MATHEMATICAL TOOLS IN OPTIMAL SEMICONDUCTOR DESIGN

BY

MICHAEL HINZE AND RENÉ PINNAU

Abstract

This paper intends to give a comprehensive overview on the basic mathematical tools which are presently used in optimal semiconductor design. Focusing on the drift diffusion model for semiconductor devices we collect available results concerning the solvability of design problems and present for the first time results on the uniqueness of optimal designs. We discuss the construction of descent algorithms employing the adjoint state and investigate their numerical performance. The feasibility of this approach is underlined by various numerical examples.

1. Introduction

From the very beginning of semiconductor industry there has been a never–ending drive towards increased miniaturization. The rapidly increasing demand for semiconductor technology requires that the design cycle for a new device gets shorter from year to year. The original aim was to produce more devices per unit area, e.g. the Semiconductor Industry Association (SIA) projects that by 2009 the leading edge MOS device will employ a $0.05\,\mu$m length scale and an oxide thickness of 1.5 nm or less. So far, numerical simulations proved to be the main tool for the reduction of the time of a design cycle. These require enhanced models which are capable of describing

the electronic behavior of the device appropriately. This reveals several challenging problems for electrical engineers and applied mathematicians, too. In fact, they all have to contribute to the three topics *Modeling, Simulation and Optimal Design.*

Concerning the first point, there exists a hierarchy of models, which ranges from microscopic, like the Boltzmann–Poisson or the Wigner–Poisson model, to macroscopic models, like the energy transport, the hydrodynamic and the drift diffusion (DD) model [17, 25, 28, 33]. Having these models at hand it was possible to shorten the design cycle significantly. Numerical simulations helped to plan experiments and the computer could even replace expensive experiment set–ups. Clearly, also the development of new numerical techniques yielded an important speed up of the required simulation times. There was a strong need for new discretization schemes and non-linear iterations exploiting the structure of the underlying model equations, since they typically posed severe numerical problems. As the models consist of nonlinear elliptic, parabolic or hyperbolic equations, which are generally strongly coupled, this led to the development of many new mathematical techniques. Most popular and widely used in commercial simulation packages is the DD model, which allows for a very efficient numerical study of the charge transport in many cases of practical relevance. There exists a large amount of literature on this model (cf. [7] and the references therein), which covers questions of the mathematical analysis [12, 24, 26] as well as of the numerical discretization and simulation [13, 18].

The increasing computing power and the availability of fast simulation tools made it even possible to compute the optimal design of semiconductors. First approaches in the engineering literature were based on black box optimization tools or nonlinear least–squares methods applied to the well understood DD model [6, 8, 19, 20, 22, 27, 30, 31, 32]. Unfortunately, they require hundreds of solves of the underlying model equations. Nevertheless, they gave reasonable results and sometimes even unexpected answers to special design questions. Only recently the applied mathematics community did begin to investigate methods from optimization theory and optimal control of partial differential equations as well as specially designed optimization algorithms, which made it possible to speed up the numerical optimization tools significantly [1, 2, 3, 5, 9, 10, 11, 14, 15].

The main objective in optimal semiconductor design is to get an improved current flow at a specific contact of the device, e.g. focusing on the

reduction of the leakage current in MOSFET devices or maximizing the drive current [30]. In both cases a certain working point is fixed and one tries to achieve the objective by a change of the doping profile.

This paper intends to give a comprehensive overview on the basic mathematical tools which are presently used in optimal semiconductor design. We will focus on the DD model which is presented in Section 1.1 and embed the design question in the context of optimization with constraints given by partial differential equations (cf. Section 1.2). The question of uniqueness for optimal designs is investigated in Section 2, where a non–uniqueness result is proven. Section 3 is devoted to the discussion of the first–order optimality system leading to the adjoint state equations. These are used in Section 4 for the construction of a descent algorithm whose performance is studied for a new cost functional. Concluding remarks are given in Section 5.

1.1. The drift diffusion model

The stationary standard DD model for semiconductor devices stated on a bounded domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2$, or 3 reads

$$J_n = q \left(D_n \nabla n + \mu_n \, n \, \nabla V \right), \tag{1.1a}$$

$$J_p = -q \left(D_p \nabla p - \mu_p \, p \, \nabla V \right), \tag{1.1b}$$

$$\mathrm{div} \, J_n = 0, \tag{1.1c}$$

$$\mathrm{div} \, J_p = 0, \tag{1.1d}$$

$$-\epsilon \Delta V = q(n - p - C). \tag{1.1e}$$

The unknowns are the densities of electrons $n(x)$ and holes $p(x)$, the current densities of electrons $J_n(x)$ and holes $J_p(x)$, respectively, and the electrostatic potential $V(x)$. The total current density is given by

$$J = J_n + J_p. \tag{1.1f}$$

The doping profile is denoted by $C(x)$. The parameters D_n, D_p, μ_n, μ_p are the diffusion coefficients and mobilities of electrons and holes respectively. The physical constants are the elementary charge q and the permittivity constant ϵ.

Remark 1.1. For notational simplicity we assume that the device is operated near thermal equilibrium such that no generation–recombination

processes are present, but it is possible to extend the forthcoming ideas also to cases where recombination models of *Shockley–Read–Hall* or *Auger* type as well as *impact ionization* play a role [33].

