TECHNISCHE UNIVERSITÄT WIEN

D I S S E R T A T I O N

Stochastic PDEs for Modeling Transport in Nanoscale Devices

ausgeführt zum Zwecke der Erlangung des akademischen Grades

eines Doktors der technischen wissenschaften unter der Leitung von

Associate Prof. Dipl.-Ing. Dr. techn. Clemens Heitzinger

E101 – Institut für Analysis und Scientific Computing, TU Wien

eingereicht an der Technischen Universität Wien

Fakultät für Mathematik und Geoinformation

von

Leila Taghizadeh M.Sc.

Matrikelnummer: 1228218

Diese Dissertation haben begutachtet:

1. **Prof. Dr. Clemens Heitzinger**
   Institut für Analysis und Scientific Computing, TU Wien

2. **Prof. Dr. Christian Ringhofer**
   Arizona State University

3. **Prof. Dr. Roger Ghanem**
   University of Southern California

Wien, 2019

Kurzfassung

Das Verständnis des Ladungstransports spielt eine wesentliche Rolle bei der Entwicklung vieler elektronischer Geräte. und nanoskalige Geräte wie Sensoren für die Elektroimpedanztomographie (EIT), Nanodrähte, etc. Feldeffekt-(Bio- und Gas-)Sensoren und Nanoporensensoren, z.B. in medizinischen Anwendungen und Nanotechnologie. So ist eine sorgfältige und realistische Modellierung und Analyse des Ladungstransports möglich. in nanoskaligen Geräten sind von großer Bedeutung. In diesem Zusammenhang erweitern wir das Transportmodell, nämlich das Drift-Diffusions-Poisson-System in den Frequenzbereich und analysieren die Existenz von und die lokale Einzigartigkeit ihrer Lösung im Wechselstrom (AC)-Kleinsignalregime, die erst kürzlich experimentell demonstriert wurden. Um das Modell weiter zu verbessern, haben wir die Entwicklung des stochastischen Drift-Diffusions-Poisson-Systems, um die Unsicherheit in der nanoskalige Geräte. Zu diesem Zweck analysieren wir zunächst das stochastische PDE-System, indem wir Folgendes vorstellen Existenz und lokale Einzigartigkeit seiner Lösung, und dann entwickeln optimale stochastische numerische numerische Verfahren wie mehrstufiges Monte-Carlo und mehrstufiges randomisiertes quasi Monte-Carlo. Finite-Element-Methoden zur Modellierung der Zufälligkeit beim Ladungstransport. Tatsächlich ist der Gesamtfehler des vorgestellten stochastischen Methoden, die verschiedene (statistische und Diskretisierungs-)Quellen beinhalten, haben ausgeglichen werden, um die rechnerische Effizienz der Verfahren zu verbessern. Dies führt zu optimale Diskretisierungsparameter und Anzahl der Proben und damit optimale Ergebnisse zu finden. stochastische Methoden. Auch die realitätsnahe Modellierung von medizinischen und elektronischen Geräten wie EIT-Sensoren ist unerlässlich. in diesem Bereich. In dieser Dissertation entwickeln wir ein inverses Modellproblem des EIT in einem unendlichen Bereich. dimensionale Einstellung durch Erweiterung des Standard-Vorwärtsmodells auf ein nichtlineares elliptisches Teilmodell Differentialgleichung. Die Unsicherheit im vorgestellten nichtlinearen EIT-Modell ist auf die Tatsache zurückzuführen, dass die Material- und Einschlusseigenschaften wie Permittivitäten, Chargen und Größen von Einschlüssen in die den Hauptkörper, die in der Medizin, im EIT und in der Bioimpedanztomographie unerlässlich sind, um die den inneren Körper abzuschirmen und Tumore zu erkennen oder die Körperzusammensetzung zu bestimmen. Diese geometrische und physikalische Steuerungsparameter werden gleichzeitig extrahiert, indem sie gelöst werden. das daraus resultierende inverse Problem des EIT mittels einer adaptiven Markov-Kette Monte-Carlo Finite-Elemente-Methode (MCMC-FEM), einschließlich einer MCMC-Probenahmetechnik für die Wahrscheinlichkeitsraum und eine Galerkin-Finite-Elemente-Approximation für die Diskretisierung der physischer Raum. Darüber hinaus formulieren wir das inverse EIT-Modell in einer messentheoretischen Form. Gerüst und beweisen die Klarheit und Souveränität der hinteren Maßnahme und der Bayes'sche Inversion. Die Bayes'sche Inferenz beweist auch ihre Fähigkeit, die statistische Variabilität in der die gemessen Ergebnisse von Wachstum und Abbau von Biofilmen. Zu diesem Zweck stellen wir eine System von PDEs als mathematisches Modell für Biofilme,

das die zeitabhängige Entwicklung der Größe des Biofilms einschließlich Quorum Sensing und Kooperation von Bakterien gegen Antibiotika. Die Ergebnisse der Biofilm inversen Problem beweisen die Fähigkeit der vorgeschlagenen Unsicherheitsquantifizierungsverfahren zur genauen Schätzung relevanter Systemparameter in der Modell.

Abstract

The understanding of charge transport plays an essential role in the design of many electronic and nanoscale devices such as electrical-impedance tomography (EIT) sensors, nanowire field-effect (bio- and gas) sensors and nanopore sensors for instance in medical applications and nanotechnology. Thus, carefully and realistic modeling and analysis of charge transport in nanoscale devices are of great importance. In this regard, we extend the transport model, namely the drift-diffusion-Poisson system to the frequency domain and analyze the existence and local uniqueness of its solution in the alternating-current (AC) small-signal regime, which were only demonstrated experimentally recently. To further improve the model, we develop the stochastic drift-diffusion-Poisson system in order to model uncertainty in the nanoscale devices. To this end, we first analyze the stochastic PDE system by presenting existence and local uniqueness of its solution, and then develop optimal stochastic numerical methods such as multilevel Monte-Carlo and multilevel randomized quasi Monte-Carlo finite-element methods to model randomness in charge transport. In fact the total errors of the presented stochastic methods including different (statistical and discretization) sources have to be balanced in order to improve the computational efficiency of the methods. This leads to finding the optimal discretization parameters and number of samples and consequently optimal stochastic methods. Realistic modeling of medical and electronic devices such as EIT sensors is also essential in this field. In this dissertation, we develop an EIT inverse model problem in an infinite-dimensional setting by extending the standard forward model to a nonlinear elliptic partial differential equation. The uncertainty in the presented nonlinear EIT model is due to the material and inclusion properties such as permittivities, charges and sizes of inclusions in the main body, which are essential in medicine, EIT and bioimpedance tomography to screen the interior body and to detect tumors or to determine body composition. These geometrical and physical governing parameters are extracted simultaneously by solving the resulting EIT inverse problem by means of an adaptive Markov-chain Monte-Carlo finite-element method (MCMC-FEM), including an MCMC sampling technique for the probability space and a Galerkin finite-element approximation for the discretization of the physical space. Furthermore, we formulate the EIT inverse model in a measure-theoretic framework and prove well-definedness and well-posedness of the posterior measure and the Bayesian inversion. The Bayesian inference also proves its ability to interpret the statistical variability in the measured outputs of biofilms growth and degradation. To this end, we present a system of PDEs as a mathematical model for biofilms, which describes the time dependent evolution of the size of the biofilm including quorum sensing and cooperation of bacteria against antibiotics. The results of biofilm inverse problem prove the ability of the proposed uncertainty quantification method to accurately estimate relevant system parameter in the model.

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor, Prof. Clemens Heitzinger, for his guidance, patience, knowledge, and invaluable assistance and of course for giving me the opportunity to attend many scientific conferences and many other opportunities in my scientific life that have been fundamental for my growth both as an individual and as a mathematician in the past years. I am, and always will be proud of working in his research group.

I would like to thank Prof. Christian Ringhofer and Prof. Roger Ghanem for their reports on this dissertation and their interest in my work.

I wish to thank all the colleagues, friends and members of our research group for the great work atmosphere and their kind help during my PhD studies: Jose Morales Escalante, Boaz Blankrot, Amirreza Khodadadian, Benjamin Stadlbauer, Gregor Mitscha-Baude, Gerhard Tulzer, Stefan Rigger, Gudmund Pammer and Daniel Pasterk.

Deepest gratitude is also due to my family. I would like to thank my parents and my sister Dr. Roya Taghizadeh for their constant support and encouragement throughout my studies.

Finally, I wish to thank my beloved husband Ahmad Karimi for his endless encouragement, understanding, and patience, even during hard times. This dissertation is dedicated to him and our son Ryan.

Vienna, 2019 *Leila Taghizadeh*

FWF

Der Wissenschaftsfonds.

Contents

1. Introduction

This research is focused on different aspects of nonlinear partial differential equations (PDEs), which model charge transport in devices such as nanowire, nanopore and tomography sensors with wide range of applications in industry, medicine and life sciences. These aspects include analysis, numerical solution to the PDE models as well as PDE inverse problems. The main PDE models include Poisson-Boltzmann equation and the drift-diffusion-Poisson (DDP) system in deterministic and stochastic regimes.

The primary interest here in the context of analysis is in studying existence and uniqueness of the solutions to these elliptic PDE models. Particularly, analysis of the models in the alternating-current regime which is a more realistic platform for such problems is studied.

A question which then arises is on the numerical analysis and the way that the uncertainty in the models are numerically dealt with. The stochastic numerical methods such as multilevel Monte-Carlo and randomized quasi-Monte-Carlo methods are sampling-based techniques which are successfully used to solve the stochastic elliptic PDE models numerically. These techniques perform even more efficiently in encountering the uncertainties in the target models if they are optimized. The optimal Monte-Carlo methods perform similarly to the vanilla ones but with less computational cost. Reducing the cost of calculations is of great importance as in such problems the numerical method is iterated for a (large) number of samples. The applications include for example modeling charge transport in random environments, such as nanoscale devices.

The third aspect of the elliptic PDE models which this research is focused is PDE inverse problems. There are unknown parameters in the models, which cannot be determined neither from the experiments nor from the computational models. Bayesian estimation techniques are strong uncertainty quantification tools which are capable of dealing with nonlinear and ill-posed inverse problems in order to recover the quantities of interest. To this end, Markov-chain Monte-Carlo (MCMC) algorithms are used to find confidence interval of the unknown parameters by means of estimating their posterior probability density. These algorithms are combination of Monte Carlo as a sampling method and Markov chains which search the parameter space efficiently. The applications include electrical-impedance tomography (EIT) and modeling biofilms in medicine.

The Bayesian analysis in the measure-theoretic framework is also studied for the Poisson-Boltzmann equation, which models the electrical-impedance tomography (EIT) sensors with applications in medicine. This analysis focuses on the well-definedness and well-posedness of the posterior measure and the Bayesian estimation for the EIT inverse problem.

In the following sections, we introduce aforementioned models, methods and applications of this research in detail.

1.1. Drift-diffusion-Poisson System in the Alternating-current Regime

In this work, the mathematical modeling of affinity based field-effect sensors is considered in the alternating-current (AC) regime. While nanowire field-effect bio- and gas sensors operating in the direct-current (DC) regime have been established experimentally in the past decade, their use in the AC regime is more recent and has been shown to yield more physical insight and additional sensing information. Sensor based on nanopores and the principle of the Coulter counter are governed by the model equations in this work as well.

The working principle of nanowire field-effect sensors [1–3] is that the target molecules to be detected in a liquid change the charge concentration at the sensor surface after selectively binding to receptor molecules, which in turn modulates the conductance of the semiconducting nanowire. The currents through the nanowires are recorded and indicate the number of target molecules present. A schematic diagram is shown in Figure 1.1. The main advantage of this type of affinity based sensor compared to currently employed technology is its label-free operation: No fluorescent or radioactive markers are required. Further advantages are high sensitivity, real-time operation, and high selectivity. The concept is a very general one, since DNA and RNA oligomers as well as antigens with known antibodies can be detected. The functioning of nanowire gas sensors [4–8] is analogous.
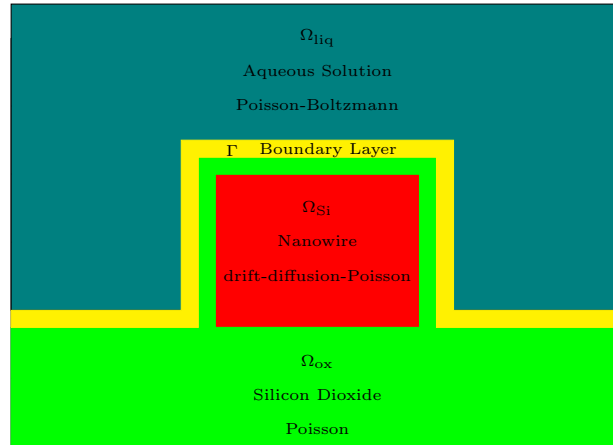
Figure 1.1.: Schematic diagram of a transversal cross section of a nanowire field-effect sensor.

The mathematical modeling of affinity based field-effect sensors and semiconductor devices in general has been addressed previously only in the DC regime [9–12]. Furthermore, in [13], a multiscale problem inherent in nanowire sensors was solved and lead to the interface conditions used below. In [14], an effective equation for the covariance was derived after homogenization of a random charge distribution at a sensor surface. In [15], existence and uniqueness for the drift-diffusion-Poisson system with interface conditions were shown for the DC regime. A parallel numerical method was developed in [16]. These mathematical results have then been used to provide quantitative understanding and to optimize sensor design [17, 18].

More recently, such nanowire field-effect sensors have been fabricated for use in the AC

regime and characterized [19–23]. In the experiments, the electric potentials around the DC equilibrium are small and the frequencies are low enough to ensure that the free charge carriers in the liquid are equilibrated, avoiding spurious signals.

Nanopore sensors are also described by the same transport equations. Here the particles that move in a self-consistent manner are anions, cations, and target molecules. The principle of a nanopore sensor is similar to the Coulter counter. A schematic diagram is shown in Figure 1.2.
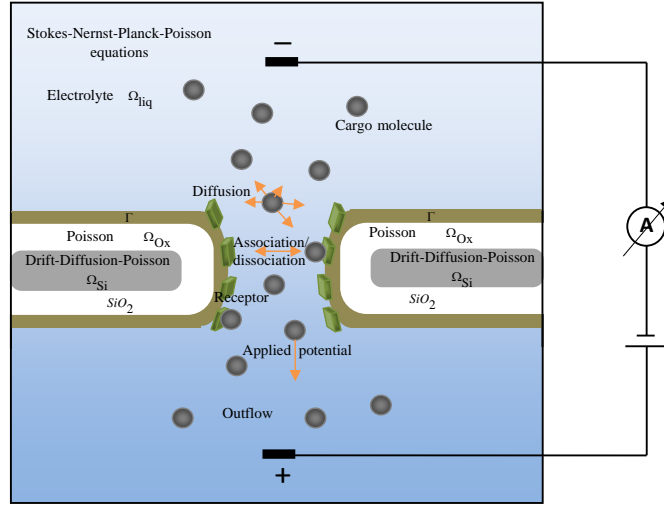
Figure 1.2.: Schematic diagram of a nanopore sensor.

1.2. Multilevel Monte-Carlo Methods for Stochastic Drift-diffusion-Poisson System

In this work, we consider the system consisting of the drift-diffusion-Poisson equations coupled with the Poisson-Boltzmann equation, all with random coefficients. We show existence and local uniqueness of weak solutions for the stationary problem. This system is a general model for transport processes, where a stochastic process determines the coefficients. Furthermore, we develop a multi-level (ML) Monte-Carlo (MC) finite-element method (FEM) for the system of equations. The different types of errors in the numerical approximation must be balanced and the optimal approach is found here.

In the system of equations considered here, both the operators and the forcing terms are stochastic, and therefore this system has numerous applications (see Figure 1.3). A deterministic and simplified version, without the Poisson-Boltzmann equation, is the standard model for semiconductor devices. Nowadays, randomness due to the location of impurity atoms is the most important effect limiting the design of integrated circuits. This application area is included in the present model equations. Furthermore, the full system of equations

considered here describes a very general class of field-effect sensors including their most recent incarnation, nanowire bio- and gas sensors. While previous mathematical modeling has focused on the deterministic problem and stochastic surface reactions [8, 13–16, 18], the present model describes how various stochastic processes propagate through a PDE model and result in noise and fluctuations in a transport model. Quantifying noise and fluctuations in sensors is important, since they determine the detection limit and the signal-to-noise ratio. Noise and fluctuations are of great importance especially in nanometer-scale devices, as any random effect becomes proportionally more important as devices are shrunk.

Various sources of noise and fluctuations are included in the model equations here. Doping of semiconductor devices is inherently random and results in a random number of impurity atoms placed at random positions, each one changing the charge concentration and the mobility at its location. In field-effect sensors, target molecules bind to randomly placed probe molecules in a stochastic process, so that the detection mechanism is inherently stochastic. The Brownian motion of the target molecules also results in changes in charge concentration and permittivity. This randomness at the sensor surface propagates through the self-consistent transport equations and finally results in noise in the sensor output.

In summary, there are many applications where both the operators and the forcing terms in the drift-diffusion-Poisson system are random. The probability distributions of permittivities and charge concentrations can be calculated from physical models [24].

Figure 1.3.: Schematic diagram of a nanowire field-effect sensor (top) and a longitudinal cross section (bottom) used for the finite-element discretization, displaying the different subdomains as well as the source and drain contacts.

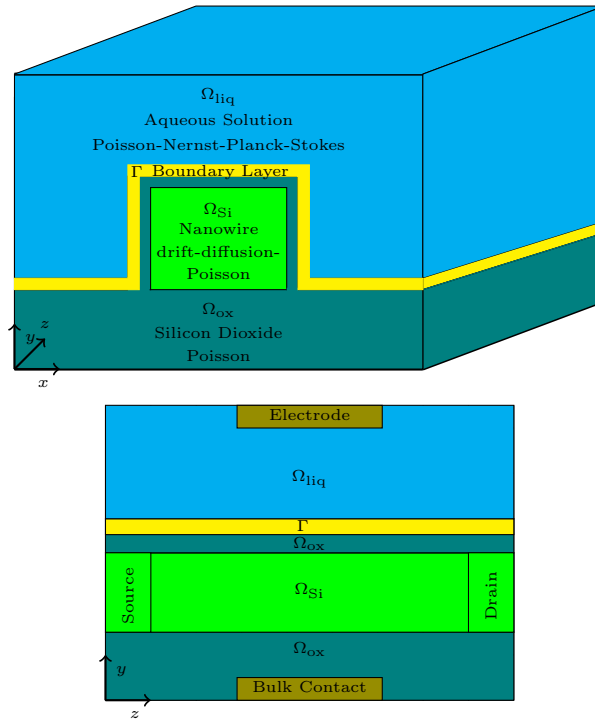In many realistic situations, the probability space is high-dimensional. For example, each probe molecule, each target molecule, and each probe-target complex needs to be modeled in sensors. In transistors, the number impurities and their positions are random. The large number of dimensions favors the use of Monte-Carlo (MC) methods: It is well-known that the convergence rate of standard MC methods is independent of the number of dimensions. On the other hand, it is inversely proportional to the square root of the number of evaluations and here each evaluation requires solving a two- or three-dimensional system of elliptic equations.

These considerations motivate the development of a multi-level Monte-Carlo (MLMC) algorithm. In [25], after earlier work [26] on numerical quadrature, it was shown that a multi-level approach and a geometric sequence of timesteps can reduce the order of computational complexity of MC path simulations for estimating the expected value of the solution of a stochastic ordinary differential equation. This is done by reducing the variance and leaving the bias unchanged due to the Euler discretization used as the ODE solver. In [27], the Milstein scheme was used as the ODE solver to improve the convergence rate of the MLMC method for scalar stochastic ordinary differential equations and the method was made more efficient. The new method has the same weak order of convergence, but an improved first-order strong convergence, and it is the strong order of convergence which is central to the efficiency of MLMC methods. In [28], the MLMC method was combined with quasi-Monte-Carlo (QMC) integration using a randomized rank-1 lattice rule and the asymptotic order of convergence of MLMC was improved and a lower computational cost was achieved as well.

In [29], an MLMC finite-element method was presented for elliptic partial differential equations with stochastic coefficients. In this problem, the source of randomness lies in the coefficients inside the operator and the coefficient fields are bounded uniformly from above and away from zero. The MLMC error and work estimates were given for the expected values of the solutions and for higher moments. Also, in [30], the same problem was considered and numerical results indicate that the MLMC estimator is not limited to smooth problems. In [31], a multi-level quasi-Monte-Carlo finite-element method for a class of elliptic PDEs with random coefficients was presented. The error analysis of QMC was generalized to a multi-level scheme with the number of QMC points dependent on the discretization level and with a level-dependent dimension truncation strategy.

In [32], uniform bounds on the finite-element error were shown in standard Bochner spaces. These new bounds can be used to perform a rigorous analysis of the MLMC method for elliptic problems, and a rigorous bound on the MLMC complexity in a more general case was found. In [33], the finite-element error analysis was extended for the same type of equations posed on non-smooth domains and with discontinuities in the coefficient. In [34], a general optimization of the parameters in the MLMC discretization hierarchy based on uniform discretization methods with general approximation orders and computational costs was developed. In current work, we define a global optimization problem which minimizes the computational complexity such that the error bound is less or equal to a given tolerance level.

1.3. Multilevel Randomized Quasi Monte-Carlo Methods for Stochastic Drift-diffusion-Poisson System

Calculating the expected value of the solution of the stochastic drift-diffusion-Poisson system poses a computational challenge due to the large number of stochastic dimensions in realistic applications. In order to speed up the convergence of the standard Monte-Carlo method, variance-reduction methods such as the multilevel Monte-Carlo method have been developed [25–30, 35] and have also been applied to the stochastic drift-diffusion-Poisson system [36]. In [36], the parameters of the numerical approach were also optimized such that the total computational work is minimized, while an estimate of the total error is kept below a prescribed tolerance.

The idea, developed into an optimal numerical method here, is to improve the choice of samples or evaluation points, meaning that random sequences are replaced by quasi-random sequences with better uniformity. Quasi-Monte-Carlo methods with low discrepancy lead to faster convergence than the standard Monte-Carlo method, while – on the other hand – the disadvantage of quasi-Monte-Carlo methods is that the low-discrepancy sequences used are deterministic. Hence, a quasi-Monte-Carlo method is a deterministic algorithm with error bounds that are difficult to estimate.

To overcome this problem, one can randomize the deterministic sequences by using a random shift, e.g., a uniformly distributed vector. The idea of random shifting was first introduced in [37] in the context of good lattice rules and later applied to the idea of general lattice rules in [38]. Later Tuffin [39] considered random shifting of any low-discrepancy sequence and studied the discrepancy of the shifted sequence. If a randomized low-discrepancy sequence such as a randomly shifted lattice rule is used, a new method called a randomized quasi-Monte-Carlo (RQMC) method results. Using the idea of stratification, we can improve the single-level RQMC method to multilevel randomized quasi-Monte-Carlo (MLRQMC) method.

The MLRQMC method was first introduced in [28] by combining the multilevel Monte-Carlo method with quasi-Monte-Carlo integration using a randomized rank-1 lattice rule. In [40], the variance estimation of the method and its convergence rate were investigated theoretically and numerically. In [35], the method was applied to elliptic partial differential equations (PDE) with random coefficients and a finite-element discretization was used. The main goal of the current work is to develop an optimal MLRQMC method for solving a system of stochastic PDE, namely the stochastic drift-diffusion-Poisson (DDP) system. Here, we analyze the convergence of the numerical method considering the discretization and statistical errors as the main sources of error. The system of PDE has many applications including all situations where charge transport occurs in a random environment. We calculate the computational cost of the MLRQMC approach applied to this system of equations. The function modeling the computational work is minimized such that the estimated total error of the procedure is less than or equal to a prescribed error tolerance. By solving this optimization problem, optimal values for parameters such as the mesh sizes in the spatial discretization and the optimal number of quasi-points are obtained in a natural manner.

1.4. Electrical Impedance Tomography

Tomography is one of the most important techniques in imaging and could be used in medical monitoring such as monitoring for internal bleeding, screening for breast cancer, detection of pulmonary emboli and blood clots in lungs, provided that the corresponding inverse problems can be reliably solved. Imaging the internal organs to diagnose diseases is one of the most important aspects of modern medicine. Many medical problems could be easily diagnosed by information about the distribution of electrical properties inside the body. Although tomography has originally been used solely in medical imaging, in its general sense it has now become much more diverse and is used in a wide range of fields including industrial and geophysical applications. It is used in industrial-process tomography to control industrial processes such as curing and cooking. Geophysical surveying is another application of the technique which is used for determining the location of mineral and oil deposits, leakage detection in pipes, etc. Furthermore, atmospheric and forensic imaging, archeology, and land-mine detection are other fields where tomography imaging is used.

Since modern medicine relies on imaging methods, the mathematics of tomography have become one of the most important applications of mathematics in the areas of healthcare, medicine, and life sciences in general. In soft-field tomographic techniques, a sensing field is applied to an object and the responses to this field are measured. Processing of these responses allows reconstruction of the distributions of physical properties inside the object, if the nonlinear inverse problem can be solved. For instance, electrical-impedance tomography (EIT) [41] in medicine, electrical-resistance tomography (ERT) [42] in geophysical applications, and electrical-capacitance tomography (ECT) [43] in industrial process monitoring are soft-field tomography techniques. The nature of soft-field tomography techniques is much more complex than the nature of hard-field ones and requires considerably more computational analysis and algorithms to reconstruct the image.

In medical applications, computer tomography (CT) and EIT are the two main methods used in imaging and reconstruction. The CT reconstruction problem is a linear problem since X-rays propagate in straight lines through the object and their absorption at any point inside the object is independent of the absorption at any other point. Therefore, the attenuation is measured along each collimated beam direction, which is a linear problem and leads to sparse and well-conditioned sensitivity matrices (Radon transforms). In this method, the device is usually very complicated and provides very good resolution. On the other hand, EIT reconstruction is a nonlinear and ill-posed inversion problem. It is nonlinear since as the field is modified inside the body, the potential measured at the boundary of the domain is a nonlinear function of the the distribution of the electrical properties through the body. Furthermore, this problem is ill-posed, since large changes in the interior can correspond to very small changes in the measured data [44]. The resulting resolution of EIT is lower than the one of CT, but practical advantages such as simple, cheap, portable, and radiation free devices have put this method into the center of attention in recent years.

Tomography is defined as measuring the propagation of energy or particle motion in order to reconstruct information about the interior of a system. Usually this information is a reconstruction of physical and geometrical properties of the system such as permittivity, charge, and size. Hence parameter reconstruction in soft-field tomography is based on an inverse problem, where a forward model is fitted to the data. Electrical-impedance

tomography is an imaging technique for detecting (and imaging) internal properties such as the permittivity distribution inside an object by means of measuring the electrical properties at exterior electrodes. A set of contact electrodes is attached on the surface of the body and prescribed electric potentials are applied to the body through the electrodes. The corresponding electrical currents needed to maintain these potentials are measured on these electrodes. This process may be repeated with various potentials. The aim is to estimate the permittivity and charge distributions in the body from these data [45].
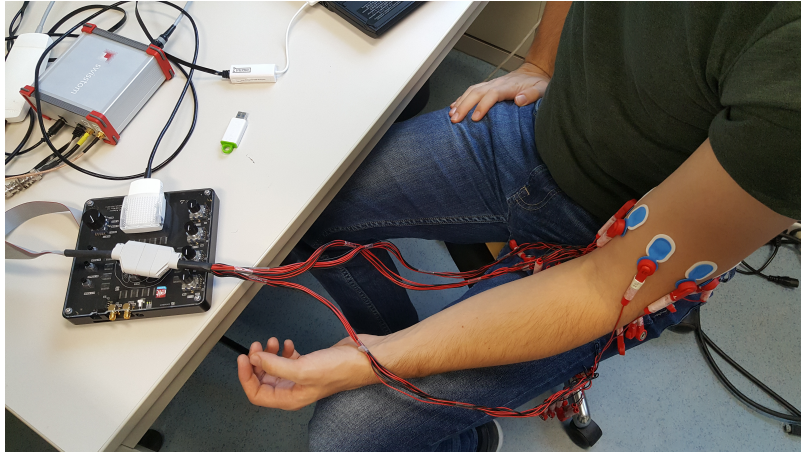
Figure 1.4.: The EIT device. Photo by Florian Thürk (TU Vienna).

The idea behind the EIT technology is not new as the idea of electrical-resistive tomography was proposed in 1978, independently by Henderson and Webster for medical imaging, and by Lytle and Dines for geophysical imaging [46]. Since that time, computational power has increased and more efficient inverse algorithms have been proposed. Hence, due to this fact and as it includes crucial applications, looking at the idea again and applying modern and powerful inverse algorithms such as Bayesian-estimation methods to analyze the results is of great importance.

The EIT poses a nonlinear and ill-posed inverse problem, which can be solved by means of either classical or so-called Bayesian-inversion methods. There are fundamental differences between classical and Bayesian inverse modeling. The first difference is in the nature of the solution; the solution of Bayesian estimation is a probability distribution of the model parameter, which is a random variable. Hence a confidence interval for each of the quantities of interest can be found, which is crucial in applications such as in medicine. The other difference is using prior knowledge in the Bayesian approach to update the current information about the parameters of interest. The connection between the two level of knowledge is made by means of Bayes' Theorem. In Bayesian estimation techniques, Markov-chain Monte-Carlo (MCMC) methods are used to deal with inverse problems in order to circumvent the need to calculate high-dimensional integrals (appearing in Bayes' Theorem). Moreover they sample the space efficiently and completely search the space. Bayesian approaches to the EIT inverse problem have previously been studied, for example in [47–50]. Bayesian inference for inverse problems, especially in infinite dimensions is a new approach, which has been applied to linear (e.g. [51]) and recently nonlinear problems

(e.g. [52,53]).

In this work, we focus on electrical-impedance tomography as a soft-field tomography technique and one of the main imaging techniques applied in medicine. We present and solve an elliptic partial differential equation, namely a nonlinear Poisson-Boltzmann equation as the EIT model to find the electrical currents flowing in and out of the electrodes in a bioimpedance tomography device. This is the first time that the nonlinear model is used for the EIT problem in order to take the free charges present in the problem into account and gives the related Bayesian formulation in an infinite-dimensional setting in order to solve the corresponding inverse problem. We solve the EIT forward problem by means of the first order Galerkin finite-element (GFE) approximation and a mesh generated by the GMSH package [54], and the corresponding inverse problem is solved using the delayed-rejection adaptive-Metropolis (DRAM) algorithm, which we have implemented in Julia [55]. This algorithm is one of the most efficient MCMC algorithms, as the proposal scale (covariance matrix) is updated to an optimal one automatically and periodically in order to achieve better convergence. The simulation results of the forward model are used as mathematical solutions in the DRAM algorithm for statistical analysis and to estimate posterior distributions for each of the parameters to be extracted. Numerical results using the state of the art GFE and the adaptive MCMC illustrate accurate extraction of the quantities of interest in the EIT inverse problem. In fact, confidence intervals for the parameters can be calculated, which is essential, for example, in medical applications. We also discuss Bayesian inversion for the presented nonlinear elliptic PDE model for EIT in the measure-theoretic framework and prove that the posterior measure is Lipschitz continuous in the data to conclude well-definedness and well-posedness of the resulting posterior measures obtained by the Bayesian technique.

1.5. Biofilms

Infection of material and devices implanted into patients' tissues and bones is associated with considerable morbidity and costs [56–58]. The use of all kinds of implants, e.g., osteosyntheses, joint prosthesis, cardiac valves and devices, percutaneous intravascular catheters, invasive monitoring to sustain life at intensive care units, and other implants is increasing. Dependent on the site of implantation, the infection rates range from 0.2% to 5% in orthopaedic and trauma surgery and up to 40% in artificial hearts [59]. Given the high incidence of fracture stabilization devices of two million per year, the number of implant infections amounts to up to 100 000 per year [58, 60]. The major pathogens of implant related infections are Staphylococcus aureus and coagulase negative staphylococci, primarily Staphylococcus epidermidis [61], followed by enterobacteria, Pseudomonas aeruginosa, and enterococci. These organisms have in common that they are difficult to eradicate by standard antibiotic therapy due to their intrinsic resistance and exposure to antimicrobials.

In implant surfaces, these organisms grow in biofilms and thus cause persistent or recurrent infections [62]. The simple definition of a biofilm is microorganisms attached to a surface. A more comprehensive definition is that a biofilm consists of a structured community of bacterial cells enclosed in a self-produced polymeric matrix and adherent to a surface. Biofilms are highly individual based on the characteristics of the microbe,

environmental conditions, nutrients, implant surfaces, and host immune reaction in case of implant infections (see Fig 1.5).
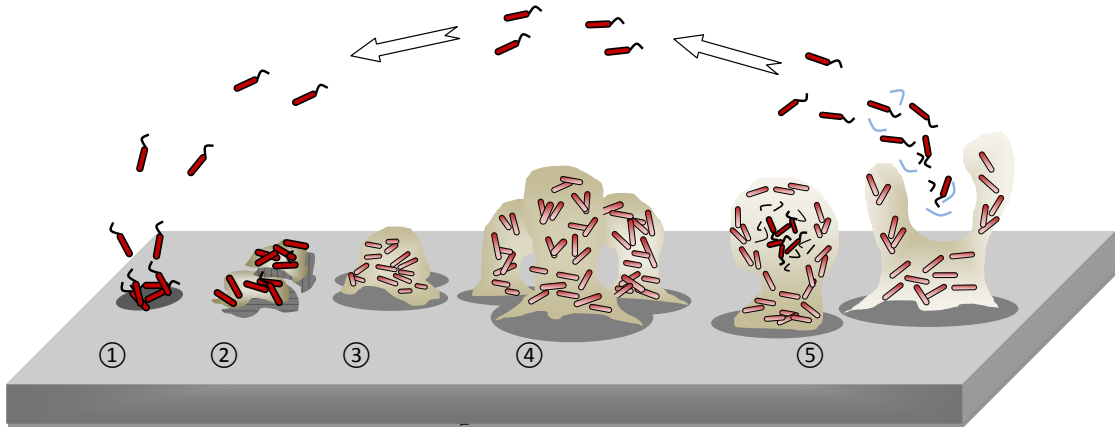
Figure 1.5.: The life cycle of bacteria in biofilms. Phase 1: initial attachment, phase 2: permanent attachment, phase 3: primary maturation, phase 4: secondary maturation, and phase 5: dispersion.

Biofilm associated infections are frequently resistant to conventional antimicrobial therapy, because the bacterial biofilm on the surface serves as a reservoir where bacteria are quasi inaccessible to antibiotics and the host defenses [63, 64]. In the clinical routine, antibiotic susceptibility is tested by determining the minimal inhibitory concentration (MIC) of the antibiotic on free floating bacteria in the growth phase. A low concentration of the MIC indicates the susceptibility of the microorganism and it is a rough approximation of the efficacy of the treatment.

There are numerous models of medical biofilms. All models as well as all visualization methods like staining or preparation for electron microscopy or confocal laser scanning microscopy (CLSM) have limitations. However, initial steps into modeling include simple models like static biofilms on microtiter plates or coverslips or dynamic biofilms within flow-cells. CLSM is a good tool to visualize metabolic and/or structural changes within biofilms and a reliable starting point for more complex biofilm models [65].

In previous publications, we demonstrated that clinically achievable antibiotic concentrations do not reduce biofilms and bacterial growth. Increasing the antibiotic concentration may only reduce the biofilm thickness and reduce bacterial growth [66, 67]. Moreover, we could demonstrate that changing environmental factors, such as increasing the environmental temperature or changing the composition of the implant material resulted in the reduction of biofilm mass and bacterial load [68, 69].

Antibiotic therapy of these chronic biofilm associated infections is generally effective only when the infection is acute and the implanted material is removed. When it is necessary to keep the implanted material, the success of therapy is mediocre. Even the exchange of the infected implant material and the necessity to implant a new prosthesis is associated with higher recurrence rates. International recommendations for the treatment

of implant infections propose the use of quinolones, high-dose beta-lactams, or daptomycin in combination with rifampicin [70]. However, antimicrobial resistance and intolerability is common in chronic infection. Thus, antibiotic therapy of these very complicated infections has to be tailored to the individual patient and the characteristics of the pathogen.

1.6. Remarks

This research focuses on different aspects of nonlinear partial differential equations (PDEs), which model charge transport in devices such as nanowire, nanopore and tomography sensors with wide range of applications in industry, medicine and life sciences. These aspects include analysis, numerical solution to the PDE models as well as PDE inverse problems. The common goal is uncertainty quantification via two different directions, namely, stochastic numerical methods for PDEs and PDE inverse problems.

The dissertation has two parts: the first part is devoted to PDE models and methods in nanotechnology, which includes modeling and analyzing the nanoscale devices and stochastic numerical methods. In the second part, the focus is on the PDE inverse modeling and the Bayesian analysis techniques in order to uncertainty quantification. More specifically, the dissertation after giving introductions in Chapter 1, is organized as follows.

- Chapter 2: In this chapter, the drift-diffusion-Poisson-Boltzmann system in the alternating-current (AC) regime is analyzed. The new modification is based on extending the transport model to the frequency domain and writing the variables as periodic functions of the frequency in a small-signal approximation. In this work, first the AC model equations are derived and then existence and local uniqueness of the solution of the model equations are shown. Real-world applications include nanowire field-effect bio- and gas sensors operating in the AC regime, which were only demonstrated experimentally recently. Furthermore, nanopore sensors are governed by the system of model equations and the analysis as well.

- Chapter 3: In this chapter, an optimal multi-level Monte-Carlo finite-element (MLMC-FE) method for the numerical approximation of the expected value of the solution of the stochastic DDP system is developed. The rate of convergence and computational complexity of the presented method is analyzed and an optimal choice of discretization and sampling parameters are found. These optimal choices allow to reduce the computation cost, which is so important in sampling-based methods. The presented optimal methods are successfully used to solve the stochastic DDP system and in the study of the effect of random dopants in nanoscale semiconductor devices, namely nanowire field-effect sensors. Furthermore efficiency of the optimal vanilla Monte-Carlo and multilevel Monte-Carlo in this application are compared. Moreover in this chapter, another optimal sampling method namely multilevel randomized quasi-Monte-Carlo finite-element (MLRQMC-FE) method is presented. In this method, quasi points, i.e., randomly shifted low-discrepancy sequences are used to estimate the expected value of the solution of the stochastic DDP system and to calculate random-dopant effects in nanoscale field-effect transistors. Using these points improves the efficiency of the method; computational cost of the optimal MLRQMC method is one order

of magnitude smaller than that of the optimal MLMC method and five orders of magnitude smaller than that of the standard Monte-Carlo method.

- Chapter 4: In this chapter, the Bayesian inversion is used for a tomography problem namely electrical impedance tomography (EIT), which is modeled by Poisson-Boltzmann equation (PBE). In this work, the Metropolis-Hastings (MH), a Markov-chain Monte-Carlo (MCMC) algorithm in the context of Bayesian estimation methods is used, which is a strong statistical tool for uncertainty quantification (UQ). In this technique, the posterior distribution of the parameters of interest is estimated and confidence intervals for each of the quantities of interest in the EIT inverse problem is found. The Bayesian analysis in the measure-theoretic framework is also studied for the PBE. In this analysis, the focus is on the well-definedness and well-posedness of the posterior measure and the Bayesian estimation for the EIT inverse problem by showing that the solution of the given model is bounded by a function of the parameters and that it is Lipschitz continuous with respect to the parameters.

- Chapter 5: The Bayesian inference also proves its ability in solving the inverse PDE problem in which growth and degradation of biofilms are modeled. In Chapter 5, a system of PDEs as a mathematical model for biofilms in order to study their growth and their response to antibiotic therapy in vivo is presented. This model describes the time dependent evolution of the size of the biofilm, which depends on environmental factors such as temperature or surface composition of the implant material. Quorum sensing is also included in the mathematical model and cooperation of bacteria against antibiotics is a quantity of interest. The results of our experiments for biofilm formation over 24 hours are also presented in this work in order to compare with the results of the computational model. In this work, the Bayesian inversion is used to solve the biofilm inverse problem and extract the quantities of interest in the model, which cannot be determined from the experiments and the computational model.

- Chapter 6: Finally, the conclusions are drawn in Chapter 6.

Part I.

PDE Models in Nanotechnology

2. Analysis of the drift-diffusion-Poisson-Boltzmann System in the Alternating-current Regime

In this chapter, we derive the basic model equations for the alternating-current (AC) small-signal regime for affinity-based field-effect sensors from the drift-diffusion-Poisson system. Whenever the frequency ω of the applied current is sufficiently small, the solution of the system of model equations can be written in terms of the real part of $\exp(\mathrm{i}\omega t)$. This makes it possible to derive the AC model equations, which are the drift-diffusion-Poisson system governing charge transport coupled to the Poisson-Boltzmann equation for the liquid. The unknowns are the electric potential and the concentrations of the positive and negative charge carriers. After the derivation of the system of model equations, the main results are existence and local uniqueness of solutions to the presented AC model equations.

This chapter is organized as follows. The starting point, namely the DC model equations, is presented in Section 2.1. Based on this discussion, the AC model equations are derived in Section 2.2 and then the main mathematical results, primarily existence and local uniqueness for the AC system, are shown in this section. Finally, we summarize the chapter in Section 2.3.

This chapter is based on the author's publication [71].

2.1. The DC Model Equations

In this section, we start with a summary of the DC model equations, i.e., the drift-diffusion-Poisson system. In the next section, the AC model equations will be derived from these equations.

The domain $\Omega \subset \mathbb{R}^d$ ($d \leq 3$) is assumed to be bounded and convex with smooth boundary $\partial\Omega$. In order to model an affinity based sensor, the domain is divided into three subdomains: the semiconductor Ω_{Si} (usually silicon), the dielectric Ω_{ox} (usually silicon oxide), and the electrolyte Ω_{liq} (a liquid containing free ions).

The Poisson equation for the electrostatic potential V is solved on the whole domain Ω and provides self-consistency. More precisely, in the subdomain Ω_{Si}, the drift-diffusion-Poisson system models charge transport. In the second subdomain Ω_{ox}, the Poisson equation holds with a vanishing right side, and in the third domain Ω_{liq}, the nonlinear Poisson-Boltzmann equation holds (see Figure 1.1).

The biofunctionalized layer at the manifold Γ between the subdomains Ω_{liq} and Ω_{ox} is responsible for the recognition of the target molecules. The charges there can be included as immobile charges in ρ_i and free charges in ρ_f, or – since they are many and small – they can be viewed as giving rise to a multiscale problem. Homogenization of a layer at a

manifold in elliptic problems yields two interface conditions [13]. The rapidly oscillating charge concentration near Γ is replaced, as $\epsilon \to 0+$, by the two interface conditions

$$V(0+, y) - V(0-, y) = \alpha(y), \tag{2.1a}$$

$$A(0+)\partial_x V(0+, y) - A(0-)\partial_x V(0-, y) = \gamma(y), \tag{2.1b}$$

where α and γ are essentially given by the dipole-moment and the surface-charge densities of the boundary layer. The local (macroscopic) coordinates are chosen so that $x = 0$ corresponds to Γ and the y-coordinate is normal to Γ. The values of α and γ in the interface conditions result from microscopic models M_α and M_γ for the dipole-moment density and surface charge density of the boundary layer. They generally depend on the electrostatic potential and can be modeled by various means [13, 72].

In summary, the model equations for the DC case (see for example [15]) are the transient boundary-value problem

$$-\nabla \cdot (A\nabla V) = q(C_{\text{dop}} + Z_p p + Z_n n) \qquad \text{in } \Omega_{\text{Si}}, \tag{2.2a}$$

$$-\nabla \cdot (A\nabla V) = 0 \qquad \text{in } \Omega_{\text{ox}}, \tag{2.2b}$$

$$-\nabla \cdot (A\nabla V) = -2\eta \sinh(\beta(V - \Phi)) \qquad \text{in } \Omega_{\text{liq}}, \tag{2.2c}$$

$$\frac{\partial p}{\partial t} + R = \frac{-1}{Z_p q}\nabla \cdot J_p \qquad \text{in } \Omega_{\text{Si}}, \tag{2.2d}$$

$$\frac{\partial n}{\partial t} + R = \frac{-1}{Z_n q}\nabla \cdot J_n \qquad \text{in } \Omega_{\text{Si}}, \tag{2.2e}$$

$$J_p = q(-D_p \nabla p - \mu_p p \nabla V) \tag{2.2f}$$

$$J_n = q(D_n \nabla n - \mu_n n \nabla V) \tag{2.2g}$$

$$V(0+, y) - V(0-, y) = \alpha(y) \qquad \text{on } \Gamma, \tag{2.2h}$$

$$A(0+)\partial_x V(0+, y) - A(0-)\partial_x V(0-, y) = \gamma(y) \qquad \text{on } \Gamma, \tag{2.2i}$$

$$\alpha = M_\alpha(V) \qquad \text{on } \Gamma, \tag{2.2j}$$

$$\gamma = M_\gamma(V) \qquad \text{on } \Gamma, \tag{2.2k}$$

$$p(t = 0) = p_I, \; n(t = 0) = n_I \qquad \text{in } \Omega_{\text{Si}} \tag{2.2l}$$

$$V = V_D \qquad \text{on } \partial\Omega_D, \tag{2.2m}$$

$$\nabla V \cdot \mathbf{n} = 0 \qquad \text{on } \partial\Omega_N, \tag{2.2n}$$

$$p = p_D, \quad n = n_D \qquad \text{on } \partial\Omega_{D,\text{Si}}, \tag{2.2o}$$

$$J_p \cdot \mathbf{n} = 0, \quad J_n \cdot \mathbf{n} = 0 \qquad \text{on } \partial\Omega_{N,\text{Si}}, \tag{2.2p}$$

for almost all $t \in [0, T_0]$, where V is the electrostatic potential, A is the electric permittivity, $q > 0$ is the elementary charge, and C_{dop} is the doping concentration. D_p and D_n are the diffusion coefficients, and μ_p and μ_n are the mobilities of the charge carriers. In the case of a nanowire sensor, $p = p(t, x)$ and $n = n(t, x)$ are the concentrations of holes and electrons, respectively: $Z_p := 1$ for holes and $Z_n := -1$ for electrons. In the case of a nanopore sensor, the charge carriers are cations and anions and the constants Z_p and Z_n correspond to their valences. Furthermore, Φ is the Fermi potential, η is the bulk concentration and $\beta := q/(k_B T)$, where k_B is the Boltzmann constant and T is temperature. Finally, J_p and J_n are the current densities, and R is the recombination rate.

2.1.1. Existence and Local Uniqueness for the Transient DC Model Equations.

Existence and local uniqueness for the transient system (2.2) is an extension of the theory of the stationary drift-diffusion-Poisson system [73, 74], while extending to the additional subdomain Ω_{liq}, where the nonlinear Poisson-Boltzmann equation (2.2c) holds, and including the interface conditions (2.2h)–(2.2i). The interface conditions necessitate an estimate for the electrostatic potential V so that the Schauder fixed-point theory can eventually be applied. The details of the stationary case can be found in [15]. Based on this result, the following result for the transient case, i.e., a system of parabolic and elliptic equations, can be shown. Note that in transient problems, the existence of a solution can only be established for a sufficiently small time interval by a fixed-point argument.

We consider the general form of the elliptic part of the problem

$$
\begin{aligned}
-\nabla \cdot (A(x)\nabla V) + g(x, V) &= f & \forall x \in \Omega \setminus \Gamma, & \quad \text{(2.3a)} \\
V &= V_D & \forall x \in \partial\Omega_D, & \quad \text{(2.3b)} \\
\nabla V \cdot \mathbf{n} &= 0 & \forall x \in \partial\Omega_N, & \quad \text{(2.3c)} \\
V(0+, y) - V(0-, y) &= \alpha(y) & \forall x \in \Gamma, & \quad \text{(2.3d)} \\
A(0+)\partial_x V(0+, y) - A(0-)\partial_x V(0-, y) &= \gamma(y) & \forall x \in \Gamma, & \quad \text{(2.3e)}
\end{aligned}
$$

and the coefficients and boundary conditions in system (2.2) must satisfy the following assumptions.

Assumptions 1.

1. *The bounded domain $\Omega \subset \mathbb{R}^3$ has a C^2 Dirichlet boundary $\partial\Omega_D$, where $|\partial\Omega_D| > 0$. The Neumann boundary $\partial\Omega_N$ with $|\partial\Omega_N| > 0$ consists of C^2 segments.*

   The C^2 manifold $\Gamma \subset \Omega$ splits the domain Ω into two nonempty domains $\Omega^+ = \Omega_{\mathrm{ox}} \cup \Omega_{\mathrm{Si}}$ and $\Omega^- = \Omega_{\mathrm{Si}}$ so that $\mathrm{meas}(\Gamma \cap \partial\Omega) = 0$ and $\Gamma \cap \partial\Omega \subset \partial\Omega_N$ hold.

2. *The coefficient function $A(x)$ is uniformly elliptic, bounded function with*

   $$
   0 < A^- \leq \operatorname{ess\,inf}_{x \in \Omega} A(x) \leq \|A(x)\|_{L^\infty(\Omega)} \leq A^+ < \infty \qquad \text{a.e. in } \Omega, \qquad \text{(2.4)}
   $$

   where $A(x)|_{\Omega^+} \in C^1(\Omega^+, \mathbb{R}^{3 \times 3})$ and $A(x)|_{\Omega^-} \in C^1(\Omega^-, \mathbb{R}^{3 \times 3})$.

   Furthermore, the inclusions $V_D(x) \in H^{1/2}(\partial\Omega) \cap L^\infty(\Gamma)$, $p_D(x, t), n_D(x, t) \in L^2((0, T_0); H^{1/2}(\partial\Omega_{\mathrm{Si}}))$, and $f(x) \in L^\infty(\Omega)$ hold.

3. *The doping concentration $C_{\mathrm{dop}}(x)$ is bounded above and below, and we define*

   $$
   \underline{C} := \inf_{x \in \Omega} C_{\mathrm{dop}}(x) \leq C(x) \leq \sup_{x \in \Omega} C_{\mathrm{dop}}(x) =: \overline{C}.
   $$

4. *The quasi Fermi-level $\Phi(x)$ is bounded above and below, and we define*

   $$
   \underline{\Phi} := \inf_{x \in \Omega} \Phi(x) \leq \Phi(x) \leq \sup_{x \in \Omega} \Phi(x) =: \overline{\Phi}.
   $$

5. *There are constants $0 \leq R_0 < \infty$ and $T_0 > 0$ such that*

$$\|(p_D, n_D)\|_X \leq R_0 \qquad \forall (x, t) \in \partial\Omega_{\mathrm{Si}} \times [0, T_0],$$

   where

$$X := \left\{ (p_D, n_D) \mid p_D, n_D \in L^4((0, T_0); H^{1/2}(\partial\Omega_{\mathrm{Si}})) \right\}$$

   with the norm

$$\|(p_D, n_D)\|_X := \|p_D\|_{L^4((0,T_0); H^{1/2}(\partial\Omega_{\mathrm{Si}}))} + \|n_D\|_{L^4((0,T_0); H^{1/2}(\partial\Omega_{\mathrm{Si}}))}.$$

6. *The microscopic models M_α and M_γ depend continuously in $H^1(\Omega)$ on the potential $V(x)$, and for every potential $V(x)$ in $H^1(\Omega) \cap L^\infty(\Omega)$, the inclusions $\alpha(y) = M_\alpha(V(y)) \in H^{1/2}(\Gamma) \cap L^\infty(\Gamma)$ and $\gamma(y) = M_\gamma(V(y)) \in L^\infty(\Gamma)$ hold.*

Using the assumptions, we can formulate the existence of weak solutions of the system (2.2). Proof of this theorem is straightforward using the Schauder fixed-point theorem.

Proposition 1 (Existence of weak solutions to the DC model)**.** *Under Assumptions 1, there exist a time $T_0 \in (0, \infty)$ and a weak solution*

$$\begin{aligned}
(V, p, n, \alpha, \gamma) \in \big(& H^1(\Omega) \cap L^\infty(\Omega) \big) \times \\
& \big(L^2((0, T_0); H^1(\Omega_{\mathrm{Si}})) \cap L^\infty(\Omega_{\mathrm{Si}} \times [0, T_0]) \big)^2 \times \\
& \big(L^\infty(\Gamma) \cap H^{1/2}(\Gamma) \big) \times L^2(\Gamma)
\end{aligned}$$

of the system (2.2). The function V satisfies the L^∞-estimate

$$\underline{V} \leq V(x) \leq \overline{V} \qquad in \ \Omega, \tag{2.5}$$

where

$$\underline{V} := \min\left(\inf_{\partial\Omega_D} V_D, \underline{\Phi} - \sup_\Omega V_L, U_T \sinh^{-1}\left(\frac{\underline{C}}{2n_i}\right) + \underline{\Phi} - \sup_\Omega V_L \right), \tag{2.6a}$$

$$\overline{V} := \max\left(\sup_{\partial\Omega_D} V_D, \overline{\Phi} - \inf_\Omega V_L, U_T \sinh^{-1}\left(\frac{\overline{C}}{2n_i}\right) + \overline{\Phi} - \inf_\Omega V_L \right), \tag{2.6b}$$

$\underline{\Phi} \leq \Phi(x) \leq \overline{\Phi}$ *holds. $U_T = 1/\beta$ is the thermal voltage and V_L is the solution of the linear elliptic equation (equation (2.3) with $g = 0$), for which the estimate*

$$\|V_L\|_{H^1(\Omega)} \leq C\left(\|f\|_{L^2(\Omega)} + \|V_D\|_{H^{1/2}(\partial\Omega)} + \|\alpha\|_{H^{1/2}(\Gamma)} + \|\gamma\|_{L^2(\Gamma)} \right)$$

holds.

