Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 329 (2018) 480–497
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Optimal multilevel randomized quasi-Monte-Carlo method for the stochastic drift–diffusion-Poisson system

Amirreza Khodadadian[a,*], Leila Taghizadeh[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 26 May 2017; received in revised form 22 September 2017; accepted 12 October 2017
Available online 26 October 2017

Abstract

In this paper, an optimal multilevel randomized quasi-Monte-Carlo method to solve the stationary stochastic drift–diffusion-Poisson system is developed. We calculate the optimal values of the parameters of the numerical method such as the mesh sizes of the spatial discretization and the numbers of quasi-points in order to minimize the overall computational cost for solving this system of stochastic partial differential equations. This system has a number of applications in various fields, wherever charged particles move in a random environment. It is shown that the computational cost of the optimal multilevel randomized quasi-Monte-Carlo method, which uses randomly shifted low-discrepancy sequences, is one order of magnitude smaller than that of the optimal multilevel Monte-Carlo method and five orders of magnitude smaller than that of the standard Monte-Carlo method. The method developed here is applied to a realistic transport problem, namely the calculation of random-dopant effects in nanoscale field-effect transistors.

Keywords: Multilevel randomized quasi-Monte-Carlo; Multilevel Monte-Carlo; Randomized quasi-Monte-Carlo; Optimal numerical method; Stochastic partial differential equation; Field-effect transistor

1. Introduction

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 [1–7] and have also been applied to the stochastic drift–diffusion-Poisson system [8]. In [8], 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.

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

https://doi.org/10.1016/j.cma.2017.10.015
0045-7825/

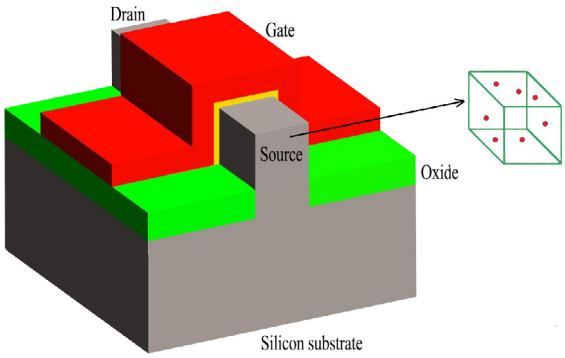

Fig. 1. Schematic structure of a three-dimensional FinFET. The random dopants are distributed in the source and drain regions.

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 [9] in the context of good lattice rules and later applied to the idea of general lattice rules in [10]. Later Tuffin [11] 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 [4] by combining the multilevel Monte-Carlo method with quasi-Monte-Carlo integration using a randomized rank-1 lattice rule. In [12], the variance estimation of the method and its convergence rate were investigated theoretically and numerically. In [7], 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.

We apply this approach to a problem of large computational cost and of practical importance. The scaling of conventional planar MOSFETs faced problems such as subthreshold swing degradation, significant drain-induced barrier lowering (DIBL), fluctuation of device characteristics, and current leakage [13–15]. To solve these problems, three-dimensional device structures were introduced. Fin field-effect transistors (FinFETs) are now state of the art in nanoscale CMOS technology and high-density memory applications [16–18]. In these devices, the current variation due to statistical fluctuations in the number and position of dopant atoms is the most serious problem when they are scaled into the deca-nanometer regime. In fact, the randomness of the dopant position and number make the fluctuation of device characteristics difficult to model and mitigate [19]. Here, we apply our numerical method to a realistic numerical example, i.e., a FinFET with a random number of randomly positioned dopants in the source and drain regions and we demonstrate its usefulness by comparing its computational effort to other methods. A schematic diagram of a three-dimensional FinFET is shown in Fig. 1. We assume that the domain $D \subset \mathbb{R}^3$ is bounded and convex, and it is partitioned into two main subdomains with different physical properties and consequently different

model equations in each of them. The first subdomain D_{Si} consists of the silicon channel (transducer), which is connected to the doped source and drain regions. The drift–diffusion-Poisson system describes charge transport in this region. The transducer is surrounded by the second subdomain D_{ox}, which is the oxide layer. In this subdomain, there are no charge carriers and hence just the Poisson equation is used. Dirichlet boundary conditions are applied at the source, drain, and gate contacts, and zero Neumann boundary conditions are imposed everywhere else.

The rest of this work is organized as follows. In Section 2, the model equations are presented. A randomized quasi-Monte-Carlo method including an error estimate is developed in Section 3. Based on this step, the multilevel version including again an error estimate is developed in Section 4. Then Section 5 is devoted to finding the optimal method. Finally, the numerical example is discussed in Section 6. Conclusions are drawn in Section 7.

2. The stochastic drift–diffusion-poisson system

The stationary drift–diffusion-Poisson system couples two elliptic transport equations for two charge carriers with the Poisson equation to ensure electrostatic self-consistency. The basic equation for the electrostatic potential here is the stochastic Poisson equation,

$$-\nabla\cdot(A(x)\nabla V(x,\omega))=\rho(x,\omega), \tag{1}$$

which is solved on the Lipschitz and bounded domain $D\subset\mathbb{R}^n$ ($n\in\{2,3\}$). Here, $x\in D$ is the spatial variable and $\omega\in\Omega$ is a realization in the d-dimensional sample space Ω, which is part of the probability space $(\Omega,\mathbb{A},\mathbb{P})$, where $\mathbb{A}$ is the σ-algebra of events and $\mathbb{P}:\mathbb{A}\to[0,1]$ is the probability measure. In (1), V denotes the electrostatic potential, A is the permittivity, and ρ is the charge density.

If a semiconductor is modeled, the charge density ρ is given by the electron and hole densities n and p and the doping concentration C, i.e.,

$$\rho=q(p-n+C), \tag{2}$$

where q is the elementary charge.

The stochastic drift–diffusion-Poisson equations are the system

$$-\nabla\cdot(A(x)\nabla V(x,\omega))=q\left(C(x,\omega)+p(x,\omega)-n(x,\omega)\right), \tag{3a}$$
$$\nabla\cdot J_n(x,\omega)=qR(x,\omega), \tag{3b}$$
$$-\nabla\cdot J_p(x,\omega)=qR(x,\omega), \tag{3c}$$
$$J_n(x,\omega)=q(D_n\nabla n(x,\omega)-\mu_n n(x,\omega)\nabla V(x,\omega)), \tag{3d}$$
$$J_p(x,\omega)=q(-D_p\nabla p(x,\omega)-\mu_p p(x,\omega)\nabla V(x,\omega)), \tag{3e}$$
$$R=\frac{np-n_i^2}{\tau_n(p+n_i)+\tau_p(n+n_i)}, \tag{3f}$$

where J_n and J_p are the current densities, μ_n and μ_p are the mobilities, and D_n and D_p are the diffusion coefficients. We assume that they are related by the Einstein relations $D_p=U_T\mu_p$ and $D_n=U_T\mu_n$. Furthermore, R is the recombination rate, whose precise form is not important here. We use the common Shockley–Read–Hall recombination rate where n_i is the intrinsic carrier density and τ_n and τ_p are the lifetimes of the free carriers.