In the following we will only consider regimes in which we can assume the Einstein relations

$$D_n = U_T\,\mu_n, \quad D_p = U_T\,\mu_p,$$

where $U_T = k_B\,T/q$ is the thermal voltage of the device and T denotes its temperature and k_B the Boltzmann constant. Further, let the mobilities be constant.

System (1.1) is supplemented with the following boundary conditions: We assume that the boundary $\partial\Omega$ of the domain Ω splits into two disjoint parts Γ_D and Γ_N, where Γ_D models the Ohmic contacts of the device and Γ_N represents the insulating parts of the boundary. Let ν denote the unit outward normal vector along the boundary. First, assuming charge neutrality and thermal equilibrium at the Ohmic contacts Γ_D and, secondly, zero current flow and vanishing electric field at the insulating part Γ_N yields the following set of boundary data

$$n = n_D, \quad p = p_D, \quad V = V_D \quad \text{on } \Gamma_D, \tag{1.1g}$$

$$J_n \cdot \nu = J_p \cdot \nu = \nabla V \cdot \nu = 0 \quad \text{on } \Gamma_N, \tag{1.1h}$$

where n_D, p_D, V_D are the $H^1(\Omega)$–extensions of

$$n_D = \frac{C + \sqrt{C^2 + 4\,n_i^2}}{2},$$

$$p_D = \frac{-C + \sqrt{C^2 + 4\,n_i^2}}{2},$$

$$V_D = -U_T\,\log\left(\frac{n_D}{n_i}\right) + U, \quad \text{on } \Gamma_D.$$

Here, U denotes the applied voltage and n_i the intrinsic carrier density.

We employ the following scaling

$$n \to C_m\,\tilde{n}, \qquad p \to C_m\,\tilde{p}, \qquad x \to L\,\tilde{x},$$

$$C \to C_m\,\tilde{C}, \qquad V \to U_T\,\tilde{V}, \qquad J_{n,p} \to \frac{q\,U_T\,C_m\,\mu_{n,p}}{L}\tilde{J}_{n,p}$$

where L denotes a characteristic device length, C_m the maximal absolute value of the background doping profile and $\mu_{n,p}$ a characteristic value for the respective mobilities. Defining the dimensionless *Debye length*

$$\lambda^2 = \frac{\epsilon \, U_T}{q \, C_m \, L^2}$$

the scaled equations read

$$\text{div}(\nabla n + n \, \nabla V) = 0, \qquad (1.2\text{a})$$

$$\text{div}(\nabla p - p \, \nabla V) = 0, \qquad (1.2\text{b})$$

$$-\lambda^2 \Delta V = n - p - C, \qquad (1.2\text{c})$$

where we eliminated the current densities and omitted the tilde for notational convenience. The Dirichlet boundary conditions transform to

$$n_D = \frac{C + \sqrt{C^2 + 4 \, \delta^4}}{2}, \qquad (1.2\text{d})$$

$$p_D = \frac{-C + \sqrt{C^2 + 4 \, \delta^4}}{2}, \qquad (1.2\text{e})$$

$$V_D = -\log\left(\frac{n_D}{\delta^2}\right) + U, \quad \text{on } \Gamma_D, \qquad (1.2\text{f})$$

where $\delta^2 = n_i / C_m$ denotes the scaled intrinsic density.

This system is analytically well understood, e.g. the solvability of the state equations for every $C \in H^1(\Omega)$ is a consequence of the following result whose proof can be found in [23, 26].

Proposition 1.2. *Let $\partial\Omega$ be regular. Then for each $C \in H^1(\Omega)$ and all boundary data $(n_D, p_D, V_D) \in H^1(\Omega)$ with*

$$\frac{1}{K} \leq n_D(x), p_D(x) \leq K, \quad x \in \Omega, \quad \text{and } \|V_D\|_{L^\infty(\Omega)} \leq K$$

for some $K \geq 1$, there exists a solution $(n, p, V) \in \left(H^1(\Omega) \cap L^\infty(\Omega)\right)^3$ of system (1.2) *fulfilling*

$$\frac{1}{L} \leq n(x), p(x) \leq L, \quad x \in \Omega, \quad \text{and } \|V\|_{L^\infty(\Omega)} \leq L$$

for some constant $L = L(\Omega, K, \|C\|_{L^p(\Omega)}) \geq 1$, where the embedding $H^1(\Omega) \hookrightarrow L^p(\Omega)$ holds.

Note, that in general one cannot expect uniqueness of solutions to the DD model. This will only hold near to the thermal equilibrium [25].

1.2. The optimization problem

The objective of the optimization, the current flow over a contact Γ, is given by

$$I = \int_\Gamma J \cdot \nu \, ds = \int_\Gamma (J_n + J_p) \cdot \nu \, ds. \tag{1.3}$$

This can be done minimizing the functional

$$Q(n,p,V,C) \stackrel{\text{def}}{=} Q_1(n,p,V) + \frac{\gamma}{2} \int_\Omega \left| \nabla(C - \bar{C}) \right|^2 \, dx, \tag{1.4}$$

where $\bar{C}$ is a given reference doping profile and the parameter γ allows to adjust the deviations from $\bar{C}$. One is mainly interested in functionals Q_1, which depend only on the values of the outflow current density on some contact Γ

$$Q_1(n,p,V) = R(J \cdot \nu|_\Gamma). \tag{1.5}$$

In [15], the functional under investigation was

$$R(J \cdot \nu|_\Gamma) = \frac{1}{2} \left\| (J - J^*) \cdot \nu \right\|^2_{(H_{00}^{1/2}(\Gamma))^*}, \tag{1.6}$$

corresponding to the objective of finding an outflow current density $J \cdot \nu$ close to a desired density $J^* \cdot \nu$. In [5] the total current flow on a contact is studied, i.e.

$$R(J \cdot \nu|_\Gamma) = \frac{1}{2} \left| \int_\Gamma J \cdot \nu \, ds - I^* \right|^2 \tag{1.7}$$

(for some desired current flow I^*). Note that especially in the one dimensional setting these two functionals are equivalent.

