Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 318 (2017) 739–761
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

The optimal multilevel Monte-Carlo approximation of the stochastic drift–diffusion-Poisson system

Leila Taghizadeh[a,*], Amirreza Khodadadian[a], Clemens Heitzinger[a,b]

[a] *Institute for Analysis and Scientific Computing, Vienna University of Technology (TU Wien), Wiedner Hauptstraße 8–10, 1040 Vienna, Austria*
[b] *School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA*

Received 31 August 2016; received in revised form 12 January 2017; accepted 6 February 2017
Available online 15 February 2017

Highlights

- Optimal multilevel Monte-Carlo method for the coupled stochastic system is developed.
- Computational complexity is reduced significantly for a given error tolerance.
- Existence and local-uniqueness theorems for solutions of the system are presented.
- The method has many applications in nanoscale devices e.g. FinFETs and biosensors.

Abstract

Existence and local-uniqueness theorems for weak solutions of a system consisting of the drift–diffusion-Poisson equations and the Poisson–Boltzmann equation, all with stochastic coefficients, are presented. For the numerical approximation of the expected value of the solution of the system, we develop a multi-level Monte-Carlo (MLMC) finite-element method (FEM) and we analyze its rate of convergence and its computational complexity. This allows to find the optimal choice of discretization parameters. Finally, numerical results show the efficiency of the method. Applications are, among others, noise and fluctuations in nanoscale transistors, in field-effect bio- and gas sensors, and in nanopores.

Keywords: Stochastic drift–diffusion-Poisson system; Existence and uniqueness; Multi-level Monte-Carlo finite-element method; Optimal method

1. Introduction

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

* Corresponding author.
E-mail addresses: Leila.Taghizadeh@TUWien.ac.at (L. Taghizadeh), Amirreza.Khodadadian@TUWien.ac.at (A. Khodadadian), Clemens.Heitzinger@TUWien.ac.at (C. Heitzinger).

http://dx.doi.org/10.1016/j.cma.2017.02.014
0045-7825/

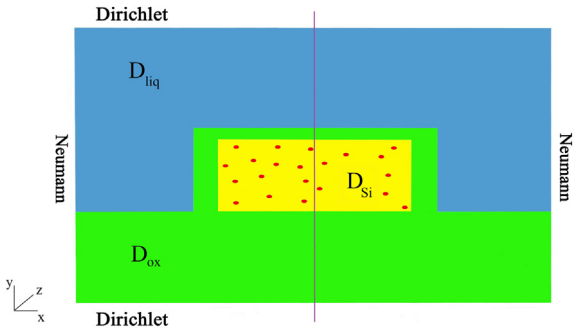

Fig. 1. A schematic diagram showing a transversal cross section for leading applications. A longitudinal cross section is indicated by the vertical line and is used for the finite-element discretization. In a field-effect transistor, the dopant atoms are randomly distributed (there is no electrolyte). In field-effect sensors, there are randomly distributed dopant atoms as well as randomly distributed charged molecules in the electrolyte.

(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 Fig. 1). 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 [1–6], 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 [7].

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 [8], after earlier work [9] 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 [10], 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 [11], 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 [12], 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 [13], the same problem was considered and numerical results indicate that the MLMC estimator is not limited to smooth problems. In [14], 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 [15], 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 [16], 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 [17], 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 than or equal to a given tolerance level.

The rest of this paper is organized as follows. In Section 2, we present the system of model equations with stochastic coefficients in detail. In Section 3, we define weak solutions of the model equations and prove existence and local-uniqueness theorems. Section 4.1 collects results about the FEM for later use. In Section 4, we introduce a multi-level Monte-Carlo finite-element method for the system and analyze its rate of convergence. In Section 5, we discuss the computational complexity and find the optimal MLMC method. In Section 6, 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, conclusions are drawn in Section 7.

2. 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 [1]. In summary, the domain is partitioned into

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

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

$$-\nabla \cdot (A(x,\omega)\nabla V(x,\omega)) = q(C_{\mathrm{dop}}(x,\omega) + p(x,\omega) - n(x,\omega)), \tag{1a}$$

$$\nabla \cdot J_n(x,\omega) = qR(n(x,\omega), p(x,\omega)), \tag{1b}$$

$$\nabla \cdot J_p(x,\omega) = -qR(n(x,\omega), p(x,\omega)), \tag{1c}$$

$$J_n(x,\omega) = q(D_n \nabla n(x,\omega) - \mu_n n(x,\omega)\nabla V(x,\omega)), \tag{1d}$$

$$J_p(x,\omega) = q(-D_p \nabla p(x,\omega) - \mu_p p(x,\omega)\nabla V(x,\omega)) \tag{1e}$$

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. Eqs. (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

$$\begin{aligned} n(x,\omega) &=: n_i \mathrm{e}^{V(x,\omega)/U_T} u(x,\omega), \\ p(x,\omega) &=: n_i \mathrm{e}^{-V(x,\omega)/U_T} v(x,\omega). \end{aligned}$$

The system (1) then becomes

$$\begin{aligned} -\nabla\cdot(A\nabla V(x,\omega)) &= q n_i (e^{-V(x,\omega)/U_T} v(x,\omega) - e^{V(x,\omega)/U_T} u(x,\omega)) + q C_{\text{dop}}(x,\omega), \\ U_T\nabla\cdot(\mu_n e^{V(x,\omega)/U_T}\nabla u(x,\omega)) &= \frac{u(x,\omega)v(x,\omega) - 1}{\tau_p(e^{V(x,\omega)/U_T}u(x,\omega)+1) + \tau_n(e^{-V(x,\omega)/U_T}v(x,\omega)+1)}, \\ U_T\nabla\cdot(\mu_p e^{-V(x,\omega)/U_T}\nabla v(x,\omega)) &= \frac{u(x,\omega)v(x,\omega) - 1}{\tau_p(e^{V(x,\omega)/U_T}u(x,\omega)+1) + \tau_n(e^{-V(x,\omega)/U_T}v(x,\omega)+1)}, \end{aligned}$$

where the continuity equations are self-adjoint.

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

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

At Ohmic contacts the space charge vanishes, i.e., $C_{\text{dop}} + p_D - n_D = 0$, 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 $\partial D_{\text{Si},D}$ 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

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

where

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

hold [18, Chapter 3]. Here, $C_{\text{dop}} := C_D^+ - C_A^-$ 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 [1]

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

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 [19], to solving the nonlinear Poisson–Boltzmann equation [7], or to systems of ordinary differential equations for surface reactions [20,21].

