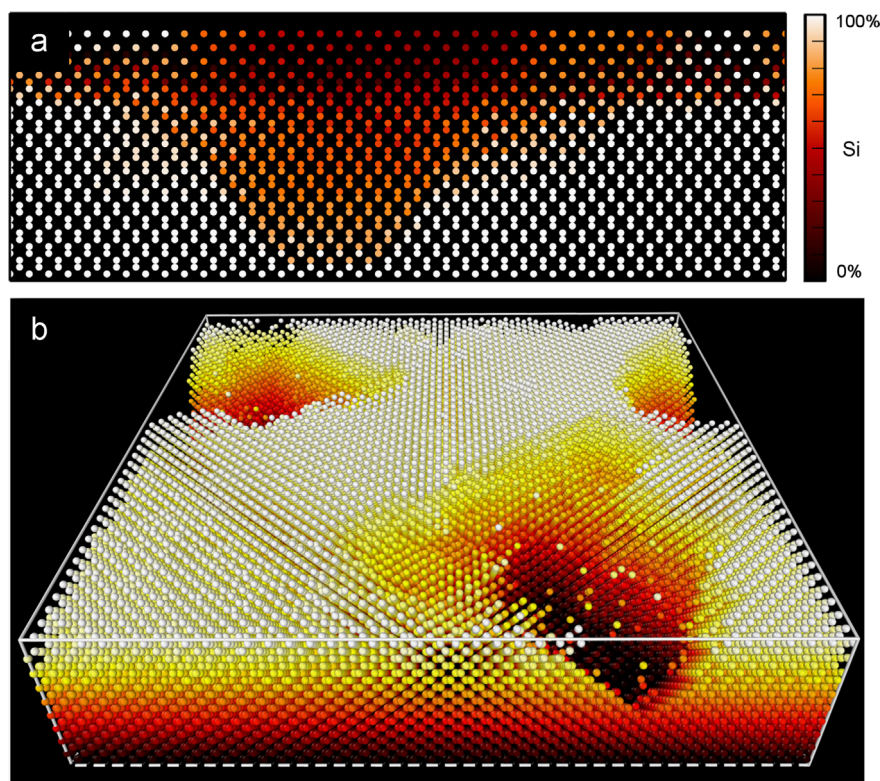

Fig. 2. This figure shows the kinetic Monte Carlo simulation results. In both parts only the Si content is shown and Ga is hidden. 10 ML of Ga were grown during a simulation time of 30 s at a simulation temperature of 477 °C. Periodic boundary conditions were applied in the horizontal axes. The KMC bond energies were: Si–Si=(0.3 eV, 0.15 eV), Si–Ga=(0.3 eV, 0.15 eV), Ga–Ga=(0.1 eV, 0.25 eV). (a) Shows the projected, integrated Si content along the {1 1 0} direction, with white columns being 100% Si and decreasing Si content with darker colours. The Si substrate after growth in 3D is shown in (b). Higher atoms with respect to the growth direction are coloured in bright colours, dark colors correspond to deep atoms. (For interpretation of the references to color, the reader is referred to the online version of this article.)

structures are predominantly limited by Si {1 1 1} planes, in complete agreement with the experimentally observed behaviour. The size of the structures is sensitive to the growth time of Ga or the temperature, with increasing temperatures and longer growth times leading to larger pyramidal voids. The structures are filled with bulk Ga, not shown here, which also corresponds to the experimental results. The depth of the structure in the simulation result was limited by the fixed two layers of Si at the sample bottom, causing the flat {0 0 1} boundary.

During the Ga deposition, formation of Ga–Si intermixed phases on the substrate was observed. These phases are several MLs thick and cover the surface area between the etching structures, i.e. they appear everywhere except in or above the inverted pyramids below the Ga droplets. The Ga–Si phases remain stable and are present until the simulation terminates. They are clearly visible in Fig. 2(a), where a finite Si concentration is present only in every second ML within the top few layers of the simulated sample. The dynamics of the Ga droplet formation on the Si surface is not yet well understood theoretically. So far, the intermediate growth phases have been studied only for Ga droplet formation on GaAs surfaces, where surface defects have been suggested as playing a role for etching structure formation [18].