Furthermore, the estimate

$$\|(p, n)\|_X \leq R_0$$

holds for every $x \in \Omega_{\mathrm{Si}}$ and for almost all $t \in [0, T_0]$, where the constant R_0 is radius of the ball

$$B_{R_0} := \{(p, n) \in X \mid \|(p, n)\|_X \leq R_0\} \subset X$$

and $X := \left\{ (p, n) \mid p, n \in L^4((0, T_0); L^2(\Omega_{\mathrm{Si}})) \right\}$ with the norm

$$\|(p, n)\|_X := \|p\|_{L^4((0,T_0); L^2(\Omega_{\mathrm{Si}}))} + \|n\|_{L^4((0,T_0); L^2(\Omega_{\mathrm{Si}}))}.$$

Proposition 1 shows the existence of weak solutions of the DC model system (2.2). Then the question arises whether the solutions are unique. Based on numerical evidence, the solution starts to oscillate between two functions when large voltages are applied as Dirichlet boundary conditions. Furthermore, it is well-known that in the derivation of the drift-diffusion equations one assumes that particle velocities or momenta are distributed according to a Maxwellian distribution, i.e., that they are in thermal equilibrium. However, large applied voltages result in fast particles which are modeled badly by a Maxwellian distribution.

These two reasons, namely the numerical evidence and the shortcomings of the drift-diffusion equations as transport models, suggest that the solutions are not unique in the case of large applied voltages. Indeed, the next theorem states that the solution is unique in a neighborhood around thermal equilibrium and hence the uniqueness is only a local one for sufficiently small Dirichlet boundary conditions.

Since the uniqueness is only local, the maximum principle is not helpful for proving uniqueness. The local uniqueness of solutions of the problem (2.2) is straightforward and can be shown using the implicit-function theorem around the equilibrium solution $(V_e, p_e, n_e, \alpha_e, \gamma_e)$. To apply the implicit-function theorem, one should show that the Fréchet derivative of the problem has a bounded inverse at the equilibrium solution. To this end, the norm estimation of the inverse of the linearization of the equations (2.2) at the equilibrium solution is needed. As mentioned before, the main assumption that we should make is that the applied voltages, i.e., the Dirichlet boundary conditions for the potential, are sufficiently small. We assume that the Dirichlet boundary conditions for the potential V are constants on each of r segments or contacts of the the Dirichlet boundary $\partial\Omega_D$ and denote the potentials there by the vector $(U_1, U_2, \cdots, U_r)$.

In order to state local uniqueness for small applied voltages, we need the following assumptions:

Assumptions 2.

1. *The domain $\Omega \subset \mathbb{R}^3$ is open and bounded and the boundary $\partial\Omega$ is as smooth as necessary (see Assumptions 1).*

2. *The Dirichlet data (V_D, p_D, n_D) are a Lipschitz-continuously differentiable map of $U := (U_1, U_2, \cdots, U_r)$, $\mathbb{R}^r \to H^2(\Omega) \times L^2((0, T_0); H^2(\Omega_{\mathrm{Si}}))^2$.*

3. *The Fréchet derivatives M'_α and M'_γ of the interface models M_α and M_γ with respect to V exist, they are in $H^{1/2}(\Gamma)$ and $L^2(\Gamma)$, respectively, and they satisfy the inequality*

$$\|M'_\alpha(V)\|_{H^{1/2}(\Gamma)} + \|M'_\gamma(V)\|_{L^2(\Gamma)} \le C\|V\|_{H^2(\Omega)} \qquad (2.7)$$

   in a neighborhood of the equilibrium potential V_e with a sufficiently small constant C.

4. *The Shockley-Read-Hall recombination rate R has the form $R = (np - n_i^2)\kappa(t, x, n, p)$, where $\kappa(t, x, \cdot, \cdot) \in C^2(\mathbb{R}_+^2)$ holds for $x \in \Omega$ and for almost all $t \in [0, T_0]$, and where the derivatives $\partial^\nu_{(n,p)}\kappa(\cdot, \cdot, n, p)$ are bounded uniformly for all (n, p) in bounded subsets of $\mathbb{R}_+^2$ and for all multiindices ν with $|\nu| \le 2$. Furthermore, there are constants $\underline{\kappa}$ and $\overline{\kappa}$ such that either $0 < \underline{\kappa} \le \kappa(t, x, n_e, p_e) \le \overline{\kappa}$ or $\kappa(t, x, n_e, p_e) = 0$ for all $x \in \Omega_{\mathrm{Si}}$ and for almost all $t \in [0, T_0]$.*

Proposition 2 (Local uniqueness of weak solutions to the DC model). *Under Assumptions 1 and 2, there exists a sufficiently small $\sigma \in \mathbb{R}$ with $|U| < \sigma$ such that the problem (2.2), has a locally unique solution*

$$\left(V^*(U), p^*(U), n^*(U), \alpha^*(U)\gamma^*(U) \right) \in H^2(\Omega) \times L^2((0, T_0); H^2(\Omega_{\mathrm{S}i}))^2$$
$$\times H^{1/2}(\Gamma) \times L^2(\Gamma).$$

The solution satisfies

$$\left(V^*(0), p^*(0), n^*(0), \alpha^*(0), \gamma^*(0) \right) = (V_e, p_e, n_e, \alpha_e, \gamma_e)$$

and it depends continuously differentiable on the Dirichlet boundary data U as a map from $\{U \in \mathbb{R}^r \mid |U| < \sigma\}$ into $H^2(\Omega) \times L^2((0, T_0); H^2(\Omega_{\mathrm{S}i}))^2 \times H^{1/2}(\Gamma) \times L^2(\Gamma)$.

2.2. Main results: the AC Model Equations

The system (2.2) also models the specific physical situation when an alternating current is applied at the contacts, i.e., when the Dirichlet boundary conditions are transient linear combinations of sine and cosine functions. For these boundary conditions, we derive model equations in this section. These model equations will be stationary, albeit with double the number of unknowns. This means that in the case when the frequency is low enough to justify this model, no computationally expensive solutions of the full transient system (2.2) are necessary, but the stationary system derived here can be used instead.

To find the model equations in the frequency domain, we start by writing the AC solution (V, n, p) as what could be called a periodic perturbation of the DC solution (V_*, n_*, p_*) of system (2.2) as

$$V = V_* + \tilde{V} \exp(\mathrm{i}\omega t), \tag{2.8a}$$
$$n = n_* + \tilde{n} \exp(\mathrm{i}\omega t), \tag{2.8b}$$
$$p = p_* + \tilde{p} \exp(\mathrm{i}\omega t), \tag{2.8c}$$

where $\tilde{V}$, $\tilde{n}$, and $\tilde{p}$ are the complex amplitudes of the AC potential V, the AC electron concentration n, and the AC hole concentration p, respectively. The variables V_*, n_*, and p_* are the DC potential and charge concentrations.

The form (2.8) of the AC unknowns V, n, and p already contains physical assumptions. The main assumption is that the relaxation time of the charge carriers is small compared to the frequency of the sinusoidal applied potentials. The relaxation time here is in the order of the momentum relaxation time. If this were not the case, the solution (V, n, p) could not be written in this form and transport models more complicated than the drift-diffusion-Poisson system would have to be used. The small-frequency assumption is the first part of Assumptions 3, and it is satisfied in the experiments performed so far (e.g., [19–21]).

Assumptions 3.

1. *The frequency ω in (2.8) of the applied AC potentials (i.e., the Dirichlet boundary conditions) is sufficiently small such that the free charge carriers are essentially in equilibrium at each point in time.*

2. *The amplitudes of the applied AC potentials are sufficiently small, and hence the amplitudes $\tilde{V}$, $\tilde{n}$, and $\tilde{p}$ in (2.8) are also small. More precisely, it is assumed that they are small compared to the thermal voltage $k_B T/q$, i.e.,*

$$\|\tilde{n}\|_\infty \ll \frac{k_B T}{q}, \quad \|\tilde{p}\|_\infty \ll \frac{k_B T}{q}, \quad and \quad \|\tilde{V}\|_\infty \ll \frac{k_B T}{q}. \tag{2.9}$$

3. *$\|\tilde{V}\|_{H^1}$ is sufficiently small.*

The second part of the assumptions means that the applied AC potentials are not too large. If they were too large, the validity of the drift-diffusion-Poisson model would be questionable to begin with, since only local uniqueness holds (see Subsection 2.1.1).

The third part of the assumptions is satisfied because of the regularity of solutions of elliptic problems. It is noted here since it plays an important role in the derivation of the model equations.

In order to find the model equations for the AC case, the next step is of course to substitute V, n, and p in (2.8) into the transient model equations (2.2).

2.2.1. The Poisson Equation.

We start with the Poisson equation (2.2a). Substituting (2.8), we get

$$\begin{aligned}
&\nabla \cdot (A\nabla V_*) + \nabla \cdot \big(A\nabla(\tilde{V}\exp(\mathrm{i}\omega t))\big) + qC_{\mathrm{dop}} \\
&+ q(Z_p p_* + Z_n n_* + Z_p \tilde{p}\exp(\mathrm{i}\omega t) + Z_n \tilde{n}\exp(\mathrm{i}\omega t)) = 0.
\end{aligned} \tag{2.10}$$

Then, we subtract the DC Poisson equation for V_* and can cancel the exponential terms to obtain the AC Poisson equation

$$\nabla \cdot (A\nabla\tilde{V}) + q(Z_p \tilde{p} + Z_n \tilde{n}) = 0 \tag{2.11}$$

that holds in in Ω_{liq} and Ω_{Si}. In Ω_{ox}, the AC Poisson equation is simply

$$\nabla \cdot (A\nabla\tilde{V}) = 0. \tag{2.12}$$

2.2.2. The Current Equations.

Regarding the transport part of the model, we consider the DC transport model (2.2d)–(2.2g) and substitute solutions of the form (2.8). The time derivative becomes

$$\frac{\partial p}{\partial t} = \frac{\partial p_*}{\partial t} + \mathrm{i}\omega\tilde{p}\exp(\mathrm{i}\omega t) \tag{2.13}$$

and the current is

$$J_p = -q\Big(D_p\nabla p + Z_p\mu_p p\nabla V\Big), \tag{2.14}$$

where $Z_p = 1$. Then using (2.8), we have

$$\begin{aligned}
J_p = -q\Big(&D_p\nabla p_* + D_p\exp(\mathrm{i}\omega t)\nabla\tilde{p} + \mu_p p_*\nabla V_* + \mu_p p_*\exp(\mathrm{i}\omega t)\nabla\tilde{V} \\
&+ \mu_p\tilde{p}\exp(\mathrm{i}\omega t)\nabla V_* + \mu_p\tilde{p}\exp(2\mathrm{i}\omega t)\nabla\tilde{V}\Big).
\end{aligned} \tag{2.15}$$

The last term is of second order in $\exp(\mathrm{i}\omega t)$. We will neglect it, because it is sufficiently small due to Assumptions 3. Therefore, the transport equation becomes

$$\frac{\partial p_*}{\partial t} + \mathrm{i}\omega \tilde{p} \exp(\mathrm{i}\omega t) + R = \nabla \cdot \Big(D_p \nabla p_* + D_p \exp(\mathrm{i}\omega t) \nabla \tilde{p} + \mu_p p_* \nabla V_* \\ + \mu_p p_* \exp(\mathrm{i}\omega t) \nabla \tilde{V} + \mu_p \tilde{p} \exp(\mathrm{i}\omega t) \nabla V_* \Big). \tag{2.16}$$

After subtracting the DC transport model (2.2d) and (2.14) and canceling the exponential terms, the first equation of the AC transport model is obtained as

$$\nabla \cdot (D_p \nabla \tilde{p}) + \nabla \cdot \big(\mu_p \nabla V_* \tilde{p}\big) + \nabla \cdot \big(\mu_p p_* \nabla \tilde{V}\big) - \mathrm{i}\omega \tilde{p} = 0. \tag{2.17}$$

Following the analogous calculation starting from the DC equations (2.2e) and (2.2g), we obtain the second equation of the AC transport model as

$$\nabla \cdot (D_n \nabla \tilde{n}) + \nabla \cdot \big(\mu_n \nabla V_* \tilde{n}\big) + \nabla \cdot \big(\mu_n n_* \nabla \tilde{V}\big) - \mathrm{i}\omega \tilde{n} = 0. \tag{2.18}$$

2.2.3. Existence and Local Uniqueness for the AC Model Equations

In this section, existence and uniqueness of the solution of the AC model equations (2.11), (2.17), and (2.18), which were derived in the previous section, will be shown. As the AC model is a system of complex valued equations, we consider the real and imaginary parts of the equations separately. To this end, we write the complex amplitudes as

$$\tilde{p} = \Re(\tilde{p}) + \mathrm{i}\Im(\tilde{p}),$$
$$\tilde{n} = \Re(\tilde{n}) + \mathrm{i}\Im(\tilde{n}),$$
$$\tilde{V} = \Re(\tilde{V}) + \mathrm{i}\Im(\tilde{V}),$$

where $\Re$ and $\Im$ denote the real and imaginary parts. Then the complex valued equations become the real valued equations

$$-\nabla \cdot (A\nabla\Re(\tilde{V})) = \begin{cases} q(\Re(\tilde{p}) - \Re(\tilde{n})) & \text{in } \Omega_{\mathrm{Si}} \cup \Omega_{\mathrm{liq}}, \\ 0 & \text{in } \Omega_{\mathrm{ox}}, \end{cases} \tag{2.19a}$$

$$-\nabla \cdot (A\nabla\Im(\tilde{V})) = \begin{cases} q(\Im(\tilde{p}) - \Im(\tilde{n})) & \text{in } \Omega_{\mathrm{Si}} \cup \Omega_{\mathrm{liq}}, \\ 0 & \text{in } \Omega_{\mathrm{ox}}, \end{cases} \tag{2.19b}$$

$$-\nabla \cdot (D_p \nabla\Re(\tilde{p})) - \nabla \cdot \big(\mu_p \nabla V_* \Re(\tilde{p})\big) = \nabla \cdot \big(\mu_p p_* \nabla\Re(\tilde{V})\big) + \omega\Im(\tilde{p}) \qquad \text{in } \Omega_{\mathrm{Si}}, \tag{2.19c}$$

$$-\nabla \cdot (D_p \nabla\Im(\tilde{p})) - \nabla \cdot \big(\mu_p \nabla V_* \Im(\tilde{p})\big) = \nabla \cdot \big(\mu_p p_* \nabla\Im(\tilde{V})\big) - \omega\Re(\tilde{p}) \qquad \text{in } \Omega_{\mathrm{Si}}, \tag{2.19d}$$

$$-\nabla \cdot (D_n \nabla\Re(\tilde{n})) + \nabla \cdot \big(\mu_n \nabla V_* \Re(\tilde{n})\big) = -\nabla \cdot \big(\mu_n n_* \nabla\Re(\tilde{V})\big) + \omega\Im(\tilde{n}) \qquad \text{in } \Omega_{\mathrm{Si}}, \tag{2.19e}$$

$$-\nabla \cdot (D_n \nabla\Im(\tilde{n})) + \nabla \cdot \big(\mu_n \nabla V_* \Im(\tilde{n})\big) = -\nabla \cdot \big(\mu_n n_* \nabla\Im(\tilde{V})\big) - \omega\Re(\tilde{n}) \qquad \text{in } \Omega_{\mathrm{Si}}. \tag{2.19f}$$

The interface between the liquid and the semiconductor in the AC model is described by the equations

$$\Re(\tilde{V})(0+,y) - \Re(\tilde{V})(0-,y) = \alpha(y),$$
$$A(0+)\partial_x\Re(\tilde{V})(0+,y) - A(0-)\partial_x\Re(\tilde{V})(0-,y) = \gamma(y),$$
$$\Im(\tilde{V})(0+,y) - \Im(\tilde{V})(0-,y) = 0,$$
$$A(0+)\partial_x\Im(\tilde{V})(0+,y) - A(0-)\partial_x\Im(\tilde{V})(0-,y) = 0$$

(2.20)

on Γ. Moreover, the Dirichlet boundary conditions on $\partial\Omega_D$ in the AC model read as

$$\Re(\tilde{V}) = \Re(\tilde{V}_D), \qquad \Im(\tilde{V}) = \Im(\tilde{V}_D),$$
$$\Re(\tilde{p}) = \Re(\tilde{p}_D), \qquad \Im(\tilde{p}) = \Im(\tilde{p}_D),$$
$$\Re(\tilde{n}) = \Re(\tilde{n}_D), \qquad \Im(\tilde{n}) = \Im(\tilde{n}_D),$$

(2.21)

and the Neumann boundary conditions on $\partial\Omega_N$ as

$$\nabla\Re(\tilde{V}) \cdot \mathbf{n} = 0, \qquad \nabla\Im(\tilde{V}) \cdot \mathbf{n} = 0,$$
$$\nabla\Re(\tilde{p}) \cdot \mathbf{n} = 0 \qquad \nabla\Im(\tilde{p}) \cdot \mathbf{n} = 0,$$
$$\nabla\Re(\tilde{n}) \cdot \mathbf{n} = 0 \qquad \nabla\Im(\tilde{n}) \cdot \mathbf{n} = 0.$$

(2.22)

To discuss the system (2.19)–(2.22), we note that all the equations are special cases of the general form

$$-\nabla \cdot (A\nabla u) \mp \nabla \cdot (\mathbf{b}u) = f \qquad \forall x \in \Omega \setminus \Gamma, \qquad (2.23\text{a})$$
$$u = u_D \qquad \forall x \in \partial\Omega_D, \qquad (2.23\text{b})$$
$$\nabla u \cdot \mathbf{n} = 0 \qquad \forall x \in \partial\Omega_N, \qquad (2.23\text{c})$$
$$u(0+,y) - u(0-,y) = \alpha(y) \qquad \forall x \in \Gamma, \qquad (2.23\text{d})$$
$$A(0+)\partial_x u(0+,y) - A(0-)\partial_x u(0-,y) = \gamma(y) \qquad \forall x \in \Gamma, \qquad (2.23\text{e})$$

where u represents either $\Re(\tilde{V})$ and $\Im(\tilde{V})$. Because of Assumptions 1, A is uniformly bounded.

The second term on the left-hand side in (2.23a) is additional to fixed-point arguments for the usual drift-diffusion-Poisson system [74], [15, Lemma 3.1]. The complications arising from this term are dealt with in the following lemma and the existence and uniqueness results below. The following lemma shows that under a certain condition, the above problem has a unique solution.

Lemma 1 (Existence and uniqueness of solutions of (2.23))**.** *Under Assumptions 1, if*

$$BC_P \leq A^-$$

holds, where $A^- \leq \|A\|_{L^\infty} \leq A^+$, $\|\mathbf{b}\|_{L^\infty} \leq B$, and C_P is a Poincaré constant, then the boundary-value problem with interface conditions (2.23) has a unique solution $u \in H^1_{u_D}(\Omega)$. Furthermore, the estimate

$$\|u\|_{H^1(\Omega)} \leq C(\|f\|_{L^2(\Omega)} + \|u_D\|_{H^{1/2}(\partial D_{\text{Si},D})} + \|\alpha\|_{H^{1/2}(\Gamma)} + \|\gamma\|_{L^2(\Gamma)}) \qquad (2.24)$$

holds true.

Proof. Multiplying the equations (2.23) by a test function $v \in H_0^1(\Omega)$ and integrating by parts, we obtain the weak formulation

$$\int_\Omega A\nabla w \cdot \nabla v \pm \int_\Omega \mathbf{b}w \cdot \nabla v = \int_\Omega fv - \int_\Omega A\nabla(\overline{u}_D + \overline{\alpha}) \cdot \nabla v$$
$$\mp \int_\Omega \mathbf{b}(\overline{u}_D + \overline{\alpha}) \cdot \nabla v + \int_\Gamma \gamma v \qquad \forall v \in H_0^1(\Omega), \quad (2.25)$$

where $w := u - \overline{u}_D - \overline{\alpha}$ and $\overline{u}_D$ is the Dirichlet lift of u_D in Ω. Also, $\overline{\alpha} \in L^2(\overline{\Omega})$ is an extension of $\alpha \in H^{1/2}(\Gamma)$. Therefore, the weak solution of the boundary-value problem (2.23) is $u \in H_{u_D}^1(\Omega)$ which satisfies

$$a(u, v) = \ell(v) \qquad \forall v \in H_0^1(\Omega), \quad (2.26)$$

where

$$a(u, v) := \int_\Omega A\nabla u \cdot \nabla v \pm \int_\Omega \mathbf{b}u \cdot \nabla v \quad (2.27)$$

and

$$\ell(v) := \int_\Omega fv + \int_\Gamma \gamma v. \quad (2.28)$$

Equivalently, the weak solution of (2.23) is the $w \in H_0^1(\Omega)$ which satisfies

$$a(w, v) = \hat{\ell}(v) \quad \forall v \in H_0^1(\Omega), \quad (2.29)$$

where

$$a(w, v) := \int_\Omega A\nabla w \cdot \nabla v \pm \int_\Omega \mathbf{b}w \cdot \nabla v \quad (2.30)$$

and

$$\hat{\ell}(v) := \int_\Omega fv - \int_\Omega A\nabla(\overline{u}_D + \overline{\alpha}) \cdot \nabla v \mp \int_\Omega \mathbf{b}(\overline{u}_D + \overline{\alpha}) \cdot \nabla v + \int_\Gamma \gamma v. \quad (2.31)$$

In order to apply the Lax-Milgram Lemma, we first show that the bilinear form a is coercive. In other words, we show there exists a constant $C_A \in \mathbb{R}^+$ such that for every $w \in H^1(\Omega)$ the bilinear form a satisfies $|a(w, w)| \geq C_A \|w\|_{H^1(\Omega)}^2$.

Based on the Assumptions 1, we have

$$A^- \|\nabla w\|_{L^2(\Omega)}^2 \leq \int_\Omega A\nabla w \cdot \nabla w$$
$$= a(w, w) \mp \int_\Omega \mathbf{b}w \cdot \nabla w$$
$$\leq |a(w, w)| + \|\mathbf{b}\|_{L^\infty} \int_\Omega |w||\nabla w|,$$

which is equivalent to

$$A^- \|\nabla w\|_{L^2(\Omega)}^2 - \|\mathbf{b}\|_{L^\infty} \int_\Omega |w||\nabla w| \leq |a(w, w)| \qquad \forall w \in H^1(\Omega).$$

Thus, for coercivity, it is sufficient to show that there exists a positive constant C_A such that

$$C_A\|w\|^2_{H^1(\Omega)} \le A^-\|\nabla w\|^2_{L^2(\Omega)} - \|\mathbf{b}\|_{L^\infty}\int_\Omega |w||\nabla w| \qquad \forall w \in H^1(\Omega).$$

Using the inequality

$$xy \le \delta x^2 + \frac{1}{4\delta}y^2 \qquad \forall \delta \in \mathbb{R}^+ \quad \forall x \in \mathbb{R} \quad \forall x \in \mathbb{R}, \tag{2.32}$$

it is sufficient to find a positive constant C_A such that

$$C_A(\|w\|^2_{L^2(\Omega)} + \|\nabla w\|^2_{L^2(\Omega)}) + \|\mathbf{b}\|_{L^\infty}\int_\Omega |w||\nabla w|$$

$$\le C_A(\|w\|^2_{L^2(\Omega)} + \|\nabla w\|^2_{L^2(\Omega)}) + \delta\|\mathbf{b}\|_{L^\infty}\|\nabla w\|^2_{L^2(\Omega)} + \frac{1}{4\delta}\|\mathbf{b}\|_{L^\infty}\|w\|^2_{L^2(\Omega)}$$

$$\le A^-\|\nabla w\|^2_{L^2(\Omega)}. \tag{2.33}$$

We have to discern two cases. If $\|\nabla w\|_{L^2(\Omega)} = 0$, then $\|w\|_{L^2(\Omega)} = 0$ by the Poincaré inequality and the bilinear form a is coercive. If on the other hand $\|\nabla w\|_{L^2(\Omega)} \ne 0$, then dividing inequality (2.33) by $\|\nabla w\|^2_{L^2(\Omega)}$ and using the Poincaré inequality, we find

$$C_A(1 + C_P^2) + B\delta + \frac{B}{4\delta}C_P^2 \le A^-, \tag{2.34}$$

where $B := \|\mathbf{b}\|_{L_\infty}$ and C_P is the Poincaré constant. This inequality always holds for A^- large enough and C_P and B small enough. Maximizing the range of values when it holds leads to the minimization problem

$$\delta := \arg\min_{\delta \in \mathbb{R}^+} B\delta + \frac{B}{4\delta}C_P^2 \tag{2.35}$$

with the solution $\delta := C_P/2$. Substituting δ in (2.34) results in the inequality

$$C_A \le \frac{A^- - BC_P}{1 + C_P^2}. \tag{2.36}$$

Since C_A must be a positive constant, the nominator must be positive, i.e.,

$$BC_P \le A^-, \tag{2.37}$$

which is satisfied by assumption.

Next, the continuity of the bilinear form a is shown. Using the Cauchy-Schwarz inequality we have

$$|a(w, v)| \le \|A\|_{L^\infty}\|\nabla w\|_{L^2(\Omega)}\|\nabla v\|_{L^2(\Omega)} + \|\mathbf{b}\|_{L^\infty}\|w\|_{L^2(\Omega)}\|\nabla v\|_{L^2(\Omega)}$$

$$\le C_0\|w\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)}, \tag{2.38}$$

where $C_0 := A^+ + B$. The continuity of the functional $\hat{\ell}$ is proved similarly by calculating

$$
\begin{aligned}
|\hat{\ell}(v)| &\leq \|f\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} + \|\gamma\|_{L^2(\Gamma)}\|v\|_{L^2(\Omega)} \\
&\quad + \|A\|_{L^\infty}\|\nabla(\overline{u}_D + \overline{\alpha})\|_{L^2(\Omega)}\|\nabla v\|_{L^2(\Omega)} \\
&\quad + \|\mathbf{b}\|_{L^\infty}\|\overline{u}_D + \overline{\alpha}\|_{L^2(\Omega)}\|\nabla v\|_{L^2(\Omega)} \\
&\leq \|f\|_{L^2(\Omega)}\|v\|_{H^1(\Omega)} + \|\gamma\|_{L^2(\Gamma)}\|v\|_{H^1(\Omega)} \\
&\quad + A^+\|\overline{u}_D + \overline{\alpha}\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)} + B\|\overline{u}_D + \overline{\alpha}\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)} \\
&\leq C_1\|v\|_{H^1(\Omega)},
\end{aligned}
$$

where

$$
C_1 := \|f\|_{L^2(\Omega)} + \|\gamma\|_{L^2(\Gamma)} + C_T(A^+ + B)(\|u_D\|_{H^{1/2}(\partial\Omega)} + \|\alpha\|_{H^{1/2}(\Gamma)}).
$$

Here, we used trace inequality (there exists a constant $C_T > 0$ such that $\|\overline{u}_D\|_{H^1(\Omega)} \leq C_T\|u_D\|_{H^{1/2}(\partial\Omega)}$) to obtain the constant C_1.

Hence all the assumptions of the Lax-Milgram Lemma are satisfied. Therefore it proves the existence and uniqueness of the solution of equation (2.23).

Finally, we prove the estimate (2.24). We can write $H^1_{u_D}(\Omega) \ni u = w + \overline{u}_D + \overline{\alpha}$ with $w \in H^1_0(\Omega)$, and all $w \in H^1_0(\Omega)$ satisfy

$$
\begin{aligned}
a(w, w) = a(u - \overline{u}_D - \overline{\alpha}, w) &= a(u, w) - a(\overline{u}_D + \overline{\alpha}, w) \\
&= \ell(w) - a(\overline{u}_D + \overline{\alpha}, w) = \hat{\ell}(w).
\end{aligned}
\tag{2.39}
$$

Here $\overline{u}_D \in H^1(\Omega)$ is the Dirichlet lift of u_D in Ω and thus satisfies $T\overline{u}_D = u_D$, where T is the trace operator. Due to coercivity and (2.39), we have

$$
\|w\|^2_{H^1(\Omega)} \leq \frac{1}{C_A}a(w, w) = \frac{1}{C_A}\hat{\ell}(w).
\tag{2.40}
$$

Using the bound for $\hat{\ell}$, we can write

$$
\|w\|_{H^1(\Omega)} \leq \frac{1}{C_A}|\hat{\ell}(w)| \leq \frac{C_1}{C_A}.
\tag{2.41}
$$

Using the triangle and trace inequalities as well as (2.41), we have

$$
\begin{aligned}
\|u\|_{H^1(\Omega)} &\leq \|w\|_{H^1(\Omega)} + \|\overline{u}_D\|_{H^1(\Omega)} + \|\overline{\alpha}\|_{H^1(\Omega)} \\
&\leq \frac{1}{C_A}\big(\|f\|_{L^2(\Omega)} + \|\gamma\|_{L^2(\Gamma)} + C_T(A^+ + B)(\|u_D\|_{H^{1/2}(\partial\Omega)} + \|\alpha\|_{H^{1/2}(\Gamma)})\big) \\
&\quad + C_T\big(\|u_D\|_{H^{1/2}(\partial\Omega)} + \|\alpha\|_{H^{1/2}(\partial\Omega)}\big) \\
&\leq \frac{1}{C_A}\left(\|f\|_{L^2(\Omega)} + \|\gamma\|_{L^2(\Gamma)}\right) + C_T\left(\frac{A^+ + B}{C_A} + 1\right)(\|u_D\|_{H^{1/2}(\partial\Omega)} + \|\alpha\|_{H^{1/2}(\partial\Omega)})\big) \\
&\leq C\big(\|f\|_{L^2(\Omega)} + \|\gamma\|_{L^2(\Gamma)} + \|u_D\|_{H^{1/2}(\partial\Omega)} + \|\alpha\|_{H^{1/2}(\partial\Omega)}\big),
\end{aligned}
\tag{2.42}
$$

where $C := \max\{\frac{1}{C_A}, C_T(1 + \frac{A^+ + B}{C_A})\}$. $\qquad\square$

Next, we prove the existence and local uniqueness of solutions of the AC model system (2.19)–(2.22). For the existence proof, we apply the Schauder fixed-point theorem to show the existence of a fixed point for a fixed-point map defined by the system of equations. Since a fixed point of this map is a solution of the equations, we thus prove the existence of a solution. To show local uniqueness, we apply the implicit-function theorem.

Proposition 3 (Existence of weak solutions of the AC model). *Under Assumptions 1, the system of boundary-value problems* (2.19)–(2.22) *has a weak solution*

$$\big(\Re(\tilde{V}), \Im(\tilde{V}), \Re(\tilde{p}), \Im(\tilde{p}), \Re(\tilde{n}), \Im(\tilde{n}), \alpha, \gamma\big)$$
$$\in \big(L^\infty(\Omega) \cap H^1(\Omega)\big)^2 \times \big(L^\infty(\Omega_{\mathrm{Si}}) \cap H^1(\Omega_{\mathrm{Si}})\big)^4 \times \big(L^\infty(\Gamma) \cap H^1(\Gamma)\big)^2.$$

Proof. We show the existence of solutions using the Schauder fixed-point theorem. To this end, we define a fixed-point map $G\colon M \to M$, where

$$
\begin{aligned}
M := \Big\{ &\big(\Re(\tilde{V}), \Im(\tilde{V}), \Re(\tilde{p}), \Im(\tilde{p}), \Re(\tilde{n}), \Im(\tilde{n}), \alpha, \gamma\big) \\
&\in \big(L^2(\Omega)\big)^2 \times \big(L^2(\Omega_{\mathrm{Si}})\big)^4 \times \big(L^2(\Gamma)\big)^2 \mid \\
&\alpha, \gamma \text{ bounded a.e. on } \Gamma, \quad \Re(\tilde{V}), \Im(\tilde{V}) \text{ bounded a.e. on } \Omega, \\
&\Re(\tilde{p}), \Im(\tilde{p}), \Re(\tilde{n}), \Im(\tilde{n}) \text{ bounded a.e. on } \Omega_{\mathrm{Si}}\Big\}
\end{aligned}
$$

is a closed and convex subset of $\big(L^2(\Omega)\big)^2 \times \big(L^2(\Omega_{\mathrm{Si}})\big)^4 \times \big(L^2(\Gamma)\big)^2$. For a given solution

$$\big(\Re(\tilde{V}_0), \Im(\tilde{V}_0), \Re(\tilde{p}_0), \Im(\tilde{p}_0), \Re(\tilde{n}_0), \Im(\tilde{n}_0), \alpha_0, \gamma_0\big), \tag{2.43}$$

the map G is defined by

$$
\begin{aligned}
G\big((\Re(\tilde{V}_0), \Im(\tilde{V}_0), &\Re(\tilde{p}_0), \Im(\tilde{p}_0), \Re(\tilde{n}_0), \Im(\tilde{n}_0), \alpha_0, \gamma_0)\big) \\
&:= \big(\Re(\tilde{V}_1), \Im(\tilde{V}_1), \Re(\tilde{p}_1), \Im(\tilde{p}_1), \Re(\tilde{n}_1), \Im(\tilde{n}_1), \alpha_1, \gamma_1)\big).
\end{aligned}
$$

The functions on the right-hand side are the solutions of the following boundary-value problems, which taken together are the given system of equations.

1. Solve the boundary-value problem

$$-\nabla \cdot (A\nabla\Re(\tilde{V})) = \begin{cases} q\big(\Re(\tilde{p}_0) - \Re(\tilde{n}_0)\big) & \text{in } \Omega_{\mathrm{Si}} \cup \Omega_{\mathrm{liq}}, \\ 0 & \text{in } \Omega_{\mathrm{ox}} \end{cases} \tag{2.44}$$

   with the boundary conditions

$$\Re(\tilde{V}) = \Re(\tilde{V}_D) \quad \text{in } \partial\Omega_D \qquad \text{and} \qquad \nabla\Re(\tilde{V}) \cdot \mathbf{n} = 0 \quad \text{in } \partial\Omega_N$$

   and the interface conditions

$$\Re(\tilde{V})(0+, y) - \Re(\tilde{V})(0-, y) = \alpha_0(y) \qquad \text{on } \Gamma,$$
$$A(0+)\partial_x\Re(\tilde{V})(0+, y) - A(0-)\partial_x\Re(\tilde{V})(0-, y) = \gamma_0(y) \qquad \text{on } \Gamma$$

   to obtain the solution $\Re(\tilde{V}) = \Re(\tilde{V}_1)$.

2. Solve the boundary-value problem

$$
-\nabla \cdot (A\nabla\Im(\tilde{V})) = \begin{cases} q\Big(\Im(\tilde{p}_0) - \Im(\tilde{n}_0)\Big) & \text{in } \Omega_{\text{Si}} \cup \Omega_{\text{liq}}, \\ 0 & \text{in } \Omega_{\text{ox}} \end{cases} \tag{2.45}
$$

with the boundary conditions

$$
\Im(\tilde{V}) = \Im(\tilde{V}_D) \quad \text{in } \partial\Omega_D \qquad \text{and} \qquad \nabla\Im(\tilde{V}) \cdot \mathbf{n} = 0 \quad \text{in } \partial\Omega_N
$$

and the interface conditions

$$
\Im(\tilde{V})(0+, y) - \Im(\tilde{V})(0-, y) = 0 \qquad \text{on } \Gamma,
$$
$$
A(0+)\partial_x\Im(\tilde{V})(0+, y) - A(0-)\partial_x\Im(\tilde{V})(0-, y) = 0 \qquad \text{on } \Gamma
$$

to obtain the solution $\Im(\tilde{V}) = \Im(\tilde{V}_1)$.

3. Solve the boundary-value problem

$$
-\nabla \cdot (D_p\nabla\Re(\tilde{p})) - \nabla \cdot \big(\mu_p\nabla V_*\Re(\tilde{p})\big) = \nabla \cdot \big(\mu_p p_*\nabla\Re(\tilde{V}_1)\big) + \omega\Im(\tilde{p}_0) \tag{2.46}
$$

in Ω_{Si} with the boundary conditions

$$
\Re(\tilde{p}) = \Re(\tilde{p}_D) \quad \text{in } \partial\Omega_{D,\text{Si}} \qquad \text{and} \qquad \nabla\Re(\tilde{p}) \cdot \mathbf{n} = 0 \quad \text{in } \partial\Omega_{N,\text{Si}}
$$

to obtain the solution $\Re(\tilde{p}) = \Re(\tilde{p}_1)$.

4. Solve the boundary-value problem

$$
-\nabla \cdot (D_p\nabla\Im(\tilde{p})) - \nabla \cdot \big(\mu_p\nabla V_*\Im(\tilde{p})\big) = \nabla \cdot \big(\mu_p p_*\nabla\Im(\tilde{V}_1)\big) - \omega\Re(\tilde{p}_0) \tag{2.47}
$$

in Ω_{Si} with the boundary conditions

$$
\Im(\tilde{p}) = \Im(\tilde{p}_D) \quad \text{in } \partial\Omega_{D,\text{Si}} \qquad \text{and} \qquad \nabla\Im(\tilde{p}) \cdot \mathbf{n} = 0 \quad \text{in } \partial\Omega_{N,\text{Si}}
$$

to obtain the solution $\Im(\tilde{p}) = \Im(\tilde{p}_1)$.

5. Solve the boundary-value problem

$$
-\nabla \cdot (D_n\nabla\Re(\tilde{n})) + \nabla \cdot \big(\mu_n\nabla V_*\Re(\tilde{n})\big) = -\nabla \cdot \big(\mu_n n_*\nabla\Re(\tilde{V}_1)\big) + \omega\Im(\tilde{n}_0) \tag{2.48}
$$

in Ω_{Si} with the boundary conditions

$$
\Re(\tilde{n}) = \Re(\tilde{n}_D) \quad \text{in } \partial\Omega_{D,\text{Si}} \qquad \text{and} \qquad \nabla\Re(\tilde{n}) \cdot \mathbf{n} = 0 \quad \text{in } \partial\Omega_{N,\text{Si}}
$$

to obtain the solution $\Re(\tilde{n}) = \Re(\tilde{n}_1)$.

6. Solve the boundary-value problem

$$
-\nabla \cdot (D_n\nabla\Im(\tilde{n})) + \nabla \cdot \big(\mu_n\nabla V_*\Im(\tilde{n})\big) = -\nabla \cdot \big(\mu_n n_*\nabla\Im(\tilde{V}_1)\big) - \omega\Re(\tilde{n}_0) \tag{2.49}
$$

in Ω_{Si} with the boundary conditions

$$
\Im(\tilde{n}) = \Im(\tilde{n}_D) \quad \text{in } \partial\Omega_{D,\text{Si}} \qquad \text{and} \qquad \nabla\Im(\tilde{n}) \cdot \mathbf{n} = 0 \quad \text{in } \partial\Omega_{N,\text{Si}}
$$

to obtain the solution $\Im(\tilde{n}) = \Im(\tilde{n}_1)$.

7. Update the surface-charge density and dipole-moment density by the microscopic model

$$
\begin{aligned}
\alpha_1(y) &:= M_\alpha(\Re(\tilde{V}_1) + \mathrm{i}\Im(\tilde{V}_1)), \\
\gamma_1(y) &:= M_\gamma(\Re(\tilde{V}_1) + \mathrm{i}\Im(\tilde{V}_1)).
\end{aligned}
\tag{2.50}
$$

Note that variables indexed with stars are the real valued DC model solutions which are known from Subsection 2.1.1.

All of these boundary-value problems are of the general form (2.23). Therefore existence and uniqueness of the solutions of all problems immediately follow from Lemma 1. Therefore, if we show that

$$
\big(\Re(\tilde{V}_1), \Im(\tilde{V}_1), \Re(\tilde{p}_1), \Im(\tilde{p}_1), \Re(\tilde{n}_1), \Im(\tilde{n}_1), \alpha_1, \gamma_1\big) \in M,
$$

then the fixed-point map G is well-defined. By Lemma 1, there exists a unique and bounded solution

$$
\begin{aligned}
&\big(\Re(\tilde{V}_1), \Im(\tilde{V}_1), \Re(\tilde{p}_1), \Im(\tilde{p}_1), \Re(\tilde{n}_1), \Im(\tilde{n}_1), \alpha_1, \gamma_1\big) \\
&\qquad \in \big(L^\infty(\Omega) \cap H^1(\Omega)\big)^2 \times \big(L^\infty(\Omega_{\mathrm{Si}}) \cap H^1(\Omega_{\mathrm{Si}})\big)^4 \times \big(L^\infty(\Gamma) \cap H^1(\Gamma)\big)^2
\end{aligned}
$$

for a given $\big(\Re(\tilde{V}_0), \Im(\tilde{V}_0), \Re(\tilde{p}_0), \Im(\tilde{p}_0), \Re(\tilde{n}_0), \Im(\tilde{n}_0), \alpha_0, \gamma_0\big)$ in M. This implies

$$
\big(\Re(\tilde{V}_1), \Im(\tilde{V}_1), \Re(\tilde{p}_1), \Im(\tilde{p}_1), \Re(\tilde{n}_1), \Im(\tilde{n}_1), \alpha_1, \gamma_1\big) \in M,
$$

which shows the self-mapping property of G.

According to Lemma 1, continuous dependence of the solution

$$
\big(\Re(\tilde{V}_1), \Im(\tilde{V}_1), \Re(\tilde{p}_1), \Im(\tilde{p}_1), \Re(\tilde{n}_1), \Im(\tilde{n}_1), \alpha_1, \gamma_1\big)
$$

on the data of the problems (2.44)–(2.50) leads us to the estimate

$$
\begin{aligned}
&\|\Re(\tilde{V}_1)\|_{H^1(\Omega)} + \|\Im(\tilde{V}_1)\|_{H^1(\Omega)} + \|\Re(\tilde{p}_1)\|_{H^1(\Omega_{\mathrm{Si}})} + \|\Im(\tilde{p}_1)\|_{H^1(\Omega_{\mathrm{Si}})} \\
&\quad + \|\Re(\tilde{n}_1)\|_{H^1(\Omega_{\mathrm{Si}})} + \|\Im(\tilde{n}_1)\|_{H^1(\Omega_{\mathrm{Si}})} + \|\alpha_1\|_{H^1(\Gamma)} + \|\gamma_1\|_{L^2(\Gamma)} \\
&\leq C\big(\|\Re(\tilde{V}_D)\|_{H^{1/2}(\partial\Omega)} + \|\Im(\tilde{V}_D)\|_{H^{1/2}(\partial\Omega)} + \|\Re(\tilde{p}_0)\|_{H^1(\Omega_{\mathrm{Si}})} + \|\Im(\tilde{p}_0)\|_{H^1(\Omega_{\mathrm{Si}})} \\
&\qquad + \|\Re(\tilde{n}_0)\|_{H^1(\Omega_{\mathrm{Si}})} + \|\Im(\tilde{n}_0)\|_{H^1(\Omega_{\mathrm{Si}})} + \|\Re(\tilde{p}_D)\|_{H^{1/2}(\partial\Omega_{\mathrm{Si}})} + \|\Im(\tilde{p}_D)\|_{H^{1/2}(\partial\Omega_{\mathrm{Si}})} \\
&\qquad + \|\Re(\tilde{n}_D)\|_{H^{1/2}(\partial\Omega_{\mathrm{Si}})} + \|\Im(\tilde{n}_D)\|_{H^{1/2}(\partial\Omega_{\mathrm{Si}})} \\
&\qquad + \|V_*\|_{H^1(\Omega)} + \|p_*\|_{H^1(\Omega_{\mathrm{Si}})} + \|n_*\|_{H^1(\Omega_{\mathrm{Si}})} + \|\alpha_0\|_{H^{1/2}(\Gamma)} + \|\gamma_0\|_{L^2(\Gamma)}\big),
\end{aligned}
$$

using (2.42), where C is a constant. Furthermore, $\|\alpha_1\|_{H^1(\Gamma)}$ and $\|\gamma_1\|_{L^2(\Gamma)}$ are bounded due to the assumptions on M_α and M_γ. Thus, the image $G(M)$ is bounded as a subset of $H^1(\Omega)^2 \times H^1(\Omega_{\mathrm{Si}})^4 \times H^1(\Gamma)^2$, which is compactly embedded in $L^2(\Omega)^2 \times L^2(\Omega_{\mathrm{Si}})^4 \times L^2(\Gamma)^2$ according to the Rellich-Kondrachov theorem. This implies that the closure of $G(M)$ is compact and thus $G(M)$ is precompact. Therefore, G is compact. Compactness and continuity of G lead to the complete continuity of G. Finally, the Schauder fixed-point theorem implies the existence of a weak solution of the AC system of equations (2.19)–(2.22). $\qquad\square$

Proposition 3 means that the AC model system (2.19)–(2.22) always has at least one weak solution. As mentioned before, the applied AC potentials are not too large. Otherwise, the validity of the drift-diffusion-Poisson model would be questionable to begin with, since only local uniqueness holds as discussed in Subsection 2.1.1. Therefore, similarly to the DC case, to show the local uniqueness, we use the implicit-function theorem. We denote the equilibrium solution using an index e.

The equilibrium solution is given by the solution

$$\big(\Re(\tilde{V}_e), \Im(\tilde{V}_e), \Re(\tilde{p}_e), \Im(\tilde{p}_e), \Re(\tilde{n}_e), \Im(\tilde{n}_e), \alpha_e, \gamma_e\big), \tag{2.51}$$

of the equilibrium boundary-value problem

$$
\begin{aligned}
-\nabla \cdot (A\nabla \Re(\tilde{V}_e)) &= -2qn_i \sinh(\beta(\Re(\tilde{V}_e) - \Phi)) && \text{in } \Omega_{\text{Si}}, \\
-\nabla \cdot (A\nabla \Re(\tilde{V}_e)) &= 0 && \text{in } \Omega_{\text{ox}}, \\
-\nabla \cdot (A\nabla \Re(\tilde{V}_e)) &= -2q\eta \sinh(\beta(\Re(\tilde{V}_e) - \Phi)) && \text{in } \Omega_{\text{liq}}, \\
-\nabla \cdot (A\nabla \Im(\tilde{V}_e)) &= -2qn_i \sinh(\beta(\Im(\tilde{V}_e) - \Phi)) && \text{in } \Omega_{\text{Si}}, \\
-\nabla \cdot (A\nabla \Im(\tilde{V}_e)) &= 0 && \text{in } \Omega_{\text{ox}}, \\
-\nabla \cdot (A\nabla \Im(\tilde{V}_e)) &= -2q\eta \sinh(\beta(\Im(\tilde{V}_e) - \Phi)) && \text{in } \Omega_{\text{liq}}, \\
\Re(\tilde{V}_e)(0+, y) - \Re(\tilde{V}_e)(0-, y) &= \alpha_e(y) && \text{on } \Gamma, \\
A(0+)\partial_x \Re(\tilde{V}_e)(0+, y) - A(0-)\partial_x \Re(\tilde{V}_e)(0-, y) &= \gamma_e(y) && \text{on } \Gamma, \\
\Im(\tilde{V}_e)(0+, y) - \Im(\tilde{V}_e)(0-, y) &= 0 && \text{on } \Gamma, \\
A(0+)\partial_x \Im(\tilde{V}_e)(0+, y) - A(0-)\partial_x \Im(\tilde{V}_e)(0-, y) &= 0 && \text{on } \Gamma, \\
\Re(\tilde{V}_e) &= \Re(\tilde{V}_D)(0) && \text{on } \partial\Omega_D, \\
\Im(\tilde{V}_e) &= \Im(\tilde{V}_D)(0) && \text{on } \partial\Omega_D, \\
\nabla \Re(\tilde{V}_e) \cdot \mathbf{n} &= 0 && \text{on } \partial\Omega_N, \\
\nabla \Im(\tilde{V}_e) \cdot \mathbf{n} &= 0 && \text{on } \partial\Omega_N.
\end{aligned}
$$

This solution exists uniquely due to [15, Lemma 3.2]. Furthermore, if all Fermi potentials are equal to Φ, the equilibrium concentrations of charge carriers in Ω_{Si}, i.e., $(\Re(\tilde{p}_e), \Im(\tilde{p}_e), \Re(\tilde{n}_e), \Im(\tilde{n}_e))$ are obtained by

$$\tilde{p}_e = n_i \exp\big(-\beta(\tilde{V}_e - \Phi)\big), \tag{2.53a}$$

$$\tilde{n}_e = n_i \exp\big(\beta(\tilde{V}_e - \Phi)\big). \tag{2.53b}$$

As mentioned, we apply the implicit-function theorem to show the local uniqueness of the solutions of the problem (2.19)–(2.22) around the equilibrium solution (2.51). To this end, similarly to the DC case we show that the Fréchet derivative of the problem has a bounded inverse at the equilibrium solution. Again we assume that the Dirichlet boundary conditions for the potential, are sufficiently small and constant on each of r contacts that partition the whole Dirichlet boundary $\partial\Omega_D$ and denote the potentials there by the vector

$$\tilde{U} := (\tilde{U}_1, \ldots, \tilde{U}_r) := \big((\Re(\tilde{U}_1), \Im(\tilde{U}_1)), \ldots, (\Re(\tilde{U}_r), \Im(\tilde{U}_r))\big) \in \mathbb{C}^r.$$

In summary, the following assumptions are made to show local uniqueness for small applied voltages.

Assumptions 4.

1. *The Dirichlet data* $(\Re(\tilde{V}_D), \Im(\tilde{V}_D), \Re(\tilde{p}_D), \Im(\tilde{p}_D), \Re(\tilde{n}_D), \Im(\tilde{n}_D))$ *are a Lipschitz-continuously differentiable map of* $\tilde{U} := (\tilde{U}_1, \cdots, \tilde{U}_r)$ *from* $\mathbb{C}^r$ *into* $(H^2(\Omega))^2 \times (H^2(\Omega_{\mathrm{Si}}))^4$.

2. *The Fréchet derivatives* M'_α *and* M'_γ *of the interface models* M_α *and* M_γ *with respect to* $\tilde{V}$ *exist, they are in* $H^{1/2}(\Gamma)$ *and* $L^2(\Gamma)$, *respectively, and they satisfy the inequality*

$$\|M'_\alpha(\tilde{V})\|_{H^{1/2}(\Gamma)} + \|M'_\gamma(\tilde{V})\|_{L^2(\Gamma)} \leq C\|\tilde{V}\|_{H^2(\Omega)} \tag{2.54}$$

in a neighborhood of the equilibrium potential $\tilde{V}_e$ *with a sufficiently small constant* C.

Proposition 4 (Local uniqueness of weak solutions of the AC model)**.** *Under Assumptions 1 and 4, there exists a sufficiently small* $\delta \in \mathbb{R}$ *with* $|\tilde{U}| < \delta$ *such that the AC system of equations* (2.19)–(2.22) *has a locally unique solution*

$$\left(\Re(\tilde{V}^*)(\tilde{U}), \Im(\tilde{V}^*)(\tilde{U}), \Re(\tilde{p}^*)(\tilde{U}), \Im(\tilde{p}^*)(\tilde{U}), \Re(\tilde{n}^*)(\tilde{U}), \Re(\tilde{n}^*)(\tilde{U}), \alpha^*(\tilde{U}), \gamma^*(\tilde{U})\right)$$
$$\in (H^2(\Omega))^2 \times (H^2(\Omega_{\mathrm{Si}}))^4 \times H^{1/2}(\Gamma) \times L^2(\Gamma).$$

Furthermore, this solution satisfies

$$\left(\Re(\tilde{V}^*)(0), \Im(\tilde{V}^*)(0), \Re(\tilde{p}^*)(0), \Im(\tilde{p}^*)(0), \Re(\tilde{n}^*)(0), \Re(\tilde{n}^*)(0), \alpha^*(0), \gamma^*(0)\right)$$
$$= \left(\Re(\tilde{V}_e), \Im(\tilde{V}_e), \Re(\tilde{p}_e), \Im(\tilde{p}_e), \Re(\tilde{n}_e), \Im(\tilde{n}_e), \alpha_e, \gamma_e\right),$$

and it depends continuously differentiably on $\tilde{U}$ *as a map from* $\left\{\tilde{U} \in \mathbb{C}^r \mid |\tilde{U}| < \delta, \ r \leq 3\right\}$ *into* $(H^2(\Omega))^2 \times (H^2(\Omega_{\mathrm{Si}}))^4 \times H^{1/2}(\Gamma) \times L^2(\Gamma)$.

