Robust Optimal Design of Quantum Electronic Devices

OCIEL MORALES,* FRANCISCO PERIAGO† and JOSÉ A. VALLEJO‡

July 27, 2018

Abstract

We consider the optimal design of a sequence of quantum barriers in order to manufacture an electronic device at the nanoscale such that the dependence of its transmission coefficient on the bias voltage is linear. The technique presented here is easily adaptable to other response characteristics. The transmission coefficient is computed using the Wentzel-Kramers-Brillouin (WKB) method, so we can explicitly compute the gradient of the objective function. In contrast with earlier treatments, manufacturing uncertainties are incorporated in the model through random variables and the optimal design problem is formulated in a probabilistic setting. As a measure of robustness, a weighted sum of the expectation and the variance of a least-squares performance metric is considered. Several simulations illustrate the proposed approach.

Keywords: Robust Optimal Design, Nanoelectronics, Stochastic Collocation Methods, WKB Approximation.

1 Introduction

Nanoelectronic devices operate with extremely low intensity currents. Under these circumstances, it is desirable to have at our disposal mechanisms to produce and control electronic currents with a high precision. Electronic beams are relatively easy to produce, but their filtering to obtain nanocurrents with specified properties is much more difficult. A widely used approach consists in directing the beam on a sequence of quantum barriers with an externally adjustable bias voltage applied throughout the device. One expects to be able to control the response of the device in the form of a current whose intensity depends, say, linearly on the applied bias. This setting naturally leads to an optimal design problem: What must be the width and height of the layers composing the barriers, supposed fixed in number, in order to achieve this linear response? (Of course, the problem is quite general, admitting a more complex relation between the external voltage and the current, but here we deal with the linear case just for simplicity).

There should be no need to stress the importance of the solution to this problem from a practical point of view, but it must be noticed right from the start that a closed-form, analytic solution is impossible to obtain in most cases. The use of numerical computations at some stage is unavoidable, and this leads to the question of which method to use in order to obtain a good approximation to the solution. In [9] the non–constant potential energy profile is approximated by piecewise constant

---

*Facultad de Ciencias. Universidad Autónoma de San Luís Potosí (México). Email: `omorales@fc.uaslp.mx`

†Departamento de Matemática Aplicada y Estadística. Universidad Politécnica de Cartagena (Spain). Email: `f.periago@upct.es`

‡Facultad de Ciencias. Universidad Autónoma de San Luís Potosí (México). Email: `jvallejo@fc.uaslp.mx`

potentials. Then, the propagation matrix method [5, 8] is applied to compute the transmission coefficient and, finally, the gradient of a least–squares–type objective function (which is required by the numerical solution method) is computed using the adjoint method. It is important to point out that different discretization processes, which are used to approximate objective functions and/or its gradients, may lead to very different results. Moreover, it has been observed in some optimization problems [7] that first approximating a cost functional, and then computing the gradient of the approximated one, in general differs from approximating the gradient of the exact cost functional. That is, the schemes 'first discretize, then optimize' and 'first optimize, then discretize', do not commute in general. Also, as it will be showed later on in this paper, optimizing for the same cost functional via its exact gradient gives different solutions than using an approximate one.

Another issue, which cannot be obviated in a realistic mathematical model, is the presence of uncertainties. There are several sources of uncertainty in the problem under consideration, one of the most important regarding the influence on the computed optimal design being the manufacturing uncertainties. Due to the smallness of the currents involved, and the narrow width of the quantum barriers needed, methods such as MBE (Molecular Beam Epitaxy) or CVD (Chemical Vapor Deposition) are used to growth thin layers (in many cases, monolayers) of some material to build the barriers, two of the preferred ones being MoS_2 and $GaAs$ (see [2, 6], for example). These methods allow the growth even of monolayers, but the difficulties inherent to the manufacturing process at a semi–commercial scale lead almost inevitably to inaccuracies that ultimately lead to a potential configuration that may be different from the numerically computed, optimal one [11]. For these reasons, the problem of computing an optimal quantum profile which, in addition, is robust against those uncertainties is an important one. If there is some statistical information about the uncertainties, then the machinery of probability theory gives a framework in which to we can accomodate uncertainties (by using random variables and/or random fields), and model objective functions (by means of expectation and variance operators, among others choices). In [14], this approach has been used for the case in which the cost functional only includes the averaging of a least–squares performance metric, and by using the standard Monte–Carlo method for its numerical resolution.

The present work addresses the problem of the optimal design of a quantum potential profile (modeling a nanoelectronic device) in order to obtain a transmission coefficient linearly depending on an externally applied bias voltage, in the presence of manufacturing uncertainties. The transmission coefficient is explicitly computed by using the WKB method. As a consequence, an explicit formula for the gradient of the cost functional is obtained. A weighted sum of expectation and variance of a random least–squares performance metric is considered as a measure of robustness. The inclusion of the second order statistical moment in the cost functional amounts to a reduction the dispersion of the random transmission coefficient and hence an increase in the robustness of the optimal design. Since the resulting integrand in the cost functional is smooth with respect to a random parameter, a sparse grid stochastic collocation method is used for the numerical approximation of the involved integrals in the random domain. This method preserves the parallelizable character of Monte–Carlo sampling. However, in contrast to Monte–Carlo (which is computationally very expensive, of order $O\left(M^{-1/2}\right)$, with M the number of random sampling points), the stochastic collocation method shows an exponential convergence with respect to the number of sampling points. Several simulations illustrate the proposed approach, which shows itself to be an improvement in accuracy over previous ones of about a 69.4%.

2 Setting of the Optimal Design Problems

Considered is a nanoscale semiconductor electronic device composed of N layers occupying positions $x_0 = 0 < x_1 < \cdots < x_N = L$. The local potential energy at the ith layer is denoted by U_i, $i = 1, 2, \cdots, N$. For $x < x_0$, the potential energy is denoted by U_0 and for $x > x_N$ it is U_{N+1}. It is assumed that a single electron propagating from $-\infty$ is incident at x_0 and that a voltage bias V_{bias} is applied across the device. A linear approximation of the underlying Poisson's equation [9, 13] leads to the following expression for the resulting potential energy profile

$$V(x) = V(x, U, V_{\text{bias}}) = \begin{cases} U_0, & -\infty < x < x_0 \\ \sum_{j=1}^{N} U_j \mathcal{X}_j(x) - V_{\text{bias}} \dfrac{x - x_0}{L}, & x_0 \leq x \leq x_N \\ U_{N+1} - V_{\text{bias}}, & x_N < x < +\infty, \end{cases} \tag{1}$$

where $U = (U_1, \cdots, U_N)$ is the vector of local layer potentials in the device, and

$$\mathcal{X}_j(x) = \begin{cases} 1, & x_{j-1} \leq x < x_j \\ 0, & \text{otherwise}, \end{cases}$$

is the characteristic function of the interval $[x_{j-1}, x_j[$, $1 \leq j \leq N$.