3.3. TMGa deposition

Aside from TEGa, TMGa is a Ga precursor often used in MOCVD due to its high vapour pressure and its great stability at room temperature [38]. Based on the different reaction pathways for the decomposition, differing residual organic groups occur for TEGa and TMGa. While TEGa presumably decomposes primarily via β-hydride elimination at high temperatures yielding Ga, ethylene and hydrogen, it is assumed that TMGa decomposes by hydrogen transfer processes generating Ga, methane, hydrogen and carbon [39]. It is known that the carbon incorporation, for example for the growth of GaAs, can be significantly decreased by the use of the Ga precursor TEGa rather than TMGa [40]. However, based on the metallic character of the deposited Ga, the growth processes leading to an incorporation of the residual organic groups might be thoroughly different here. No formation of polyethylene chains is expected from the methane resulting from the TMGa decomposition. Since TMGa, having a 50% decomposition temperatures of approximately 450 °C in H_2, only decomposes completely at higher reactor temperatures than TEGa, which has a 50% decomposition temperatures of approximately 280 °C in H2 [38,41], the Ga partial pressure and for some samples also the temperature for the deposition using TMGa as precursor were chosen higher than when TEGa was used. The partial pressure for the Ga precursor was increased by a factor of 130 for TMGa compared to TEGa to achieve an similar deposition. This is due to the different reaction pathways for the decomposition. At the chosen growth temperature, the TEGa molecules presumably partly decompose in the gas phase and intermediate precursor molecules form, which adsorb and fully decompose on the hydrogen passivated Si surface. For TMGa the decomposition does not seem to occur for the individual molecules. An AFM image of a sample with a Ga deposition for 10 s at a growth temperature of 500 °C is shown in Fig. 3(a). It clearly reveals that, in comparison to the samples using TEGa as precursor, the number of mounds formed on the Si surface is decreased while the mound size is drastically increased. Between the mounds the stepped Si surface is still clearly visible, remaining unaffected by the Ga deposition. This indicates that the decomposition of the TMGa molecules is not induced by the hydrogen passivated Si surface but takes place only in the presence of multiple precursor molecules. Due to the low partial pressure of the Ga precursor in comparison to the H_2 carrier gas, the probability of clustering of the molecules is low and most TMGa

molecules pass through the reactor without decomposing. An AFM image of the same sample after selective etching with hydrochloric acid (Fig. 3(b)) displays that, in contrast to the results using TEGa, the structures deposited with TMGa can be etched off, indicating that, for this precursor structures form, consisting mostly of Ga, which are not covered by a passivating layer. In addition holes are visible in the Si surface after the selective etching. Those are due to Ga etching the Si surface during the deposition, as Ga was then removed by HCl. This result is also supported by the cross-sectional HAADF STEM image (Fig. 3(c)) of a sample deposited with growth parameters equivalent to those used for the sample shown in Fig. 3(a), showing bright pyramidal structures, too, for this precursor. Again, the structure's boundaries are on the most stable Si {1 1 1} planes, but, as already observed for the mounds forming on the surface, here the structures etched into Si are considerably larger, but in much lower density, than those found for TEGa. As before, EDX measurements, performed at one of the bright pyramidal structures (Fig. 3(d)), evidence that the difference in contrast is based on the diffusion of Ga into Si. In comparison with Fig. 1, it is noticeable that in the centre of the pyramidal structure the Si crystal lattice cannot be seen anymore for this sample. This is presumably due to the clearly increased size of the Ga structure in Si in relation to the surrounding intact Si lattice in the direction of the electron beam. As a preferably very thin area of the electron transparent sample was chosen for investigation by STEM, the sample thickness here is in the range of the pyramidal structure size of about 50 nm at maximum. Between the very bright centre of the Ga containing structure and the pure Si lattice an area with a reduced contrast to the Si lattice can be seen. This is due to the pyramidal geometry, which causes the Ga containing area penetrated by the electron beam to be larger in the centre in comparison to the boundary area. Additionally, the Si crystal structure can clearly be detected in this area of reduced contrast. However, the HAADF image reveals stacking faults here, as can be seen in the inset in Fig. 3(c). Those are presumably caused by crystalline Si dissolving in the liquid Ga at growth temperature and recrystallizing at the boundaries of the Ga containing structure during the cooling down after growth due to the decreasing solubility of Si in Ga with decreasing temperature. EDX measurements demonstrate that the structures are covered by a mound containing significant amounts of oxygen. Since no oxygen is present during the deposition, this mound has to form after the growth as the samples are handled in air with no further precaution then, which supposedly leads to a rapid oxidation of the metallic Ga.