Proof. We rewrite the system of equations (2.19)–(2.22) using the substitutions

$$\Re(\hat{V}) := \Re(\tilde{V}) - \Re(\tilde{V}_D)(\tilde{U}), \qquad \Im(\hat{V}) := \Im(\tilde{V}) - \Im(\tilde{V}_D)(\tilde{U}),$$
$$\Re(\hat{p}) := \Re(\tilde{p}) - \Re(\tilde{p}_D)(\tilde{U}), \qquad \Im(\hat{p}) := \Im(\tilde{p}) - \Im(\tilde{p}_D)(\tilde{U}),$$
$$\Re(\hat{n}) := \Re(\tilde{n}) - \Re(\tilde{n}_D)(\tilde{U}), \qquad \Im(\hat{n}) := \Im(\tilde{n}) - \Im(\tilde{n}_D)(\tilde{U}).$$

Then, the system becomes

$$-\nabla \cdot (A\nabla\Re(\hat{V} + \tilde{V}_D)) = q\Big(\Re(\hat{p} + \tilde{p}_D) - \Re(\hat{n} + \tilde{n}_D)\Big) \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.55a)$$

$$-\nabla \cdot (A\nabla\Re(\hat{V} + \tilde{V}_D)) = 0 \qquad \text{in } \Omega_{\text{ox}}, \qquad (2.55b)$$

$$-\nabla \cdot (A\nabla\Re(\hat{V} + \tilde{V}_D)) = -2\eta \sinh(\beta(\Re(\hat{V} + \tilde{V}_D) - \Phi)) \qquad \text{in } \Omega_{\text{liq}}, \qquad (2.55c)$$

$$-\nabla \cdot (A\nabla\Im(\hat{V} + \tilde{V}_D)) = q\Big(\Im(\hat{p} + \tilde{p}_D) - \Im(\hat{n} + \tilde{n}_D)\Big) \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.55d)$$

$$-\nabla \cdot (A\nabla\Im(\hat{V} + \tilde{V}_D)) = 0 \qquad \text{in } \Omega_{\text{ox}}, \qquad (2.55e)$$

$$-\nabla \cdot (A\nabla\Im(\hat{V} + \tilde{V}_D)) = -2\eta \sinh(\beta(\Im(\hat{V} + \tilde{V}_D) - \Phi)) \qquad \text{in } \Omega_{\text{liq}}, \qquad (2.55f)$$

$$-\nabla \cdot (D_p\nabla\Re(\hat{p} + \tilde{p}_D)) - \nabla \cdot \big(\mu_p\nabla V_*\Re(\hat{p} + \tilde{p}_D)\big)$$
$$= \nabla \cdot \big(\mu_p p_*\nabla\Re(\hat{V} + \tilde{V}_D)\big) + \omega\Im(\hat{p} + \tilde{p}_D) \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.55g)$$

$$-\nabla \cdot (D_p\nabla\Im(\hat{p} + \tilde{p}_D)) - \nabla \cdot \big(\mu_p\nabla V_*\Im(\hat{p} + \tilde{p}_D)\big)$$
$$= \nabla \cdot \big(\mu_p p_*\nabla\Im(\hat{V} + \tilde{V}_D)\big) - \omega\Re(\hat{p} + \tilde{p}_D) \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.55h)$$

$$-\nabla \cdot (D_n\nabla\Re(\hat{n} + \tilde{n}_D)) + \nabla \cdot \big(\mu_n\nabla V_*\Re(\hat{n} + \tilde{n}_D)\big)$$
$$= -\nabla \cdot \big(\mu_n n_*\nabla\Re(\hat{V} + \tilde{V}_D)\big) + \omega\Im(\hat{n} + \tilde{n}_D) \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.55i)$$

$$-\nabla \cdot (D_n\nabla\Im(\hat{n} + \tilde{n}_D)) + \nabla \cdot \big(\mu_n\nabla V_*\Im(\hat{n} + \tilde{n}_D)\big)$$
$$= -\nabla \cdot \big(\mu_n n_*\nabla\Im(\hat{V} + \tilde{V}_D)\big) - \omega\Re(\hat{n} + \tilde{n}_D) \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.55j)$$

$$\Re(\hat{V})(0+, y) - \Re(\hat{V})(0-, y) = \alpha(y) \qquad \text{on } \Gamma, \qquad (2.55k)$$

$$A(0+)\partial_x\Re(\hat{V})(0+, y) - A(0-)\partial_x\Re(\hat{V})(0-, y) = \gamma(y) \qquad \text{on } \Gamma, \qquad (2.55l)$$

$$\Im(\hat{V})(0+, y) - \Im(\hat{V})(0-, y) = 0 \qquad \text{on } \Gamma, \qquad (2.55m)$$

$$A(0+)\partial_x\Im(\hat{V})(0+, y) - A(0-)\partial_x\Im(\hat{V})(0-, y) = 0 \qquad \text{on } \Gamma, \qquad (2.55n)$$

$$\alpha = M_\alpha(\Re(\hat{V} + \tilde{V}_D) + \Im(\hat{V} + \tilde{V}_D)) \qquad \text{on } \Gamma, \qquad (2.55o)$$

$$\gamma = M_\gamma(\Re(\hat{V} + \tilde{V}_D) + \Im(\hat{V} + \tilde{V}_D)) \qquad \text{on } \Gamma, \qquad (2.55p)$$

$$\Re(\hat{V}) = \Im(\hat{V}) = 0 \qquad \text{on } \partial\Omega_D, \qquad (2.55q)$$

$$\Re(\hat{p}) = \Im(\hat{p}) = 0 \qquad \text{on } \partial\Omega_{D,\text{Si}}, \qquad (2.55r)$$

$$\Re(\hat{n}) = \Im(\hat{n}) = 0 \qquad \text{on } \partial\Omega_{D,\text{Si}}, \qquad (2.55s)$$

$$\nabla\Re(\hat{V}) \cdot \mathbf{n} = \nabla\Im(\hat{V}) \cdot \mathbf{n} = 0 \qquad \text{on } \partial\Omega_N, \qquad (2.55t)$$

$$\nabla\Re(\hat{p}) \cdot \mathbf{n} = \nabla\Im(\hat{p}) \cdot \mathbf{n} = 0 \qquad \text{on } \partial\Omega_{N,\text{Si}}, \qquad (2.55u)$$

$$\nabla\Re(\hat{n}) \cdot \mathbf{n} = \nabla\Im(\hat{n}) \cdot \mathbf{n} = 0 \qquad \text{on } \partial\Omega_{N,\text{Si}}. \qquad (2.55v)$$

The above system can be written as the operator equation

$$Q\big((\Re(\hat{V}), \Im(\hat{V}), \Re(\hat{p}), \Im(\hat{p}), \Re(\hat{n}), \Im(\hat{n}), \alpha, \gamma), \tilde{U}\big) = 0,$$

where the operator

$$Q\colon D \times S_{\delta_1}(0) \to (L^2(\Omega))^2 \times (L^2(\Omega_{\text{Si}}))^4 \times H^{1/2}(\Gamma) \times L^2(\Gamma)$$

is given by solving the above system. D is an open subset of $H_{\partial}^2(\Omega \setminus \Gamma) \times H_{\partial}^2(\Omega_{\mathrm{Si}})^2 \times H^{1/2}(\Gamma) \times L^2(\Gamma)$, where H_{∂}^2 is defined by

$$H_{\partial}^2(\Omega) := \{\phi \in H^2(\Omega) \mid \nabla\phi \cdot \mathbf{n} = 0 \text{ on } \partial\Omega_N, \ \phi = 0 \text{ on } \partial\Omega_D\}, \tag{2.56}$$

and $S_{\delta_1}(0) \subset \mathbb{C}^r$, $r \leq 3$, is a sphere with radius δ_1 and center 0.

Since

$$Q\big((\Re(\hat{V}), \Im(\hat{V}), \Re(\hat{p}), \Im(\hat{p}), \Re(\hat{n}), \Im(\hat{n}), \alpha, \gamma), \tilde{U}\big)$$
$$\in (L^2(\Omega))^2 \times (L^2(\Omega_{\mathrm{Si}}))^4 \times H^{1/2}(\Gamma) \times L^2(\Gamma)$$

results in $(\Re(\hat{V}), \Im(\hat{V}), \Re(\hat{p}), \Im(\hat{p}), \Re(\hat{n}), \Im(\hat{n}), \alpha, \gamma) \in D$ and $\tilde{U} \in S_{\delta_1}(\tilde{U})$ and since any product of functions in D is in $L^2(\Omega)$ because of the inequality $\|uv\|_{L^2(\tilde{U})} \leq C\|u\|_{H^1(\tilde{U})}\|v\|_{H^1(\tilde{U})}$ for all u and $v \in H^1(\tilde{U})$, the operator Q is well-defined.

To apply the implicit-function theorem, the Fréchet derivative

$$D_{(\Re(\hat{V}), \Im(\hat{V}), \Re(\hat{p}), \Im(\hat{p}), \Re(\hat{n}), \Im(\hat{n}), \alpha, \gamma)}Q$$

must have a bounded inverse at the equilibrium solution

$$\big(\Re(\tilde{V}_e) - \Re(\tilde{V}_D)(0), \Im(\tilde{V}_e) - \Im(\tilde{V}_D)(0), \Re(\tilde{p}_e) - \Re(\tilde{p}_D)(0), \Im(\tilde{p}_e) - \Im(\tilde{p}_D)(0),$$
$$\Re(\tilde{n}_e) - \Re(\tilde{n}_D)(0), \Im(\tilde{n}_e) - \Im(\tilde{n}_D)(0), \alpha_e, \gamma_e, 0\big),$$

which is by definition also a solution of the equation $Q = 0$.

Suppose $(g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8) \in L^2(\Omega) \times L^2(\Omega_{\mathrm{Si}})^4 \times H^{1/2}(\Gamma) \times L^2(\Gamma)$. To find the inverse of the Fréchet derivative, we have to solve the equation

$$D_{(\Re(\hat{V}), \Im(\hat{V}), \Re(\hat{p}), \Im(\hat{p}), \Re(\hat{n}), \Im(\hat{n}), \alpha, \gamma)}Q(a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8)$$
$$= (g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8) \tag{2.57}$$

for a_i, $i \in \{1, \ldots, 8\}$, where $(g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8) \in L^2(\Omega) \times L^2(\Omega_{\mathrm{Si}})^4 \times H^{1/2}(\Gamma) \times L^2(\Gamma)$ and Q are calculated at the equilibrium solution. To write this equation as a boundary-value problem, we have to linearize the original equation (2.55) around the equilibrium solution first. The derivative $D_{(\Re(\hat{V}), \Im(\hat{V}), \Re(\hat{p}), \Im(\hat{p}), \Re(\hat{n}), \Im(\hat{n}), \alpha, \gamma)}Q$ at the equilibrium solution is

$$\begin{pmatrix} -\nabla \cdot (A\nabla) & 0 & -q & 0 & q & 0 & 0 & 0 \\ 0 & -\nabla \cdot (A\nabla) & 0 & -q & 0 & q & 0 & 0 \\ 0 & 0 & L_1 & -\omega & 0 & 0 & 0 & 0 \\ 0 & 0 & \omega & L_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & L_3 & -\omega & 0 & 0 \\ 0 & 0 & 0 & 0 & \omega & L_4 & 0 & 0 \\ -M_{\alpha}'(\Re(\tilde{V}_e)) & -M_{\alpha}'(\Im(\tilde{V}_e)) & 0 & 0 & 0 & 0 & 1 & 0 \\ -M_{\gamma}'(\Re(\tilde{V}_e)) & -M_{\gamma}'(\Im(\tilde{V}_e)) & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

in Ω_{Si}, where

$$
\begin{aligned}
L_1 &:= -\nabla \cdot (D_p \nabla) - \nabla \cdot (\mu_p \nabla V_*), \\
L_2 &:= -\nabla \cdot (D_p \nabla) - \nabla \cdot (\mu_p \nabla V_*), \\
L_3 &:= -\nabla \cdot (D_n \nabla) + \nabla \cdot (\mu_n \nabla V_*), \\
L_4 &:= -\nabla \cdot (D_n \nabla) + \nabla \cdot (\mu_n \nabla V_*).
\end{aligned}
$$

In the matrix above, the first and second rows pertain to the linearizations of the Poisson equation with respect to $\Re(\tilde{V})$ and $\Im(\tilde{V})$, respectively, and hence they depend on the subdomain. In the three subdomains Ω_{Si}, Ω_{ox}, and Ω_{liq}, the linearizations with respect to $\Re(\tilde{V})$ are given by the rows of the matrix

$$
\begin{pmatrix}
-\nabla \cdot (A\nabla) & 0 & -q & 0 & q & 0 & 0 & 0 \\
-\nabla \cdot (A\nabla) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\nabla \cdot (A\nabla) + 2\eta\beta \cosh(\beta(\Re(\tilde{V}_e) - \Phi)) & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

and the linearizations with respect to $\Im(\tilde{V})$ are given by

$$
\begin{pmatrix}
0 & -\nabla \cdot (A\nabla) & 0 & -q & 0 & q & 0 & 0 \\
0 & -\nabla \cdot (A\nabla) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\nabla \cdot (A\nabla) + 2\eta\beta \cosh(\beta(\Im(\tilde{V}_e) - \Phi)) & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.
$$

Therefore, equation (2.57) is equivalent to the boundary-value problem

$$-\nabla \cdot (A\nabla a_1) = q(a_3 - a_5) + g_1 \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.58\text{a})$$

$$-\nabla \cdot (A\nabla a_1) = g_2 \qquad \text{in } \Omega_{\text{ox}}, \qquad (2.58\text{b})$$

$$-\nabla \cdot (A\nabla a_1) = -2\eta \cosh(\beta(\Re(\tilde{V}_e) - \Phi))a_1 + g_1 \qquad \text{in } \Omega_{\text{liq}}, \qquad (2.58\text{c})$$

$$-\nabla \cdot (A\nabla a_2) = q(a_4 - a_6) + g_2 \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.58\text{d})$$

$$-\nabla \cdot (A\nabla a_2) = g_2 \qquad \text{in } \Omega_{\text{ox}}, \qquad (2.58\text{e})$$

$$-\nabla \cdot (A\nabla a_2) = -2\eta \cosh(\beta(\Im(\tilde{V}_e) - \Phi))a_2 + g_2 \qquad \text{in } \Omega_{\text{liq}}, \qquad (2.58\text{f})$$

$$-\nabla \cdot (D_p \nabla a_3) - \nabla \cdot (\mu_p \nabla V_* a_3) = \omega a_4 + g_3 \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.58\text{g})$$

$$-\nabla \cdot (D_p \nabla a_4) - \nabla \cdot (\mu_p \nabla V_* a_4) = -\omega a_3 + g_4 \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.58\text{h})$$

$$-\nabla \cdot (D_n \nabla a_5) + \nabla \cdot (\mu_n \nabla V_* a_5) = \omega a_6 + g_5 \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.58\text{i})$$

$$-\nabla \cdot (D_n \nabla a_6) + \nabla \cdot (\mu_n \nabla V_* a_6) = -\omega a_5 + g_6 \qquad \text{in } \Omega_{\text{Si}}, \qquad (2.58\text{j})$$

$$a_1(0+, y) - a_1(0-, y) = a_7 \qquad \text{on } \Gamma, \qquad (2.58\text{k})$$

$$A(0+)\partial_x a_1(0+, y) - A(0-)\partial_x a_1(0-, y) = a_8 \qquad \text{on } \Gamma, \qquad (2.58\text{l})$$

$$a_2(0+, y) - a_2(0-, y) = 0 \qquad \text{on } \Gamma, \qquad (2.58\text{m})$$

$$A(0+)\partial_x a_2(0+, y) - A(0-)\partial_x a_2(0-, y) = 0 \qquad \text{on } \Gamma, \qquad (2.58\text{n})$$

$$a_7 = M'_\alpha(\Re(\tilde{V}_e))a_1 + M'_\alpha(\Im(\tilde{V}_e))a_2 + g_7 \qquad \text{on } \Gamma, \qquad (2.58\text{o})$$

$$a_8 = M'_\gamma(\Re(\tilde{V}_e))a_1 + M'_\gamma(\Im(\tilde{V}_e))a_2 + g_8 \qquad \text{on } \Gamma, \qquad (2.58\text{p})$$

$$a_1 = a_2 = 0 \qquad \text{on } \partial\Omega_D, \qquad (2.58\text{q})$$

$$a_3 = a_4 = a_5 = a_6 = 0 \qquad \text{on } \partial\Omega_{D,\text{Si}}, \qquad (2.58\text{r})$$

$$\nabla a_1 \cdot \mathbf{n} = \nabla a_2 \cdot \mathbf{n} = 0 \qquad \text{on } \partial\Omega_N, \qquad (2.58\text{s})$$

$$\nabla a_3 \cdot \mathbf{n} = \nabla a_4 \cdot \mathbf{n} = \nabla a_5 \cdot \mathbf{n} = \nabla a_6 \cdot \mathbf{n} = 0 \qquad \text{on } \partial\Omega_{N,\text{Si}}. \qquad (2.58\text{t})$$

There exist unique solutions a_3, a_4, a_5, and a_6 of the semilinear elliptic equations (2.58g)–(2.58j) due to [15, Lemma 3.2], and the estimates

$$\|a_3\|_{H^2(\Omega_{\text{Si}})} + \|a_4\|_{H^2(\Omega_{\text{Si}})} + \|a_5\|_{H^2(\Omega_{\text{Si}})} + \|a_6\|_{H^2(\Omega_{\text{Si}})}$$
$$\leq C(\|g_3\|_{L^2(\Omega_{\text{Si}})} + \|g_4\|_{L^2(\Omega_{\text{Si}})} + \|g_5\|_{L^2(\Omega_{\text{Si}})} + \|g_6\|_{L^2(\Omega_{\text{Si}})})$$

holds in Ω_{Si}. Substituting a_3, a_4, a_5, and a_6 in (2.58a)–(2.58f) and then summing the estimates, we find

$$\|a_1\|_{H^2(\Omega)} + \|a_2\|_{H^2(\Omega)} \leq C_1(\|g_1\|_{L^2(\Omega)} + \|g_2\|_{L^2(\Omega)} + \|g_3\|_{L^2(\Omega_{\text{Si}})}$$
$$+ \|g_4\|_{L^2(\Omega_{\text{Si}})} + \|g_5\|_{L^2(\Omega_{\text{Si}})} + \|g_6\|_{L^2(\Omega_{\text{Si}})} + \|g_7\|_{H^{1/2}(\Gamma)}$$
$$+ \|g_8\|_{L^2(\Gamma)} + \|a_7\|_{H^{1/2}(\Gamma)} + \|a_8\|_{L^2(\Gamma)}).$$

Using equations (2.58o)–(2.58p) and Assumption 4 on M_α and M_γ, there exists a sufficiently small constant C_2 such that the estimate

$$\|a_7\|_{H^{1/2}(\Gamma)} + \|a_8\|_{L^2(\Gamma)} \leq C_2(\|a_1\|_{H^2(\Omega)} + \|a_2\|_{H^2(\Omega)}) + \|g_7\|_{H^{1/2}(\Gamma)} + \|g_8\|_{L^2(\Gamma)}$$

holds. For sufficiently small C_2, i.e., if $C_1 C_2 < 1$, we obtain from the last two inequalities

$$(1 - C_1 C_2)\big(\|a_1\|_{H^2(\Omega)} + \|a_2\|_{H^2(\Omega)}\big) \leq C_1\big(\|g_1\|_{L^2(\Omega)} + \|g_2\|_{L^2(\Omega)} + \|g_3\|_{L^2(\Omega_{\mathrm{Si}})}$$
$$+ \|g_4\|_{L^2(\Omega_{\mathrm{Si}})} + \|g_5\|_{L^2(\Omega_{\mathrm{Si}})} + \|g_6\|_{L^2(\Omega_{\mathrm{Si}})} + 2\|g_7\|_{H^{1/2}(\Gamma)} + 2\|g_8\|_{L^2(\Gamma)}\big).$$

Therefore, the Fréchet derivative of Q at the equilibrium solution has a bounded inverse, i.e., there is a constant C such that

$$\|(D_{(\Re(\hat{V}),\Im(\hat{V}),\Re(\hat{p}),\Im(\hat{p}),\Re(\hat{n}),\Im(\hat{n}),\alpha,\gamma)}Q)^{-1}\| \leq C, \tag{2.59}$$

where the norm is the operator norm of

$$H^2(\Omega) \times H^2(\Omega_{\mathrm{Si}})^4 \times H^{1/2}(\Gamma) \times L^2(\Gamma) \to L^2(\Omega) \times L^2(\Omega_{\mathrm{Si}})^4 \times H^{1/2}(\Gamma) \times L^2(\Gamma). \tag{2.60}$$

Finally, the implicit-function theorem proves the local uniqueness of the solution. $\qquad\square$

2.3. Summary

Many devices, especially sensors, can be used in the AC regime. Still, these model equations have barely been analyzed mathematically. We started from a general model for affinity based field-effect sensors, namely the transient drift-diffusion-Poisson system coupled with the Poisson-Boltzmann equation. Assuming sufficiently low frequencies and sufficiently small signals, the model equations for the AC regime were derived based on the DC regime. The main results are existence and uniqueness for the AC model equations.

3. Optimal Methods for the Stochastic Drift-diffusion-Poisson System

This chapter has two parts, where two optimal methods for a coupled stochastic system, namely the stochastic drift-diffusion-Poisson system are developed. The source of randomness is due to the stochastic coefficients in the mathematical model and the goal is to find an optimal methods to approximate the expected value of the solution of the system. To this end, we develop a multilevel Monte-Carlo and a multilevel randomized quasi Monte-Carlo finite-element method with analyzing their convergence rate and computational complexity. The presented optimal methods are achieved by means of solving optimization problems, which lead to find the optimal discretization parameters and optimal number of samples. Furthermore, computational cost of the two optimal methods is compared in order to find the most efficient stochastic method. The applications include many areas such as transport problems in nanoscale devices.

The chapter is organized as follows. In Section 3.1, we present the system of model equations with stochastic coefficients in detail. In Section 3.2, we define weak solutions of the model equations and prove existence and local-uniqueness theorems. In Section 3.3, we introduce a multi-level Monte-Carlo finite-element method for the system and analyze its rate of convergence. In Section 3.4, we discuss the computational complexity and find the optimal MLMC method. In Section 3.5, we present numerical results for random impurity atoms in nanowire field-effect sensors. The MLMC-FEM method is illustrated there and the computational costs of various numerical techniques are compared as well. Finally, summary of the chapter on MLMC-FEM is presented in Section 3.6 to close the first part of the chapter. The rest of this chapter is organized as follows. A randomized quasi-Monte-Carlo method including an error estimate is developed in Section 3.7. Based on this step, the multilevel version including again an error estimate is developed in Section 3.8. Then Section 3.9 is devoted to finding the optimal method. Finally, the numerical example is discussed in Section 3.10 and a summary is presented in Section 3.11.

This chapter is based on the author's publications [36, 75].

3.1. The Stochastic Model Equations

Suppose that the domain $D \subset \mathbb{R}^d$ is bounded and convex, and that $d \leq 3$. The whole domain D is partitioned into three subdomains with different physical properties and hence different model equations in order to include a large range of applications. The first subdomain D_{Si} consists of the (silicon) nanowire and acts as the transducer of the sensor; in this subdomain, the drift-diffusion-Poisson system describes charge transport. The transducer is surrounded by two materials. First, the oxide layer D_{ox} protects the semiconductor. In D_{ox}, there are no charge carriers and hence simply the Poisson equation

holds. Second, D_{liq} is the aqueous solution containing cations and anions and the Poisson-Boltzmann equation holds.

Also, the boundary layer at the sensor surface is responsible for the recognition of the target molecules. In the case of field-effect sensors, solving a homogenization problem gives rise to two interface conditions for the Poisson equation [13]. In summary, the domain is partitioned into

$$D = D_{\text{Si}} \cup D_{\text{ox}} \cup D_{\text{liq}}.$$

In the subdomain D_{Si}, the stationary drift-diffusion-Poisson system

$$-\nabla \cdot (A(x,\omega)\nabla V(x,\omega)) = q(C_{\text{dop}}(x,\omega) + p(x,\omega) - n(x,\omega)), \tag{3.1a}$$

$$\nabla \cdot J_n(x,\omega) = qR(n(x,\omega), p(x,\omega)), \tag{3.1b}$$

$$\nabla \cdot J_p(x,\omega) = -qR(n(x,\omega), p(x,\omega)), \tag{3.1c}$$

$$J_n(x,\omega) = q(D_n\nabla n(x,\omega) - \mu_n n(x,\omega)\nabla V(x,\omega)), \tag{3.1d}$$

$$J_p(x,\omega) = q(-D_p\nabla p(x,\omega) - \mu_p p(x,\omega)\nabla V(x,\omega)) \tag{3.1e}$$

models charge transport, where $A(x,\omega)$, the permittivity, is a random field with $x \in \mathbb{R}^d$ and a random parameter $\omega \in \Omega$ in a probability space $(\Omega, \mathbb{A}, \mathbb{P})$. Ω denotes the set of elementary events, i.e., the sample space, $\mathbb{A}$ the σ-algebra of all possible events, and $\mathbb{P} \colon \mathbb{A} \to [0,1]$ is a probability measure. $V(x,\omega)$ is the electrostatic potential and $q > 0$ is the elementary charge, $C_{\text{dop}}(x,\omega)$ is the doping concentration, $n(x,\omega)$ and $p(x,\omega)$ are the concentrations of electrons and holes, respectively, $J_n(x,\omega)$ and $J_p(x,\omega)$ are the current densities, D_n and D_p are the diffusion coefficients, μ_n and μ_p are the mobilities, and $R(n(x,\omega), p(x,\omega))$ is the recombination rate. We use the Shockley-Read-Hall recombination rate

$$R(n(x,\omega), p(x,\omega)) := \frac{n(x,\omega)p(x,\omega) - n_i^2}{\tau_p(n(x,\omega) + n_i) + \tau_n(p(x,\omega) + n_i)}$$

here, where the constant n_i is the intrinsic charge density and τ_n and τ_p are the lifetimes of the free carriers, although the mathematical results here hold for many expressions for the recombination rate. Equations (3.1) include the convection terms $-n\nabla V$ and $-p\nabla V$, which prohibit the use of the maximum principle in a simple way.

We assume that the Einstein relations $D_n = U_T\mu_n$ and $D_p = U_T\mu_p$ hold, where the constant U_T is the thermal voltage. Therefore, it is beneficial to change variables from the concentrations n and p to the Slotboom variables u and v defined by

$$n(x,\omega) =: n_i e^{V(x,\omega)/U_T} u(x,\omega),$$

$$p(x,\omega) =: n_i e^{-V(x,\omega)/U_T} v(x,\omega).$$

The system (3.1) then becomes

$$-\nabla \cdot (A\nabla V(x,\omega)) = qn_i(e^{-V(x,\omega)/U_T}v(x,\omega) - e^{V(x,\omega)/U_T}u(x,\omega)) + qC_{\text{dop}}(x,\omega),$$

$$U_T\nabla \cdot (\mu_n e^{V(x,\omega)/U_T}\nabla u(x,\omega)) = \frac{u(x,\omega)v(x,\omega) - 1}{\tau_p(e^{V(x,\omega)/U_T}u(x,\omega) + 1) + \tau_n(e^{-V(x,\omega)/U_T}v(x,\omega) + 1)},$$

$$U_T\nabla \cdot (\mu_p e^{-V(x,\omega)/U_T}\nabla v(x,\omega)) = \frac{u(x,\omega)v(x,\omega) - 1}{\tau_p(e^{V(x,\omega)/U_T}u(x,\omega) + 1) + \tau_n(e^{-V(x,\omega)/U_T}v(x,\omega) + 1)},$$

where the continuity equations are self-adjoint.

The boundary ∂D is partitioned into Dirichlet and Neumann boundaries. For the Ohmic contacts we have

$$V(x,\omega)|_{\partial D_D} = V_D(x), \quad u(x,\omega)|_{\partial D_{\mathrm{Si},D}} = u_D(x) \quad \text{and} \quad v(x,\omega)|_{\partial D_{\mathrm{Si},D}} = v_D(x).$$

At Ohmic contacts the space charge vanishes, i.e., $C_{\mathrm{dop}} + p_D - n_D = 0$, and the system is in thermal equilibrium, i.e., $n_D p_D = n_i^2$ on ∂D_D. Furthermore, at each contact, the quasi Fermi potential levels of silicon are aligned with an external applied voltage $U(x)$. Therefore, by using the quasi Fermi potential, we determine the boundary condition on $\partial D_{\mathrm{Si},D}$ using

$$V_1(x) := U(x) + U_T \ln\left(\frac{n_D(x)}{n_i}\right) = U(x) - U_T \ln\left(\frac{p_D(x)}{n_i}\right).$$

The boundary values $u_D(x)$ and $v_D(x)$ are found to be

$$u_D(x) := n_i^{-1} e^{-V_1(x)/U_T} n_D(x),$$
$$v_D(x) := n_i^{-1} e^{V_1(x)/U_T} p_D(x),$$

where

$$n_D(x) := \frac{1}{2}\left(C_{\mathrm{dop}} + \sqrt{C_{\mathrm{dop}}^2 + 4n_i^2}\right),$$
$$p_D(x) := \frac{1}{2}\left(-C_{\mathrm{dop}} + \sqrt{C_{\mathrm{dop}}^2 + 4n_i^2}\right)$$

hold [74, Chapter 3]. Here, $C_{\mathrm{dop}} := C_D^+ - C_A^-$ is the net doping concentration, where C_D^+ and C_A^- are the donor and acceptor concentrations, respectively.

The zero Neumann boundary conditions

$$\mathbf{n} \cdot \nabla V(x,\omega) = 0, \quad \mathbf{n} \cdot \nabla u(x,\omega) = 0, \quad \mathbf{n} \cdot \nabla v(x,\omega) = 0$$

hold on the rest of the boundary ∂D. Here $\mathbf{n}$ denotes the unit outward normal vector on the boundary.

A jump in the permittivity A always gives rise to two continuity conditions: the continuity of the potential and the continuity of the electric displacement field. Homogenization of an elliptic problem with a periodic boundary layer at a manifold Γ yields the two interface conditions [13]

$$V(0+, y, \omega) - V(0-, y, \omega) = \alpha(y, \omega),$$
$$A(0+)\partial_x V(0+, y, \omega) - A(0-)\partial_x V(0-, y, \omega) = \gamma(y, \omega)$$

between the semiconductor and the liquid. Here we denote the one-dimensional coordinate orthogonal to the manifold Γ by x and the remaining $(d-1)$-dimensional coordinates by y. α and γ are essentially given by the dipole-moment and the surface-charge densities of the boundary layer; in general, we write them as the functional $M_\alpha(V)$ and $M_\gamma(V)$ of the potential V. They may correspond to the Metropolis Monte-Carlo method [72], to

solving the nonlinear Poisson-Boltzmann equation [24], or to systems of ordinary differential equations for surface reactions [7, 76].

In the subdomain D_{ox}, there are no charge carriers and the Poisson equation is simply

$$-\nabla \cdot (A\nabla V(x,\omega)) = 0.$$

In the subdomain D_{liq}, the nonlinear Poisson-Boltzmann equation

$$-\nabla \cdot (A(x,\omega)\nabla V(x,\omega)) + 2\eta \sinh(\beta(V(x,\omega) - \Phi(x,\omega))) = 0$$

holds and models screening by free charges. Here η is the ionic concentration, the constant β equals $\beta := q/(k_B T)$ in terms of the Boltzmann constant k_B and the temperature T, and Φ is the Fermi level.

In summary, for all $\omega \in \Omega$, the model equations are the boundary-value problem

$$- \nabla \cdot (A(x,\omega)\nabla V(x,\omega)) \tag{3.2a}$$
$$= qC_{\mathrm{dop}}(x,\omega) - qn_i(\mathrm{e}^{V(x,\omega)/U_T}u(x,\omega) - \mathrm{e}^{-V(x,\omega)/U_T}v(x,\omega)) \text{in } D_{\mathrm{Si}},$$

$$- \nabla \cdot (A(x,\omega)\nabla V(x,\omega)) = 0 \qquad\qquad\qquad \text{in } D_{\mathrm{ox}}, \tag{3.2b}$$

$$- \nabla \cdot (A(x,\omega)\nabla V(x,\omega)) = -2\eta \sinh(\beta(V(x,\omega) - \Phi(x,\omega))) \text{ in } D_{\mathrm{liq}}, \tag{3.2c}$$

$$V(0+,y,\omega) - V(0-,y,\omega) = \alpha(y,\omega) \qquad\qquad \text{on } \Gamma, \tag{3.2d}$$

$$A(0+)\partial_x V(0+,y,\omega) - A(0-)\partial_x V(0-,y,\omega) = \gamma(y,\omega) \qquad \text{on } \Gamma, \tag{3.2e}$$

$$U_T \nabla \cdot (\mu_n \mathrm{e}^{V(x,\omega)/U_T} \nabla u(x,\omega)) \tag{3.2f}$$
$$= \frac{u(x,\omega)v(x,\omega) - 1}{\tau_p(\mathrm{e}^{V(x,\omega)/U_T}u(x,\omega) + 1) + \tau_n(\mathrm{e}^{-V(x,\omega)/U_T}v(x,\omega) + 1)} \quad \text{in } D_{\mathrm{Si}},$$

$$U_T \nabla \cdot (\mu_p \mathrm{e}^{-V(x,\omega)/U_T} \nabla v(x,\omega)) \tag{3.2g}$$
$$= \frac{u(x,\omega)v(x,\omega) - 1}{\tau_p(\mathrm{e}^{V(x,\omega)/U_T}u(x,\omega) + 1) + \tau_n(\mathrm{e}^{-V(x,\omega)/U_T}v(x,\omega) + 1)} \quad \text{in } D_{\mathrm{Si}},$$

$$\alpha(y,\omega) = M_\alpha(V(y,\omega)) \qquad\qquad\qquad\qquad\qquad \text{in } \Gamma, \tag{3.2h}$$

$$\gamma(y,\omega) = M_\gamma(V(y,\omega)) \qquad\qquad\qquad\qquad\qquad \text{in } \Gamma, \tag{3.2i}$$

$$V(x,\omega) = V_D(x) \qquad\qquad\qquad\qquad\qquad\qquad \text{on } \partial D_D, \tag{3.2j}$$

$$\mathbf{n} \cdot \nabla V(x,\omega) = 0 \qquad\qquad\qquad\qquad\qquad\qquad \text{on } \partial D_N, \tag{3.2k}$$

$$u(x,\omega) = u_D(x), \quad v(x,\omega) = v_D(x) \qquad\qquad \text{on } \partial D_{D,\mathrm{Si}}, \tag{3.2l}$$

$$\mathbf{n} \cdot \nabla u(x,\omega) = 0, \quad \mathbf{n} \cdot \nabla v(x,\omega) = 0 \qquad\qquad \text{on } \partial D_{N,\mathrm{Si}}. \tag{3.2m}$$

3.2. Existence and Local Uniqueness

In order to state the main theoretical results, we first record the assumptions on the data of the system (3.2). The assumptions are moderate in the sense that similar ones are necessary for the deterministic system of equations. Then weak solutions and Bochner spaces are defined. Using the assumptions and definitions, existence and local uniqueness are shown.

3.2.1. Assumptions

The following assumptions are required.

Assumptions 5. *1. The bounded domain $D \subset \mathbb{R}^3$ has a C^2 Dirichlet boundary ∂D_D, the Neumann boundary ∂D_N consists of C^2 segments, and the Lebesgue measure of the Dirichlet boundary ∂D_D is nonzero. The C^2 manifold $\Gamma \subset D$ splits the domain D into two nonempty domains D^+ and D^- so that $\mathrm{meas}(\Gamma \cap \partial D) = 0$ and $\Gamma \cap \partial D \subset \partial D_N$ hold.*

2. $(\Omega, \mathbb{A}, \mathbb{P})$ is a probability space, where Ω denotes the set of elementary events (sample space), $\mathbb{A}$ the σ-algebra of all possible events, and $\mathbb{P} \colon \mathbb{A} \to [0,1]$ is a probability measure.

3. The diffusion coefficient $A(x,\omega)$ is assumed to be a strongly measurable mapping from Ω into $L^\infty(D)$. It is uniformly elliptic and bounded function of position $x \in D$ and the elementary event $\omega \in \Omega$, i.e., there exist constants $0 < A^- < A^+ < \infty$ such that

$$0 < A^- \leq \mathrm{ess\,inf}_{x \in D}\, A(x,\omega) \leq \|A(\cdot,\omega)\|_{L^\infty(D)} \leq A^+ < \infty \quad \forall \omega \in \Omega.$$

Furthermore, $A(x,\omega)|_{D^+ \times \Omega} \in C^1(D^+ \times \Omega, \mathbb{R}^{3\times3})$ and $A(x,\omega)|_{D^- \times \Omega} \in C^1(D^- \times \Omega, \mathbb{R}^{3\times3})$.

4. The doping concentration $C_{\mathrm{dop}}(x,\omega)$ is bounded above and below with the bounds

$$\underline{C} := \inf_{x \in D} C_{\mathrm{dop}}(x,\omega) \leq C_{\mathrm{dop}}(x,\omega) \leq \sup_{x \in D} C_{\mathrm{dop}}(x,\omega) =: \overline{C} \quad \forall \omega \in \Omega.$$

5. There is a constant $\mathbb{R} \ni K \geq 1$ satisfying

$$\frac{1}{K} \leq u_D(x), v_D(x) \leq K \quad \forall x \in \partial D_{\mathrm{Si},D}.$$

6. The functionals $M_\alpha : L^2(\Omega; H^1(D)) \cap L^\infty(D \times \Omega) \to L^2(\Omega; H^{1/2}(\Gamma)) \cap L^\infty(\Gamma \times \Omega)$ and $M_\gamma : L^2(\Omega; H^1(D)) \cap L^\infty(D \times \Omega) \to L^\infty(\Gamma \times \Omega)$ are continuous.

7. The mobilities $\mu_n(x,\omega)$ and $\mu_p(x,\omega)$ are uniformly bounded functions of $x \in D$ and $\omega \in \Omega$, i.e.,

$$0 < \mu_n^- \leq \mu_n(x,\omega) \leq \mu_n^+ < \infty \qquad \forall x \in D, \quad \forall \omega \in \Omega,$$
$$0 < \mu_p^- \leq \mu_p(x,\omega) \leq \mu_p^+ < \infty \qquad \forall x \in D, \quad \forall \omega \in \Omega,$$

where $\mu_p(x,\omega), \mu_n(x,\omega) \in C^1(D_{\mathrm{Si}} \times \Omega, \mathbb{R}^{3\times3})$.

Furthermore, the inclusions $f(x,\omega) \in L^2(\Omega; L^2(D)) \cap L^\infty(D \times \Omega)$, $V_D(x) \in H^{1/2}(\partial D) \cap L^\infty(\Gamma)$, $u_D, v_D(x) \in H^{1/2}(\partial D_{\mathrm{Si}})$, $\alpha(x,\omega) \in L^2(\Omega; H^{1/2}(\Gamma))$, and $\gamma(x,\omega) \in L^2(\Omega; L^2(\Gamma))$ hold.

Assumptions 3 and 7 guarantee the uniform ellipticity of the Poisson and the continuity equations, respectively.

3.2.2. Weak Solution of the Model Equations

In order to define the weak formulation of the stochastic boundary-value problem (3.2), it suffices to consider the semilinear boundary-value problem

$$
\begin{aligned}
-\nabla \cdot (A^*(x,\omega)\nabla w(x,\omega)) + h(x, w(x,\omega)) = f(x,\omega) &\quad \forall x \in D \setminus \Gamma \quad \forall \omega \in \Omega, &\text{(3.3a)}\\
w(x,\omega) = w_D(x) &\quad \forall x \in \partial D_D \quad \forall \omega \in \Omega, &\text{(3.3b)}\\
\mathbf{n} \cdot \nabla w(x,\omega) = 0 &\quad \forall x \in \partial D_N \quad \forall \omega \in \Omega, &\text{(3.3c)}\\
w(0+, y, \omega) - w(0-, y, \omega) = \alpha(y, \omega) &\quad \forall x \in \Gamma \quad \forall \omega \in \Omega, &\text{(3.3d)}\\
A^*(0+)\partial_x w(0+, y, \omega) - A^*(0-)\partial_x w(0-, y, \omega) = \gamma(y, \omega) &\quad \forall x \in \Gamma \quad \forall \omega \in \Omega, &\text{(3.3e)}
\end{aligned}
$$

which is a semilinear Poisson equation with interface conditions. Here (3.3a) includes (3.2a)–(3.2c) if A^* is replaced by the permittivity A, and it includes (3.2f) and (3.2g) if A^* is replaced by $\mu_n e^{V/U_T}$ and $\mu_p e^{-V/U_T}$, respectively. Uniform ellipticity holds in each of these cases per Assumption 5.

For the weak formulation, we define the Hilbert space

$$
X := H_g^1(D) = \left\{ w \in H^1(D) \mid Tw = g \right\} \tag{3.4}
$$

as the solution space, where T is the trace operator defined such that $Tw = g$, where g is Dirichlet lift of $w_D := w|_{\partial D_D}$. The operator T is well-defined and continuous from $H^1(D)$ onto $H^{1/2}(\partial D)$ for the Lipschitz domain D. For $g = 0$, we define the test space

$$
X_0 := H_0^1(D) = \left\{ w \in H^1(D) \mid Tw = 0 \right\}. \tag{3.5}
$$

Definition 1 (Bochner spaces). *Given a Banach space $(X, \|\cdot\|_X)$ and $1 \leq p \leq +\infty$, the Bochner space $L^p(\Omega; X)$ is defined to be the space of all measurable functions $w \colon \Omega \to X$ such that for every $\omega \in \Omega$ the norm*

$$
\|w\|_{L^p(\Omega;X)} := \begin{cases} \left(\int_\Omega \|w(\cdot,\omega)\|_X^p \, d\mathbb{P}(\omega) \right)^{1/p} = \mathbb{E}\left[\|w(\cdot,\omega)\|_X^p \right]^{1/p} < \infty, & 1 \leq p < \infty, \\ \operatorname{ess\,sup}_{\omega \in \Omega} \|w(\cdot,\omega)\|_X < \infty, & p = \infty \end{cases} \tag{3.6}
$$

is finite.

To derive the variational formulation of our model (3.3), we fix the event $\omega \in \Omega$ at first, multiply (3.3a) by a test function $\phi \in L^2(\Omega; X_0)$, and integrate by parts in D to obtain the relation

$$
\int_D A^* \nabla w \cdot \nabla \phi + \int_D h(w)\phi = \int_D f\phi + \int_\Gamma \gamma\phi \qquad \forall \phi \in L^2(\Omega; X_0).
$$

Definition 2 (Weak solution on $D \times \Omega$). *Suppose that A^* satisfies Assumptions 5 and that $f(x,\omega) \in L^2(\Omega; L^2(D))$, $w_D(x) \in H^{1/2}(\partial D_D)$, and $\gamma(x,\omega) \in L^2(\Omega; L^2(\Gamma))$ holds. A function $w \in L^2(\Omega; X)$ is called a* weak solution *of the boundary-value problem (3.3), if it satisfies*

$$
a(w, \phi) = \ell(\phi) \qquad \forall \phi \in L^2(\Omega; X_0), \tag{3.7}
$$

where $a\colon L^2(\Omega; X) \times L^2(\Omega; X_0) \to \mathbb{R}$ and $\ell\colon L^2(\Omega; X_0) \to \mathbb{R}$ are defined by

$$a(w, \phi) := \mathbb{E}\left[\int_D A^* \nabla w \cdot \nabla \phi \mathrm{d}x\right] + \mathbb{E}\left[\int_D h(w)\,\phi \mathrm{d}x\right]$$

and

$$\ell(\phi) := \mathbb{E}\left[\int_D f\phi \mathrm{d}x\right] + \mathbb{E}\left[\int_\Gamma \gamma\phi \mathrm{d}x\right].$$

3.2.3. Existence and Local Uniqueness of the Solution

In the next step, we prove existence and local uniqueness of solutions of system of stochastic elliptic boundary-value problems with interface conditions (3.2) using the Schauder fixed-point theorem and the implicit-function theorem similarly to [15, Theorem 2.2 and 5.2].

Theorem 1 (Existence). *Under Assumptions 5, for every $f(x, \omega) \in L^2(\Omega; L^2(D))$ and $V_D, u_D, v_D \in H^{1/2}(\partial D)$, there exists a weak solution*

$$\begin{aligned}(V(x,\omega), u(x,\omega), v(x,\omega), \alpha(x,\omega), \gamma(x,\omega)) \in \,& \left(L^2(\Omega; H^1_{V_D}(D) \cap L^\infty(D \times \Omega)\right) \\ \times \left(L^2(\Omega; H^1_{u_D}(D_{\mathrm{Si}})) \cap L^\infty(D_{\mathrm{Si}} \times \Omega)\right) \times &\left(L^2(\Omega; H^1_{v_D}(D_{\mathrm{Si}})) \cap L^\infty(D_{\mathrm{Si}} \times \Omega)\right) \\ & \times \left(L^2(\Omega; H^1(\Gamma)) \cap L^\infty(\Gamma \times \Omega)\right)^2\end{aligned}$$

of the stochastic boundary-value problem (3.2), and for every $\omega \in \Omega$ it satisfies the L^∞-estimate

$$\begin{aligned}\underline{V} &\leq V(x,\omega) \leq \overline{V} && in \ D, \\ \frac{1}{K} &\leq u(x,\omega) \leq K && in \ D_{\mathrm{Si}}, \\ \frac{1}{K} &\leq v(x,\omega) \leq K && in \ D_{\mathrm{Si}},\end{aligned}$$

where

$$\underline{V} := \min(\inf_{\partial D_D} V_D, \Phi - \sup_D V_L, U_T \ln(\frac{1}{2K n_i}(\underline{C} + \sqrt{\underline{C}^2 + 4n_i^2})) - \sup_D V_L),$$

$$\overline{V} := \max(\sup_{\partial D_D} V_D, \Phi - \inf_D V_L, U_T \ln(\frac{K}{2n_i}(\overline{C} + \sqrt{\overline{C}^2 + 4n_i^2})) - \inf_D V_L).$$

Here $V_L(x,\omega)$ is the solution of the linear problem (i.e., problem (3.3) with $h \equiv 0$), for which the estimate

$$\|V_L\|_{L^2(\Omega; H^1_{V_D}(D))} \leq C\left(\|f\|_{L^2(\Omega; L^2(D))} + \|V_D\|_{H^{1/2}(\partial D_D)} + \|\alpha\|_{L^2(\Omega; H^{1/2}(\Gamma))} + \|\gamma\|_{L^2(\Omega; L^2(\Gamma))}\right)$$

holds, where C is a positive constant.

Proof. The existence of the solution is proved using the Schauder fixed-point theorem and the estimates are obtained from a maximum principle. First, we define a suitable space

$$N := \left\{ (V, u, v, \alpha, \gamma) \in L^2(\Omega; H^1(D)) \times L^2(\Omega; H^1(D_{\mathrm{Si}}))^2 \times L^2(\Omega; H^1(\Gamma))^2 \mid \right.$$

$$\underline{V} \le V(x, \omega) \le \overline{V} \quad \text{a.e. in } D \times \Omega, \quad \frac{1}{K} \le u(x, \omega), v(x, \omega) \le K \quad \text{a.e. in } D_{\mathrm{Si}} \times \Omega,$$

$$\left. \alpha, \gamma \text{ bounded a.e. on } \Gamma \times \Omega \right\},$$

which is closed and convex. Then we define a fixed-point map $F \colon N \to N$ by

$$F(V_0, u_0, v_0, \alpha_0, \gamma_0) := (V_1, u_1, v_1, \alpha_1, \gamma_1),$$

where the elements of the vector $(V_1, u_1, v_1, \alpha_1, \gamma_1)$ are the solutions of the following equations for given data $(V_0, u_0, v_0, \alpha_0, \gamma_0)$.

1. Solve the elliptic equation

$$\begin{aligned}
-\nabla \cdot (A \nabla V_1) &= q n_i (e^{-V_1/U_T} v_0 - e^{V_1/U_T} u_0) + q C_{\mathrm{dop}} && \text{in } D, \\
\mathbf{n} \cdot \nabla V_1 &= 0 && \text{on } \partial D_N, \\
V_1 &= V_D && \text{on } \partial D_D
\end{aligned}$$

for V_1.

2. Solve the elliptic equation

$$U_T \nabla \cdot (\mu_n e^{V_1/U_T} \nabla u_1)$$
$$-\frac{u_1 v_0 - 1}{\tau_p(e^{V_1/U_T} u_0 + 1) + \tau_n(e^{-V_1/U_T} v_0 + 1)} = 0 \quad \text{in } D_{\mathrm{Si}},$$
$$\mathbf{n} \cdot \nabla u_1 = 0 \quad \text{on } \partial D_{\mathrm{Si},N},$$
$$u_1 = u_D \text{ on } \partial D_{\mathrm{Si},D}$$

for u_1.

3. Solve the elliptic equation

$$U_T \nabla \cdot (\mu_p e^{-V_1/U_T} \nabla v_1)$$
$$-\frac{u_0 v_1 - 1}{\tau_p(e^{V_1/U_T} u_0 + 1) + \tau_n(e^{-V_1/U_T} v_0 + 1)} = 0 \quad \text{in } D_{\mathrm{Si}}$$
$$\mathbf{n} \cdot \nabla v_1 = 0 \quad \text{on } \partial D_{\mathrm{Si},N},$$
$$v_1 = v_D \text{ on } \partial D_{\mathrm{Si},D},$$

for v_1.

4. Update the surface-charge density and dipole-moment density according to the microscopic model

$$\alpha_1(y, \omega) := M_\alpha(V_0),$$
$$\gamma_1(y, \omega) := M_\gamma(V_0).$$

Using Lemmata on the existence and uniqueness of solutions of elliptic boundary-value problems with interface conditions [77], every equation present in the model is uniquely solvable. Therefore the map F is well-defined. Furthermore, continuity and the self-mapping property of F as well as the precompactness of $F(N)$ can be shown similarly to [15, Theorem 2.2] and [78, Theorem 1]. Therefore, applying the Schauder fixed-point theorem yields a fixed-point of F, which is a weak solution of (3.2). □

In general, the solution in Theorem 1 is not unique; uniqueness of the solution only holds in a neighbourhood around thermal equilibrium. This necessitates sufficiently small Dirichlet boundary conditions. The following theorem yields local uniqueness of the solution of our system (3.2) of model equation. The proof is based on the implicit-function theorem.

Theorem 2 (Local uniqueness). *Under Assumption 5, for every $f(x,\omega) \in L^2(\Omega; L^2(D))$, $V_D, u_D, v_D \in H^{1/2}(\partial D)$, $\alpha \in L^2(\Omega; H^{1/2}(\Gamma))$, and $\gamma \in L^2(\Omega; L^2(\Gamma))$, there exists a sufficiently small $\sigma \in \mathbb{R}$ with $|U| < \sigma$ such that the stochastic problem in the existence theorem 1 has a locally unique solution*

$$\Big(V^*(U), u^*(U), v^*(U), \alpha^*(U), \gamma^*(U)\Big) \in L^2(\Omega; H^2(D \setminus \Gamma)) \times L^2(\Omega; H^2(D_{\mathrm{Si}}))^2$$
$$\times L^2(\Omega; H^{1/2}(\Gamma)) \times L^2(\Omega; L^2(\Gamma)).$$

The solution satisfies

$$\Big(V^*(0), u^*(0), v^*(0), \alpha^*(0), \gamma^*(0)\Big) = (V_e, 1, 1, \alpha_e, \gamma_e)$$

and it depends continuously differentiably on U as a map from $\{U \in \mathbb{R}^k, |U| < \sigma\}$ into $L^2(\Omega; H^2(D \setminus \Gamma)) \times L^2(\Omega; H^2(D_{\mathrm{Si}}))^2 \times L^2(\Omega; H^{1/2}(\Gamma)) \times L^2(\Omega; L^2(\Gamma))$.

Proof. We call the equilibrium potential $V_e(x,\omega)$ and the equilibrium surface densities $\alpha_e(x,\omega)$ and $\gamma_e(x,\omega)$. $(V_e, 1, 1, \alpha_e, \gamma_e)$ is a solution of the stochastic equilibrium boundary-value problem, which has a unique solution due to the existence and uniqueness of solutions of stochastic semilinear elliptic boundary-value problems of the form

$$
\begin{aligned}
-\nabla \cdot (A(x,\omega)\nabla V_e(x,\omega)) &= qC_{\mathrm{dop}}(x,\omega) - qn_i(e^{V_e(x,\omega)/U_T} - e^{-V_e(x,\omega)/U_T}) && \text{in } D_{\mathrm{Si}}, \\
-\nabla \cdot (A(x,\omega)\nabla V_e(x,\omega)) &= 0 && \text{in } D_{\mathrm{ox}}, \\
-\nabla \cdot (A(x,\omega)\nabla V_e(x,\omega)) &= -2\eta \sinh(\beta(V_e(x,\omega) - \Phi(x,\omega))) && \text{in } D_{\mathrm{liq}}, \\
V_e(0+,y,\omega) - V_e(0-,y,\omega) &= \alpha_e(y,\omega) && \text{on } \Gamma, \\
A(0+)\partial_x V_e(0+,y,\omega) - A(0-)\partial_x V_e(0-,y,\omega) &= \gamma_e(y,\omega) && \text{on } \Gamma, \\
V_e(x,\omega) &= V_D(x) && \text{on } \partial D_D, \\
\mathbf{n} \cdot \nabla V_e(x,\omega) &= 0 && \text{on } \partial D_N.
\end{aligned}
$$

To apply the implicit-function theorem, we define the map

$$G \colon B \times S_{\sigma_1(0)} \to L^2(\Omega; L^2(D)) \times L^2(\Omega; L^2(D_{\mathrm{Si}}))^2 \times L^2(\Omega; L^2(\Gamma))^2,$$
$$G(V, u, v, \alpha, \gamma, U) = 0,$$

where G is given by the boundary-value problem (3.2) after substituting $\overline{V} := V - V_D(U)$, $\overline{u} := u - u_D(U)$, and $\overline{v} := v - v_D(U)$. B is an open subset of $L^2(\Omega; H_\partial^2(D)) \times L^2(\Omega; H_\partial^2(D_{\text{S}i}))^2 \times L^2(\Omega; L^2(\Gamma))^2$ with

$$H_\partial^2(D) := \{\phi \in H^2(D) \mid \mathbf{n} \cdot \nabla\phi = 0 \text{ on } \partial D_N, \ \phi = 0 \text{ on } \partial D_D\},$$

and the sphere S_{σ_1} with radius σ_1 and center 0 is a subset of $\mathbb{R}^d$. The equilibrium solution $(V_e - V_D(0), 0, 0, \alpha_e, \gamma_e, 0)$ is a solution of the equation $G = 0$. One can show that the Fréchet derivative $D_{(V,u,v,\alpha,\gamma)}G(V_e - V_D(0), 0, 0, \alpha_e, \gamma_e, 0)$ has a bounded inverse (see, e.g., [15, Theorem 2.2]). Then the implicit-function theorem implies uniqueness of the solution of (3.2). $\qquad\square$

3.3. Multi-Level Monte-Carlo Finite-Element Method

We start by briefly recapitulating the finite-element approximation of the system of model equations considered here. Then we review the types of error in the Monte-Carlo approximation of solutions of stochastic partial differential equations in Section 3.3.2. In Section 3.3.3, a multi-level Monte-Carlo (MLMC) finite-element (FE) method for the solution of the system of stochastic equations (3.2) is developed. We give an error bound for MLMC-FEM approximation and discuss the computational complexity.

3.3.1. The Finite-Element Method

In this subsection, we briefly recapitulate the Galerkin finite-element approximation and fix some notation. It provides the foundation for the following section.

We suppose that the domain D can be partitioned into quasi-uniform triangles or tetrahedra such that sequences $\{\tau_{h_\ell}\}_{\ell=0}^\infty$ of regular meshes are obtained. For any $\ell \geq 0$, we denote the mesh size of τ_{h_ℓ} by

$$h_\ell := \max_{K \in \tau_{h_\ell}} \{\text{diam } K\}.$$

To ensure that the mesh quality does not deteriorate as refinements are made, shape-regular meshes can be used.

