J Comput Electron (2014) 13:753–762
DOI 10.1007/s10825-014-0597-5

Determining bound states in a semiconductor device with contacts using a nonlinear eigenvalue solver

William G. Vandenberghe · Massimo V. Fischetti · Roel Van Beeumen · Karl Meerbergen · Wim Michiels · Cedric Effenberger

Published online: 23 July 2014

Abstract We present a nonlinear eigenvalue solver enabling the calculation of bound solutions of the Schrödinger equation in a system with contacts. We discuss how the imposition of contacts leads to a nonlinear eigenvalue problem and discuss the numerics for a one- and two-dimensional potential. We reformulate the problem so that the eigenvalue problem can be efficiently solved by the recently proposal rational Krylov method for nonlinear eigenvalue problems, known as NLEIGS. In order to improve the convergence of the method, we propose a holomorphic extension such that we can easily deal with the branch points introduced by a square root. We use our method to determine the bound states of the one-dimensional Pöschl–Teller potential, a two-dimensional potential describing a particle in a canyon and the multi-band Hamiltonian of a topological insulator.

Keywords Bound states · Nonlinear eigenvalue problem · Contacts

1 Introduction

The study of electronic transport in semiconductor devices requires the simultaneous solution of the Poisson, Schrödinger, and transport equations. When collisions only have a minor impact on the determination of the device current, the ballistic picture is satisfactory. The current can thus be determined by solving the single-particle Schrödinger equation with "open" (transmitting) boundary conditions [1]. As a result, the imposition of open boundary conditions has become more or less standard practice when studying quantum transport and forms the basis for the widely used ballistic non-equilibrium Green's function formalism [2] or other open-boundary condition formalisms [3–5].

However, when imposing open boundary conditions almost all the attention has been focused on the extended states, *i.e.*, states with incoming components and very little attention has been paid to the bound states, *i.e.*, states with no incoming components. Furthermore, there is no general method available to determine bound states and bound states have sometimes been treated in a similar way as quasi-bound states while the latter are not bound but extended states. The focus of this paper is to provide a general method to determine the bound states in a system with contacts numerically. But in the following paragraphs, we elaborate on the importance of bound states and give an overview of previous work on bound states in literature.

The greatest practical importance of accounting for bound states lies in the determination of the charge density. Indeed, Frensley [6] demonstrated that determining the charge density by weighing the states at each contact with the corresponding Fermi-Dirac distribution can result in a non-physical charge density. And although one can point to the absence of interactions between the extended states and the bound states which would otherwise populate the bound states, the non-physical charge density fundamentally stems from the unjustified implicit assumption that all relevant states have at least one incoming component from a contact. Since a non-physical charge density precludes a self-

W. G. Vandenberghe (✉) · M. V. Fischetti
Department of Materials Science and Engineering,
University of Texas at Dallas, Richardson, TX 75080, USA
e-mail: wxv101020@utdallas.edu

R. Van Beeumen · K. Meerbergen · W. Michiels
Department of Computer Science, KU Leuven,
3001 Heverlee, Belgium

C. Effenberger
EPFL, Lausanne, Switzerland

consistent solution of the potential, ad-hoc solutions dealing with bound states were often employed to obtain convergence [3,7].

Mamaluy et al. [7] presented the contact-block reduction (CBR) method enabling the determination of the charge associated with the extended states in the open system from the solutions of the Schrödinger equation of the closed system. But since the CBR method only yields the charge associated with the extended states, Mamaluy et al. complement the charge of the extended states with the charge associated with a selected subset of bound states of the closed system. In this way, a self-consistent solution of Poisson and Schrödinger equations is possible but from a physical and mathematical point of view the picture is not entirely satisfactory as the bound states of the open and not those of the closed system should complement the extended states.

Taylor et al. [8] dealt with the ab initio modeling of quantum transport in molecular electronic devices. They recognize the presence of bound states but refrain from calculating the bound states since they consider their determination a very time-consuming task. Taylor et al. circumvent the problem of determining the bound states by assuming that the energies of the bound states are below the Fermi level of all contacts, determining the spectral functions of the complete system, and by performing a complex contour integration they avoid the poles of the Green's function due to bound states. Li et al. [9] extends the approach of Taylor et al. to bound states anywhere in the spectrum. Frensley and Bowen et al. [10–13] have previously presented methods to calculate the resonant states for a 1D system with contacts.

We note that the determination of the number of bound states of the Schrödinger equation has also drawn interest in the mathematical community. The problem is referred to as "bound states of the Schrödinger operator" and determining the limit on the number of bound states is one of the problems of interest [14,15]. In the condensed matter community, on the other hand, bound states have been shown to have very interesting physical properties. Bound states are responsible for the quantum Hall effect and the edge states in topological insulators are also bound states [16].

The objective of the paper is to determine bound states by using a recently proposed method for solving nonlinear eigenvalue problems [17]. This subspace method approximates the nonlinear matrix-valued function by polynomial or rational interpolation and then applies a rational Krylov algorithm to the linearization. In order to improve the convergence of the method, we perform a transformation of variables so that the interval of interest gets rid of branch points and branch cuts. The code to reproduce the experiments conducted in this paper is available from http://twr.cs.kuleuven.be/research/software/nleps/bound-states.html.

In the paper, we first discuss the discretization of the Schrödinger equation in one and two dimensions in the presence of contacts. In both cases, we obtain a nonlinear eigenvalue problem with branch points where the real eigenvalues are the energies and the eigenvectors are the wavefunctions of the bound states. Next, we put forward our numerical method to solve the nonlinear eigenvalue problem describing the bound states. Then, we calculate the energy and the wavefunctions of the bound states in a one-dimensional potential, a two-dimensional potential and a one-dimensional potential with a multiband Hamiltonian. We finish with our conclusion.

2 Bound states

We cast the Schrödinger eigenvalue problem in the form

$$\left[-\nabla^2 + U(\mathbf{r})\right]\psi(\mathbf{r}) = \lambda\psi(\mathbf{r}) \tag{1}$$

having used Rydberg units ($\hbar = 2m = 1$). $U(\mathbf{r})$ is the potential energy and we indicate the eigenvalues and eigenvectors as $(\lambda_i, \psi_i(\mathbf{r}))$. The eigenvalues λ will always be real since Eq. (1) is Hermitian.