3.4. Annealing

Several samples were annealed after the growth process and a complete cool down to study the influence of high temperatures on the grown structures. The temperature for each annealing was chosen in the range between 400 °C and 800 °C, a temperature regime that could be suitable for the synthesis of various crystal structures. An annealing time of 5 min was chosen for all runs. After annealing at 800 °C no Ga was detected anymore at the surface of any of the samples. For the sample grown with TMGa, the AFM image of the as grown sample is shown in Fig. 3(a); voids were left in the Si surface after annealing at 800 °C, which are displayed by AFM imaging (Fig. 4(a)). The AFM image of the sample deposited with TEGa as precursor, which is presented as grown in Fig. 1(a), displays mounds with a large density even after annealing at 800 °C (Fig. 4(b)). For this sample, only the size and distribution of mounds on the Si surface was altered by the annealing. However, the STEM measurement of the sample after annealing (Fig. 4(c)) reveals no bright pyramidal structures in Si
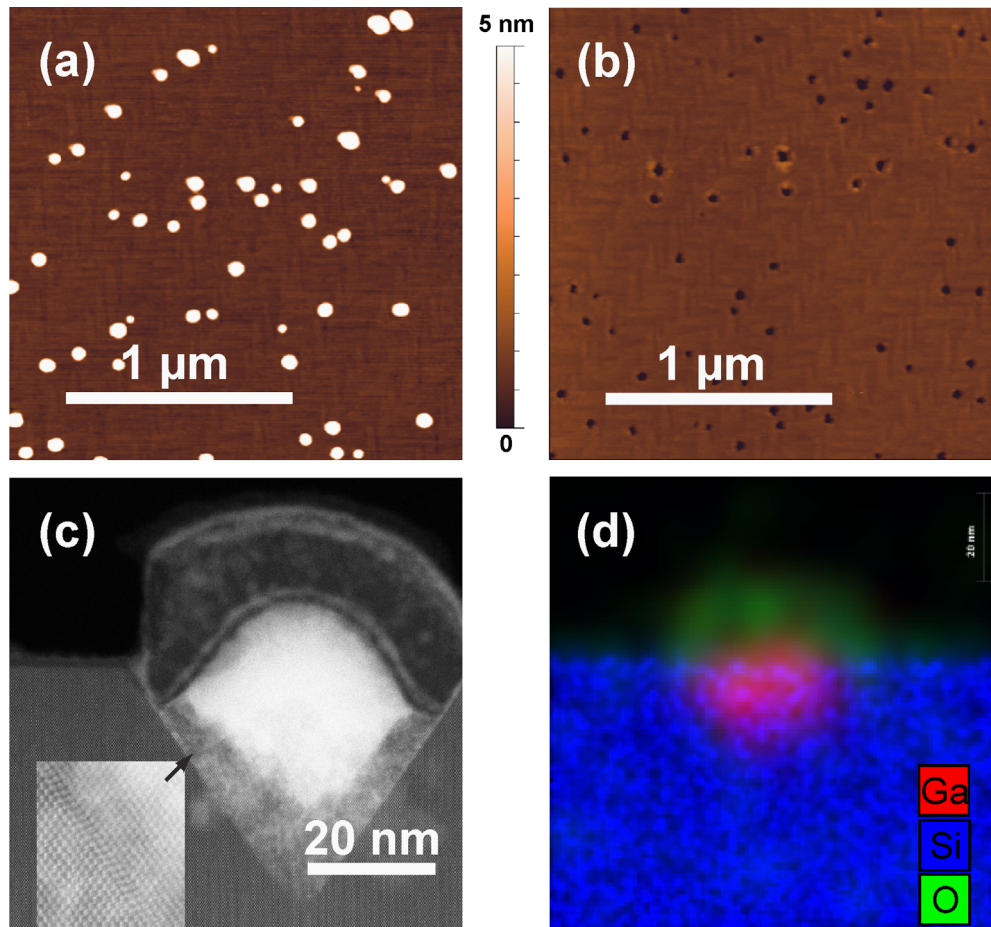
Fig. 3. AFM images of the sample surface after 10 s gallium deposition using TMGa as precursor at a growth temperature of 500 °C after growth (a) and after etching in 30% hydrochloric acid for 30 s (b) showing deposited gallium droplets, and holes remaining in the silicon surface after etching, respectively. Bright Si {1 1 1} limited pyramidal structures were found in the cross-sectional HAADF STEM image of a sample grown under equivalent growth conditions (c). The silicon lattice inside the pyramidal structure exhibits stacking faults, as can be seen in the inset. The EDX map generated at a pyramidal structure in the same sample shows that the bright contrast originates from gallium (represented in red) incorporated in Si (blue) which is covered by an oxygen (green) containing film (d). (For interpretation of the references to color, the reader is referred to the online version of this article.)