Since the current density J is given by a solution of the DD model this yields altogether a constrained optimization problem, which is well–known in the context of optimal control of nonlinear partial differential equations [16]. To get a solution to this problem one can follow two different ideas. On the one hand one can discretize the overall problem and then use nonlinear programming techniques [22]. On the other hand one can formulate the optimization algorithms on the continuous level and use then internal

approximations for each step. This is also the method we want to use since it inherits more structure of the problem and allows finally for different discretization techniques. For this purpose we introduce the state $x \stackrel{\text{def}}{=} (n, p, V)$ and an admissible set of controls $\mathcal{C} \subset H^1(\Omega)$ and rewrite the state equations (1.2) shortly as $e(x, C) = 0$. Now one defines function spaces $X \stackrel{\text{def}}{=} x_D + X_0$, where $x_D \stackrel{\text{def}}{=} (n_D, p_D, V_D)$ denotes the boundary data introduced in (1.2) and $X_0 \stackrel{\text{def}}{=} \left(H^1_{0,\Gamma_D}(\Omega) \cap L^\infty(\Omega) \right)^3$, where we define

$$H^1_{0,\Gamma_D}(\Omega) \stackrel{\text{def}}{=} \left\{ \phi \in H^1(\Omega) : \phi|_{\Gamma_D} = 0 \right\},$$

as well as $Z \stackrel{\text{def}}{=} [H^1(\Omega)]^3$. Then, $e : X \times H^1(\Omega) \to Z^*$ is well–defined. These preliminaries allow for the exact mathematical setting of our minimization problem

$$\min_{X \times \mathcal{C}} Q(n, p, V, C) \quad \text{such that} \quad e(n, p, V, C) = 0. \tag{1.8}$$

Remark 1.3. This concept of constrained optimization is very general, since it allows to study various cost functionals and all forthcoming techniques can also be adopted to other constraints, e.g. different semiconductor models.

In [15] the existence of a minimizer is proved under mild assumptions on the cost functional Q.

Theorem 1.4. *The constrained minimization problem* (1.8) *admits at least one solution* $(n^*, p^*, V^*, C^*) \in X \times \mathcal{C}$.

In general, we cannot expect the uniqueness of the minimizer since the optimization problem is non–convex due to the nonlinear constraint. Further, already the state system might admit multiple solutions. In the next section we present analytical and numerical results showing that the optimization problem has in fact multiple solutions for some special devices.

2. Uniqueness and Non–uniqueness

The question of uniqueness or possible non–uniqueness of the minimizer is so far not studied in the existing literature. This question is challenging since already the state system may admit for multiple solutions [25] and also for the adjoint system uniqueness could be only established for small current

flows [15]. In the following we present analytical and numerical results that show that there exist indeed multiple solutions to the above minimization problem at least in some cases.

2.1. Multiple solutions in the symmetric case

In this section we consider the optimal design problem for a one dimensional symmetric np–diode, i.e. we assume that $\bar{C}(x) = -\bar{C}(1-x)$. Then the following non–uniqueness result holds.

Theorem 2.1. *There exist at least two minimizers $C_i \in H^1(\Omega)$, $i = 1, 2$, to the optimal design problem (1.8) with $Q_1(n, p, V) = |J - \bar{J}|^2$. Especially, let $(n_1, p_1, V_1, C_1) \in [H^1(\Omega)]^3 \times \mathcal{C}$ be a minimizer. Then, there exists a second minimizer $(n_2, p_2, V_2, C_2) \in [H^1(\Omega)]^3 \times \mathcal{C}$ given by*

$$C_2(x) = -C_1(1-x), \quad n_2(x) = p_1(1-x) \tag{2.1a}$$

$$p_2(x) = n_1(1-x), \quad V_2(x) = -V_1(1-x) + U. \tag{2.1b}$$

Furthermore, it holds $J_{n2} = J_{p1}$, $J_{p2} = J_{n1}$ and $Q(n_1, p_1, V_1, C_1) = Q(n_2, p_2, V_2, C_2)$.

Proof. Let $(n_1, p_1, V_1, C_1) \in [H^1(\Omega)]^3 \times \mathcal{C}$ be a minimizer. First, we show that (n_2, p_2, V_2, C_2) defined by (2.1) fulfills the state system. We define the new variable $\xi = 1 - x$ and calculate

$$J_{n2} = \partial_x n_2(x) + n_2(x)\partial_x V_2(x) = -\partial_\xi p_1(\xi) + p_1(\xi)\,\partial_\xi V_1(\xi) = J_{p1},$$

$$J_{p2} = -\partial_x p_2(x) + p_2(x)\partial_x V_2(x) = \partial_\xi n_1(\xi) + n_1(\xi)\,\partial_\xi V_1(\xi) = J_{n1}.$$

Hence, it holds $\partial_x J_{n2} = \partial_x J_{p2} = 0$. Further, we have

$$-\lambda^2 \partial_{xx} V_2(x) = \lambda^2 \partial_{\xi\xi} V_1(\xi)$$
$$= -n_1(\xi) + p_1(\xi) + C_1(1-\xi)$$
$$= n_2(x) - p_2(x) - C_2(x).$$

One easily verifies that (n_2, p_2, V_2, C_2) also satisfies the boundary conditions such that it is indeed a solution of the state system and the total currents fulfill $J_1 = J_2$.

