

A Globally Convergent Gummel Map for Optimal Dopant Profiling

Martin Burger

Institut für Numerische und Angewandte Mathematik
Westfälische Wilhelms-Universität Münster
Einsteinstr. 62
D 48149 Münster, Germany
martin.burger@uni-muenster.de

René Pinnau

Fachbereich Mathematik
Technische Universität Kaiserslautern
Erwin Schrödinger Str.
D-67663 Kaiserslautern, Germany
pinnau@mathematik.uni-kl.de

We study a generalized Gummel iteration for the solution of an abstract optimal semiconductor design problem, which covers a wide range of semiconductor models. The algorithm is exploiting the special structure of the KKT system and can be interpreted as a descent algorithm for an appropriately defined cost functional. This allows for a convergence proof which does not need the assumption of small biasing voltages. The algorithm is explicitly stated for the (quantum) drift diffusion model, the energy transport model and the microscopic Schrödinger-Poisson model.

Keywords. Semiconductors, optimal design, dopant profiling, generalized Gummel iteration, convergence, drift diffusion model, energy transport model.

Subject (MSC). 35J50, 49J20, 49K20.

1. Introduction

Optimal semiconductor design gained considerable attention during the last decade *Stea98, St00, HP02a*. The ongoing miniaturization of semiconductor devices and the reduced time for a product cycle require an accelerated development of new designs meeting certain performance criteria. This is an important and challenging task in modern microelectronics and the increasing computing power makes these problems tractable using mathematical optimization based on sophisticated models for the device behavior. The most prominent design variable (and correspondingly the unknown in the associated optimization problems) is the device doping profile, which describes the (charge) density of ion impurities in the device. Typical objectives of the optimization are certain device characteristics, like current-voltage or capacitance-voltage characteristics.

In order to solve such optimal design problems it is important to find suitable objec-

tive functionals to be minimized, so that a reasonable compromise between conflicting design goals (e.g. maximizing current and keeping the doping profile close to the reference state) can be achieved. On the other hand one has to choose an appropriate mathematical model to describe the physical device behavior. Nowadays, there is a whole hierarchy of semiconductor device models available, ranging from the microscopic Schrödinger–Poisson model to the classical (quantum) energy transport or the classical (quantum) drift diffusion model ^{MRS90,Jun01}. These are (more or less) well understood from the analytical point of view and can build the basis for automated optimization software. Since the classical drift diffusion model is most common in modern simulation tools, it is not astonishing that most mathematical optimization approaches are based on this model ^{HP02a,HP02b,BP03,HP05,HP06}. Recently, also the classical energy transport model was used for the solution of semiconductor design problems ^{DP08,AnDr04,Dra07}.

Naturally, one has finally to choose an appropriate optimization algorithm for the solution of the respective design problem. In the engineering community, black-box algorithms are most frequently used, since they only require a simulation tool for the forward problem ^{Stea98,St00}. More sophisticated optimization strategies are based on the mathematical structure of the first-order optimality system. Using the adjoint variables one can construct gradient descent ^{HP02a,HP02b} or Newton-type algorithms ^{HP05,HP06}. Clearly, such an approach can be used for each of the above mentioned semiconductor models, but it requires the derivation and implementation of adjoint systems.

Although the (formal) application of the variational calculus to a specific semiconductor model is straight forward, one should also consider the special structure of the underlying partial differential system. This structure is typically exploited in the construction of numerical methods, e.g. Gummel-type iterations, for the forward simulation ^{Gum64,Ker86}. Hence, the construction of a well-suited optimization algorithm should follow these guidelines. This observation was first made in ^{BP03}, where a special, generalized Gummel iteration for the solution of the optimization problem based on the bipolar drift diffusion model was suggested and tested.

In this paper we will apply this kind of Gummel iteration to general semiconductor models fulfilling some specific structural requirements. This allows us to prove the convergence of the generalized Gummel iteration in an abstract setting.

1.1. Models for Optimal Dopant Profiling

First, we discuss the general optimization problem. Semiconductor device models are usually composed of two basic ingredients, namely the electric potential V and a set of field variables $\rho = (\rho_1, \dots, \rho_n)$ (e.g., the electron and hole densities and/or their respective temperatures), which together satisfy a nonlinear system of the

form

$$\lambda^2 \Delta V = Q(\rho, V) - C, \quad (1.1a)$$

$$F(\rho, V) = 0. \quad (1.1b)$$

Here, $\lambda > 0$ denotes the so-called scaled *Debye length*, $Q(\rho, V)$ is the total charge density generated by ρ and V , C is the doping profile (modeled as a function of space) and F symbolizes nonlinear (partial) differential equations for ρ . All equations are to be solved in a regular domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2$ or 3 , modeling the device geometry. Moreover, suitable boundary conditions need to be specified, which we do not further discuss here. For an excellent overview of semiconductor device models and their asymptotic relations we refer to ^{JP01, MRS90, Sel84}. Typical models fitting in the above regime are the macroscopic (quantum) drift diffusion model and the (quantum) energy transport model, as well as the microscopic Schrödinger–Poisson system.

A typical design goal is the increase of the on-state current. Hence, the objective of the optimization can be modeled as a functional of the densities and the voltage, i.e.,

$$R(\rho, V) \rightarrow \min_{(\rho, V, C) \text{ satisfying (1.1)}}. \quad (1.2)$$

The functional R could, e.g., be the negative current outflow on a contact $\Gamma \subset \partial\Omega$,

$$R(\rho, V) = - \int_{\Gamma} J \cdot \nu \, ds,$$

in order to maximize the current (cf. ^{PSS98, St00, Stea98}) or the square of current minus a target current (cf. ^{HP02a, HP02b}), i.e.,

$$R(\rho, V) = \int_{\Gamma} |(J - J^*) \cdot \nu|^2 \, ds.$$

Assuming the solvability of the state system (1.1) for a given doping profile C , the natural design variable is the doping profile C itself. Since the minimization problem (1.2) is in general not well-posed, one considers instead a regularized version

$$R(\rho, V) + \frac{\alpha}{2} \|C - C^*\|^2 \rightarrow \min_{(\rho, V, C) \text{ satisfying (1.1)}},$$

where C^* is a given reference doping profile (cf. ^{HP02a, HP02b}) and $\alpha > 0$ a parameter which allows for the penalization of deviations from C^* . This is also necessary from the engineering point of view, since the basic device structure should not change during the optimization process.

Exploiting carefully the structure of the state system (1.1) and observing that typical objectives only depend on (ρ, V) , one can also use the total charge in the device $W := Q(\rho, V) - C$ as a design variable. This approach has been first introduced in ^{BP03} and it turns out that the corresponding optimality system has a significantly simpler structure than the ones derived in ^{HP02a, HP02b}.