The transmission coefficient of the device $T = T(V_{\text{bias}}, U)$ is defined as the ratio of current density transmitted from the device at $x = x_N$ and the incident one at $x = x_0$. As explained in detail in [9], T may be expressed as

$$T(V_{\text{bias}}, U) = \frac{\kappa_{N+1}}{\kappa_0} |\psi(x_N)|, \tag{2}$$

where $\kappa_0 = \sqrt{2me(E - U_0)}/\hbar$ and $\kappa_{N+1} = \sqrt{2me(E - U_{N+1} + V_{\text{bias}})}/\hbar$, for values of the energy $E > \max\{U_0, U_{N+1} - V_{\text{bias}}\}$. The cases $E \leq U_0$ and $E \leq U_{N+1} - V_{\text{bias}}$ may be treated in a similar way. Here m is the effective mass of the electron, e denotes the electron charge, $\hbar$ is Planck's constant, E is the electron energy, and finally $\psi(x)$ solves the following boundary-value problem for the Schrödinger equation

$$\begin{cases} -\dfrac{\hbar^2}{2m} \dfrac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x), & x_0 < x < x_N \\ \mathbf{i}\kappa_0\psi(x_0) + \dfrac{d\psi}{dx}(x_0) = 2\mathbf{i}\kappa_0 A_0 e^{\mathbf{i}\kappa_0 x_0}, \\ \mathbf{i}\kappa_{N+1}\psi(x_N) - \dfrac{d\psi}{dx}(x_N) = 0. \end{cases} \tag{3}$$

Here $\mathbf{i}$ denotes the unit imaginary number and A_0 the amplitude of the transmitted wave at x_0.

2.1 Deterministic Optimal Design

The (deterministic) optimal design problem considered in this paper is formulated as the following nonlinear data-fitting problem: Given a desired transmission coefficient $T_0(V_{\text{bias}})$, which is defined for $V_{\text{min}} \leq V_{\text{bias}} \leq V_{\text{max}}$, and lower, U_L, and upper, U_R, bounds for the local layer potentials, with $0 \leq U_L < U_R < \infty$,

$$\begin{cases} \text{Minimize in} & U = (U_1, \cdots, U_N) : J(U) = \sum_{i=1}^{M} |T_0(V_i) - T(V_i, U)|^2 \\ \text{subject to} & U_L \leq U_j \leq U_R, \quad j = 1, \cdots, N, \end{cases} \tag{4}$$

where $T(V_i, U)$ is given by (2) with $V_{\text{bias}} = V_i$ and $V_{\text{min}} \leq V_i \leq V_{\text{max}}$, $i = 1, \cdots, M$.

2.2 Optimal Design Under Manufacturing Uncertainties

As indicated in the introduction, it is very convenient to analyse the robustness of optimal designs with respect to manufacturing uncertainties. These may be modelled by adding a vector of random variables

$$X(\omega) = (X_1(\omega), \cdots, X_N(\omega)) \tag{5}$$

to the vector U of local layer potentials. Here ω represents a random event and thus $X_j(\omega)$ is a small unknown error in manufacturing the local potential U_j. Hence, the cost functional J considered in problem (4) becomes a random variable given by

$$J(U, \omega) = \sum_{i=1}^{M} |T_0(V_i) - (T(V_i, U + X(\omega)))|^2. \tag{6}$$

In order to obtain a design of the potential energy profile U less sensitive with respect to fabrication unknown fluctuations, the new cost functional is considered:

$$J_\alpha(U) = \mathbb{E}(J(U, \cdot)) + \alpha \mathrm{Var}(J(U, \cdot)), \tag{7}$$

with $\alpha \geq 0$ a weighting parameter. Here $\mathbb{E}$ and Var denote the expectation and variance operators, respectively. Then, the robust optimization problem is formulated as

$$\begin{cases} \text{Minimize in} & U = (U_1, \cdots, U_N) : J_\alpha(U) \\ \text{subject to} & U_L \leq U_j \leq U_R, \quad j = 1, \cdots, N. \end{cases} \tag{8}$$

where $J_\alpha(U)$ is given by (7).

3 Solving the Optimal Design problems

The numerical resolution of the optimal design problems stated in the preceding section requires the computation of the transmission coefficient (2) and, therefore, the resolution of the boundary-value problem (3). This problem may be numerically approximated by standard numerical methods such as finite differences or finite elements. Another approach is proposed in [9] where, after approximating the potential $V(x)$, as given by (1), by piecewise constant potentials, problem (3) is transformed into a two–dimensional linear non–autonomous difference equation. Here we propose a different approach based on the so–called WKB method [10]. From the point of view of optimization, WKB method is very appealing since, within its range of validity, it provides an explicit form for the solution to (3). From this, explicit expressions for the gradients of the cost functionals considered in problems (4) and (8) are derived. In addition, having an explicit expression for $J(U, \omega)$ allows us to prove its smoothness with respect to U and ω. From this, both existence of solutions to (4) and (8), as well as designing a computationally very efficient numerical resolution method, will be derived in this section.

We begin by explicitly computing the transmission coefficient (2) and then describe the numerical resolution methods for problems (4) and (8).

3.1 Explicit Computation of Transmission Coefficient

3.1.1 Case of a single potential barrier

For the sake of clarity, consider first the case of a single potential barrier as illustrated in Figure 1.
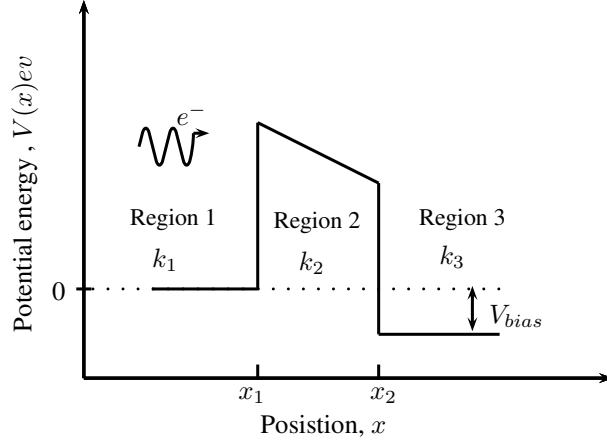
Figure 1: Sketch of the onde-dimensional potential energy barrier $V(x) = U - V_{\text{bias}}\frac{x-x_1}{x_2-x_1}$, $x_1 \leq x \leq x_2$. An electron of mass m, charge e and energy E, incident from left, has wave vector k_j in region j.