We use λ instead of E for the energy since when dealing with the numerical aspects (Sect. 3), scalars are denoted by lowercase italic symbols (w), matrices are denoted with a uppercase italic symbols (A) while vectors are denoted as uppercase boldface letters ($\mathbf{X}$).

If the domain of the Schrödinger equation is infinite—as is the case in a device with contacts—the spectrum of the eigenvalues will be a combination of a continuous spectrum and a discrete spectrum. A well-known example of bound states in a system with infinite dimensions are the orbitals of the hydrogen atom. The states in the discrete spectrum are referred to as bound states.

In the one-dimensional case, it can be proved that if the domain is infinite and the potential energy has a limiting value $u(x) = 0$ for $|x| \to \infty$, at least one bound state will be found for $\lambda < 0$ if $u(x) < 0$ for any x [14]. Solving the higher dimensional Schrödinger equation, a criterion for the presence of bound states is no longer as easily determined. Nevertheless, the problem of finding bound states in the continuum is also relevant for the higher dimensional Schrödinger equation as bound states can naturally arise in any system with contacts.

We proceed with the conversion of the physical problem in a one- and a two-dimensional system with contacts to a numerical nonlinear eigenvalue problem.

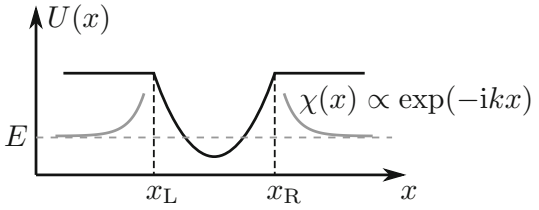

Fig. 1 Illustration of a one-dimensional potential with contacts starting at $x = x_\mathrm{L}$ and $x = x_\mathrm{R}$, the decaying nature of the wavefunction is illustrated in *grey*

2.1 One-dimensional potential

2.1.1 Transmitting boundary conditions

A one-dimensional system with a left and right contact is illustrated in Fig. 1 and has a potential energy

$$U(x) = \begin{cases} u_\mathrm{L} & x < x_\mathrm{L} \\ U_\mathrm{D}(x) & x_\mathrm{L} \le x \le x_\mathrm{R} \\ u_\mathrm{R} & x_\mathrm{R} < x \end{cases}. \tag{2}$$

The one-dimensional Schrödinger equation reads

$$\left[-\frac{\mathrm{d}^2}{\mathrm{d}x^2} + U(x)\right]\chi(x) = \lambda\chi(x) \tag{3}$$

and the homogeneous solutions inside the left and the right contact are

$$\chi^{\pm}_{\mathrm{L,R}}(x) = e^{\pm \mathrm{i}k_{\mathrm{L,R}}x} \tag{4}$$

where

$$k_{\mathrm{L,R}}(\lambda) = \sqrt{(\lambda - u_{\mathrm{L,R}})}. \tag{5}$$

To determine the bound states, we require that only decaying waves penetrate into the contact:

$$\chi(x) = \begin{cases} e^{-\mathrm{i}k_\mathrm{L}(\lambda_i)(x-x_\mathrm{L})}\chi(x_\mathrm{L}) & x < x_\mathrm{L} \\ e^{\mathrm{i}k_\mathrm{R}(\lambda_i)(x-x_\mathrm{R})}\chi(x_\mathrm{R}) & x_\mathrm{R} < x \end{cases} \tag{6}$$

where k_L and k_R are imaginary since $\lambda < \min(u_\mathrm{L}, u_\mathrm{R})$.

Depending on the numerical implementation, either Eq. (6) can be used immediately to determine the boundary conditions for the discretized Schrödinger equation. Or, Eq. (6) can be used to determine a boundary condition for the derivative at $x = x_{\mathrm{L,R}}$.

Because of the nonlinear relation between $k_{\mathrm{L,R}}$ and λ (Eq. 5), the boundary conditions for the Schrödinger equation will be nonlinear and the bound states will be the solution of a nonlinear eigenvalue problem. Furthermore, two (possibly coinciding) branch points will be introduced for $\lambda = u_{\mathrm{L,R}}$.

2.1.2 Numerical implementation

In order to determine the states numerically, we discretize $U(x)$ in the device region as a vector $\mathbf{U}$. We use a mesh with n_x points with fixed spacing Δx and we discretize the second derivative using the finite difference approximation. After discretization, the calculation of the bound states can be written as the nonlinear eigenvalue problem

$$\left[-D + \mathrm{diag}(\mathbf{U}) + \Sigma(\lambda) - \lambda\mathbb{1}\right]\mathbf{X} = 0 \tag{7}$$

where D is a matrix with $-2/\Delta x^2$ on its diagonal and $1/\Delta x^2$ on its sub- and superdiagonal

$$D = \frac{1}{\Delta x^2}\begin{bmatrix} -2 & 1 & 0 & \dots \\ 1 & -2 & 1 & \dots \\ 0 & 1 & -2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{8}$$

$\mathrm{diag}(\mathbf{U})$ is a matrix with $\mathbf{U}$ on its diagonal, $\Sigma(\lambda)$ is the matrix incorporating the transmitting boundary conditions and is determined by

$$\Sigma(\lambda) = \Sigma_\mathrm{L}(\lambda)\begin{bmatrix} 1 & 0 & \dots \\ 0 & 0 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} + \Sigma_\mathrm{R}(\lambda)\begin{bmatrix} \ddots & \vdots & \vdots \\ \dots & 0 & 0 \\ \dots & 0 & 1 \end{bmatrix}, \tag{9}$$

$$\Sigma_\mathrm{L}(\lambda) = -\frac{1}{\Delta x^2}e^{\mathrm{i}\Delta x\sqrt{\lambda - u_\mathrm{L}}} \tag{10}$$

$$\Sigma_\mathrm{R}(\lambda) = -\frac{1}{\Delta x^2}e^{\mathrm{i}\Delta x\sqrt{\lambda - u_\mathrm{R}}} \tag{11}$$

while $\mathbf{X}$ is the discretized version of the wavefunction $\chi(x)$.