Note that, if one minimizes with respect to (ρ, V, W) , then one can reconstruct the doping profile C uniquely from the formula $C = Q(\rho, V) - W$. Moreover, the Poisson equation simplifies to

$$\lambda^2 \Delta V = W, \quad (1.3)$$

i.e., we have no self-consistent coupling in the new variables. Hence, for given W , we can solve (1.3) and (1.1b) consecutively.

Remark 1.1. We point out that the self-consistent coupling with the Poisson equation for the electrostatic potential V , in general introduces severe difficulties in the analysis of semiconductor models ^{MRS90,Jun01}. This also holds for the analysis of the optimality system for the minimization problem (1.2) (cf. ^{HP02a,HP02b}).

Again, the primary optimization goal can be modeled as a functional of the field variables and the potential, i.e.,

$$H(\rho, V, W) := R(\rho, V) + \frac{\beta}{2} \|W - W^*\|^2 \rightarrow \min_{(\rho, V, W) \text{ satisfying (1.1b), (1.3)}}, \quad (1.4)$$

where W^* is a given total charge, describing the underlying device, and $\beta > 0$ allows to adjust the penalization.

In ^{BP03} a generalized Gummel iteration for the solution of (1.4) is suggested, where the operator F describes the stationary, bipolar drift diffusion model. This iteration proves to be numerically very efficient, in fact the effort for the solution of the optimization problem is just twice as much as the one for the solution of the state system.

In the following we provide a convergence proof for this generalized Gummel iteration in an abstract setting, which is independent of the specific semiconductor model. The proof is based on weak assumptions on the solvability of (1.1b) and heavily relies on the observation that the iteration yields in fact a descent algorithm for the reduced cost functional $\hat{H}(W) := H(\rho(W), V(W), W)$. This allows to prove the global convergence of the generalized Gummel iteration for arbitrarily applied biasing voltages. This is most astonishing, since the convergence of the Gummel iteration for the drift diffusion model itself can be only proved near to the thermal equilibrium state ^{Gum64,Ker86}. Here, the special structure of the optimality system gives the additional information which is needed.

The paper is organized as follows. In Section 2 we introduce the abstract setting of the minimization problem (1.4) and prove its solvability. The generalized Gummel iteration is introduced in Section 3, where also its global convergence is proved. Its application to some specific semiconductor models, like (quantum) drift diffusion, energy transport and Schrödinger–Poisson, are discussed in Section 4. Finally, we give concluding remarks in Section 5.

2. Abstract Optimal Dopant Profiling

In this section we provide the abstract mathematical setting for the design problem (1.4) and prove it solvability. For this purpose let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{J}$ denote Banach spaces and assume that $\tilde{\mathcal{Y}}$ and \mathcal{U} are Hilbert spaces with $\mathcal{Y} \subset \tilde{\mathcal{Y}} \subset \mathcal{U}$. We shall discuss duality based on \mathcal{U} , i.e., identify \mathcal{U}^* with its dual space. In the following, \mathcal{X} is the space of field variables ρ and we have $V \in \mathcal{Y}$. Further, let the *field operator* $F : \mathcal{X} \times \tilde{\mathcal{Y}} \rightarrow \mathcal{Z}^*$ be given. Restricting the designs, i.e., $W \in \mathcal{U}$, we first need to ensure that the Poisson equation (1.3) is uniquely solvable. To this end we assume

A.1 The Laplacian Δ is in $\mathcal{L}(\tilde{\mathcal{Y}}, \tilde{\mathcal{Y}}^*)$ and has a bounded inverse, which is compact as an operator from \mathcal{U} to \mathcal{Y} .

Further, to handle (1.1b) we assume

A.2 The operator $F : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}^*$ is continuously Fréchet-differentiable and $\left(\frac{\partial F}{\partial \rho}\right)^{-1}$ is regular. Moreover, $\frac{\partial F}{\partial V}(\rho, V)^*$ is a bounded linear operator from \mathcal{Z} to $\tilde{\mathcal{Y}}^*$ for (ρ, V) satisfying $F(\rho, V) = 0$.

Remark 2.1. Assumption **A.1** can be easily fulfilled by the appropriate choice of the spaces $\mathcal{U}, \tilde{\mathcal{Y}}$, and \mathcal{Y} , e.g., $\mathcal{U} = L^2(\Omega)$ and $\tilde{\mathcal{Y}} = H^1(\Omega)$ and some mild requirements on the smoothness of the boundary of Ω , e.g., $\partial\Omega \in C^{0,1}$ *Tro87*. A requirement for the space \mathcal{Y} is the possibility to handle the nonlinearity in the continuity equations, typically $\mathcal{Y} \subset L^\infty(\Omega) \cap H^1(\Omega)$. In spatial dimension one the choice $\mathcal{Y} = H^1(\Omega)$ would then still be possible due to the compact embedding, while in higher dimensions a regularity assumption for the Laplacian (and the corresponding boundary values) is needed to fulfill the assumptions for a space like $\mathcal{Y} = H^2(\Omega)$ compactly embedded in $L^\infty(\Omega)$. Note that Assumption **A.2** ensures in combination with the implicit function theorem that, for given $V \in \mathcal{Y}$, we can solve $F(\rho, V)$ locally for $\rho \in \mathcal{X}$. The differentiability of F is necessary for the derivation of the first-order optimality system. Moreover, the mapping property of the adjoint is necessary for the well-definedness of a Lagrange multiplier (adjoint state), which is determined from a Poisson equation with right-hand side involving $\frac{\partial F}{\partial V}(\rho, V)^*$ as we shall see below. Since the mapping property is only needed on the zero set of F , this again allows to use regularity of ρ and V . This issue will become clear in the application to specific models, in particular for the case of the Poisson-drift-diffusion model.

For the definition of the abstract optimization goal we define the *current-operator* $I : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{J}$, $J = I(\rho, V)$ and the *observation-operator* $R : \mathcal{J} \rightarrow \mathbb{R}$, $R = R(J)$. Again, we need some smoothness and assume

A.3 The operators $I : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{J}$ and $R : \mathcal{J} \rightarrow \mathbb{R}$ are continuously Fréchet-differentiable.

Remark 2.2. The functional R could, e.g., be the negative current outflow on a contact (in order to maximize the current, cf. *PSS98,St00,Stea98*) or the square of

current minus a target current (cf. *HP02a,HP02b*).

The solvability of the minimization problem

$$\min_{(\rho,V,W) \in (\mathcal{X} \times \mathcal{Y}) \times \mathcal{U}} H(\rho, V, W) := R(I(\rho, V)) + \frac{\beta}{2} \|W - W^*\|_{\mathcal{U}}^2, \text{ s.t. (1.1b), (1.3) holds,} \quad (2.1)$$

is the content of the following result.