anymore. Rather, for this sample similar structures with a darker contrast in comparison to the Si lattice were found. Fig. 4(d) presents the EDX line profiles acquired along the growth direction at one of those dark structures. It evidences that no detectable amount of Ga is left at the surface, but that the dark structures arise from voids occurring in Si. The AFM as well as the STEM measurements indicate that the amorphous carbon containing layer covering the embedded pyramidal Ga structures after growth was barely affected by annealing at those temperatures, though, and remained on the surface covering the voids. There are two possible explanations for the disappearance of Ga from the sample surface. On the one hand it was found that significant amounts of Ga desorb from a Si (0 0 1) surface at temperatures above 900 K [40]. An annealing for 5 min at 800 °C might be sufficient to desorb all deposited Ga from the sample. This mechanism is not consistent with the results on the annealed sample that was deposited with TEGa though, since here the passivating layer is still intact while none of the underlying Ga can be observed anymore. A second possible explanation is that Ga diffuses further into Si due to the chosen temperature, distributing so widely that it cannot be detected anymore by the EDX measurement. Since, judging from these results, we did not expect finding any Ga left at the sample surface after annealing at temperatures higher than 800 °C, no annealing was carried out at temperatures above 800 °C. For an annealing at lower temperatures, i.e. temperatures between 400 °C and 700 °C, mounds remained on the surface for