For normalization purposes, the integral of $|\chi(x)|^2$ has to be calculated over the entire domain $]-\infty, \infty[$. For $x \in]-\infty, x_\mathrm{L}]$ and $x \in [x_\mathrm{R}, \infty[$, the integral can be performed analytically

$$\int_{-\infty}^{x_\mathrm{L}} \chi_i^*(x)\chi_i(x)\mathrm{d}x = \frac{|\chi(x_\mathrm{L})|^2}{2\sqrt{\lambda - u_\mathrm{L}}} \tag{12}$$

$$\int_{x_\mathrm{R}}^{\infty} \chi_i^*(x)\chi_i(x)\mathrm{d}x = \frac{|\chi(x_\mathrm{R})|^2}{2\sqrt{\lambda - u_\mathrm{R}}} \tag{13}$$

while the integral for $x \in [x_\mathrm{L}, x_\mathrm{R}]$ must be performed numerically, for example using the trapezoidal rule

$$\int_{x_\mathrm{R}}^{x_\mathrm{L}} \chi_i^*(x)\chi_i(x)\mathrm{d}x \approx \mathbf{X}^\dagger \mathrm{diag}(\Delta x[\tfrac{1}{2}, 1, \dots, 1, \tfrac{1}{2}])\mathbf{X}. \tag{14}$$

The integral of $|\chi(x)|^2$ of the entire domain is then the sum of Eqs. (12–14), where $\chi(x_{\mathrm{L,R}})$ are the first and last element of $\mathbf{X}$.

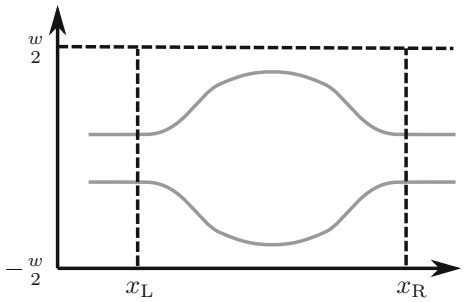

Fig. 2 Illustration of a two-dimensional potential with contacts starting at $x = x_{\mathrm{L}}$ and $x = x_{\mathrm{R}}$

2.2 Two-dimensional potential

2.2.1 Transmitting boundary conditions

To study the bound states in a two-dimensional potential, we study a rectangular structure with a contact on the left ($x < x_{\mathrm{L}}$) and right ($x > x_{\mathrm{R}}$) and the device region extending from $z = -w/2$ to $z = w/2$ with boundary condition $\chi(x, -w/2) = \chi(x, w/2) = 0$, as illustrated in Fig. 2. The two-dimensional potential in such a structure is given by

$$U(x,z) = \begin{cases} U_{\mathrm{L}}(z) & x < x_{\mathrm{L}} \\ U_{\mathrm{D}}(x,z) & x_{\mathrm{L}} \le x \le x_{\mathrm{R}} \\ U_{\mathrm{R}}(z) & x_{\mathrm{R}} < x \end{cases} \tag{15}$$

and the two-dimensional Schrödinger equation reads

$$\left(-\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{\mathrm{d}^2}{\mathrm{d}z^2}\right) + U(x,z)\right)\chi(x,z) = \lambda\chi(x,z). \tag{16}$$

In the contacts, the subband energies and wavefunctions $\phi_j(z)$ must be determined as the solutions of the one-dimensional Schrödinger equation

$$\left[-\frac{\mathrm{d}}{\mathrm{d}z^2} + U_{\mathrm{L,R}}(z)\right]\phi_{\mathrm{L,R}\,j}(z) = w_{\mathrm{R,L}\,j}\phi_{\mathrm{R,L}\,j}(z). \tag{17}$$

where $w_{\mathrm{L,R}}$ are the subband eigenenergies. The transmitting boundary conditions for the two-dimensional case become

$$\chi(x,z) = \sum_j \phi_{\mathrm{R,L}\,j}(z) e^{\mathrm{i}\sqrt{\lambda - w_{\mathrm{L,R}\,j}}x} \int \mathrm{d}z \phi^*_{\mathrm{R,L}\,j}(z)\chi(x_{\mathrm{L,R}}, z) \tag{18}$$

assuming

$$\int_0^w \mathrm{d}z \phi^*_{\mathrm{L,R}\,j}(z)\phi_{\mathrm{L,R}\,j}(z) = 1. \tag{19}$$

2.2.2 Numerical implementation

In two dimensions, we use a square mesh with $n_x n_z$ points and fixed spacing Δx and Δz. We extend our finite difference approximation to two dimensions by taking the tensor product of the second derivative in x and z direction with the identity matrix in the perpendicular direction.

The two-dimensional nonlinear eigenvalue problem yielding the two-dimensional bound states now reads

$$\Big(-D_{xx}\otimes \mathbb{1}_z - \mathbb{1}_x \otimes D_{zz} + \mathrm{diag}(\mathbf{U}) + \Sigma(\lambda) - \lambda \mathbb{1}_x \otimes \mathbb{1}_z\Big)\mathbf{X} = 0 \tag{20}$$

where $A \otimes B$ denotes the tensor product of matrices A and B. $\mathbb{1}_z$ and $\mathbb{1}_x$ denote the identity matrix of dimension n_z and n_x respectively. D_{xx} and D_{zz} denote the second derivative in the finite difference approximation and the contact terms are given by

$$\Sigma(\lambda) = \begin{bmatrix} 1 & 0 & \dots \\ 0 & 0 & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \otimes \Sigma_{\mathrm{L}}(\lambda) + \begin{bmatrix} \ddots & \vdots & \vdots \\ \dots & 0 & 0 \\ \dots & 0 & 1 \end{bmatrix} \otimes \Sigma_{\mathrm{R}}(\lambda)$$

with

$$\Sigma_{\mathrm{L,R}}(\lambda) = -\frac{1}{\Delta x^2} Y_{\mathrm{L,R}} \mathrm{diag}\left(e^{\mathrm{i}\sqrt{\lambda - \mathbf{W}_{\mathrm{L,R}}}\Delta x}\right) Y^{-1}_{\mathrm{L,R}} \tag{21}$$

where the columns of $Y_{\mathrm{L,R}}$ are the discretized subband wavefunctions $\mathbf{Y}_{\mathrm{L,R}}$ which are determined by performing the eigendecomposition of the Hamiltonian in the contacts

$$\left(-D_{zz} + \mathrm{diag}(\mathbf{U}_{\mathrm{L,R}})\right)\mathbf{Y}_{\mathrm{L,R}} = w_{\mathrm{L,R}}\mathbf{Y}_{\mathrm{L,R}}. \tag{22}$$

$\mathbf{W}_{\mathrm{L,R}}$ is the vector containing the subband eigenenergies $w_{\mathrm{L,R}}$.