Definition 3 (Shape regular mesh)**.** *A sequence $\{\tau_{h_\ell}\}_{l=0}^\infty$ of meshes is* shape regular *if there exists a constant $\kappa < \infty$ independent of ℓ such that*

$$\frac{h_K}{\rho_K} \leq \kappa \quad \forall K \in \tau_{h_\ell}.$$

Here ρ_K is the radius of the largest ball that can be inscribed into any $K \in \tau_{h_\ell}$.

Uniform refinement of the mesh can be achieved by regular subdivision. This results in the mesh size

$$h_\ell = r^{-\ell}h_0, \tag{3.8}$$

where h_0 denotes the mesh size of the coarsest triangulation and $r > 1$ is independent of ℓ. The nested family $\{\tau_{h_\ell}\}_{\ell=0}^\infty$ of regular triangulations obtained in this way is shape regular.

The Galerkin approximation is the discrete version of the weak formulation in (3.7) of the stochastic elliptic boundary-value problem (3.2). We consider finite-element discretizations with approximations $u_h \in X_{h_\ell}$ of $u \in X$. Given a mesh τ_{h_ℓ}, X is the solution space (3.4) and $X_{h_\ell} \subset X$ is the discretized space. For all $k \geq 1$, it is defined as

$$X_{h_\ell} := \mathbb{P}^k(\tau_{h_\ell}) := \{u \in X \mid u|_K \in \mathbb{P}^k(K) \ \forall K \in \tau_{h_\ell}\}, \tag{3.9}$$

where $\mathbb{P}^k(K) := \operatorname{span}\{x^\alpha \mid |\alpha| \leq k\}$ is the space of polynomials of total degree less equal k. The space X_0 is the space (3.5) of test functions. The discretized test space $X_{0h_\ell} \subset X_0$ is defined analogously to (3.9).

After introducing the finite-element spaces, everything is ready to define the Galerkin approximation.

Definition 4 (Galerkin approximation). *Suppose $X_{h_\ell} \subset X$ and $X_{0h_\ell} \subset X_0$. The Galerkin approximation of* (3.3) *is the function*

$$w_{h_\ell} \in L^2(\Omega; X_{h_\ell})$$

that satisfies

$$B(w_{h_\ell}, \phi_{h_\ell}) = F(\phi_{h_\ell}) \qquad \forall \phi_{h_\ell} \in L^2(\Omega; X_{0h_\ell}), \tag{3.10}$$

where B and F are defined in (3.7).

3.3.2. Monte-Carlo Finite-Element Approximation

The straightforward Monte-Carlo method for a stochastic PDE approximates the expectation $\mathbb{E}[u]$ of the solution u by the sample mean of a (large) number of evaluations. Since we use the same finite-element mesh τ with the mesh size h for all samples, we drop the index ℓ in this subsection for the MC-FEM. We approximate $\mathbb{E}[u]$ by $\mathbb{E}[u_h]$, where u_h is again the FE approximation of u using a mesh of size h. The standard MC estimator $\mathrm{E_{MC}}$ for $\mathbb{E}[u_h]$ is the sample mean

$$\mathrm{E_{MC}}[u_h] := \hat{u}_h := \frac{1}{M} \sum_{i=1}^{M} u_h^{(i)}. \tag{3.11}$$

where $u_h^{(i)} = u_h(x, \omega^{(i)})$ is the ith sample of the solution.

The following lemma shows the error of the MC estimator for a random variable u which is not discretized in space is of order $O(M^{-1/2})$.

Lemma 2. *For any number of samples $M \in \mathbb{N}$ and for a random variable $u \in L^2(\Omega; X)$, the inequality*

$$\|\mathbb{E}[u] - \mathrm{E_{MC}}[u]\|_{L^2(\Omega; X)} = M^{-1/2}\sigma[u] \tag{3.12}$$

holds for the MC error, where $\sigma[u] := \|\mathbb{E}[u] - u\|_{L^2(\Omega; X)}$.

Proof. The result follows from the calculation

$$\|\mathbb{E}[u] - \mathrm{E}_{\mathrm{MC}}[u]\|_{L^2(\Omega;X)}^2 = \mathbb{E}\Big[\Big\|\mathbb{E}[u] - \frac{1}{M}\sum_{i=1}^M u^{(i)}\Big\|_X^2\Big]$$

$$= \frac{1}{M^2}\sum_{i=1}^M \mathbb{E}\Big[\|\mathbb{E}[u] - u^{(i)}\|_X^2\Big]$$

$$= \frac{1}{M}\mathbb{E}\Big[\|\mathbb{E}[u] - u\|_X^2\Big] = M^{-1}\sigma^2[u].$$

$\square$

Therefore, the variance of the MC estimator is

$$\sigma^2[\mathrm{E}_{\mathrm{MC}}[u]] = \|\mathbb{E}[\mathrm{E}_{\mathrm{MC}}[u]] - \mathrm{E}_{\mathrm{MC}}[u]\|_{L^2(\Omega;X)}^2 = M^{-1}\sigma^2[u]. \tag{3.13}$$

Next we generalize the result to the finite-element solution by using the MC estimator to approximate the expectation $\mathbb{E}[u]$ of a solution u of an SPDE, which is discretized in space by the finite-element method. In other words, if u_h and $\hat{u}_h$ are the finite-element and MC solutions of the SPDE, respectively, then we have

$$\mathbb{E}[u] \approx \mathbb{E}[u_h] \approx \hat{u}_h.$$

Therefore, the MC-FEM method involves two approximations and hence there are two sources of error.

Discretization error The approximation of $\mathbb{E}[u]$ by $\mathbb{E}[u_h]$ gives to the discretization error, which stems from the spatial discretization.

Statistical error The approximation of the expected value $\mathbb{E}[u_h]$ by the sample mean $\hat{u}_h$ gives rise to the statistical error, which is caused by the MC estimator.

Lemma 2 takes care of the statistical error. The order of the discretization error depends on the order of the finite-element method.

Recalling that $\hat{u}_h = \mathrm{E}_{\mathrm{MC}}$, we first obtain the mean square error of the Monte-Carlo FEM in the L^2-norm in the following proposition. Later we also show a theorem for the error in the H^1-norm.

Proposition 5. *Let $\hat{u}_h$ be the Monte-Carlo estimator with M samples to approximate the expectation $\mathbb{E}[u]$ of a solution $u(\cdot,\omega) \in L^2(D)$ of an SPDE by using a FE solution $u_h(\cdot,\omega)$ with mesh size h. Then the mean square error of the Monte-Carlo estimator satisfies*

$$\|\hat{u}_h - \mathbb{E}[u]\|_{L^2(\Omega;L^2(D))}^2 = M^{-1}\sigma^2[u_h] + \|\mathbb{E}[u_h] - \mathbb{E}[u]\|_{L^2(\Omega;L^2(D))}^2. \tag{3.14}$$

Proof. Starting from the mean square error, we calculate

$$\mathrm{MSE} := \|\hat{u}_h - \mathbb{E}[u]\|_{L^2(\Omega;L^2(D))}^2$$

$$= \mathbb{E}\Big[\|\hat{u}_h - \mathbb{E}[u]\|_{L^2(D)}^2\Big] = \mathbb{E}\Big[\int_D (\hat{u}_h - \mathbb{E}[u])^2 \mathrm{d}x\Big]$$

$$= \int_D \mathbb{E}[(\hat{u}_h - \mathbb{E}[u])^2]\mathrm{d}x, \tag{3.15}$$

where the last equation holds due to Fubini's Theorem. Add and subtracting the term $\mathbb{E}[\hat{u}_h]$, we find

$$
\begin{aligned}
\text{MSE} &= \int_D \mathbb{E}[(\hat{u}_h - \mathbb{E}[\hat{u}_h] + \mathbb{E}[\hat{u}_h] - \mathbb{E}[u])^2]\mathrm{d}x \\
&= \int_D \mathbb{E}[(\hat{u}_h - \mathbb{E}[\hat{u}_h])]\mathrm{d}x + \int_D \mathbb{E}[(\mathbb{E}[\hat{u}_h] - \mathbb{E}[u])^2]\mathrm{d}x \\
&= \|\hat{u}_h - \mathbb{E}[\hat{u}_h]\|^2_{L^2(\Omega;L^2(D))} + \|\mathbb{E}[u_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))} \\
&= \sigma^2[\hat{u}_h] + \|\mathbb{E}[u_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))} \\
&= M^{-1}\sigma^2[u_h] + \|\mathbb{E}[u_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))},
\end{aligned}
\tag{3.16}
$$

where we used $\mathbb{E}[\hat{u}_h] = \mathbb{E}[u_h]$, because the Monte-Carlo estimator is unbiased, and $\sigma^2[\hat{u}_h] = M^{-1}\sigma^2[u_h]$ due to equation (3.13). $\qquad\square$

Next we extend this result to H^1. In the following theorems, the finite-element space X is H^1 (see (3.4)).

Theorem 3. *Suppose $\alpha, C_0, C_1 \in \mathbb{R}^+$. Let $\hat{u}_h$ be the Monte-Carlo estimator with M samples to approximate the expectation $\mathbb{E}[u]$ of a solution $u(\cdot, \omega) \in X$ of an SPDE by using a FE solution $u_h(\cdot, \omega) \in X_h$ with mesh size h. Suppose that the discretization error converges with order α, i.e.,*

$$
\|\mathbb{E}[u_h - u]\|_{L^2(\Omega;X)} \le C_1 h^\alpha,
\tag{3.17}
$$

and that the estimate

$$
\sigma^2[u_h] \le C_0
\tag{3.18}
$$

holds. Then the mean square error of the MC estimator satisfies

$$
\|\hat{u}_h - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} = O(h^{2\alpha}) + O(M^{-1}).
\tag{3.19}
$$

Proof. We use the mean square error and calculate

$$
\begin{aligned}
\text{MSE} &:= \|\hat{u}_h - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= \|\hat{u}_h - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))} + \|\nabla\hat{u}_h - \mathbb{E}[\nabla u]\|^2_{L^2(\Omega;L^2(D))} \\
&= \|\hat{u}_h - \mathbb{E}[\hat{u}_h]\|^2_{L^2(\Omega;L^2(D))} + \|\mathbb{E}[\hat{u}_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))} \\
&\quad + \|\nabla\hat{u}_h - \mathbb{E}[\nabla\hat{u}_h]\|^2_{L^2(\Omega;L^2(D))} + \|\mathbb{E}[\nabla\hat{u}_h] - \mathbb{E}[\nabla u]\|^2_{L^2(\Omega;L^2(D))} \\
&= \|\hat{u}_h - \mathbb{E}[\hat{u}_h]\|^2_{L^2(\Omega;X)} + \|\mathbb{E}[\hat{u}_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= \sigma^2[\hat{u}_h] + \|\mathbb{E}[\hat{u}_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= M^{-1}\sigma^2[u_h] + \|\mathbb{E}[u_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)}.
\end{aligned}
\tag{3.20}
$$

In the last expression, $\mathbb{E}[\hat{u}_h] = \mathbb{E}[u_h]$ holds again because the Monte-Carlo estimator is unbiased, and $\sigma^2[\hat{u}_h] = M^{-1}\sigma^2[u_h]$ holds due to (3.13). Therefore, using the assumptions (3.17) and (3.18), we have

$$
\text{MSE} \le C_0 M^{-1} + (C_1 h^\alpha)^2 = O(M^{-1}) + O(h^{2\alpha}),
\tag{3.21}
$$

which concludes the proof. $\qquad\square$

3.3.3. Multi-Level Monte-Carlo Finite-Element Approximation

In this section, we first present the MLMC FE method and an its error. In this method, several levels of meshes are used and the MC estimator is employed to approximate the solution on each level independently. We start by discretizing the variational formulation (3.7) on the sequence

$$X_{h_0} \subset X_{h_1} \subset \cdots \subset X_{h_L} \subset X$$

of finite-dimensional sub-spaces, where $X_{h_\ell} := \mathbb{P}^1(\tau_{h_\ell})$ for all $\ell \in \{0, 1, 2, \ldots, L\}$ (see Section 3.3.1). The finite-element approximation at level L can be written as the telescopic sum

$$u_{h_L} = u_{h_0} + \sum_{\ell=1}^{L}(u_{h_\ell} - u_{h_{\ell-1}}),$$

where each u_{h_ℓ} is the solution on the mesh τ_{h_ℓ} at level ℓ. Therefore, the expected value of u_{h_L} is given by

$$\mathbb{E}[u_{h_L}] = \mathbb{E}[u_{h_0}] + \mathbb{E}\left[\sum_{\ell=1}^{L}(u_{h_\ell} - u_{h_{\ell-1}})\right] = \mathbb{E}[u_{h_0}] + \sum_{\ell=1}^{L}\mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}]. \tag{3.22}$$

In the MLMC FEM, we estimate $\mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}]$ by a level dependent number M_ℓ of samples. The MLMC estimator $\mathbb{E}[u]$ is defined as

$$\mathrm{E}_{\mathrm{MLMC}}[u] := \hat{u}_{h_L} := \mathrm{E}_{\mathrm{MC}}[u_{h_0}] + \sum_{\ell=1}^{L}\mathrm{E}_{\mathrm{MC}}[u_{h_\ell} - u_{h_{\ell-1}}], \tag{3.23}$$

where E_{MC} is the Monte-Carlo estimator defined in (3.11). Therefore, we find

$$\hat{u}_{h_L} = \frac{1}{M_0}\sum_{i=1}^{M_0} u_{h_0}^{(i)} + \sum_{\ell=1}^{L}\frac{1}{M_\ell}\sum_{i=1}^{M_\ell}(u_{h_\ell}^{(i)} - u_{h_{\ell-1}}^{(i)}). \tag{3.24}$$

It is important to note that the approximate solutions $u_{h_\ell}^{(i)}$ and $u_{h_{\ell-1}}^{(i)}$ correspond to the same sample i, but are computed on different levels of the mesh, i.e., on the meshes M_ℓ and $M_{\ell-1}$, respectively.

Recalling the two sources of error constituting the MC-FE error, the following result holds for the MLMC-FEM error.

Theorem 4. *Suppose $\alpha, \beta, C_{00}, C_0, C_1 \in \mathbb{R}^+$. Let $\hat{u}_{h_L}$ be the multi-level Monte-Carlo estimator to approximate the expectation $\mathbb{E}[u]$ of a solution $u(\cdot, \omega) \in X$ of an SPDE by using a FE solution $u_{h_\ell}(\cdot, \omega) \in X_{h_\ell}$ with M_ℓ samples in level ℓ, $\ell \in \{0, 1, 2, \ldots, L\}$ and with mesh size h_ℓ. Suppose that the convergence order α for the discretization error, i.e.,*

$$\|\mathbb{E}[u_{h_\ell}] - \mathbb{E}[u]\|_{L^2(\Omega;X)} \leq C_1 h_\ell^\alpha, \tag{3.25}$$

the convergence order β for

$$\sigma^2[u_{h_\ell} - u_{h_{\ell-1}}] \leq C_0 h_{\ell-1}^\beta, \tag{3.26}$$

and assume that the estimate

$$\sigma^2[u_{h_0}] \leq C_{00} \tag{3.27}$$

holds. Then the mean-square error of the MLMC estimator satisfies

$$\|\mathbb{E}[u] - \hat{u}_{h_L}\|^2_{L^2(\Omega;X)} = O(h_L^{2\alpha}) + O(M_0^{-1}) + \sum_{\ell=1}^{L} O(M_\ell^{-1}) O(h_{\ell-1}^\beta). \tag{3.28}$$

Proof. Analogously to the MC case, the MSE is used to assess the accuracy of the MLMC FE estimator. We calculate

$$
\begin{aligned}
\text{MSE} &:= \|\hat{u}_{h_L} - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= \|\hat{u}_{h_L} - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))} + \|\nabla \hat{u}_{h_L} - \mathbb{E}[\nabla u]\|^2_{L^2(\Omega;L^2(D))} \\
&= \|\hat{u}_{h_L} - \mathbb{E}[\hat{u}_{h_L}]\|^2_{L^2(\Omega;L^2(D))} + \|\mathbb{E}[\hat{u}_{h_L}] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))} \\
&\quad + \|\nabla \hat{u}_{h_L} - \mathbb{E}[\nabla \hat{u}_{h_L}]\|^2_{L^2(\Omega;L^2(D))} + \|\mathbb{E}[\nabla \hat{u}_{h_L}] - \mathbb{E}[\nabla u]\|^2_{L^2(\Omega;L^2(D))} \\
&= \|\hat{u}_{h_L} - \mathbb{E}[\hat{u}_{h_L}]\|^2_{L^2(\Omega;X)} + \|\mathbb{E}[\hat{u}_{h_L}] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= \sigma^2[\hat{u}_{h_L}] + \|\mathbb{E}[\hat{u}_{h_L}] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)}.
\end{aligned}
\tag{3.29}
$$

Expanding as in (3.20), using the relation $\sigma^2[\hat{u}_{h_L}] = \sum_{\ell=0}^{L} M_\ell^{-1} \sigma^2[u_{h_\ell} - u_{h_{\ell-1}}]$ [30], and finally applying the assumptions (3.25)–(3.27), we obtain the asserted estimate by observing that

$$
\begin{aligned}
\text{MSE} &= M_0^{-1} \sigma^2[u_{h_0}] + \sum_{\ell=1}^{L} M_\ell^{-1} \sigma^2[u_{h_\ell} - u_{h_{\ell-1}}] + \|\mathbb{E}[u_{h_L}] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&\leq C_{00} M_0^{-1} + C_0 \sum_{\ell=1}^{L} M_\ell^{-1} h_{\ell-1}^\beta + (C_1 h_L^\alpha)^2 \\
&= O(M_0^{-1}) + \sum_{\ell=1}^{L} O(M_\ell^{-1}) O(h_{\ell-1}^\beta) + O(h_L^{2\alpha}),
\end{aligned}
\tag{3.30}
$$

which concludes the proof. $\qquad\square$

3.4. Optimal Monte-Carlo and Multi-Level Monte-Carlo Methods

In this section, we first estimate the computational cost of the MLMC FE method to achieve a given accuracy and compare it with the MC FE method. Based on these considerations, the computational work is then minimized for a given accuracy to be achieved in order to find the optimal number of samples and the optimal mesh size.

As the model equations (3.2) are a system of PDEs, the work estimate consists of the sum of the work for all equations, i.e., the Poisson equation for V and the two drift-diffusion equations for u and v. Therefore, the total computational work is given by

$$W := W_P + 2W_D = W_{P,a} + W_{P,s} + 2W_{D,a} + 2W_{D,s}, \tag{3.31}$$

where the index P indicates the Poisson equation, the index D indicates the two drift-diffusion equations, the index a denotes assembly of the system matrix, and the index s denotes solving the system matrix. We assume that the necessary number of fixed-point or Newton iterations to achieve numerical convergence is constant; this is supported by the numerical results. For each of these four parts the work per sample in level ℓ is given by

$$W_{\ell,P,a} = \mu_1 h_\ell^{-\gamma_1}, \tag{3.32a}$$

$$W_{\ell,P,s} = \mu_2 h_\ell^{-\gamma_2}, \tag{3.32b}$$

$$W_{\ell,D,a} = \mu_3 h_\ell^{-\gamma_3}, \tag{3.32c}$$

$$W_{\ell,D,s} = \mu_4 h_\ell^{-\gamma_4} \tag{3.32d}$$

with all $\mu_k > 0$ and $\gamma_k > 0$. Here M_ℓ is the number of samples used at level ℓ, and h_ℓ is the corresponding mesh size. Therefore the work per sample is given by

$$W_\ell = W_{\ell,P,a} + W_{\ell,P,s} + 2(W_{\ell,D,a} + W_{\ell,D,s}). \tag{3.33}$$

Analogously, in the case of the vanilla Monte-Carlo method, the computational work is obtained without stratification, i.e., there is only one level. In this case, we will drop the index ℓ.

The exponents (and constants) in equations (3.32) are determined by the algorithm used for assembling the FE matrix in the case of $W_{P,a}$ and $W_{D,a}$ (see, e.g., [79] for an efficient algorithm) and by the order of the FE discretization in the case of $W_{P,s}$ and $W_{D,s}$ (see Section 3.3.2). The constants $\mu_i > 0$ depend on the implementation.

3.4.1. The Optimal Monte-Carlo Finite-Element Method

In the case of the Monte-Carlo method, there is only one level so that the index ℓ will be dropped. We will choose the optimal M and h such that the total computational cost W is minimized given an error bound ϵ to be achieved. This optimization problem with inequality constraints can be solved using the Karush-Kuhn-Tucker (KKT) conditions, which are generalization of Lagrange multipliers in the presence of inequality constraints.

In view of (3.32) and (3.20), the most general problem is the following. We minimize the computational work subject to the accuracy constraint MSE $\leq \varepsilon^2$ so that the root-mean-square error RMSE $\leq \varepsilon$. To this end, we solve the optimization problem

$$\begin{aligned}
\underset{M,h}{\text{minimize}} \quad & f(M,h) := MW \\
\text{subject to} \quad & g(M,h) := \frac{C_0}{M} + (C_1 h^\alpha)^2 - \varepsilon^2 \leq 0,
\end{aligned} \tag{3.34}$$

where the optimization is over $M > 1$ and $h > 0$. To simplify the problem, we introduce the new variable θ with $0 < \theta < 1$ such that

$$\frac{C_0}{M} = \theta \varepsilon^2 \qquad \text{and} \qquad (C_1 h^\alpha)^2 = (1-\theta)\varepsilon^2. \tag{3.35}$$

By viewing h and M as functions of θ, (3.34) becomes a one-dimensional convex optimization problem. Due to the exponents of h and M, it is a nonlinear constraint optimization problem.

Our goal is to formulate the inequality constrained problem as an equality constrained problem to which Newton's method can be applied. In order to solve the optimization problem, we use the interior-point method [80, 81].

For each $\mu > 0$, we replace the non-negativity constraints with logarithmic barrier terms in the objective function

$$\underset{\chi,s}{\text{minimize}} \quad f_\mu(\chi, s) := f(\chi) - \mu \sum_i \ln(s_i) \tag{3.36}$$
$$\text{subject to} \quad g(\chi) - s = 0.$$

Here χ, a vector, denotes (M, h) and the vectors g and s represent the $g_i(x)$ and s_i, respectively. The s_i are restricted to be positive away from zero to ensure that the $\ln(s_i)$ are bounded. As μ decreases to zero, the minimum of f_μ approaches the minimum of f. After denoting the Lagrange multiplier for the system (3.36) by y, the system

$$\nabla f(\chi) - \nabla g(\chi)^T y = 0,$$
$$SYe = \mu e,$$
$$g(\chi) - s = 0$$

is obtained, where S is a diagonal matrix with elements s_i, e is a vector of all ones, and ∇g denotes the Jacobian of the constraint g. Now we apply Newton's method to compute the search directions $\Delta\chi$, Δs, Δh via

$$\begin{pmatrix} H(\chi, y) & 0 & -A(\chi)^T \\ 0 & Y & S \\ A(x) & -I & 0 \end{pmatrix} \begin{pmatrix} \Delta\chi \\ \Delta s \\ \Delta h \end{pmatrix} = \begin{pmatrix} -\nabla f(\chi) + A(\chi)^T y \\ \mu e - SYe \\ -g(\chi) + s \end{pmatrix}. \tag{3.37}$$

The Hessian matrix is given by

$$H(\chi, y) = \nabla^2 f(\chi) - \sum_i y_i \nabla^2 g_i(\chi)$$

and $A(\chi)$ is the Jacobian matrix of the constraint (3.34). The second equation is used to calculate Δs. By substituting into the third equation, we obtain the reduced KKT system

$$\begin{pmatrix} -H(\chi, y) & A(\chi)^T \\ A(\chi) & SY^{-1} \end{pmatrix} \begin{pmatrix} \Delta\chi \\ \Delta s \end{pmatrix} = \begin{pmatrix} \nabla f(\chi) - A(\chi)^T y \\ -h(\chi) + \mu Y^{-1} e \end{pmatrix}. \tag{3.38}$$

Now we use iteration to update the solutions by

$$\chi^{(k+1)} := \chi^{(k)} + \alpha^{(k)} \Delta\chi^{(k)},$$
$$s^{(k+1)} := s^{(k)} + \alpha^{(k)} \Delta s^{(k)},$$
$$y^{(k+1)} := y^{(k)} + \alpha^{(k)} \Delta y^{(k)},$$

where $(\chi^{(0)}, s^{(0)}, y^{(0)})$ is the initial guess and $\alpha^{(k)}$ is chosen to ensure both that $s^{(k+1)} > 0$ and the objective function

$$\Psi_{v,\mu}(\chi, s) = f_\mu(\chi, s) + \frac{v}{2} \|g(\chi) - s\|,$$

is sufficiently reduced [82]. The parameter v may increase with the iteration number to force the solution toward feasibility.

3.4.2. The Optimal Multi-Level Monte-Carlo Finite-Element Method

For an optimal multi-level Monte-Carlo finite-element method, our goal is to determine the optimal hierarchies $(L, \{M_\ell\}_{\ell=0}^{L}, h_0, r)$ which minimize the computational work subject to the given accuracy constraint $\mathrm{MSE} \leq \varepsilon^2$. The optimal number L of levels is also unknown a priori. To this end, we solve the optimization problem

$$
\begin{aligned}
\underset{M_\ell, h_0, r}{\text{minimize}} \quad & f(M_\ell, h_0, r, L) := \sum_{\ell=0}^{L} M_\ell W_\ell \\
\text{subject to} \quad & g(M_\ell, h_0, r, L) := \frac{C_{00}}{M_0} + C_0 \sum_{\ell=1}^{L} \frac{h_{\ell-1}^{\beta}}{M_\ell} + (C_1 h_L^{\alpha})^2 \leq \varepsilon^2.
\end{aligned}
\tag{3.39}
$$

Again, the problem is over $M_\ell > 1$, $h_0 > 0$, and $r > 1$. To obtain the optimal number M_ℓ of samples for $\ell \in \{0, \ldots, L\}$, we calculate

$$
\frac{\partial}{\partial M_\ell}(f + \xi^2 g) = 0,
\tag{3.40}
$$

where ξ^2 is the Lagrange multiplier. This leads to

$$
M_\ell = \xi \sqrt{V_\ell / W_\ell},
\tag{3.41}
$$

where $V_0 = C_{00}$ and $V_\ell = C_0 h_{\ell-1}^{\beta}$. Similarly to (3.35), the equations

$$
\frac{C_{00}}{M_0} + C_0 \sum_{\ell=1}^{L} \frac{h_{\ell-1}^{\beta}}{M_\ell} = \theta \varepsilon^2 \qquad \text{and} \qquad (C_1 h_L^{\alpha})^2 = (1 - \theta) \varepsilon^2
\tag{3.42}
$$

hold. Hence, the Lagrange multiplier is given by

$$
\xi = \left(\theta \epsilon^2\right)^{-1} \sum_{\ell=0}^{L} \sqrt{V_\ell W_\ell}.
\tag{3.43}
$$

Additionally, according to (3.42), h_0 is calculated by

$$
h_0 = \left(\frac{\sqrt{1 - \theta}\, \varepsilon}{C_1} \right)^{1/\alpha} r^L.
\tag{3.44}
$$

Thus we arrive at a two-dimensional optimization problem for the unknowns θ and r. Similarly to the vanilla Monte-Carlo case, we use the interior-point method to solve this nonlinear problem and optimize the hierarchies. In problems with two or three physical/spatial dimensions, the optimal determination of the mesh sizes h_ℓ is a crucial factor in the optimization problem specifically if the exponents γ_k are greater than 1.

There are two options: one is to choose the h_ℓ as a geometric progression according to (3.8). In this case, we solve the minimization problem (3.39). The other is to choose the mesh sizes h_ℓ freely such that they only satisfy the natural condition

$$
h_0 \geq h_1 \geq h_2 \geq \cdots \geq h_L.
$$

We will explore both options in Section 3.5.

In the second case, when the mesh sizes are freely chosen, we write them as

$$h_\ell := \frac{h_0}{r_\ell}, \qquad \ell = 1, \ldots, L,$$

where

$$r_\ell := \prod_{i=1}^{\ell} r_i \qquad \text{and} \qquad r_i \geq 1. \tag{3.45}$$

It is clear that $r_L \geq r_{L-1} \ldots \geq r_1 \geq 1$. Here the optimization problem is an $(L+1)$-dimensional problem for the unknowns θ and $r_1, \ldots, r_L$. The same procedure can be applied to solve the problem.

In the next section, we apply these two approaches to a MLMC FE method and discuss their efficiency.

3.5. Numerical Results

In this section, we present numerical results for the Monte-Carlo and multi-level Monte-Carlo methods for the drift-diffusion-Poisson system. We also investigate the choices of the FE mesh sizes on each level, namely as geometric progressions or freely chosen. The random coefficients in the drift-diffusion-Poisson system considered here stem from a real-world application, namely the effect of random dopants in nanoscale semiconductor devices, which is of great importance in its own right.

3.5.1. The Leading Example

Random-dopant effects are called discrete-dopant fluctuation effects [83–85]. In nanoscale semiconductor devices, the charge profile of the dopant atoms cannot be validly modeled as a continuum anymore, but the random location of each dopant needs to be taken into account. This means that each device is a realization of a random process and corresponds to an event ω. In this manner, the potential and carrier-density fluctuations due to the discreteness and randomness of the dopants are clearly captured.

Here the silicon lattice is doped with boron as the impurity atoms. The domain $D \subset \mathbb{R}^2$ is depicted in Figure 3.1. The thickness of the oxide layer is $8\,\mathrm{nm}$, the thickness of the nanowire is $40\,\mathrm{nm}$, its width is $60\,\mathrm{nm}$ and the nanowire length is $60\,\mathrm{nm}$. Regarding the geometry, Dirichlet boundary conditions are used at the contacts with a back-gate voltage of $-1\,\mathrm{V}$ (at the bottom of the device) and an electrode voltage of $0\,\mathrm{V}$ (at the top of the device). Zero Neumann boundary conditions are used everywhere else. The relative permittivities in the subdomains are $A_{\mathrm{Si}} = 11.7$, $A_{\mathrm{ox}} = 3.9$, $A_{\mathrm{liq}} = 78$, and $A_{\mathrm{dop}} = 4.2$. The number of dopants placed randomly in the device corresponds to a doping concentration of $4 \cdot 10^{16}\,\mathrm{cm}^{-3}$. According to its volume, the silicon subdomain hence contains 6 negative impurity atoms when $C_{\mathrm{dop}} = 5 \cdot 10^{15}\,\mathrm{cm}^{-3}$ and 600 dopants when $C_{\mathrm{dop}} = 5 \cdot 10^{17}\,\mathrm{cm}^{-3}$.

Regarding the doping concentration C_{dop}, the discrete dopants are approximated by Gaussian distributions

$$C_{\mathrm{dop}}(x) := \sum_j \frac{C_j}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{(x-x_j)^2}{2\sigma^2}\right),$$

where σ is the influence parameter, C_j is the charge of the j-th dopant atom, and x_j its position [86].

Figure 3.1 shows a cross section of the domain in the longitudinal direction. The longitudinal direction accounts for the transport of charge carriers through the nanowire (black meshes) connecting the source and drain contacts. At least two spatial dimensions are required for this type of problem: the potentials applied at the top and bottom require one dimension and the transport between the source and drain contacts requires another one. The drawback of 3D simulations is the large computational cost. In order to reduce the overall computational cost, a two-dimensional implementation was chosen for the numerical results presented here. A three-dimensional implementation would of course be a more faithful idealization of the three-dimensional reality, not leading to a constant, infinite extension of the two dimensions into three.

In order to solve the system of equations, we use Scharfetter-Gummel iteration. In spite of the quadratic convergence of Newton's method for the system, Scharfetter-Gummel iteration has advantages for the problem at hand. First of all, Scharfetter-Gummel iteration is much less sensitive to the choice of the initial guess than Newton's method. Another important feature is the reduced computational effort and memory requirement, since in each iteration it requires the successive solution of three much smaller elliptic problems.

The calculations are performed using MATLAB version 2015a on an Intel Core i5-4430 3.00 GHz 4-core processor with 8 GB of main memory.
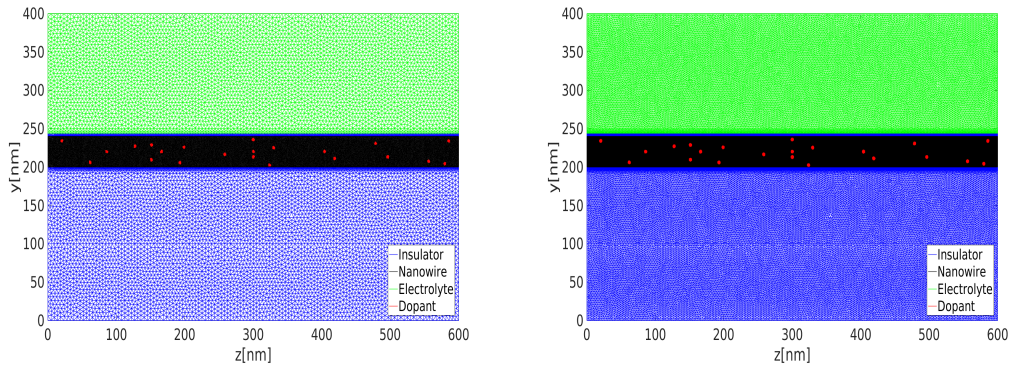
Figure 3.1.: Meshes for the random distribution of impurity atoms (red circles) in a nanowire field-effect sensor for levels $\ell = 0$ (left) and $\ell = 1$ (right), where $h_0 = 4.02$, $r = 2$, and $C_{\mathrm{dop}} = 4 \times 10^{16}\,\mathrm{cm}^{-3}$. Additionally, oxide subdomain (D_{ox}), transducer (D_{Si}) and the electrolyte (D_{liq}) are depicted with blue, black and green meshes respectively.
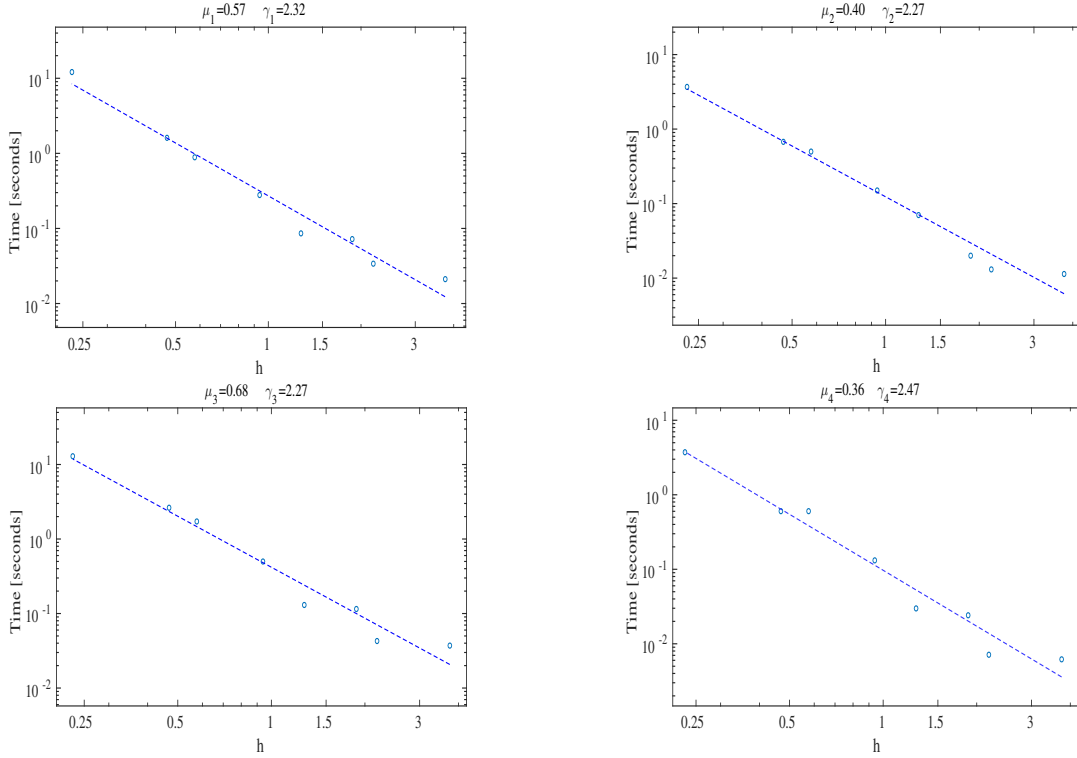
Figure 3.2.: Computational work for matrix assembly (top) and solving the system (bottom), both for the Poisson equation (left) and the drift-diffusion-equations (right).

3.5.2. The Computational Work

As the first step, we calculate the coefficients in the expressions (3.32) for the computational work. To that end, we solve the system for various mesh sizes and measure elapsed wall-clock time spent on matrix assembly and solving the resulting system, both for the Poisson equation and the drift-diffusion equations. Figure 3.2 shows the results for the coefficients in the expressions for the computational work.

The coefficients α and C_1 in the FE discretization error

$$\|\mathbb{E}[V - \hat{V}_h]\|_X + \|\mathbb{E}[u - \hat{u}_h]\|_X + \|\mathbb{E}[v - \hat{v}_h]\|_X \leq C_1 h^\alpha$$

of the system are given in Figure 3.3. The exponent $\alpha = 1.926$ found here agrees very well with the order of the discretization used here, i.e., P_1 finite elements.

For the statistical error, we determine the coefficients in the inequality

$$(\sigma[\Delta V_{h_0}] + \sigma[\Delta V_{h_\ell}]) + (\sigma[\Delta u_{h_0}] + \sigma[\Delta u_{h_\ell}]) + (\sigma[\Delta v_{h_0}] + \sigma[\Delta v_{h_\ell}]) \leq C_{00} + C_0 h_{\ell-1}^\beta.$$

Here $C_{00} = 0.07$ and the rest of the coefficients are shown in Figure 3.3.

3.5.3. Optimization

Having determined the coefficients in the expressions for the computational work, it is now possible to numerically solve the optimization problems. As described in Section 3.4, we

Figure 3.3.: Discretization error (left) and statistical (right) error as a function of h.

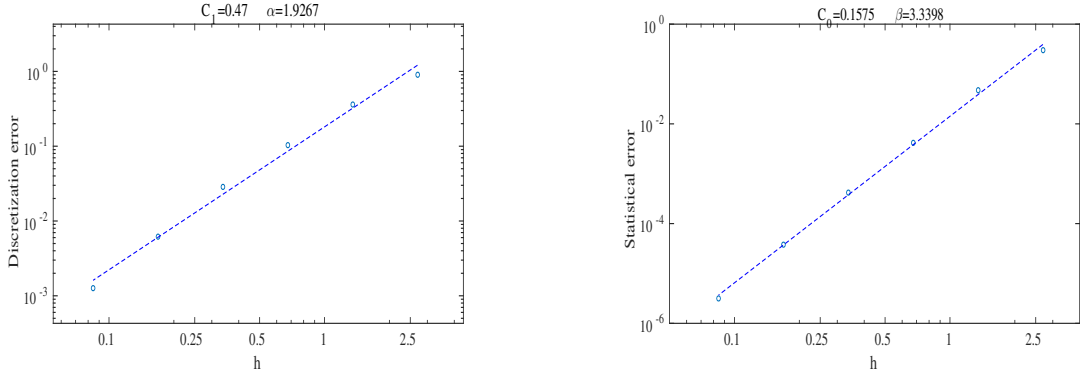apply an iterative interior-point method to optimize both the number of samples and mesh sizes.

Monte Carlo

First of all, we solve the optimization problem (3.34) for the MC-FE method. Because there is only one level, it is straightforward to solve. The optimal values for the MC FE method are summarized in Table 3.1 for given ε.

ε	0.1	0.05	0.02	0.01	0.005	0.002	0.001
h	0.348	0.243	0.151	0.105	0.074	0.046	0.032
M	12	46	282	1 130	4 519	28 268	113 130

Table 3.1.: Optimal MC FE method parameters for various given error tolerances.

Multi-Level Monte Carlo

In the MLMC-FE method, determining the optimal number of levels is an important part of the calculation. This is achieved here by solving the optimization problem for several levels starting with a single level and noting that the computational work increases above a certain number of levels. More precisely, we solve the optimization problem (3.39) for $0 \le L \le 7$ levels as well as for various given error bounds.

Since the number of samples in each level is a continuous variable in the optimization problem, the optimal number of samples is – in general – not an integer and hence we choose M_ℓ, $\ell = 0, \ldots, L$, as the final numbers of levels.

The results of the optimization problems provide insight into the MLMC procedure. Figure 3.4 shows the minimized computational work as a function of the number of levels and as a function of the given tolerance. It shows that for smaller tolerances ε, a larger L is required.

In Figure 3.5, the two approaches to multi-level Monte Carlo are compared, namely choosing the h_ℓ as a geometric progressions or freely. Due to generality of the second option,
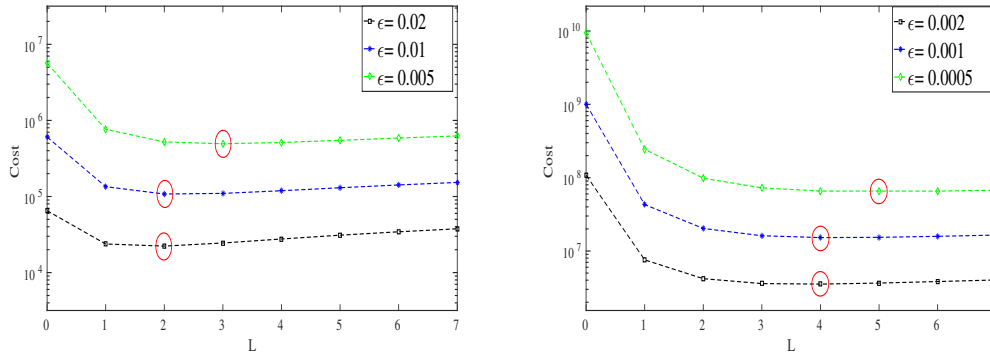
Figure 3.4.: The minimized computational work for the MLMC-FE method as a function of the number of levels and as a function of the given error tolerance. The results for a geometric progression for h (left) and general h (right) are shown. The number of levels yielding the minimal overall computational work is indicated by red circles.
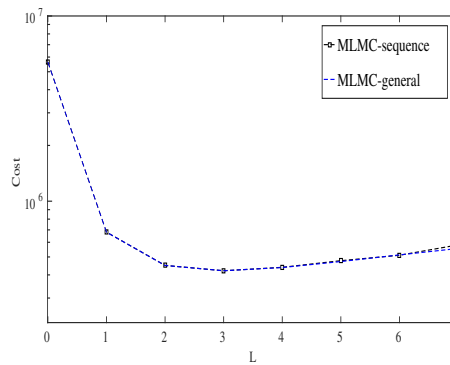
Figure 3.5.: Comparison between the two different approaches to MLMC FE method for $\varepsilon = 0.015$.

the total work when choosing the h_ℓ freely is lower compared to the first option. However, only a small reduction in computational cost is achieved by the choosing meshes freely. The results for both approaches to MLMC-FEM are summarized in Tables 3.2 and 3.3 for various given error tolerances. Both figures show additionally that more than two levels (i.e., $L > 2$) only yield a relatively small reduction in computational cost even for small tolerance levels. In practice, it should hence be considered that the interior-point method requires more time as the number of levels increases.

Comparison

Finally, as Figure 3.6 shows, the computational work for the multi-level Monte-Carlo method is approximately two times lower than the one for the Monte-Carlo method for larger tolerance levels such as $\varepsilon = 0.1$. The effectiveness of the MLMC-FE method is more

ε	h_0	r	M_0	M_1	M_2	M_3	M_4
0.1	0.5171	1.7381	16	4	–	–	–
0.05	0.4433	2.1319	63	9	–	–	–
0.02	0.4620	2.0079	471	70	11	–	–
0.01	0.4222	2.2940	1 882	207	22	–	–
0.005	0.4515	2.0765	8 534	1 167	161	23	–
0.002	0.4549	2.0033	58 203	8 536	1296	197	30
0.001	0.4407	2.1669	232 299	29 141	3 575	438	54

Table 3.2.: Optimal levels for the MLMC-FE method with h_ℓ chosen as a geometric progression for given error tolerances ε.

ε	h_0	r_1	r_2	r_3	r_4
0.1	0.5171	1.738	–	–	–
0.05	0.4433	2.131	–	–	–
0.02	0.4618	2.020	1.990	–	–
0.01	0.4201	2.270	2.320	–	–
0.005	0.4507	2.080	2.070	2.060	–
0.002	0.4587	2.038	2.017	1.990	1.950
0.001	0.4412	2.140	2.157	2.170	2.196

ε	M_0	M_1	M_2	r_3	r_4
0.1	16	4	–	–	–
0.05	63	9	–	–	–
0.02	471	69	11	–	–
0.01	1 884	210	22	–	–
0.005	8 531	1 162	160	22	–
0.002	58 100	8 327	1 221	185	29
0.001	232 539	29 564	3701	455	55

Table 3.3.: Optimal levels for the MLMC-FE method with general h_ℓ for given error tolerances ε.

pronounced for smaller error bounds; for $\varepsilon = 0.001$, the computational work is about a factor 10^2 lower than the Monte-Carlo work. The results agree with Giles' standard complexity theorem [30] in the sense that the estimated exponents α, β, and γ satisfy the assumption of the theorem, i.e., $\alpha \geq \frac{1}{2}\min(\beta, \gamma)$. Therefore, according to the theorem, the computational cost of the MLMC-FEM is $O(\varepsilon^{-2})$. Additionally, according to Figure 3.6, the total cost of the MC-FEM is $O(\varepsilon^{-2-\gamma/\alpha})$, which agrees with [32]. The optimal distribution of the samples among the levels in the multi-level method leads to more evaluations in the first levels (which are cheaper) and to fewer evaluations in the higher levels. On the other hand, to satisfy the first constraint of (3.34), the Monte-Carlo method needs a smaller mesh size compared to the multi-level method, which greatly increases the total computational work although the total number of samples is lower.
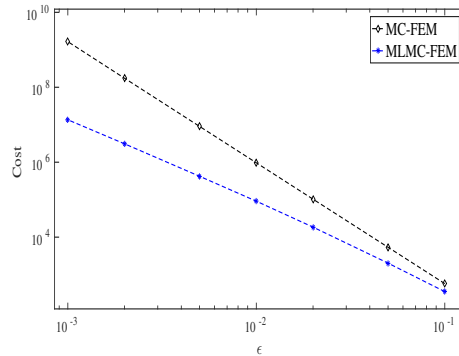
Figure 3.6.: Comparison of total computational work for MC-FEM and the two approaches to MLMC-FEM for various given tolerances.

3.6. Summary

In this work, we considered the stochastic drift-diffusion-Poisson equations as the main model equation for describing transport in random environments with many applications. We presented existence and local uniqueness theorems for the weak solution of the system. We also developed MC- and MLMC-FE methods for this system of stochastic PDEs.

Additionally, we balanced the various parameters in the numerical methods by viewing this problem as a global optimization problem. The goal is to determine the numerical parameters such that the computational work to achieve a total error, i.e., discretization error plus statistical error, less than or equal to a given error tolerance is minimized.

Although the exponential terms in the constraints make the optimization problems nonlinear, the optimization problems can be solved by an interior-point method with sufficient iterations. The solution of the constrained optimization problem leads to optimal (M, h) in the case of the vanilla MC method and to hierarchies consisting of $(L, \{M_\ell\}_{\ell=0}^{L}, h_0, r)$ in the case of the MLMC method.

Moreover, we investigated two different options to the mesh refinement in the multi-level method. Although less computational effort is needed by choosing the mesh sizes freely, the difference is negligible. In the comparison of the MC with the MLMC method, the MLMC method was found to decrease the total computational effort by four orders of magnitude for small error tolerances. The speed-up becomes better as the error tolerance decreases.

3.7. Randomized Quasi-Monte-Carlo Method

In the second part of this chapter, we develop another optimal method, namely optimal multilevel randomized quasi Monte-Carlo method for approximating the expected value of the solution of the stochastic drift-diffusion-Poisson system. Here we use mathematical model (3.2), which models nanoscale devices as described in Section 3.1 in detail.

In this section, first a randomized quasi-Monte-Carlo method is developed and serves as a basic building block for the multilevel method developed in the next section. As aforementioned, the MLRQMC method was first introduced in [28] and developed for a

PDE in [35]. Here, the multi-level and QMC ideas are applied to a system of equations, and based thereon, an optimal method is developed.

3.7.1. The Koksma-Hlawka Inequality

In quasi-Monte-Carlo (QMC) methods, quasi-random sequences are used instead of random sequences. Correlations between the points provide greater uniformity and speedup the computations and therefore the convergence rate, which is generally of a higher order than that of the standard Monte-Carlo method. This approach avoids the problem of clumping in the standard Monte-Carlo method; about $\sqrt{N}$ out of N points lie in clumps in the standard Monte-Carlo method [87]. The reason for clumping in the standard Monte-Carlo method is, of course, the independence of the random points, while the points are correlated in quasi-Monte-Carlo methods and thus clumping is avoided. The convergence rate of QMC methods is $\mathcal{O}((\log N)^d N^{-1})$, where d is the dimension of the random variable. Therefore, QMC methods have a smaller error and converge faster than the standard Monte-Carlo method. Still, a large number of dimensions limits the effectiveness of quasi-Monte-Carlo sequences [35, 87].

It is convenient to describe QMC methods in the context of numerical quadrature rules. QMC methods approximate an integral on a d-dimensional unit cube by an N-point equal-weight quadrature rule of the form

$$\int_{[0,1]^d} f(\omega)\mathrm{d}\omega \approx \frac{1}{N}\sum_{j=1}^{N} f(\omega_j). \tag{3.46}$$

Rather than choosing the points ω_j uniformly from the unit cube, as is the case with the standard Monte-Carlo method, QMC methods choose the points in a deterministic manner.

The basis for analyzing QMC quadrature error is the Koksma-Hlawka inequality.

Theorem 5 (Koksma-Hlawka theorem)**.** *For any sequence $\{\omega_j\}_{j\geq 1}$ and any function f with bounded variation, the integration error due to (3.46) is bounded by*

$$\left|\frac{1}{N}\sum_{j=1}^{N} f(\omega_j) - \int_{[0,1]^d} f(\omega)\mathrm{d}\omega\right| \leq V_{\mathrm{HK}}(f)D_N^*(\omega), \tag{3.47}$$

where $V_{\mathrm{HK}}(f)$ is the Hardy-Krause variation of f defined by

$$V_{\mathrm{HK}}(f) := \int_{[0,1]^d} \left|\frac{\partial^d f}{\partial \omega^1 \cdots \partial \omega^d}\right|\mathrm{d}\omega \tag{3.48}$$

for sufficiently differentiable f.

The first factor $V_{\mathrm{HK}}(f)$ in (3.47) is the variation of f in the sense of Hardy and Krause [88]. This term measures the variability of the function values, whereas the discrepancy term $D_N^*(\omega_j)$, the second factor, measures the variability of the underlying sequence, i.e., the quadrature nodes, from the ideal distribution.

Unfortunately, QMC methods have drawbacks as well. In fact, when the dimension d is too large, the calculation of the integral in (3.46) is computationally extremely expensive.

In other words, for large d, the number N of samples has to be considerably large for $(\log N)^d N^{-1}$ to be smaller than $N^{-1/2}$. In [89], it is proved that there exist lattice rules such that the optimal rate of convergence for QMC rules is $\mathcal{O}(N^{-\alpha/2+\delta})$ for any $\delta > 0$ and with a parameter $\alpha > 1$. This convergence rate is independent of the dimension d. In [90], Kuo showed that there exist shifted rank-1 lattice rules (constructed by the CBC algorithm) that achieve the optimal convergence of $\mathcal{O}(N^{-1+\delta})$ for any $\delta > 0$. The value of δ depends on the problem and is estimated in Section 3.10 below.

Furthermore, $V_{\mathrm{HK}}(f)$ and D_N^* are difficult to compute. In order to overcome these difficulties, randomized QMC (RQMC) methods have been developed [91].

The accuracy of a QMC method can be improved by rewriting the function so that the variation term is reduced [92] or by constructing sequences that have smaller discrepancy [93]. Using RQMC methods with very low discrepancy sequences such as rank-1 lattice rules helps to increase accuracy and gives a useful error bound.

3.7.2. Randomized Quasi-Monte-Carlo Finite-Element Method (RQMC-FE-M)

In order to analyze and estimate the variance and to find an error estimate, QMC methods can be randomized. Randomized quasi-Monte-Carlo (RQMC) methods can also be considered as a variance reduction technique for the standard Monte-Carlo method. The simplest method of randomizing is to use a uniformly distributed d-dimensional shift $\Delta \sim U[0,1]^d$.

In particular, a randomized rank-1 lattice rule [94] can be constructed as

$$\omega_j^{(i)} := \frac{j}{N}\lambda + \Delta^{(i)} \mod 1, \quad j \in \{1,\dots,N\}, \quad i \in \{1,\dots,M\}, \tag{3.49}$$

where N is the number of quasi-random points, $\Delta \in [0,1]^d$ is the random shift, which is uniformly distributed over $[0,1]^d$, M is the number of random shifts, and $\lambda \in \mathbb{R}^d$ is a d-dimensional deterministic generating vector. Choosing λ carefully is important in order to achieve uniformity. The quality of a randomly shifted lattice rule is determined by the choice of the generating vector λ. This essential question is addressed, e.g., in [95, Section 4].

If the system (3.2) has a solution (V, u, v), we denote finite-element numerical approximations by $(V_h(x,\omega), u_h(x,\omega), v_h(x,\omega))$ for a given $\omega \in \Omega$. Since all three components of the solution are in $H^1(D)$ for a given $\omega \in \Omega$, the variable u may denote any of the three components from now on to simplify notation. We define the Hilbert space

$$X := H_g^1(D) = \{u \in H^1(D) \mid Tu = g\} \tag{3.50}$$

as the solution space, where T is the trace operator defined such that $Tu = g$, where g is Dirichlet lift of $u_D := u|_{\partial D_D}$. The operator T is well-defined and continuous from $H^1(D)$ onto $H^{1/2}(\partial D)$ for the Lipschitz domain D.