Finally, we note that

$$\int_\Omega \left|\partial_x(\bar{C}(x) - C_2(x))\right|^2 \ dx = \int_\Omega \left|\partial_x(\bar{C}(1-x) - C_1(1-x))\right|^2 \ dx$$
$$= \int_\Omega \left|\partial_x(\bar{C}(x) - C_1(x))\right|^2 \ dx,$$

which finishes the proof due to $Q(n_1, p_1, V_1, C_1) = Q(n_2, p_2, V_2, C_2)$, i.e. (n_2, p_2, V_2, C_2) is indeed a second minimizer. $\qquad\square$

From the proof we learn that the main reason for the multiplicity of optimal designs is the prescription of the total current density J, which allows for an interchange of roles of the electron and hole current density. This can be clearly seen in the following numerical example. In Figure 2.1 we present the two optimal designs for a symmetric np-diode and the corresponding reference doping profile. The computations were performed on a uniform grid with 300 points and the scaled parameters were set to $\lambda^2 = 10^{-3}$, $\delta^2 = 10^{-2}$ and $U = 10$, i.e. ten times the thermal voltage. For the parameter γ we chose 10^{-3}. The state system and the adjoint system were discretized by an exponentially fitted scheme [4, 29]. The first minimizer is computed using a descent algorithm (see Algorithm 1 in Section 4, which is used here with a constant stepsize 1), while the second is given by (2.1). That this is also a stationary point for the descent algorithm can be seen from Figure 2.2, which shows the evolution of the cost functional if the algorithm is initialized with (2.1). The respective electron and hole densities can be found in Figure 2.3 and Figure 2.4.
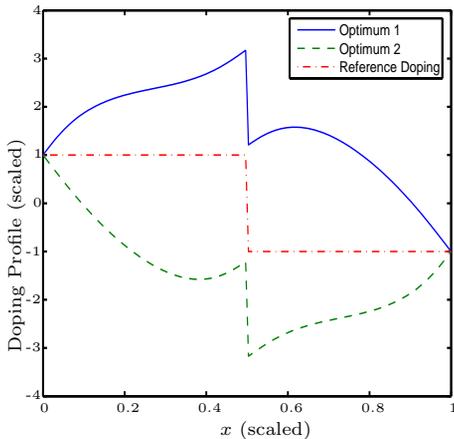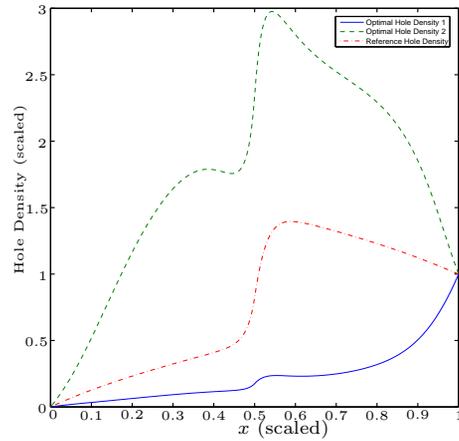

Figure 2.1. Optimized doping profiles.
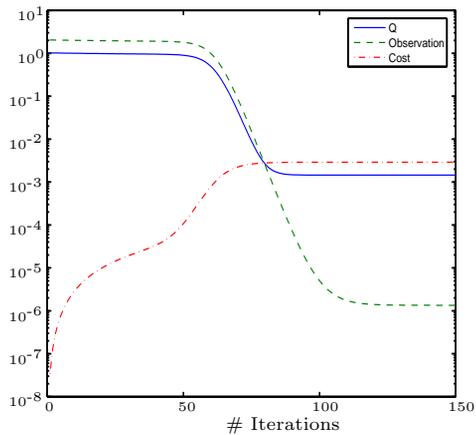
Figure 2.2. Evolution of the cost functional.
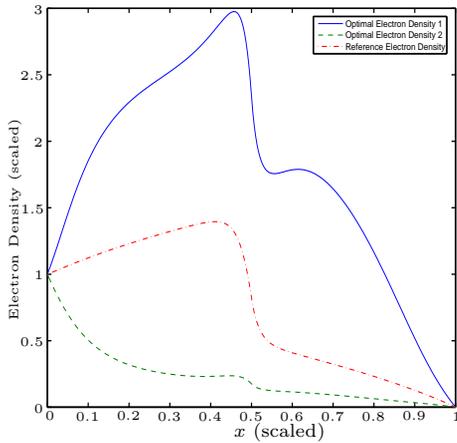
**Figure 2.3.** Optimized electron densities.

Figure 2.4. Optimized hole densities.

## 2.2. Numerical study of the unsymmetric case

Now, we study the case of an unsymmetric np–diode numerically. The main reason is that the above proof cannot be directly extended to the unsymmetric case, since it crucially exploits the symmetry of the reference doping profile. The reference doping profile, depicted in Figure 2.5, is also the starting point for the gradient algorithm which will be discussed in Section 4. The parameters are the same as in Section 2.1.