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

$$-\nabla \cdot (A(x, \omega)\nabla V(x, \omega)) = qC_{\text{dop}}(x, \omega) - qn_i(\mathrm{e}^{V(x,\omega)/U_T} u(x, \omega) - \mathrm{e}^{-V(x,\omega)/U_T} v(x, \omega)) \quad \text{in } D_{\text{Si}}, \tag{2a}$$
$$-\nabla \cdot (A(x, \omega)\nabla V(x, \omega)) = 0 \quad \text{in } D_{\text{ox}}, \tag{2b}$$
$$-\nabla \cdot (A(x, \omega)\nabla V(x, \omega)) = -2\eta \sinh(\beta(V(x, \omega) - \Phi(x, \omega))) \quad \text{in } D_{\text{liq}}, \tag{2c}$$
$$V(0+, y, \omega) - V(0-, y, \omega) = \alpha(y, \omega) \quad \text{on } \Gamma, \tag{2d}$$
$$A(0+)\partial_x V(0+, y, \omega) - A(0-)\partial_x V(0-, y, \omega) = \gamma(y, \omega) \quad \text{on } \Gamma, \tag{2e}$$
$$U_T \nabla \cdot (\mu_n \mathrm{e}^{V(x,\omega)/U_T} \nabla u(x, \omega)) = \frac{u(x, \omega)v(x, \omega) - 1}{\tau_p(\mathrm{e}^{V(x,\omega)/U_T} u(x, \omega) + 1) + \tau_n(\mathrm{e}^{-V(x,\omega)/U_T} v(x, \omega) + 1)} \quad \text{in } D_{\text{Si}}, \tag{2f}$$
$$U_T \nabla \cdot (\mu_p \mathrm{e}^{-V(x,\omega)/U_T} \nabla v(x, \omega)) = \frac{u(x, \omega)v(x, \omega) - 1}{\tau_p(\mathrm{e}^{V(x,\omega)/U_T} u(x, \omega) + 1) + \tau_n(\mathrm{e}^{-V(x,\omega)/U_T} v(x, \omega) + 1)} \quad \text{in } D_{\text{Si}}, \tag{2g}$$
$$\alpha(y, \omega) = M_\alpha(V(y, \omega)) \quad \text{in } \Gamma, \tag{2h}$$
$$\gamma(y, \omega) = M_\gamma(V(y, \omega)) \quad \text{in } \Gamma, \tag{2i}$$
$$V(x, \omega) = V_D(x) \quad \text{on } \partial D_D, \tag{2j}$$
$$\mathbf{n} \cdot \nabla V(x, \omega) = 0 \quad \text{on } \partial D_N, \tag{2k}$$
$$u(x, \omega) = u_D(x), \quad v(x, \omega) = v_D(x) \quad \text{on } \partial D_{D,\text{Si}}, \tag{2l}$$
$$\mathbf{n} \cdot \nabla u(x, \omega) = 0, \quad \mathbf{n} \cdot \nabla v(x, \omega) = 0 \quad \text{on } \partial D_{N,\text{Si}}. \tag{2m}$$

3. Existence and local uniqueness

In order to state the main theoretical results, we first record the assumptions on the data of the system (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.

3.1. Assumptions

The following assumptions are required.

Assumptions 1. 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 $meas(\Gamma \cap \partial D) = 0$ 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

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

   Furthermore, $A(x, \omega)|_{D^+ \times \Omega} \in C^1(D^+ \times \Omega, \mathbb{R}^{3\times 3})$ and $A(x, \omega)|_{D^- \times \Omega} \in C^1(D^- \times \Omega, \mathbb{R}^{3\times 3})$.
4. The doping concentration $C_{\mathrm{dop}}(x, \omega)$ is bounded above and below with the bounds

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

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

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

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.,

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

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

   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. Weak solution of the model equations

In order to define the weak formulation of the stochastic boundary-value problem (2), it suffices to consider the semilinear boundary-value problem

$$-\nabla \cdot (A^*(x, \omega)\nabla w(x, \omega)) + h(x, w(x, \omega)) = f(x, \omega) \quad \forall x \in D \setminus \Gamma \quad \forall \omega \in \Omega, \tag{3a}$$
$$w(x, \omega) = w_D(x) \quad \forall x \in \partial D_D \quad \forall \omega \in \Omega, \tag{3b}$$
$$\mathbf{n} \cdot \nabla w(x, \omega) = 0 \quad \forall x \in \partial D_N \quad \forall \omega \in \Omega, \tag{3c}$$
$$w(0+, y, \omega) - w(0-, y, \omega) = \alpha(y, \omega) \quad \forall x \in \Gamma \quad \forall \omega \in \Omega, \tag{3d}$$
$$A^*(0+)\partial_x w(0+, y, \omega) - A^*(0-)\partial_x w(0-, y, \omega) = \gamma(y, \omega) \quad \forall x \in \Gamma \quad \forall \omega \in \Omega, \tag{3e}$$

which is a semilinear Poisson equation with interface conditions. Here (3a) includes (2a)–(2c) if A^* is replaced by the permittivity A, and it includes (2f) and (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 1.

For the weak formulation, we define the Hilbert space

$$X := H^1_g(D) = \left\{ w \in H^1(D) \mid Tw = g \right\} \tag{4}$$

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

$$X_0 := H_0^1(D) = \left\{ w \in H^1(D) \mid Tw = 0 \right\}. \tag{5}$$

Definition 3.1 (*Bochner Spaces*)**.** Given a Banach space $(X, \|\cdot\|_X)$ and $1 \le p \le +\infty$, 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

$$\|w\|_{L^p(\Omega;X)} := \begin{cases} \left(\int_\Omega \|w(\cdot,\omega)\|_X^p \mathrm{d}\mathbb{P}(\omega) \right)^{1/p} = \mathbb{E}\Big[\|w(\cdot,\omega)\|_X^p \Big]^{1/p} < \infty, & 1 \le p < \infty, \\ \operatorname*{ess\,sup}_{\omega\in\Omega} \|w(\cdot,\omega)\|_X < \infty, & p = \infty \end{cases} \tag{6}$$

is finite.

To derive the variational formulation of our model (3), we fix the event $\omega \in \Omega$ at first, multiply (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 3.2 (*Weak Solution on $D \times \Omega$*)**.** Suppose that A^* satisfies Assumption 1 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), if it satisfies

$$a(w,\phi) = \ell(\phi) \qquad \forall \phi \in L^2(\Omega; X_0), \tag{7}$$

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

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

3.3. 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 (2) using the Schauder fixed-point theorem and the implicit-function theorem similarly to [2, Theorem 2.2 and 5.2].