Having chosen a finite-element mesh τ_h and having fixed $k \in \mathbb{N}$ with $k \geq 1$, the space

$$X_h := \mathcal{P}^k(\tau_h) := \{u \in X \mid u|_K \in \mathcal{P}^k(K) \ \forall K \in \tau_h\} \subset X \tag{3.51}$$

is the discretization space, where $\mathcal{P}^k(K) := \mathrm{span}\{x^\alpha \mid |\alpha| \leq k\}$ is the space of polynomials of total degree less than or equal to k. The expected value of the solution u is the integral

$$\mathbb{E}[u] = \int_{[0,1]^d} u(x,\omega)\mathrm{d}\mathbb{P}(\omega). \tag{3.52}$$

63

The RQMC estimator to approximate $\mathbb{E}[u_h]$ is then defined by

$$Q_{N,M}(u_h) := \frac{1}{M} \sum_{i=1}^{M} \frac{1}{N} \sum_{j=1}^{N} u_h(x, \omega_j^{(i)}) \tag{3.53}$$

using the quasi-random points defined in (3.49).

3.7.3. Error Bound for the RQMC-FE Method

As aforementioned, in order to overcome the difficulty of finding an error bound for the QMC approach, we use a RQMC method. In this method, the standard assumption is that u_h has bounded variation $V_{\mathrm{HK}}(u_h)$ in the sense of Hardy and Krause and behaves like the variance [40]. Therefore, we assume that

$$V_{\mathrm{HK}}(u_h) \leq C_0, \tag{3.54}$$

where C_0 is a positive constant. Similar to the standard MC method, the mean square error (MSE) can be written as the sum of the variance of the estimator plus the square of the discretization error [36]. As in [36], a prescribed accuracy is to be achieved, i.e., MSE $\leq \varepsilon^2$.

Using the Koksma-Hlawka inequality (3.47) by calculating

$$
\begin{aligned}
\sigma^2[Q_{N,M}(u_h)] &= \sigma^2 \left[\frac{1}{M} \sum_{i=1}^{M} \frac{1}{N} \sum_{j=1}^{N} u_h(x, \omega_j^{(i)}) \right] \\
&= \frac{1}{M} \sigma^2 \left[\frac{1}{N} \sum_{j=1}^{N} u_h(x, \omega_j) \right] \\
&= \frac{1}{M} \int_{[0,1]^d} \left(\frac{1}{N} \sum_{j=1}^{N} u_h(x, \omega_j) - \mathbb{E}[u_h] \right)^2 \mathrm{d}\mathbb{P}(\omega) \\
&\leq \frac{1}{M} \int_{[0,1]^d} \left(V_{\mathrm{HK}}(u_h) D_N^*(\omega_j) \right)^2 \mathrm{d}\mathbb{P}(\omega) \\
&= \mathcal{O}(V_{\mathrm{HK}}^2(u_h) N^{-2+\delta}) \qquad \forall \delta > 0,
\end{aligned}
\tag{3.55}
$$

the variance of the RQMC estimator (3.53) is estimated following [40]. In fact, in rank-1 lattice rules, the discrepancy satisfies

$$D_N^*(\omega_j) = \mathcal{O}(N^{-1+\delta}) \qquad \forall \delta > 0 \tag{3.56}$$

for any number of points $N > 1$, any shift of the lattice, and for any dimension $d \geq 1$ [96]. The above result is obtained by using component-by-component (CBC) construction, i.e., the components of the generating vector λ are constructed one at a time to minimize the worst-case error in certain weighted function spaces [94].

Using the boundedness assumption (3.54) for $V_{\mathrm{HK}}(u_h)$ in (3.55), we obtain the estimate

$$\sigma^2[Q_N(u_h)] \leq C_0 M^{-1} N^{-2+\delta} \qquad \forall \delta > 0 \tag{3.57}$$

for the variance of the RQMC method, where C_0 is estimated using (3.54). Furthermore, δ is estimated in Section 3.10 (see Figure 3.7). The inequality (3.57) will be used later for an error estimate.

The variable u may represent any of the three components of the solution (V, u, v) of the system (3.2) in order to simplify notation, since all three components are in $H^1(D)$ for a given $\omega \in \Omega$.

Proposition 6. *Suppose that $Q_{N,M}(u_h)$ is the RQMC estimator to approximate the expectation $\mathbb{E}[u]$ of the solution $u(x,\omega) \in X$ of (3.2). Assume further that the spatial discretization error converges with order α, i.e.,*

$$\|\mathbb{E}[u - u_h]\|_{L^2(\Omega;X)} \leq C_1 h^\alpha \qquad \exists C_1 \in \mathbb{R}^+, \qquad (3.58)$$

where $u_h(x,\omega) \in X_h$ is the FE approximation with mesh size h and it has bounded variation. Then the mean square error of the RQMC estimator $Q_{N,M}$ satisfies

$$\|Q_{N,M}(u_h) - \mathbb{E}[u]\|_{L^2(\Omega;X)}^2 = \mathcal{O}(M^{-1}N^{-2+\delta}) + \mathcal{O}(h^{2\alpha}) \qquad \forall \delta > 0. \qquad (3.59)$$

Proof. We estimate the mean square error (MSE). Using inequality (3.57) and assumption (3.58), we find that

$$
\begin{aligned}
\text{MSE} &:= \|Q_{N,M}(u_h) - \mathbb{E}[u]\|_{L^2(\Omega;X)}^2 \\
&= \|Q_{N,M}(u_h) - \mathbb{E}[Q_{N,M}(u_h)]\|_{L^2(\Omega;X)}^2 + \|\mathbb{E}[Q_{N,M}(u_h)] - \mathbb{E}[u]\|_{L^2(\Omega;X)}^2 \\
&= \sigma^2[Q_{N,M}(u_h)] + \|\mathbb{E}[u - u_h]\|_{L^2(\Omega;X)}^2 \\
&\leq C_0 M^{-1} N^{-2+\delta} + C_1 h^{2\alpha} \\
&= \mathcal{O}(M^{-1}N^{-2+\delta}) + \mathcal{O}(h^{2\alpha})
\end{aligned}
\qquad (3.60)
$$

for every $\delta > 0$. $\qquad\qquad\square$

This means that the error behaves like (3.57).

3.8. Multilevel Randomized Quasi-Monte-Carlo Finite-Element Method (MLRQMC-FE Method)

Based on the RQMC method in the previous section, a multilevel version is summarized here.

3.8.1. The Levels

In a multilevel approach, several mesh levels are used, and on each level, the RQMC estimator (3.53) is employed to approximate the solution. The domain D is partitioned into quasi-uniform triangles such that sequences $\{\tau_{h_\ell}\}_{\ell=0}^L$ of regular meshes are obtained. For any $\ell \geq 0$, we denote the mesh size of τ_{h_ℓ} by

$$h_\ell := \max_{K \in \tau_{h_\ell}} \operatorname{diam}(K).$$

Uniform refinement of the mesh to obtain a nested family $\{\tau_{h_\ell}\}_{\ell=0}^{\infty}$ of regular triangulations can be achieved by regular subdivision yielding the mesh size

$$h_\ell = r^{-\ell}h_0, \tag{3.61}$$

where h_0 denotes the mesh size of the coarsest triangulation and $r > 1$ is independent of ℓ. In this method, the finite-dimensional sequence $X_{h_0} \subset X_{h_1} \subset \cdots \subset X_{h_L} \subset X$ of discretization spaces are used, where $X_{h_\ell} := \mathbb{P}^k(\tau_{h_\ell})$ with $\ell \in \{0, 1, 2, \ldots, L\}$, and $k \in \mathbb{N}$ is fixed (cf. (3.51)).

The finite-element approximation at level L can be written as the telescoping sum

$$u_{h_L} = u_{h_0} + \sum_{\ell=1}^{L} (u_{h_\ell} - u_{h_{\ell-1}}),$$

where u_{h_ℓ} is the approximation on the mesh τ_{h_ℓ} at level ℓ. Furthermore, $\mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}]$ can be estimated using N_ℓ quasi-random points and M_ℓ random shifts on each level ℓ. Therefore the multilevel RQMC FE estimator with respect to one or more random shift is defined as

$$Q_{L,N_\ell,M_\ell}(u_{h_L}) := \frac{1}{M_0}\sum_{i=1}^{M_0}\frac{1}{N_0}\sum_{j=1}^{N_0} u_{h_0}(x,\omega_j^{(i)}) + \sum_{\ell=1}^{L}\frac{1}{M_\ell}\sum_{i=1}^{M_\ell}\frac{1}{N_\ell}\sum_{j=1}^{N_\ell}\left(u_{h_\ell}(x,\omega_j^{(i)}) - u_{h_{\ell-1}}(x,\omega_j^{(i)})\right).$$
$$\tag{3.62}$$

The sample points $\omega_j^{(i)}$ are obtained using (3.49), for example, and their total number is $M_\ell N_\ell$.

3.8.2. Error Bound for the MLRQMC-FE Method

In order to state a proposition for the mean square error of the multilevel RQMC approximation, we first make the following necessary assumptions.

Assumptions 6. *The assumptions on the boundedness of the variations of the FEM approximation and on the convergence order of the discretization error are*

1. $V_{\mathrm{HK}}(u_{h_0}) \leq C_{00} \quad \exists C_{00} \in \mathbb{R}^+,$

2. $V_{\mathrm{HK}}(u_{h_\ell} - u_{h_{\ell-1}}) \leq C_0 h_{\ell-1}^{\beta} \quad \exists C_0 \in \mathbb{R}^+, \exists \beta \in \mathbb{R}^+,$

3. $\|\mathbb{E}[u - u_{h_\ell}]\|_{L^2(\Omega;X)} \leq C_1 h_\ell^{\alpha} \quad \exists C_1 \in \mathbb{R}^+, \exists \alpha \in \mathbb{R}^+.$

By using the multilevel approach, the difference between u_{h_ℓ} and $u_{h_{\ell-1}}$ decreases for higher levels and therefore $V_{\mathrm{HK}}(u_{h_\ell})$ is reduced. Hence, it is a decent assumption that the Hardy-Krause variation behaves similarly to the variance of $u_{h_\ell} - u_{h_{\ell-1}}$.

Proposition 7. *Suppose Assumptions 6 hold and $Q_{L,N_\ell,M}(u_{h_L})$ is the multilevel randomized quasi-Monte-Carlo estimator with $N_\ell M_\ell$ sample points in level ℓ, $\ell \in \{0, 1, 2, \ldots, L\}$, to approximate the expectation $\mathbb{E}[u]$ of the solution $u(\cdot, \omega) \in X$ of (3.2) using FEM approximations $u_{h_\ell}(\cdot, \omega) \in X_{h_\ell}$ with mesh size h_ℓ.*

Then the mean square error of the multilevel RQMC estimator satisfies

$$\|\mathbb{E}[u] - Q_{L,N_\ell,M_\ell}(u_{h_L})\|^2_{L^2(\Omega;X)} = \mathcal{O}(h_L^{2\alpha}) + \mathcal{O}(M_0^{-1}N_0^{-2+\delta}) + \sum_{\ell=1}^{L}\mathcal{O}(h_{\ell-1}^{2\beta}M_\ell^{-1}N_\ell^{-2+\delta}) \qquad \forall \delta > 0.$$

(3.63)

Proof. First, following [40], we estimate the variance of the multilevel RQMC estimator using inequality (3.47) by calculating

$$\sigma^2[Q_{L,N_\ell,M_\ell}(u_{h_L})] = \sigma^2\Big[\frac{1}{M_0}\sum_{i=1}^{M_0}\frac{1}{N_0}\sum_{j=1}^{N_0}u_{h_0}(x,\omega_j^{(i)}) + \sum_{\ell=1}^{L}\frac{1}{M_\ell}\sum_{i=1}^{M_\ell}\frac{1}{N_\ell}\sum_{j=1}^{N_\ell}\big(u_{h_\ell}(x,\omega_j^{(i)}) - u_{h_{\ell-1}}(\omega_j^{(i)})\big)\Big]$$

$$= \frac{1}{M_0}\sigma^2\Big[\frac{1}{N_0}\sum_{j=1}^{N_0}u_{h_0}(x,\omega_j)\Big] + \sum_{\ell=1}^{L}\frac{1}{M_\ell}\sigma^2\Big[\frac{1}{N_\ell}\sum_{j=1}^{N_\ell}\big(u_{h_\ell}(x,\omega_j) - u_{h_{\ell-1}}(x,\omega_j)\big)\Big]$$

$$= \frac{1}{M_0}\int_{[0,1]^d}\Big(\frac{1}{N_0}\sum_{j=1}^{N_0}u_{h_0}(x,\omega_j) - \mathbb{E}[u_{h_0}]\Big)^2\mathrm{d}\mathbb{P}(\omega)$$

$$+ \sum_{\ell=1}^{L}\frac{1}{M_\ell}\int_{[0,1]^d}\Big(\frac{1}{N_\ell}\sum_{j=1}^{N_\ell}\big(u_{h_\ell}(x,\omega_j)$$

$$- u_{h_{\ell-1}}(x,\omega_j)\big) - \mathbb{E}[u_{h_\ell}(x,\omega_j) - u_{h_{\ell-1}}(x,\omega_j)]\Big)^2\mathrm{d}\mathbb{P}(\omega)$$

$$\leq \frac{1}{M_0}\int_{[0,1]^d}\Big(V_{HK}(u_{h_0})D_{N_0}^*(\omega_j)\Big)^2\mathrm{d}\mathbb{P}(\omega)$$

$$+ \sum_{\ell=1}^{L}\frac{1}{M_\ell}\int_{[0,1]^d}\Big(V_{HK}\big(u_{h_\ell}(x,\omega_j) - u_{h_{\ell-1}}(x,\omega_j)\big)D_{N_\ell}^*(\omega_j)\Big)^2\mathrm{d}\mathbb{P}(\omega)$$

$$= \mathcal{O}(V_{HK}^2(u_{h_0})N_0^{-2+\delta}) + \mathcal{O}(V_{HK}^2\big(u_{h_\ell}(x,\omega_j) - u_{h_{\ell-1}}(x,\omega_j)\big)N_\ell^{-2+\delta}),$$

(3.64)

where we used the estimate (3.56).

Therefore, we have

$$\sigma^2[Q_{L,N_\ell,M_\ell}(u_{h_L})] \leq C_{00}M_0^{-1}N_0^{-2+\delta} + C_0\sum_{\ell=1}^{L}h_{\ell-1}^{2\beta}M_\ell^{-1}N_\ell^{-2+\delta},$$

(3.65)

using the assumptions of bounded variations, i.e., Assumptions 6.1 and 6.2. This estimate shows how the error of the method behaves in terms of the number of samples (same as (3.57)), as we will see in the following.

Similarly to the RQMC estimator, the MSE assesses the accuracy of the MLRQMC-FE

estimator. Using Assumptions 6.3 and the variance estimate (3.64), we find

$$
\begin{aligned}
\mathrm{MSE} &:= \|Q_{L,N_\ell,M_\ell}(u_{h_L}) - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= \|Q_{L,N_\ell,M_\ell}(u_{h_L}) - \mathbb{E}[Q_{L,N_\ell,M_\ell}(u_{h_L})]\|^2_{L^2(\Omega;X)} + \|\mathbb{E}[Q_{L,N_\ell,M_\ell}(u_{h_L})] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= \sigma^2[Q_{L,N_\ell,M_\ell}(u_{h_L})] + \|\mathbb{E}[u - u_{h_L}]\|^2_{L^2(\Omega;X)} \\
&\leq C_{00}M_0^{-1}N_0^{-2+\delta} + C_0 \sum_{\ell=1}^{L} h_{\ell-1}^{2\beta}M_\ell^{-1}N_\ell^{-2+\delta} + (C_1 h_L^\alpha)^2 \\
&= \mathcal{O}(M_0^{-1}N_0^{-2+\delta}) + \sum_{\ell=1}^{L} \mathcal{O}(h_{\ell-1}^{2\beta})\mathcal{O}(M_\ell^{-1}N_\ell^{-2+\delta}) + \mathcal{O}(h_L^{2\alpha})
\end{aligned}
$$

(3.66)

for every $\delta > 0$. □

3.9. Optimal Multilevel Randomized Quasi-Monte-Carlo Method

Since both the spatial and the stochastic dimensions are to be discretized, the question how to distribute the computational work between the spatial and stochastic dimensions poses itself. In other words, various parameters in the numerical approaches outlined so far must still be determined. These parameters include the mesh size of the finite-element discretization, the number of levels in the multilevel approach, and the samples to be used on each level. Because of the computational challenge of solving a system of stochastic partial differential equations, efficient computational methods are crucial. Therefore, we develop an optimal method based on the previous section here.

The error bound found in Proposition 7 is used as an estimate of the total error. The total error is prescribed and the unknown parameters are chosen such that the computational work is minimized. Hence the computational work must be modeled as a function of the unknown parameters. It consists of the sum of work necessary to solve each of three equations in the coupled system, i.e., the total computational work is given by

$$
W := W_{P,a} + W_{P,s} + 2W_{D,a} + 2W_{D,s}, \tag{3.67}
$$

where the index P denotes the work due to the Poisson equation and the index D denotes the work due to the two drift-diffusion equations. Furthermore, since the steps for solving each equation exhibit different scaling of the computational cost, there are separate terms for the computational work for assembling the system matrices (index a) and for solving the resulting systems (index s). Each of four work terms above has the form of $\mu_k h_\ell^{-\gamma_k}$, where the constants $\mu_k > 0$ and $\gamma_k > 0$ depend on the implementation and will be measured. If an appropriate linear solver is used to calculated the finite-element approximation u_{h_ℓ}, then we expect that $\gamma_k \approx n$ holds, where n is the number of spatial dimension [97]. In Section 3.10, we will see that our numerical results agree with this estimate.

Using this model for the computational work for one sample, the total work for solving

the system (3.2) is modeled as

$$
\begin{aligned}
W &:= \sum_{\ell=0}^{L} M_\ell N_\ell W_\ell \\
&= \sum_{\ell=0}^{L} M_\ell N_\ell (W_{\ell,P,a} + W_{\ell,P,s} + 2W_{\ell,D,a} + 2W_{\ell,D,s}) \\
&= \sum_{\ell=0}^{L} M_\ell N_\ell (\mu_1 h_\ell^{-\gamma_1} + \mu_2 h_\ell^{-\gamma_2} + \mu_3 h_\ell^{-\gamma_3} + \mu_4 h_\ell^{-\gamma_4}).
\end{aligned} \tag{3.68}
$$

Having modeled the computational work, we can now state the optimization problem in the sense that we want to minimize the total computational work for a prescribed error tolerance ε. The minimization problem is

$$
\underset{M_\ell, N_\ell, h_0, r}{\text{minimize}} \quad f(M_\ell, N_\ell, h_0, r, L) := \sum_{\ell=0}^{L} M_\ell N_\ell \sum_{k=1}^{4} \mu_k h_0^{-\gamma_k} r^{\gamma_k}
$$

$$
\text{subject to} \quad g(N_\ell, h_0, r, L) := C_{00} M_0^{-1} N_0^{-2+\delta} + C_0 \sum_{\ell=1}^{L} h_{\ell-1}^{2\beta} M_\ell^{-1} N_\ell^{-2+\delta} + (C_1 h_L^\alpha)^2 \le \varepsilon^2,
$$

$$
\tag{3.69}
$$

for every $\delta > 0$, where $h_0 > 0$, $r > 1$, N_ℓ, and $N_\ell \ge 1$. The given maximal total error ε^2 is an upper bound for (3.66), i.e., MSE $\le \varepsilon^2$. The goal is to determine optimal values h_ℓ (by calculating optimal values for h_0 and r and using their relation (3.61)) and N_ℓ, $\ell \in \{0, 1, \dots, L\}$. For all levels, the number M_ℓ of shift realizations is an integer, i.e., $M_\ell \in \mathbb{N}$.

This nonlinear constrained optimization problem can be solved numerically; the nonlinearity of the constraints g and the objective function f due to the exponents motivates the use of sequential quadratic programming (SQP) [98] as a generalization of Newton's method for unconstrained optimization. The method is iterative and solves quadratic subproblems; it can be used in both the line-search and trust-region frameworks. SQP is well-suited for solving problems with significant nonlinearities.

We denote the parameters found in step s, $s \in \mathbb{N}$, by $\chi_s := (N_{\ell,s}, h_{0,s}, r_s, L_s)$. In each iteration, χ_s is found by solving a quadratic programming (QP) subproblem, whose solution is then used in the next iteration. The subproblems are of course constructed such that the sequence χ_s converges to a local minimum χ as $s \to \infty$. The QP subproblems are based on a quadratic approximation of the Lagrangian function

$$
\mathcal{L}(\chi, \zeta) := f(\chi) + \sum_{i=1}^{m} \zeta^\top g_i(\chi),
$$

where the vector ζ contains the Lagrange multipliers. In order to solve the optimization problem (3.69), the objective function is replaced by its local quadratic approximation

$$
f(\chi) \approx f(\chi_s) + \nabla f(\chi_s)(\chi - \chi_s) + \frac{1}{2}(\chi - \chi_s) H f(\chi_s)(\chi - \chi_s),
$$

where H is the Hessian matrix. The term $f(\chi^s)$ in the expression above can be eliminated from the minimization problem since it is constant. The nonlinear constraint g is replaced by its linearization $g(\chi) \approx g(\chi_s) + \nabla g(\chi_s)(\chi - \chi_s)$.

Hence the minimization problem (3.69) yields the simplified, linearized QP subproblems

$$\text{minimize} \quad \frac{1}{2}\nu(\chi)^\top H f(\chi_s)\nu(\chi) + \nabla f(\chi_s)^\top \nu(\chi)$$
$$\text{subject to} \quad \nabla g\left(\chi_s\right)^\top \nu(\chi) + g(\chi_s) \leq 0, \tag{3.70}$$

where $\nu(\chi) = \chi - \chi_s$. The next approximation is given by

$$\chi_{s+1} := \chi_s + \alpha_s \nu_s,$$

where ν_s is obtained by (3.70) and the step-length parameter α_s is determined by line search [99]. Also, H can be updated by any of the quasi-Newton methods, e.g., by the BFGS method [100].

3.10. Numerical Results

A numerical example is presented here in order to illustrate the advantages of the method developed in Section 3.9. The well-known deterministic version of the model equations, namely the drift-diffusion-Poisson system, describes charge transport in many situations; the stochastic version makes it possible to describe charge transport in random environments.

3.10.1. Computational Cost

As discussed in Section 3.9, the optimal parameters are found by solving the minimization problem that minimizes the computational work for a prescribed total error. This procedure yields the mesh sizes and numbers of samples in the multilevel approach. Before the minimization problem can be solved, the constants and the exponents in (3.69) must be measured.

As already mentioned, the statistical error depends on the mesh size h and the number N of samples. Figure 3.7 (left) depicts the error for different mesh sizes ($h_0 = 5$, $r = 2$, and $N = 100$) with a decay of variance of the order $\beta = 1.652$. Shifted rank-1 lattice rules give rise to the convergence rate $\mathcal{O}(N^{-2+\delta})$ for a $\delta > 0$. However, the value of δ is crucial for the optimization problem. As seen in the figure, the variance of MLRQMC-FEM decays as $\mathcal{O}(N^{-1.88})$ (i.e., $\delta = 0.12$), while in the case of MC-FEM a rate of $\mathcal{O}(N^{-1})$ is achieved. These values are obtained using $h = 5$ with respect to different numbers of quasi points. Additionally, Figure 3.8 illustrates the discretization error for different mesh sizes, where the parameters were estimated using 100 samples by comparing the variance of the multilevel estimator (3.62) for different mesh sizes. The numerically determined exponent $\alpha = 1.731$ agrees very well with the order of the P_1 FE discretization used here. The coefficients in the model for the computational work were also found numerically. For matrix assembly and solving the system, we recorded the CPU time used as a function of different mesh sizes, and hence the values of μ_k and γ_k are found. A summary of the coefficients and exponents is given in Table 3.4.
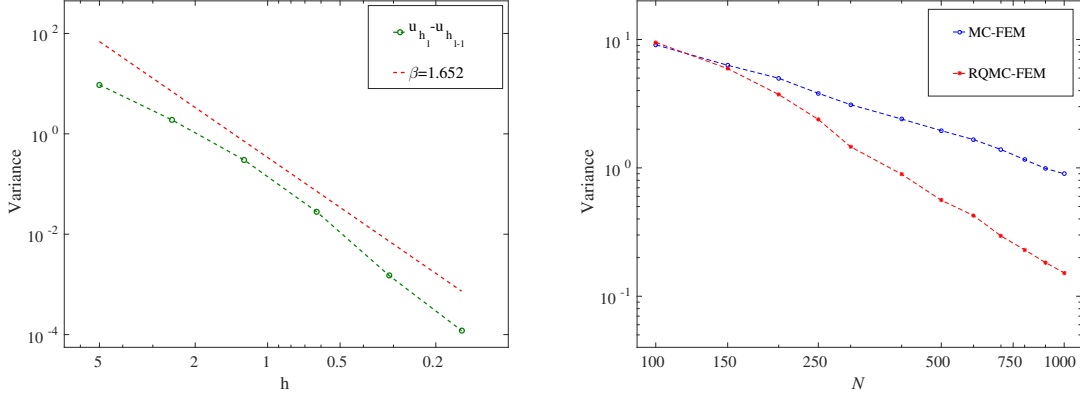
Figure 3.7.: The decay of variance of the solution as a function of different mesh sizes (left) and number of samples (right). The values $C_{00} = 9.45$, $C_0 = 0.338$, and $\delta = 0.12$ are found

.
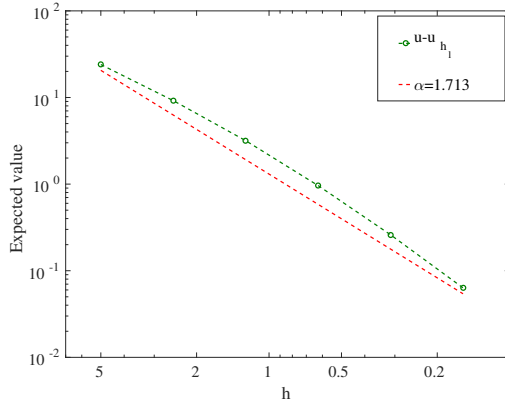
Figure 3.8.: The expected value of the solution as a function of different mesh sizes with $C_1 = 1.304$.

Coefficient	μ_1	γ_1	μ_2	γ_2	μ_3	γ_3	μ_4	γ_4
Value	0.51	3.07	0.63	3.06	0.38	2.98	0.34	2.93

Table 3.4.: The estimated coefficients and exponents in (3.68).

Since (3.69) is a continuous optimization problem, the solutions N_ℓ are generally no integers. We therefore round the values N_ℓ up to the next integer. Regarding the number of shift realizations, the value $M_\ell = 10$ is used in all the QMC estimators. Summaries of the parameter values (h, N), (h_0, r, N_ℓ) and (h_0, r, M_ℓ) for the QMC-FE and MLRQMC-FE and MLMC-FE methods are given in Table 3.5, Table 3.6 and Table 3.7 respectively.

We compare a previously developed optimal MLMC-FE method [36] with the optimal MLRQMC-FE method developed in Section 3.9. Figure 3.9 shows the computational work

ε	0.100	0.050	0.030	0.020	0.010	0.005	0.003	0.001
h	0.427	0.208	0.154	0.122	0.081	0.054	0.071	0.041
N	65	135	231	356	744	1 554	3 844	9 913

Table 3.5.: Optimal mesh size h and number N of samples for the QMC-FE method for different prescribed total errors ε.

ε	h_0	r	N_0	N_1	N_2	N_3	N_4	N_5	N_6
0.100	2.192	2.270	209	31	6	2	–	–	–
0.050	2.651	2.144	633	124	24	5	2	–	–
0.030	2.174	2.204	1 015	154	28	5	2	–	–
0.020	2.943	2.094	2 363	534	106	21	5	2	–
0.010	2.213	2.149	4 395	697	131	25	5	2	–
0.005	2.899	2.115	14 139	3 107	603	117	25	5	2

Table 3.6.: Optimal hierarchies in the MLRQMC-FE method for different prescribed total errors ε.

ε	h_0	r	M_0	M_1	M_2	M_3	M_4	M_5	M_6
0.100	1.303	2.151	3 920	363	33	3	–	–	–
0.050	1.370	2.020	18 046	1 993	215	24	3	–	–
0.030	1.430	1.908	56 136	7 344	957	125	17	3	–
0.020	1.390	1.987	126 266	14 749	1 688	193	23	3	–
0.010	1.459	1.9829	545 840	73 009	9 182	1 154	165	21	3

Table 3.7.: Optimal hierarchies in the MLMC-FE method for different prescribed total errors ε.
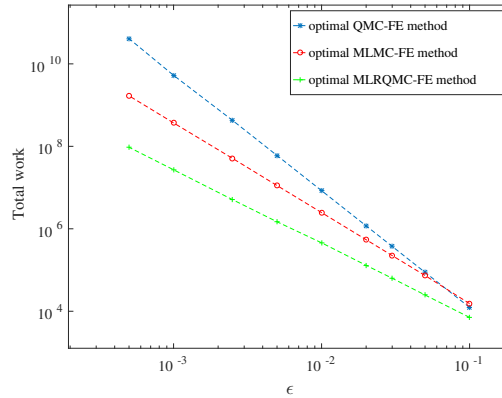

Figure 3.9.: Comparison of the total computational work required for the optimal ML-RQMC and MLMC methods. For smaller total errors, the effectiveness of the randomized method is more pronounced.
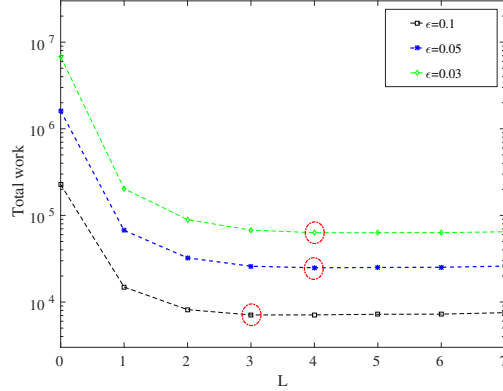
Figure 3.10.: The comparison of the total work of MLRQMC for different levels (between $L = 0$ and $L = 7$) for three different total errors $\varepsilon = 0.1$, $\varepsilon = 0.05$, and $\varepsilon = 0.03$. For each prescribed total error, the optimal number of levels is indicated by a red circle.

for the optimal quasi-Monte-Carlo method and the multilevel methods. It shows that $\mathcal{O}(\varepsilon^{-2.75})$ is roughly constant for the standard QMC method. In the MLMC-FE method, the assumptions of the standard complexity theorem [30] are satisfied, i.e., $\alpha \geq \frac{1}{2}\min(\beta, \gamma)$, so that the computational cost is $\mathcal{O}(\varepsilon^{-2.2})$. The faster convergence rate of the RQMC points results in less computational work for a given total error. In the MLRQMC-FE method, the RQMC aspect yields a computational complexity of $\mathcal{O}(\varepsilon^{-1.82})$, which results in additional savings of a factor between 2 and 17 (relative to MLMC) and 2 and 500 (relative to QMC). Therefore, the efficiency increase of the multilevel RQMC method is more pronounced for smaller prescribed total errors.

Additionally, choosing the optimal number L of levels is another important consideration. Figure 3.10 depicts the optimal number of levels for three different prescribed total errors. Using only one level ($L := 0$) results in the standard Monte-Carlo method. Distributing the samples among several levels $\ell \in \{0, \dots, L\}$ results in significant savings in computational cost. For smaller error bounds, a larger number of levels is necessary to obtain the minimum of computational cost.

3.10.2. The Effect of Random Dopants

The main source of randomness inside the nanoscale devices such as nanowire field-effect sensors and transistors is the random motion of dopant atoms through the semiconductor during the fabrication steps of implantation and annealing resulting in their random locations. To define the microscopic doping profile of individual randomly distributed dopants, a point doping model such as

$$C(x, \omega) = \sum_{j=1} C_j \delta(x - x_j(\omega)) \tag{3.71}$$

can be used, where x_j and C_j are the position and the charge of the j-th dopant and $\delta(x - x_j(\omega))$ is the Dirac delta distribution at point $x_j(\omega)$. The Gaussian model [86, 101]

$$C(x,\omega) := \sum_j \frac{C_j}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{(x - x_j(\omega))^2}{2\sigma^2}\right) \tag{3.72}$$

is a smoothed version of the point doping model (3.71), where σ is the so called influence parameter. Here we study the effect of randomness in the position of the dopants, whose number is constant. The fluctuation of the current in the discrete model for $V_g = 0.1\,\text{V}$ and $V_g = 0.2\,\text{V}$ is shown in Figure 3.11.
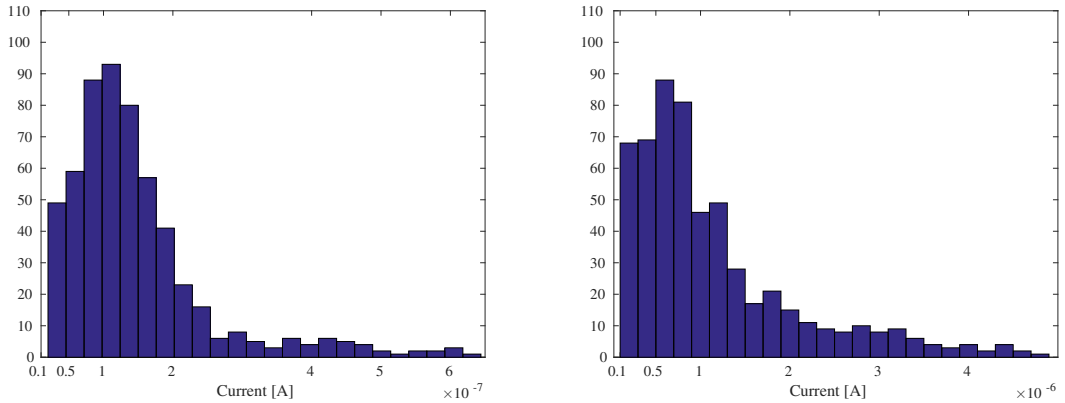

Figure 3.11.: Histogram of the current for $V_g = 0.1\,\text{V}$ (left) and $V_g = 0.2\,\text{V}$ (right), $N_{\text{dop}} = 25$ and 563 simulations. Here $\mathbb{E}(I) = 1.49 \cdot 10^{-7}\,\text{A}$ for the lower gate voltage and $\mathbb{E}(I) = 1.14 \cdot 10^{-6}\,\text{A}$ for the higher gate voltage.

Next, we compare the expected value of the current for different numbers N_{dop} of dopants with the continuum model. Figure 3.12 shows the expected value of the current for different numbers of dopants, varying from 5 to 50, for various gate voltages. The total charge of the dopants is kept constant to allow the comparison. According to the figure, the presence of more than 10 atoms in the regions results in a higher current compared to the deterministic model at the same gate voltage. In this figure, it is observed that the variation in the number of dopants ($\Delta N_{\text{dop}} \neq 0$) gives rise to a noticeable current fluctuation. In other words, the variations decrease gradually when there are more dopants in the region, which is consistent with convergence to the continuum model as the number of dopants tends to infinity. Here, a comparison between two numbers of dopants ($N_{\text{dop}} = 5$ and $N_{\text{dop}} = 50$) is made in Figure 3.13, where the histograms show that more dopants lead to less fluctuations. Finally, an interesting result of the simulations is that considering the discrete nature of the dopants in the devices results in a decrease of the threshold voltage.

3.11. Summary

We have used the stochastic drift-diffusion-Poisson system to model and simulate charge transport in random environments. We have developed an optimal multilevel randomized
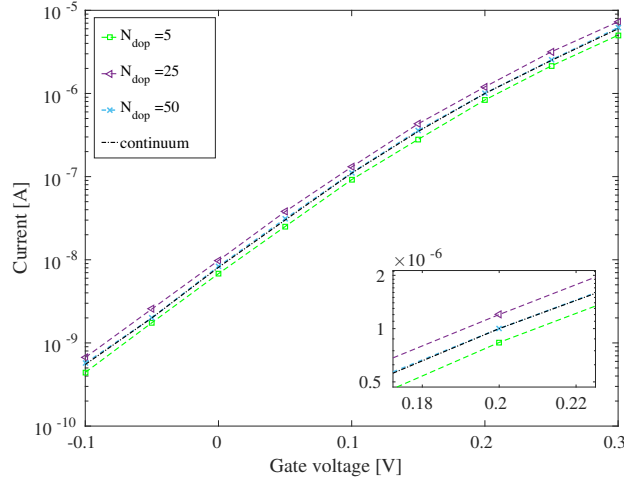
Figure 3.12.: The *I-V* characteristics for different numbers of dopants. The results for the continuum model are shown as well.
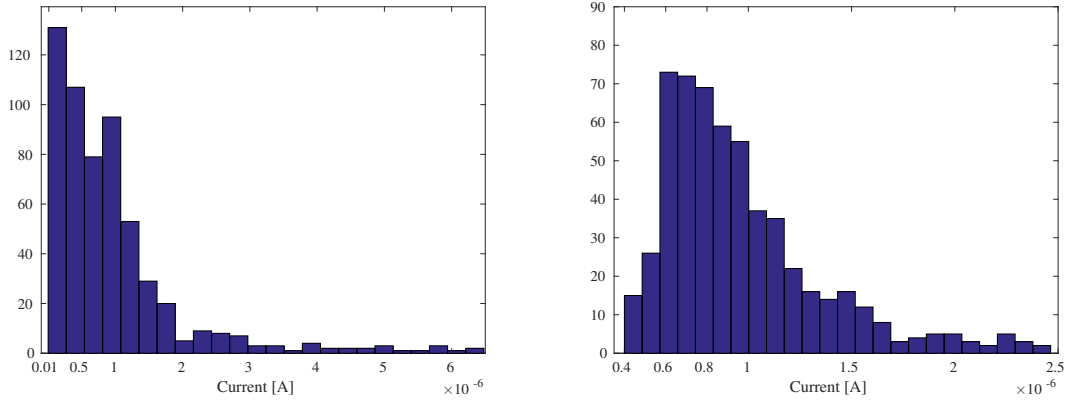

Figure 3.13.: Histogram of the current for $V_g = 0.2\,\mathrm{V}$ and for 563 simulations. Left: $N_{\mathrm{dop}} = 5$ resulting in $\mathbb{E}(I) = 9.22 \cdot 10^{-7}\,\mathrm{A}$. Right: $N_{\mathrm{dop}} = 50$, resulting in $\mathbb{E}(I) = 9.66 \cdot 10^{-7}\,\mathrm{A}$. The current obtained by the continuum model is $I = 9.66 \cdot 10^{-7}\,\mathrm{A}$.

quasi-Monte-Carlo method to calculate the expected value of the solution. We have compared the new method with the optimal multilevel Monte-Carlo method, where a reduction in the computational cost of the new method by more than one order of magnitude is found. In order to obtain the parameters of the numerical method and to solve the resulting optimization problem, we have used an SQP method as a generalization of Newton's method and approximated the nonlinear objective function by its local quadratic approximation. In summary, a computational complexity of $\mathcal{O}(\varepsilon^{-1.82})$ is achieved.

Here the effects of random dopants in nanoscale devices and variations due to the location and the number of dopants have been considered and compared to the continuum model.

As the number of dopants in the discrete model goes to infinity, the continuum model is obtained as the limit as expected. The variations are significant for a realistically small number of dopants.

Part II.

Inverse Modeling and Bayesian Inversion

4. Bayesian Inversion for Electrical-Impedance Tomography

Tomography is one of the most important imaging techniques in vast areas such as industrial, geophysical and medical application. Many problems in these applications could be easily diagnosed by information about the distribution of electrical properties inside the object of study. In this chapter, we focus on electrical-impedance tomography (EIT) as a soft-field tomography technique for imaging in medical applications. The goal is to extract internal properties of the body using surface measurements on the electrodes. We present and solve an elliptic partial differential equation, namely a nonlinear Poisson-Boltzmann equation as the EIT model to find the electrical currents flowing in and out of the electrodes in a bioimpedance tomography device.

We solve the EIT forward problem by means of the first order Galerkin finite-element (GFE) approximation and a mesh generated by the GMSH package [54], and the corresponding inverse problem is solved using the delayed-rejection adaptive-Metropolis (DRAM) algorithm in the context of Beyasian inversion techniques, which we have implemented in Julia [55]. We also discuss Bayesian inversion for the presented nonlinear elliptic PDE model for EIT in the measure-theoretic framework and prove that the posterior measure is Lipschitz continuous in the data to conclude well-definedness and well-posedness of the resulting posterior measures obtained by the Bayesian technique.

The chapter is organized as follows: In Section 4.1, the geometry of the physical domain is defined, and the related mathematical forward model, namely the standard linear and the nonlinear model are introduced. Section 4.2 is devoted to the EIT inverse problem. In this section, first we give a brief review of Bayesian estimation as a strong uncertainty quantification (UQ) tool for solving the EIT inverse problem and then formulate it in the measure theoretic framework. The Markov-chain Monte-Carlo method in the context of Bayesian inversion and the adaptive Metropolis-Hastings algorithm, namely the DRAM algorithm, are presented in this section as well. Furthermore, we prove well-definedness and well-posedness of Bayesian inversion for the EIT PDE model. In Section 4.3, numerical experiments including their results are illustrated. These results show the capability of the Bayesian approach to identify the quantities of interest in the EIT inverse problem, which is defined for the presented nonlinear forward model and support our theoretical findings. Finally, summary of the chapter is presented in Section 4.4.

This chapter is based on the author's work [102].

4.1. The EIT Forward Problem

The problem formulation for the EIT is based on [44]. First, we briefly review derivation of the standard linear EIT model, and then extend it to the nonlinear one. We give a weak

formulation of the new model together with the required assumptions for the existence and uniqueness results.

4.1.1. The linear Model

First, we describe the geometry of the physical domain. We assume a square domain as a cross-section of the main object under consideration containing an inclusion consisting of a different material than the background medium. This EIT device has eight electrodes equidistantly attached to the surface of the main body. In Figure 4.1, a schematic diagram of the device with one of the measurement patterns is shown. In this pattern, a potential is applied between two electrodes, and the resulting electrical current is measured at the rest of the electrodes. The forward problem is to find the electrostatic potential in the physical
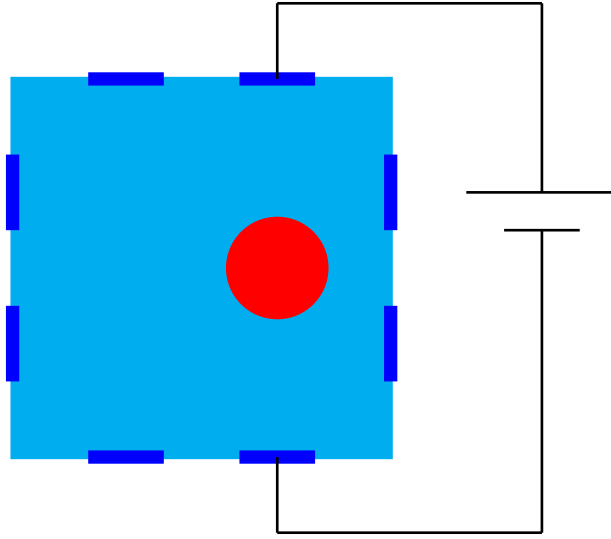
Figure 4.1.: Schematic diagram of a cross-section of an EIT device with eight electrodes attached to the boundary of the main object including one inclusion.

domain and then to calculate the electrical current flowing through the electrodes. Assume that $D \subset \mathbb{R}^n, n \in \{2,3\}$, is a closed and bounded domain with a smooth boundary ∂D. Using the quasi-static approximation, the electrical field E can be represented in terms of a scalar potential u by

$$E(x) = -\nabla u(x), \tag{4.1}$$

where $x \in D$. For simplicity we assume direct current or sufficiently low-frequency current such that the magnetic field can be neglected.

In the case of direct currents, in which the applied voltage is independent of time, the derivation is simple. The electric potential u results in the current density J, which satisfies the continuum Ohm's law $J = -\gamma \nabla u$, where γ is the admittivity [103]. In the absence of current sources in the interior of the body, the continuum version of Kirchoff's law $\nabla \cdot J = 0$ results in the elliptic PDE

$$\nabla \cdot (\gamma \nabla u) = 0.$$

In the case of alternating currents, we consider the time-harmonic Maxwell's equations at a fixed angular frequency ω, and it is assumed that the transient components of all fields are negligible [103]. Ignoring magnetic effects, we substitute equation (4.1) into the time-harmonic Maxwell's equations, while assuming no internal current source in EIT problems, and take the divergence on the both sides to obtain the (complex-valued) linear model

$$\nabla \cdot (\gamma \nabla u) = 0, \tag{4.2}$$

where

$$\gamma(x, \omega) := \sigma(x, \omega) + i\omega\epsilon(x, \omega) \tag{4.3}$$

is the admittivity, and σ and ϵ are the electric conductivity and permittivity, respectively. Also, ω is the frequency of the electrical current. Since we restrict the present discussion to static fields, i.e., $\omega \to 0+$, the admittivity is real and coincides with the static conductivity. Moreover, the impedivity $\rho = 1/\gamma$ is just the resistivity of the body [45]. The linear model (4.2) is widely used for modeling EIT; the leading EIT software EIDORS (http://eidors.org) is based on this model equation.

To find the boundary conditions, we assume that there are L contact electrodes e_ℓ, which are attached to the surface of the body, i.e.,

$$e_\ell \subset \partial D, \quad 1 \leq \ell \leq L, \tag{4.4}$$

such that $\overline{e}_\ell \cap \overline{e}_k = \emptyset$ for $\ell \neq k$.

Definition 5 (Voltage pattern). *Let U_ℓ denote the voltage applied to the ℓ-th electrode, the ground voltage being chosen such that $\sum_{\ell=1}^{L} U_\ell = 0$. Then the vector $U := [U_1, \ldots, U_L]^T \in \mathbb{R}^L$ is called a voltage pattern.*

We assume that the electrodes conduct perfectly, and thus the tangential electrical field vanishes along the electrodes. Then possible boundary conditions on the electrodes are the Dirichlet boundary conditions

$$u(x) = U_\ell, \quad x \in e_\ell, \quad 1 \leq \ell \leq L. \tag{4.5}$$

We also assume that no current flows in and out of the body between the electrodes, which leads to zero Neumann boundary condition

$$\frac{\partial u(x)}{\partial \mathbf{n}} = 0, \quad x \in \partial D \backslash \bigcup_{\ell=1}^{L} e_\ell. \tag{4.6}$$

4.1.2. The nonlinear Model

To extend the model, we consider free charges in the equation. To this end, we assume that ions in the background medium are pointlike charges. In this case, the free charges contribute to the electrical potential u in the body. Assume that bulk concentration of positive and negative free charges are denoted by η^+ and η^-, respectively. Furthermore assume that charges of single positive and negative charge carriers are denoted by q^+ and q^-. Assuming that the energies of all ions in the electric field are distributed according

to a Boltzmann distribution, then the charge density of free charges in the system can be defined by

$$f_{\text{free}} := \eta^+ \exp(-\frac{q^+ u(x)}{k_B T}) - \eta^- \exp(-\frac{q^- u(x)}{k_B T}).$$

Assuming $\eta^+ = \eta^- =: \eta$ and $q^+ = q^- =: q$, the above equation can be simplified as

$$f_{\text{free}} := \eta(\exp(-\beta u) - \exp(\beta u)) = -2\eta \sinh(\beta u).$$

Adding charge density of free charges f_{free} to the fixed charges f_{fixed} (for simplicity we denote it by f in the equation), we arrive at the nonlinear Poisson-Boltzmann equation

$$-\nabla \cdot (A(x)\nabla u(x)) = f(x) - 2\eta(x)\sinh(\beta u(x))$$

as the extended model for the EIT problem. Therefore, the forward problem describing EIT is to find the potential u in the main object D, given the permittivity A, the voltage pattern $U = [U_1, \ldots, U_L]^T$, the bulk ionic concentration η, and the concentration f of fixed charges, that solves the (real-valued) nonlinear elliptic PDE

$$-\nabla \cdot (A(x)\nabla u(x)) + 2\eta(x)\sinh(\beta u(x)) = f(x) \qquad \forall x \in D, \tag{4.7a}$$

$$u(x) = U_\ell \qquad \forall x \in e_\ell, \tag{4.7b}$$

$$\frac{\partial u(x)}{\partial \mathbf{n}} = 0 \qquad \forall x \in \partial D \backslash \bigcup_{\ell=1}^{L} e_\ell, \tag{4.7c}$$

where $\beta := q/k_B T$. Furthermore, the sinh term stems from the Boltzmann distributions for two species of ions.

As mentioned before, in every measurement pattern a potential is applied to the electrodes and the resulting electrical current on the rest of the electrodes is measured. The electrical current flowing through the electrodes in the EIT problem [44] is calculated by

$$I_\ell = \int_{e_\ell} A \frac{\partial u(x)}{\partial \mathbf{n}} ds, \qquad \ell = 1, 2, \ldots, L. \tag{4.8}$$

Weak Formulation and Computational Method

Here, we first present the required assumptions for existence of a unique solution to the model (4.7), which is postponed to Section 4.2. Then the weak formulation of the problem is stated, which are then used to prove the main theoretical results in Section 4.2.

Assumptions 7. *The permittivity* $A \colon D \to \mathbb{R}^{2 \times 2}$ *and the voltage simulation pattern* $\{U_\ell\}_{\ell=1}^{L} \in \mathbb{R}^L$ *satisfy the following assumptions:*

1. *The coefficient* $A \colon D \to \mathbb{R}^{2 \times 2}$ *is a piecewise constant-valued matrix satisfying*

$$A \in L^\infty(D; \mathbb{R}^{2 \times 2}), \quad \operatorname{ess\,inf}_{x \in D} A(x) = A^- > 0 \tag{4.9}$$

and containing the permittivity of the inclusion or background medium, as the two materials are different in their physical properties.

2. *The voltage applied to the ℓ-th electrode is chosen such that $\sum_{\ell=1}^{L} U_\ell = 0$.*

Definition 6. *Under Assumption 7, for every $f(x) \in L^2(D)$ and $U_\ell \in H^{1/2}(\partial D), x \in e_\ell,$ the function $u \in H^1(D)$ is a weak solution of the boundary-value problem (4.7), if it satisfies*

$$a(u, \phi) = l(\phi) \qquad \forall \phi \in H_0^1(D), \tag{4.10}$$

where $a \colon H^1(D) \times H_0^1(D) \to \mathbb{R}$ and $l \colon H_0^1(D) \to \mathbb{R}$ are defined by

$$a(u, \phi) := \int_D A\nabla u \cdot \nabla \phi \mathrm{d}x + \int_D (-2\eta)\sinh(\beta u)\phi \mathrm{d}x$$

and

$$l(\phi) := \int_D f\phi \mathrm{d}x.$$

The PDE model (4.7) defined on the physical domain is solved using the first order Galerkin finite-element method. The mesh is generated using the GMSH package, and it is aligned with the inclusion such that each element has a constant value for the coefficient A. Figure 4.2 displays a simulated computational domain of the EIT with FEM mesh with 1335 triangular elements and 657 nodes, and mesh size 0.1. The results of the FEM solver for the solution of the nonlinear forward model are presented in Section 4.3.
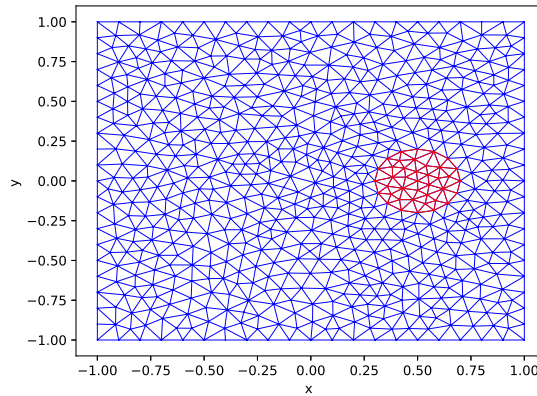
Figure 4.2.: Simulated EIT domain containing a circle-shaped inclusion centred at $(0.5, 0)$ with radius of 0.2.

4.2. The EIT Inverse Problem

The EIT inverse problem is to reconstruct the electrical and physical properties of the body interior, given the electrical current measurements on its surface. Bayesian inferences [104–107] applied to the EIT inverse problem corresponding to the standard linear forward model (4.2) has previously been studied. This inversion method for nonlinear inverse problems in infinite dimensions is a new approach. In this work, we apply Bayesian

estimation to the EIT inverse problem corresponding to the presented nonlinear forward model (4.7), as this inversion technique is capable of dealing with the ill-posedness and nonlinearity of the problem at hand successfully. We first briefly review Bayes' Theorem as a connection between inverse problems and probability theory. One of the goals in this section is to show well-definedness of the posterior measure obtained by the Bayesian techniques fo the EIT inverse model which arises from the nonlinear forward model that we presented in the previous section. Well-posedness of the Bayesian inversion is also proved. To this end, we first present the mathematical formulation of the Bayesian analysis in a measure-theoretic framework and in the infinite-dimensional setting. Additionally we collect required assumptions as well as Proposition 8 that finds a pointwise estimate for solution of the nonlinear forward model. Finally, our main theoretical results are stated and proved in Proposition 9. These results include boundedness and Lipschitz continuity of the solution of the physical model by functions of the parameters, which lead to well-definedness and well-posedness of the applied Bayesian estimation method for the prescribed model in the EIT technology. At the rest of the section, we review the Markov-chain Monte-Carlo (MCMC) methods and the delayed-rejection adaptive-Metropolis (DRAM) algorithm in the context of Bayesian inversion. MCMC methods are sampling methods based on non-i.i.d. and correlated samples, and the DRAM algorithm is of interest as an adaptive MCMC method since it chooses the optimal proposal scale for defining the Markov-chain states automatically. We will discuss it in Subsection 4.2.5 in detail.

4.2.1. Bayesian Analysis

In fact the solution of the inverse problem is the posterior density that best reflects the distribution of the parameter based on the observations. As the observations or measurements are subject to noise and the observational noise, i.e., the error e due to modeling and measurement, is unbiased and iid, it can be represented by random variables as

$$M = G(Q) + e, \tag{4.11}$$

where e is a mean-zero random variable and M is a given random variable representing observed data or measurements, for which we have a model $G(Q)$ (observation operator) dependent on a random variable Q with realizations $q = Q(\omega)$ representing parameters to be estimated [104].