both precursors used for the Ga deposition. The AFM image of the sample grown with TMGa, which was annealed at 700 °C (Fig. 4 (e)), shows the remaining structures on the Si surface. No noticeable change with respect to the as grown sample (Fig. 3(a)) could be detected by the AFM measurement after the annealing. However, in contrast to the Ga structures without annealing (the AFM image is shown in Fig. 3(a) as grown and in Fig. 3(b) after selective etching) for this sample treatment the selective etching did not have a significant effect on the structures. This can be seen in the AFM image of the same sample after etching in HCl for 30 s (Fig. 4(d)). Although a few voids formed in the surface where former Ga droplets could be etched off, most of the structures detected after the annealing remain unchanged by the selective etching. The same was seen for an annealing at lower temperatures. This alteration of the properties of the deposited structures due to annealing can only be explained by dissolving of Si in the liquid Ga at suitable temperatures during the annealing. As a result of cooling an oversaturated Ga–Si alloy may form, which may lead to Si reaching the surface of the Ga droplets and forming a passivating film, preventing those structures from being selectively etched off by HCl. No clear statement can be given here for the samples grown with TEGa, due to the fact that the Ga structures were already covered with a protective film after deposition. The influence of an annealing at temperatures considerably lower than 800 °C on those structures can therefore not be analyzed.
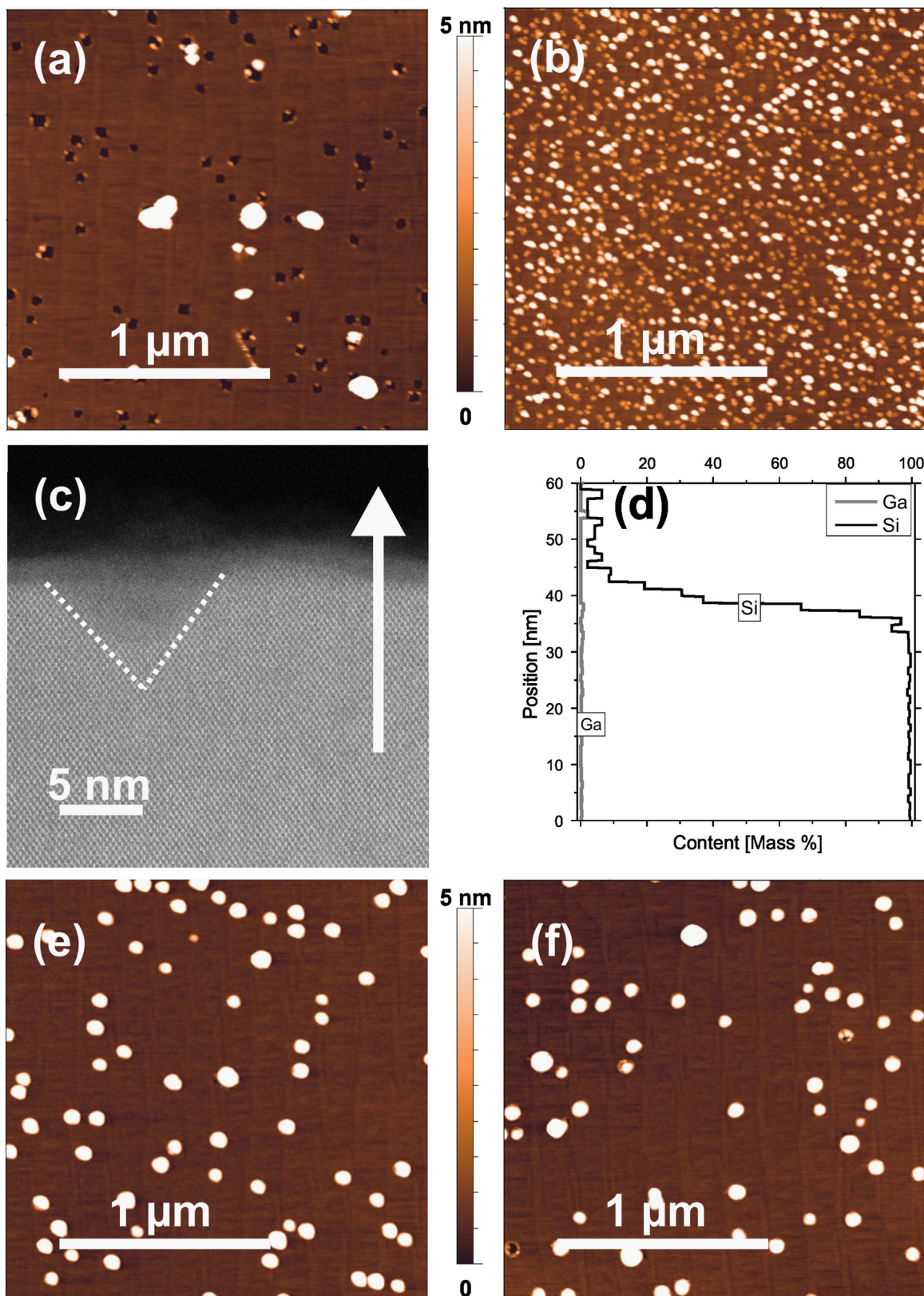
Fig. 4. AFM images taken after annealing at 800 °C of samples grown with TMGa (a) and TEGa (b). Cross-sectional HAADF STEM image of the TEGa deposited sample after annealing at 800 °C (c), as shown in the AFM image (b), revealing a structure with a dark contrast in Si. The EDX measurement clarifies that virtually no Ga is left at the Si surface, as can be seen in the line profile along the growth direction acquired at one of the voids (d). To obtain a sufficient EDX intensity a larger STEM spot size was chosen for the measurement. Additionally, the signal was averaged over the width of the pyramid along the horizontal axis of the image. The AFM image of a sample grown with TMGa after annealing at 700 °C, taken directly after the annealing (e) and after etching for 30 s in 30% hydrochloric acid (f), reveals that after annealing at lower temperatures a Si–Ga alloy has formed, which is more stable to HCl. The annealing was carried out for 5 min in each case.

4. Conclusion

We found that for the deposition of small amounts of TEGa as well as of TMGa on Si, pyramidal Ga containing structures with boundaries on the Si {1 1 1} lattice planes form in Si. The formed etching structures were found in KMC simulation results of the Ga growth, in full agreement with the experimental observations. When TMGa is the Ga precursor the size of the structures formed is clearly increased while the density is decreased with respect to the structures formed when TEGa is the Ga precursor. The structures deposited with TEGa are, in the temperature range examined, covered with a protective layer presumably formed by the residual organic groups, which was not observed for the samples deposited using TMGa as precursor. After annealing at high temperatures, i.e. temperatures above 800 °C, for a sufficient time, no Ga can be detected anymore at the Si surface, while an annealing at lower temperatures leads to the formation of an alloy more stable to hydrochloric acid.