Theorem 2.1. *Assume (A.1)–(A.3). Additionally, let $F : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}^*$ be continuous. If H is weakly lower semi-continuous and bounded from below, then the minimization problem (2.1) has a solution $(\rho_0, V_0, W_0) \in (\mathcal{X} \times \mathcal{Y}) \times \mathcal{U}$.*

Proof. For the proof, we introduce the reduced cost functional

$$\hat{H}(W) := H(\rho(V(W)), V(W), W),$$

where $\rho(V) : \mathcal{Y} \rightarrow \mathcal{X}$ is a well-defined and continuous solution operator due to Assumption (A.2) in combination with the implicit function theorem. Further, $V(W) : \mathcal{U} \rightarrow \mathcal{Y}$ is well-defined and continuous on account of Assumption (A.1). Now, the minimization problem (2.1) is equivalent to

$$\min_{W \in \mathcal{U}} \hat{H}(W) \quad (2.2)$$

We choose a minimizing sequence $(W^k)_{k \in \mathbb{N}} \subset \mathcal{U}$ with

$$\lim_{k \rightarrow \infty} \hat{H}(W^k) = \inf_{W \in \mathcal{U}} \hat{H}(W) =: H_0.$$

The boundedness of H yields $H_0 > -\infty$. By definition, \hat{H} is coercive with respect to W , which yields the boundedness of $\|W^k\|_{\mathcal{U}}$. Since \mathcal{U} is a Hilbert space, we can extract a weakly convergent subsequence, again denoted by $(W^k)_{k \in \mathbb{N}}$, such that $W^k \rightharpoonup W_0 \in \mathcal{U}$ as $k \rightarrow \infty$. Due to Assumption (A.1) there exists a corresponding (sub)sequence $(V^k)_{k \in \mathbb{N}} := (V(W^k))_{k \in \mathbb{N}} \subset \mathcal{Y}$ which is strongly converging to $V_0 = V(W_0)$. Now, the continuity of the solution operator $\rho(V) : \mathcal{Y} \rightarrow \mathcal{X}$ gives a strongly converging sequence $(\rho^k)_{k \in \mathbb{N}} := (\rho(V^k))_{k \in \mathbb{N}}$ with $\lim_{k \rightarrow \infty} \rho^k = \rho(V_0) =: \rho_0$. The continuity of F implies now

$$F(\rho^k, V^k) \rightarrow F(\rho_0, V_0) \quad \text{for } k \rightarrow \infty.$$

Finally, we use the weak lower semi-continuity of \hat{H} which gives

$$H_0 \leq \hat{H}(W_0) \leq \liminf_{k \rightarrow \infty} \hat{H}(W^k) = H_0.$$

Hence, (ρ_0, V_0, W_0) is in fact the desired minimizer of (2.1). \square

Remark 2.3. The solvability of the optimization problem for certain, specific semiconductor models has been already established for the drift diffusion as well as for the energy transport model, compare *BP03,Dra07*.

2.1. The First-Order Optimality System

The first-order necessary condition for a minimizer of (2.1) is derived using the Lagrangian $L : (\mathcal{X} \times \mathcal{Y}) \times \mathcal{U} \times (\mathcal{Y} \times \mathcal{Z}) \rightarrow \mathbb{R}$ associated to (2.1), which is given by

$$L(\rho, V, W; p, q) = H(\rho, V, W) + \langle -\lambda^2 \Delta V + W, p \rangle_{\tilde{\mathcal{Y}}^*, \tilde{\mathcal{Y}}} + \langle F(\rho, V), q \rangle_{\mathcal{Z}^*, \mathcal{Z}}.$$

It is straight forward to prove the following representation of the first-order optimality condition $L' = 0$ for a minimizer.

Lemma 2.1. *Assume (A.1)–(A.3). Then, the first-order optimality condition $L' = 0$ can be written in the more concise form*

$$0 = \frac{\partial L}{\partial p} = \lambda^2 \Delta V - W \quad \text{in } \tilde{\mathcal{Y}}^*, \quad (2.3a)$$

$$0 = \frac{\partial L}{\partial q} = F(\rho, V) \quad \text{in } \mathcal{Z}^*, \quad (2.3b)$$

$$0 = \frac{\partial L}{\partial \rho} = R'(I) \frac{\partial I}{\partial \rho} + \left(\frac{\partial F}{\partial \rho} \right)^* q \quad \text{in } \mathcal{X}^*, \quad (2.3c)$$

$$0 = \frac{\partial L}{\partial V} = R'(I) \frac{\partial I}{\partial V} - \lambda^2 \Delta p + \left(\frac{\partial F}{\partial V} \right)^* q \quad \text{in } \tilde{\mathcal{Y}}^*, \quad (2.3d)$$

$$0 = \frac{\partial L}{\partial W} = \beta(W - W^*) + p \quad \text{in } \mathcal{U}. \quad (2.3e)$$

Remark 2.4. The structure of the optimality system for (2.1) turns out to be very convenient for optimization purposes. For a given design variable $W \in \mathcal{U}$ one can subsequently solve the Poisson equation (2.3a) for $V \in \mathcal{Y}$, followed by the equation (2.3b) for the field variables $\rho \in \mathcal{X}$. Finally, one can consecutively solve the adjoint equations (2.3c) and (2.3d) for $q \in \mathcal{Z}$ and $p \in \tilde{\mathcal{Y}} \subset \mathcal{U}$. The solvability of (2.3a) and (2.3c) is a consequence of Assumption (A.2), while (2.3a) and (2.3d) can be solved (with the additional regularity $V \in \mathcal{Y}$) due to Assumption (A.1)

3. The Generalized Gummel Iteration

In this Section we propose an iterative algorithm for the computation of a minimizer of (2.1). The algorithm is just exploiting of the first-order optimality system (2.3) in the spirit of the well-known Gummel iteration for the solution of the classical drift diffusion model. There, the idea is a decoupling of Poisson's equation for the electrostatic potential V from the equations for the charge densities ρ . We extend this idea by incorporating now also the design variable and the adjoint variables into this iterative procedure. This results in the following *generalized Gummel iteration*:

Algorithm 1.

- (1) Choose an admissible $W^0 \in \mathcal{U}$.
- (2) For $k = 0, 1, 2, \dots$ solve consecutively
 - (a) $\lambda^2 \Delta V^k = W^k$ for $V^k \in \mathcal{Y}$
 - (b) $F(\rho^k, V^k) = 0$ for $\rho^k \in \mathcal{X}$
 - (c) $R'(I(\rho^k, V^k)) \frac{\partial I(\rho^k, V^k)}{\partial \rho} + \left(\frac{\partial F(\rho^k, V^k)}{\partial \rho} \right)^* q^k = 0$ for $q^k \in \mathcal{Z}$
 - (d) $R'(I(\rho^k, V^k)) \frac{\partial I(\rho^k, V^k)}{\partial V} + \left(\frac{\partial F(\rho^k, V^k)}{\partial V} \right)^* q^k - \lambda^2 \Delta p^k = 0$ for $p^k \in \tilde{\mathcal{Y}}$
 - (e) $\beta(W^{k+1} - W^*) + \tau_k(W^{k+1} - W^k) + p^k = 0$ for $W^{k+1} \in \mathcal{U}$

Here, $(\tau_k)_{k \in \mathbb{N}}$ is a sequence of appropriately chosen, positive relaxation parameters, e.g., $\tau_k \equiv \tau > 0$, like in ^{BP03}.