We use the Slotboom variables u and v, which are defined by

$$p(x)=:n_i e^{-V/U_T}v, \tag{4}$$
$$n(x)=:n_i e^{V/U_T}u, \tag{5}$$

where $U_T:=K_BT/q$ denotes the thermal voltage, which is roughly 26 mV at room temperature, K_B is the Boltzmann constant, and T is the absolute temperature.

Using Slotboom variables, the drift–diffusion equations take the form

$$U_T n_i\nabla\cdot(\mu_p e^{-V/U_T}\nabla u(x,\omega))=R(x,\omega), \tag{6a}$$
$$U_T n_i\nabla\cdot(\mu_n e^{V/U_T}\nabla v(x,\omega))=R(x,\omega) \tag{6b}$$

with the recombination rate

$$R(x,\omega) = n_i \frac{u(x,\omega)v(x,\omega) - 1}{\tau_p(e^{V/U_T}u(x,\omega) + 1) + \tau_n(e^{-V/U_T}v(x,\omega) + 1)}.$$

Existence and local uniqueness of the solution of (3) are discussed in detail in [20]. Finally, the total current is found by integrating the current densities J_n and J_p along a cross section.

In certain applications such as field-effect sensors, the interface Γ between the device and a liquid can be characterized by interface conditions such as

$$\begin{aligned} V(0+,y,\omega) - V(0-,y,\omega) &= 0 \qquad \text{on } \Gamma, \\ A(0+)\partial_x V(0+,y,\omega) - A(0-)\partial_x V(0-,y,\omega) &= 0 \qquad \text{on } \Gamma, \end{aligned}$$

which are due to discontinuous material properties or a homogenization problem [21].

The boundary ∂D of the domain D is partitioned into two main parts, namely the Dirichlet part ∂D_D and the Neumann part ∂D_N. The conditions

$$V|_{\partial D_D} = V_D, \quad u|_{\partial D_{D,\mathrm{Si}}} = u_D, \quad \text{and} \quad v|_{\partial D_{D,\mathrm{Si}}} = v_D \tag{7}$$

are employed on the Dirichlet boundary, where $\partial D_{D,\mathrm{Si}}$ denotes the Dirichlet boundary of D_{Si}. These conditions are used for the potential V at the source, drain, and gate contacts, i.e., $V = V_S$, $V = V_D$, and $V = V_G$, respectively. Zero Neumann boundary conditions are applied everywhere else, i.e.,

$$\begin{aligned} \mathbf{n}\cdot\nabla V(x) &= 0 \quad \text{for } x \in \partial D_N, \\ \mathbf{n}\cdot\nabla u(x,\omega) &= 0 \quad \text{for } x \in \partial D_{N,\mathrm{Si}}, \\ \mathbf{n}\cdot\nabla v(x,\omega) &= 0 \quad \text{for } x \in \partial D_{N,\mathrm{Si}}, \end{aligned}$$

where $\partial D_{N,\mathrm{Si}}$ denotes the Neumann boundary of D_{Si}.

In summary, the stochastic transport model equations considered here with random input $C(x,\omega)$ are the system

$$-\nabla\cdot(A(x)\nabla V(x,\omega)) = q\left(C(x,\omega) - n_i\left(\mathrm{e}^{V(x,\omega)/U_T}u(x,\omega) - \mathrm{e}^{-V(x,\omega)/U_T}v(x,\omega)\right)\right) \quad \text{in } D_{\mathrm{Si}}, \tag{8a}$$

$$-\nabla\cdot(A(x)\nabla V(x,\omega)) = 0 \quad \text{in } D_{\mathrm{ox}}, \tag{8b}$$

$$V(0+,y,\omega) - V(0-,y,\omega) = 0 \quad \text{on } \Gamma, \tag{8c}$$

$$A(0+)\partial_x V(0+,y,\omega) - A(0-)\partial_x V(0-,y,\omega) = 0 \quad \text{on } \Gamma, \tag{8d}$$

$$U_T\nabla\cdot(\mu_n \mathrm{e}^{V(x,\omega)/U_T}\nabla u(x,\omega)) = n_i \frac{u(x,\omega)v(x,\omega) - 1}{\tau_p(e^{V/U_T}u(x,\omega) + 1) + \tau_n(e^{-V/U_T}v(x,\omega) + 1)} \quad \text{in } D_{\mathrm{Si}}, \tag{8e}$$

$$U_T\nabla\cdot(\mu_p \mathrm{e}^{-V(x,\omega)/U_T}\nabla v(x,\omega)) = n_i \frac{u(x,\omega)v(x,\omega) - 1}{\tau_p(e^{V/U_T}u(x,\omega) + 1) + \tau_n(e^{-V/U_T}v(x,\omega) + 1)} \quad \text{in } D_{\mathrm{Si}}, \tag{8f}$$

$$V(x,\omega) = V_D(x) \quad \text{on } \partial D_D, \tag{8g}$$

$$\mathbf{n}\cdot\nabla V(x,\omega) = 0 \quad \text{on } \partial D_N, \tag{8h}$$

$$u(x,\omega) = u_D(x), \quad v(x,\omega) = v_D(x) \quad \text{on } \partial D_{D,\mathrm{Si}}, \tag{8i}$$

$$\mathbf{n}\cdot\nabla u(x,\omega) = 0, \quad \mathbf{n}\cdot\nabla v(x,\omega) = 0 \quad \text{on } \partial D_{N,\mathrm{Si}}, \tag{8j}$$

which is solved for $V(x,\omega)$, $u(x,\omega)$, and $v(x,\omega)$.

3. Randomized quasi-Monte-Carlo method

In this section, 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 [4] and developed for a PDE in [7]. Here, the multi-level and QMC ideas are applied to a system of equations, and based thereon, an optimal method is developed.

3.1. 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 [22]. 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 limit the effectiveness of quasi-Monte-Carlo sequences [7,22].