References

[1] S. Liebich, M. Zimprich, A. Beyer, C. Lange, D.J. Franzbach, S. Chatterjee, et al., Laser operation of Ga(NAsP) lattice-matched to (0 0 1) silicon substrate, Appl. Phys. Lett. 99 (2011) 071109. http://dx.doi.org/10.1063/1.3624927.

[2] M. Razeghi, M. Defour, R. Blondeau, F. Omnes, P. Maurel, O. Acher, et al., First cw operation of a $Ga_{0.25}In_{0.75}As_{0.5}P_{0.5}$–InP laser on a silicon substrate, Appl. Phys. Lett. 53 (1988) 2389. http://dx.doi.org/10.1063/1.100239.

[3] J.R. Lang, J. Faucher, S. Tomasulo, K. Nay Yaung, M. Larry Lee, Comparison of GaAsP solar cells on GaP and GaP/Si, Appl. Phys. Lett. 103 (2013) 092102. http://dx.doi.org/10.1063/1.4819456.

[4] T.J. Grassman, M.R. Brenner, M. Gonzalez, A.M. Carlin, R.R. Unocic, R.R. Dehoff, et al., Characterization of metamorphic GaAsP/Si materials and devices for photovoltaic applications, IEEE Trans. Electron Devices 57 (2010) 3361–3369. http://dx.doi.org/10.1109/TED.2010.2082310.

[5] S. Guha, N.A. Bojarczuk, Ultraviolet and violet GaN light emitting diodes on silicon, Appl. Phys. Lett. 72 (1998) 415. http://dx.doi.org/10.1063/1.120775.

[6] C.A. Tran, A. Osinski, R.F. Karlicek, I. Berishev, Growth of InGaN/GaN multiple-quantum-well blue light-emitting diodes on silicon by metalorganic vapor phase epitaxy, Appl. Phys. Lett. 75 (1999) 1494. http://dx.doi.org/10.1063/1.124733.

[7] H. Kroemer, Polar-on-nonpolar epitaxy, J. Cryst. Growth 81 (1987) 193–204. http://dx.doi.org/10.1016/0022-0248(87)90391-5.

[8] A. Beyer, B. Haas, K.I. Gries, K. Werner, M. Luysberg, W. Stolz, et al., Atomic structure of (1 1 0) anti-phase boundaries in GaP on Si(0 0 1), Appl. Phys. Lett. 103 (2013) 032107. http://dx.doi.org/10.1063/1.4815985.

[9] K. Volz, A. Beyer, W. Witte, J. Ohlmann, I. Németh, B. Kunert, et al., GaP-nucleation on exact Si (0 0 1) substrates for III/V device integration, J. Cryst. Growth 315 (2011) 37–47. http://dx.doi.org/10.1016/j.jcrysgro.2010.10.036.

[10] N. Sawaki, T. Hikosaka, N. Koide, S. Tanaka, Y. Honda, M. Yamaguchi, Growth and properties of semi-polar GaN on a patterned silicon substrate, J. Cryst. Growth 311 (2009) 2867–2874. http://dx.doi.org/10.1016/j.jcrysgro.2009.01.032.

[11] D. Zhu, D.J. Wallis, C.J. Humphreys, Prospects of III-nitride optoelectronics grown on Si, Rep. Prog. Phys 76 (2013) 106501. http://dx.doi.org/10.1088/0034-4885/76/10/106501.

[12] B. Zhang, Y. Liu, A review of GaN-based optoelectronic devices on silicon substrate, Chin. Sci. Bull 59 (2014) 1251–1275. http://dx.doi.org/10.1007/s11434-014-0169-x.

[13] Z. Liliental-Weber, M.A. O'Keefe, J. Washburn, Inversion boundaries in GaAs grown on Si, Ultramicroscopy 30 (1989) 20–26. http://dx.doi.org/10.1016/0304-3991(89)90168-X.

[14] W. Stolz, F.E.G. Guimaraes, K. Ploog, Optical and structural properties of GaAs grown on (1 0 0) Si by molecular-beam epitaxy, J. Appl. Phys. 63 (1988) 492. http://dx.doi.org/10.1063/1.340269.