The good performance of the algorithm can be seen from Figure 2.6, where the evolution of the cost functional is depicted. The algorithm terminates with the minimizer C_2 given by the dashed line in Figure 2.5.

Secondly, we initialize the algorithm with $C_0(x) = 2 \cdot \bar{C}(x) - C_2(x)$. Again, the algorithm terminates and the computed solution can be found as the solid line in Figure 2.5. But now, we only know that this is another critical point of the first–order optimality system, which still might be a saddle point. Nevertheless, this example underlines that one has to interpret the numerical results carefully, although the gradient algorithm behaves very robust. Clearly, the convergence properties and also the limit strongly depends on the starting point for the iteration.
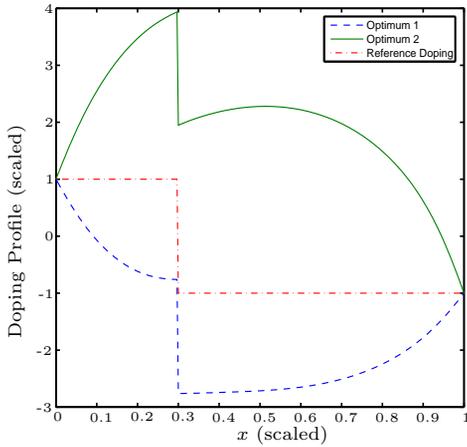
Figure 2.5. Optimized doping profiles.
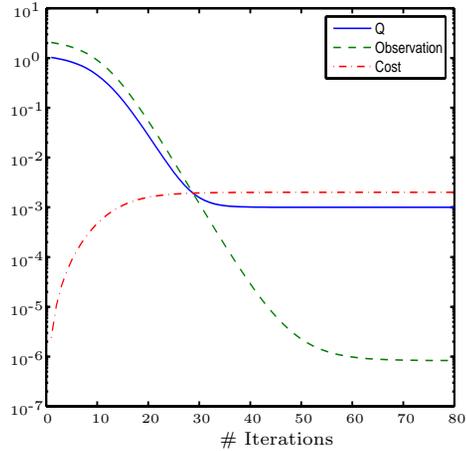
Figure 2.6. Evolution of the cost functional.

3. The First–Order Optimality System

In this section we want to discuss the first–order optimality system which is somehow the basis for all optimization methods seeking at least a stationary point. Since we have a constrained optimization problem, we write the first–oder optimality system using the Lagrangian $\mathcal{L} : X \times \mathcal{C} \times Z \to \mathbb{R}$ associated to problem (1.8) defined by

$$\mathcal{L}(x, C, \xi) \stackrel{\text{def}}{=} Q(x, C) + \langle e(x, C), \xi \rangle_{Z^*, Z},$$

where $\xi \stackrel{\text{def}}{=} (\xi^n, \xi^p, \xi^V)$ denotes the adjoint variable. For the existence of a Lagrange multiplier associated to an optimal solution (x^*, C^*) of (1.8) it is sufficient that the operator $e'(x^*, C^*)$ is surjective. Note the equivalence

$$e'(x, C)[(v, \tilde{C})] = g \quad \text{in } Z^* \quad \Leftrightarrow \quad e_x(x, C)[v] = g - e_C(x, C)[\tilde{C}] \quad \text{in } Z^*.$$

For the DD model this does in general not hold, but one can ensure the bounded invertibility of $e'(x^*, C^*)$ for small current densities [25]. This idea was used in [15] to prove the unique existence of adjoint states.

Theorem 3.1. *There exists a constant $j = j(\Omega, \lambda, U) > 0$ such that for*

each state $x \in X$ with

$$\left\| \frac{J_n^2}{n} \right\|_{L^\infty(\Omega)} + \left\| \frac{J_p^2}{p} \right\|_{L^\infty(\Omega)} \leq j$$

there exists an adjoint state $\xi \in Z$ fulfilling $(e_x(x,C))^*\xi = -Q_x(x,C)$.

Hence, at least for small current densities there exists a unique Lagrange multiplier ξ^* such that together with an optimal solution (x^*, C^*) it fulfills the first–order optimality system

$$\nabla_{(x,C,\xi)}\mathcal{L}(x^*, C^*, \xi^*) = 0. \tag{3.1}$$

In fact one can rewrite this equations in a more concise form [15]:

$$e(x^*, C^*) = 0 \quad \text{in } Z^*,$$
$$e_x^*(x^*, C^*)\xi^* + Q_x(x^*, C^*) = 0 \quad \text{in } X^*,$$
$$e_C(x^*, C^*)\xi^* + Q_C(x^*, C^*) = 0 \quad \text{in } \mathcal{C}^*.$$

I.e., a critical point of the Lagrangian has to satisfy the state system (1.1) with boundary data given in (1.2), as well as the adjoint system

$$\Delta \xi^n - \nabla V \, \nabla \xi^n = \xi^V, \tag{3.2a}$$
$$\Delta \xi^p + \nabla V \, \nabla \xi^p = -\xi^V, \tag{3.2b}$$
$$-\lambda^2 \Delta \xi^V + \text{div}\,(n \, \nabla \xi^n) - \text{div}\,(p \, \nabla \xi^p) = 0, \tag{3.2c}$$

supplemented with appropriate boundary data and the optimality condition

$$\gamma \Delta \left(C - \bar{C} \right) = \xi^V \quad \text{in } \Omega, \tag{3.3a}$$
$$C = \bar{C} \quad \text{on } \Gamma_D, \quad \nabla C \cdot \nu = \nabla \bar{C} \cdot \nu \quad \text{on } \Gamma_N. \tag{3.3b}$$

Remark 3.2. The specific form of the boundary data for the adjoint system (3.2) depends on the special choice of the cost functional Q_1 (cf. [5, 15]).