Assume a given probability space (Ω, F, P), where Ω is the set of elementary events (sample space), F a σ-algebra of events, and P a probability measure. Furthermore assume that all the random variables are absolutely continuous.

Bayes' Theorem in terms of probability densities can be written as

$$\pi(q|y) = \frac{\pi_0(q)\pi(y|q)}{\pi(y)}, \tag{4.12}$$

with

$$\pi(y) := \int_{\mathbb{R}^p} \pi_0(q)\pi(y|q)\mathrm{d}q \neq 0, \tag{4.13}$$

where the unknown parameters $q = (q_1, \ldots, q_p) \in \mathbb{R}^p$ and the observed data y are realizations of the random variables Q and M, respectively. Furthermore, $\pi_0(q)$, $\pi(q|y)$, and $\pi(y|q)$ are

the probability density functions of the prior, posterior, and (data) sampling distributions, respectively. The density $\pi(y|q)$ of the data provides information from the measurement data to update the prior knowledge, and it is well-known as the likelihood density function. The goal of Bayesian inversion is to estimate the posterior probability density function $\pi(q|y)$, which reflects the uncertainty about the quantity of interest q using measurement data y.

Equation (4.12) gives the posterior density and summarizes our beliefs about q after we have observed y. Therefore, Bayes' Theorem for inverse problems can be stated as follows.

Theorem 6 (Bayes' Theorem for inverse problems [45, 104])**.** *Let $\pi_0(q)$ be the prior probability density function of the realizations q of the random parameter Q. Let y be a realization or measurement of the random observation variable M. Then the posterior density of Q given the measurements y is*

$$\pi(q|y) = \frac{\pi_0(q)\pi(y|q)}{\pi(y)} = \frac{\pi_0(q)\pi(y|q)}{\int_{\mathbb{R}^p} \pi_0(q)\pi(y|q)\mathrm{d}q}. \tag{4.14}$$

Computing the integral appearing in Bayes' Theorem 6 is costly especially if the parameter space $\mathbb{R}^p$ is high-dimensional. Another problem with quadrature rules is that they require a relatively good knowledge of the support of the probability distribution, which is usually part of the information that we seek [45, 104]. In Section 4.2.4 we shortly discuss the algorithms for Bayesian estimation, which do not require evaluations of the integral and that are used to achieve the numerical results for the nonlinear model equation.

4.2.2. Bayesian Approach in Measure-Theoretic Framework

To describe the Bayesian approach on function spaces, we formulate Bayes' Theorem in a measure-theoretic framework, which is suitable for problems on infinite-dimensional spaces. To this end, assume that $(X, \|\cdot\|_X)$ (infinite-dimensional) and $(Y, \|\cdot\|_Y)$ (possibly infinite-dimensional) are separable Banach spaces and that $G\colon X \to Y$ is the observation operator. Therefore, let $q \in X$ be a random variable distributed according to measure μ_0 on X, in which our prior beliefs about the unknown parameter q are described. We assume the distribution of the measurement error e (data likelihood) is defined by

$$\pi(y|q) := \pi(y - G(q)) \tag{4.15}$$

to calculate the posterior probability measure μ^y for $q \in X$ given $y \in Y$, which leads to

$$\pi(q|y) = \frac{\pi_0(q)\pi(y - G(q))}{\int_{\mathbb{R}^p} \pi_0(q)\pi(y - G(q))\mathrm{d}q} \tag{4.16}$$

using Bayes' formula (4.14), where π_0 and π are the prior and posterior density functions and correspond to the probability measures μ_0 and μ^y, respectively. Thus we have

$$\pi(q|y) \propto \pi_0(q)\pi(y - G(q)) \tag{4.17}$$

with a constant of proportionality depending only on y.

First we define absolutely continuous measures and state the Radon-Nikodym Theorem, which are needed in the rest of the paper.

Definition 7 (Absolutely continuous measures)**.** *If μ and ν are two measures on the same measure space (Ω, F), then μ is absolutely continuous with respect to ν if $\nu(A) = 0$ implies $\mu(A) = 0$ for every $A \in F$. This concept is denoted by $\mu \ll \nu$.*

Theorem 7 (Radon-Nikodym Theorem)**.** *Assume that μ and ν are two measures on the same measure space (Ω, F). If $\mu \ll \nu$ and ν is σ-finite, then there exists a ν-measurable function $f \colon \Omega \to [0, \infty]$ such that for all ν-measurable sets $A \in F$*

$$\mu(A) = \int_A f(x) \mathrm{d}\nu(x) \tag{4.18}$$

holds.

Remark 1. *The function f is known as the Radon-Nikodym derivative of μ with respect to ν and it can be written as*

$$\frac{\mathrm{d}\mu}{\mathrm{d}\nu} = f(x). \tag{4.19}$$

As in infinite-dimensional spaces there is no density with respect to the Lebesgue measure, Bayes' rule should be interpreted as providing the Radon-Nikodym derivative between the posterior measure $\mu^y(\mathrm{d}q) = P(\mathrm{d}q|y)$ (with density $\pi(q|y)$) and the prior measure $\mu_0(\mathrm{d}q) = P(\mathrm{d}q)$ (with density π_0), yielding

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(q) \propto \pi(y - G(q)). \tag{4.20}$$

Without loss of generality, we can view the right-hand side as the exponential of the negative of $\Phi(q, y)$, where $\Phi \colon X \times Y \to \mathbb{R}$ is called a potential. Hence equation (4.20) can be rewritten as

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(q) \propto \exp\left(-\Phi(q, y)\right), \tag{4.21}$$

since the density π is nonnegative [52]. The following theorem justifies the formula rigorously.

Theorem 8 ([52], Theorem 6.31)**.** *Suppose that $G \colon X \to \mathbb{R}^m$ is continuous, that π has support equal to $\mathbb{R}^m$, and that $\mu_0(X) = 1$. Then $q|y$ is distributed according to the measure $\mu^y(\mathrm{d}q)$, which is absolutely continuous with respect to $\mu_0(\mathrm{d}q)$ and has the Radon-Nikodym derivative*

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(q) \propto \exp\left(-\Phi(q, y)\right). \tag{4.22}$$

Furthermore, the posterior measure μ^y in some PDE inverse problems can be formulated as

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(q) = \frac{1}{C(y)} \exp\left(-\Phi(q, y)\right), \tag{4.23}$$

where $C(y)$ is a normalization constant and chosen such that μ^y is a probability measure, i.e.,

$$C(y) := \int_X \exp\left(-\Phi(q, y)\right) \mathrm{d}\mu_0(q). \tag{4.24}$$

Furthermore, we assume that $\mu_0(X) = 1$ holds for the infinite-dimensional separable Banach space X.

The goal is to show that the posterior measure μ^y of the form (4.23) is well-defined and that the problem is well-posed with respect to its dependence on the data. To this end, the function $\Phi\colon X \times Y \to \mathbb{R}$ should have essential properties, namely lower and upper bounds and the Lipschitz property in q and y. As this function is defined in terms of the function $G\colon X \to \mathbb{R}^m$, it is sufficient to prove the following properties of the function G corresponding to the inverse problem of interest. This implies that $\Phi\colon X \times \mathbb{R}^m \to \mathbb{R}$ satisfies Assumption 2.6 in [52] with $(Y, \|\cdot\|_Y) = (\mathbb{R}^m, |\cdot|)$.

Assumptions 8. *The function $G\colon X \to \mathbb{R}^m$ has the following properties.*

1. *For every $\varepsilon > 0$, there exists an $M(\varepsilon) \in \mathbb{R}$ such that the inequality*

$$|G(q)| \leq \exp(\varepsilon\|q\|_X^2 + M(\varepsilon)) \tag{4.25}$$

   holds for all $q \in X$.

2. *For every $r > 0$, there exists a $K(r) > 0$ such that the inequality*

$$|G(q_1) - G(q_2)| \leq K(r)\|q_1 - q_2\|_X \tag{4.26}$$

   holds for all $q_1, q_2 \in X$ with $\max(\|q_1\|_X, \|q_2\|_X) < r$.

First, we check if the bounds and Lipschitz properties in Assumptions 8 can be shown when G is given by the solution of the (real-valued) linear elliptic PDE

$$\begin{align} -\nabla \cdot (A(x)\nabla u(x)) &= f & \forall x \in D, \tag{4.27a}\\ u(x) &= g & \forall x \in \partial D_D, \tag{4.27b}\\ \frac{\partial u(x)}{\partial \mathbf{n}} &= 0 & \forall x \in \partial D_N \tag{4.27c} \end{align}$$

and by the solution of the (real-valued) nonlinear elliptic PDE

$$\begin{align} -\nabla \cdot (A(x)\nabla u(x)) + 2\eta(x)\sinh(\beta u(x)) &= 0 & \forall x \in D, \tag{4.28a}\\ u(x) &= g & \forall x \in \partial D_D, \tag{4.28b}\\ \frac{\partial u(x)}{\partial \mathbf{n}} &= 0 & \forall x \in \partial D_N, \tag{4.28c} \end{align}$$

where ∂D_D and ∂D_N denote the Dirichlet and Neumann boundaries. In (4.27), the solution u can be the real or the imaginary part of the solution of the complex-valued model (4.2) − (4.3), i.e., either $\Re(u)$ or $\Im(u)$. Here, for the sake of simplicity we denote the solution of the real-valued model (4.27) by u as well.

First we state a proposition [15], which ensures the existence and uniqueness of the solution of the nonlinear Poisson-Boltzmann equation (4.28) and gives a pointwise estimate for the solution to be used later.

Proposition 8 (Existence, uniqueness, and a pointwise estimate)**.** *Suppose that the domain $D \subset \mathbb{R}^n$, $n \in \{2,3\}$ is bounded, the matrix-valued function $A \in L^\infty(D; \mathbb{R}^{d\times d})$ is uniformly*

elliptic, $\eta \in L^\infty(D; \mathbb{R}_0^+)$, $\beta \in \mathbb{R}^+$, and $g \in H^{1/2}(\partial D)$. Then the semilinear boundary-value problem (4.28) has a unique solution $u \in H^1(D) \cap L^\infty(D)$. Furthermore, the estimate

$$\kappa \leq u(x) \leq \lambda$$

holds for almost all $x \in \overline{D}$, where

$$\kappa := \min\left(\inf_{\partial D} g, -\sup_D u_L\right),$$

$$\lambda := \max\left(\sup_{\partial D} g, -\inf_D u_L\right)$$

and u_L is the solution of the associated linear equation

$$\begin{aligned}
-\nabla \cdot (A(x)\nabla u_L(x)) &= 0 && \forall x \in D, \\
u_L(x) &= g && \forall x \in \partial D_D, \\
\frac{\partial u_L(x)}{\partial \mathbf{n}} &= 0 && \forall x \in \partial D_N.
\end{aligned}$$

4.2.3. Main Results

In the following proposition, the main results are a parameter dependent bound for the solution of the nonlinear model equation (4.28) and the Lipschitz property of the solution. Furthermore, a similar bound for and the Lipschitz property of the solution of the linear model equation (which can also be found in [108]) are proved in this proposition as well, as they are the foundation for the theory for the nonlinear equation.

Proposition 9 (Bounds for real-valued linear and nonlinear models). *Suppose the real-valued linear and nonlinear equations (4.27) and (4.28) hold in the bounded domain $D \subset \mathbb{R}^n$, $n \in \{2, 3\}$, with a smooth boundary ∂D and $A := e^q =: \eta$, where $q \in L^\infty(D)$. Then the following estimates hold.*

1. *The estimate*

$$\|u\|_{H^1(D)} \leq F e^{2\|q\|_{L^\infty(D)}} \tag{4.30}$$

   holds for all $q \in L^\infty(D)$, and the estimate

$$\|u_1 - u_2\|_{H^1(D)} \leq F e^{4\max(\|q_1\|_{L^\infty(D)}, \|q_2\|_{L^\infty(D)})} \|q_1 - q_2\|_{L^\infty(D)} \tag{4.31}$$

   holds for all q_1 and $q_2 \in L^\infty(D)$ for equation (4.27), where $F = F(\|\nabla \overline{g}\|_{L^2(D)}, \|f\|_{L^2(D)})$ is a function and $\overline{g} \in L^2(D)$ is the Dirichlet lift of g.

2. *The estimate*

$$\|u\|_{H^1} \leq H e^{2\|q\|_{L^\infty}} \tag{4.32}$$

   holds for all $q \in L^\infty(D)$, and the estimate

$$\|u_1 - u_2\|_{H^1} \leq S e^{4\max(\|q_1\|_{L^\infty(D)}, \|q_2\|_{L^\infty(D)})} \|q_1 - q_2\|_{L^\infty(D)} \tag{4.33}$$

   holds for all q_1 $q_2 \in L^\infty(D)$ hold for equation (4.28), where $H = H(\|\nabla \overline{g}\|_{L^2(D)})$ and $S = S(\|\nabla \overline{g}\|_{L^2(D)})$ are functions and $\overline{g} \in L^2(D)$ is the Dirichlet lift of g.

Proof. We substitute $v := u - \overline{g}$ in (4.27), where $\overline{g}$ is the Dirichlet lift of g. It is defined by

$$\overline{g} := \begin{cases} g & \text{on } \partial D, \\ \text{arbitrary} & \text{in } D \end{cases} \tag{4.34}$$

such that $\overline{g} \in L^2(D)$.

In the first step, we show the estimates for the linear equation. In order to find estimates (4.30) and (4.31) for the linear model (4.27), we take the inner product with any $v \in H_0^1(D)$ to obtain

$$(-\nabla \cdot (A\nabla v), v) = (\nabla \cdot (A\nabla \overline{g}), v) + (f, v) \qquad \forall v \in H_0^1(D),$$

which leads to

$$(A\nabla v, \nabla v) = (-A\nabla \overline{g}, \nabla v) + (f, v)$$

and

$$I := \left| \int A\nabla v \cdot \nabla v \right| = \left| -\int A\nabla \overline{g} \cdot \nabla v + \int fv \right|.$$

Hence using the Cauchy-Schwarz inequality we find

$$e^{-\|q\|_{L^\infty}} \|\nabla v\|_{L^2}^2 \leq I \leq e^{\|q\|_{L^\infty}} \|\nabla \overline{g}\|_{L^2} \|\nabla v\|_{L^2} + \|f\|_{L^2} \|v\|_{L^2},$$

where $A = e^q$. Using the Poincaré inequality, the inequality on the right-hand side can be written as

$$I \leq e^{\|q\|_{L^\infty}} \|\nabla \overline{g}\|_{L^2} \|\nabla v\|_{L^2} + C_p \|f\|_{L^2} \|\nabla v\|_{L^2},$$

where C_p is a Poincaré constant. This leads to

$$\|\nabla v\|_{L^2} \leq e^{2\|q\|_{L^\infty}} \|\nabla \overline{g}\|_{L^2} + C_p e^{\|q\|_{L^\infty}} \|f\|_{L^2}.$$

Furthermore, we calculate

$$\begin{aligned} \|\nabla u\|_{L^2} \leq \|\nabla v\|_{L^2} + \|\nabla \overline{g}\|_{L^2} &\leq e^{2\|q\|_{L^\infty}} \|\nabla \overline{g}\|_{L^2} + C_p e^{\|q\|_{L^\infty}} \|f\|_{L^2} + \|\nabla \overline{g}\|_{L^2} \quad (4.35) \\ &= (1 + e^{2\|q\|_{L^\infty}}) \|\nabla \overline{g}\|_{L^2} + C_p e^{\|q\|_{L^\infty}} \|f\|_{L^2} \\ &\leq 2e^{2\|q\|_{L^\infty}} \|\nabla \overline{g}\|_{L^2} + C_p e^{\|q\|_{L^\infty}} \|f\|_{L^2} \\ &\leq 2C_p e^{2\|q\|_{L^\infty}} (\|\nabla \overline{g}\|_{L^2} + \|f\|_{L^2}) \\ &= F_0 e^{2\|q\|_{L^\infty}}, \end{aligned}$$

where $F_0 := 2C_p(\|\nabla \overline{g}\|_{L^2} + \|f\|_{L^2})$. Using

$$\frac{1}{1 + C_p^2} \|u\|_{H^1}^2 \leq \|\nabla u\|_{L^2}^2, \tag{4.36}$$

which is obtained by applying the Poincaré inequality in the definition of the H^1-norm and using inequality (4.35), we can write

$$\frac{1}{(1+C_p^2)^{1/2}}\|u\|_{H^1} \leq \|\nabla u\|_{L^2} \leq F_0 e^{2\|q\|_{L^\infty}},$$

where C_p is a Poincaré constant. This leads to

$$\|u\|_{H^1} \leq F e^{2\|q\|_{L^\infty}},$$

where $F = F_0(1 + C_p^2)^{1/2}$, which proves inequality (4.30).

To prove inequality (4.31), we assume that u_1 and u_2 satisfy (4.27) with coefficients $A_1 = e^{q_1}$ and $A_2 = e^{q_2}$. Therefore we have

$$\nabla \cdot (e^{q_1} \nabla u_1) = \nabla \cdot (e^{q_2} \nabla u_2).$$

Subtracting the term $\nabla \cdot (e^{q_1} \nabla u_2)$, we see that the difference $u_1 - u_2$ satisfies

$$\nabla \cdot (e^{q_1} \nabla (u_1 - u_2)) = \nabla \cdot ((e^{q_2} - e^{q_1}) \nabla u_2).$$

Now taking the inner product with $u_1 - u_2$, we obtain

$$(\nabla \cdot (e^{q_1} \nabla (u_1 - u_2)), u_1 - u_2) = (\nabla \cdot ((e^{q_2} - e^{q_1}) \nabla u_2), u_1 - u_2)$$

and hence

$$(e^{q_1} \nabla (u_1 - u_2), \nabla (u_1 - u_2)) = ((e^{q_2} - e^{q_1}) \nabla u_2, \nabla (u_1 - u_2)).$$

This yields

$$I := \left| \int e^{q_1} \nabla (u_1 - u_2) \cdot \nabla (u_1 - u_2) \right| = \left| \int (e^{q_2} - e^{q_1}) \cdot \nabla u_2 \nabla (u_1 - u_2) \right|.$$

Using the Cauchy-Schwarz inequality, we find

$$e^{-\|q_1\|_{L^\infty}} \|\nabla (u_1 - u_2)\|_{L^2} \leq I \leq \|e^{q_2} - e^{q_1}\|_{L^\infty} \|\nabla u_2\|_{L^2}. \tag{4.37}$$

Since $e^{q(x)}$ $(x \in D)$ is continuously differentiable, it is Lipschitz continuous by the Weierstrass Theorem. Thus we have

$$\|e^{q_2} - e^{q_1}\|_{L^\infty} \leq \|q_1 - q_2\|_{L^\infty} e^{\max(\|q_1\|_{L^\infty}, \|q_2\|_{L^\infty})}. \tag{4.38}$$

Substituting (4.38) into (4.37) and using inequality (4.35), we have

$$\|\nabla (u_1 - u_2)\|_{L^2} \leq C_p (\|\nabla \bar{g}\|_{L^2} + \|f\|_{L^2}) e^{4\max(\|q_1\|_{L^\infty}, \|q_2\|_{L^\infty})} \|q_1 - q_2\|_{L^\infty}.$$

Therefore we can write

$$\begin{aligned}
\|u_1 - u_2\|_{H^1} &\leq (1 + C_p^2)^{1/2} \|\nabla (u_1 - u_2)\|_{L^2} \\
&\leq C_p (\|\nabla \bar{g}\|_{L^2} + \|f\|_{L^2})(1 + C_p^2)^{1/2} e^{4\max(\|q_1\|_{L^\infty}, \|q_2\|_{L^\infty})} \|q_1 - q_2\|_{L^\infty} \\
&= F e^{4\max(\|q_1\|_{L^\infty}, \|q_2\|_{L^\infty})} \|q_1 - q_2\|_{L^\infty}
\end{aligned}$$

using inequality (4.36), which completes the proof for the linear equation (4.27).

To obtain similar estimates for the nonlinear model (4.28), we generalize the ideas used for equation (4.27). To find the bound (4.32), we consider the inner product

$$(A\nabla v, \nabla v) = (-A\nabla \overline{g}, \nabla v) + (-2\eta \sinh(\beta(\overline{g}+v)), v) \quad \forall v \in H_0^1(D).$$

Therefore we can write

$$I := \left| \int A\nabla v \cdot \nabla v \right| = \left| -\int A\nabla \overline{g} \cdot \nabla v - \int 2\eta \sinh(\beta(\overline{g}+v))v \right|,$$

where $A = e^q$. Thus, we have

$$e^{-\|q\|_{L^\infty}} \|\nabla v\|_{L^2}^2 \leq \left| \int e^q \nabla v \cdot \nabla v \right| = I.$$

Furthermore, using $\eta = e^q$ and $\sinh(\beta(\overline{g}+v)) = (e^{\beta(\overline{g}+v)} - e^{-\beta(\overline{g}+v)})/2$ as well as the triangle inequality, we find

$$I = \left| -\int e^q \nabla \overline{g} \cdot \nabla v - \int e^q (e^{\beta(\overline{g}+v)} + e^{-\beta(\overline{g}+v)})v \right|$$
$$\leq \left| \int e^q \nabla \overline{g} \cdot \nabla v \right| + \left| \int e^q e^{\beta(\overline{g}+v)} v \right| + \left| \int e^q e^{-\beta(\overline{g}+v)} v \right|.$$

Using the Cauchy-Schwarz and Poincaré inequalities, we can write

$$I \leq e^{\|q\|_{L^\infty}} \|\nabla \overline{g}\|_{L^2} \|\nabla v\|_{L^2} + C_p e^{\|q\|_{L^\infty}} e^{|\beta\lambda|} \|\nabla v\|_{L^2} + C_p e^{\|q\|_{L^\infty}} e^{|\beta\kappa|} \|\nabla v\|_{L^2},$$

where C_p is a Poincaré constant and $\kappa \leq u = \overline{g} + v \leq \lambda$ is a pointwise estimate obtained from Proposition 8 for the solution of the Poisson-Boltzmann equation, where κ and λ are constants.

This yields

$$e^{-\|q\|_{L^\infty}} \|\nabla v\|_{L^2}^2 \leq I \leq e^{\|q\|_{L^\infty}} (\|\nabla \overline{g}\|_{L^2} + C_p e^{|\beta\lambda|} + C_p e^{|\beta\kappa|} \|\nabla v\|_{L^2})$$

and thus

$$\|\nabla v\|_{L^2} \leq e^{2\|q\|_{L^\infty}} (\|\nabla \overline{g}\|_{L^2} + C_p e^{|\beta\lambda|} + C_p e^{|\beta\kappa|}).$$

Furthermore, we calculate

$$\begin{aligned}
\|\nabla u\|_{L^2} &\leq \|\nabla v\|_{L^2} + \|\nabla \overline{g}\|_{L^2} \qquad\qquad\qquad\qquad\qquad\qquad (4.39)\\
&\leq e^{2\|q\|_{L^\infty}} (\|\nabla \overline{g}\|_{L^2} + C_p e^{|\beta\lambda|} + C_p e^{|\beta\kappa|}) + \|\nabla \overline{g}\|_{L^2}\\
&= (1 + e^{2\|q\|_{L^\infty}}) \|\nabla \overline{g}\|_{L^2} + C_p e^{2\|q\|_{L^\infty}} (e^{|\beta\lambda|} + e^{|\beta\kappa|})\\
&\leq 2e^{2\|q\|_{L^\infty}} \|\nabla \overline{g}\|_{L^2} + C_p e^{2\|q\|_{L^\infty}}) (e^{|\beta\lambda|} + e^{|\beta\kappa|})\\
&\leq 2C_p e^{2\|q\|_{L^\infty}} (\|\nabla \overline{g}\|_{L^2} + e^{|\beta\lambda|} + e^{|\beta\kappa|})\\
&= H_0 e^{2\|q\|_{L^\infty}},
\end{aligned}$$

where $H_0 := 2C_p(\|\nabla \overline{g}\|_{L^2} + e^{|\beta\lambda|} + e^{|\beta\kappa|})$. Now using inequality (4.36), the estimate (4.32)

$$\|u\|_{H^1} \le He^{2\|q\|_{L^\infty}} \tag{4.40}$$

can be reached, where $H = H_0(1 + C_p^2)^{1/2}$.

To prove inequality (4.33), we assume that u_1 and u_2 satisfy (4.28) with coefficients $A_1 = e^{q_1} = \eta_1$ and $A_2 = e^{q_2} = \eta_2$. Hence we have

$$-\nabla \cdot (e^{q_1}\nabla u_1) + e^{q_1}(e^{\beta u_1} - e^{-\beta u_1}) = -\nabla \cdot (e^{q_2}\nabla u_2) + e^{q_2}(e^{\beta u_2} - e^{-\beta u_2}).$$

Subtracting the term $\nabla \cdot (e^{q_1}\nabla u_2)$, we observe that the difference $u_2 - u_1$ satisfies the equation

$$\nabla \cdot (e^{q_1}\nabla(u_2 - u_1)) = \nabla \cdot ((e^{q_1} - e^{q_2})\nabla u_2) + e^{q_2}(e^{\beta u_2} - e^{-\beta u_2}) - e^{q_1}(e^{\beta u_1} - e^{-\beta u_1}).$$

Taking the inner product of this equation with $u_2 - u_1$ yields

$$(e^{q_1}\nabla(u_2 - u_1), \nabla(u_2 - u_1)) = ((e^{q_1} - e^{q_2})\nabla u_2, \nabla(u_2 - u_1))$$
$$+ (e^{q_2}(e^{\beta u_2} - e^{-\beta u_2}) - e^{q_1}(e^{\beta u_1} - e^{-\beta u_1}), u_2 - u_1).$$

This leads to

$$I := \left| \int e^{q_1}\nabla(u_2 - u_1) \cdot \nabla(u_2 - u_1) \right|$$
$$= \left| \int (e^{q_1} - e^{q_2})\nabla u_2 \cdot \nabla(u_2 - u_1) + (e^{q_2}(e^{\beta u_2} - e^{-\beta u_2}) - e^{q_1}(e^{\beta u_1} - e^{-\beta u_1}))(u_2 - u_1) \right|,$$

where the inequalities

$$e^{-\|q_1\|_{L^\infty}}\|\nabla(u_2 - u_1)\|_{L^2}^2 \le \left| \int e^{q_1}\nabla(u_2 - u_1) \cdot \nabla(u_2 - u_1) \right| = I$$

and

$$I = \left| \int (e^{q_1} - e^{q_2})\nabla u_2 \cdot \nabla(u_2 - u_1) + (e^{q_2}(e^{\beta u_2} - e^{-\beta u_2}) - e^{q_1}(e^{\beta u_1} - e^{-\beta u_1}))(u_2 - u_1) \right|$$
$$\le \left| \int (e^{q_1} - e^{q_2})\nabla u_2 \cdot \nabla(u_2 - u_1) \right| + \left| \int (e^{q_2}(e^{\beta u_2} - e^{-\beta u_2}) - e^{q_1}(e^{\beta u_1} - e^{-\beta u_1}))(u_2 - u_1) \right|$$
$$\le \left| \int (e^{q_1} - e^{q_2})\nabla u_2 \cdot \nabla(u_2 - u_1) \right| + \left| \int (e^{q_2} - e^{q_1})(e^{\beta \max(\lambda_1, \lambda_2)} - e^{-\beta \max(\lambda_1, \lambda_2)})(u_2 - u_1) \right|$$

hold because of the triangle inequality and the pointwise estimates $\kappa_i \le u_i \le \lambda_i$, $i \in \{1, 2\}$, for the solution of the nonlinear Poisson-Boltzmann equation obtained by Proposition 8, where κ_i and λ_i are constants.

Now we use the Cauchy-Schwarz and Poincaré inequalities to find

$$I \le \|e^{q_1} - e^{q_2}\|_{L^\infty}\|\nabla u_2\|_{L^2}\|\nabla(u_2 - u_1)\|_{L^2}$$
$$+ C_p|e^{\beta \max(\lambda_1, \lambda_2)} - e^{-\beta \max(\lambda_1, \lambda_2)}|\|e^{q_1} - e^{q_2}\|_{L^\infty}\|\nabla(u_2 - u_1)\|_{L^2},$$

where C_p is a Poincaré constant. Therefore, we can write

$$e^{-\|q_1\|_{L^\infty}}\|\nabla(u_2-u_1)\|_{L^2} \leq I \leq \|e^{q_1}-e^{q_2}\|_{L^\infty}(\|\nabla u_2\|_{L^2}+C_p|e^{\beta\max(\lambda_1,\lambda_2)}-e^{-\beta\max(\lambda_1,\lambda_2)}|).$$

Using equations (4.38) and (4.39), we obtain

$$e^{-\|q_1\|_{L^\infty}}\|\nabla(u_2-u_1)\|_{L^2} \leq I \leq \|q_1-q_2\|_{L^\infty}e^{\max(\|q_1\|_\infty,\|q_2\|_\infty)}\times$$
$$\left(e^{2\|q_2\|_\infty}\left(\|\nabla\overline{g}\|_{L^2}+C_pe^{|\beta\lambda_2|}+C_pe^{|\beta\kappa_2|}\right)+C_p|e^{\beta\max(\lambda_1,\lambda_2)}-e^{-\beta\max(\lambda_1,\lambda_2)}|\right),$$

which leads to

$$\|\nabla(u_2-u_1)\|_{L^2} \leq \|q_1-q_2\|_{L^\infty}e^{4\max(\|q_1\|_\infty,\|q_2\|_\infty)}\times$$
$$\left(\|\nabla\overline{g}\|_{L^2}+C_pe^{|\beta\lambda_2|}+C_pe^{|\beta\kappa_2|}+C_p|e^{\beta\max(\lambda_1,\lambda_2)}-e^{-\beta\max(\lambda_1,\lambda_2)}|\right)$$
$$= S_0\|q_1-q_2\|_{L^\infty}e^{4\max(\|q_1\|_\infty,\|q_2\|_\infty)},$$

where $S_0 := \|\nabla\overline{g}\|_{L^2}+C_pe^{|\beta\lambda_2|}+C_pe^{|\beta\kappa_2|}+C_p|e^{\beta\max(\lambda_1,\lambda_2)}-e^{-\beta\max(\lambda_1,\lambda_2)}|$. Therefore, using inequality (4.36), we can write

$$\|u_1-u_2\|_{H^1} \leq (1+C_p^2)^{1/2}\|\nabla(u_1-u_2)\|_{L^2}$$
$$\leq (1+C_p^2)^{1/2}\left(\|\nabla\overline{g}\|_{L^2}+C_pe^{|\beta\lambda_2|}+C_pe^{|\beta\kappa_2|}+C_p|e^{\beta\max(\lambda_1,\lambda_2)}-e^{-\beta\max(\lambda_1,\lambda_2)}|\right)\times$$
$$e^{4\max(\|q_1\|_\infty,\|q_2\|_\infty)}\|q_1-q_2\|_{L^\infty}$$
$$= Se^{4\max(\|q_1\|_\infty,\|q_2\|_\infty)}\|q_1-q_2\|_{L^\infty},$$

where $S := S_0(1+C_p^2)^{1/2}$. This completes the proof for the nonlinear equation (4.28). $\square$

Remark 2. *The quantity $\|\nabla\overline{g}\|_{L^2(D)}$ is non-zero even if the Dirichlet datum g is a constant. The new variable $v := u - \overline{g}$ is defined such that $\overline{g} = g$ on ∂D and $\overline{g}$ is arbitrary in D. Therefore in the (realistic) case of non-constant g, $\|\nabla\overline{g}\|_{L^2(D)}$ is non-zero in D. Hence, even if $f = 0$, the quantities F, H, and S are non-zero.*

We summarize the above results for linear and nonlinear elliptic inverse problems in the following theorem for both kinds of observation operator G.

Theorem 9. *Assume $G(q) = u$ is the observation operator representing the solution u of the real-valued linear equation (4.27) and the nonlinear equation (4.28) in the bounded domain $D \subset \mathbb{R}^n$, $n \in \{2,3\}$, with a smooth boundary ∂D and $A = \exp(q) = \eta$ with $q \in L^\infty(D)$. Then the estimates*

$$|G(q)| \leq Z\exp(2\|q\|_{L^\infty(D)}) \tag{4.41}$$

and

$$|G(q_1)-G(q_2)| \leq Z\exp(4\max\{\|q_1\|_{L^\infty(D)},\|q_2\|_{L^\infty(D)}\})\|q_1-q_2\|_{L^\infty(D)} \tag{4.42}$$

hold, where $Z =: F = F(\|\nabla\overline{g}\|_{L^2},\|f\|_{L^2})$ in these two inequalities becomes $Z =: H = H(\|\nabla\overline{g}\|_{L^2})$ for the linear equation (4.27) and becomes $Z =: S = S(\|\nabla\overline{g}\|_{L^2})$ for the nonlinear equation (4.28), which have been defined in Proposition 9.

The estimates (4.41) and (4.42) for real and imaginary parts of the solution yield estimates for the complex-valued equation (4.27) as well. Therefore, Assumptions 8 are satisfied for the general EIT model equations (4.27), where g is constant at each contact at the surface of the device. This will lead to the main results of the paper, i.e., well-definedness and well-posedness of the Bayesian inversion problem for linear and nonlinear elliptic problems including the EIT inverse problem.

The posterior probability measure μ^y defined by (4.23) is well-defined if we show that the measure is normalizable. For well-posedness, the continuity of the posterior measure in the Hellinger metric with respect to the data must be shown. We first define Hellinger distance and then state the two main theorems resulting from Proposition 9. The reader is referred to [52] for the details of the following proofs.

Definition 8 (Hellinger distance)**.** *Let μ and μ' be two absolutely continuous probability measures with respect to measure ν. Then the Hellinger distance between μ and μ' is defined as*

$$d_{\text{Hell}}(\mu, \mu') := \sqrt{\left(\frac{1}{2} \int \left(\sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\nu}} - \sqrt{\frac{\mathrm{d}\mu'}{\mathrm{d}\nu}}\right)^2 \mathrm{d}\nu\right)}, \tag{4.43}$$

where $\mathrm{d}\mu/\mathrm{d}\nu$ and $\mathrm{d}\mu'/\mathrm{d}\nu$ are the Radon-Nikodym derivatives of μ and μ', respectively, with respect to ν.

Theorem 10 (Well-definedness of the posterior measure [52, Theorem 4.1])**.** *Let G satisfy Assumptions 8 and assume that the prior measure μ_0 is a Gaussian measure satisfying $\mu_0(X) = 1$. Then the posterior measure μ^y given by (4.23) is a well-defined probability measure.*

Proof. Since G satisfies Assumptions 8, the function Φ satisfies Assumptions 2.6 in [52]. Using these assumptions and the Fernique Theorem, the normalization constant $C(y)$ in (4.23) is bounded from below and above and hence the posterior measure is normalizable and well-defined. $\qquad\square$

The following theorem states the well-posedness for inverse problems by showing Lipschitz continuity of the posterior measure in the Hellinger metric with respect to changes in the data.

Theorem 11 (Well-posedness of the Bayesian inverse problem [52, Theorem 4.2])**.** *Let G satisfy Assumptions 8. Assume also that the prior measure μ_0 is a Gaussian measure satisfying $\mu_0(X) = 1$ and that the measure is absolutely continuous, $\mu^y \ll \mu_0$, with its Radon-Nikodym derivative given by (4.23) for each $y \in Y$.*

Then the posterior measure μ^y is Lipschitz continuous in the data y with respect to the Hellinger distance, i.e., if μ^y and $\mu^{y'}$ are two measures corresponding to data y and y', then there exists $\alpha = \alpha(r) > 0$ such that the inequality

$$d_{\text{Hell}}(\mu^y, \mu^{y'}) \le \alpha \|y - y'\|_Y$$

holds for all y and y' with $\max\{\|y\|_Y, \|y'\|_Y\} < r$.

Proof. Let $C = C(y)$ and $C' = C(y')$ denote the normalization constants

$$
\begin{aligned}
C &= \int_X \exp\left(-\Phi(q, y)\right) \mathrm{d}\mu_0(q), \\
C' &= \int_X \exp\left(-\Phi(q, y')\right) \mathrm{d}\mu_0(q)
\end{aligned}
$$

for the posterior measures μ^y and $\mu^{y'}$, respectively, which can be bounded from below using Assumptions 2.6 in [52]. Applying the Fernique Theorem and using the assumptions on the function Φ, we find

$$
|C - C'| \le \alpha \|y - y'\|_Y, \tag{4.44}
$$

where $\alpha > 0$ is a constant. From the definition of the Hellinger distance, we have

$$
\begin{aligned}
2d_{\text{Hell}}(\mu^y, \mu^{y'})^2 &= \int_X \left(C^{-1/2} \exp\left(-\frac{1}{2}\Phi(q, y)\right) - (C)'^{-1/2} \exp\left(-\frac{1}{2}\Phi(q, y')\right) \right)^2 \mathrm{d}\mu_0(q) \\
&\le I_1 + I_2,
\end{aligned}
$$

where

$$
\begin{aligned}
I_1 &:= \frac{2}{C} \int_X \left(\exp\left(-\frac{1}{2}\Phi(q, y)\right) - \exp\left(-\frac{1}{2}\Phi(q, y')\right) \right)^2 \mathrm{d}\mu_0(q), \\
I_2 &:= 2|C^{-1/2} - (C')^{-1/2}|^2 \int_X \exp\left(-\Phi(q, y')\right) \mathrm{d}\mu_0(q).
\end{aligned}
$$

Now again using Assumptions 2.6 in [52] and the Fernique Theorem, we obtain the estimate

$$
\frac{C}{2} I_1 \le \alpha \|y - y'\|_Y^2
$$

and that the integral in I_2 is finite. Furthermore, we find

$$
|C^{-1/2} - (C')^{-1/2}|^2 \le \alpha \|y - y'\|_Y^2
$$

using the lower bounds for C and C', which results in a similar bound for I_2. Combining the bounds gives the desired continuity in the Hellinger metric and completes the proof. $\square$

4.2.4. Markov-Chain Monte-Carlo (MCMC) Methods

As the computational cost of the integral in (4.14) in Bayes' Theorem is expensive, we need techniques such as Monte-Carlo methods when the parameter space is high-dimensional. Efficiently sampling the parameter space is essential in the sense that in order to specify parameter values to evaluate the density, we need algorithms which search the geometry of the distribution efficiently. Otherwise most of the space is empty in high dimensions. To this end, we use Markov chains with the desired stationary distribution. Markov-chain Monte-Carlo methods are a class of Monte-Carlo methods with the general idea of constructing (time-reversible) Markov chains whose stationary distribution is the posterior density [104].

The Metropolis-Hastings algorithm is an MCMC algorithm to draw samples from a desired distribution. In this algorithm, the first state of the chain q_0 is chosen and then the

new state q_k, $k = 1, 2, \ldots, N$, of the chain is constructed based on the previous state q_{k-1}. To this end, a new value q^* using the proposal density function $J(q^*|q_{k-1})$ is proposed. Admissibility of this proposed value is tested by means of calculating the acceptance ratio $\alpha(q^*|q_{k-1})$. If the proposed value is admissible, it is accepted as q_k, otherwise the old value is kept and a new proposal is made.

The main question in constructing the MCMC methods is the evolution method of the chain: Can we find a transition rule r going from q_{k-1} to q_k such that the stationary distribution exists and equals the posterior distribution? One answer to this question is the Metropolis transition rule, which is defined by

$$r(q_{k-1} \to q_k) := \frac{\pi(q_k)}{\pi(q_{k-1})}. \tag{4.45}$$

The rule states if $r > 1$, accept q_k. Otherwise if $0 < r < 1$, reject it and keep the old value so that $q_k = q_{k-1}$. Although this rule works in a continuous parameter space with infinitely many points q, when new candidate points are selected at random, but to concentrate the sampling in good areas, the next point is selected dependent on the last point according to a proposal/jumping function $J(q_k|q_{k-1})$. Hence the ratio of probabilities (4.45) can be modified as

$$r(q_{k-1} \to q_k) = \frac{\pi(q_k)}{\pi(q_{k-1})} \cdot \frac{J(q_{k-1}|q_k)}{J(q_k|q_{k-1})}, \tag{4.46}$$

which leads to the Metropolis-Hastings algorithm.

The number of samples or length of the chain needed in MCMC methods depends on the problem. One usually discards as many parameter values as needed to be sure that the starting point is more or less random and a stationary distribution has been reached, which is called the burn-in period. For more details about MCMC methods see for example [109–112]. Although the convergence speed is determined by choice of a good proposal distribution, at least tens or hundreds of thousands samples are necessary to converge to the target distribution. Choosing the optimal proposal scaling is a crucial issue and affects the MCMC results; if the covariance of the proposal distribution is too small, the generated Markov chain moves too slowly, and if it is too large, the proposals are rejected. Hence, optimal proposal values should be found to avoid both extremes, which leads to adaptive MCMC methods [113–115]. In the following section, we will consider an adaptive algorithm that helps us with sampling from potentially complicated distributions.

4.2.5. Delayed-Rejection Adaptive-Metropolis (DRAM) Algorithm

Searching for a good proposal value can be done manually through trial and error, but this becomes intractable in high dimensions. Therefore, adaptive algorithms that find optimal proposal scales automatically are advantageous.

The delayed-rejection adaptive-Metropolis (DRAM) algorithm is an efficient adaptive MCMC algorithm [114]. It is based on the combination of two powerful ideas to modify the Markov-chain Monte-Carlo method, namely adaptive Metroplolis (AM) [116, 117] and delayed-rejection (DR) [118, 119], which are used as global and local adaptive algorithms, respectively. AM updates the proposal covariance matrix with an optimal scale and ensures that information leaned about the posterior distribution is remembered as the chain

progresses, while DR updates the proposal scale when the proposed value is rejected to improve mixing and avoid stagnation of the chain.

The adaptive Metropolis algorithm is a global adaptive strategy, where a recursive relation is used to update the proposal covariance matrix. In this algorithm, we take the Gaussian proposal centered at the current state of the chain q_k and update the chain covariance matrix at the k-th step using

$$V_k = s_p \text{Cov}(q_0, q_1, \ldots, q_{k-1}) + \varepsilon I_p, \tag{4.47}$$

where s_p is a design parameter and depends only on the dimension p of the parameter space. This parameter is specified as $s_p := 2.38^2/p$ as the common choice for Gaussian targets and proposals [120], as it optimizes the mixing properties of the Metropolis-Hastings search in the case of Gaussians. Furthermore, I_p denotes the p-dimensional identity matrix, and $\varepsilon > 0$ is a very small constant to ensure that V_k is not singular theoretically, and in most cases it can be set to zero [114].

The adaptive Metropolis algorithm employs the recursive relation

$$V_{k+1} := \frac{k-1}{k} V_k + \frac{s_p}{k} \left(k \overline{q}_{k-1} \overline{q}_{k-1}^\top - (k+1) \overline{q}_k \overline{q}_k^\top + q_k q_k^\top \right)$$

to update the proposal covariance matrix, where the sample mean $\overline{q}_k$ is calculated recursively by

$$\begin{aligned}
\overline{q}_k &= \frac{1}{k+1} \sum_{i=0}^{k} q_k \\
&= \frac{k}{k+1} \cdot \frac{1}{k} \sum_{i=1}^{k} q_{k-1} + \frac{1}{k+1} q_k \\
&= q_k + \frac{k}{k+1} (\overline{q}_{k-1} - q_k).
\end{aligned}$$

The basic idea of the DR algorithm is that, if the proposed value q^* for q_k is rejected, this algorithm provides a mechanism for constructing an alternative candidate q^{**} instead of retaining the previous chain value q_{k-1} as in the standard Metropolis algorithms. This process is called delaying rejection, which can be done for one or many stages. Furthermore, the acceptance probability of the new candidate(s) is calculated. Therefore, in the DR process, the previous state of the chain is updated using the optimal parameter scale or proposal covariance matrix that has been calculated via the AM algorithm. The DRAM algorithm has been summarized in Algorithm 1.

If q^{**} is rejected, a third-stage proposal (and corresponding acceptance condition) can be built. This process can be continued and a jth-stage proposal can be constructed if the $(j-1)$th-stage proposal is rejected [114]. We have written a package in Julia [55] in order to implement the DRAM as an adaptive Markov-chain Monte-Carlo method for Bayesian analysis. The Bayesian inference results of the EIT inverse problem by means of the adaptive algorithm are presented in Section 4.3.

Algorithm 1 The DRAM algorithm

Initialization:

Choose the first state of the chain q_0 such that $\pi_0(q_0) > 0$.

Choose the number N_{samples} of samples or iterations.

Choose the parameter ε.

Choose the initial proposal covariance matrix V_0 (diagonal or symmetric).

Choose the factor γ (often $\gamma := 1/5$) for the second-stage proposal distribution.

for $k = 1 : N_{\text{samples}}$ **do**

   1.   (Adaptivity:) The covariance matrix V_k in the k-th step is updated by (4.47).

   2.   A first-stage proposal q^* is generated from $J(q^*|q_{k-1}) := N(q_{k-1}, V_k)$.

   3.   The new value q^* is accepted with probability

$$\alpha(q^*|q_{k-1}) = \min\left(1, \frac{\pi(q^*)}{\pi(q_{k-1})} \cdot \frac{J(q_{k-1}|q^*)}{J(q^*|q_{k-1})}\right). \tag{4.48}$$

   4.   If the new state is accepted, we set $q_k = q^*$. Otherwise:

      a)   (Delayed rejection:) A second-stage proposal q^{**} is generated from proposal density

$$J_2(q^{**}|q_{k-1}, q^*) := N(q_{k-1}, \gamma^2 V_k), \tag{4.49}$$

          where V_k is the adapted covariance matrix.

      b)   The new value q^{**} is accepted with probability

$$\alpha_2(q^{**}|q_{k-1}, q^*) := \min\left(1, \frac{\pi(q^{**}|y)J(q^*|q^{**})(1 - \alpha(q^*|q^{**}))}{\pi(q_{k-1}|y)J(q^*|q_{k-1})(1 - \alpha(q^*|q_{k-1}))}\right). \tag{4.50}$$

      c)   If the new state is accepted, we set $q_k := q^{**}$, otherwise $q_k := q_{k-1}$.

end for

4.3. Numerical Experiments

In this section, we show numerical results for the EIT problem presented in this work including both forward and inverse problems. The numerical results for the forward problem include solving the nonlinear Poisson-Boltzmann equation on the prescribed physical domain and the inverse problem results include simultaneous reconstruction of multiple parameters of the EIT system. Given a sufficient number of electrical current measurements and corresponding voltage patterns, we aim to identify the EIT parameters including electrical and physical parameters, namely the electrical permittivity, the electrical charge, and the size of the inclusion inside the physical domain. To solve this inverse problem, we will use Bayesian techniques, where prior information about the parameters of interest is updated using measurement data (either real world measurements or other computational models such as more refined ones) to obtain posterior information about the parameters.

4.3.1. Results for the Forward Model

Here, we present numerical results of solving the presented EIT nonlinear model (4.7) for an 8-electrode set-up as illustrated in Figure 4.1. In this set-up, we consider a square domain $D = [-1, 1]^2$ as the main body with eight contact electrodes disjointly attached to the surface of the body, two on each side of the square domain. Moreover, the electrodes are labeled counterclockwise starting from the left bottom one, which is labeled as electrode 1, to the bottom left electrode, which is labeled as electrode 8. We assume that the background medium is a continuum dielectric of permittivity 80. The background medium contains an inclusion centered at $(0.5, 0)$ in the domain, and made of a different material and thus different permittivity, whose size, charge and permittivity are quantities of interest in solving the EIT inverse problem here.

The applied voltage patterns are shown in Table 4.1. The choice of the voltage patterns or Dirichlet boundary conditions for the forward problem is important, since it effects the resulting currents and consequently the conservation of current is affected. For instance, if the voltages applied on the contacts 2 and 5 are the same with different signs, say $U_2 = 0.01$ and $U_5 = -0.01$, then the inclusion which is located in between acts as a capacitor and stores potential energy in the electric field. This leads to a reduction in the level of the resulting current values. Here, we have chosen the voltage pattern by skipping one contact and applying ± 0.01 V. Changing the locations of the applied voltages and always skipping one contact, three more voltage patterns can be made. The four patterns can be seen in Table 4.1.

Pattern	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8
1	0.01	0	-0.01	0	0.01	0	-0.01	0
2	0	0.01	0	-0.01	0	0.01	0	-0.01
3	-0.01	0	0.01	0	-0.01	0	0.01	0
4	0	-0.01	0	0.01	0	-0.01	0	0.01

Table 4.1.: Four different applied voltage patterns (in Volt) in an 8-electrode EIT device configuration. The indices of the voltages are the labels of the electrodes.

The results of the presented forward EIT model (4.7) on the prescribed geometry are shown in Figure 4.3. This figure illustrates the potential u calculated for four different settings of boundary conditions (see Table 4.1). The solution to the physical model is approximated using the first order finite-element and the mesh generation for the prescribed geometry of the device is done using GMSH package. The mesh is generated in such a way that it is aligned with the inclusion in the sense that each element has a constant value for the coefficient A, and the leading equation for the inclusion and the background medium is respectively linear and nonlinear defined.

4.3.2. Bayesian Inference Results

The numerical results for the nonlinear EIT inverse problem are presented here. Our goal is simultaneous extraction of multiple parameters which reflect physical and electrical
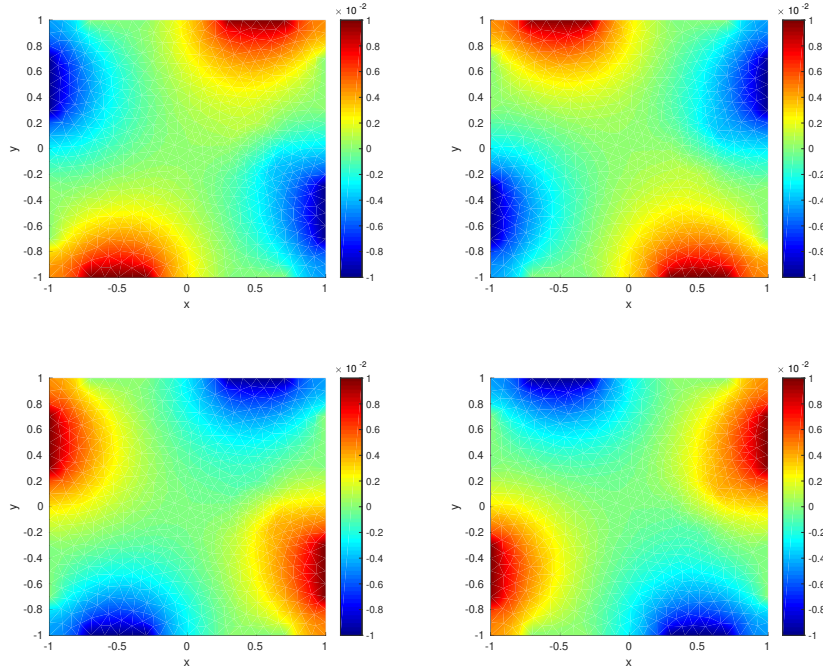
Figure 4.3.: Electrical potential for the 8-electrode EIT device configuration with an inclusion centred at $(0.5, 0)$. The figures illustrate voltage patterns 1-4 (clockwise, starting from top left) mentioned in Table 4.1.

properties of the body interior in the framework of our theoretical results. To this end, we implement an adaptive MCMC method, namely the delayed-rejection adaptive Metropolis (DRAM) algorithm as a package in Julia as this inversion algorithm is capable of dealing with the nonlinearity and ill-posedness of our problem. As mentioned in Section 4.2, the EIT inverse problem is to identify physical and electrical properties of the body interior given a sufficiently large number of electrical current measurements and corresponding voltage patterns. These parameters include size R, permittivity A, and charge Q of the inclusion contained in the body and surrounded by the medium, which have been modeled by nonlinear model (4.7). The results are for the 8-electrode EIT device configuration, whose solution of the corresponding forward problem is shown in Figure 4.3. In this work, the electrical currents calculated on the eight electrodes in four different modes (voltage patterns or boundary conditions) by the nonlinear forward model solver for the parameter values $A = 3, R = 0.2$, and $Q = 1$ are assumed to be the exact parameter values.

In Figure 4.4, sample paths obtained using the DRAM algorithm are shown, which display convergence of the Markov chains of the three parameters. The chain of the middle parameter R has tiny oscillations around the exact value 0.2 of the parameter and converges quickly to this value. The chains of the other two parameters converge to their exact values, too. Regardless of the chosen initial state, the produced chains jump to around the exact

parameter values in a short burn-in period and jump around it using the proposed scale (covariance) by the algorithm till the end of the iterations, which shows robustness of the applied algorithm to the EIT inverse problem. We examine the convergence of the generated chains in the next section by looking at how independent the produced samples are and how good the mixing of the chains is.
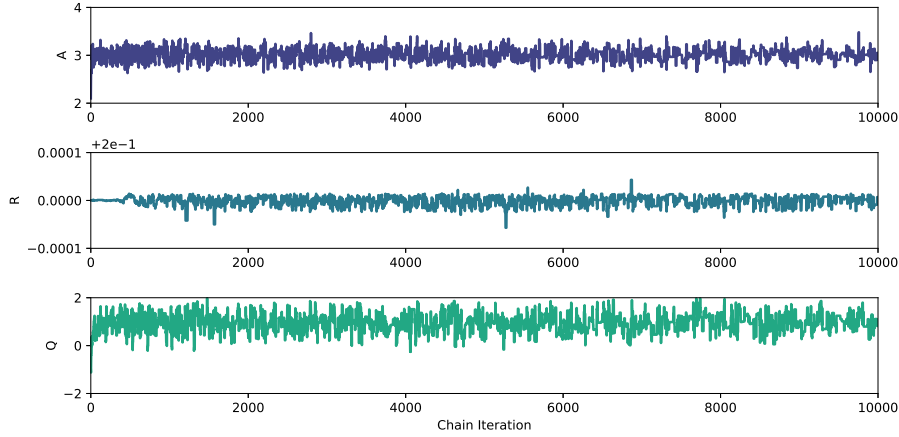

Figure 4.4.: Trace plots for the 8-electrode EIT device configuration obtained by the DRAM algorithm for the parameters permittivity A, size R, and charge Q. The figures display convergence of the parameters to their exact values, i.e., $A = 3$, $R = 0.2$, and $Q = 1$.