[15] A.A. Baski, Gallium growth and reconstruction on the Si(1 0 0) surface, J. Vac. Sci. Technol., A 8 (1990) 245. http://dx.doi.org/10.1116/1.577076.

[16] H. Sakama, A. Kawazu, T. Sueyoshi, T. Sato, M. Iwatsuki, Scanning tunneling microscopy on Ga/Si(1 0 0), Phys. Rev. B: Condens. Matter 54 (1996) 8756–8760. http://dx.doi.org/10.1103/PhysRevB.54.8756.

[17] B. Bourguignon, K.L. Carleton, S.R. Leone, Surface structures and growth mechanism of Ga ON Si(1 0 0) determined by LEED and Auger electron spectroscopy, Surf. Sci 204 (1988) 455–472. http://dx.doi.org/10.1016/0039-6028(88)90226-9.

[18] A.A. Lyamkina, D.V. Dmitriev, Y.G. Galitsyn, V.G. Kesler, S.P. Moshchenko, A.I. Toropov, The investigation of intermediate stage of template etching with metal droplets by wetting angle analysis on (0 0 1) GaAs surface, Nanoscale Res. Lett. 6 (2010) 42. http://dx.doi.org/10.1007/s11671-010-9790-z.

[19] K. Reyes, P. Smereka, D. Nothern, J. Millunchick, S. Bietti, C. Somaschini, et al., Unified model of droplet epitaxy for compound semiconductor nanostructures: experiments and theory, Phys. Rev. B: Condens. Matter 87 (2013) 165406. http://dx.doi.org/10.1103/PhysRevB.87.165406.

[20] D.-S. Lin, T.-S. Ku, T.-J. Sheu, Thermal reactions of phosphine with Si(1 0 0): a combined photoemission and scanning-tunneling-microscopy study, Surf. Sci 424 (1999) 7–18. http://dx.doi.org/10.1016/S0039-6028(98)00943-1.

[21] E. García-Tabarés, I. García, D. Martín, I. Rey-Stolle, Influence of PH 3 exposure on silicon substrate morphology in the MOVPE growth of III–V on silicon multijunction solar cells, J. Phys. D: Appl. Phys 46 (2013) 445104. http://dx.doi.org/10.1088/0022-3727/46/44/445104.

[22] T. Soga, T. Jimbo, M. Umeno, Epitaxial growth of a two-dimensional structure of GaP on a Si substrate by metalorganic chemical vapor deposition, Appl. Surf. Sci 82-83 (1994) 64–69. http://dx.doi.org/10.1016/0169-4332(94)90196-1.

[23] J.P. André, J. Hallais, C. Schiller, Heteroepitaxial growth of GaP on silicon, J. Cryst. Growth 31 (1975) 147–157. http://dx.doi.org/10.1016/0022-0248(75)90124-4.

[24] H. Ishikawa, K. Yamamoto, T. Egawa, T. Soga, T. Jimbo, M. Umeno, Thermal stability of GaN on (1 1 1) Si substrate, J. Cryst. Growth 189-190 (1998) 178–182. http://dx.doi.org/10.1016/S0022-0248(98)00223-1.

[25] A. Dadgar, M. Poschenrieder, J. Bläsing, O. Contreras, F. Bertram, T. Riemann, et al., MOVPE growth of GaN on Si(1 1 1) substrates, J. Cryst. Growth 248 (2003) 556–562. http://dx.doi.org/10.1016/S0022-0248(02)01894-8.

[26] K. Yamane, T. Kobayashi, Y. Furukawa, H. Okada, H. Yonezu, A. Wakahara, Growth of pit-free GaP on Si by suppression of a surface reaction at an initial growth stage, J. Cryst. Growth 311 (2009) 794–797. http://dx.doi.org/10.1016/j.jcrysgro.2008.09.097.

[27] K. Brixius, A. Beyer, J. Güdde, M. Dürr, K. Volz, U. Höfer, Second-harmonic Generation as Probe for Structural and Electronic Properties of Buried GaP/Si (0 0 1) Interfaces, Manuscript in Preparation, (n.d.).

[28] Y. Cui, C.M. Lieber, Functional nanoscale electronic devices assembled using silicon nanowire building blocks, Science 291 (2001) 851–853. http://dx.doi.org/10.1126/science.291.5505.851.