Remark 3.1. This generalized Gummel iteration can be easily integrated in modern semiconductor device simulation tools, since it is based on the algorithm which is most commonly implemented in these tools. For this reason we do not employ sophisticated optimization algorithms like SQP or Newton-type models, which are based on second order information. These would again result in fully coupled linear systems (see also the discussion in ^{BP03}).

Remark 3.2. We want to mention that the approach originally introduced in ^{BP03} was using a slightly different damping. In this setup $p = -\beta(W - W^*)$ was eliminated from the optimality condition before the iteration, so that step (e) was missing and step (d) became an effective update for W , where the damping term $\tau_k(W^{k+1} - W^k)$ was added. This approach can also be written equivalently in the above form by changing (e) to

$$(e') \quad \beta(W^{k+1} - W^*) - \tau_k \Delta^{-1}(W^{k+1} - W^k) + p^k = 0 \text{ for } W^{k+1} \in \mathcal{U}.$$

The convergence proof for the classical Gummel iteration is based on Banach's fixed point theorem ^{Ker86}. There the contractivity of the fixed-point mapping is ensured by the assumption of small biasing voltages, i.e., only near to the thermal equilibrium state. Here we follow a different idea and exploit in detail the structure of the minimization problem (2.1). Especially, we show that the generalized Gummel iteration results in a descent algorithm for the solution of (2.1). This additional information allows for a proof of convergence without any restriction on the applied voltage, which is most astonishing, since we also solve the forward problem during our iterative procedure.

It is most convenient to consider the reduced cost functional

$$\hat{H}(W) \stackrel{\text{def}}{=} R(I(S(W))) + \frac{\beta}{2} \|W - W^*\|_{\mathcal{U}}^2, \quad (3.1)$$

where $S : \mathcal{U} \rightarrow \mathcal{X} \times \mathcal{Y}$, $S(W) = (\rho(W), V(W))$, is the solution operator given by

$$\lambda^2 \Delta V(W) = W, \quad F(\rho(W), V(W)) = 0.$$

Note that S is well-defined on account of Assumption **(A.1)** and **(A.2)**.
 First, we derive a representation of the derivative of the reduced cost functional \hat{H} .

Lemma 3.1. *Assume **(A.1)**–**(A.3)**. Then, \hat{H} is continuously Frechet-differentiable and the derivative is given in the sensitivity form*

$$\hat{H}'(W) = R'(I) \frac{\partial I}{\partial S} \frac{\partial S}{\partial W} + \beta(W - W^*) \in \mathcal{U}^*$$

or in adjoint form

$$\hat{H}'(W) = \left(\frac{\partial S}{\partial W} \right)^* \left(\frac{\partial I}{\partial S} \right)^* R'(I) + \beta(W - W^*) \in \mathcal{U}^*.$$

In infinite-dimensional optimization it is most common to apply a descent algorithm for the minimization of the reduced cost functional (3.1), which only uses first-order derivative information (see also ^{Kel95}). The convergence of such a method crucially relies on the appropriate choice of the descent direction, where clearly the negative gradient is most popular. In the following we will show that the direction $d^k = W^{k+1} - W^k$ provided by the generalized Gummel iteration is in fact the gradient direction.

Theorem 3.1. *Assume **(A.1)**–**(A.3)**. For the derivative of $\hat{H}(W)$ it holds*

$$\hat{H}'(W^k) = -(\beta + \tau_k)d^k.$$

Proof. Let $\theta^k \in \mathcal{U}$ be arbitrarily given. We calculate

$$\begin{aligned} \frac{\partial R(I(S))(W^k)}{\partial W}[\theta^k] &= \left\langle \left(\frac{\partial S}{\partial W} \right)^* \left(\frac{\partial I}{\partial S} \right)^* R'(I^k), \theta^k \right\rangle_{\mathcal{U}^*, \mathcal{U}} \\ &= \left\langle \left(\frac{\partial I}{\partial S} \right)^* R'(I^k), \frac{\partial S}{\partial W}[\theta^k] \right\rangle_{(\mathcal{X} \times \mathcal{Y})^*, \mathcal{X} \times \mathcal{Y}} \\ &= \left\langle \left(\frac{\partial I}{\partial S} \right)^* R'(I^k), (\tilde{\rho}^k, \tilde{V}^k) \right\rangle_{(\mathcal{X} \times \mathcal{Y})^*, \mathcal{X} \times \mathcal{Y}} \end{aligned}$$

$$\begin{aligned}
 \text{where } \lambda^2 \Delta \tilde{V}^k &= \theta^k \text{ and } \tilde{\rho}^k = - \left(\frac{\partial F}{\partial \rho} \right)^{-1} \frac{\partial F}{\partial V} [\tilde{V}^k], \\
 &= \left\langle R'(I^k), \frac{\partial I}{\partial S} [(\tilde{\rho}^k, \tilde{V}^k)] \right\rangle_{\mathcal{J}^*, \mathcal{J}} \\
 &= \left\langle R'(I^k), \frac{\partial I}{\partial \rho} [\tilde{\rho}^k] + \frac{\partial I}{\partial V} [\tilde{V}^k] \right\rangle_{\mathcal{J}^*, \mathcal{J}} \\
 &= - \left\langle \left(\frac{\partial F}{\partial \rho} \right)^* [q^k], \tilde{\rho}^k \right\rangle_{\mathcal{X}^*, \mathcal{X}} - \left\langle \left(\frac{\partial F}{\partial V} \right)^* [q^k], \tilde{V}^k \right\rangle_{\tilde{\mathcal{Y}}^*, \tilde{\mathcal{Y}}} + \left\langle \lambda^2 \Delta p^k, \tilde{V}^k \right\rangle_{\tilde{\mathcal{Y}}^*, \tilde{\mathcal{Y}}} \\
 &= - \overbrace{\left\langle \frac{\partial F}{\partial \rho} [\tilde{\rho}^k] + \frac{\partial F}{\partial V} [\tilde{V}^k], q^k \right\rangle}_{=0}^{\mathcal{Z}^*, \mathcal{Z}} + \left\langle p^k, \lambda^2 \Delta \tilde{V}^k \right\rangle_{\tilde{\mathcal{Y}}^*, \tilde{\mathcal{Y}}} \\
 &= \left\langle -\beta(W^{k+1} - W^*) - \tau_k(W^{k+1} - W^k), \lambda^2 \tilde{\Delta} V^k \right\rangle_{\tilde{\mathcal{Y}}^*, \tilde{\mathcal{Y}}}
 \end{aligned}$$

Hence, we get

$$\begin{aligned}
 \hat{H}'(W^k)[\theta^k] &= \langle -\beta(W^{k+1} - W^*) - \tau_k(W^{k+1} - W^k), \theta^k \rangle_{\mathcal{U}, \mathcal{U}^*} + \beta \langle W^k - W^*, \theta^k \rangle_{\mathcal{U}^*, \mathcal{U}} \\
 &= -(\beta + \tau_k) \langle d^k, \theta^k \rangle_{\mathcal{U}^*, \mathcal{U}}
 \end{aligned}$$

and, since this holds for all $\theta^k \in \mathcal{U}$, we can identify

$$\hat{H}'(W^k) = -(\beta + \tau_k) d^k. \quad \square$$

Remark 3.3. The above proof can be adapted to the damping via (e') corresponding to the original setup in ^{BP03}. In this case the update is not the gradient of the reduced functional, but one still obtains a descent step with analogous properties, which suffices also for the results shown below.