4. Numerical Methods

The construction of numerical algorithms for the solution of (1.8) mainly relies on the knowledge of derivatives of the cost functional and the con-

straint, since one wants to construct first order or second order convergent algorithms, like gradient descent or Newton-like methods. The differentiability of the cost functional is easy to achieve by its convenient definition, while the differentiability of the state mapping is the content of the following result [15].

Theorem 4.1. *The mapping e is infinitely often Fréchet–differentiable with derivatives vanishing for order greater than* 2.

An adequate and easy to implement numerical method for the solution of (1.8) is the following gradient algorithm.

Algorithm 1.

1. *Choose $C_0 \in \mathcal{C}$.*
2. *For $k = 1, 2, \ldots$ compute $C_k = C_{k-1} - \alpha_k \hat{Q}'(C_{k-1})$.*

Here, $\hat{Q}(C) \stackrel{\text{def}}{=} Q(x(C), C)$ denotes the reduced cost functional, which can be introduced near to the thermal equilibrium state, and $\hat{Q}'(C)$ is the Riesz representative of its first variation. The evaluation of

$$\hat{Q}'(C) = Q_C(x, C) + e_C^* \xi$$

requires the solution of the nonlinear state system for x as well as a solution of the linear adjoint system for ξ and finally a linear solve of a Poisson problem to get the correct Riesz representative.

Remark 4.2. There exist various choices for the parameters α_k ensuring the convergence of this algorithm to a critical point. The overall numerical performance of this algorithm relies on an appropriate choice of the step–size rule for α_k, since these methods require in general consecutive evaluations of the cost functional requiring additional solves of the nonlinear state system [21].

4.1. Numerical examples

In this section we apply Algorithm 1 for the optimal design of an unsymmetric n–p–diode (for the reference doping profile see Figure 4.1). We already learned that the cost functional employed so far might admit multiple minimizers. For this reason we study here a slightly different functional

of the form

$$R(J_n \cdot \nu|_\Gamma, J_p \cdot \nu|_\Gamma) = \frac{1}{2} \left| \int_\Gamma J_n \cdot \nu \, ds - I_n^* \right|^2 + \frac{1}{2} \left| \int_\Gamma J_p \cdot \nu \, ds - I_p^* \right|^2.$$

This allows to adjust the electron and hole current separately. The computations were performed on a uniform grid with 1000 points and the scaled parameters were set to $\lambda^2 = 10^{-3}$, $\delta^2 = 10^{-2}$ and $U = 10$. For the parameter γ we chose $2 \cdot 10^{-2}$. The step–size α_k is computed by an exact one dimensional linesearch

$$\alpha_k = \text{argmin}_\alpha \hat{Q} \left(C_{k-1} - \alpha \hat{Q}'(C_{k-1}) \right).$$

The iteration terminates when the relative error $\left\| \hat{Q}'(C_k) \right\|_{H^1} / \left\| \hat{Q}'(C_0) \right\|_{H^1}$ is less than $5 \cdot 10^{-4}$.
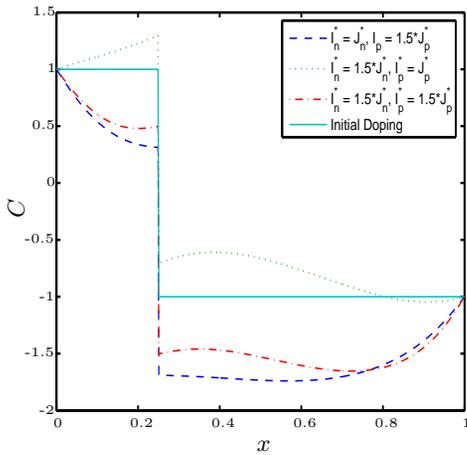

Figure 4.1. Optimized doping profiles.
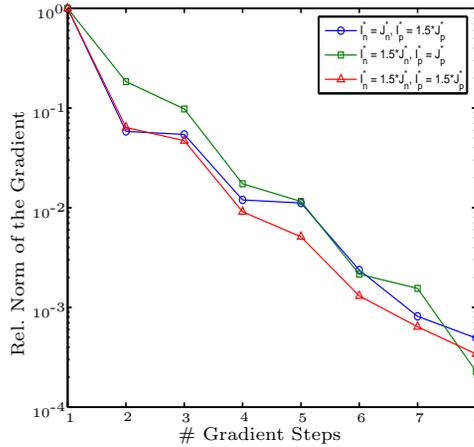
Figure 4.2. Evolution of the relative error.

In Figure 4.1 we present the optimized doping profiles for different choices of I_n^*, I_p^*, i.e. we are seeking an amplification of either the hole current ($I_n^* = J_n^*, I_p^* = 1.5 \cdot J_p^*$) or of the electron current ($I_n^* = 1.5 \cdot J_n^*, I_p^* = J_p^*$) or of both of them $I_n^* = 1.5 \cdot J_n^*, I_p^* = 1.5 \cdot J_p^*$) by 50%. The evolution of the relative error of the gradient can be found in Figure 4.2. Note that the number of gradient steps is independent of the cost functional, which is at a first glance astonishing. A possible explanation is the fixed parameter γ which introduces somehow the same amount of convexity into the three problems.