Theorem 1 (*Existence*)**.** *Under Assumption 1, 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*

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

of the stochastic boundary-value problem (2), and for every $\omega \in \Omega$ it satisfies the L^∞-estimate

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

where

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

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

Here $V_L(x,\omega)$ is the solution of the linear problem (i.e., problem (3) with $h \equiv 0$), for which the estimate

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

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

$$\begin{aligned} N := \big\{&(V,u,v,\alpha,\gamma) \in L^2(\Omega;H^1(D)) \times L^2(\Omega;H^1(D_{\mathrm{Si}}))^2 \times L^2(\Omega;H^1(\Gamma))^2 \big| \\ &\underline{V} \le V(x,\omega) \le \overline{V} \quad \text{a.e. in } D\times\Omega, \quad \frac{1}{K} \le u(x,\omega), v(x,\omega) \le K \quad \text{a.e. in } D_{\mathrm{Si}}\times\Omega, \\ &\alpha,\gamma \text{ bounded a.e. on } \Gamma\times\Omega\big\}, \end{aligned}$$

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

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

   for V_1.
2. Solve the elliptic equation

$$\begin{aligned} -\frac{U_T\nabla\cdot(\mu_n e^{V_1/U_T}\nabla u_1)}{\dfrac{u_1v_0-1}{\tau_p(e^{V_1/U_T}u_0+1)+\tau_n(e^{-V_1/U_T}v_0+1)}} &= 0 && \text{in } D_{\mathrm{Si}}, \\ \mathbf{n}\cdot\nabla u_1 &= 0 && \text{on } \partial D_{\mathrm{Si},N}, \\ u_1 &= u_D && \text{on } \partial D_{\mathrm{Si},D} \end{aligned}$$

   for u_1.
3. Solve the elliptic equation

$$\begin{aligned} -\frac{U_T\nabla\cdot(\mu_p e^{-V_1/U_T}\nabla v_1)}{\dfrac{u_0v_1-1}{\tau_p(e^{V_1/U_T}u_0+1)+\tau_n(e^{-V_1/U_T}v_0+1)}} &= 0 && \text{in } D_{\mathrm{Si}} \\ \mathbf{n}\cdot\nabla v_1 &= 0 && \text{on } \partial D_{\mathrm{Si},N}, \\ v_1 &= v_D && \text{on } \partial D_{\mathrm{Si},D}, \end{aligned}$$

   for v_1.
4. Update the surface-charge density and dipole-moment density according to the microscopic model

$$\begin{aligned} \alpha_1(y,\omega) &:= M_\alpha(V_0), \\ \gamma_1(y,\omega) &:= M_\gamma(V_0). \end{aligned}$$

Using Lemmata on the existence and uniqueness of solutions of elliptic boundary-value problems with interface conditions [22], 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 [2, Theorem 2.2] and [23, Theorem 1]. Therefore, applying the Schauder fixed-point theorem yields a fixed-point of F, which is a weak solution of (2). □

In general, the solution in Theorem 1 is not unique; uniqueness of the solution only holds in a neighborhood around thermal equilibrium. This necessitates sufficiently small Dirichlet boundary conditions. The following theorem yields local uniqueness of the solution of our system (2) of model equation. The proof is based on the implicit-function theorem.

Theorem 2 (*Local Uniqueness*). *Under Assumption 1, 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

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

where G is given by the boundary-value problem (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 $L^2(\Omega; H^2_\partial(D)) \times L^2(\Omega; H^2_\partial(D_{\mathrm{Si}}))^2 \times L^2(\Omega; L^2(\Gamma))^2$ with

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

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., [2, Theorem 2.2]). Then the implicit-function theorem implies uniqueness of the solution of (2). □

4. 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 4.2. In Section 4.3, a multi-level Monte-Carlo (MLMC) finite-element (FE) method for the solution of the system of stochastic equations (2) is developed. We give an error bound for MLMC-FEM approximation and discuss the computational complexity.

4.1. 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

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

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

Definition 4.1 (*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

$$h_\ell = r^{-\ell} h_0, \tag{8}$$

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 (7) of the stochastic elliptic boundary-value problem (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 (4) and $X_{h_\ell} \subset X$ is the discretized space. For all $k \geq 1$, it is defined as

$$X_{h_\ell} := \mathbb{P}^k(\tau_{h_\ell}) := \{u \in X \mid u|_K \in \mathbb{P}^k(K) \ \ \forall K \in \tau_{h_\ell}\}, \tag{9}$$

where $\mathbb{P}^k(K) := \operatorname{span}\{x^\alpha \mid |\alpha| \leq k\}$ is the space of polynomials of total degree less than or equal to k. The space X_0 is the space (5) of test functions. The discretized test space $X_{0h_\ell} \subset X_0$ is defined analogously to (9).

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

Definition 4.2 (*Galerkin Approximation*)**.** Suppose $X_{h_\ell} \subset X$ and $X_{0h_\ell} \subset X_0$. The *Galerkin approximation* of (3) is the function

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

that satisfies

$$B(w_{h_\ell}, \phi_{h_\ell}) = F(\phi_{h_\ell}) \qquad \forall \phi_{h_\ell} \in L^2(\Omega; X_{0h_\ell}), \tag{10}$$

where B and F are defined in (7).

4.2. 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 E_{MC} for $\mathbb{E}[u_h]$ is the sample mean

$$\mathrm{E}_{\mathrm{MC}}[u_h] \coloneqq \hat{u}_h \coloneqq \frac{1}{M}\sum_{i=1}^{M} u_h^{(i)} \tag{11}$$

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

$$\|\mathbb{E}[u] - \mathrm{E}_{\mathrm{MC}}[u]\|_{L^2(\Omega;X)} = M^{-1/2}\sigma[u] \tag{12}$$

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

Proof. The result follows from the calculation

$$\begin{aligned}
\|\mathbb{E}[u] - \mathrm{E}_{\mathrm{MC}}[u]\|^2_{L^2(\Omega;X)} &= \mathbb{E}\Big[\Big\|\mathbb{E}[u] - \frac{1}{M}\sum_{i=1}^{M} u^{(i)}\Big\|_X^2\Big] \\
&= \frac{1}{M^2}\sum_{i=1}^{M}\mathbb{E}\Big[\|\mathbb{E}[u] - u^{(i)}\|_X^2\Big] \\
&= \frac{1}{M}\mathbb{E}\Big[\|\mathbb{E}[u] - u\|_X^2\Big] = M^{-1}\sigma^2[u]. \quad \square
\end{aligned}$$