The WKB method proposes a solution of the Schrödinger equation in the form

$$\psi(x) = \begin{cases} \psi_1(x) = A_1 e^{\frac{\mathbf{i}}{\hbar}\kappa_1 x} + B_1 e^{-\frac{\mathbf{i}}{\hbar}\kappa_1 x}, & x < x_1 \\ \psi_2(x) = \frac{A_2}{\sqrt{\kappa_2(x)}} e^{\frac{\mathbf{i}}{\hbar}\int_{x_1}^{x}\kappa_2(s)ds} + \frac{B_2}{\sqrt{\kappa_2(x)}} e^{-\frac{\mathbf{i}}{\hbar}\int_{x_1}^{x}\kappa_2(s)ds}, & x_1 \leq x \leq x_2 \\ \psi_3(x) = A_3 e^{\frac{\mathbf{i}}{\hbar}\kappa_3 x} + B_3 e^{-\frac{\mathbf{i}}{\hbar}\kappa_3 x}, & x > x_2, \end{cases} \tag{9}$$

with $\kappa_1 = \sqrt{2meE}$, $\kappa_2(x) = \sqrt{2me(E - V(x))}$ if $E > V(x)$ for all $x_1 \leq x \leq x_2$, and $\kappa_3(x) = \sqrt{2me(E + V_{\text{bias}})}$ for all $x > x_2$.

In the context of quantum electronic devices, solutions to (3) admit a smooth representative in their L^2 classes (of regularity class C^1). Hence, continuity of ψ and its first derivative at the interface point x_1 leads to

$$\begin{bmatrix} e^{\frac{\mathbf{i}}{\hbar}\kappa_1 x_1} & e^{-\frac{\mathbf{i}}{\hbar}\kappa_1 x_1} \\ \frac{\mathbf{i}}{\hbar}\kappa_1 e^{\frac{\mathbf{i}}{\hbar}\kappa_1 x_1} & -\frac{\mathbf{i}}{\hbar}\kappa_1 e^{-\frac{\mathbf{i}}{\hbar}\kappa_1 x_1} \end{bmatrix} \begin{bmatrix} A_1 \\ B_1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{\kappa_2(x_1)}} & \frac{1}{\sqrt{\kappa_2(x_1)}} \\ -\frac{C}{2\kappa_2^2(x_1)\sqrt{\kappa_2(x_1)}} + \frac{\mathbf{i}\kappa_2(x_1)}{\sqrt{\kappa_2(x_1)}} & -\frac{C}{2\kappa_2^2(x_1)\sqrt{\kappa_2(x_1)}} - \frac{\mathbf{i}\kappa_2(x_1)}{\sqrt{\kappa_2(x_1)}} \end{bmatrix} \begin{bmatrix} A_2 \\ B_2 \end{bmatrix} \tag{10}$$

where

$$C = C\left(V_{\text{bias}}\right) = \frac{meV_{\text{bias}}}{x_2 - x_1}. \tag{11}$$

By writing each 2×2 matrix in (10) as the product of two matrices as follows

$$\underbrace{\begin{bmatrix} 1 & 1 \\ \frac{\mathbf{i}}{\hbar}\kappa_1 & -\frac{\mathbf{i}}{\hbar}\kappa_1 \end{bmatrix}}_{K(\kappa_1)} \underbrace{\begin{bmatrix} e^{\frac{\mathbf{i}}{\hbar}\kappa_1 x_1} & 0 \\ 0 & e^{-\frac{\mathbf{i}}{\hbar}\kappa_1 x_1} \end{bmatrix}}_{E(\kappa_1, x_1)} \begin{bmatrix} A_1 \\ B_1 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 1 & 1 \\ \frac{-C}{2\kappa_2(x_1)} + \frac{\mathbf{i}\kappa_2(x_1)}{\hbar} & \frac{-C}{2\kappa_2(x_1)} - \frac{\mathbf{i}\kappa_2(x_1)}{\hbar} \end{bmatrix}}_{\mathcal{K}(\kappa_2, x_1)} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{\kappa_2(x_1)}} & 0 \\ 0 & \frac{1}{\sqrt{\kappa_2(x_1)}} \end{bmatrix}}_{\mathcal{E}(\kappa_2, x_1)} \begin{bmatrix} A_2 \\ B_2 \end{bmatrix} \tag{12}$$

5

and solving (12) for A_1 and B_1,

$$\begin{bmatrix} A_1 \\ B_1 \end{bmatrix} = E^{-1}(\kappa_1, x_1) K^{-1}(\kappa_1) \mathcal{K}(\kappa_2, x_1) \mathcal{E}(\kappa_2, x_1) \begin{bmatrix} A_2 \\ B_2 \end{bmatrix}. \tag{13}$$

Proceeding in the same way at the point x_2, one obtains

$$\begin{bmatrix} A_2 \\ B_2 \end{bmatrix} = \mathcal{E}^{-1}(\kappa_2, x_2) \mathcal{K}^{-1}(\kappa_2, x_2) K(\kappa_3) E(\kappa_3, x_2) \begin{bmatrix} A_3 \\ B_3 \end{bmatrix}, \tag{14}$$

where

$$\mathcal{K}(\kappa_j, x) = \begin{bmatrix} 1 & 1 \\ -\frac{C}{2\kappa_j^2(x)} + \frac{\mathbf{i}}{\hbar}\kappa_j(x) & -\frac{C}{2\kappa_j^2(x)} - \frac{\mathbf{i}}{\hbar}\kappa_j(x) \end{bmatrix}$$

and

$$\mathcal{E}(\kappa_j, x) = \begin{bmatrix} \dfrac{e^{\frac{\mathbf{i}}{\hbar}\int_{x_{j-1}}^{x} \kappa_j(s)\,ds}}{\sqrt{\kappa_j(x)}} & 0 \\ 0 & \dfrac{e^{-\frac{\mathbf{i}}{\hbar}\int_{x_{j-1}}^{x} \kappa_j(s)\,ds}}{\sqrt{\kappa_j(x)}} \end{bmatrix}.$$

Substituting (14) into (13), we get

$$\begin{bmatrix} A_1 \\ B_1 \end{bmatrix} = E^{-1}(\kappa_1, x_1) K^{-1}(\kappa_1) \mathcal{K}(\kappa_2, x_1) \mathcal{E}(\kappa_2, x_1) \mathcal{E}^{-1}(\kappa_2, x_2) \mathcal{K}^{-1}(\kappa_2, x_2) K(\kappa_3) E(\kappa_3, x_2) \begin{bmatrix} A_3 \\ B_3 \end{bmatrix}. \tag{15}$$