To get an impression of the overall performance of the method we also have to consider the nonlinear solves needed for the exact one dimensional linesearch. These are presented in Figure 4.3 and one realizes that this is indeed the numerically most expensive part. In Figure 4.4 we present the evolution of the observation $R(J_n \cdot \nu|_\Gamma, J_p \cdot \nu|_\Gamma)$, where one observes that a good approximation of the minimizer is already attained after a few gradient steps, which is characteristic for this method. Further, we note that the third case has the worst match which can on the one hand be explained by the unsymmetry of the device and on the other hand by the fixed choice of γ.
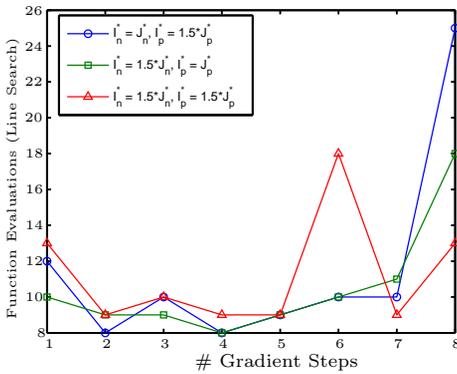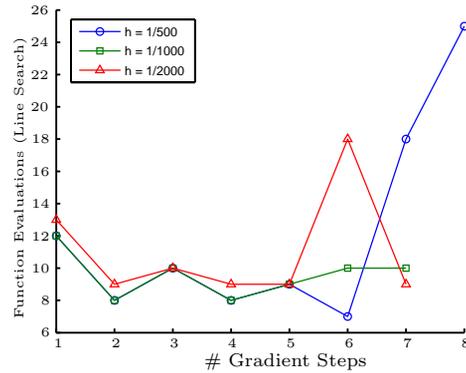

Figure 4.3. Function evaluations for the line search.
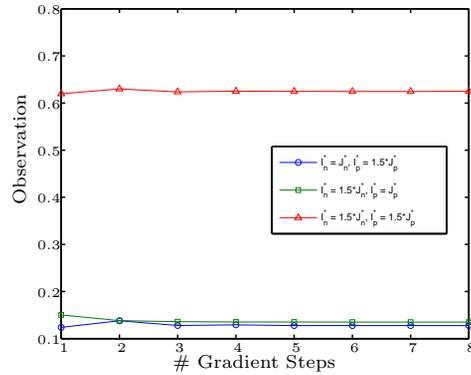
Figure 4.4. Evolution of the observation.
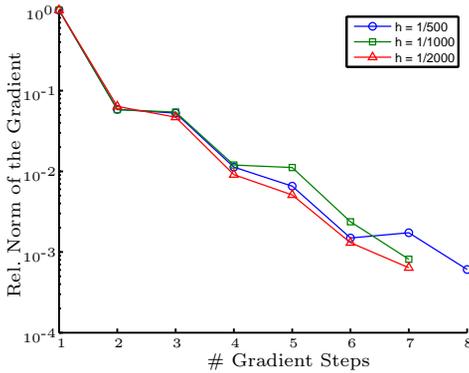
Figure 4.5. Evolution the relative error.

Figure 4.6. Function evaluations for the line search.

The dependence of the algorithm on the grid spacing is studied in Figure 4.5, where we solve the second test case for various grid sizes. One sees that

the number of gradient steps is not affected by the grid size, whereas the function evaluations for the line search are indeed depending on the grid. Note that it is essential that the stopping criterion is related to the discretization error, else it might happen that the iteration does not terminate.

5. Conclusion

In this paper we gave an overview of the mathematical tools which are employed in optimal semiconductor design. Using the DD model as the underlying semiconductor model it was possible to apply techniques from variational calculus to prove existence of optimal designs. There is analytical and numerical evidence that such optimal designs are not unique which should be encountered by the construction of numerical algorithms. The introduction of the adjoint state yields an elegant possibility to formulate numerical algorithms on the continuous level and allows for an adequate consecutive discretization. Gradient–based descent algorithms yield a robust performance. Future work will focus on the development of second–order methods, multi–objective optimization as well as additional constraints on the design variable or the state.

Acknowledgments

References

1. M. Burger, H. W. Engl, and P. Markowich, Inverse doping problems for semiconductor devices, In *Recent Progress in Computational and Applied PDEs*, T. Tang, J. A. Xu, L. A. Ying, T. F. Chan nad Y. Huang, editor, 39-54. Kluwer, 2002.

2. M. Burger, H. W. Engl, P. A. Markowich and P. Pietra, Identification of doping profiles in semiconductor devices, *Inverse Problems*, **17**(2001), 1765-1795.

3. S. Busenberg, W. Fang and K. Ito, Modeling and analysis of laser-beam-induced current images in semiconductors. *SIAM J. Appl. Math.*, **53**(1993), 187-204.