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

$$\int_{[0,1]^d} f(\omega)\mathrm{d}\omega \approx \frac{1}{N}\sum_{j=1}^{N} f(\omega_j). \tag{9}$$

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 1 (*Koksma–Hlawka Theorem*)**.** *For any sequence* $\{\omega_j\}_{j\geq 1}$ *and any function* f *with bounded variation, the integration error due to* (9) *is bounded by*

$$\left|\frac{1}{N}\sum_{j=1}^{N} f(\omega_j) - \int_{[0,1]^d} f(\omega)\mathrm{d}\omega\right| \leq V_{\mathrm{HK}}(f) D_N^*(\omega), \tag{10}$$

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

$$V_{\mathrm{HK}}(f) := \int_{[0,1]^d} \left|\frac{\partial^d f}{\partial\omega^1\cdots\partial\omega^d}\right| \mathrm{d}\omega \tag{11}$$

for sufficiently differentiable f.

The first factor $V_{\mathrm{HK}}(f)$ in (10) is the variation of f in the sense of Hardy and Krause [23]. 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 (9) 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 [24], 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 [25], 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 6.

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 [26].

The accuracy of a QMC method can be improved by rewriting the function so that the variation term is reduced [27] or by constructing sequences that have smaller discrepancy [28]. Using RQMC methods with very low discrepancy sequences such as rank-1 lattice rules helps to increase accuracy and gives a useful error bound.

3.2. 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 [29] can be constructed as

$$\omega_j^{(i)} := \frac{j}{N}\lambda + \Delta^{(i)} \bmod 1, \quad j \in \{1,\dots,N\}, \quad i \in \{1,\dots,M\}, \tag{12}$$

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 [30, Section 4].

If the system (8) 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

$$X := H_g^1(D) = \{u \in H^1(D) \mid Tu = g\} \tag{13}$$

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

$$X_h := \mathcal{P}^k(\tau_h) := \{u \in X \mid u|_K \in \mathcal{P}^k(K) \ \ \forall K \in \tau_h\} \subset X \tag{14}$$

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

$$\mathbb{E}[u] = \int_{[0,1]^d} u(x,\omega)\mathrm{d}\mathbb{P}(\omega). \tag{15}$$

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

$$Q_{N,M}(u_h) := \frac{1}{M}\sum_{i=1}^{M}\frac{1}{N}\sum_{j=1}^{N} u_h(x,\omega_j^{(i)}) \tag{16}$$

using the quasi-random points defined in (12).

3.3. 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 [12]. Therefore, we assume that

$$V_{\mathrm{HK}}(u_h) \leq C_0, \tag{17}$$

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 [8]. As in [8], a prescribed accuracy is to be achieved, i.e., $\mathrm{MSE} \leq \varepsilon^2$.

Using the Koksma–Hlawka inequality (10) by calculating

$$\begin{aligned}
\sigma^2[Q_{N,M}(u_h)] &= \sigma^2\left[\frac{1}{M}\sum_{i=1}^{M}\frac{1}{N}\sum_{j=1}^{N} u_h(x,\omega_j^{(i)})\right] \\
&= \frac{1}{M}\sigma^2\left[\frac{1}{N}\sum_{j=1}^{N} u_h(x,\omega_j)\right] \\
&= \frac{1}{M}\int_{[0,1]^d}\left(\frac{1}{N}\sum_{j=1}^{N} u_h(x,\omega_j) - \mathbb{E}[u_h]\right)^2 \mathrm{d}\mathbb{P}(\omega)
\end{aligned} \tag{18}$$

$$\leq \frac{1}{M}\int_{[0,1]^d}\left(V_{\mathrm{HK}}(u_h)D_N^*(\omega_j)\right)^2 \mathrm{d}\mathbb{P}(\omega)$$
$$= \mathcal{O}(V_{\mathrm{HK}}^2(u_h)N^{-2+\delta}) \qquad \forall\delta>0,$$

the variance of the RQMC estimator (16) is estimated following [12]. In fact, in rank-1 lattice rules, the discrepancy satisfies

$$D_N^*(\omega_j) = \mathcal{O}(N^{-1+\delta}) \qquad \forall\delta>0 \tag{19}$$

for any number of points $N > 1$, any shift of the lattice, and for any dimension $d \geq 1$ [31]. 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 [29].

Using the boundedness assumption (17) for $V_{\mathrm{HK}}(u_h)$ in (18), we obtain the estimate

$$\sigma^2[Q_N(u_h)] \leq C_0 M^{-1}N^{-2+\delta} \qquad \forall\delta>0 \tag{20}$$

for the variance of the RQMC method, where C_0 is estimated using (17). Furthermore, δ is estimated in Section 6 (see Fig. 3). The inequality (20) 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 (8) in order to simplify notation, since all three components are in $H^1(D)$ for a given $\omega \in \Omega$.

Proposition 1. *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 (8). Assume further that the spatial discretization error converges with order α, i.e.,*

$$\|\mathbb{E}[u-u_h]\|_{L^2(\Omega;X)} \leq C_1 h^\alpha \qquad \exists C_1 \in \mathbb{R}^+, \tag{21}$$

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

$$\|Q_{N,M}(u_h) - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} = \mathcal{O}(M^{-1}N^{-2+\delta}) + \mathcal{O}(h^{2\alpha}) \qquad \forall\delta>0. \tag{22}$$

Proof. We estimate the mean square error (MSE). Using inequality (20) and assumption (21), we find that

$$\begin{aligned}
\mathrm{MSE} &:= \|Q_{N,M}(u_h) - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= \|Q_{N,M}(u_h) - \mathbb{E}[Q_{N,M}(u_h)]\|^2_{L^2(\Omega;X)} + \|\mathbb{E}[Q_{N,M}(u_h)] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)} \\
&= \sigma^2[Q_{N,M}(u_h)] + \|\mathbb{E}[u-u_h]\|^2_{L^2(\Omega;X)} \\
&\leq C_0 M^{-1}N^{-2+\delta} + C_1 h^{2\alpha} \\
&= \mathcal{O}(M^{-1}N^{-2+\delta}) + \mathcal{O}(h^{2\alpha})
\end{aligned} \tag{23}$$

for every $\delta > 0$. □

This means that the error behaves like (20).

4. 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.

4.1. The levels

In a multilevel approach, several mesh levels are used, and on each level, the RQMC estimator (16) 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

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

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

$$h_\ell = r^{-\ell} h_0, \tag{24}$$

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. (14)).

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

$$Q_{L,N_\ell,M_\ell}(u_{h_L}) := \frac{1}{M_0}\sum_{i=1}^{M_0}\frac{1}{N_0}\sum_{j=1}^{N_0} u_{h_0}(x,\omega_j^{(i)}) + \sum_{\ell=1}^{L}\frac{1}{M_\ell}\sum_{i=1}^{M_\ell}\frac{1}{N_\ell}\sum_{j=1}^{N_\ell}\big(u_{h_\ell}(x,\omega_j^{(i)}) - u_{h_{\ell-1}}(x,\omega_j^{(i)})\big). \tag{25}$$

The sample points $\omega_j^{(i)}$ are obtained using (12), for example, and their total number is $M_\ell N_\ell$.