Denoting by M the product of matrices from $E^{-1}(\kappa_1, x_1)$ to $E(\kappa_3, x_2)$ the transmission coefficient (2) takes the form $T = |\frac{1}{M_{11}}|^2$, where M_{11} is the first entry of M. More precisely, the following explicit formula for that transmission coefficient, in the case $E > V(x)$ for all $x_1 \le x \le x_2$, is obtained:

$$\begin{aligned} T(V_{\text{bias}}, U) \quad =& \frac{\kappa_2(x_1)}{\kappa_2(x_2)} \left[\left(\frac{\hbar C \cos\frac{I}{\hbar}}{4\kappa_1\kappa_2^2(x_1)} - \frac{\hbar\kappa_2(x_1)C\cos\frac{I}{\hbar}}{4\kappa_1\kappa_2^3(x_2)} - \frac{\kappa_2(x_1)\sin\frac{I}{\hbar}}{2\kappa_1} - \frac{\hbar^2 C^2 \sin\frac{I}{\hbar}}{8\kappa_1\kappa_2^2(x_1)\kappa_2^3(x_2)} - \frac{\kappa_3 \sin\frac{I}{\hbar}}{2\kappa_2(x_2)} \right)^2 \right. \\ &\left. + \left(\frac{\cos\frac{I}{\hbar}}{2} + \frac{\kappa_3\kappa_2(x_1)\cos\frac{I}{\hbar}}{2\kappa_1\kappa_2(x_2)} - \frac{\hbar C \sin\frac{I}{\hbar}}{4\kappa_2^3(x_2)} + \frac{\hbar C\kappa_3 \sin\frac{I}{\hbar}}{4\kappa_1\kappa_2^2(x_1)\kappa_2(x_2)} \right)^2 \right]^{-1}, \end{aligned} \tag{16}$$

where C is given by (11) and

$$I = I(V_{\text{bias}}, U) = \frac{2\sqrt{2me}(x_2 - x_1)}{3mV_{\text{bias}}} \left[(E - U + V_{\text{bias}})^{3/2} - (E - U)^{3/2} \right] \tag{17}$$

The case $E < V(x)$ for all $x_1 \le x \le x_2$ is completely analogous. Denoting by

$$\overline{\kappa}_2(x) = \sqrt{2me(V(x) - E)}, \quad x_1 \le x \le x_2, \tag{18}$$

the transmission coefficient is given by

$$\begin{aligned} T(V_{\text{bias}}, U) \quad =& \frac{\overline{\kappa}_2(x_1)}{\overline{\kappa}_2(x_2)} \left[\left(\frac{\overline{\kappa}_2(x_1)\sinh\frac{\overline{I}}{\hbar}}{2\kappa_1} - \frac{\hbar C \cosh\frac{\overline{I}}{\hbar}}{4\kappa_1\overline{\kappa}_2^2(x_1)} + \frac{\hbar C\overline{\kappa}_2(x_1)\cosh\frac{\overline{I}}{\hbar}}{4\kappa_1\overline{\kappa}_2^3(x_2)} - \frac{\hbar^2 C^2 \sinh\frac{\overline{I}}{\hbar}}{8\kappa_1\overline{\kappa}_2^2(x_1)\overline{\kappa}_2^3(x_2)} - \frac{\kappa_3 \sinh\frac{\overline{I}}{\hbar}}{2\overline{\kappa}_2(x_2)} \right)^2 \right. \\ &\left. + \left(\frac{\cosh\frac{\overline{I}}{\hbar}}{2} + \frac{\overline{\kappa}_2(x_1)\kappa_3\cosh\frac{\overline{I}}{\hbar}}{2\kappa_1\overline{\kappa}_2(x_2)} + \frac{\hbar C \sinh\frac{\overline{I}}{\hbar}}{4\overline{\kappa}_2^3(x_2)} - \frac{\hbar C\kappa_3 \sinh\frac{\overline{I}}{\hbar}}{4\kappa_1\overline{\kappa}_2(x_1)^2\overline{\kappa}_2(x_2)} \right)^2 \right]^{-1} \end{aligned} \tag{19}$$

where

$$\overline{I} = \overline{I}(V_{\text{bias}}, U) = \frac{2\sqrt{2me}(x_2 - x_1)}{3mV_{\text{bias}}} \left[(U - E)^{3/2} - (U - E - V_{\text{bias}})^{3/2} \right] \tag{20}$$

6

3.1.2 Range of validity for WKB method

The condition for the validity of WKB method and therefore for formulas (16) and (19) is that the change in the potential energy over the decay length be smaller than the magnitude of the kinetic energy (see, for instance, [13, p. 483]). For the case $E > V(x)$, this condition can be expressed as

$$\left| \hbar \frac{dV/dx}{\kappa_2(x)} \right| < |V(x) - E| \quad \forall x_1 \leq x \leq x_2, \tag{21}$$

and for $E < V$,

$$\left| \hbar \frac{dV/dx}{\overline{\kappa}_2(x)} \right| < |V(x) - E| \quad \forall x_1 \leq x \leq x_2. \tag{22}$$

In particular, (21) and (22) constraint the values of U and V_{bias} for which the WKB methods applies. Assume that $U > V_{\text{bias}}$: Since $V(x) \geq U - V_{\text{bias}}$, $\kappa_2(x) \geq \sqrt{2me(E - U)}$ for all $x_1 \leq x \leq x_2$. Hence, by introducing the function

$$F_{E>V}(V_{\text{bias}}, U) = |E - U| - \frac{\hbar V_{\text{bias}}}{(x_2 - x_1)\sqrt{2me|E - U|}}, \tag{23}$$