Therefore, the variance of the MC estimator is

$$\sigma^2[\mathrm{E}_{\mathrm{MC}}[u]] = \|\mathbb{E}[\mathrm{E}_{\mathrm{MC}}[u]] - \mathrm{E}_{\mathrm{MC}}[u]\|^2_{L^2(\Omega;X)} = M^{-1}\sigma^2[u]. \tag{13}$$

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 3 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 1. *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*

$$\|\hat{u}_h - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))} = M^{-1}\sigma^2[u_h] + \|\mathbb{E}[u_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))}. \tag{14}$$

Proof. Starting from the mean square error, we calculate

$$\begin{aligned}\text{MSE} &:= \|\hat{u}_h - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))}\\ &= \mathbb{E}\Big[\|\hat{u}_h - \mathbb{E}[u]\|^2_{L^2(D)}\Big] = \mathbb{E}\Big[\int_D (\hat{u}_h - \mathbb{E}[u])^2 \mathrm{d}x\Big]\\ &= \int_D \mathbb{E}[(\hat{u}_h - \mathbb{E}[u])^2]\mathrm{d}x, \end{aligned} \tag{15}$$

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

$$\begin{aligned}\text{MSE} &= \int_D \mathbb{E}[(\hat{u}_h - \mathbb{E}[\hat{u}_h] + \mathbb{E}[\hat{u}_h] - \mathbb{E}[u])^2]\mathrm{d}x\\ &= \int_D \mathbb{E}[(\hat{u}_h - \mathbb{E}[\hat{u}_h])]\mathrm{d}x + \int_D \mathbb{E}[(\mathbb{E}[\hat{u}_h] - \mathbb{E}[u])^2]\mathrm{d}x\\ &= \|\hat{u}_h - \mathbb{E}[\hat{u}_h]\|^2_{L^2(\Omega;L^2(D))} + \|\mathbb{E}[u_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))}\\ &= \sigma^2[\hat{u}_h] + \|\mathbb{E}[u_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))}\\ &= M^{-1}\sigma^2[u_h] + \|\mathbb{E}[u_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))}, \end{aligned} \tag{16}$$

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 Eq. (13). □

Next we extend this result to H^1. In the following theorems, the finite-element space X is H^1 (see (4)).

Theorem 4. *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.,*

$$\|\mathbb{E}[u_h - u]\|_{L^2(\Omega;X)} \leq C_1 h^\alpha, \tag{17}$$

and that the estimate

$$\sigma^2[u_h] \leq C_0 \tag{18}$$

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

$$\|\hat{u}_h - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} = O(h^{2\alpha}) + O(M^{-1}). \tag{19}$$

Proof. We use the mean square error and calculate

$$\begin{aligned}\text{MSE} &:= \|\hat{u}_h - \mathbb{E}[u]\|^2_{L^2(\Omega;X)}\\ &= \|\hat{u}_h - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))} + \|\nabla\hat{u}_h - \mathbb{E}[\nabla u]\|^2_{L^2(\Omega;L^2(D))}\\ &= \|\hat{u}_h - \mathbb{E}[\hat{u}_h]\|^2_{L^2(\Omega;L^2(D))} + \|\mathbb{E}[\hat{u}_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))}\\ &\quad + \|\nabla\hat{u}_h - \mathbb{E}[\nabla\hat{u}_h]\|^2_{L^2(\Omega;L^2(D))} + \|\mathbb{E}[\nabla\hat{u}_h] - \mathbb{E}[\nabla u]\|^2_{L^2(\Omega;L^2(D))}\\ &= \|\hat{u}_h - \mathbb{E}[\hat{u}_h]\|^2_{L^2(\Omega;X)} + \|\mathbb{E}[\hat{u}_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)}\\ &= \sigma^2[\hat{u}_h] + \|\mathbb{E}[\hat{u}_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)}\\ &= M^{-1}\sigma^2[u_h] + \|\mathbb{E}[u_h] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)}. \end{aligned} \tag{20}$$

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 (13). Therefore, using the assumptions (17) and (18), we have

$$\text{MSE} \leq C_0 M^{-1} + (C_1 h^\alpha)^2 = O(M^{-1}) + O(h^{2\alpha}), \tag{21}$$

which concludes the proof. □

4.3. Multi-level Monte-Carlo finite-element approximation

In this section, we first present the MLMC FE method and 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 (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 4.1). 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

$$\mathbb{E}[u_{h_L}] = \mathbb{E}[u_{h_0}] + \mathbb{E}\left[\sum_{\ell=1}^{L} (u_{h_\ell} - u_{h_{\ell-1}})\right] = \mathbb{E}[u_{h_0}] + \sum_{\ell=1}^{L} \mathbb{E}[u_{h_\ell} - u_{h_{\ell-1}}]. \tag{22}$$

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

$$\mathrm{E}_{\mathrm{MLMC}}[u] := \hat{u}_{h_L} := \mathrm{E}_{\mathrm{MC}}[u_{h_0}] + \sum_{\ell=1}^{L} \mathrm{E}_{\mathrm{MC}}[u_{h_\ell} - u_{h_{\ell-1}}], \tag{23}$$

where E_{MC} is the Monte-Carlo estimator defined in (11). Therefore, we find

$$\hat{u}_{h_L} = \frac{1}{M_0} \sum_{i=1}^{M_0} u_{h_0}^{(i)} + \sum_{\ell=1}^{L} \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} (u_{h_\ell}^{(i)} - u_{h_{\ell-1}}^{(i)}). \tag{24}$$

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 5. *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.,*

$$\|\mathbb{E}[u_{h_\ell}] - \mathbb{E}[u]\|_{L^2(\Omega;X)} \leq C_1 h_\ell^\alpha, \tag{25}$$

the convergence order β for

$$\sigma^2[u_{h_\ell} - u_{h_{\ell-1}}] \leq C_0 h_{\ell-1}^\beta, \tag{26}$$

and assume that the estimate

$$\sigma^2[u_{h_0}] \leq C_{00} \tag{27}$$

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

$$\|\mathbb{E}[u] - \hat{u}_{h_L}\|^2_{L^2(\Omega;X)} = O(h_L^{2\alpha}) + O(M_0^{-1}) + \sum_{\ell=1}^{L} O(M_\ell^{-1}) O(h_{\ell-1}^\beta). \tag{28}$$