4.2. 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 1. The assumptions on the boundedness of the variations of the FEM approximation and on the convergence order of the discretization error are

1. $V_{\mathrm{HK}}(u_{h_0}) \le C_{00} \quad \exists C_{00} \in \mathbb{R}^+$,
2. $V_{\mathrm{HK}}(u_{h_\ell} - u_{h_{\ell-1}}) \le C_0 h_{\ell-1}^{\beta} \quad \exists C_0 \in \mathbb{R}^+, \exists \beta \in \mathbb{R}^+$,
3. $\|\mathbb{E}[u - u_{h_\ell}]\|_{L^2(\Omega;X)} \le C_1 h_\ell^{\alpha} \quad \exists C_1 \in \mathbb{R}^+, \exists \alpha \in \mathbb{R}^+$.

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 2. *Suppose Assumptions 1 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 (8) 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

$$\|\mathbb{E}[u] - Q_{L,N_\ell,M_\ell}(u_{h_L})\|^2_{L^2(\Omega;X)} = \mathcal{O}(h_L^{2\alpha}) + \mathcal{O}(M_0^{-1}N_0^{-2+\delta}) + \sum_{\ell=1}^{L}\mathcal{O}(h_{\ell-1}^{2\beta}M_\ell^{-1}N_\ell^{-2+\delta}) \qquad \forall \delta > 0. \tag{26}$$

Proof. First, following [12], we estimate the variance of the multilevel RQMC estimator using inequality (10) by calculating

$$\begin{aligned}\sigma^2[Q_{L,N_\ell,M_\ell}(u_{h_L})] &= \sigma^2\Big[\frac{1}{M_0}\sum_{i=1}^{M_0}\frac{1}{N_0}\sum_{j=1}^{N_0} u_{h_0}(x,\omega_j^{(i)}) + \sum_{\ell=1}^{L}\frac{1}{M_\ell}\sum_{i=1}^{M_\ell}\frac{1}{N_\ell}\sum_{j=1}^{N_\ell}\big(u_{h_\ell}(x,\omega_j^{(i)}) - u_{h_{\ell-1}}(\omega_j^{(i)})\big)\Big] \\ &= \frac{1}{M_0}\sigma^2\Big[\frac{1}{N_0}\sum_{j=1}^{N_0} u_{h_0}(x,\omega_j)\Big] + \sum_{\ell=1}^{L}\frac{1}{M_\ell}\sigma^2\Big[\frac{1}{N_\ell}\sum_{j=1}^{N_\ell}\big(u_{h_\ell}(x,\omega_j) - u_{h_{\ell-1}}(x,\omega_j)\big)\Big]\end{aligned} \tag{27}$$

$$
\begin{aligned}
&= \frac{1}{M_0}\int_{[0,1]^d}\Big(\frac{1}{N_0}\sum_{j=1}^{N_0} u_{h_0}(x,\omega_j) - \mathbb{E}[u_{h_0}]\Big)^2 \mathrm{d}\mathbb{P}(\omega)\\
&\quad + \sum_{\ell=1}^{L}\frac{1}{M_\ell}\int_{[0,1]^d}\Bigg(\frac{1}{N_\ell}\sum_{j=1}^{N_\ell}\big(u_{h_\ell}(x,\omega_j)\\
&\quad - u_{h_{\ell-1}}(x,\omega_j)\big) - \mathbb{E}[u_{h_\ell}(x,\omega_j) - u_{h_{\ell-1}}(x,\omega_j)]\Bigg)^2 \mathrm{d}\mathbb{P}(\omega)\\
&\le \frac{1}{M_0}\int_{[0,1]^d}\Big(V_{HK}(u_{h_0})D^*_{N_0}(\omega_j)\Big)^2 \mathrm{d}\mathbb{P}(\omega)\\
&\quad + \sum_{\ell=1}^{L}\frac{1}{M_\ell}\int_{[0,1]^d}\Big(V_{HK}\big(u_{h_\ell}(x,\omega_j) - u_{h_{\ell-1}}(x,\omega_j)\big)D^*_{N_\ell}(\omega_j)\Big)^2 \mathrm{d}\mathbb{P}(\omega)\\
&= \mathcal{O}(V^2_{HK}(u_{h_0})N_0^{-2+\delta}) + \mathcal{O}(V^2_{HK}\big(u_{h_\ell}(x,\omega_j) - u_{h_{\ell-1}}(x,\omega_j)\big)N_\ell^{-2+\delta}),
\end{aligned}
$$

where we used the estimate (19).

Therefore, we have

$$
\sigma^2[Q_{L,N_\ell,M_\ell}(u_{h_L})] \le C_{00}M_0^{-1}N_0^{-2+\delta} + C_0\sum_{\ell=1}^{L} h_{\ell-1}^{2\beta}M_\ell^{-1}N_\ell^{-2+\delta} \tag{28}
$$

using the assumptions of bounded variations, i.e., Assumptions 1.1 and 1.2 This estimate shows how the error of the method behaves in terms of the number of samples (same as (20)), as we will see in the following.

Similarly to the RQMC estimator, the MSE assesses the accuracy of the MLRQMC-FE estimator. Using Assumptions 1.3 and the variance estimate (27), we find

$$
\begin{aligned}
\mathrm{MSE} &:= \|Q_{L,N_\ell,M_\ell}(u_{h_L}) - \mathbb{E}[u]\|^2_{L^2(\Omega;X)}\\
&= \|Q_{L,N_\ell,M_\ell}(u_{h_L}) - \mathbb{E}[Q_{L,N_\ell,M_\ell}(u_{h_L})]\|^2_{L^2(\Omega;X)} + \|\mathbb{E}[Q_{L,N_\ell,M_\ell}(u_{h_L})] - \mathbb{E}[u]\|^2_{L^2(\Omega;X)}\\
&= \sigma^2[Q_{L,N_\ell,M_\ell}(u_{h_L})] + \|\mathbb{E}[u - u_{h_L}]\|^2_{L^2(\Omega;X)}\\
&\le C_{00}M_0^{-1}N_0^{-2+\delta} + C_0\sum_{\ell=1}^{L} h_{\ell-1}^{2\beta}M_\ell^{-1}N_\ell^{-2+\delta} + (C_1h_L^\alpha)^2\\
&= \mathcal{O}(M_0^{-1}N_0^{-2+\delta}) + \sum_{\ell=1}^{L}\mathcal{O}(h_{\ell-1}^{2\beta})\mathcal{O}(M_\ell^{-1}N_\ell^{-2+\delta}) + \mathcal{O}(h_L^{2\alpha})
\end{aligned} \tag{29}
$$