For two dimensions the nonlinear eigenvalue problem has n_z branch points and bound states can be found between two branch points as will be shown in Sect. 4.

3 Nonlinear eigenvalue problem

We define the nonlinear eigenvalue problem as follows

$$A(\lambda)\mathbf{X} = \left(\sum_{i=1}^{k} B_i f_i(\lambda)\right)\mathbf{X} = 0, \tag{23}$$

where $A : \mathbb{C} \to \mathbb{C}^{n\times n}$, $B_i \in \mathbb{C}^{n\times n}$ are constant matrices, f_i are scalar functions of λ, and $k \le n^2$. Note that the nonlinear eigenvalue problem (23) is nonlinear in the eigenvalue λ but linear in the eigenvector $\mathbf{X} \in \mathbb{C}^n \backslash \{0\}$. Therefore, the number of eigenvalues is unbounded in general.

The bound states satisfy

$$A(\lambda)\mathbf{X} = 0, \tag{24}$$

where $\mathbf{X}$ is the eigenvector and

$$A(\lambda) = H - \lambda\mathbb{1} + \Sigma(\lambda), \tag{25}$$

with H a constant matrix. Note that Eq. (7) is applicable in the one-dimensional case and Eq. (20) in the two-dimensional case.

For solving the problem (24)–(25), we cannot use standard techniques for linear eigenvalue problems. Therefore, we discuss a method in the next section for computing simultaneously several eigenvalues of the nonlinear eigenvalue problem (24)–(25).

3.1 Rational Krylov method (NLEIGS)

In the recent literature we find several subspace based nonlinear eigensolvers [17–20]. This type of methods have the advantage that they are able to compute several eigenpairs at once. Therefore, they are more reliable than local methods, such as Newton's method.

In this paper, we will use the (fully) rational Krylov method [17,19], abbreviated as NLEIGS, since this method also applies to problems with eigenvalues nearby singularities. NLEIGS is based on rational interpolation and dynamically constructs a rational interpolant of the nonlinear matrix-valued function $A(\lambda)$. This results in a rational eigenvalue problem which is solved by using a companion-type structured linearization.

The NLEIGS software can both deal with nonlinear eigenvalue problems where $A(\lambda)$ is given as a function as well as where $A(\lambda)$ is expressed as a summation of constant matrices B_i times scalar (non)linear functions $f_i(\lambda)$, such as in Eq. (23). The latter is favourable since in this case we only need to approximate scalar functions. Moreover, the possible low rank structure of the matrices B_i can be exploited, which leads to a reduced memory cost.

One of the main requirements for subspace methods is that the functions $f_i(\lambda)$ in (23) are analytic in the region of interest in the complex plane. The following section explains how this requirement can be fulfilled for computing real eigenvalues of the nonlinear eigenvalue problem (24)–(25).

3.2 Holomorphic extension

We first rewrite (25) in the notation of (23)

$$A(\lambda) = H - \lambda\mathbb{1} + \sum_{i=0}^{k} \exp\left(\mathrm{i}\sqrt{\lambda - s_i}\right) S_i, \tag{26}$$

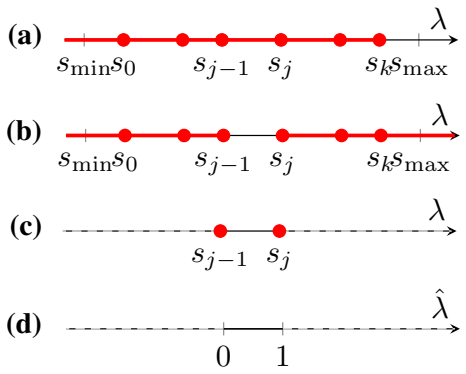

Fig. 3 Graphical illustration of holomorphic extension

where $\sqrt{\cdot}$ denotes the principal branch of the complex square root, defined by

$$\sqrt{z} = \begin{cases} \sqrt{|z|} \cdot \exp\left(\frac{1}{2}\mathrm{i}\arg(z)\right), & z \neq 0 \\ 0, & z = 0 \end{cases},$$

with $\arg(z) \in (-\pi, \pi]$ denoting the argument of the complex number z. Moreover, H and $S_0, \ldots, S_k$ are given constant matrices, and $s_0 < \cdots < s_k$ are given real numbers. However, to apply NLEIGS to the problem (24),(26), a holomorphic (also known as analytic) extension of $A(\lambda)$ to a neighborhood of the real axis is required. Unfortunately, such an extension does not exist in the vicinity of the branch points $s_0, \ldots, s_k$. To overcome this obstacle, we will confine ourselves to the subproblem of finding all eigenvalues inside the real interval $[s_{j-1}, s_j]$ for some $j \in \{1, \ldots, k\}$. The remaining intervals $[s_{\min}, s_0)$ and (s_k, ∞) will be discussed further.

A graphical illustration of the holomorphic extension of the interval $[s_{j-1}, s_j]$ is given in Fig. 3. In this figure, the red dots and lines represent the branch points and branch cuts, respectively.

The standard choice for the branch cuts in Eq. (26) is illustrated in case (a). Next in case (b), the branch cuts corresponding to the branch points $s_j, \ldots, s_k$ are flipped in order to make the interval $[s_{j-1}, s_j]$ branch cut free. Then in case (c), we restrict the eigenvalue computation problem to the subproblem on this interval. Note that it still contains branch cuts on the endpoints. Finally in case (d), we map the interval $[s_{j-1}, s_j]$ to the interval $[0, 1]$ by a transformation from λ to $\hat{\lambda}$ which eliminates the branch points. We now discuss this holomorphic extension in more detail.