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

$$\begin{aligned}
\text{MSE} &:= \|\hat{u}_{h_L} - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= \|\hat{u}_{h_L} - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))} + \|\nabla\hat{u}_{h_L} - \mathbb{E}[\nabla u]\|^2_{L^2(\Omega;L^2(D))} \\
&= \|\hat{u}_{h_L} - \mathbb{E}[\hat{u}_{h_L}]\|^2_{L^2(\Omega;L^2(D))} + \|\mathbb{E}[\hat{u}_{h_L}] - \mathbb{E}[u]\|^2_{L^2(\Omega;L^2(D))} \\
&\quad + \|\nabla\hat{u}_{h_L} - \mathbb{E}[\nabla\hat{u}_{h_L}]\|^2_{L^2(\Omega;L^2(D))} + \|\mathbb{E}[\nabla\hat{u}_{h_L}] - \mathbb{E}[\nabla u]\|^2_{L^2(\Omega;L^2(D))} \\
&= \|\hat{u}_{h_L} - \mathbb{E}[\hat{u}_{h_L}]\|^2_{L^2(\Omega;X)} + \|\mathbb{E}[\hat{u}_{h_L}] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= \sigma^2[\hat{u}_{h_L}] + \|\mathbb{E}[\hat{u}_{h_L}] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)}. \qquad (29)
\end{aligned}$$

Expanding as in (20), using the relation $\sigma^2[\hat{u}_{h_L}] = \sum_{\ell=0}^{L} M_\ell^{-1}\sigma^2[u_{h_\ell} - u_{h_{\ell-1}}]$ [13], and finally applying the assumptions (25)–(27), we obtain the asserted estimate by observing that

$$\begin{aligned}
\text{MSE} &= M_0^{-1}\sigma^2[u_{h_0}] + \sum_{\ell=1}^{L} M_\ell^{-1}\sigma^2[u_{h_\ell} - u_{h_{\ell-1}}] + \|\mathbb{E}[u_{h_L}] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&\leq C_{00}M_0^{-1} + C_0\sum_{\ell=1}^{L} M_\ell^{-1}h_{\ell-1}^{\beta} + (C_1h_L^{\alpha})^2 \\
&= O(M_0^{-1}) + \sum_{\ell=1}^{L} O(M_\ell^{-1})O(h_{\ell-1}^{\beta}) + O(h_L^{2\alpha}), \qquad (30)
\end{aligned}$$

which concludes the proof. □

5. 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 (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

$$W := W_P + 2W_D = W_{P,a} + W_{P,s} + 2W_{D,a} + 2W_{D,s}, \qquad (31)$$

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

$$W_{\ell,P,a} = \mu_1 h_\ell^{-\gamma_1}, \qquad (32a)$$

$$W_{\ell,P,s} = \mu_2 h_\ell^{-\gamma_2}, \qquad (32b)$$

$$W_{\ell,D,a} = \mu_3 h_\ell^{-\gamma_3}, \qquad (32c)$$

$$W_{\ell,D,s} = \mu_4 h_\ell^{-\gamma_4} \qquad (32d)$$

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

$$W_\ell = W_{\ell,P,a} + W_{\ell,P,s} + 2(W_{\ell,D,a} + W_{\ell,D,s}). \qquad (33)$$

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 Eqs. (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., [24] 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 4.2). The constants $\mu_i > 0$ depend on the implementation.

5.1. 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 (32) and (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

$$\begin{aligned} &\underset{M,h}{\text{minimize}} \quad f(M,h) := MW \\ &\text{subject to} \quad g(M,h) := \frac{C_0}{M} + (C_1 h^\alpha)^2 - \varepsilon^2 \leq 0, \end{aligned} \tag{34}$$

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

$$\frac{C_0}{M} = \theta\varepsilon^2 \qquad \text{and} \qquad \left(C_1 h^\alpha\right)^2 = (1-\theta)\varepsilon^2. \tag{35}$$

By viewing h and M as functions of θ, (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 [25,26].

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

$$\begin{aligned} &\underset{\chi,s}{\text{minimize}} \quad f_\mu(\chi,s) := f(\chi) - \mu \sum_i \ln(s_i) \\ &\text{subject to} \quad g(\chi) - s = 0. \end{aligned} \tag{36}$$

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 (36) by y, the system

$$\begin{aligned} \nabla f(\chi) - \nabla g(\chi)^T y &= 0, \\ SYe &= \mu e, \\ g(\chi) - s &= 0 \end{aligned}$$

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

$$\begin{pmatrix} H(\chi,y) & 0 & -A(\chi)^T \\ 0 & Y & S \\ A(x) & -I & 0 \end{pmatrix} \begin{pmatrix} \Delta\chi \\ \Delta s \\ \Delta h \end{pmatrix} = \begin{pmatrix} -\nabla f(\chi) + A(\chi)^T y \\ \mu e - SYe \\ -g(\chi) + s \end{pmatrix}. \tag{37}$$

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 (34). The second equation is used to calculate Δs. By substituting into the third equation, we obtain the reduced KKT system

$$\begin{pmatrix} -H(\chi, y) & A(\chi)^T \\ A(\chi) & SY^{-1} \end{pmatrix} \begin{pmatrix} \Delta\chi \\ \Delta s \end{pmatrix} = \begin{pmatrix} \nabla f(\chi) - A(\chi)^T y \\ -h(\chi) + \mu Y^{-1} e \end{pmatrix}. \tag{38}$$

Now we use iteration to update the solutions by

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

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

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

is sufficiently reduced [27]. The parameter υ may increase with the iteration number to force the solution toward feasibility.

5.2. 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 MSE $\leq \varepsilon^2$. The optimal number L of levels is also unknown a priori. To this end, we solve the optimization problem

$$\begin{aligned} &\underset{M_\ell, h_0, r}{\text{minimize}} \quad f(M_\ell, h_0, r, L) := \sum_{\ell=0}^{L} M_\ell W_\ell \\ &\text{subject to} \quad g(M_\ell, h_0, r, L) := \frac{C_{00}}{M_0} + C_0 \sum_{\ell=1}^{L} \frac{h_{\ell-1}^\beta}{M_\ell} + (C_1 h_L^\alpha)^2 \leq \varepsilon^2. \end{aligned} \tag{39}$$

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

$$\frac{\partial}{\partial M_\ell}(f + \xi^2 g) = 0, \tag{40}$$

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

$$M_\ell = \xi \sqrt{V_\ell / W_\ell}, \tag{41}$$

where $V_0 = C_{00}$ and $V_\ell = C_0 h_{\ell-1}^\beta$. Similarly to (35), the equations

$$\frac{C_{00}}{M_0} + C_0 \sum_{\ell=1}^{L} \frac{h_{\ell-1}^\beta}{M_\ell} = \theta\varepsilon^2 \qquad \text{and} \qquad (C_1 h_L^\alpha)^2 = (1-\theta)\varepsilon^2 \tag{42}$$

hold. Hence, the Lagrange multiplier is given by

$$\xi = \left(\theta\epsilon^2\right)^{-1} \sum_{\ell=0}^{L} \sqrt{V_\ell W_\ell}. \tag{43}$$

Additionally, according to (42), h_0 is calculated by

$$h_0 = \left(\frac{\sqrt{1-\theta}\,\varepsilon}{C_1} \right)^{1/\alpha} r^L. \tag{44}$$

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 (8). In this case, we solve the minimization problem (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 6.