condition (21) is satisfied whenever $F_{E>V}(V_{\text{bias}}, U) > 0$. Figure 2 displays the functions $F_{E>V}(V_{\text{bias}}, U)$ and its associated transmission coefficient $T(V_{\text{bias}}, U)$, as given by (16), for $U = 0.48$ and $U = 0.55$.
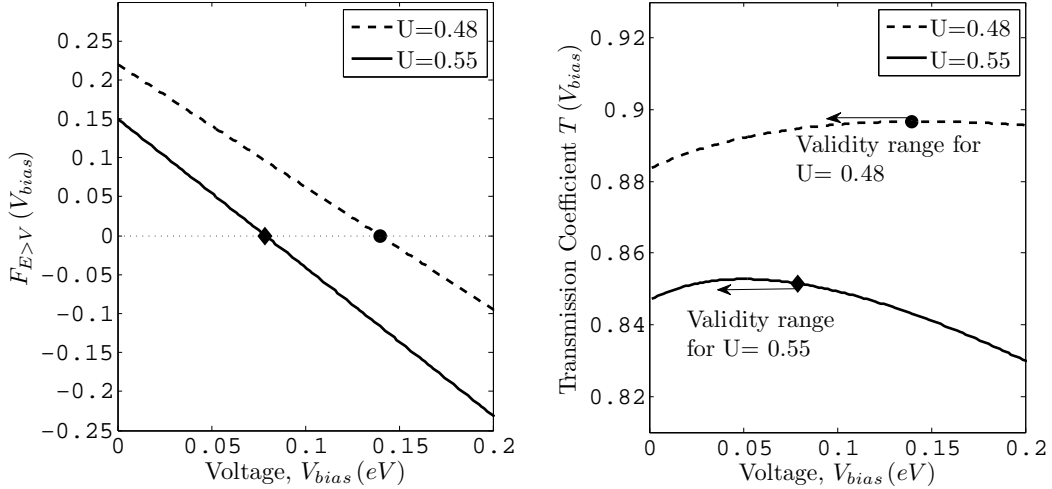

Figure 2: *(Left)* Picture of $F_{E>V}(V_{\text{bias}}, U)$ for $U = 0.48$eV (dashdotted line) and $U = 0.55$eV (continuous line). *(Right)* Transmission coefficient $T(V_{\text{bias}}, U)$ for $U = 0.48$eV (dashdotted line) and $U = 0.55$eV (continuous line). Range of validity for WKB method is highlighted. In both pictures, effective mass is $m = 0.07m_0$, where m_0 is the bare electron mass, electron's energy is $E = 0.7$eV and the thickness barrier is 1nm.

Analogously in the case $E < V$, by considering the function

$$F_{E<V}(V_{\text{bias}}, U) = |E - U + V_{\text{bias}}| - \frac{\hbar V_{\text{bias}}}{(x_2 - x_1)\sqrt{2me|E - U + V_{\text{bias}}|}}, \tag{24}$$

condition (22) holds for $F_{E<V}(V_{\text{bias}}, U) > 0$. Figure 3 plots the functions $F_{E<V}(V_{\text{bias}}, U)$ and the transmission coefficient $T(V_{\text{bias}}, U)$ given in (19) for two values of U, namely $U = 0.45$ and $U = 0.55$ with the energy of an electron $E = 0.0.26$eV.
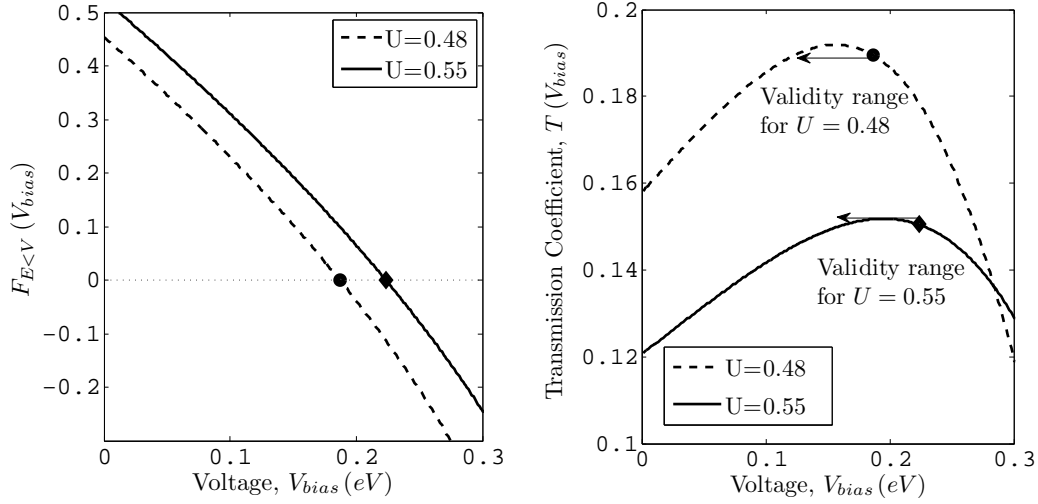
Figure 3: *(Left)* Picture of $F_{E<V}(V_{\text{bias}}, U)$ for $U = 0.48$eV (dashdotted line) and $U = 0.55$eV (continuous line). *(Right)* Transmission coefficient $T(V_{\text{bias}}, U)$ for $U = 0.48$eV (dashdotted line) and $U = 0.55$eV (continuous line). Range of validity for WKB method are highlighted. In both pictures, effective mass is $m = 0.07m_0$, where m_0 is the bare electron mass, electron's energy is $E = 0.026$eV and the thickness barrier is 1nm.

3.1.3 Case of multiple potential barriers

Consider now the configuration described at the beginning of Subsection 2.1, which is composed of N potential barriers with energy given by (1). Let us denote by

$$V_j(x) = U_j - V_{\text{bias}}\frac{x - x_0}{L}, \quad x_{j-1} \leq x < x_j, \tag{25}$$

the potential energy of the jth barrier and by

$$\kappa_j(x) = \begin{cases} \sqrt{2me(E - U_0)}, & -\infty < x < x_0, \quad j = 0 \\ \sqrt{2me(E - V_j(x))}, & x_{j-1} \leq x \leq x_j, \quad j = 1, 2, \cdots N \\ \sqrt{2me(E - U_{N+1} + V_{\text{bias}})}, & x_N < x < +\infty, \quad j = N+1. \end{cases} \tag{26}$$

The linear relationship between the coefficients A_0, B_0 of the electron wave function at x_0 and the ones A_N, B_N at x_N is expressed as

$$\begin{bmatrix} A_0 \\ B_0 \end{bmatrix} = E^{-1}(\kappa_0, x_0) K^{-1}(k_0) \cdot \prod_{j=1}^{N} M_j \cdot K(\kappa_{N+1}) E(x_N, \kappa_{N+1}) \begin{bmatrix} A_N \\ B_N \end{bmatrix}, \tag{27}$$

where the form of the transfer matrices M_j depend on the relationship between E and V_j. For the cases in which WKB method may be applied, M_j has been computed in the preceding section. For the spatial regions for which WKB method does not apply, appropriate connection formulae must be used. For instance, as in [8, 14], the linear potential $V_j(x)$ is approximated by piece-wise contant potential for which explicit expressions of M_j are well-known [5].

Finally, as in the case of a single potential barrier, the transmission coefficient $T = T(V_{\text{bias}}, U)$ is obtained from (27).