For $\lambda \in [s_{j-1}, s_j]$, $A(\lambda)$ in (26) can be rewritten as

$$\begin{aligned} A(\lambda) = H - \lambda\mathbb{1} &+ \sum_{i=0}^{j-1} \exp\left(\mathrm{i}\sqrt{\lambda - s_i}\right) S_i \\ &+ \sum_{i=j}^{k} \exp\left(-\sqrt{s_i - \lambda}\right) S_i, \end{aligned} \tag{27}$$

Note that for $\lambda \notin [s_{j-1}, s_j]$ the formulations (26) and (27) may disagree. But since we are only interested in real eigenvalues, this discrepancy is irrelevant to us. The advantage of formulation (27) over the original formulation is that it does not have a branch cut within the open interval (s_{j-1}, s_j). However, the branch points at s_{j-1} and s_j still inhibit a holomorphic extension to a neighborhood of the interval. To eliminate these branch points, we reparameterize the problem. Setting

$$\lambda = \Delta s_j \sin^2\left(\tfrac{\pi}{2}\hat{\lambda}\right) + s_{j-1}, \quad \Delta s_j = s_j - s_{j-1}, \tag{28}$$

we obtain the equivalent nonlinear eigenvalue problem

$$A_j(\hat{\lambda})\mathbf{X} = 0 \tag{29}$$

with the eigenvalue parameter $\hat{\lambda} \in [0, 1]$ and

$$\begin{aligned} A_j(\hat{\lambda}) = H &- \left(\Delta s_j \sin^2\left(\tfrac{\pi}{2}\hat{\lambda}\right) + s_{j-1}\right)\mathbb{1} \\ &+ \sum_{i=0}^{j-2} \exp\left(\mathrm{i}\sqrt{\Delta s_j \sin^2\left(\tfrac{\pi}{2}\hat{\lambda}\right) + s_{j-1} - s_i}\right) S_i \\ &+ \exp\left(\mathrm{i}\sqrt{\Delta s_j}\sin\left(\tfrac{\pi}{2}\hat{\lambda}\right)\right) S_{j-1} \\ &+ \exp\left(-\sqrt{\Delta s_j}\cos\left(\tfrac{\pi}{2}\hat{\lambda}\right)\right) S_j \\ &+ \sum_{i=j+1}^{k} \exp\left(-\sqrt{s_i - s_{j-1} - \Delta s_j \sin^2\left(\tfrac{\pi}{2}\hat{\lambda}\right)}\right) S_i. \end{aligned} \tag{30}$$

Theorem 1 *The nonlinear eigenvalue problems* (24), (26) *and* (29)–(30) *are equivalent in the sense that* $\hat{\lambda} \in [0, 1]$ *is an eigenvalue of* (29)–(30) *if and only if* $\lambda \in [s_{j-1}, s_j]$ *given by* (28) *is an eigenvalue of* (24), (26).

Proof Let $\hat{\lambda} \in [0, 1]$ be arbitrary and let λ be given by (28). Then, we have $\sin\left(\frac{\pi}{2}\hat{\lambda}\right), \cos\left(\frac{\pi}{2}\hat{\lambda}\right) \in [0, 1]$, implying that $\lambda \in [s_{j-1}, s_j]$ and

$$\sqrt{\lambda - s_{j-1}} = \sqrt{\Delta s_j}\sin\left(\tfrac{\pi}{2}\hat{\lambda}\right),$$

$$\sqrt{s_j - \lambda} = \sqrt{\Delta s_j\left(1 - \sin^2\left(\tfrac{\pi}{2}\hat{\lambda}\right)\right)} = \sqrt{\Delta s_j}\cos\left(\tfrac{\pi}{2}\hat{\lambda}\right).$$

Moreover, for $i = j, \ldots, k$,

$$\mathrm{i}\sqrt{\lambda - s_i} = -\sqrt{s_i - \lambda}$$

since $\lambda - s_i \leq 0$. Combining the above identities yields $A(\lambda) = A_j(\hat{\lambda})$, from which the claim follows.

To compute eigenvalues in the interval $[s_{\min}, s_0)$, a similar transformation may be employed. To eliminate the branch point at s_0, we substitute

$$\lambda = s_0 - \Delta s_0 \hat{\lambda}^2, \quad \Delta s_0 = s_0 - s_{\min}, \tag{31}$$

leading to the equivalent nonlinear eigenvalue problem

$$A_0(\hat{\lambda})\mathbf{X} = 0 \tag{32}$$

with the eigenvalue parameter $\hat{\lambda} \in [0, 1]$ and

$$\begin{aligned} A_0(\hat{\lambda}) = H &- \left(s_0 - \Delta s_0 \hat{\lambda}^2\right)\mathbb{1} \\ &+ \exp\left(-\sqrt{\Delta s_0}\hat{\lambda}\right) S_0 \\ &+ \sum_{i=1}^{k} \exp\left(-\sqrt{s_i - s_0 + \Delta s_0 \hat{\lambda}^2}\right) S_i. \end{aligned} \tag{33}$$

Theorem 2 *The nonlinear eigenvalue problems* (24), (26) *and* (32)–(33) *are equivalent in the sense that* $\hat{\lambda} \in [0, 1]$ *is an eigenvalue of* (32)–(33) *if and only if* $\lambda \in [s_{\min}, s_0]$ *given by* (31) *is an eigenvalue of* (24), (26).

Proof Let $\hat{\lambda} \in [0, 1]$ be arbitrary and let λ be given by (31). Then, one readily verifies that $\lambda \in [s_{\min}, s_0]$ and

$$-\sqrt{s_0 - \lambda} = -\sqrt{\Delta s_0}\hat{\lambda}.$$

Moreover, for $i = 1, \ldots, k$,

$$\mathrm{i}\sqrt{\lambda - s_i} = -\sqrt{s_i - \lambda}$$

since $\lambda - s_i \leq 0$. The proof is completed by noting that these identities imply $A(\lambda) = A_0(\hat{\lambda})$.

To compute eigenvalues in the interval (s_k, ∞), again a similar transformation may be employed. Before presenting the details, we reduce the unbounded interval (s_k, ∞) to some finite interval $(s_k, s_{\max}]$ using the following estimate.