Now, we are in the position to prove the unconditional convergence of the generalized Gummel iteration, which can in fact be interpreted as gradient descent algorithm for \hat{H} with step size $(\beta + \tau_k)^{-1}$.

Theorem 3.2. *Let $(W^k)_{k \in \mathbb{N}}$ be the sequence of design variables generated by Algorithm 1. Assume that \hat{H} is bounded from below and that the sequence of positive regularization parameters $(\tau_k)_{k \in \mathbb{N}}$ is admissible, i.e.,*

$$\hat{H}(W^k + d^k) - \hat{H}(W^k) \rightarrow 0 \quad \Rightarrow \quad (\beta + \tau_k) \|d^k\|_{\mathcal{U}} \rightarrow 0, \quad (3.2)$$

for $k \rightarrow \infty$. Then, we have

$$\lim_{k \rightarrow \infty} \hat{H}'(W^k) = 0$$

and each accumulation point of $(W^k)_{k \in \mathbb{N}}$ is a stationary point.

Proof. Let \underline{H} be the lower bound of $\hat{H}(W^k)$. Then it holds

$$\hat{H}(W^0) - \underline{H} = \sum_{k=0}^{\infty} \left(\hat{H}(W^k) - \hat{H}(W^{k+1}) \right) = \sum_{k=0}^{\infty} \left| \hat{H}(W^k + d^k) - \hat{H}(W^k) \right|$$

which yields $\hat{H}(W^k + d^k) - \hat{H}(W^k) \rightarrow 0$ for $k \rightarrow \infty$. Now the admissibility of τ_k implies

$$0 = \lim_{k \rightarrow \infty} (\beta + \tau_k) \|d^k\|_{\mathcal{U}} = \lim_{k \rightarrow \infty} \|H'(W^k)\|_{\mathcal{U}^*}.$$

Now, let W^* be an accumulation point of $(W^k)_{k \in \mathbb{N}}$. Then, there exists a subsequence, again denoted by $(W^k)_{k \in \mathbb{N}}$, such that $\lim_{k \rightarrow \infty} W^k = W^*$. The monotonicity of $\hat{H}(W^k)$ implies $\hat{H}(W^k) \geq \hat{H}(W^*)$ and by continuity we get $\hat{H}'(W^*) = \lim_{k \rightarrow \infty} \hat{H}'(W^k) = 0$. \square

Remark 3.4. The assumption on the boundedness of \hat{H} is in general no restriction, since it can be ensured by the special choice of the observation-operator R . In particular, this assumption is fulfilled for tracking-type operators.

Remark 3.5. The admissibility of the sequence of regularization parameters can be easily translated into the admissibility of the corresponding step sizes for the gradient descent algorithm. In this context it is well-known, that this can be ensured using the Armijo rule ^{Kel95}.

4. Applications of the Generalized Gummel Iteration

In the following sections we apply the generalized Gummel iteration to several specific semiconductor models which are used in modern simulation tools. In particular, we consider the classical drift diffusion and energy transport model, as well as the microscopic Schrödinger–Poisson model and the macroscopic quantum drift diffusion model. This wide range of models allows for the incorporation of many relevant physically phenomena, like non constant mobilities, temperature dependencies or quantum effects.

4.1. Optimization Based on the Drift Diffusion Model

First, we consider the most frequently used *bipolar drift diffusion model*, where $\rho = (u, v)$ are the so-called *Slotboom variables* ^{MRS90, Sel84}. They are related to the electron density n and the hole density p via $n = e^V u$ and $p = e^{-V} v$, respectively. In this case the charge density is simply $Q(\rho, V) = n - p = e^V u - e^{-V} v$ and the differential equations defining the operator F are given by

$$F(u, v, V) = \begin{pmatrix} \operatorname{div}(e^V \nabla u) \\ \operatorname{div}(e^{-V} \nabla v) \end{pmatrix}.$$

To set up the analytical framework for optimal dopant profiling with the drift diffusion model we choose the function spaces are $\mathcal{X} = [H^1(\Omega) \cap L^\infty(\Omega)]^2$, $\mathcal{Y} = H^2(\Omega)$, $\hat{\mathcal{Y}} = H^1(\Omega)$, $\mathcal{Z} = H^1(\Omega)$ and $\mathcal{J} = [L^2(\Omega)]^d$, as well as the space of designs $\mathcal{U} = L^2(\Omega)$. This choice of the function spaces will ensure (under a standard regularity assumption on the Laplacian) Assumptions **(A.1)** and **(A.2)** (see ^{MRS90}). Note that $\frac{\partial F}{\partial \rho}$ is then globally invertible ^{BP03}.

The related current-operator is defined by

$$I = I(u, v, V) = e^V \nabla u - e^{-V} \nabla v,$$

which is well-posed on $\mathcal{X} \times \mathcal{Y}$ with values in \mathcal{J} .

For the observation-operator several choices are possible, depending on the desired design goal. One example is given by

$$R(I) = \left| \int_{\Gamma} I \cdot \nu \, ds - I^* \right|^2, \quad (4.1)$$

where we consider the current over a contact $\Gamma \subset \partial\Omega$ and I^* is the desired target current. Here, ν denotes the outward unit normal vector along the boundary.