3.2 Numerical Resolution Method

Before describing the numerical methods proposed in this paper for solving (4) and (8), let us briefly consider the question concerning the existence of solution for such problems.

From the explicit expressions (16) and (19), and taking into the computations in Subsection 3.1.3, it is clear that the functionals $J(U)$ and $J_\alpha(U)$, which have been considered in problems (4) and (8), respectively, are smooth (of regularity class C^∞). In addition, the set of admissible designs $\mathcal{U}_{\text{ad}} = \left\{ U = (U_1, \cdots, U_N) \in \mathbb{R}^N : U_L \leq U_j \leq U_L, \quad 1 \leq j \leq N \right\}$ is a compact set of $\mathbb{R}^N$. As a consequence, the following existence result holds.

THEOREM 3.1 *Problems* (4) *and* (8) *have, at least, one solution.*

The nonlinear mathematical programming problem (4) is standard and may be solved by several methods, typically by a gradient-based method. In this paper, a subspace trust region method, which is based on the interior-reflective Newton method, as proposed in [3, 4], is used. This algorithm is implemented in the MatLab constrained optimization routine *fmincon*.

More challenging is the robust optimal design problem (8). The brute–force sampling Monte Carlo is the most commonly method used to solve this kind of problems. However, for smooth (with respect to the random parameter) functions, sparse grid, stochastic collocation methods are able to keep the same accuracy as Monte Carlo and, in addition, are computationally much more efficient [1]. In our case, again from (16) and (19) it is not hard to show that $J(U, \omega)$, as defined in (6), is analytic with respect to the random variable $X(\omega)$. For this reason, we propose an adaptive, isotropic, sparse grid, stochastic collocation method to approximate both, the cost functional $J_\alpha(U)$ and its gradient. Then, the robust optimal design problem (8) can be solved as in the deterministic case.

In order to explain the method for approximating integrals in the random domain used in this paper, let us first introduce some notation. By $(\Omega, \mathcal{F}, \mathbb{P})$ we denote a complete probability space. Ω is the set of outcomes, $\mathcal{F}$ is the σ-algebra of events and $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a probability measure. $\Gamma_j = X_j(\Omega)$, $1 \leq j \leq N$, are the images spaces of the sample space Ω through the real-valued random variables X_j considered in (5), and $\Gamma = \prod_{j=1}^N \Gamma_j$ is the product space. Assuming that the distribution measure of $X(\Omega)$ is absolutely continuous with respect to the Lebesgue measure, there exists a joint probability density function $\rho : \Gamma \to \mathbb{R}_+$ for $X = (X_1, \cdots, X_N)$. Hence, $(\Omega, \mathcal{F}, \mathbb{P})$ is mapped to $(\Gamma, \mathcal{B}, \rho(z) \, dz)$, where $\mathcal{B}$ is the σ-algebra of Borel sets on Γ, and dz is the Lebesgue measure. Finally, the expectation and variance of $J(U, \omega)$ take the form

$$\mathbb{E}[J(U, \cdot)] = \int_\Gamma J(U, z) \, \rho(z) \, dz, \tag{28}$$

$$\text{Var}[J(U, \cdot)] = \int_\Gamma J^2(U, z) \, \rho(z) \, dz - \left(\int_\Gamma J(U, z) \, \rho(z) \, dz \right)^2. \tag{29}$$

Following [1, 12], the isotropic sparse grid of sampling quadrature nodes is defined as follows. Starting from an integer ℓ (called the level), the index set

$$\mathbb{I}(\ell, N) = \left\{ i = (i_1, \cdots, i_N) \in \mathbb{N}_+^N : \sum_{n=1}^N (i_n - 1) \leq \ell \right\}$$

is considered, with $\mathbb{N}_+ = \{1, 2, 3, \cdots\}$. The level ℓ determines the number of collocation points R_{i_n} in the nth stochastic direction, which for the case of Smolyak rule is given by

$$R_{i_n} = \begin{cases} 1, & \text{for } i_n = 1 \\ 2^{i_n - 1} + 1, & \text{for } i_n > 1. \end{cases}$$

Smolyak quadrature rule applied to a generic function $F : \Gamma \to \mathbb{R}$ gives

$$\int_\Gamma F(z) \rho(z) \, dz \approx \sum_{i \in \mathbb{I}(\ell, N)} \left(\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_N} \right) F = \sum_{r_1=1}^{R_{i_1}} \cdots \sum_{r_N=1}^{R_{i_N}} F\left(z_1^{r_1}, \cdots, z_N^{r_N} \right) w_1^{r_1} \cdots w_N^{r_N}, \quad (30)$$

where $\Delta^{i_n} = \mathcal{Q}^{i_n} - \mathcal{Q}^{i_n - 1}$, with $\mathcal{Q}^0 = 0$, is a quadrature rule in which the coordinates $z_n^{r_n}$ of the nodes are the same as those for the 1D quadrature formula $\mathcal{Q}^{i_n}$ and its associated weights $w_n^{r_n}$ are the difference between those for the i_n and $i_n - 1$ levels.

It remains to analyze the question on how to properly choose the quadrature level ℓ. For this:

(1) A positive, large enough, integer $\bar{\ell}$, and a tolerance level $0 < \varepsilon \ll 1$ are fixed.

(2) The first and second order statistical moments of $J(U, z)$, as given by (28) and the first term in the right-hand side of (29), are approximated by using (30), with level $\bar{\ell}$. These two approximations, which are denoted by $\mathcal{M}_{1,\bar{\ell}}(J(U))$ and $\mathcal{M}_{2,\bar{\ell}}(J(U))$, respectively, play the role of *enriched or reference* values for the exact values of the first two statistical moments of $J(U, z)$, which, obviously, cannot be explicitly computed.

(3) Finally, the level ℓ is linearly increased from $\ell = 1$ to $\ell_{\text{opt}} < \bar{\ell}$ until the stopping criterion

$$\max \left\{ \frac{|\mathcal{M}_{1,\bar{\ell}}(J(U)) - \mathcal{M}_{1,\ell}(J(U))|}{\mathcal{M}_{1,\bar{\ell}}(J(U))} \quad , \quad \frac{|\mathcal{M}_{2,\bar{\ell}}(J(U)) - \mathcal{M}_{2,\ell}(J(U))|}{\mathcal{M}_{2,\bar{\ell}}(J(U))} \right\} \leq \varepsilon, \quad (31)$$

is satisfied. Here $\mathcal{M}_{1,\ell}(J(U))$ and $\mathcal{M}_{2,\ell}(J(U))$ denote, respectively, approximations, by using (30) with level ℓ, of the first two statistical moments of $J(U, z)$. If (31) is not satisfied for the tolerance ε and the initial $\bar{\ell}$, then the reference level $\bar{\ell}$ is increased.