Theorem 3 *Let* $\lambda \geq s_k$ *be a real eigenvalue of the nonlinear eigenvalue problem* (24),(26). *Then,* $\lambda \leq s_{\max}$, *where, for any matrix norm* $\|\cdot\|$,

$$s_{\max} = \|H\| + \sum_{i=0}^{k} \|S_i\|. \tag{34}$$

Proof By definition of the matrix-valued function A in (26), λ is an eigenvalue of (24) if and only if it is an eigenvalue of the matrix

$$A_\lambda = H + \sum_{i=0}^{k} \exp\left(\mathrm{i}\sqrt{\lambda - s_i}\right) S_i.$$

Moreover, $\lambda \geq s_k$ implies that $\sqrt{\lambda - s_i}$ is real for all $i = 0, \ldots, k$. Consequently, $\left|\exp\left(\mathrm{i}\sqrt{\lambda - s_i}\right)\right| = 1, i = 0, \ldots, k$, and

$$\lambda \leq \|A_\lambda\| \leq \|H\| + \sum_{i=0}^{k} \left|\exp\left(\mathrm{i}\sqrt{\lambda - s_i}\right)\right| \|S_i\| = s_{\max}.$$

To eliminate the branch point at s_k, we substitute

$$\lambda = \Delta s_{k+1}\hat{\lambda}^2 + s_k, \quad \Delta s_{k+1} = s_{\max} - s_k, \tag{35}$$

leading to the equivalent nonlinear eigenvalue problem

$$A_{k+1}(\hat{\lambda})\mathbf{X} = 0 \tag{36}$$

with the eigenvalue parameter $\hat{\lambda} \in [0, 1]$ and

$$A_{k+1}(\hat{\lambda}) = H - \left(\Delta s_{k+1}\hat{\lambda}^2 + s_k\right)\mathbb{1} + \sum_{i=0}^{k-1} \exp\left(\mathrm{i}\sqrt{\Delta s_{k+1}\hat{\lambda}^2 + s_k - s_i}\right) S_i + \exp\left(\mathrm{i}\sqrt{\Delta s_{k+1}}\hat{\lambda}\right) S_k. \tag{37}$$

Theorem 4 *The nonlinear eigenvalue problems* (24), (26) *and* (36)–(37) *are equivalent in the sense that* $\hat{\lambda} \in [0, 1]$ *is an eigenvalue of* (36)–(37) *if and only if* $\lambda \in [s_k, s_{\max}]$ *given by* (35) *is an eigenvalue of* (24), (26).

Proof Let $\hat{\lambda} \in [0, 1]$ be arbitrary and let λ be given by (35). Then, one readily verifies that $\lambda \in [s_k, s_{\max}]$ and

$$\mathrm{i}\sqrt{\lambda - s_k} = \mathrm{i}\sqrt{\Delta s_{k+1}}\hat{\lambda}.$$

The proof is completed by noting that the latter identity implies $A(\lambda) = A_{k+1}(\hat{\lambda})$.

In this section we showed that the transformed eigenvalue problems (29)–(30), (32)–(33), and (36)–(37) possess holomorphic extensions to a neighbourhood of the intervals $[s_{j-1}, s_j]$, $j \in \{1, \ldots, k\}$, $[s_{\min}, s_0]$, and $[s_k, s_{\max}]$, respectively. Thus, we can solve these problems and consequently also the nonlinear eigenvalue problem (24), (26) by using NLEIGS.

4 Results

In this section we show the application of our method to three potentials which have bound states. A one-dimensional effective mass case: (i) the Pöschl–Teller potential, (ii) a two-dimensional effective mass case: a particle in a canyon and (iii) a one-dimensional multiband case: a topological insulator.

The experiments can be reproduced with the MATLAB code available from http://twr.cs.kuleuven.be/research/software/nleps/bound-states.html.

4.1 One-dimensional: Pöschl–Teller potential

As a first test for our method to determine bound states, we calculate the (known) bound states in the case of the Pöschl–Teller potential [21]

$$U(x) = -\frac{\alpha^2\lambda(\lambda - 1)}{\cosh^2(\alpha x)} \tag{38}$$

illustrated in Fig. 4. There are $\lfloor\lambda\rfloor$ solutions with an energy eigenvalue

$$-\alpha^2(\lambda - n)^2 \quad n \leq \lambda \wedge n \in \mathcal{N} \tag{39}$$

where $\mathcal{N}$ denotes the set of natural numbers.

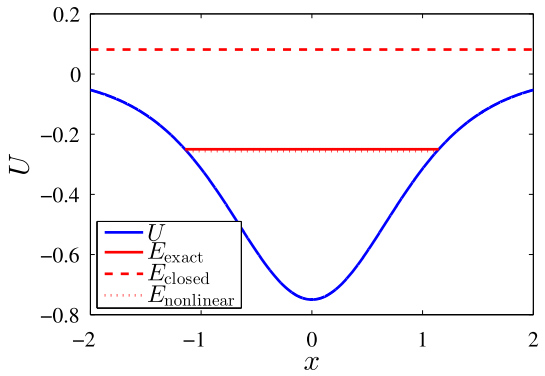

Fig. 4 Illustration of the Pöschl–Teller potential for $\lambda_P = 1.5$ and $\alpha = 1$. The position over which the potential is solved extends from $x = -2$ to $x = 2$. Imposing transmitting boundary conditions and solving the resulting nonlinear eigenvalue problem results in a good estimate (*dotted line*, almost coinciding with the exact eigenvalue) of the exact eigenvalue of the Pöschl–Teller potential (*solid line*) while solving the closed system results in a positive eigenvalue (*dashed line*)

Performing the determination of the energy of the bound state for $\alpha = 1$ and $\lambda = 1.5$ on the region extending from -2 to 2 using a discretization of $\Delta x = .001$, we obtain an eigenvalue for the bound state of -0.2564 while the exact value is -0.25. The difference between the exact eigenvalue and the calculated eigenvalue stems from the assumption that the potential in the contact is uniform, which is only an approximation in the case of the Pöschl–Teller potential. Doubling the region and the number of mesh points so it extends from -4 to 4, the exact eigenvalue is retained up to four digits.