This current-operator and the choice of the observation-operator yields that also Assumption **(A.3)** holds and Algorithm 1 results in the iteration:

- (1) Choose an admissible W^0 .
- (2) For $k = 0, 1, 2, \dots$ solve consecutively

$$\begin{aligned} \lambda^2 \Delta V^{k+1} &= W^k \\ \operatorname{div} (e^{V^{k+1}} \nabla u^{k+1}) &= 0 \\ \operatorname{div} (e^{-V^{k+1}} \nabla v^{k+1}) &= 0 \\ \operatorname{div} (e^{V^{k+1}} \nabla r^{k+1}) &= 0 \\ \operatorname{div} (e^{-V^{k+1}} \nabla s^{k+1}) &= 0 \\ \lambda^2 \Delta p^{k+1} &= e^{V^{k+1}} \nabla u^{k+1} \nabla r^{k+1} - e^{-V^{k+1}} \nabla v^{k+1} \nabla s^{k+1} \\ -\beta(W^{k+1} - W^*) - \tau_k(W^{k+1} - W^k) &= p^{k+1}. \end{aligned}$$

Hence, in the classical Gummel iteration ^{Gum64, Ker86} one needs to include two additional solves of elliptic equations for the adjoint states. Those have the same structure as the forward equations, such that they can be easily implemented in some existing simulation tool. This algorithm was already tested numerically and proved its superiority to standard optimization algorithms ^{BP03}.

Remark 4.1. Overall, the generalized Gummel iteration for optimal dopant profiling requires just a numerical effort, which is twice as much as the one for the classical Gummel iteration. This is clearly optimal for this special case! Note that this depends strongly on the specific form of the drift diffusion model written in Slotboom variables. We get the full decoupling of the equations only in this formulation. Usage of the primal variables n and p would in fact result in a coupling of the adjoint system.

4.2. Optimization Based on the Energy Transport Model

As a second example we consider the *unipolar energy transport model* including thermal effects ^{DGJ98, Jum01}. This model was already used successfully for optimal dopant profiling ^{DP08, AnDr04}. The field variables are given by $\rho = (g_1, g_2, T)$, where

g_1 and g_2 are generalized densities and T is the electron temperature. Here, the charge density is simply given by $Q(\rho, V) = n(g_1, g_2)$. The specific relation depends on the special choice of the underlying band structure of the semiconductor $DJP00$. In the sequel we address only the so-called *Chen model* for a parabolic band structure $CCJS93$, then we just have $n(g_1, g_2) = g_1$.

The operator F is defined by

$$F(g_1, g_2, T, V) = \begin{pmatrix} \operatorname{div} I_1 \\ \operatorname{div} I_2 - I_1 \cdot \nabla V - D(g_1, g_2, T, V) \\ T - f(g_1, g_2) \end{pmatrix},$$

where

$$I_i = I_i(g_i, T, V) = \nabla g_i - \frac{g_i}{T} \nabla V, \quad i = 1, 2.$$

Here, D is a relaxation term and for the Chen model we have $f(g_1, g_2) = (2g_2)/(3g_1)$. Appropriate function spaces are $\mathcal{X} = [H^1(\Omega) \cap L^\infty(\Omega)]^3$, $\mathcal{Y} = H^2(\Omega)$, $\tilde{\mathcal{Y}} = H^1(\Omega)$, $\mathcal{Z} = H^1(\Omega)$ and $\mathcal{J} = [L^2(\Omega)]^d$, as well as the space of designs $\mathcal{U} = L^2(\Omega)$.

The current-operator given by

$$I = I(g_1, T, V) = I_1$$

is well-defined as an operator from $\mathcal{X} \times \mathcal{Y}$ to \mathcal{J} , if we assume an uniformly positive temperature T . This is physically reasonable and also observed in numerical simulations $HJP03$. But so far there is no general proof for this fact available. Choosing an observation-operator R by (4.1) or by

$$R = R(I) = - \int_{\Gamma} I \cdot \nu \, ds, \quad (4.2)$$

yields again that Assumption **(A.3)** is fulfilled. The special choice of the function spaces directly ensures **(A.1)**. The invertibility of $\frac{\partial F}{\partial \rho}$ is so far not investigated analytically. Due to the fact that the energy transport model behaves near to the thermal equilibrium state like the drift diffusion model, we expect that the regularity of the linearization holds at least for such states, i.e., then also **(A.2)** holds $Jun01$. For the energy transport model Algorithm 1 results in the iteration:

- (1) Choose an admissible W^0 .
- (2) For $k = 0, 1, 2, \dots$ solve consecutively

$$\lambda^2 \Delta V^{k+1} = W^k$$

$$\operatorname{div} I_1(g_1^{k+1}, T^{k+1}, V^{k+1}) = 0$$

$$\operatorname{div} I_2(g_2^{k+1}, T^{k+1}, V^{k+1}) = I_1(g_1^{k+1}, T^{k+1}, V^{k+1}) \cdot \nabla V^{k+1} - D(g_1^{k+1}, g_2^{k+1}, T^{k+1}, V^{k+1})$$

$$T^{k+1} = f(g_1^{k+1}, g_2^{k+1})$$

$$\begin{aligned}
 \Delta r_1^{k+1} + \operatorname{div} (r_1^{k+1} \nabla V^{k+1}) + \frac{r_1^{k+1}}{T^{k+1}} |\nabla V^{k+1}|^2 + \frac{1}{T^{k+1}} \nabla V^{k+1} \cdot \nabla r_1^{k+1} &= -D_{g_1} r_2^{k+1} + f_{g_1} r_3^{k+1} \\
 \Delta r_2^{k+1} + \frac{r_2^{k+1}}{T^{k+1}} |\nabla V^{k+1}|^2 + \frac{1}{T^{k+1}} \nabla V^{k+1} \cdot \nabla r_2^{k+1} &= -D_{g_2} r_2^{k+1} + f_{g_2} r_3^{k+1} \\
 -D_T r_2^{k+1} - \frac{g_1^{k+1}}{T^{k+1}} \nabla V^{k+1} \cdot (\nabla r_1^{k+1} + r_2^{k+1} \nabla V^{k+1}) - \frac{g_2^{k+1}}{T^{k+1}} \nabla V^{k+1} \cdot \nabla r_1^{k+1} &= r_3^{k+1} \\
 \lambda^2 \Delta p^{k+1} = \operatorname{div} \left(\frac{g_1^{k+1}}{T^{k+1}} \nabla r_1^{k+1} + \frac{g_2^{k+1}}{T^{k+1}} \nabla r_2^{k+1} - r_2^{k+1} I_2^{k+1} + 2 \frac{r_2^{k+1} g_1^{k+1}}{T^{k+1}} \nabla V^{k+1} \right) + D_V r_2^{k+1} \\
 p^{k+1} &= -\beta(W^{k+1} - W^*) - \tau_k(W^{k+1} - W^k).
 \end{aligned}$$

Hence, after the solution of Poisson's equation, one needs to solve the state system for the generalized densities and the temperature. Note that the temperature can be eliminated using its constitutive relation. Nevertheless, one still has to find the solution of a nonlinear system. Also for the adjoints, one has to solve a linear system. Here, one can also eliminate the adjoint variable r_3^{k+1} .

Remark 4.2. Compared to a standard gradient algorithm one has again the gain of the decoupling from Poisson's equation (compare ^{DP08}). First numerical test of the generalized Gummel iteration for the energy transport model can be found in ^{Dra07}. Note that there a different set of variables is used, which are known as dual entropy variables (for a discussion of different formulations of the energy transport model see also ^{DGJ98,Jun01,DP08}).