for every $\delta > 0$. □

5. 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 2 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

$$W := W_{P,a} + W_{P,s} + 2W_{D,a} + 2W_{D,s}, \tag{30}$$

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 calculate the finite-element approximation u_{h_ℓ}, then we expect that $\gamma_k \approx n$ holds, where n is the number of spatial dimension [32]. In Section 6, 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 (8) is modeled as

$$\begin{aligned} W &:= \sum_{\ell=0}^{L} M_\ell N_\ell W_\ell \\ &= \sum_{\ell=0}^{L} M_\ell N_\ell (W_{\ell,P,a} + W_{\ell,P,s} + 2W_{\ell,D,a} + 2W_{\ell,D,s}) \\ &= \sum_{\ell=0}^{L} M_\ell N_\ell (\mu_1 h_\ell^{-\gamma_1} + \mu_2 h_\ell^{-\gamma_2} + \mu_3 h_\ell^{-\gamma_3} + \mu_4 h_\ell^{-\gamma_4}). \end{aligned} \tag{31}$$

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

$$\begin{aligned} &\underset{M_\ell, N_\ell, h_0, r}{\text{minimize}} \quad f(M_\ell, N_\ell, h_0, r, L) := \sum_{\ell=0}^{L} M_\ell N_\ell \sum_{k=1}^{4} \mu_k h_0^{-\gamma_k} r^{\gamma_k} \\ &\text{subject to} \quad g(N_\ell, h_0, r, L) := C_{00} M_0^{-1} N_0^{-2+\delta} + C_0 \sum_{\ell=1}^{L} h_{\ell-1}^{2\beta} M_\ell^{-1} N_\ell^{-2+\delta} + (C_1 h_L^{\alpha})^2 \le \varepsilon^2, \end{aligned} \tag{32}$$

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 (29), 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 (24)) and N_ℓ, $\ell \in \{0, 1, \ldots, L\}$. 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) [33] 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 (32), 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 $g(\chi) \approx g(\chi_s) + \nabla g(\chi_s)(\chi - \chi_s)$.

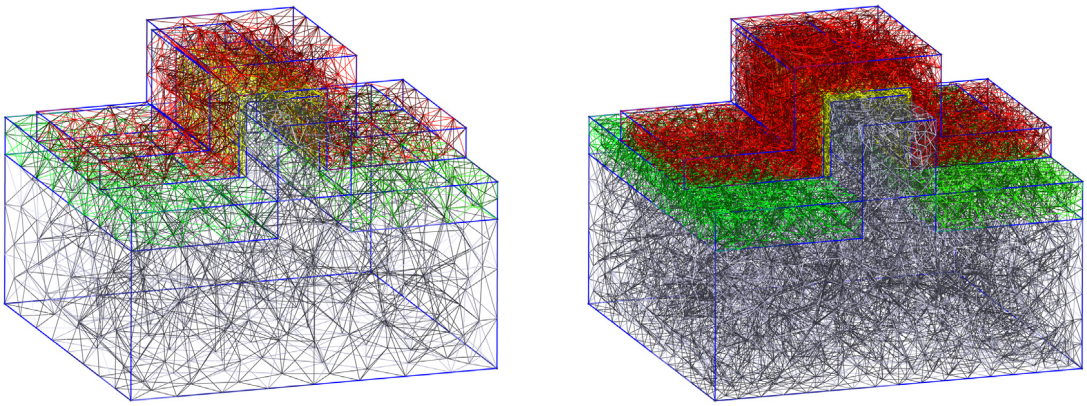

Fig. 2. The 3D meshes corresponding to the FinFET for $\ell = 0$ (left) and $\ell = 1$ (right).

Hence the minimization problem (32) yields the simplified, linearized QP subproblems

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2}\nu(\chi)^\top Hf(\chi_s)\nu(\chi) + \nabla f(\chi_s)^\top \nu(\chi) \\ &\text{subject to} \quad \nabla g(\chi_s)^\top \nu(\chi) + g(\chi_s) \leq 0, \end{aligned} \tag{33}$$

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 (33) and the step-length parameter α_s is determined by line search [34]. Also, H can be updated by any of the quasi-Newton methods, e.g., by the BFGS method [35].

6. Numerical example

A numerical example is presented here in order to illustrate the advantages of the method developed in Section 5. 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. In this section, by using the stochastic model, we study the effect of random-dopant fluctuations on the current through a FinFET.

6.1. The leading example and random dopants

The basic structure of FinFET is a channel controlled on three sides. A typical FinFET is shown in Fig. 1 and the 3D corresponding meshes for two different levels are illustrated in Fig. 2. This FinFET structure consists of a thin (vertical) silicon fin on a substrate, as mentioned in Section 1.

The numerical example discussed here is a realistic one: the dopant atoms in nanoscale transistors are distributed randomly resulting in unavoidable device variations between the many transistors in an integrated circuit. These random-dopant effects are of great importance in nanoscale devices.

The device parameters are the following. The permittivities are $A_{\text{Si}} = 11.7A_0$, $A_{\text{ox}} = 3.9A_0$, and the vacuum permittivity (dielectric constant) is $A_0 = 8.85 \cdot 10^{-12}$ F m^{-1}. Moreover, the gate length is 60 nm, and it is separated from the silicon channel by a 1.2 nm thick oxide layer. This channel is connected to the n-type doped source and drain regions of lengths $L_{\text{SD}} = 15$ nm and $V_{\text{SD}} = 0.1$ V. Regarding the boundary conditions in (8), Dirichlet boundary conditions are employed at the gate, source, and drain contacts. Zero Neumann boundary conditions are applied everywhere else.

The main source of randomness inside the device 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

$$C(x,\omega) = \sum_{j=1} C_j \delta(x - x_j(\omega)) \tag{34}$$

can be used, where x_j and C_j are the position and the charge of the jth dopant and $\delta(x - x_j(\omega))$ is the Dirac delta distribution at point $x_j(\omega)$. The Gaussian model [36,37]

$$C(x,\omega) := \sum_j \frac{C_j}{\left(2\pi\sigma^2\right)^{3/2}} \exp\left(-\frac{\left(x - x_j(\omega)\right)^2}{2\sigma^2}\right) \tag{35}$$

is a smoothed version of the point doping model (34), where σ is the so called influence parameter. Also, $\sigma := 0.3$ nm relates to the diameter of the direct electrostatic influence of the dopant; the results are not very sensitive to the value of σ. To make the results comparable between continuous and discrete doping models, the total doping must match. Hence, the integrals over a continuous doping concentration C_{dop} and over a discrete doping concentration must agree (for the fixed random variable ω^*), i.e.,

$$\int_{D_{\text{Si}}} C_{\text{dop}}(x)\mathrm{d}x = \int_{D_{\text{Si}}} C(x,\omega^*)\mathrm{d}x.$$

In the simulations, a three-dimensional domain is used and therefore the dimension d of random variables is equal to three times the number of dopants. In order to determine the location of each dopant, three random points are used and a translation is applied such that the dopants are in the source and drain regions. As an example, for x-axis, y-axis and z-axis, $\omega^* = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ indicates exactly the center of a region (source or drain). Also, we assume that the numbers of dopants in the source and drain regions are equal. Finally, it is noted that the standard continuum model is deterministic and that it cannot model any randomness; it is an important feature of the stochastic drift–diffusion-Poisson system that it includes various kinds of randomness.