[29] M.a. Bangar, W. Chen, N.V. Myung, A. Mulchandani, Conducting polymer 1-dimensional nanostructures for FET sensors, Thin Solid Films 519 (2010) 964–973. http://dx.doi.org/10.1016/j.tsf.2010.08.023.

[30] Z.W. Pan, S. Dai, D.B. Beach, N.D. Evans, D.H. Lowndes, Gallium-mediated growth of multiwall carbon nanotubes, Appl. Phys. Lett 82 (2003) 1947. http://dx.doi.org/10.1063/1.1563727.

[31] M.K. Sunkara, S. Sharma, R. Miranda, G. Lian, E.C. Dickey, Bulk synthesis of silicon nanowires using a low-temperature vapor–liquid–solid method, Appl. Phys. Lett. 79 (2001) 1546. http://dx.doi.org/10.1063/1.1401089.

[32] Z.W. Pan, Z.R. Dai, C. Ma, Z.L. Wang, Molten gallium as a catalyst for the large-scale growth of highly aligned silica nanowires, J. Am. Chem. Soc. 124 (2002) 1817–1822. http://dx.doi.org/10.1021/ja017284n.

[33] R.L. Barns, W.C. Ellis, Whisker crystals of gallium arsenide and gallium phosphide grown by the vapor–liquid–solid mechanism, J. Appl. Phys. 36 (1965) 2296. http://dx.doi.org/10.1063/1.1714466.

[34] P. Knoll, Z. Rainer, Method for preparation of thin, oriented GaAs crystals, J. Appl. Phys. 37 (1966) 5006. http://dx.doi.org/10.1063/1.1708191.

[35] A. Ishizaka, Y. Shiraki, Low temperature surface cleaning of silicon and its application to silicon MBE, J. Electrochem. Soc. 133 (1986) 666–671. (http://dx.doi.org/10.1149/1.2108651).

[36] A.B. Bortz, M.H. Kalos, J.L. Lebowitz, A new algorithm for Monte Carlo simulation of Ising spin systems, J. Comput. Phys. 17 (1975) 10–18. http://dx.doi.org/10.1016/0021-9991(75)90060-1.

[37] K. Sato, M. Shikida, Y. Matsushima, T. Yamashiro, K. Asaumi, Y. Iriye, et al., Characterization of orientation-dependent etching properties of single-crystal silicon: effects of KOH concentration, Sens. Actuators, A 64 (1998) 87–93. http://dx.doi.org/10.1016/S0924-4247(97)01658-0.

[38] G.B. Stringfellow, Organometallic Vapour-Phase Epitaxy: Theory and Practice, Academic Press, San Diego, 1999.

[39] R. Lin, The decomposition of triethylgallium on Si(1 0 0), J. Vac. Sci. Technol., B 7 (1989) 725. http://dx.doi.org/10.1116/1.584633.

[40] T.F. Kuech, E. Veuhoff, T.S. Kuan, V. Deline, R. Potemski, The influence of growth chemistry on the MOVPE growth of GaAs and $Al_xGa_{1-x}As$ layers and heterostructures, J. Cryst. Growth 77 (1986) 257–271. http://dx.doi.org/10.1016/0022-0248(86)90310-6.

[41] P.W. Lee, T.R. Omstead, D.R. McKenna, K.F. Jensen, In situ mass spectroscopy and thermogravimetric studies of GaAs MOCVD gas phase and surface reactions, J. Cryst. Growth 85 (1987) 165–174. http://dx.doi.org/10.1016/0022-0248(87)90218-1.

Wissenschaftlicher Werdegang

Name:	Jan Oliver Oelerich
Geburtsdatum:	25.02.1987
Geburtsort:	Köln

1999–2006 Besuch des Gymnasiums im Schloss, Wolfenbüttel
Abitur 06/2006

2007–2010 Bachelor Studium der Physik, Philipps-Universität Marburg
Thesis: *On the validity of the Einstein relation for disordered systems*

2010–2012 Master Studium der Physik, Philipps-Universität Marburg
Thesis: *Theoretical description of charge transport in organic semiconductors*

2012–2015 Doktorand, Fachbereich Physik, Philipps-Universität Marburg
Kollegiat des Graduiertenkollegs *Functionalization of Semiconductors.*
Dissertation: *Theoretical Modeling of Kinetic Phenomena of Atoms and Charge Carriers in Disordered Materials*

/>