4.3. Optimization Based on the Schrödinger–Poisson Model

The next example is devoted to the microscopic *Schrödinger–Poisson model* ^{MRS90}, which is — up to the authors knowledge — so far not used in optimal dopant profiling. Here, the field variable is just the wave function, i.e., $\rho = \psi$. Then, the charge density is given by the square of the modulus of ψ , i.e., $Q(\rho) = |\psi|^2$. The operator F is defined by

$$F(\psi) = -\frac{\varepsilon^2}{2} \Delta \psi + V \psi + V_{ext} \psi,$$

where ε is the scaled Planck constant and $V_{ext} \in L^\infty(\Omega)$ is a given external potential. The corresponding function spaces are $\mathcal{X} = H^1(\Omega; \mathbb{C})$, $\mathcal{Y} = \tilde{\mathcal{Y}} = H^1(\Omega)$, $\mathcal{Z} = H^1(\Omega; \mathbb{C})$ and $\mathcal{J} = [L^p(\Omega)]^d$ with appropriate $p \in (1, 2)$, as well as the space of designs $\mathcal{U} = L^2(\Omega)$. Note that the choice of \mathcal{J} does not use regularity results for solutions of ψ so far. This choice is sufficient for most functionals such as linear ones (negative total current) or quadratic ones in appropriate norms ($H^{-1/2}$). For certain objective functionals, such as quadratic fitting in the L^2 -norm, regularity results for ψ and appropriate embedding have to be employed in order to obtain a reasonable functional-analytic framework.

The current-operator given by

$$I = I(\psi) = \varepsilon \operatorname{Im}(\bar{\psi} \nabla \psi)$$

is well-defined as an operator from \mathcal{X} to \mathcal{J} . Using the previous observation operator (4.2), Assumption **(A.3)** is again fulfilled and the special choice of the function spaces ensures **(A.1)**. Assumption **(A.2)** trivially holds, since F is a linear operator. For the Schrödinger–Poisson model Algorithm 1 results in the iteration:

- (1) Choose an admissible W^0 .
- (2) For $k = 0, 1, 2, \dots$ solve consecutively

$$\begin{aligned} \lambda^2 \Delta V^{k+1} &= W^k \\ -\frac{\varepsilon^2}{2} \Delta \psi^{k+1} + V^{k+1} \psi^{k+1} + V_{ext} \psi^{k+1} &= 0 \\ -\frac{\varepsilon^2}{2} \Delta q^{k+1} + V^{k+1} q^{k+1} + V_{ext} q^{k+1} &= 0 \\ \lambda^2 \Delta p^{k+1} &= \operatorname{Re}(\bar{\psi}^{k+1} q^{k+1}) \\ -\beta(W^{k+1} - W^*) - \tau_k(W^{k+1} - W^k) &= p^{k+1}. \end{aligned}$$

Hence, also for this microscopic model the iteration yields a full decoupling of the equations and on each iteration level we just need to solve two linear Schrödinger equations and two Poisson equations. Note that since ψ and q are complex-valued, a real part is needed in the right-hand side of the fourth equation, which arises from a real part in the Lagrange-functional (cf. *BSV07*).

Remark 4.3. Note that we did not consider the special structure of the boundary conditions or the external potential, which is clearly essential for a correct device simulation *Ar08*. Hence, the above discussion is only formal and only gives insight in the overall structure of the generalized Gummel iteration. Although the analytical understanding of the forward problem is well developed *Ar08, Ba00, BMP05*, there exists so far no analysis on optimization based on the coupled Schrödinger–Poisson model. Nevertheless, first results on bilinear optimal control of Schrödinger equations can be found in *BSV07, ItKu07, TR03, MT03*.

4.4. Optimization Based on the Quantum Drift Diffusion Model

The last example is the *unipolar quantum drift diffusion model* *AnTi87, BaUn98, PiUn99, Pin01*, which is so far also not used for optimal dopant profiling. Here, the field variables are $\rho = (s, G)$, where s is the square root of the electron density and G denotes the generalized quantum quasi Fermi level. Hence, the charge density is given by $Q(\rho) = s^2$. The operator F is defined by

$$F(s, G, V) = \begin{pmatrix} \operatorname{div}(s^2 \nabla G) \\ -\varepsilon^2 \Delta s + s(\log(s^2) + V - G) \end{pmatrix},$$

where ε denotes the scaled Planck constant. The appropriate function spaces are the same ones as for the classical drift diffusion model, i.e., $\mathcal{X} = [H^1(\Omega) \cap L^\infty(\Omega)]^2$,

$\mathcal{Y} = H^2(\Omega)$, $\tilde{\mathcal{Y}} = H^1(\Omega)$, $\mathcal{Z} = H^1(\Omega)$ and $\mathcal{J} = [L^2(\Omega)]^d$, as well as the space of designs $\mathcal{U} = L^2(\Omega)$. This choice of the function spaces ensures Assumptions **(A.1)** and **(A.2)**, at least near to the thermal equilibrium state and for small scaled Planck constants ε (compare also *Pin99*).

The related current–operator is defined by

$$I = I(s, G) = s^2 \nabla G,$$

which is well–posed on $\mathcal{X} \times \mathcal{Y}$ with values in \mathcal{J} .

This current–operator and, e.g., the observation–operator (4.1) yield that also Assumption **(A.3)** holds and Algorithm 1 results in the iteration:

- (1) Choose an admissible W^0 .
- (2) For $k = 0, 1, 2, \dots$ solve consecutively

$$\lambda^2 \Delta V^{k+1} = W^k$$

$$\begin{aligned} \operatorname{div}((s^{k+1})^2 \nabla G^{k+1}) &= 0 \\ -\varepsilon^2 \Delta s^{k+1} + s^{k+1}(\log((s^{k+1})^2) + V^{k+1} - G^{k+1}) &= 0 \end{aligned}$$

$$\begin{aligned} \operatorname{div}((s^{k+1})^2 \nabla q_1^{k+1}) &= 0 \\ -\varepsilon^2 \Delta q_2^{k+1} + (\log((s^{k+1})^2) + V^{k+1} - G^{k+1} + 2^{k+1})q_2^{k+1} + 2s^{k+1} \nabla G^{k+1} \cdot \nabla q_1^{k+1} &= 0 \end{aligned}$$

$$\begin{aligned} \lambda^2 \Delta p^{k+1} &= s^{k+1} q_2^{k+1} \\ -\beta(W^{k+1} - W^*) - \tau_k(W^{k+1} - W^k) &= p^{k+1}. \end{aligned}$$

Hence, apart from Poisson’s equation for the potential V , one needs to solve one nonlinear forward system and a linear adjoint system. But these are much simpler compared to the ones which are needed to solve for the standard gradient algorithm.