Figure 4.5 illustrates correlation between each pair of the parameters and their marginal posterior distributions. Bayesian inference results show that the size of the inclusion can be extracted very accurately. Permittivity and charge (and charge sign) of the inclusion are the other parameters of interest that can be recovered nicely using Bayesian inversion. To better interpret the histogram plots, we calculate confidence intervals for each parameter.

The whole computational method including the forward model solver and the MCMC calculations is computationally expensive since the physical model must be evaluated for thousands of samples iteratively. The intensive work is spent on the generation of the mesh due to the fact that the size of the inclusion is one of the quantities of interest and therefore the process of mesh generation must be iterated for each sample.

Autocorrelation Function

To statistically testing the convergence of the generated Markov chains, we look at the correlation between the samples with calculating the autocorrelation function (ACF), which is defined by

$$\text{ACF}\,(N, t) := \frac{\sum_{k=1}^{N-t}(q_k - \overline{q})(q_{k+t} - \overline{q})}{\sum_{k=1}^{N}(q_k - \overline{q})^2}, \tag{4.51}$$

where q_k and $\overline{q}$ are respectively the k-th state of the Markov chain and the sample mean [121]. This Formula is based on the idea of computing the correlation between the subchains of
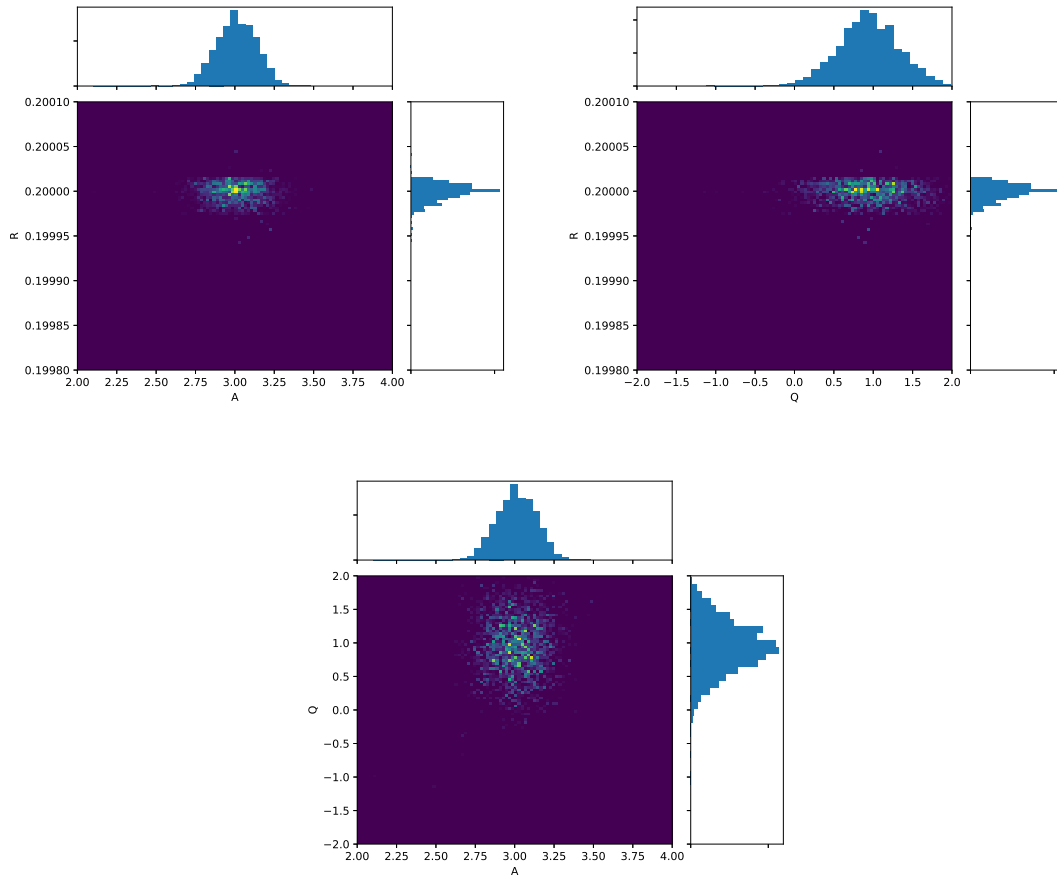
Figure 4.5.: Two-dimensional histograms and marginal posterior distributions for the 8-electrode EIT device configuration and for 10 000 samples obtained by the DRAM algorithm for the parameters A, Q and R. The figures show correlations between each pair of the parameters.

length N with lag t. The numerator of Equation (4.51) is an estimate of the autocovariance and the denominator is an estimate of the variance of the chain in order to normalize the output. Low autocorrelation means that samples are independent and mixing in the produced chain is good. Figure 4.6 displays plots of autocorrelation function for the three Markov chains which are shown in Figure 4.4.

Confidence Interval

As an advantage of the statistical method of Bayesian inference, we can find confidence interval containing admissible and optimal values for each parameter of interest if we could reconstruct it. Figure 4.7 displays the marginal histograms of the posterior distributions of the parameters A, Q and R as the quantities of interest in the EIT problem at hand.
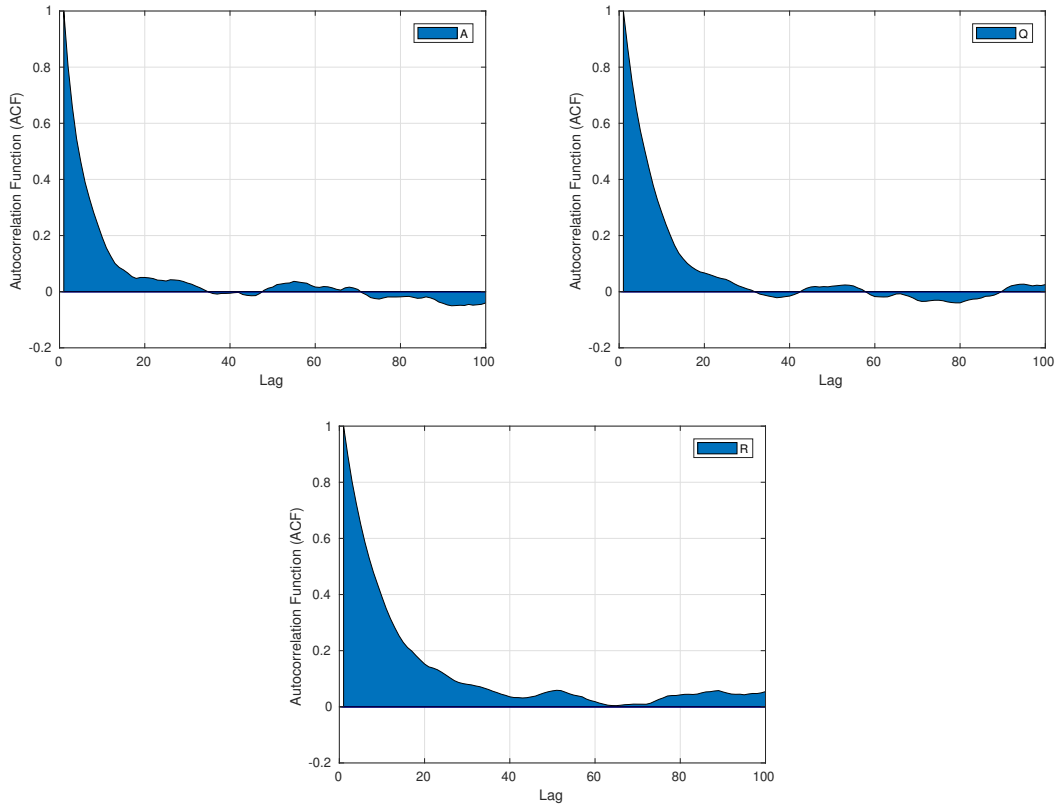
Figure 4.6.: Autocorrelation functions of the generated Markov chains by the DRAM algorithm for the three parameters of interest.

The histograms have been plotted according to the posterior Markov chain using $10\,000$ samples for the 8-electrode EIT device configuration. The mean of the generated chains and exact values of the parameters are indicated on the figures. The estimated mean value of the generated posterior chains for the parameter A, R and Q are 3.01, 0.1999988 and 0.94, respectively. Having computed the mean and standard deviation of the chains, we can calculate the interval of admissible values of the parameters with 95% of confidence. According to Bayesian estimation results, this interval for parameters A, Q and R respectively are $A \in [2.76, 3.26]$, $R \in [0.199979, 0.200019]$, and $Q \in [0.18, 1.69]$. Therefore, the size of the inclusion can be recovered more accurately in the defined EIT problem.

4.4. Summary

In this work, we present a comprehensive physical model for the electrical-impedance tomography sensors with applications in medicine and healthcare. In the EIT applications for instance in medicine, there are big data sets of measurements which need to be statistically analyzed to infer the electrical properties in order to help diagnosis and treatment. We formulate the Bayesian estimation for the new governing model in a measure-theoretic
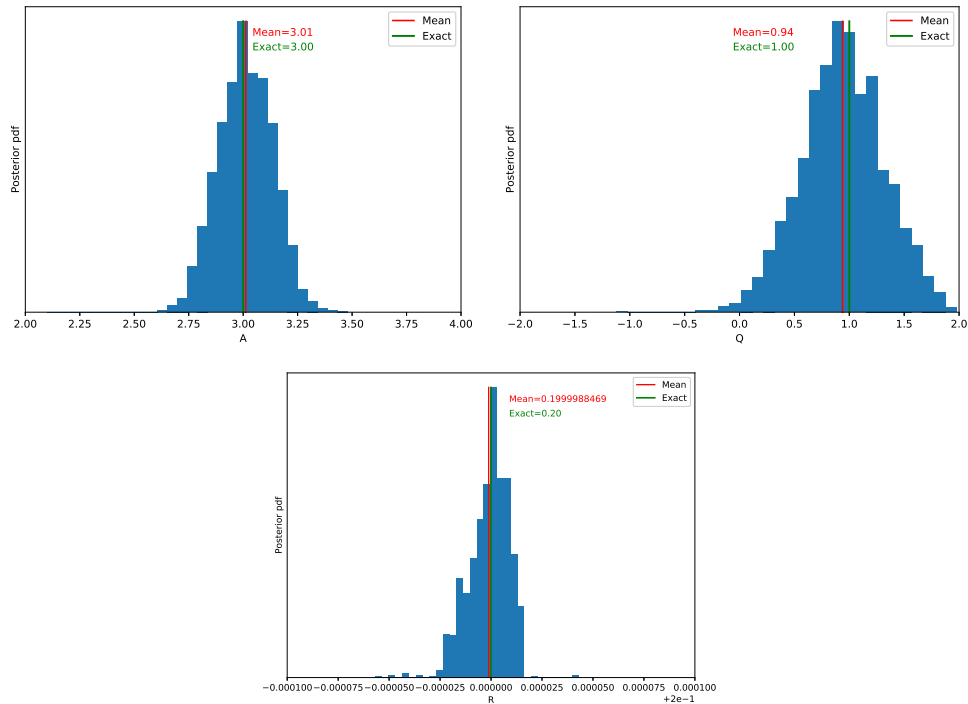
Figure 4.7.: Marginal histograms of posterior distributions of the parameters A, Q and R produced by means of the DRAM algorithm using 10 000 samples. The mean of the produced chain and the exact value of the parameters are indicated on the figures.

framework and in an infinite-dimensional setting and present the first application of the Bayesian inference for the simultaneous extraction of multiple electrical and geometrical parameters in the EIT technology. We also prove well-definedness of the posterior measure and well-posedness of the Bayesian inversion for the presented nonlinear model. Numerical results of the computational MCMC-FEM include calculating electrical potential in the EIT sensor using FEM, and calculation of the sizes, charges, and permittivities of inclusions in the body interior under study of the EIT technology using an adaptive MCMC method, which support our theoretical findings.

5. Bayesian Inversion for a Biofilm Model Including Quorum Sensing

This chapter focuses on usually monomicrobial biofilms spreading on surfaces of material implanted in bones or tissues allowing little space for complex three-dimensional biofilms. These biofilms are the ones that are most relevant in the clinic.

In order to understand the growth of biofilms and their response to antibiotic therapy in vivo, we have developed a mathematical model based on parabolic partial differential equations (PDE) that describes the time-dependent evolution of the size of the biofilm. The model contains parameters such as growth rate and cooperation that cannot be determined directly from experimental data. We therefore solve the corresponding inverse problem to estimate these important parameter values. In particular, we use Bayesian inversion and Markov-chain Monte-Carlo techniques as they are capable of dealing with ill-posed inverse problems. The ultimate goal is to provide validated and predictive models for biofilm formation under conditions relevant in the clinical setting. Such predictive models are invaluable in decision making for optimal treatment.

This chapter is organized as follows. In Section 5.1, we introduce the proposed PDE model equations as well as the required initial and boundary conditions. In Section 5.2, existence and uniqueness of solutions to the proposed model equations are discussed. Simulation results of the growth of biofilms as the forward model under consideration are presented in Section 5.3. Section 5.4 is devoted to the experimental findings. In Section 5.5, we define the biofilms inverse problem and describe two different approaches, namely a genetic algorithm as a global, deterministic method and Bayesian inversion as a tool from uncertainty quantification. Numerical results of solving the biofilm inverse problem using these two approaches and hence determining the parameter values are presented in this section as well. Finally, summary of the chapter is in Section 5.6.

This chapter is based on the author's work [122].

5.1. The Biofilm Model

Mathematical models come in many forms that can range from very simple empirical correlations to sophisticated and computationally intensive algorithms that describe three-dimensional biofilm morphology and activity. The model we propose captures two main characteristics of a biofilm: It tracks its growth and degradation depending on environmental factors and it includes the emergence of resistance (i.e., cooperation against environmental factors).

We first introduce the model equations and then discuss the meaning of the various terms and how they relate to different behaviors of bacteria forming a biofilm.

5.1.1. The Model Equations

The mathematical model is a system that consists of a parabolic partial differential equation, which is well-known as a reaction-diffusion equation describing biofilm growth, coupled to an ordinary differential equation (ODE) that describes cooperation among the bacteria including quorum sensing [123]. This type of system of equations is well-known as a reaction-diffusion-ODE model and has been applied to the modeling of, for example, interactions between cellular processes such as cell growth in mathematical biology [124]. As mentioned above, the model has two outputs: the concentration of bacteria, whose evolution in time shows the growth and degradation of the biofilm, and the cooperation of the bacteria in the biofilm exhibiting the strength of resistance and protection against environmental factors.

Our model equations are the reaction-diffusion-ODE system

$$\partial_t u = \nabla \cdot (A\nabla u) + \alpha u(1 - u/\beta) - \gamma(t, x, y)(1 - v)u \quad \text{in } \mathbb{R}^+ \times D, \qquad (5.1a)$$

$$\partial_t v = \rho \max\big(0, \arctan\big(q(u) - \nu\big)\big)v - \kappa v^2 \qquad \qquad \text{in } \mathbb{R}^+ \times D, \qquad (5.1b)$$

where

$$
\begin{aligned}
q(u)(t, x, y) &:= \big(u(t, \cdot, \cdot) * G(\cdot, \cdot)\big)(x, y) \\
&= \int_{\mathbb{R}^2} u(t, \xi, \eta)G(x - \xi, y - \eta)\mathrm{d}\xi\mathrm{d}\eta \quad \text{in } \mathbb{R}^+ \times D
\end{aligned}
\qquad (5.2)
$$

and

$$G(x, y) := \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad \text{in } \mathbb{R}^+ \times D. \qquad (5.3)$$

Here $u(t, x, y)$ is the concentration of bacteria. The variables x and $y \in D$ denote position, where D is a bounded domain in $\mathbb{R}^2$ with the boundary ∂D and $t \in \mathbb{R}^+$ denotes time. Since each bacterium has a finite size, the value of $u(t, x, y)$ corresponds to the thickness of the biofilm at position (x, y). The variable $v(t, x, y)$ denotes cooperation. A value of $v = 0$ means no cooperation, while $v = 1$ means maximal cooperation.

We call $q(u)$ the quorum functional. We define it as the convolution of the bacteria concentration $u(t, x, y)$ with a two-dimensional Gaussian $G(x, y)$ with variance σ^2. It measures the population density in a neighbourhood of (x, y) at time t and is used in the quorum sensing model.

Finally, A, α, β, ρ, κ, and ν denote positive constants and $\gamma(t, x, y)$ is a function from $\mathbb{R}_0^+ \times \mathbb{R}^2$ to $\mathbb{R}_0^+$. These constants depend on the bacteria strain and on environmental factors. They are predictive factors that will be determined and that will make it possible to predict the development of a patient's biofilm thus allowing to choose the optimal treatment.

In the following, the rationale behind the model equations is explained.

The term $\nabla \cdot (A\nabla u)$ describes the spreading of the bacteria according to a diffusion process. Different strategies for the bacteria, given the same energy budget, correspond to different relative sizes of the constants A and γ: Spreading faster (larger A) in a thinner biofilm that provides less protection because there is less cooperation (smaller γ) on the one hand, or spreading slower (smaller A) and building a better protected biofilm with more cooperation (larger γ) on the other hand.

The term αu describes the increase of the number of bacteria using a constant growth rate $\alpha > 0$. We assume that enough nutrients are present to sustain the growth of the biofilm.

Hence this means that a reduction of the amount of nutrients left is not explicitly included in this model. (Of course, coupling the growth of the biofilm to a variable amount of nutrients is a possible model extension. Another model extension is to model the spreading of the bacteria depending on the local food concentration, and to model the food concentration by an additional equation.)

The term $-\alpha u^2/\beta$ implies that there is an upper limit for the thickness of the biofilm. $\beta > 0$ is a constant and its value best corresponding to measurements will be found in Subsection 5.5.1. This value also gives the maximum depth of the biofilm. If we consider a logistic ordinary differential equation

$$\partial_t u(t) = \alpha u(t)(1 - u(t)/\beta),$$

then

$$\lim_{t \to \infty} u(t) = \beta$$

holds for the long-term limit.

The term $-\gamma(t)(1-v)u$ with a function $\gamma \colon \mathbb{R}_0^+ \times \mathbb{R}^2 \to \mathbb{R}_0^+$ describes the degradation of the biofilm due to environmental factors and antimicrobial agents. The function γ depends on the time of adding antibiotic substances. γ is not only a constant, but a function, in order to be able to investigate the solution after an antibiotic substance has been added after the biofilm has grown for some time. In many situations, γ is a piecewise constant function. The factor $1 - v$ means that the adverse effect of the environment on the biofilm is reduced whenever there is more cooperation represented by v. Of course, $1 - v$ could be replaced by any function of v that vanishes for $v = 1$, that equals 1 for $v = 0$, and that is monotonically decreasing.

The variable $v(t, x, y)$ is proportional to the extent of cooperation among the bacteria in the biofilm. There is increasing empirical evidence that suggests that bacteria cooperate by producing exoproducts or public goods such as components of the matrix in the biofilm [125, 126]. Hence $v(t, x, y)$ denotes the strength of the resistance or protection that the bacteria develop against environmental influences. If their concentration u is low, they do not use energy to develop resistance. However, if their local average concentration $q(u)$ is above a threshold $\nu > 0$, the quorum threshold (i.e., $q(u) - \nu > 0$), the bacteria start to cooperate and spend some of their excess energy on cooperative measures such as strengthening the matrix of the biofilm. v is in the interval $[0, 1]$, where 0 denotes no cooperation and 1 means maximum cooperation. Otherwise the term including q is zero.

The concept of quorum sensing has been included in biofilm models and its different aspects have been studied by many authors (see e.g. [127–129]). For example, in [129], the authors very recently formulated a mathematical model of QS as a stress response mechanism that increases resistance against antibiotics. In the current paper, we present a new model for biofilms describing their growth and degradation by means of a system of equations including a PDE for biofilm concentration and an ODE for biofilm cooperation. By the second equation, we define cooperativity as a new concept (or variable), and we investigate the concept of QS in the framework of the newly defined variable, i.e., the cooperation. There is much room for further extensions of the present model, e.g., by incorporating further equations for signaling.

5.1.2. The Initial and Boundary Conditions

In order to complete the proposed model equations (5.1)–(5.3), we discuss the initial and boundary conditions. We consider the initial conditions

$$u_0(x,y) := u(t=0,x,y) = \exp(-(x^2+y^2)/10) \qquad \forall (x,y) \in D, \qquad (5.4\text{a})$$

$$v_0(x,y) := v(t=0,x,y) = \exp(-(x^2+y^2)/10) \qquad \forall (x,y) \in D. \qquad (5.4\text{b})$$

The equations (5.4) mean that the biofilm starts to grow at the origin $(0,0)$ with initial concentration 1.

Furthermore, if the solution u is radially symmetric, then we transform the model equation to polar coordinates (r,ϕ) so that the derivatives with respect to the angle ϕ vanish. This yields the independent variables t and r and reduces the number of dimensions. We will discuss this situation in more detail below.

To supplement the model, for every point (x,y) on the boundary ∂D, we use the zero Neumann boundary condition

$$\frac{\partial u}{\partial \mathbf{n}} = 0 \qquad \forall t \in \mathbb{R}^+ \quad \forall x,y \in \partial D, \qquad (5.5)$$

where $\mathbf{n}$ denotes the unit outwards normal vector on the boundary.

5.2. Existence and Uniqueness of Solutions of the Model

In this section, we study the existence and regularity of solutions to the initial-boundary value problem (5.1)–(5.5). To this end, we consider the initial-boundary value problem (IBVP)

$$\partial_t u = d_u \Delta u + f(u,v) \qquad \text{in } \mathbb{R}^+ \times D, \qquad (5.6\text{a})$$

$$\partial_t v = g(u,v) \qquad \text{in } \mathbb{R}^+ \times D, \qquad (5.6\text{b})$$

$$\frac{\partial u}{\partial \mathbf{n}} = 0 \qquad \text{on } \mathbb{R}^+ \times \partial D, \qquad (5.6\text{c})$$

$$u(0,x,y) = u_0(x,y) \qquad \text{in } D, \qquad (5.6\text{d})$$

$$v(0,x,y) = v_0(x,y) \qquad \text{in } D. \qquad (5.6\text{e})$$

First some assumptions are required. Then, we state a result on local-in-time existence and uniqueness of solutions to the above IBVP.

Assumptions 9. *We assume:*

1. *The domain $D \subset \mathbb{R}^d$ is bounded with sufficiently regular boundary ∂D.*

2. *The initial conditions u_0 and v_0 are bounded and symmetric functions on D.*

3. *The nonlinearities $f = f(u,v)$ and $g = g(u,v)$ are locally Lipschitz continuous functions on $u\colon \mathbb{R}^+ \times D \to \mathbb{R}^+$ and $v\colon \mathbb{R}^+ \times D \to (0,1] \subset \mathbb{R}^+$.*

Theorem 12 (Local-in-time existence and uniqueness [124])**.** *Under assumptions 9, there exists $T = T(\|u_0\|_\infty, \|v_0\|_\infty) > 0$ such that the initial-boundary value problem (5.6) has a unique local-in-time mild solution $(u, v) \in L^\infty([0, T], L^\infty(D))$.*

We recall the definition of a mild solution and then give a brief sketch for the proof of Theorem 12.

Definition 9. *A mild solution of the initial-boundary value problem (5.6) is a couple of measurable functions $u\colon [0, T] \times D \to \mathbb{R}^+$ and $v\colon [0, T] \times D \to (0, 1] \subset \mathbb{R}^+$ satisfying the system of integral equations*

$$u(t, x, y) = S(t)u_0(x, y) + \int_0^t S(t - s)f(u(s, x, y), v(s, x, y))\,ds, \tag{5.7a}$$

$$v(t, x, y) = v_0(x, y) + \int_0^t g(u(s, x, y), v(s, x, y))\,ds, \tag{5.7b}$$

where S is the semigroup of linear operators generated by the Laplacian with homogeneous Neumann boundary condition.

Proof. The system (5.1)–(5.5) satisfies all assumptions: It is straightforward to check that the initial conditions are bounded and symmetric, since they are Gaussian functions. Furthermore, the nonlinearities f and g in the model are locally Lipschitz continuous functions as can be seen using the mean-value theorem and due to the boundedness of the computational domain and the derivatives of the functions f and g on the domain. The function $\tan^{-1}$ has a bounded derivative and the derivative of a polynomial (as a C^1 function) is continuous and thus bounded.

The proof is based on an application of the Banach fixed-point theorem in order to construct a local-in-time unique solution for system (5.7), as the nonlinearities f and g are locally Lipschitz continuous functions. Details of the approach and the proof of Theorem 12 for even more general reaction-diffusion equations can be found in [130, Theorem 1, p. 111]. The construction of nonnegative solutions of a particular reaction-diffusion-ODE models also can be found in [131, Chapter 3]. $\qquad\square$

Remark 3. *To obtain more regular solutions, we can use more regular initial conditions, i.e., if $u_0 \in H_0^2(D) := \{u \in H^2(D) \mid \mathbf{n} \cdot \nabla u = 0 \ \text{on } \partial D\}$, then $(u, v) \in C(D; H_0^2(D) \times L^\infty(D))$. The proof is based on the theory of strongly continuous semigroups [132]. Furthermore, if $u_0 \in C^{2+\eta}(D)$ and $v_0 \in C^\eta(D)$ for some $\eta \in (0, 1)$, and the compatibility condition $\mathbf{n} \cdot \nabla u_0 = 0$ holds on ∂D, then the mild solutions of the IBVP (5.6) are smooth and satisfy $u \in C^{1+\eta/2, 2+\eta}([0, T] \times D)$ and $v \in C^{1,\eta}([0, T] \times D)$. For more details and studies of general reaction-diffusion-ODE systems in Hölder spaces, we refer the reader to [130, Theorem 1, p. 112] and [133].*

5.3. The Forward Problem

In this section, we solve the IBVP (5.1)–(5.5) numerically in order to discuss its behavior and to observe the evolution of concentration and cooperation in time.

5.3.1. The Model in Polar Coordinates

As mentioned, the solution u is radially symmetric for radially symmetric initial conditions. Therefore, we can transform the model equation to polar coordinates (r, ϕ) and the derivatives with respect to the angle ϕ vanish. This results in a problem in the independent variables t and r, and thus reduces the number of dimensions. The original model can be rewritten as

$$\frac{\partial u}{\partial t} = A\frac{\partial^2 u}{\partial r^2} + \frac{A}{r}\frac{\partial u}{\partial r} + \alpha u(1 - u/\beta) - \gamma(t)(1 - v)u, \tag{5.8a}$$

$$\frac{\partial v}{\partial t} = \max\big(0, \arctan\big(q(u) - \nu\big)\big)v - \kappa v^2, \tag{5.8b}$$

where

$$q(u)(t, r) := \big(u(t, \cdot) * rG(\cdot)\big)(r) \tag{5.9}$$

and

$$G(r) := \frac{1}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \tag{5.10}$$

with the initial conditions

$$u(t = 0, r) = \exp(-r^2/10), \tag{5.11a}$$

$$v(t = 0, r) = \exp(-r^2/10), \tag{5.11b}$$

which are radially symmetric, and with zero Neumann boundary condition for u. We have implemented the PDE model (5.8)–(5.11) of growth and degradation of biofilms including quorum sensing by means of the method of lines (MOL) and present the results in the following section. MOL is a technique for solving partial differential equations by discretizing in all but one dimension and then integrating the semi-discrete problem as a system of ODEs. Here, the dicretization in space is done by means of the finite-difference method, and the resulting ODE system in time is solved using a multistep solver based on numerical differentiation formulas (NDFs).

For computing $q(u)$ in equation (5.9), we have approximated the continuous one-dimensional convolution using the discrete one-dimensional one. Using a general notation, we estimate the convolution integral

$$g(x) := (f * h)(x) = \int_{-\infty}^{\infty} f(u)h(x - u)du \tag{5.12}$$

by

$$\hat{g}(x_k) = \sum_{j=-\infty}^{\infty} f(x_j)h(x_{k-j+1}), \tag{5.13}$$

where the sum is a discrete one-dimensional convolution.

5.3.2. Numerical Results of the Forward Biofilm Problem

Here, we present simulation results for the model (5.8)–(5.11). We assume a circular dish with diameter $35\,\text{mm}$ centered at the origin as the growth space and the computational

domain for the biofilms, which are initially located at the center of the dish. Biofilm growth is monitored for about six hours. The goal is to calculate the area covered by the biofilm after four hours.

In these simulations, the initial biofilm was located at the center of the dish. We find the largest circle in the computational domain where the concentration is above a certain threshold and calculate the relative coverage using the area of this circle. Fig. 5.1 illustrates the time evolution of the concentration and the cooperation of biofilms in about six hours without the presence of any antibiotics and using the diffusion parameter $A = 1.01$ and the growth rate $\alpha = 1.03$. Furthermore, the rest of the parameter values are $\beta = 0.7361$, $\nu = 0.5210$, and $\kappa = 0.1983$. In this setup, the total area covered by the biofilm after four hours is calculated as 0.9334, i.e., almost the entire domain.
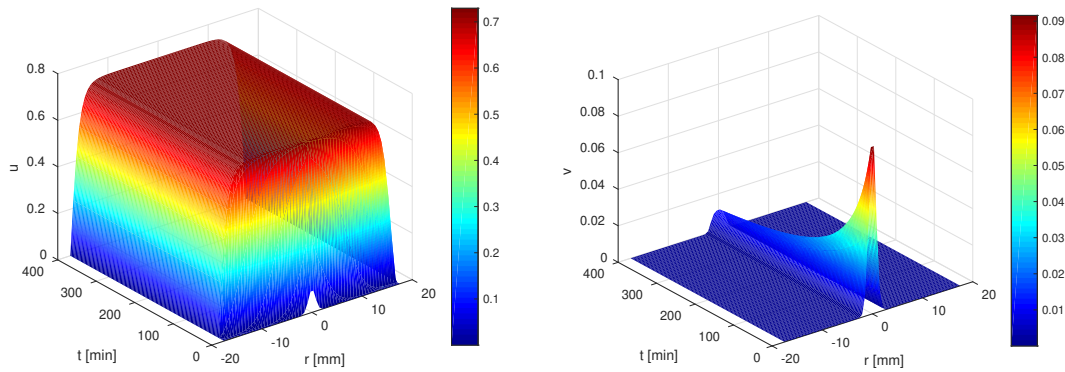

Figure 5.1.: Biofilm concentration (left) and cooperation (right) using $A = 1.01$ and $\alpha = 1.03$ and without the presence of any antibiotics. The resulting relative coverage is 0.9334.

Numerical results for the case when the antibiotic is present from the beginning are shown in Fig 5.2. In this setup, the relative coverage of the biofilm after four hours is calculated as 0.2761, since the antibiotic does not allow the biofilms to grow and spread fast.

Fig. 5.3 shows the behaviour of the biofilm in the case when an antibiotic is added after a while. The biofilm starts to grow quickly till the antibiotic is added and then it degrades due to addition of antibiotic. In this case, the resulting relative coverage is 0.6346. Here γ is a piecewise constant function defined as

$$\gamma := \begin{cases} 0.01 & \text{if } t \leq T, \\ 1 & \text{if } t > T, \end{cases} \tag{5.14}$$

where $T = 200$ is time of addition of the antibiotic.

Next, we discuss numerical results of biofilms growth and cooperation using different parameter values. In these numerical experiments, we assume that the antibiotic is present from the beginning and we study the effect of parameter variations on the solutions of the biofilm model, namely the concentration and the cooperation, while keeping the rest of the parameters constant. The default parameter values are the values that yield the best agreement with measurements, which will be explained in Subsection 5.5.1 in detail.
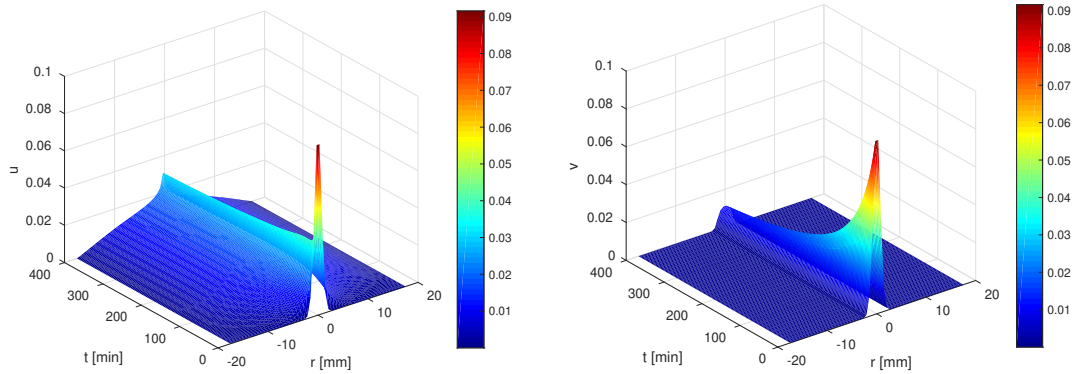
Figure 5.2.: Biofilm concentration (left) and cooperation (right) using $A = 1.01$ and $\alpha = 1.03$ and in the presence of antibiotic from the beginning. The resulting relative coverage is 0.2761.
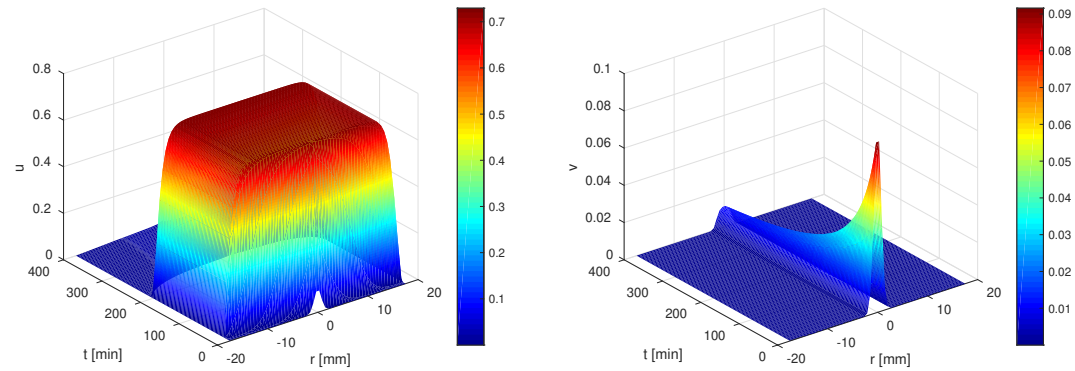

Figure 5.3.: Biofilm concentration (left) and cooperation (right) using $A = 1.01$ and $\alpha = 1.03$ and in the presence of antibiotic at $t = 200$. The resulting relative coverage is 0.6346.

Fig. 5.4 displays the resulting concentrations of biofilms using diffusion constants $A = 0.8$ and 0.02. These results show that when the concentration spreads faster (i.e., when the diffusion constant is larger), the biofilm is thinner and the covered area is larger. For a smaller diffusion constant, the biofilm is thicker and more concentrated with smaller coverage, as expected. This result agrees with our qualitative assessment in Section 5.1.1.

Fig. 5.5 shows concentrations of biofilms for three different growth rates $\alpha = 1.01$, 1.03 and 1.05. In fact, if two populations of bacteria start to grow with the same diffusion constant but different growth rates, the one with higher growth rate creates a much more concentrated and thicker biofilm. Furthermore, the numerical results show larger coverage area for a bigger growth rate.

We observe the effect of different values of β on the concentration and the cooperation of the biofilms in Fig 5.6. According to these results, using a larger value for the parameter
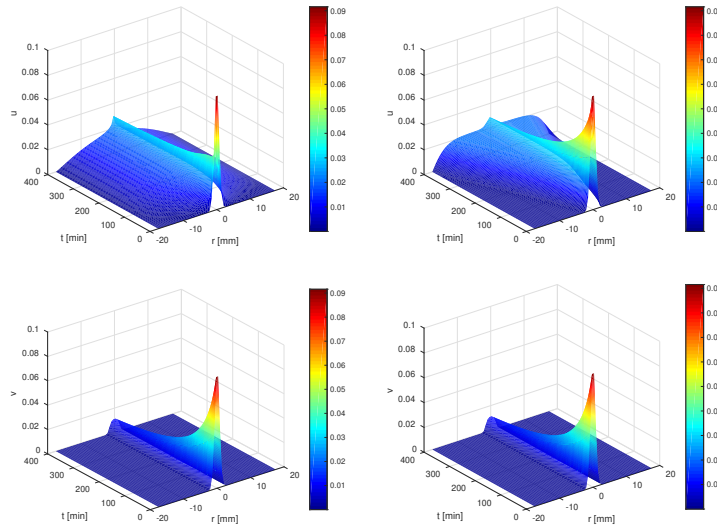
Figure 5.4.: Biofilm concentration (top) and cooperation (bottom) for $A = 0.8$ (left) and $A = 0.02$ (right).
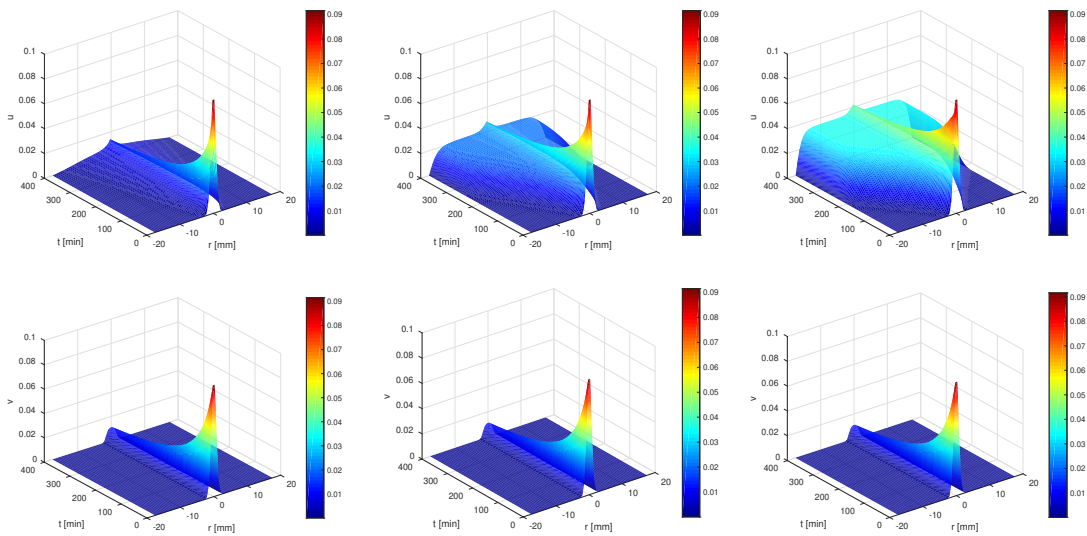

Figure 5.5.: Biofilm concentration (top) and cooperation (bottom) for $\alpha = 1.01$ (left), $\alpha = 1.03$ (middle), and $\alpha = 1.05$ (right).

results in a thicker biofilm and bigger coverage, however slight changes in the parameter value hardly affect the relative coverage.

Moreover, in Fig. 5.7, the effect of different values of the parameter κ is studied and the corresponding numerical results, i.e., the concentration of biofilm and the cooperation are illustrated. According to these results, the cooperation among the bacteria increases with larger κ, and the coverage is almost constant.
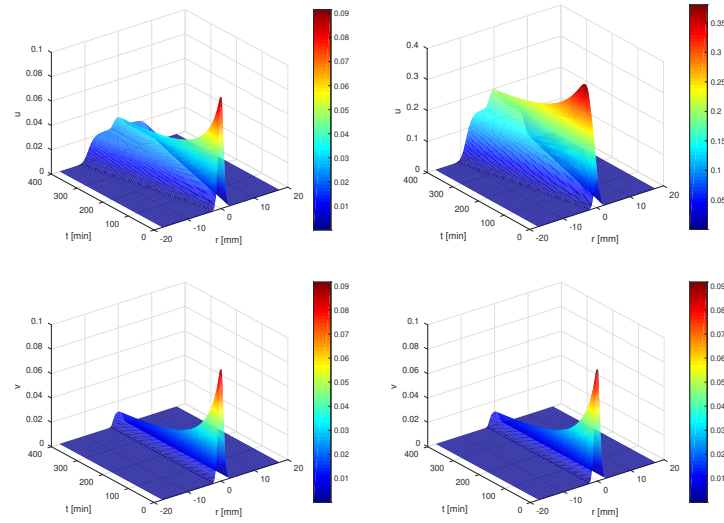
Figure 5.6.: Biofilm concentration (top) and cooperation (bottom) for $\beta = 0.74$ (left) and $\beta = 5$ (right).
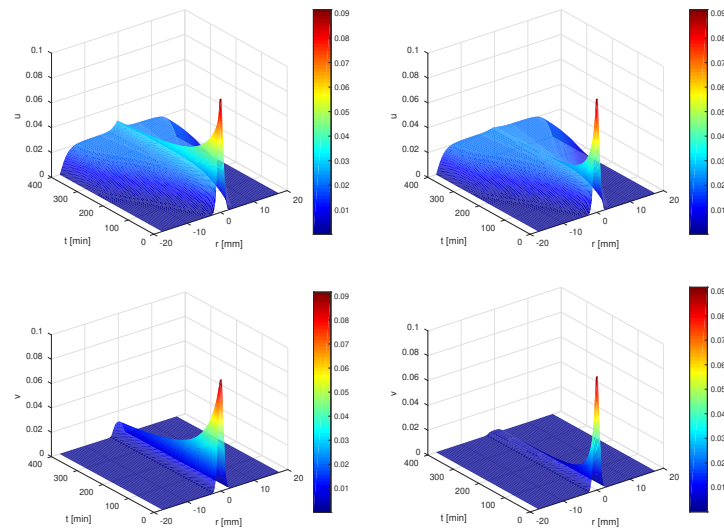

Figure 5.7.: Biofilm concentration (top) and cooperation (bottom) for $\kappa = 1$ (left) and $\kappa = 0.2$ (right).

According to these numerical results, solving the problem without having proper parameter values at hand may result in arbitrary, highly parameter dependent solutions. In Section 5.5, we will fix this issue by means of solving the inverse problem and determining the parameter values. But before we can do that, the experimental results are described in the next section.

5.4. Experimental Results

In this section, we present the results of our experiments and observations for biofilm growth. In these experiments, we aimed to monitor the development of biofilms during 24 hours. To this end, $1\,\text{mL}$ of each reference (diluted in Brain Heart Infusion Medium $1.5 \cdot 10^7$) was cultivated in single 24-well Ibidi μ-Dishes (Ibidi Treat 1, 5 polymer coverslip, tissue culture treated; Ibidi GmbH, Planegg/Martinsried, Germany). Biofilms were grown at $37°C$ for 24 hours on an orbital shaker. Every hour one well was taken off for further investigation. Furthermore, biofilms were washed two times in PBS and fixed with 4% glutaraldehyde. We observed bacteria in biofilms and the matrix (EPS) structure at different times of biofilm formation. To observe the dense DNA of the dead bacteria (Molecular Probes$^{®}$; Thermo Fisher Scientific), propidium iodide was used. Polysaccharides, representing the most characteristic fraction of the extrapolymeric substances, were stained using concanavalin-A (ConA) (Sigma-Aldrich Corp, St. Louis, MO, USA). Fig. 5.8 displays the formation of *S. epidermidis* and *S. aureus* biofilms after six hours of incubation.
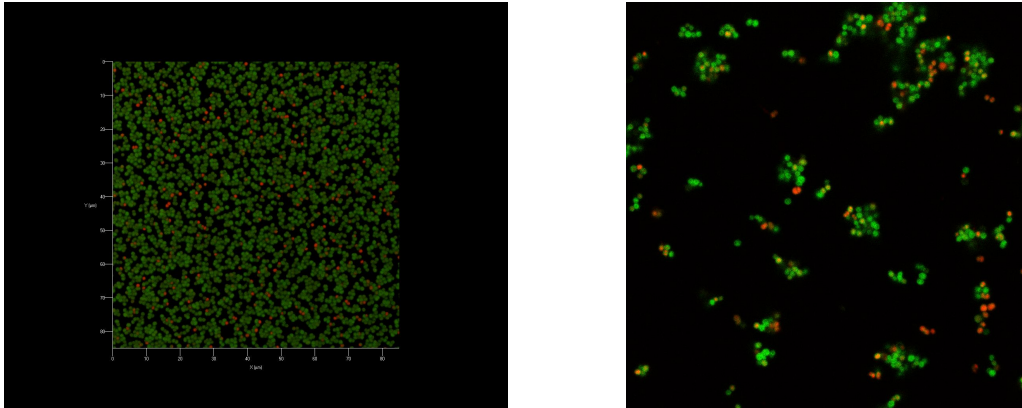

Figure 5.8.: The *S. epidermidis* (left) and *S. aureus* (right) biofilms after six hours of incubation. The *S. epidermidis* film shows scattered cells over the surface, while *S. aureus* forms clusters. The images display DNA in red and polysaccharide in green.

During the 24-hour long observation of the reference strains, we found differences in the pattern of biofilm formation; *S. aureus* ATCC25923 aggregated and formed various grape-like aggregations of bacterial cells coated by single polysaccharides before spreading on the surface and forming a biofilm layer. In contrast, *S. epidermidis* biofilms started with scattered cells spreading over the surface until reaching confluence and their maximum thickness at 24 hours without forming grapes or clusters [134].

Fig. 5.9 and Fig.5.10 illustrate the area coverage by the two reference strains of biofilms, namely *S. aureus* ATCC 25923 and *S. epidermidis* DSM 3269 after 6 and 24 hours, respectively. In these figures, we show how much area is covered by polysaccharide, mix, and DNA.

In the experiments, the relative coverage at time t is defined as the relative area of the dish covered by the biofilm as measured by counting particles using the ImageJ software.

ImageJ is an open source Java image processing program for automatic particle counting and analyzing.
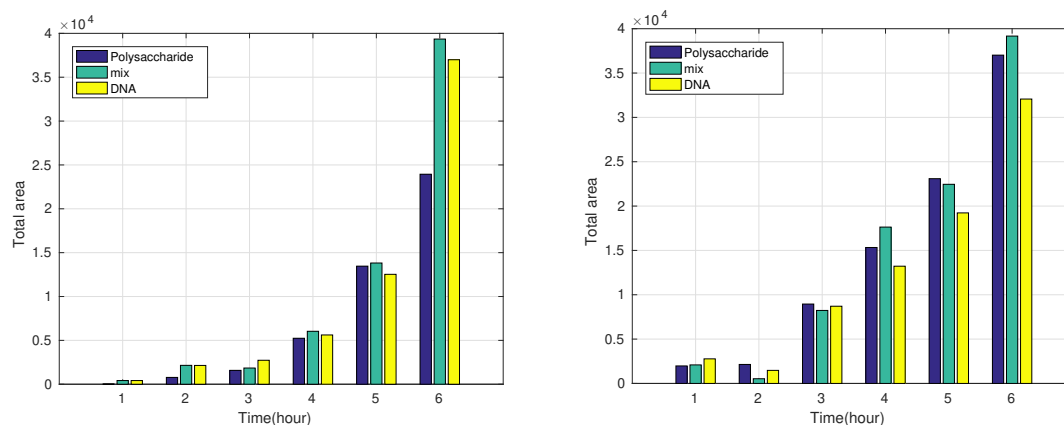

Figure 5.9.: Total area covered by the biofilms *S. epidermidis* DSM 3269 (left) and *S. aureus* ATCC 25923 (right) (polysaccharide, mix, DNA) after 6 hours based on experimental data.
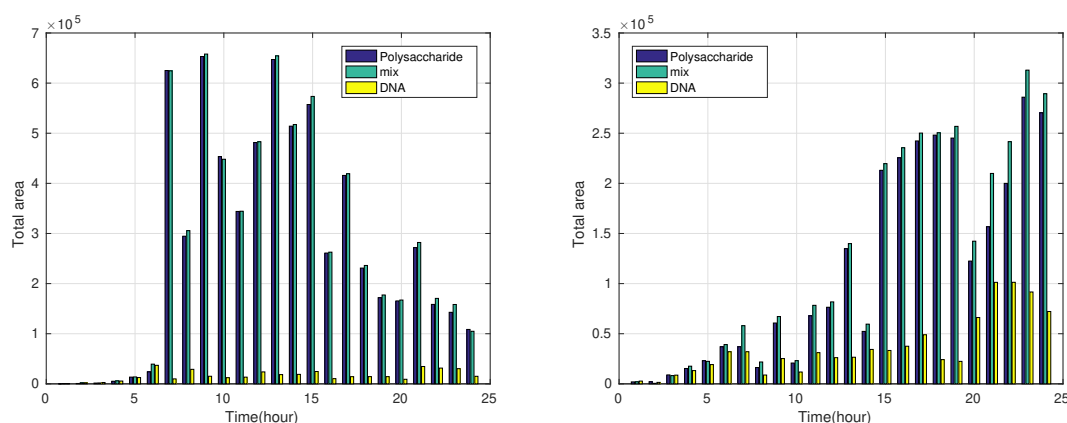

Figure 5.10.: Total area covered by the biofilms *S. epidermidis* DSM 3269 (left) and *S. aureus* ATCC 25923 (right) (polysaccharide, mix, DNA) in after 24 hours based on experimental data.

In this work, we have not included the amount of available nutrients as a variable in our model, but we assume that enough nutrients are present in order to enable biofilm growth. As already mentioned, coupling the growth of the biofilm to a variable amount of nutrients is a possible model extension. Further model extension are to model the spreading of the bacteria depending on the local food concentration and to model the food concentration by an additional equation. Furthermore, if we include the amount of nutrients in the model, new findings show that the growth of biofilms can even be affected by the communication between nearby bacterial communities [135]. The bacterial communities seem to use a

time-sharing strategy in the sense that two distinct biofilms can synchronize their growth through electrical signals. When bacteria face limited nutrients, they alternate feeding periods and each community takes turns consuming nutrients in order to reduce competition and to maximize efficiency in consumption. This behavior extends communication among functional units such as cells to biofilms. It modifies cell-to-cell signaling mediated by ion channels [136–139] to signaling between two distinct biofilms.

5.5. The Inverse Problem

The coefficients and parameters used in the mathematical model are effective factors in the growth and cooperation of biofilms, but they have unknown and uncertain values in the sense that they can only be determined by measurements that always come with measurement errors. These parameters cannot be controlled during the experiments. However, knowing the parameter values that determine the growth and cooperation of biofilms helps to react against biofilms after diagnosis in the sense that these parameter values give concrete information about the bacterial behavior. Understanding the bacterial behavior is also helpful to know which sort of antibiotic and how much should be used to kill the bacteria. Therefore any information about these effective factors improves the speed and quality of the treatment.

There are various methods for estimating the parameter values and solving the inverse problem. One can divide the statistical inference approaches into two categories: frequentist and Bayesian inference. The differences between these two approaches stem from the way the concept of probability is interpreted.

In the frequentist approach, the unknown parameter is assumed to be fixed and deterministic, and probabilities are defined as long-term frequencies of occurrences of an event. The event has to occur many times. Therefore, in this approach we collect data from a sample of the population and estimate its mean as the value which agrees best with the data.

In contrast, in the Bayesian approach, the unknown parameters are assumed to be random variables. In this technique, probabilities are rooted in degrees of belief and logical support and can be used to represent uncertainties in any event, even in non-repeatable events. In Bayesian inference, we define probability distributions over possible parameter values and use data to update the distribution, which means that beliefs are updated in response to new evidence. The updating is done by applying Bayes' theorem. In fact, the new information (e.g., experimental data) makes the probability distribution more focused around the true value of the unknown parameter. Thus confidence intervals can easily be calculated.

Here, our goal is to estimate two very effective quantities in particular, namely the diffusion constant A and the growth rate α in the biofilm model (5.1). To identify these parameters, we use two approaches from the above mentioned statistical categories: a genetic algorithm (GA) as a deterministic one and an adaptive Markov-chain Monte-Carlo (MCMC) algorithm as a stochastic (Bayesian) technique.

We apply Bayesian inversion as a stochastic approach and a genetic algorithm as a deterministic approach to estimate the parameter values in the biofilm inverse problem, also to illustrate the differences between these two approaches. Moreover, the starting point of the Markov chain in an MCMC method in the Bayesian approach is important for fast

convergence of the chain to the posteriori distribution of the parameter of interest. To this end, we use the optimal parameter values obtained from the deterministic approach as the starting points for the MCMC method. In the case of a genetic algorithm, the technique finds optimal values for the parameters by minimizing the residual between the experimental data and the numerical results. In the case of the Bayesian technique, 95% confidence interval of the parameter values are also presented.

Here, our goal is to estimate effective quantities in the biofilm model (5.1). To this end, we use two approaches from the above mentioned statistical categories: a genetic algorithm (GA) as a deterministic one and an adaptive Markov-chain Monte-Carlo (MCMC) algorithm as a stochastic (Bayesian) technique.

5.5.1. Deterministic Approach: Genetic Algorithm

In this section, we use a genetic algorithm (GA) as a global-optimization method in order to estimate the parameters of interest in the biofilm inverse problem. The genetic-algorithm based approaches are deterministic methods in the sense that the parameters of interest are assumed to be deterministic variables, whose optimal values are found by means of minimizing a cost (fitness) function. This method must be equipped with a cost function, since it enables the method to rank the individuals. Here, the cost function is defined as the residual

$$\text{cost} = \|\text{cov} - \text{cov}_{\text{exp}}\|_{\ell^2}, \tag{5.15}$$

between the experimental data and numerical results of the forward problem, i.e., the coverage data obtained by using the MOL solver for the forward model. Here cov and cov_{exp} denote the coverages obtained numerically and in the experiments, respectively.

Steps 1 to 6 show the GA used here, which has been based on [140]:

1. **Initialization.** In this step, a random initial population is generated.

2. **Evaluation.** Each individual is evaluated using a fitness function.