The source and drain regions contain n-type dopants corresponding to a uniform doping concentration of $1 \cdot 10^{19}$ cm^{-3}, and the doping concentration in the channel is $2 \cdot 10^{16}$ cm^{-3}.

The electron and hole mobilities have a similar dependence on doping. For low doping concentrations, the mobility is almost constant and is primarily limited by phonon scattering. At higher doping concentrations the mobility decreases due to ionized impurity scattering with the ionized doping atoms. The actual mobility also depends on the type of dopant [38].

6.2. Computational cost

As discussed in Section 5, 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 (32) must be measured.

As already mentioned, the statistical error depends on the mesh size h and the number N of samples. Fig. 3 (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, Fig. 4 illustrates the discretization error for different mesh sizes, where the parameters were estimated using 100 samples by comparing the variance of the multilevel estimator (25) 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 1.

Since (32) 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

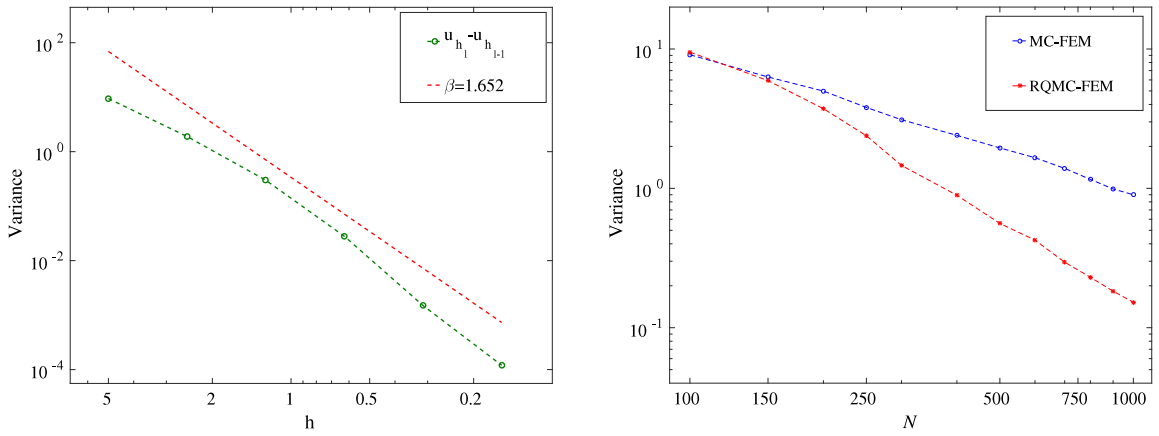

Fig. 3. The decay of variance of the solution as a function of different mesh sizes (left) and number of samples (right). The values $C_{00} = 9.45$, $C_0 = 0.338$, and $\delta = 0.12$ are found.

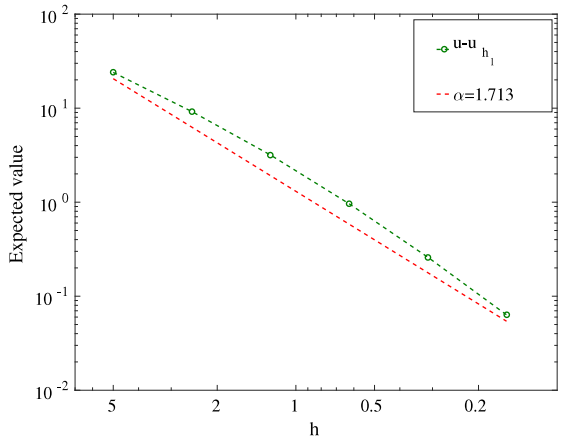

Fig. 4. The expected value of the solution as a function of different mesh sizes with $C_1 = 1.304$.

Table 1
The estimated coefficients and exponents in (31).

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 2
Optimal mesh size h and number N of samples for the QMC-FE method for different prescribed total errors ε.

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	1554	3844	9913

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 Tables 2, 3 and 4 respectively.

We compare a previously developed optimal MLMC-FE method [8] with the optimal MLRQMC-FE method developed in Section 5. Fig. 5 shows the computational work 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

Table 3
Optimal hierarchies in the MLRQMC-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	3107	603	117	25	5	2

Table 4
Optimal hierarchies in the MLMC-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	1688	193	23	3	–
0.010	1.459	1.9829	545 840	73 009	9182	1154	165	21	3

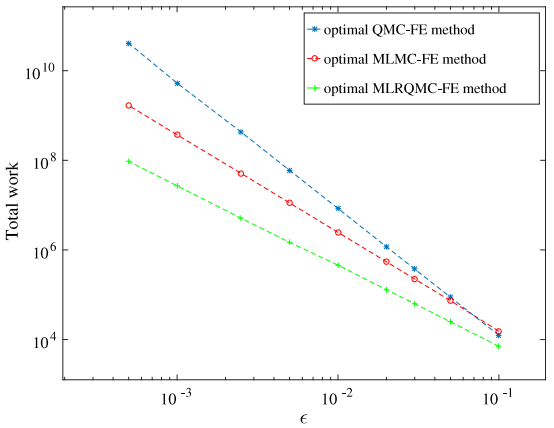

Fig. 5. Comparison of the total computational work required for the optimal MLRQMC and MLMC methods. For smaller total errors, the effectiveness of the randomized method is more pronounced.

method, the assumptions of the standard complexity theorem [6] 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. Fig. 6 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 $\ell \in \{0, \ldots, L\}$ 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.

6.3. Simulation of a FinFET

We focus on the subthreshold current, where the diffusion component of the current is larger than the drift component. First of all, we study the effect of randomness in the position of the dopants, whose number is constant.