For more details on this adaptive algorithm, including its convergence, we refer the reader to [1] and the references therein.

4   Numerical Simulations

In this section, numerical results for problems (4) and (8) are presented and discussed. In all experiments, a four layers device is considered of the same thickness (1nm), so that the total length is $L = 4$nm. The desired linear transmission coefficient is

$$T_0(V_{\text{bias}}) = 0.00002 V_{\text{bias}} + 0.0000099, \quad 0 \leq V_{\text{bias}} \leq 0.25. \quad (32)$$

Quadratic and square root transmission coefficients may be treated analogously. The design is based on 10 equally spaced bias voltages. Hence, $V_i = i \frac{0.25}{10}$, $1 \leq i \leq 10$. The electron mass is $m = 0.07 \times m_0$, with $m_0 = 9.10939 \times 10^{-31}$ Kg (which is appropriate for an electron in the conduction band of $\text{Al}_\xi \text{Ga}_{1-\xi} \text{As}$), its energy is $E = 0.026$ eV, and its charge $e = 1.602 \times 10^{-19}$ C. Planck's constant is $\hbar = 1.05457 \times 10^{-34} \, J \cdot s$. The lower and upper bounds for the design variable are taken as $U_L = 0.7$ eV and $U_H = 1.7$ eV, respectively.

The goal of this section is twofold: On the one hand, it is aimed at analyzing the differences that may occur when using exact gradients or numerical gradients in the optimization algorithm. On the other hand, we want to analyze the influence of manufacturing uncertainties on the computed designs. We deal with these issues in the following subsections.

4.1 Exact versus numerical gradient

In this experiment, the deterministic problem (4) is solved, by using the MatLab routine *fmincon*, in the cases where: (a) The exact gradient is provided, as computed from the explicit expression for the cost functional $J(U)$ in Section 3, and (b) the gradient is numerically computed by using finite differences. In both cases, the algorithm is initiated with $U_j^0 = 0.7$ eV, $1 \leq j \leq 4$, and the stopping criterion, provided by the MatLab routine *fmincon*, is fixed to 10^{-15}.

First column in Table 1 displays results for the values of the cost functional after convergence of the algorithm. The optimal energy potentials are showed in the remaining columns. Inspection of Table 1 reveals (as expected) that performance increases by using the exact gradient. Precisely, the value of the objective function obtained by using the optimal design as computed with the exact gradient improves in about 69.4% the corresponding one obtained via the numerical gradient.

Table 1: Results for (4) using exact *(first row)* and numerical gradient *(second row)*.

Gradient	$J(U)$	U_1	U_2	U_3	U_4
Exact	1.43×10^{-12}	0.70	1.31	1.54	0.70
Numerical	4.68×10^{-12}	1.03	0.96	0.96	1.02

4.2 Deterministic design versus design under uncertainty

Manufacturing uncertainties are modelled by random variables uniformly distributed in $[-a, a]$, i.e., $X_j = \mathcal{U}(-a, a)$ for all $j = 1, 2, 3, 4$. As an illustration, the cases $a = 0.05$ and $a = 0.2$ are considered. As in the preceding example, the algorithm is initiated with $U_j^0 = 0.7$ eV, $1 \leq j \leq 4$, and the stopping criterion is fixed to 10^{-15}.

Case 1: $a = 0.05$, which represents 5% of error in manufacturing each one of the potentials U_j. The algorithm described in Subsection 3.2 has been implemented by using the Sparse Grids Matlab kit 15.8 (see `http://csqi.epfl.ch` and [1]). The stopping criterion (31), with $\varepsilon = 10^{-7}$, is satisfied for $\bar{\ell} = 20$ and $\ell_{\text{opt}} = 15$, which corresponds to 895 collocation nodes in the random domain $\Gamma = [-0.05, 0.05]^4$. Figure 4 displays the rates of convergence of the isotropic sparse grid algorithm.

Table 2 shows the results, after convergence of the algorithm, for the expectation and the standard deviation of the cost functional $J(U, \omega)$ given by (6). As expected, the optimal design in mean ($\alpha = 0$) provides a solution which reduces the impact of manufacturing errors in comparison with the deterministic approach (see first and second rows in the first column of Table 2). It is also observed that increasing the value of the weighing parameter α reduces the dispersion of the computed designs. These results are more significant when the level of uncertainty increases, as pointed out in Table 3.

Case 2: $a = 0.2$, which corresponds to 20% of noise. The same procedure as in the preceding case has been applied. in this case, (31) is satisfied, with $\varepsilon = 10^{-2}$, for $\bar{\ell} = 20$ and $\ell = 16$, which corresponds to 1212 collocation nodes.

The same qualitative results as in the preceding case are observed in Table 3. However, since in this case the level of manufacturing uncertainties is higher, the differences of corresponding solutions are much more significant than in the preceding case. Indeed, for 20% of noise, the reduction in the mean value of $J(U, \cdot)$ for the mean optimal value ($\alpha = 0$), in comparison with the deterministic optimal design, is of the order of 39%. For the case of 5% of noise, this reduction
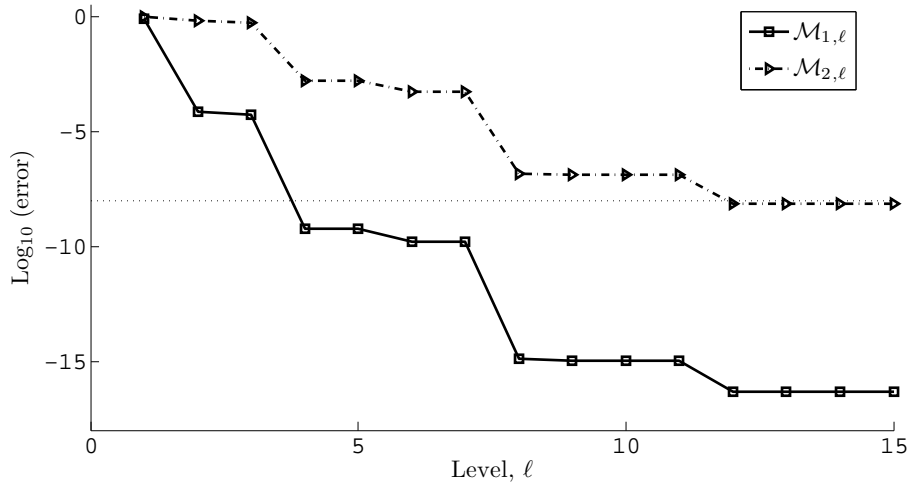
Figure 4: Case $a = 0.05$. Rates of convergence of the isotropic sparse grid algorithm. In the vertical axis, $\log_{10}$ (error) represents the two terms for relative error considered in (31). Precisely, the continuous line corresponds to relative error for the first order moment of $J(U, \cdot)$, and the dashed–dotted line is its second moment. The horizontal dashed line represents the prescribed accuracy level in the stopping criterion (31).