Determining the bound state using the same potential but now closing the system and imposing Dirichlet boundary conditions, the energy of the lowest lying state is 0.0814 and doubling the size improves the estimate to -0.22306, doubling once more improves the estimate of the closed system to -0.24957.

The imposition of transmitting boundary conditions thus yields a good estimate of the eigenvalues of the Pöschl–Teller potential which can only be matched by solving the closed eigenvalue problem on a much larger domain.

4.2 Two-dimensional: particle in a canyon

To test our method for a two-dimensional potential, we opt for a potential describing a particle in a canyon

$$U(x, z) = \begin{cases} -U_0\theta(w_1/2 - |z|) & \mathrm{abs}(x) > l/2 \\ -U_0\theta(w_2/2 - |z|) & \mathrm{abs}(x) \leq l/2 \end{cases} \tag{40}$$

where U_0 is the depth of the canyon, $\theta(x)$ is the step function and w_1 and w_2 are the width of the contact and the width of the canyon as illustrated in Fig. 5. The confinement inside the contacts is stronger than the confinement in the center of

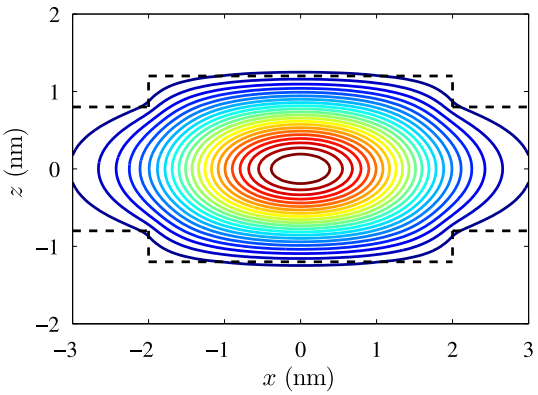

Fig. 5 Illustration of the canyon potential (*dashed line*) and contour plot of the squared amplitude of the first bound state in the canyon potential. Dimensions of the simulated region are $x_R = -x_L = 3\,\text{nm}$, $w = 4\,\text{nm}$ and the canyon width is determined by $w_1 = 1.6\,\text{nm}$, $w_2 = 2.4\,\text{nm}$, the canyon depth is $U_0 = 3\,\text{eV}$ and the electron mass is $m^* = .2\,m_0$

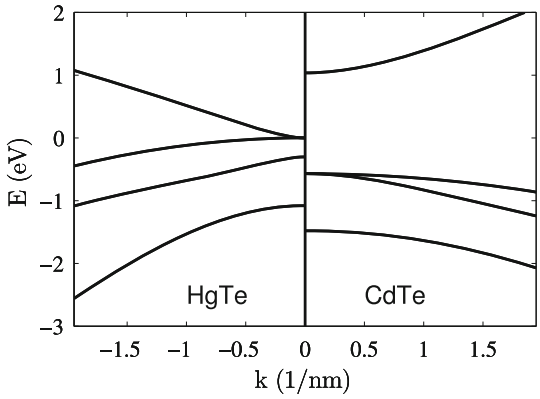

Fig. 6 The HgTe and CdTe band structure obtained from the Hamiltonian presented in Ref. [24]. HgTe is a semimetal while CdTe is a trivial insulator

the device. This will result in the presence of one or more bound states inside the device.

For $x_R = -x_L = 3\,\text{nm}$, $w = 4\,\text{nm}$, $w_1 = 1.6\,\text{nm}$, $w_2 = 2.4\,\text{nm}$, $U_0 = 3\,\text{eV}$ and $m^* = .2\,m_0$, the first two branch points are located at -2.561 and $-1.329\,\text{eV}$ while the first five bound states have an energy of -2.714, -2.581, -2.017, -1.788 and $-1.450\,\text{eV}$. The energies are calculated up to an accuracy of more 11 digits. The wavefunction of the first bound state is illustrated in Fig. 5.

For the one-dimensional potential, the energy of the bound states could also be calculated by imposing closed boundary conditions and enlarging the device. For the two-dimensional case however, a similar approach is problematic as there is no straightforward criterion to distinguishing the bound states from the states introduced by imposing closed boundary conditions.

4.3 Multi-band: topological insulator

In the previous two examples an effective mass Hamiltonian was used and as a result, the scalar functions from Eq. (23) introducing the non-linearity ($f_i(\lambda)$) were available in closed form. In general however, the energy dispersion is not available in closed form and the matrix imposing the transmitting boundary conditions has to be computed numerically for each energy. To demonstrate that our method can deal with these cases as well, we apply our method to a topological insulator structure described by a multi-band Hamiltonian.

The topological insulator phase is a material phase which was recently discovered [16] and some CdTe/HgTe heterostructures turn out to be topological insulators [22,23]. CdTe has a direct bandgap and its bandstructure is topologically classified as a trivial insulator while HgTe is a semi-metal as shown in Fig. 6. Sandwiching a layer of

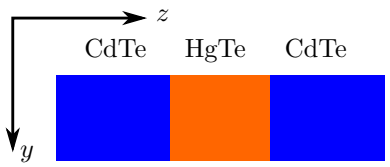

Fig. 7 Illustration of a CdTe/HgTe sandwich, the sandwich becomes a two-dimensional topological insulator when the HgTe width exceeds 6.7 nm

HgTe between two layers of CdTe as illustrated in Fig. 7, the bandgap decreases from the bulk CdTe bandgap with increasing HgTe layer thickness until a critical thickness of the HgTe layer is reached at which the bandgap vanishes. Subsequently increasing the HgTe layer thickness creates an inverted bandgap and a two-dimensional topological insulator. The transition from a normal band ordering to an inverted band ordering signals a topological transition.

Labeling $x - y$ as the direction parallel to the CdTe/HgTe interface and z as the direction perpendicular to the CdTe/HgTe interface. The CdTe/HgTe sandwich does not have translational symmetry in the z direction but the sandwich can be modeled as a system with contacts. The bandstructure of the bound states can be determined from a nonlinear eigenvalue problem similar to the one from previous section.