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

$$r_\ell := \prod_{i=1}^{\ell} r_i \qquad \text{and} \qquad r_i \geq 1. \tag{45}$$

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.

6. 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.

6.1. The leading example

Random-dopant effects are called discrete-dopant fluctuation effects [28–30]. 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 (red circles in Fig. 1). The domain $D \subset \mathbb{R}^2$ is depicted in Fig. 2. The thickness of the oxide layer is 8 nm, the thickness of the nanowire is 40 nm, its width is 60 nm and the nanowire length is 60 nm. Regarding the geometry, Dirichlet boundary conditions are used at the contacts with a back-gate voltage of -1 V (at the bottom of the device) and an electrode voltage of 0 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 $4 \cdot 10^{16}$ cm^{-3}. According to its volume, the silicon subdomain hence contains 6 negative impurity atoms when $C_{\mathrm{dop}} = 5 \cdot 10^{15}$ cm^{-3} and 600 dopants when $C_{\mathrm{dop}} = 5 \cdot 10^{17}$ cm^{-3}.

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

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

where σ is the influence parameter, C_j is the charge of the jth dopant atom, and x_j its position [31].

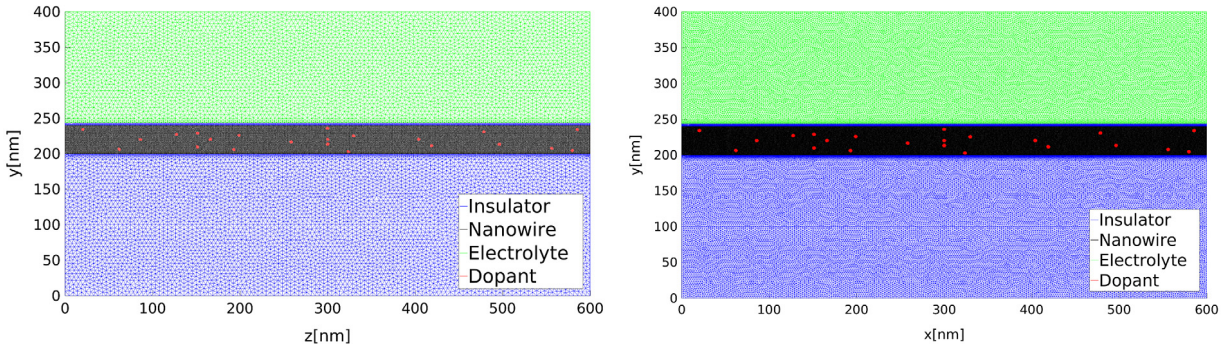

Fig. 2. Meshes for the random distribution of impurity atoms (red circles) in a nanowire field-effect sensor for levels $\ell = 0$ (left) and $\ell = 1$ (right), where $h_0 = 4.02$, $r = 2$, and $C_{\text{dop}} = 4 \times 10^{16}$ cm^{-3}. Additionally, oxide subdomain (D_{ox}), transducer (D_{Si}) and the electrolyte (D_{liq}) are depicted with blue, black and green meshes respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 2 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.

6.2. The computational work

As the first step, we calculate the coefficients in the expressions (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. Fig. 3 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 Fig. 4. 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 Fig. 4.

6.3. 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 5, we apply an iterative interior-point method to optimize both the number of samples and mesh sizes.

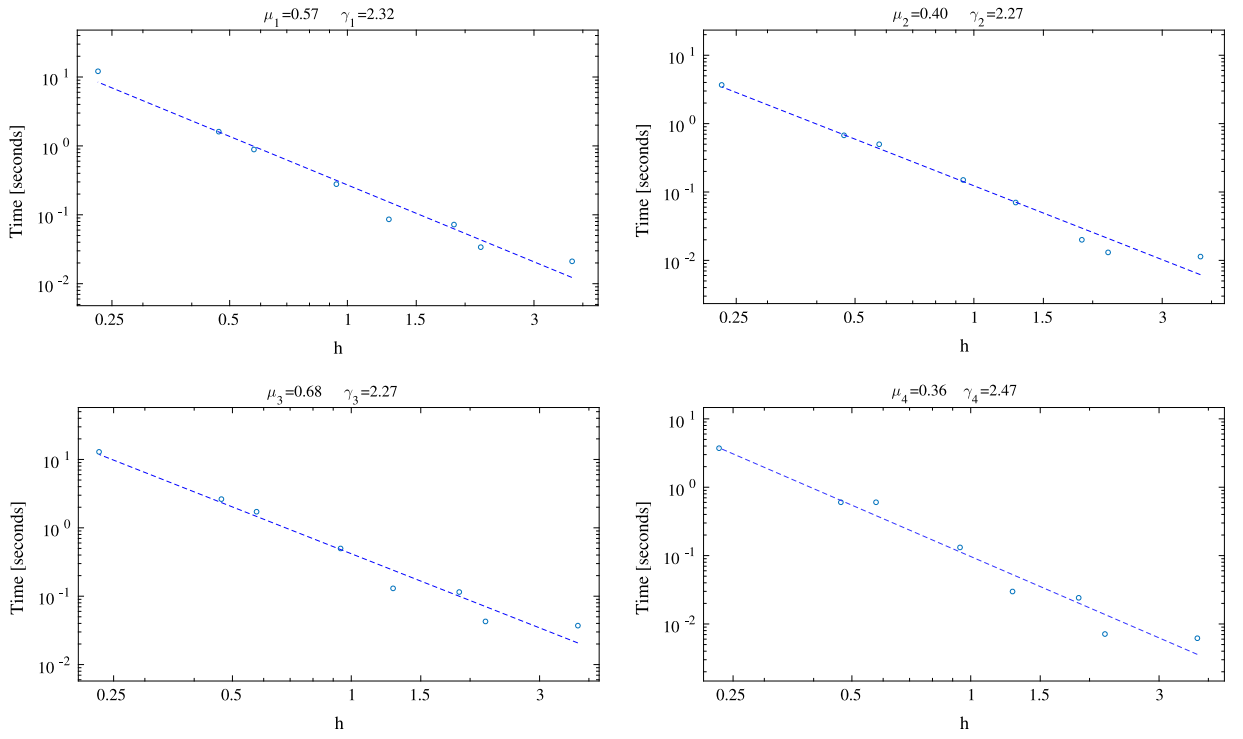

Fig. 3. Computational work for matrix assembly (top) and solving the system (bottom), both for the Poisson equation (left) and the drift–diffusion-equations (right).