4. M. Burger and R. Pinnau. Exponential fitting for adjoint equations in semiconductor design, in preparation.

5. M. Burger and R. Pinnau. Fast optimal design for semiconductor devices. *SIAM J. Appl. Math.*, **64**(2003), 108-126.

6. L. Ciampolini. Scanning capacitance microscope imaging and modelling, PhD thesis, ETH Zürich, 2001.

7. G. F. Carey, A. L. Pardhanani, and S. W. Bowa. Advanced numerical methods and software approaches for semiconductor device simulation. *VLSI Design*, **10**(2000), no.4, 391-414.

8. A. C. Diebold, M. R. Kump, J. J. Kopanski and D. G. Seiler, Characterization of two-dimensional dopant profiles: status and review, *J. Vac. Sci. Technol. B*, **14**(1996), 196-201.

9. W. Fang and E. Cumberbatch, Inverse problems for metal oxide semiconductor field-effect transistor contact resistivity. *SIAM J. Appl. Math.*, **52**(1992), 699-709.

10. W. Fang and K. Ito, Identifiability of semiconductor defects from LBIC images, *SIAM J. Appl. Math.*, **52**(1992), 1611-1625.

11. W. Fang and K. Ito, Reconstruction of semiconductor doping profile from laser-beam-induced current image. *SIAM J. Appl. Math.*, **54**(1994), 1067-1082.

12. H. Gajewski, On existence, uniqueness and asymptotic behavior of solutions of the basic equations for carrier transport in semiconductors. *Z. Angew. Math. Mech.*, **65**(1985), no.2, 101-108.

13. H. K. Gummel, A self–consistent iterative scheme for one–dimensional steady state transistor calculations, *IEEE Trans. Elec. Dev.*, ED–**11**(1964), 455-465.

14. M. Hinze and R. Pinnau, Optimal control of the drift diffusion model for semiconductor devices. In *Optimal Control of Complex Structures*, K.-H. Hoffmann, I. Lasiecka, G. Leugering, and J. Sprekels, editors, volume 139 of *ISNM*, 95-106. Birkhäuser, 2001.

15. M. Hinze and R. Pinnau. An optimal control approach to semiconductor design. *Math. Models Methods Appl. Sci.*, **12**(2002), no.1, 89-107.

16. K. Ito and K. Kunisch, Augmented Lagrangian-SQP-methods for nonlinear optimal control problems of tracking type. *SIAM J. Control Optim.*, **34**(1996), 874-891.

17. A. Jüngel, *Quasi–hydrodynamic Semiconductor Equations*, Birkhäuser, PNLDE 41, 2001.

18. T. Kerkhoven, A proof of convergence of Gummel's algorithm for realistic device geometries, *SIAM J. Numer. Anal.*, **23**(1986), no.6, 1121-1137.

19. N. Khalil, J. Faricelli, D. Bell, and S. Selberherr, The extraction of two-dimensional MOS transistor doping via inverse modeling, *IEEE Electron Device Lett.*, **16**(1995), 17-19.

20. N. Khalil, ULSI characterization with technology computer-aided design, PhD thesis, TU Vienna, 1995.

21. D. G. Luenberger, *Linear and Nonlinear Programming*, second edition, Addison-Wesley, Reading, 1989.

22. W. R. Lee, S. Wang and K. L. Teo, An optimization approach to a finite dimensional parameter estimation problem in semiconductor device design. *J. Comput. Phys.*, **156**(1999), 241-256.

23. P. A. Markowich, *The Stationary Semiconductor Device Equations*, first edition, Springer–Verlag, Wien, 1986.

24. M. S. Mock, *Analysis of Mathematical Models of Semiconductor Devices*, first edition, Boole Press, Dublin, 1983.

25. P. A. Markowich, Ch. A. Ringhofer, and Ch. Schmeiser, *Semiconductor Equations*, Springer–Verlag, Wien, first edition, 1990.

26. J. Naumann and M. Wolff, A uniqueness theorem for weak solutions of the stationary semiconductor equations. *Appl. Math. Optim.*, **24**, 223-232.

27. R. Plasun, M. Stockinger, R. Strasser, and S. Selberherr, Simulation based optimization environment and its application to semiconductor devices, In *Proceedings IASTED Intl. Conf. on Applied Modelling and Simulation*, 313-316, 1998.

28. S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Springer, Wien, New York, 1984.

29. D. L. Scharfetter and H. K. Gummel, Large signal analysis of a silicon read diode oscillator, *IEEE Trans. Electr. Dev.*, **15**(1969), 64-77.

30. M. Stockinger, R. Strasser, R. Plasun, A. Wild, and S. Selberherr, A qualitative study on optimized MOSFET doping profiles, In *Proceedings SISPAD 98 Conf.*, 77-80, 1998.

31. M. Stockinger, R. Strasser, R. Plasun, A. Wild, and S. Selberherr, Closed–loop MOSFET doping profile optimization for portable systems, In *Proceedings Intl. Conf. on Modelling and Simulation of Microsystems, Semiconductors and Sensors*, 395-398, 1999.

32. M. Stockinger, Optimization of ultra-low-power CMOS transistors, PhD thesis, TU Vienna, 2000.

33. S. M. Sze, *Physics of Semiconductor Devices*, second edition, Wiley, New York, 1981.

Institut für Numerische Mathematik, Technische Universität Dresden, D–01069 Dresden, Germany.

E-mail: hinze@math.tu-dresden.de

Fachbereich Mathematik, Technische Universität Kaiserslautern, D–67663 Kaiserslautern, Germany.

E-mail: pinnau@mathematik.uni-kl.de