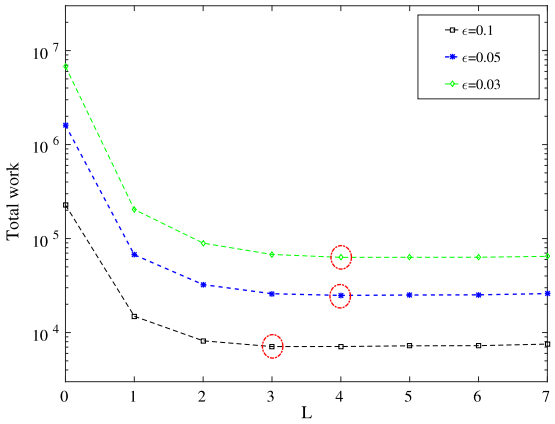

Fig. 6. The comparison of the total work of MLRQMC for different levels (between $L = 0$ and $L = 7$) for three different total errors $\varepsilon = 0.1$, $\varepsilon = 0.05$, and $\varepsilon = 0.03$. For each prescribed total error, the optimal number of levels is indicated by a red circle. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

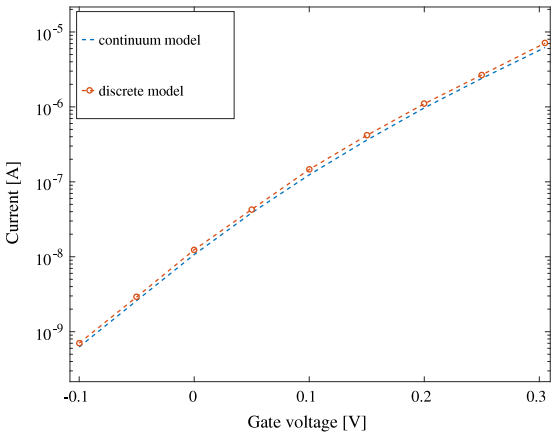

Fig. 7. The expected value of current as a function of different gate voltages calculated using continuum and discrete models.

Fig. 7 shows the comparison between the expected value of the current calculated using the discrete model (obtained for $\epsilon = 0.05$) and the continuum model for different gate voltages varying between $V_g = -0.1$ V and $V_g = 0.3$ V. The fluctuation of the current in the discrete model for $V_g = 0.1$ V and $V_g = 0.2$ V is also shown in Fig. 8.

Next, we compare the expected value of the current for different numbers N_{dop} of dopants with the continuum model. Fig. 9 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 Fig. 10, 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.

7. Conclusions

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 quasi-Monte-Carlo method to calculate the

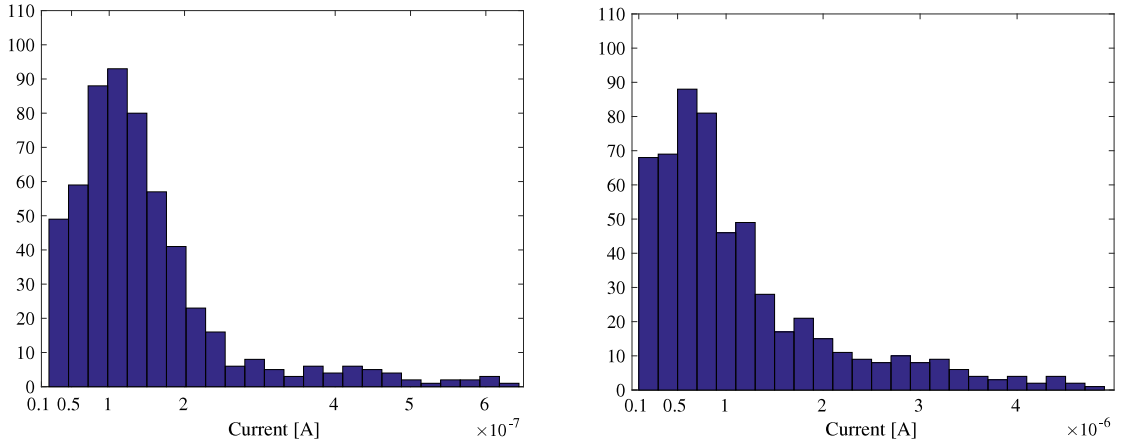

Fig. 8. Histogram of the current for $V_g = 0.1\,\mathrm{V}$ (left) and $V_g = 0.2\,\mathrm{V}$ (right), $N_{\mathrm{dop}} = 25$ and 563 simulations. Here $\mathbb{E}(I) = 1.49 \cdot 10^{-7}$ A for the lower gate voltage and $\mathbb{E}(I) = 1.14 \cdot 10^{-6}$ A for the higher gate voltage.

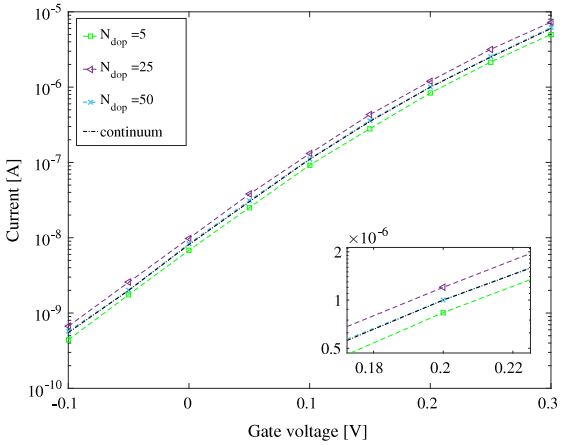

Fig. 9. The I–V characteristics for different numbers of dopants. The results for the continuum model are shown as well.

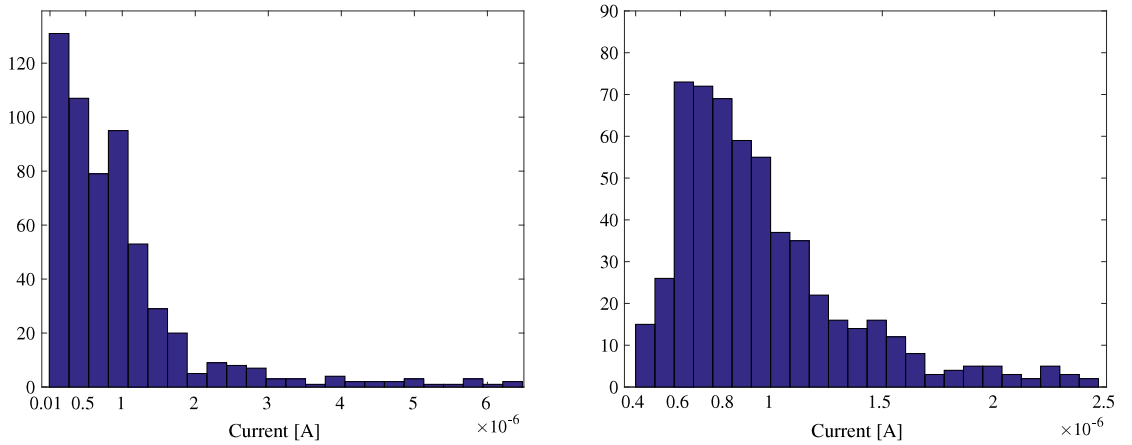

Fig. 10. Histogram of the current for $V_g = 0.2\,\mathrm{V}$ and for 563 simulations. Left: $N_{\mathrm{dop}} = 5$ resulting in $\mathbb{E}(I) = 9.22 \cdot 10^{-7}$ A. Right: $N_{\mathrm{dop}} = 50$, resulting in $\mathbb{E}(I) = 9.66 \cdot 10^{-7}$ A. The current obtained by the continuum model is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

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.