Table 2: Case $a = 0.05$. Mean *(first column)* and standard deviation *(second column)* of the functional (6) for the optimal deterministic design *(first row)*, the optimal design in mean *(second row)*, and the optimal design for $\alpha = 10^{12}$ *(third row)*.

Design	$\mathbb{E}(J(U, \cdot))$	std $(J(U, \cdot))$	U_1	U_2	U_3	U_4
Deterministic	1.737×10^{-11}	2.143×10^{-11}	0.70	1.31	1.54	0.70
$\alpha = 0$	1.664×10^{-11}	1.920×10^{-11}	0.95	0.90	1.60	0.70
$\alpha = 10^{12}$	1.908×10^{-11}	1.845×10^{-11}	1.14	0.83	1.05	0.97

Table 3: Case $a = 0.2$. Mean *(first column)* and standard deviation *(second column)* of the cost functional (6) for the optimal deterministic design *(first row)*, the optimal design in mean *(second row)*, and the optimal design for $\alpha = 10^{10}$ *(third row)*. Columns 3 to 6 show the corresponding optimal potentials.

Design	$\mathbb{E}(J(U, \cdot))$	std $(J(U, \cdot))$	U_1	U_2	U_3	U_4
Deterministic	3.48×10^{-10}	7.17×10^{-10}	0.70	1.31	1.54	0.70
$\alpha = 0$	2.12×10^{-10}	2.66×10^{-10}	1.19	0.82	1.21	0.92
$\alpha = 10^{10}$	2.35×10^{-10}	2.15×10^{-10}	1.20	0.88	1.07	1.05

is of the order of 4.2%. The decreasing if the standard deviation of the computed designs is also much more significant in the current case than in the case of 5% of manufacturing noise.

5   Conclusions

We have addressed the problem of determining optimal designs of nanoelectronic devices whose physical behavior is governed by the time–independent Schrödinger equation. Two situations are considered: A deterministic version of the problem, and the more realistic case where manufacturing uncertainties are accounted for. In both cases, the corresponding optimal design problems are formulated as the minimization of a least–squares performance metric. In the stochastic case, the variance of the random metric is incorporated in the cost functional as a measure of robustness. An explicit expression for that metric is computed using a semi–classical approximation. Having at our disposal an analytic expression for the cost functional has the following advantages:

(a) At the theoretical level, it is easily deduced that the cost functional is smooth (of regularity class C^∞) with respect to the design variables. As a consequence, the existence of solutions for both optimization problems is proved.

(b) In the deterministic case, an explicit expression for the gradient of the cost functional is obtained. Numerical simulations in Subsection 4.1 show the relevance of this issue, at least in the specific problem considered in this work, when gradient–based minimization algorithms are used as the numerical resolution method.

(c) In the stochastic optimization problem, it is also deduced that integrands, which appear in the considered cost functional, are analytic with respect to the random vector parameter. Thus, stochastic collocation methods (which possess an exponential rate convergence with respect to the number of sampling points and are, computationally, much more efficient than the classical brute-force Monte-Carlo method) are preferred.

The results obtained show an improvement of the accuracy in the linear response characteristic of about 69.4% over previous, brute–force, approaches. Moreover, the robustness of the design is manifest even under weight values of $\alpha = 10^{12}$ in the variance (physically, as seen in in Table 2, this would amount to pass from a design using Germanium, with a band gap of $\simeq 0.7$eV, to one using Silicon, whose band gap is $\simeq 1.1$eV, so that extreme case is still physically feasible).

As noticed along the paper, the semi–classical approach used in this work has the drawback that constraints the values of the applied bias voltage and those of the local energy potentials. Accordingly, the whole range of possible energies and potential profiles may be covered by using the approach proposed in this work in combination with appropriate connection formulas, for instance, the ones presented in [9].

References

[1] J. Bäck, F. Nobile, L. Tamellini and R. Tempone: *Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison.* In 'Spectral and High Order Methods for Partial Differential Equations'. Lecture Notes in Computational Science and Engineering. Vol 76, 43–66. Springer, 2011.

[2] C. Bergeron et al.: *Chemical vapor deposition of monolayer MoS_2 directly on ultrathin Al_2O_3 for low-power electronics*, Appl. Phys. Lett. **110** (2017) 053101.

[3] T. F. Coleman, Y. Li: *On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to bounds*, Mathematical Programming **67** (2) (1994) 189–224.

[4] T. F. Coleman, Y. Li: *An interior, thrust region approach for nonlinear minimization subject to bounds*, SIAM J. Optimization **6** (1996) 418–445.

[5] R. Gilmore: *Elementary Quantum Mechanics in One Dimension*, The Johns Hopkins University Press, 2004.

[6] C. Gong et al.: *Metal Contacts on Physical Vapor Deposited Monolayer MoS_2*, ACS Nano **7** (12) (2013) 11350–11357.

[7] W. W. Hager, Runge-Kutta methods in optimal control and the transformed adjoint system, Numer. Math., 87, 247-282 (2000).

[8] A. F. J. Levi, Applied Quantum Mechanics, Cambridge University Press, Cambridge, UK, 2006.

[9] A. F. J. Levi and I. G. Rosen, A novel formulation of the adjoint method in the optimal design of quantum electronic devices, SIAM J. Control Optim. 43(5), 3191-3223 (2010).

[10] V. P. Maslov and M. V. Fedoriuk, Semi-Classical Approximation in Quantum Mechanics, D. Reidel Publishing Company, 1981.

[11] P. Schmidt, S. Haas and A. F. J. Levi, Synthesis of electron transmission in nanoscale semiconductor devices, Applied Physics Letters 88, 013502 (2006).

[12] S. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Doklady Akademii Nauk SSSR, 4, 240-243 (1963).

[13] S. Wang, Fundamentals of Semiconductor Theory and Device Physics, Prentice-Hall, Englewood Cliffs, N.J., 1989.

[14] J. Zhang and R. Kosut, Robust design of quantum potential profile for electron transmission in semiconductor nanodevices, Proceedings of the European Control Conference, Kos, Greece, 2007.