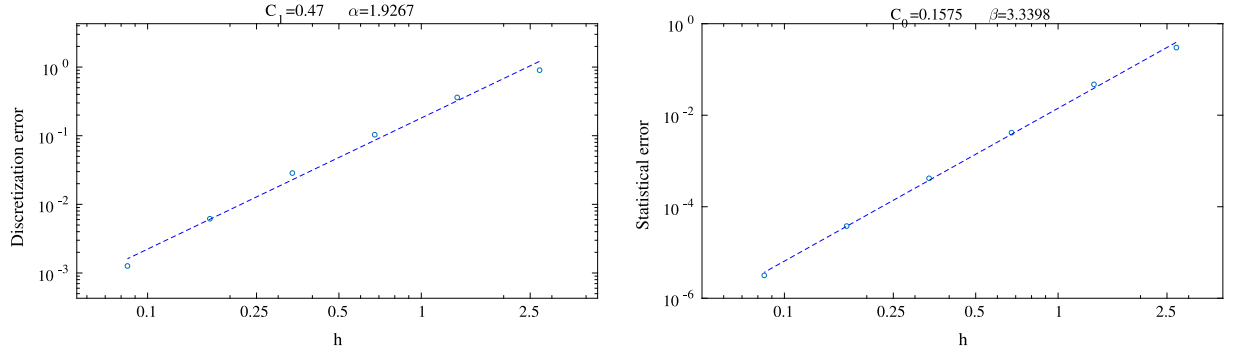

Fig. 4. Discretization error (left) and statistical (right) error as a function of h.

Table 1
Optimal MC FE method parameters for various given error tolerances.

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

6.3.1. Monte Carlo

First of all, we solve the optimization problem (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 1 for given ε.

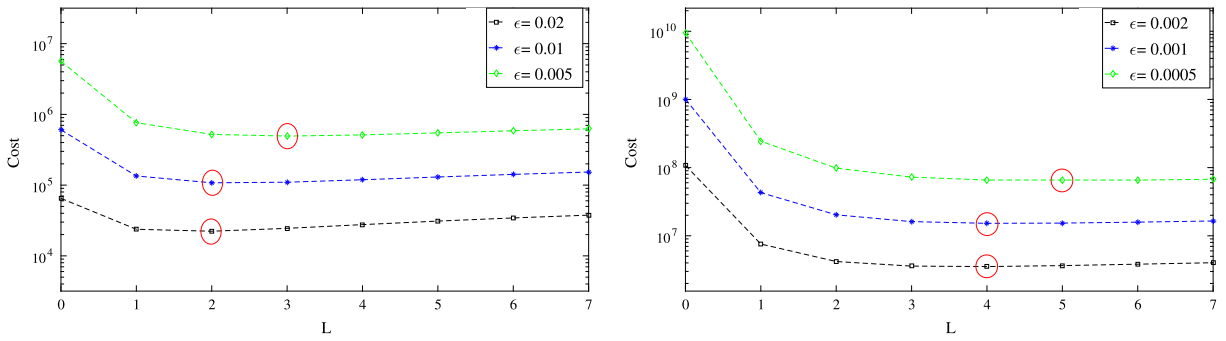

Fig. 5. The minimized computational work for the MLMC-FE method as a function of the number of levels and as a function of the given error tolerance. The results for a geometric progression for h (left) and general h (right) are shown. The number of levels yielding the minimal overall computational work is indicated by red circles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

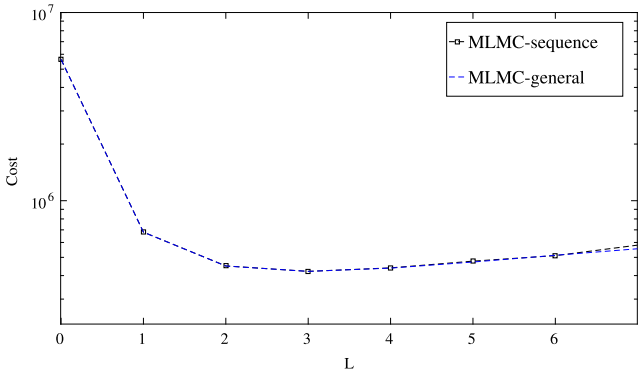

Fig. 6. Comparison between the two different approaches to MLMC FE method for $\varepsilon = 0.015$.

6.3.2. 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 (39) for $0 \leq L \leq 7$ 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 $\lceil M_\ell \rceil$, $\ell = 0, \ldots, L$, as the final numbers of levels.

The results of the optimization problems provide insight into the MLMC procedure. Fig. 5 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 Fig. 6, 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, 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 2 and 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.

6.3.3. Comparison

Finally, as Fig. 7 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 pronounced for smaller error bounds; for $\varepsilon = 0.001$, the computational work is about

Table 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	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	3575	438	54

Table 3
Optimal levels for the MLMC-FE method with general h_ℓ 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	M_3	M_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	1221	185	29
0.001	232 539	29 564	3701	455	55

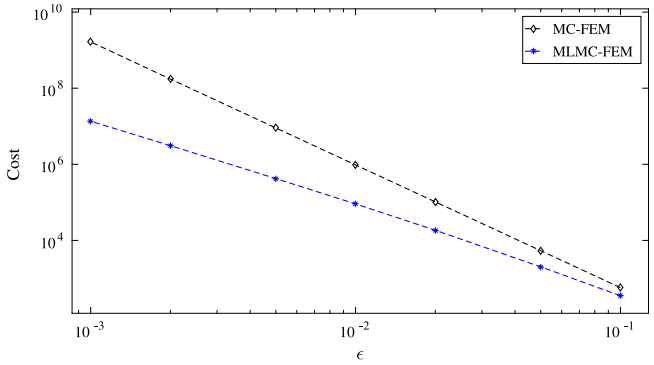

Fig. 7. Comparison of total computational work for MC-FEM and the two approaches to MLMC-FEM for various given tolerances.

a factor 10^2 lower than the Monte-Carlo work. The results agree with Giles' standard complexity theorem [13] 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 Fig. 7, the total cost of the MC-FEM is $O(\varepsilon^{-2-\gamma/\alpha})$, which agrees with [15]. 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 (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.

7. Conclusions

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.

Acknowledgments

The authors acknowledge support by FWF (Austrian Science Fund) START project no. Y660 *PDE Models for Nanotechnology*. The authors also acknowledge the helpful comments by the anonymous reviewers.