3. **Selection.** Selection is a process in which the individuals which are suitable for the next generation are chosen. In this step, our model adopts tournament selection. This step is repeated until the number of individuals selected is equal to the desired population size. In order to ensure the propagation of elite individuals, elitism is used. This mechanism selects the individuals with the best fitness values and directly places them in the next generation, while the remaining individuals must go through the selection process.

4. **Crossover and Mutation.** Using crossover and mutation, children are produced from the parents. In mutation, children are produced by making random changes to a single parent, while in crossover this is done by combining the vector entries of a pair of parents.

5. **Evaluation.** Each individual is evaluated using the fitness function.

6. **Check termination criteria.** Steps 2 to 5 are repeated until the termination criteria are satisfied. Here the algorithm is stopped when the maximum number of generations

has been exhausted or when the solution with the best fitness has not been changed from the previous generation.

Table 5.1 lists the GA settings used and can be used for reproducing the results. The selection function defines how individuals are selected to become parents. The number of elite children shows the number of individuals with the best fitness values in the current generation that are guaranteed to survive into the next generation. These individuals are called elite children. Moreover, crossover fraction is the fraction of individuals in the next generation, other than elite children, that are created by crossover. The remaining are generated by mutation.

Setting	value/method
number of generations	50
population size	60
initialization method	uniform sampling
selection function	stochastic universal sampling (SUS) [141]
number of elite children	2
crossover fraction	0.8
mutation fraction	0.2
crossover function	uniform crossover
mutation function	Gaussian mutation

Table 5.1.: Settings used in the genetic algorithm.

In this algorithm, we have calculated the mean and the minimum of the obtained residual in the populations in 50 generations and we have shown the results in Fig. 5.11. As the number of generations increases, the individuals in the population crowd around the minimum point 0.
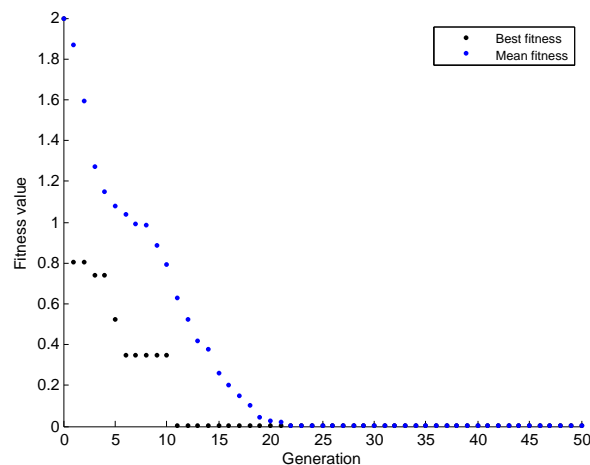

Figure 5.11.: The best and mean fitness values in 50 generations in the GA.

The GA yields the optimal parameters A, α, β, ν, and κ in the biofilm model equation (5.1) as

$$(A, \alpha, \beta, \nu, \kappa) = \text{argmin cost},$$

where the cost function is defined in (5.15). The values found are shown in Table 5.2. We will use these values as the initial values for the Markov chains of the corresponding parameters in the stochastic approach in the next section. A near optimal choice of starting states of a Markov chain helps the chain to quickly converge around the true value of the parameters of interest.

Parameter	A	α	β	ν	κ
Best value found by GA	0.048 79	1.0102	0.7361	0.521	0.1983

Table 5.2.: Best parameters values in the biofilm inverse problem found by the GA.

5.5.2. Bayesian PDE Inversion

The second approach to solve the biofilm inverse problem is Bayesian PDE inversion [104–106]. In this method, we consider the unknown parameters as random variables and calculate the posterior probability density that reflects the distribution of the parameter values based on the observations. Therefore, in this method, not a single parameter value but its probability distribution is found. This is an advantage, since probability distribution conveys information how well the parameters can be determined. Bayesian inversion method and the Markov-chain Monte-Carlo (MCMC) method as well as delayed-rejection adaptive-Metropolis (DRAM) algorithm as an adaptive MCMC algorithm have been explained in Subsections 4.2.1, 4.2.4 and 4.2.5 in detail.

Numerical Results for the Biofilm Inverse Problem Using Bayesian Inference

In this section, to provide a better insight into the biofilm model, we quantify uncertain values of the model by means of the Bayesian inversion method and by comparing and analyzing measurements and computed model results. Bayesian inference has been already applied in order to parameter identification in different models related to biofilms [142, 143]. Here the goal is to extract the as much information as possible using the measurements. However the amount of recovered information depends on many factors including the size of available experimental data. Due to a couple of restrictions such as insufficient measurements (there is no experimental data of cooperation), recovering all model parameters accurately at the same time is difficult. However, we can find confidence intervals for all the parameters. Among all parameters, the growth rate α can be extracted best. We have used the DRAM algorithm as an adaptive MCMC method in the context of the Bayesian inversion method for this analysis. To start the Markov chains, we use the optimal parameter values that we have found by means of the genetic algorithm, which are summarized in Table 5.2 in Section 5.5.1 to speed up the convergence of the generated chains. We also use the measured area covered by the *S. epidermidis* biofilm (mix) (see Section 5.4). In Fig. 5.12, marginal histograms of

the posterior distribution of the parameters in the prior defined intervals together with their mean values are shown. We can infer a confidence interval of $\alpha \in [0.3049, 0.803]$ at the 95% level and a mean of 0.5540 for the growth rate α. For the rest of the parameters, we find mean value for the posterior distributions of the parameters, but the corresponding Markov chains explore the whole, large interval and thus posterior distributions are wide-spread. The marginal histograms show that among all the growth rate α can be inferred precisely. This is due to the ill-posedness of the biofilm inverse problem, to the strong correlation among the model parameters, and in particular to the insufficient number of measurement values and their noise.

5.6. Summary

In this work, we presented a reaction-diffusion-ODE system as the model equations for biofilms including the cooperation among the bacteria. This is the first time–to the best of our knowledge–that cooperation of bacteria in biofilms is introduced and quantified. The model includes also the concept of quorum sensing, which is presented in the framework of the newly defined quantity, i.e., the cooperation in the sense that when local average mass of biofilms is above a quorum threshold, the bacteria start to cooperate, which leads to higher resistance against antibiotics. We also presented local-in-time existence and uniqueness results as well as regularity of solutions to the presented model type.

To provide better insight into the model, we quantified the unknown parameter values of the model by means of a Bayesian technique. In this method, the experiments observations are compared to the computational growth and coverage obtained from the numerical solution of the model in order to infer the uncertain values of the model, which cannot be controlled or determined through the experiments and computations. In the Bayesian inference, the output is not a single value but a distribution of the parameter density and a corresponding confidence interval. In this work, the model parameters including those describe the biofilm concentration and cooperation are estimated using an adaptive MCMC method, namely the DRAM algorithm. This adaptive algorithm leads to better mixing of the generated chain and avoids stagnations since it chooses optimal proposal scales automatically. Furthermore, the Markov chain generated by this algorithm converges faster and the burn-in period is shorter in comparison to the naive Metropolis-Hastings algorithm.

Due to the ill-posedness of the biofilm inverse problem and strong correlation among the model parameters and also due to insufficient information from experiments (a small number of data points for the growth and the coverage area and no cooperation data), the growth rate α is the parameter which can be recovered best using Bayesian analysis. According to the numerical results, the growth rate α is recovered successfully and a confidence interval of $[0.3049, 0.803]$ at the 95% level and a mean of 0.5540 for this parameter are found.

Moreover, we estimate the parameters in the biofilm model using a genetic algorithm and use these values as the starting points of the Markov chains in the DRAM algorithm. Carefully choosing the starting state ensures that the chain converges to the true value more quickly. The mathematical model equations in conjunction with Bayesian PDE inversion make it possible to assign biologically and medically relevant parameter values to various species of bacteria. This procedure therefore allows us to quantify and to compare the
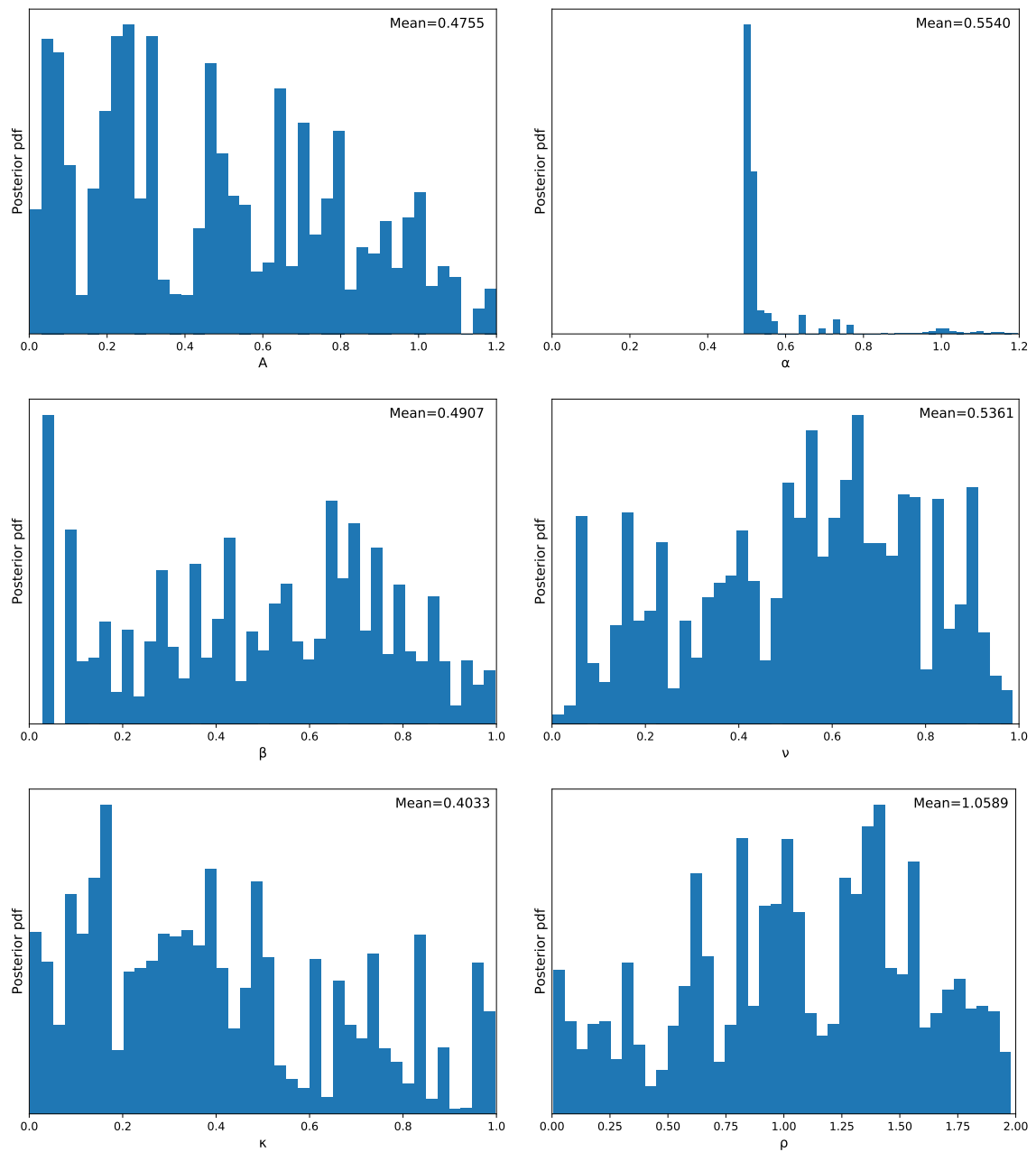
Figure 5.12.: Marginal histograms of posterior distributions of the parameters in the biofilm model using the DRAM algorithm with 10^4 samples. The mean values are also indicated.

behavior and the strategies of different species of bacteria.

6. Conclusions and Outlook

The main objective of this dissertation was uncertainty quantification in various devices including nanoscale devices such as nanopores and nanowire field-effect sensors and electronic devices such as electrical-impedance tomography sensors. To this end, we focused on different aspects of nonlinear partial differential equations (PDEs), which model charge transport in the devices, which have many applications in industry, medicine and life sciences. These aspects include analysis, numerical solution to the PDE models as well as PDE inverse problems. The main PDE models include Poisson-Boltzmann equation and the drift-diffusion-Poisson (DDP) system in deterministic and stochastic regimes.

As mentioned, we divide the work into two main parts: first, uncertainty quantification in nanoscale devices by modeling and analyzing nanoelectronic devices and second, uncertainty quantification in EIT devices and biofilms model by inverse modeling and Bayesian inference.

Following the first part of the research, we first analyzed the so-called drift-diffusion-Poisson (DDP) system in the alternating-current small-signal regime and extended the transport model to the frequency domain. We also proved existence and local uniqueness of solution of the presented model equations in the AC regime. The applications include wide range of nanoscale devices such as nanowire filed-effect bio- and gas sensors as well as nanopore sensors, which operate in the AC regime.

Then we studied uncertainty quantification in aforementioned nanoscale devices by analyzing and developing numerical methods for the stochastic version of the drift-diffusion-Poisson system. More precisely, we presented existence and local uniqueness theorems for the system and developed an optimal multilevel Monte-Carlo finite-element method (MLMC-FEM) by solving an optimization problem. In fact, we minimized the computational cost of solving the stochastic DDP system by the MLMC-FEM such that the error bound is less than or equal to a given tolerance level. This leads to find optimal discretization parameters and optimal number of samples in the stochastic numerical method. The presented optimal methods are successfully used to solve the stochastic DDP system to study the effect of random dopants in nanoscale nanowire field-effect sensors. In the comparison of the MC and MLMC methods, the MLMC method was found to decrease the total computational cost by four orders of magnitude for small error tolerances. To further improve the computational efficiency, an optimal multilevel randomized quasi Monte-Carlo finite-element method (MLRQMC-FEM) was developed and tested. Computational cost of optimal MLRQMC was one order of magnitude smaller than that of optimal MLMC and five orders of magnitude smaller than that of the vanilla Monte-Carlo method.

Following the second part of the research, we studied two PDE inverse models, namely EIT and biofilms inverse problems. First, we presented a nonlinear comprehensive physical model for the EIT sensors, which have applications for instance in medical imaging and monitoring for diagnosis purposes. Then we formulated the Bayesian inversion for the presented nonlinear model in a measure-theoretic framework and in an infinite-dimensional

setting and presented the first application of the Bayesian analysis for the simultaneous recovery of multiple parameters in the EIT inverse problem. The well-definedness of the posterior measure and well-posedness of the Bayesian inversion for the presented model were also proved. Then we illustrated numerical results of Markov-chain Monte-Carlo finite-element method (MCMC-FEM), including calculating electrical currents on the electrodes by means of FEM, and recovery of some electrical and geometrical properties of internal body using an adaptive MCMC method.

Furthermore, we presented a system of PDEs to model biofilms growth and degradation considering quorum sensing and cooperation of bacteria against antibiotics. We also solved the biofilm inverse problem by means of Bayesian inversion techniques and an adaptive MCMC method comparing the results obtained from the computational model and the biofilms experiments in order to extract parameters of interest in the mathematical model.

Bibliography

[1] Gengfeng Zheng, Fernando Patolsky, Yi Cui, Wayne U Wang, and Charles M Lieber. Multiplexed electrical detection of cancer markers with nanowire sensor arrays. *Nature Biotechnology*, 23(10):1294–1301, 2005.

[2] Eric Stern, James F Klemic, David A Routenberg, Pauline N Wyrembak, Daniel B Turner-Evans, Andrew D Hamilton, David A LaVan, Tarek M Fahmy, and Mark A Reed. Label-free immunodetection with CMOS-compatible semiconducting nanowires. *Nature*, 445(7127):519–522, 2007.

[3] Eric Stern, Aleksandar Vacic, Nitin K Rajan, Jason M Criscione, Jason Park, Bojan R Ilic, David J Mooney, Mark A Reed, and Tarek M Fahmy. Label-free biomarker detection from whole blood. *Nature Nanotechnology*, 5(2):138–142, 2010.

[4] Anton Köck, Alexandra Tischner, Thomas Maier, Michael Kast, Christian Edtmaier, Christian Gspan, and Gerald Kothleitner. Atmospheric pressure fabrication of SnO_2-nanowires for highly sensitive CO and CH_4 detection. *Sensors and Actuators B: Chemical*, 138(1):160–167, 2009.

[5] Hui Huang, CY Ong, Jun Guo, Tim White, Man Siu Tse, and Ooi Kiang Tan. Pt surface modification of SnO_2 nanorod arrays for CO and H_2 sensors. *Nanoscale*, 2(7):1203–1207, 2010.

[6] Claus Griessler, Elise Brunet, Thomas Maier, Stephan Steinhauer, Anton Köck, Teva Jordi, Franz Schrank, and Martin Schrems. Tin oxide nanosensors for highly sensitive toxic gas detection and their 3D system integration. *Microelectronic Engineering*, 88(8):1779–1781, 2011.

[7] Gerhard Tulzer, Stefan Baumgartner, Elise Brunet, Giorgio C Mutinati, Stephan Steinhauer, Anton Köck, Paolo E Barbano, and Clemens Heitzinger. Kinetic parameter estimation and fluctuation analysis of CO at SnO_2 single nanowires. *Nanotechnology*, 24(31):315501, 2013.

[8] Gerhard Tulzer and Clemens Heitzinger. Fluctuations due to association and dissociation processes at nanowire-biosensor surfaces and their optimal design. *Nanotechnology*, 26(2):025502, 2015.

[9] Naoufel Ben Abdallah, Florian Méhats, and Claudia Negulescu. Adiabatic quantum-fluid transport models. *Communications in Mathematical Sciences*, 4(3):621–650, 2006.

[10] Clemens Heitzinger and Christian Ringhofer. A transport equation for confined structures derived from the Boltzmann equation. *Commun. Math. Sci*, 9(3):829–857, 2011.

[11] Nicola Zamponi and Ansgar Juengel. Two spinorial drift-diffusion models for quantum electron transport in graphene. *Commun. Math. Sci*, 11:927–950, 2013.

[12] Paola Pietra and Nicolas Vauchelet. Numerical simulation of an energy-transport model for partially quantized particles. *Commun. Math. Sci*, 12(1):99–123, 2014.

[13] Clemens Heitzinger, Norbert J Mauser, and Christian Ringhofer. Multiscale modeling of planar and nanowire field-effect biosensors. *SIAM Journal on Applied Mathematics*, 70(5):1634–1654, 2010.

[14] Clemens Heitzinger and Christian Ringhofer. Multiscale modeling of fluctuations in stochastic elliptic PDE models of nanosensors. *Commun. Math. Sci.*, 12(3):401–421, 2014.

[15] Stefan Baumgartner and Clemens Heitzinger. Existence and local uniqueness for 3D self-consistent multiscale models for field-effect sensors. *Commun. Math. Sci*, 10(2):693–716, 2012.

[16] Stefan Baumgartner and Clemens Heitzinger. A one-level FETI method for the drift–diffusion-poisson system with discontinuities at an interface. *Journal of Computational Physics*, 243:74–86, 2013.

[17] Stefan Baumgartner, Martin Vasicek, Alena Bulyha, and Clemens Heitzinger. Optimization of nanowire DNA sensor sensitivity using self-consistent simulation. *Nanotechnology*, 22(42):425503, 2011.

[18] Stefan Baumgartner, Clemens Heitzinger, Aleksandar Vacic, and Mark A Reed. Predictive simulations and optimization of nanowire field-effect PSA sensors including screening. *Nanotechnology*, 24(22):225503, 2013.

[19] Federico Pittino, Luca Selmi, and Frans Widdershoven. Numerical and analytical models to investigate the AC high-frequency response of nanoelectrode/SAM/electrolyte capacitive sensing elements. *Solid-State Electronics*, 88:82–88, 2013.

[20] Mattia Rossi, Marco Bennati, Marta Lombardini, Peter Ashburn, Hywel Morgan, Enrico Sangiorgi, and Marco Tartagni. Hybrid system for complex AC sensing of nanowires. In *Proc. 2014 15th International Conference on Ultimate Integration on Silicon (ULIS)*, pages 41–44. IEEE, 2014.

[21] Cecilia Laborde, Federico Pittino, Harrie Verhoeven, Serge Lemay, Luca Selmi, Maarten Jongsma, and Frans Widdershoven. Real-time imaging of microparticles and living cells with CMOS nanocapacitor arrays. *Nature Nanotechnology*, 10:791–796, 2015.

[22] Federico Pittino and Luca Selmi. Improved sensitivity of nanowire bio-FETs operated at high-frequency: A simulation study. In *Nanotechnology (IEEE-NANO), 2015 IEEE 15th International Conference on*, pages 792–795. IEEE, 2015.

[23] Federico Pittino and Luca Selmi. Use and comparative assessment of the CVFEM method for Poisson–Boltzmann and Poisson–Nernst–Planck three dimensional simulations of impedimetric nano-biosensors operated in the DC and AC small signal regimes. *Computer Methods in Applied Mechanics and Engineering*, 278:902–923, 2014.

[24] Clemens Heitzinger, Yang Liu, Norbert J Mauser, Christian Ringhofer, and Robert W Dutton. Calculation of fluctuations in boundary layers of nanowire field-effect biosensors. *Journal of Computational and Theoretical Nanoscience*, 7(12):2574–2580, 2010.

[25] Michael B Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.

[26] Stefan Heinrich. Multilevel Monte Carlo methods. In *Large-Scale Scientific Computing*, pages 58–67. Springer, 2001.

[27] Mike Giles. Improved multilevel Monte Carlo convergence using the Milstein scheme. In *Monte Carlo and quasi-Monte Carlo methods 2006*, pages 343–358. Springer, 2008.

[28] Michael B Giles and Ben J Waterhouse. Multilevel quasi-Monte Carlo path simulation. *Advanced Financial Modelling, Radon Series on Computational and Applied Mathematics*, pages 165–181, 2009.

[29] Andrea Barth, Christoph Schwab, and Nathaniel Zollinger. Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients. *Numerische Mathematik*, 119(1):123–161, 2011.

[30] Andrew Cliffe, Mike Giles, Robert Scheichl, and Aretha Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Computing and Visualization in Science*, 14(1):3–15, 2011.

[31] Frances Y Kuo, Christoph Schwab, and Ian H Sloan. Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients. *SIAM Journal on Numerical Analysis*, 50(6):3351–3374, 2012.

[32] Julia Charrier, Robert Scheichl, and Aretha L Teckentrup. Finite element error analysis of elliptic pdes with random coefficients and its application to multilevel monte carlo methods. *SIAM Journal on Numerical Analysis*, 51(1):322–352, 2013.

[33] Aretha Teckentrup, Robert Scheichl, Mike Giles, and Elisabeth Ullmann. Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients. *Numerische Mathematik*, 125(3):569–600, 2013.

[34] Abdul-Lateef Haji-Ali, Fabio Nobile, Erik von Schwerin, and Raúl Tempone. Optimization of mesh hierarchies in multilevel Monte Carlo samplers. *Stochastic Partial Differential Equations: Analysis and Computations*, pages 1–37, 2015.

[35] Frances Y Kuo, Christoph Schwab, and Ian H Sloan. Multi-level quasi-Monte Carlo finite element methods for a class of elliptic PDEs with random coefficients. *Foundations of Computational Mathematics*, 15(2):411–449, 2015.

[36] Leila Taghizadeh, Amirreza Khodadadian, and Clemens Heitzinger. The optimal multilevel Monte-Carlo approximation of the stochastic drift–diffusion-Poisson system. *Computer Methods in Applied Mechanics and Engineering*, 318:739–761, 2017.

[37] R Cranley and TNL Patterson. Randomization of number theoretic methods for multiple integration. *SIAM Journal on Numerical Analysis*, 13(6):904–914, 1976.

[38] Stephen Joe. Randomization of lattice rules for numerical multiple integration. *Journal of Computational and Applied Mathematics*, 31(2):299–304, 1990.

[39] Bruno Tuffin et al. On the use of low discrepancy sequences in Monte Carlo methods. *Monte Carlo Methods and Applications*, 2:295–320, 1996.

[40] Thomas Gerstner and Marco Noll. Randomized multilevel quasi-Monte Carlo path simulation. In *Recent Developments in Computational Finance: Foundations, Algorithms and Applications*, pages 349–369. World Scientific, 2013.

[41] David S Holder. *Electrical Impedance Tomography: Methods, History and Applications*. CRC Press, 2004.

[42] Steven J Stanley and Gary T Bolton. A review of recent electrical resistance tomography (ERT) applications for wet particulate processing. *Particle & Particle Systems Characterization*, 25(3):207–215, 2008.

[43] Manuchehr Soleimani and William RB Lionheart. Nonlinear image reconstruction for electrical capacitance tomography using experimental data. *Measurement Science and Technology*, 16(10):1987, 2005.

[44] Margaret Cheney, David Isaacson, and Jonathan C Newell. Electrical impedance tomography. *SIAM Review*, 41(1):85–101, 1999.

[45] Jari Kaipio and Erkki Somersalo. *Statistical and Computational Inverse Problems*, volume 160. Springer Science & Business Media, 2006.

[46] Abelardo L Ramirez, William D Daily, and Andrew M Binley. Electrical resistance tomography. Technical report, Pacific Northwest National Lab., Richland, WA (US), 2000.

[47] A Nissinen, LM Heikkinen, and JP Kaipio. The Bayesian approximation error approach for electrical impedance tomography—experimental results. *Measurement Science and Technology*, 19(1), 2008.

[48] Jari P Kaipio, Ville Kolehmainen, Erkki Somersalo, and Marko Vauhkonen. Statistical inversion and Monte Carlo sampling methods in electrical impedance tomography. *Inverse Problems*, 16(5):1487, 2000.

[49] Lassi Roininen, Janne MJ Huttunen, and Sari Lasanen. Whittle-Matérn priors for Bayesian statistical inversion with applications in electrical impedance tomography. *Inverse Problems & Imaging*, 8(2):561–586, 2014.

[50] Matthew M Dunlop and Andrew M Stuart. The Bayesian formulation of EIT: Analysis and algorithms. *Inverse Problems & Imaging*, 10(4):1007–1036, 2016.

[51] Matti Lassas, Eero Saksman, and Samuli Siltanen. Discretization-invariant Bayesian inversion and Besov space priors. *Inverse Problems and Imaging*, 3(1):87–122, 2009.

[52] Andrew M Stuart. Inverse problems: a Bayesian perspective. *Acta Numerica*, 19:451–559, 2010.

[53] Masoumeh Dashti and Andrew M Stuart. The Bayesian approach to inverse problems. *Handbook of Uncertainty Quantification*, pages 1–118, 2016.

[54] Christophe Geuzaine and Jean-François Remacle. Gmsh: A 3-D finite element mesh generator with built-in pre-and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009.

[55] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98, 2017.

[56] Michael P Siegenthaler, Jürgen Martin, and Friedhelm Beyersdorf. Mechanical circulatory assistance for acute and chronic heart failure. *Journal of Interventional Cardiology*, 16(6):563–572, 2003.

[57] Christopher K Hebert, Rodrek E Williams, Richard S Levy, and Robert L Barrack. Cost of treating an infected total knee replacement. *Clinical Orthopaedics and Related Research*, 331:140–145, 1996.

[58] Rabih O Darouiche. Treatment of infections associated with surgical implants. *New England Journal of Medicine*, 350(14):1422–1429, 2004.

[59] Muhammad R Sohail, Daniel Z Uslan, Akbar H Khan, Paul A Friedman, David L Hayes, Walter R Wilson, James M Steckelberg, Sarah M Stoner, and Larry M Baddour. Risk factor analysis of permanent pacemaker infection. *Clinical Infectious Diseases*, 45(2):166–173, 2007.

[60] E Presterl, A Lassnigg, B Parschalk, F Yassin, H Adametz, and W Graninger. Clinical behavior of implant infections due to staphylococcus epidermidis. *The International Journal of Artificial Organs*, 28(11):1110–1118, 2005.

[61] David N Fisman, Donald T Reilly, Adolf W Karchmer, and Sue J Goldie. Clinical effectiveness and cost-effectiveness of 2 management strategies for infected total hip arthroplasty in the elderly. *Clinical Infectious Diseases*, 32(3):419–430, 2001.

[62] T Schulin and A Voss. Coagulase-negative staphylococci as a cause of infections related to intravascular prosthetic devices: limitations of present therapy. *Clinical Microbiology and Infection*, 7(s4):1–7, 2001.

[63] JW Costerton, L Montanaro, and CR Arciola. Biofilm in implant infections: its production and regulation. *The International Journal of Artificial Organs*, 28(11):1062–1068, 2005.

[64] JW Costerton, Philip S Stewart, and EP Greenberg. Bacterial biofilms: a common cause of persistent infections. *Science*, 284(5418):1318–1322, 1999.

[65] Georges Limbert, Rebecca Bryan, Ross Cotton, Philippe Young, Luanne Hall-Stoodley, Sandeep Kathju, and Paul Stoodley. On the mechanics of bacterial biofilms on non-dissolvable surgical sutures: a laser scanning confocal microscopy-based finite element study. *Acta Biomaterialia*, 9(5):6641–6652, 2013.

[66] Stefan Hajdu, Andrea Lassnigg, Wolfgang Graninger, Alexander M Hirschl, and Elisabeth Presterl. Effects of vancomycin, daptomycin, fosfomycin, tigecycline, and ceftriaxone on staphylococcus epidermidis biofilms. *Journal of Orthopaedic Research*, 27(10):1361–1365, 2009.

[67] Elisabeth Presterl, Stefan Hajdu, Andrea M Lassnigg, Alexander M Hirschl, Johannes Holinka, and Wolfgang Graninger. Effects of azithromycin in combination with vancomycin, daptomycin, fosfomycin, tigecycline, and ceftriaxone on staphylococcus epidermidis biofilms. *Antimicrobial Agents and Chemotherapy*, 53(8):3205–3210, 2009.

[68] Stefan Hajdu, Johannes Holinka, Sonja Reichmann, Alexander M Hirschl, Wolfgang Graninger, and Elisabeth Presterl. Increased temperature enhances the antimicrobial effects of daptomycin, vancomycin, tigecycline, fosfomycin, and cefamandole on staphylococcal biofilms. *Antimicrobial Agents and Chemotherapy*, 54(10):4078–4084, 2010.

[69] J Holinka, M Pilz, B Kubista, E Presterl, and R Windhager. Effects of selenium coating of orthopaedic implant surfaces on bacterial adherence and osteoblastic cell growth. *Bone & Joint Journal*, 95(5):678–682, 2013.

[70] Andrej Trampuz and Werner Zimmerli. Antimicrobial agents in orthopaedic surgery. *Drugs*, 66(8):1089–1106, 2006.

[71] Clemens Heitzinger and Leila Taghizadeh. Analysis of the drift-diffusion-Poisson–Boltzmann system for nanowire and nanopore sensors in the alternating-current regime. *Communications in Mathematical Sciences*, 15(8):2303–2325, 2017.

[72] Alena Bulyha and Clemens Heitzinger. An algorithm for three-dimensional Monte-Carlo simulation of charge distribution at biofunctionalized surfaces. *Nanoscale*, 3(4):1608–1617, 2011.

[73] Peter A Markowich. *The stationary semiconductor device equations*. Springer, 1985.

[74] Peter A Markowich, Christian A Ringhofer, and Christian Schmeiser. *Semiconductor equations*. Springer-Verlag New York, Inc., 1990.

[75] Amirreza Khodadadian, Leila Taghizadeh, and Clemens Heitzinger. Optimal multi-level randomized quasi-Monte-Carlo method for the stochastic drift–diffusion-Poisson system. *Computer Methods in Applied Mechanics and Engineering*, 329:480–497, 2018.

[76] Ada Fort, Santina Rocchi, M Belén Serrano-Santos, Roberto Spinicci, and Valerio Vignoli. Surface state model for conductance responses during thermal-modulation of SnO-based thick film sensors: Part I—model derivation. *IEEE Transactions on Instrumentation and Measurement*, 55(6):2102–2106, 2006.

[77] Clemens Heitzinger and Leila Taghizadeh. Existence and local uniqueness for the stochastic drift-diffusion-Poisson system. *In preparation.*

[78] Leila Taghizadeh and Clemens Heitzinger. Existence and local uniqueness for the Stokes–Nernst–Planck–drift–diffusion–Poisson system modeling nanopore and nanowire sensors. *At press.*

[79] François Cuvelier, Caroline Japhet, and Gilles Scarella. An efficient way to perform the assembly of finite element matrices in Matlab and Octave. Technical Report 8305, Université Paris 13 and INRIA Paris-Rocquencourt, May 2013.

[80] Anders Forsgren, Philip E Gill, and Margaret H Wright. Interior methods for nonlinear optimization. *SIAM Review*, 44(4):525–597, 2002.

[81] Margaret Wright. The interior-point revolution in optimization: history, recent developments, and lasting consequences. *Bulletin of the American Mathematical Society*, 42(1):39–56, 2005.

[82] Hande Y Benson, David F Shanno, and Robert J Vanderbei. Interior-point methods for nonconvex nonlinear programming: Jamming and comparative numerical testing. *Operations Research and Financial Engineering, Princeton University, ORFE-00-02*, 2000.

[83] Bart Van Zeghbroeck. Principles of semiconductor devices. *Colarado University*, 2004.

[84] Scott Roy and Asen Asenov. Where do the dopants go? *Science*, 309(5733):388–390, 2005.

[85] Amirreza Khodadadian and Clemens Heitzinger. Basis adaptation for the stochastic nonlinear Poisson–Boltzmann equation. *Journal of Computational Electronics*, 15(4):1393–1406, 2016.

[86] Xiang-Wei Jiang, Hui-Xiong Deng, Jun-Wei Luo, Shu-Shen Li, and Lin-Wang Wang. A fully three-dimensional atomistic quantum mechanical study on random dopant-induced effects in 25-nm mosfets. *IEEE Transactions on Electron Devices*, 55(7):1720–1726, 2008.

[87] Russel E Caflisch. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, 7:1–49, 1998.

[88] Harald Niederreiter. *Random number generation and quasi-Monte Carlo methods.* Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 1992.

[89] Ian H Sloan and Henryk Woźniakowski. Tractability of multivariate integration for weighted Korobov classes. *Journal of Complexity*, 17(4):697–721, 2001.

[90] Frances Y Kuo. Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces. *Journal of Complexity*, 19(3):301–320, 2003.

[91] Bruno Tuffin. Randomization of quasi-Monte Carlo methods for error estimation: survey and normal approximation. *Monte Carlo Methods and Applications MCMA*, 10(3-4):617–628, 2004.

[92] G Ökten. Error reduction techniques in quasi-Monte Carlo integration. *Mathematical and Computer Modelling*, 30(7):61–69, 1999.

[93] Giray Ökten and Warren Eastman. Randomized quasi-Monte Carlo methods in pricing securities. *Journal of Economic Dynamics and Control*, 28(12):2399–2426, 2004.

[94] Ronald Cools, Frances Y Kuo, and Dirk Nuyens. Constructing embedded lattice rules for multivariate integration. *SIAM Journal on Scientific Computing*, 28(6):2162–2188, 2006.

[95] Ivan G Graham, Frances Y Kuo, James A Nichols, Robert Scheichl, Ch Schwab, and Ian H Sloan. Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. *Numerische Mathematik*, 131(2):329–368, 2015.

[96] Vasile Sinescu and Pierre L'Ecuyer. Existence and construction of shifted lattice rules with an arbitrary number of points and bounded weighted star discrepancy for general decreasing weights. *Journal of Complexity*, 27(5):449–465, 2011.

[97] Susanne Brenner and Ridgway Scott. *The mathematical theory of finite element methods*, volume 15. Springer Science & Business Media, 2007.

[98] Paul T Boggs and Jon W Tolle. Sequential quadratic programming. *Acta Numerica*, 4:1–51, 1995.

[99] Pu-yan Nie. An sqp approach with line search for a system of nonlinear equations. *Mathematical and Computer Modelling*, 43(3):368–373, 2006.

[100] Dong C Liu and Jorge Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1):503–528, 1989.

[101] Duan Chen and Guo-Wei Wei. Modeling and simulation of electronic structure, material interface and random doping in nano-electronic devices. *Journal of Computational Physics*, 229(12):4431–4460, 2010.

[102] Leila Taghizadeh, Banjamin Stadlbauer, and Clemens Heitzinger. Bayesian inversion for electrical-impedance tomography. *Submitted for publication.*

[103] Otmar Scherzer. *Handbook of Mathematical Methods in Imaging.* Springer Science & Business Media, 2010.

[104] Ralph C Smith. *Uncertainty Quantification: Theory, Implementation, and Applications*, volume 12. SIAM, 2013.

[105] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian Data Analysis.* CRC Press, 2003.

[106] Albert Tarantola. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation.* Elsevier, 1987. New York.

[107] Wences P Gouveia and John A Scales. Resolution of seismic waveform inversion: Bayes versus Occam. *Inverse Problems*, 13(2):323–349, 1997.

[108] Masoumeh Dashti and Andrew M Stuart. Uncertainty quantification and weak approximation of an elliptic inverse problem. *SIAM Journal on Numerical Analysis*, 49(6):2524–2542, 2011.

[109] Walter R Gilks, Sylvia Richardson, and David Spiegelhalter. *Markov Chain Monte Carlo in Practice.* Chapman & Hall, 1996.

[110] Christian P. Robert and George Cassella. *Monte Carlo Statistical Methods.* Springer Verlag, 1999.

[111] Malcolm Sambridge and Klaus Mosegaard. Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40(3):1–29, 2002.

[112] Klaus Mosegaard and Malcolm Sambridge. Monte Carlo analysis of inverse problems. *Inverse Problems*, 18(3):R29, 2002.

[113] Jeffrey S Rosenthal et al. Optimal proposal distributions and adaptive MCMC. *Handbook of Markov Chain Monte Carlo*, 4, 2011.

[114] Heikki Haario, Marko Laine, Antonietta Mira, and Eero Saksman. DRAM: efficient adaptive MCMC. *Statistics and Computing*, 16(4):339–354, 2006.

[115] Simon L Cotter, Gareth O Roberts, Andrew M Stuart, and David White. MCMC methods for functions: modifying old algorithms to make them faster. *Statistical Science*, pages 424–446, 2013.

[116] Heikki Haario, Eero Saksman, and Johanna Tamminen. Adaptive proposal distribution for random walk Metropolis algorithm. *Computational Statistics*, 14(3):375–396, 1999.

[117] Heikki Haario, Eero Saksman, Johanna Tamminen, et al. An adaptive Metropolis algorithm. *Bernoulli*, 7(2):223–242, 2001.

[118] Luke Tierney and Antonietta Mira. Some adaptive Monte Carlo methods for Bayesian inference. *Statistics in Medicine*, 18(1718):2507–2515, 1999.

[119] Peter J Green and Antonietta Mira. Delayed rejection in reversible jump Metropolis–Hastings. *Biometrika*, 88(4):1035–1053, 2001.

[120] Andrew Gelman, Gareth O Roberts, Walter R Gilks, et al. Efficient Metropolis jumping rules. *Bayesian Statistics*, 5(599-608):42, 1996.

[121] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[122] Leila Taghizadeh, Elisabeth Presterl, and Clemens Heitzinger. Bayesian inversion for a biofilm model including quorum sensing. *In Review*.

[123] Christopher M Waters and Bonnie L Bassler. Quorum sensing: cell-to-cell communication in bacteria. *Annu. Rev. Cell Dev. Biol.*, 21:319–346, 2005.

[124] Jacek Banasiak and Mustapha Mokhtar-Kharroubi. *Evolutionary Equations with Applications in Natural Sciences*. Springer, 2015.

[125] Tamás Czárán and Rolf F Hoekstra. Microbial communication, cooperation and cheating: quorum sensing drives the evolution of cooperation in bacteria. *PLOS ONE*, 4(8):e6655, 2009.

[126] Michael A Brockhurst, Michelle GJL Habets, Ben Libberton, Angus Buckling, and Andy Gardner. Ecological drivers of the evolution of public-goods cooperation in bacteria. *Ecology*, 91(2):334–340, 2010.

[127] Mallory R Frederick, C Kuttler, BA Hense, J Müller, and HJ Eberl. A mathematical model of quorum sensing in patchy biofilm communities with slow background flow. *Can Appl Math Q*, 18(3):267–298, 2010.

[128] Maryam Ghasemi and Hermann J Eberl. Extension of a regularization based time-adaptive numerical method for a degenerate diffusion-reaction biofilm growth model to systems involving quorum sensing. *Procedia Computer Science*, 108:1893–1902, 2017.

[129] Maryam Ghasemi, Burkhard A Hense, Hermann J Eberl, and Christina Kuttler. Simulation-based exploration of quorum sensing triggered resistance of biofilms to antibiotics. *Bulletin of Mathematical Biology*, 80:1736–1775, 2018.

[130] Franz Rothe. *Global solutions of reaction-diffusion systems*, volume 1072. Springer, 2006.

[131] Anna Marciniak-Czochra, Grzegorz Karch, and Kanako Suzuki. Unstable patterns in reaction–diffusion model of early carcinogenesis. *Journal de Mathematiques Pures et Appliquees*, 99(5):509–543, 2013.

[132] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media, 2010.

[133] M. G. Garroni, V. A. Solonnikov, and M. A. Vivaldi. Schauder estimates for a system of equations of mixed type. *Rendiconti di Matematica e delle sue Applicazioni*, 29(1):117–132, 2009.

[134] Beata Zatorska, Marion Groger, Doris Moser, Magda Diab-Elschahawi, Luigi Segagni Lusignani, and Elisabeth Presterl. Does extracellular DNA production vary in Staphylococcal biofilms isolated from infected implants versus controls? *Clinical Orthopaedics and Related Research*®, 475(8):2105–2113, 2017.

[135] Jintao Liu, Rosa Martinez-Corral, Arthur Prindle, D Lee Dong-yeon, Joseph Larkin, Marçal Gabalda-Sagarra, Jordi Garcia-Ojalvo, and Gürol M Süel. Coupling between distant biofilms and emergence of nutrient time-sharing. *Science*, 356(6338):638–642, 2017.

[136] Gary M Dunny and Bettina AB Leonard. Cell-cell communication in gram-positive bacteria. *Annual Reviews in Microbiology*, 51(1):527–564, 1997.

[137] David G Davies, Matthew R Parsek, James P Pearson, Barbara H Iglewski, JW t Costerton, and EP Greenberg. The involvement of cell-to-cell signals in the development of a bacterial biofilm. *Science*, 280(5361):295–298, 1998.

[138] James A Shapiro. Thinking about bacterial populations as multicellular organisms. *Annual Reviews in Microbiology*, 52(1):81–104, 1998.

[139] Arthur Prindle, Jintao Liu, Munehiro Asally, San Ly, Jordi Garcia-Ojalvo, and Gürol M Süel. Ion channels enable electrical communication in bacterial communities. *Nature*, 527(7576):59–63, 2015.

[140] David E Goldberg. Genetic algorithms in search. *Optimization, and MachineLearning*, 1989.

[141] James E Baker. Reducing bias and inefficiency in the selection algorithm. In *Proceedings of the Second International Conference on Genetic Algorithms*, volume 206, pages 14–21, 1987.

[142] Anis Younes, Frédérick Delay, Noura Fajraoui, M Fahs, and Thierry A Mara. Global sensitivity analysis and Bayesian parameter inference for solute transport in porous media colonized by biofilms. *Journal of Contaminant Hydrology*, 191:1–18, 2016.

[143] Oluwole K Oyebamiji, Darren J Wilkinson, Pahala Gedara Jayathilake, Steve P Rushton, Ben Bridgens, Bowen Li, and Paolo Zuliani. A Bayesian approach to modelling the impact of hydrodynamic shear stress on biofilm deformation. *PloS One*, 13(4):e0195484, 2018.

Curriculum Vitae

Personal Information

Name	**Leila Taghizadeh**
Geburtsdatum	05.02.1985
Geburtsort	Khoy, Iran
Nationalität	Iran
Email	`leila.taghizadeh@tuwien.ac.at`

Education

since 02/2015	Ph.D. studies in Technical Mathematics, Institute for Analysis and Scientific Computing, Vienna University of Technology, Vienna, Austria
since 02/2015	Project Assistant (FWF), Institute for Analysis and Scientific Computing, Vienna University of Technology, Vienna, Austria
10/2012–01/2015	Software Developer, Institute for Analysis and Scientific Computing, Vienna University of Technology, Vienna, Austria
12/2009–09/2012	Lecturer, Ekbatan University, Qazvin, Iran
09/2007–12/2009	M.Sc. studies in Applied Mathematics, Tarbiat Modares University, Tehran, Iran
09/2003–09/2007	B.Sc. studies in Mathematics, Urmia University, Urmia, Iran
09/2003	High school graduation, Bentolhoda school, Khoy, Iran

Journal Publications

10. L. Taghizadeh, B. Stadlbauer and C. Heitzinger. *Bayesian inversion for electrical impedance tomography.* Submitted for publication.

9. L. Taghizadeh, E. Presterl and C. Heitzinger. *Bayesian inversion for a biofilm model including quorum sensing.* In review.

8. L. Taghizadeh and C. Heitzinger. *Existence and local uniqueness for the Stokes-Nernst-Planck-drift-diffusion-Poisson system modeling nanopore and nanowire sensors.* Accepted for publication in Communications in Mathematical Sciences, 2019.

7. E. K. Lenzi, L. R. Evangelista, L. Taghizadeh, D. Pasterk, R. S. Zola, T. Sandev, C. Heitzinger and I. Petreska. *The Reliability of Poisson–Nernst–Planck Anomalous Models for Impedance Spectroscopy*, Journal of Physical Chemistry B, 123(37), 7885-7892, 2019.

6. B. Stadlbauer, A. Cossettini, J. A. Morales E., P. Scarbolo, L. Taghizadeh, C. Heitzinger and L. Selmi. *Bayesian Estimation of physical and geometrical parameters for nanocapacitor array biosensors*, Journal of Computational Physics, 397:108874, 2019.

5. C. Budd, O. Koch, L. Taghizadeh and E. B. Weinmüller, *Asymptotic properties of the space–time adaptive numerical solution of a nonlinear heat equation*, Calcolo 55(4):43, 2018.

4. A. Khodadadian, L. Taghizadeh and C. Heitzinger. *Three-dimensional optimal multi-level Monte-Carlo approximation of the stochastic drift-diffusion-Poisson system in nanoscale devices*, Journal of Computational Electronics, 17(1):76–89, 2018.

3. A. Khodadadian, L. Taghizadeh and C. Heitzinger. *Optimal multilevel randomized quasi-Monte-Carlo method for the stochastic drift-diffusion-Poisson system*, Computer Methods in Applied Mechanics and Engineering (CMAME), 329:480–497, 2018.

2. C. Heitzinger and L. Taghizadeh. *Analysis of the drift-diffusion-Poisson-Boltzmann system for nanowire and nanopore sensors in the alternating-current regime*, Communications in Mathematical Sciences, 15(8):2303–2325, 2017.

1. L. Taghizadeh, A. Khodadadian and C. Heitzinger. *The optimal multilevel,Monte-Carlo approximation of the stochastic drift-diffusion-Poisson system*, Computer Methods in Applied Mechanics and Engineering (CMAME), 318:739–761, 2017.

Contributions to Conferences and Workshops

22. A. Cossettini, B. Stadlbauer, J. A. Morales E., L. Taghizadeh, L. Selmi, and C. Heitzinger. *Determination of micro- and nano-particle properties by multi-frequency Bayesian methods and applications to nanoelectrode array sensors*, In Proc. IEEE Sensors 2019, Montreal, Canada, 27–30 October 2019, At press.

21. L. Taghizadeh and C. Heitzinger. *Bayesian inference for two inverse problems in tomography and biofilms*, In Proc. International Conference on Scientific Computation and Differential Equations 2019 (SciCADE 2019), page 280, Innsbruck, Austria, 22–26 July 2019.

20. C. Heitzinger, A. Khodadadian, D. Pasterk, and L. Taghizadeh. *Modeling and simulation of nanotechnological sensors*, In Proc. 9th International Congress on Industrial and Applied Mathematics (ICIAM 2019), page 309, Valencia, Spain, 15–19 July 2019.

19. C. Heitzinger and L. Taghizadeh. *Bayesian PDE inversion in electrical-impedance tomography*, In Proc. 9th International Congress on Industrial and Applied Mathematics (ICIAM 2019), page 352, Valencia, Spain, 15–19 July 2019.

18. L. Taghizadeh and C. Heitzinger. *Bayesian analysis for Poisson-Boltzmann equation modeling electrical impedance tomography devices.* In Proc. 90th Annual Meeting of the International Association of Applied Mathematics and Mechanics (GAMM), page 422, Vienna, Austria, 18–20 February 2019.

17. C. Heitzinger, J. Morales E., B. Stadlbauer, and L. Taghizadeh. *Bayesian estimation and machine learning: current problems and challenges with examples in electrical impedance tomography and nanosensors*, In Proc. Workshop on Research Challenges and Opportunities at the interface of Machine Learning and Uncertainty Quantification, page online, University of Southern California, Los Angeles, CA, USA, 4–6 June 2018.

16. B. Stadlbauer, L. Taghizadeh, J. Morales E., C. Heitzinger, A. Cossettini and L. Selmi. *Bayesian estimation for transport equations for nanocapacitors*, In Proc. SIAM Conference on Uncertainty Quantification (UQ 2018), pages 69–70, Garden Grove, CA, USA, 16–19 April 2018.

15. A. Cossettini, P. Scarbolo, J. Morales E., B. Stadlbauer, N. Muhammad, L. Taghizadeh, C. Heitzinger and L. Selmi. *Calibration, compensation, parameter estimation, and uncertainty quantification for nanoelectrode array biosensors*, In Proc. SIAM Conference on Uncertainty Quantification (UQ 2018), page 81, Garden Grove, CA, USA, 16–19 April 2018.

14. L. Taghizadeh, J. Morales E., B. Stadlbauer and C. Heitzinger. *Bayesian estimation for a tomography problem*, In Proc. SIAM Conference on Uncertainty Quantification (UQ 2018), page 138, Garden Grove, CA, USA, 16–19 April 2018.

13. L. Taghizadeh, A. Khodadadian and C. Heitzinger. *The stochastic drift-diffusion-Poisson system for modeling nanowire and nanopore sensors*, In P. Quintela et al., editors, Progress in Industrial Mathematics at ECMI 2016, volume 26, pages 309–317. Mathematics in Industry, Springer, 2017.

12. L. Taghizadeh, A. Khodadadian and C. Heitzinger. *Optimal multi-level randomized-quasi-Monte-Carlo methods for the stochastic drift-diffusion-Poisson system and for stochastic homogenization*, In Proc. International Conference on Scientific Computation and Differential Equations (SciCADE 2017), page 129, Bath, UK, 11–15 September 2017.

11. L. Taghizadeh, A. Khodadadian, S. Rigger and C. Heitzinger. *Optimal multi-level Monte-Carlo methods for the stochastic drift-diffusion-Poisson system and for stochastic homogenization*, In Proc. 14th US National Congress on Computational Mechanics (USNCCM), page 321, Montreal, Canada, 17–20 July 2017.

10. A. Khodadadian, L. Taghizadeh and C. Heitzinger. *Optimal multi-level Monte Carlo*

method for the stochastic drift-diffusion-Poisson system, In Proc. 13th Austrian Numerical Analysis Day 2017, page 16, Salzburg, Austria, 4–5 May 2017.

9. C. Heitzinger, A. Khodadadian G. Pammer, S. Rigger and L. Taghizadeh. *Optimal numerical methods for stochastic PDEs*, In Proc. SIAM Conference on Computational Science and Engineering (CSE 2017), page 129, Atlanta, GA, USA, 27 February–3 March 2017.

8. L. Taghizadeh, A. Khodadadian, and Clem. Heitzinger. *The stochastic drift-diffusion-Poisson system for modeling nanowire and nanopore sensors*, In Proc. 19th European Conference on Mathematics for Industry (ECMI 2016), Mathematics in Industry, page 86, Santiago de Compostela, Spain, 13–17 June 2016. Springer.

7. C. Heitzinger, C. Geiersbach and L. Taghizadeh. *Optimal numerical approaches to the stochastic homogenization of elliptic equations and to the stochastic drift-diffusion-Poisson system*, In Proc. SIAM Conference on Mathematical Aspects of Materials Science (MS16), page 136, Philadelpha, PA, USA, 8–12 May 2016.

6. L. Taghizadeh and . Heitzinger. *Optimal method for calculating solutions of the stochastic drift-diffusion-Poisson system*, In Proc. SIAM Conference on Uncertainty Quantification 2016 (SIAM UQ16), page 9, Lausanne, Switzerland, 5–8 April 2016.

5. L. Taghizadeh, A. Khodadadian and C. Heitzinger. *Stochastic modeling of dopant atoms in nanoscale transistors using multi-level Monte Carlo.* In Proc. SIAM Conference on Uncertainty Quantification 2016 (SIAM UQ16), page 9, Lausanne, Switzerland, 5–8 April 2016.

4. L. Taghizadeh, A. Khodadadian, C. Geiersbach and C. Heitzinger. *Optimal multi-level Monte-Carlo method for a system of stochastic PDEs*, In Proc. SIAM Conference on the Analysis of Partial Differential Equations (SIAM PD15), pages 76–77, Scottsdale, AZ, USA, 7–10 December 2015.

3. L. Taghizadeh, C. Geiersbach and C. Heitzinger. *Optimal approach for the numerical stochastic homogenization of elliptic problems*, In Proc. SIAM Conference on Analysis of Partial Differential Equations (SIAM PD15), pages 101–102, Scottsdale, AZ, USA, 7–10 December 2015.

2. C. Heitzinger and L. Taghizadeh. *The stochastic drift-diffusion-Poisson system for modeling nanoscale devices and a multi-level Monte-Carlo method*, In Proc. 24th International Conference on Transport Theory (ICTT 2015), pages 115–116, Taormina, Italy, 7–11 September 2015.

1. C. Heitzinger and L. Taghizadeh. *Existence and uniqueness for the Stokes-Nernst-Planck-drift-diffusion-Poisson system for modeling nanowire sensors and nanopores*, In Proc. 24th International Conference on Transport Theory (ICTT 2015), pages 117–118, Taormina,

Italy, 7–11 September 2015.

Wien, 2019
                                        _____
                                        Leila Taghizadeh