Remark 4.4. The forward problem is meanwhile well understood from the analytical point of view, but it was so far not used for optimal dopant profiling. Nevertheless, the adjoints for the full system were derived in *VoUn06* in the context of a different optimization problem.

5. Conclusions

We have studied a generalized Gummel iteration for optimal dopant profiling in modern semiconductor design. Exploiting the special structure of general semiconductor models, i.e., the coupling with a Poisson equation for the electrostatic potential, we were able to provide a convergence proof of the iterative procedure for arbitrary applied voltages. This shows that the optimization problem has much nicer properties than the semiconductor model alone.

References

- AnTi87. M.G. Ancona, H.F. Tiersten. Macroscopic physics of the silicon inversion layer, *Phys. Rev.* 35, 7959–7965, 1987.
- Ar08. A. Arnold, Mathematical properties of quantum evolution equations, in: Lecture Notes in Mathematics, Springer, Berlin, 2008, to appear.
- Ba00. N. Ben Abdallah, On a multidimensional Schrödinger-Poisson scattering model for semiconductors, *J. Math. Phys.* 41, 2241–2261, 2000.
- BMP05. N. Ben Abdallah, F. Mehats, O. Pinaud, On a open transient Schrödinger-Poisson system, *Math. Models Methods Appl. Sci.* 15, 667–688, 2005.
- BaUn98. N. Ben Abdallah A. Unterreiter, On the stationary quantum drift diffusion model, *Z. Angew. Math. Phys.* 49, 251–275, 1998
- BSV07. A. Borzi, J. Salomon, S. Volkwein, Formulation and numerical solution of finite-level quantum optimal control problems, *J. Comp. Applied Math.*, 2007, to appear.
- BP03. M. Burger, R. Pinnau, Fast optimal design of semiconductor devices, *SIAM J. Appl. Math.* 64, 108–126, 2003.
- CCJS93. Z. Chen, B. Cockburn, J. Jerome, C. Shu, Finite element computation of the hydrodynamic model of semiconductor devices, *VLSI Design*, 3, 145–158, 1995.
- DGJ98. P. Degond, S. Génieys, A. Jüngel, A steady-state system in nonequilibrium thermodynamics including thermal and electrical effects, *Math. Meth. Appl. Sci.* 21, 1399–1413, 1998.
- DJP00. P. Degond, A. Jüngel, P. Pietra. Numerical discretization of energy-transport models for semiconductors with non-parabolic band structure, *SIAM J. Sci. Comput.* 22, 986–1007, 2000.
- Dra07. C.R. Drago, On fast optimal control energy-transport-based semiconductor design. In *SCEE06, Mathematics in Industry* 11, 347–355, Springer, 2007.
- AnDr04. C.R. Drago, A.M. Anile, An optimal control approach for an energy transport model in semiconductor design, in: *Proceedings SCEE 2004, Mathematics in Industry*, Springer Verlag, 2004.
- DP08. C. Drago, R. Pinnau, Optimal dopant profiling based on energy-transport semiconductor models, *Math. Models Meth. Appl. Sci.* 18(2):195–214, 2008.
- GS92. H. Gajewski, J. Somrrey, On the uniqueness of solutions of van Roosbroeck equations, *ZAMM* 72, 151–15, 1992.
- Gum64. H.K. Gummel, A self-consistent iterative scheme for one-dimensional steady state transistor calculations, *IEEE Trans. Elec. Dev.*, ED-11, 455–465, 1964.
- HP02a. M. Hinze, R. Pinnau, Optimal control of the drift-diffusion model for semiconductor devices, in: Hoffmann, K.H., et. al. (eds.): *Optimal Control of Complex Structures*. Birkhäuser, Basel, Berlin (2002), 95–106.
- HP02b. M. Hinze, R. Pinnau, An optimal control approach to semiconductor design, *Math. Mod. Meth. Appl. Sci.*, 12, 89–107 (2002).
- HP05. M. Hinze, R. Pinnau, Mathematical tools in optimal semiconductor design, *Bulletin of the Institute of Mathematics, Academia Sinica (New Series)* 4, 569–586, 2007.
- HP06. M. Hinze, R. Pinnau, A second order approach to optimal semiconductor design, *JOTA* 133(2):170–199 (2006).
- HJP03. S. Holst, A. Jüngel, P. Pietra, A mixed finite-element discretization of the energy-transport model for semiconductors, *SIAM J. Sci. Comput.* 24, 2058–2075, 2003.
- ItKu07. K. Ito, K. Kunisch, Optimal bilinear control of an abstract Schrödinger equation, *SIAM J. Control Optim.* 46(1):274–287 (2007).

- Jun01. A. Jüngel. *Quasi-hydrodynamic Semiconductor Equations*, Birkhäuser, PNLDE 41, 2001.
- JP01. A.Jüngel, Y.J.Peng, A model hierarchy for semiconductors and plasmas, *Nonlin. Anal.* 47 (2001), 1821-1832.
- Kel95. C.T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, Philadelphia, 1995.
- Ker86. T. Kerkhoven. A proof of convergence of Gummel's algorithm for realistic device geometries, *SIAM J. Numer. Anal.* 23, 1121-1137, 1986.
- MT03. Y. Maday, G. Turinici, New formulations of monotonically convergent quantum control algorithms, *J. Chem. Phys.*, 118, 8191-8196, 2003.
- MRS90. P.A. Markowich, C.A. Ringhofer, C. Schmeiser, *Semiconductor Equations*, Springer, Wien, New York, 1990.
- PSSS98. R. Plasun, M. Stockinger, R. Strasser, 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,.
- Pin99. R. Pinnau, The linearized transient quantum drift diffusion model — stability of stationary states, *ZAMM*, 80, 327-344, 2000.
- Pin01. R. Pinnau, A review on the quantum drift diffusion model, *Transp. Theory Stat. Phys.*, 31, 367-395, 2002.
- PiUn99. R. Pinnau, A. Unterreiter, The stationary current-voltage characteristics of the quantum drift diffusion model. *SIAM J. Num. Anal.*, 37, 211-245, 1999.
- Sel84. S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Springer, 1984.
- St00. M. Stockinger, *Optimization of ultra-low-power CMOS transistors*, PhD Thesis, Technical University Vienna, 2000.
- Stea98. M. Stockinger, R. Strasser, R. Plasun, A. Wild, S. Selberherr, A qualitative study on optimized MOSFET doping profiles, in: Proceedings SISPAD 98 Conf., Leuven, 77-80, 1998.
- Tro87. G. M. Troianello, *Elliptic Differential Equations and Obstacle Problems*, Plenum Press, New York, first edition, 1987.
- TR03. G. Turinici, H. Rabitz, Wavefunction controllability in quantum systems, *J. Phys.A.*, 36, 2565-2576, 2003.
- VoUn06. A. Unterreiter, S. Volkwein, Optimal control of the quantum drift diffusion model, *Commun. Math. Sci.* 5, 85-111, 2006.