References

[1] C. Heitzinger, N.J. Mauser, C. Ringhofer, Multiscale modeling of planar and nanowire field-effect biosensors, SIAM J. Appl. Math. 70 (5) (2010) 1634–1654.
[2] S. Baumgartner, C. Heitzinger, Existence and local uniqueness for 3D self-consistent multiscale models for field-effect sensors, Commun. Math. Sci. 10 (2) (2012) 693–716.
[3] S. Baumgartner, C. Heitzinger, A one-level FETI method for the drift-diffusion-Poisson system with discontinuities at an interface, J. Comput. Phys. 243 (2013) 74–86. http://dx.doi.org/10.1016/j.jcp.2013.02.043.
[4] S. Baumgartner, C. Heitzinger, A. Vacic, M.A. Reed, Predictive simulations and optimization of nanowire field-effect PSA sensors including screening, Nanotechnology 24 (22) (2013) 225503/1–9. http://dx.doi.org/10.1088/0957-4484/24/22/225503. URL http://stacks.iop.org/0957-4484/24/225503.
[5] C. Heitzinger, C. Ringhofer, Multiscale modeling of fluctuations in stochastic elliptic PDE models of nanosensors, Commun. Math. Sci. 12 (3) (2014) 401–421. http://dx.doi.org/10.4310/CMS.2014.v12.n3.a1.
[6] G. Tulzer, C. Heitzinger, Fluctuations due to association and dissociation processes at nanowire-biosensor surfaces and their optimal design, Nanotechnology 26 (2) (2015) 025502.
[7] C. Heitzinger, Y. Liu, N.J. Mauser, C. Ringhofer, R.W. Dutton, Calculation of fluctuations in boundary layers of nanowire field-effect biosensors, J. Comput. Theor. Nanosci. 7 (12) (2010) 2574–2580.
[8] M.B. Giles, Multilevel Monte Carlo path simulation, Oper. Res. 56 (3) (2008) 607–617.
[9] S. Heinrich, Multilevel Monte Carlo methods, in: Large-Scale Scientific Computing, Springer, 2001, pp. 58–67.
[10] M. Giles, Improved multilevel Monte Carlo convergence using the Milstein scheme, in: Monte Carlo and Quasi-Monte Carlo Methods 2006, Springer, 2008, pp. 343–358.
[11] M.B. Giles, B.J. Waterhouse, Multilevel quasi-Monte Carlo path simulation, Adv. Financ. Modelling Radon Ser. Comput. Appl. Math. (2009) 165–181.
[12] A. Barth, C. Schwab, N. Zollinger, Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients, Numer. Math. 119 (1) (2011) 123–161.
[13] A. Cliffe, M. Giles, R. Scheichl, A. Teckentrup, Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients, Comput. Vis. Sci. 14 (1) (2011) 3–15.
[14] F.Y. Kuo, C. Schwab, I.H. Sloan, Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients, SIAM J. Numer. Anal. 50 (6) (2012) 3351–3374.
[15] J. Charrier, R. Scheichl, A.L. Teckentrup, Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods, SIAM J. Numer. Anal. 51 (1) (2013) 322–352.
[16] A. Teckentrup, R. Scheichl, M. Giles, E. Ullmann, Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients, Numer. Math. 125 (3) (2013) 569–600.
[17] A.-L. Haji-Ali, F. Nobile, E. von Schwerin, R. Tempone, Optimization of mesh hierarchies in multilevel Monte Carlo samplers, Stoch. Partial Differ. Equ. Anal. Comput. (2015) 1–37.

[18] P.A. Markowich, C. Ringhofer, C. Schmeiser, Semiconductor Equations, Springer, Wien, 1990.
[19] A. Bulyha, C. Heitzinger, An algorithm for three-dimensional Monte-Carlo simulation of charge distribution at biofunctionalized surfaces, Nanoscale 3 (4) (2011) 1608–1617.
[20] A. Fort, S. Rocchi, M.B. Serrano-Santos, R. Spinicci, V. Vignoli, Surface state model for conductance responses during thermal-modulation of SnO-based thick film sensors: Part I—model derivation, IEEE Trans. Instrum. Meas. 55 (6) (2006) 2102–2106.
[21] G. Tulzer, S. Baumgartner, E. Brunet, G.C. Mutinati, S. Steinhauer, A. Köck, P.E. Barbano, C. Heitzinger, Kinetic parameter estimation and fluctuation analysis of CO at SnO_2 single nanowires, Nanotechnology 24 (31) (2013) 315501/1–10. http://dx.doi.org/10.1088/0957-4484/24/31/315501. URL http://iopscience.iop.org/0957-4484/24/31/315501/.
[22] C. Heitzinger, L. Taghizadeh, Existence and local uniqueness of the stochastic drift-diffusion-Poisson system for nanowire sensors, in preparation.
[23] C. Heitzinger, L. Taghizadeh, Existence and local uniqueness for the Stokes-Nernst-Planck-drift-diffusion-Poisson system modeling nanopore and nanowire sensors, in preparation.
[24] F. Cuvelier, C. Japhet, G. Scarella, An Efficient Way to Perform the Assembly of Finite Element Matrices in Matlab and Octave, Tech. Rep. 8305, Université Paris 13 and INRIA Paris-Rocquencour, 2013.
[25] A. Forsgren, P.E. Gill, M.H. Wright, Interior methods for nonlinear optimization, SIAM Rev. 44 (4) (2002) 525–597.
[26] M. Wright, The interior-point revolution in optimization: history, recent developments, and lasting consequences, Bull. Amer. Math. Soc. 42 (1) (2005) 39–56.
[27] H.Y. Benson, D.F. Shanno, R.J. Vanderbei, Interior-point methods for nonconvex nonlinear programming: Jamming and comparative numerical testing, in: Operations Research and Financial Engineering, Princeton University, ORFE-00-02, 2000.
[28] B. Van Zeghbroeck, Principles of Semiconductor Devices, Colarado University, 2004.
[29] S. Roy, A. Asenov, Where do the dopants go? Science 309 (5733) (2005) 388–390.
[30] A. Khodadadian, C. Heitzinger, Basis adaptation for the stochastic nonlinear Poisson–Boltzmann equation, J. Comput. Electron. 15 (4) (2016) 1393–1406.
[31] X.-W. Jiang, H.-X. Deng, J.-W. Luo, S.-S. Li, L.-W. Wang, A fully three-dimensional atomistic quantum mechanical study on random dopant-induced effects in 25-nm MOSFETs, IEEE Trans. Electron Devices 55 (7) (2008) 1720–1726.