The numerical method developed here has also been applied to a realistic problem, namely the effects of random dopants in a state-of-the-art transistor. 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, which is consistent with random dopants being the main limiting factor in today's transistor technology.

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] S. Heinrich, Multilevel Monte Carlo methods, in: Large-Scale Scientific Computing, Springer, 2001, pp. 58–67.
[2] M.B. Giles, Multilevel Monte Carlo path simulation, Oper. Res. 56 (3) (2008) 607–617.
[3] 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.
[4] M.B. Giles, B.J. Waterhouse, Multilevel quasi-Monte Carlo path simulation, in: Advanced Financial Modelling, in: Radon Series on Computational and Applied Mathematics, 2009, pp. 165–181.
[5] 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.
[6] 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.
[7] F.Y. Kuo, C. Schwab, I.H. Sloan, Multi-level quasi-Monte Carlo finite element methods for a class of elliptic PDEs with random coefficients, Found. Comput. Math. 15 (2) (2015) 411–449.
[8] L. Taghizadeh, A. Khodadadian, C. Heitzinger, The optimal multilevel Monte-Carlo approximation of the stochastic drift-diffusion-Poisson system, Comput. Methods Appl. Mech. Eng. (CMAME) 318 (2017) 739–761 URL http://dx.doi.org/10.1016/j.cma.2017.02.014.
[9] R. Cranley, T. Patterson, Randomization of number theoretic methods for multiple integration, SIAM J. Numer. Anal. 13 (6) (1976) 904–914.
[10] S. Joe, Randomization of lattice rules for numerical multiple integration, J. Comput. Appl. Math. 31 (2) (1990) 299–304.
[11] B. Tuffin, et al., On the use of low discrepancy sequences in Monte Carlo methods, Monte Carlo Methods Appl. 2 (1996) 295–320.
[12] T. Gerstner, M. Noll, Randomized multilevel quasi-Monte Carlo path simulation, in: Recent Developments in Computational Finance: Foundations, Algorithms and Applications, World Scientific, 2013, pp. 349–369.
[13] J.-P. Colinge, Multiple-gate SOI MOSFETs, Solid-State Electron. 48 (6) (2004) 897–905.
[14] A. Asenov, A.R. Brown, J.H. Davies, S. Kaya, G. Slavcheva, Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs, IEEE Trans. Electron Devices 50 (9) (2003) 1837–1852.
[15] L. Chang, K.J. Yang, Y.-C. Yeo, I. Polishchuk, T.-J. King, C. Hu, Direct-tunneling gate leakage current in double-gate and ultrathin body MOSFETs, IEEE Trans. Electron Devices 49 (12) (2002) 2288–2295.
[16] Y.-K. Choi, T.-J. King, C. Hu, Nanoscale CMOS spacer FinFET for the terabit era, IEEE Electron Device Lett. 23 (1) (2002) 25–27.
[17] J.-P. Colinge, et al., FinFETs and Other Multi-Gate Transistors, Springer, 2008.
[18] I. Ferain, C.A. Colinge, J.-P. Colinge, Multigate transistors as the future of classical metal-oxide-semiconductor field-effect transistors, Nature 479 (7373) (2011) 310–316.
[19] Y. Li, C.-H. Hwang, T.-Y. Li, Random-dopant-induced variability in nano-CMOS devices and digital circuits, IEEE Trans. Electron Devices 56 (8) (2009) 1588–1597.
[20] C. Heitzinger, L. Taghizadeh, Existence and local uniqueness for the stochastic drift-diffusion-Poisson system, submitted for publication.
[21] 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.
[22] R.E. Caflisch, Monte Carlo and quasi-Monte Carlo methods, Acta Numer. 7 (1998) 1–49.
[23] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.
[24] I.H. Sloan, H. Woźniakowski, Tractability of multivariate integration for weighted Korobov classes, J. Complexity 17 (4) (2001) 697–721.
[25] F.Y. Kuo, Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces, J. Complexity 19 (3) (2003) 301–320.
[26] B. Tuffin, Randomization of quasi-Monte Carlo methods for error estimation: survey and normal approximation, Monte Carlo Methods Appl. Mcma 10 (3–4) (2004) 617–628.
[27] G. Ökten, Error reduction techniques in quasi-Monte Carlo integration, Math. Comput. Modelling 30 (7) (1999) 61–69.

[28] G. Ökten, W. Eastman, Randomized quasi-Monte Carlo methods in pricing securities, J. Econom. Dynam. Control 28 (12) (2004) 2399–2426.
[29] R. Cools, F.Y. Kuo, D. Nuyens, Constructing embedded lattice rules for multivariate integration, SIAM J. Sci. Comput. 28 (6) (2006) 2162–2188.
[30] I.G. Graham, F.Y. Kuo, J.A. Nichols, R. Scheichl, C. Schwab, I.H. Sloan, Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients, Numer. Math. 131 (2) (2015) 329–368.
[31] V. Sinescu, P. L'Ecuyer, Existence and construction of shifted lattice rules with an arbitrary number of points and bounded weighted star discrepancy for general decreasing weights, J. Complexity 27 (5) (2011) 449–465.
[32] S. Brenner, R. Scott, The Mathematical Theory of Finite Element Methods, Vol. 15, Springer Science & Business Media, 2007.
[33] P.T. Boggs, J.W. Tolle, Sequential quadratic programming, Acta Numer. 4 (1995) 1–51.
[34] P.-y. Nie, An SQP approach with line search for a system of nonlinear equations, Math. Comput. Modelling 43 (3) (2006) 368–373.
[35] D.C. Liu, J. Nocedal, On the limited memory BFGS method for large scale optimization, Math. Program. 45 (1) (1989) 503–528.
[36] 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.
[37] D. Chen, G.-W. Wei, Modeling and simulation of electronic structure, material interface and random doping in nano-electronic devices, J. Comput. Phys. 229 (12) (2010) 4431–4460.
[38] N.D. Arora, J.R. Hauser, D.J. Roulston, Electron and hole mobilities in silicon as a function of concentration and temperature, IEEE Trans. Electron Devices 29 (2) (1982) 292–295.