Because spin-orbit coupling is responsible for the non-trivial nature of the band structure of the CdTe/HgTe heterostructure, the inclusion of spin-orbit coupling is essential to determine the transition from trivial to topological insulator. An eight-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian describing $Cd_{1-x}Hg_xTe$ is given in Ref. [24]. Because k_x does not commute with the Hg concentration, we opt for a finite element implementation rather than the finite difference implementation used in previous sections.

In Fig. 8, we show the envelope functions corresponding to the first conduction and valence band for a sandwich with a 4 and a 8 nm thick HgTe layer. The envelope functions for

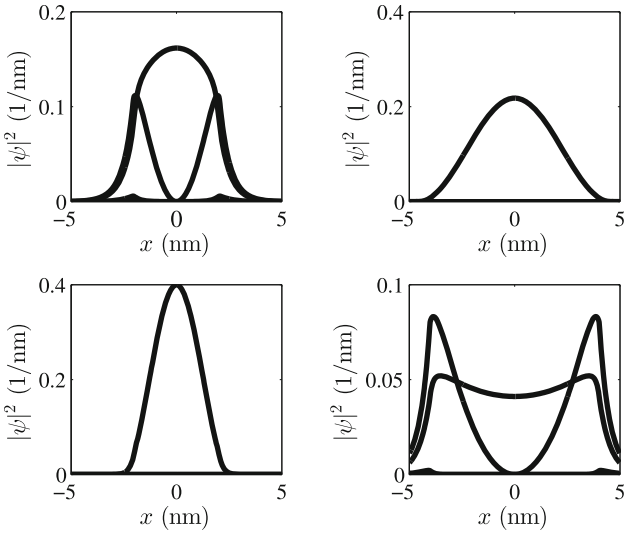

Fig. 8 $\mathbf{k}\cdot\mathbf{p}$ envelope functions of the first valence (*bottom*) and conduction band (*top*) for 4 nm (*left*) and 8 nm (*right*) sandwich

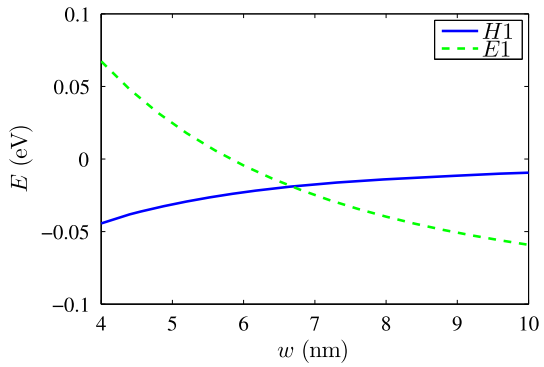

Fig. 9 First conduction band and valence band as a function of HgTe width. The label $E1$ indicates the "electron-like" band with the lowest energy and $H1$ the "heavy hole-like" band with the highest energy. The transition from topological to trivial insulator occurs at a thickness of 6.7 nm

the valence band and conduction band are switched, *i.e.*, the conduction/valence band wavefunction of the 4 nm sandwich resemble the valence/conduction band wavefunctions of the 8 nm sandwich. In Fig. 9, the valence band and the conduction band are calculated illustrating the topological transition at the critical thickness (6.7 nm).

The successful determination of the bound states in the CdTe/HgTe sandwich shows that our method to find bound states is not limited to effective mass Hamiltonians but works equally well for multi-band Hamiltonians where no closed form is available for the energy dispersion in the contacts.

5 Conclusion

We have presented the nonlinear eigenvalue problem that arises when we consider the determination of bound states in a one- and two-dimensional device with contacts. We have shown the nonlinear eigenvalue has two branch points in the one-dimensional case and a large number of branch points in the two-dimensional case. To solve the nonlinear eigenvalue problem, we have introduced a nonlinear eigenvalue solver based on a state-of-the-art rational Krylov method using a holomorphic extension to eliminate the branch points. We have used the nonlinear eigenvalue solver to obtain a solution to the problem for three different cases: a one-dimensional case presented by the Pöschl–Teller potential, a two-dimensional potential by a particle in a canyon and, finally, a topological insulator. We have shown that in each of these cases the bound states can be computed in a systematic and accurate way even if the energy of the bound states lies between two branch points.

References

1. Lent, C.S., Kirkner, D.J.: The quantum transmitting boundary method. J. Appl. Phys. **67**(10), 6353–6359 (1990). doi:10.1063/1.345156. ISSN 0021–8979.
2. Datta, S.: Electronic Transport in Mesoscopic Systems. Cambridge university press, Cambridge (1997)
3. Laux, S.E., Kumar, A., Fischetti, M.V.: Ballistic fet modeling using QDAME: quantum device analysis by modal evaluation. IEEE Trans. Nanotechnol. **1**(4), 255–259 (2002). doi:10.1109/TNANO.2002.807388. ISSN 1536–125X.
4. Polizzi, E., Ben Abdallah, N.: Self-consistent three dimensional models for quantum ballistic transport in open systems. Phys. Rev. B **66**(24), 245301–245301 (2002)
5. Vandenberghe, W., Sorée, B., Magnus, W., Groeseneken, G.: Zener tunneling in semiconductors under nonuniform electric fields. J. Appl. Phys. **107**(5), 054520 (2010)
6. Frensley, W.R.: Boundary conditions for open quantum systems driven far from equilibrium. Rev. Mod. Phys. **62**, 745–791 (1990). doi:10.1103/RevModPhys.62.745
7. Mamaluy, D., Vasileska, D., Sabathil, M., Zibold, T., Vogl, P.: Contact block reduction method for ballistic transport and carrier densities of open nanostructures. Phys. Rev. B **71**, 245321 (2005). doi:10.1103/PhysRevB.71.245321
8. Taylor, J., Guo, H., Wang, J.: Ab initio modeling of quantum transport properties of molecular electronic devices. Phys. Rev. B **63**, 245407 (2001). doi:10.1103/PhysRevB.63.245407
9. Li, R., Zhang, J., Hou, S., Qian, Z., Shen, Z., Zhao, X., Xue, Z.: A corrected negf-dft approach for calculating electronic transport through molecular devices: filling bound states and patching the non-equilibrium integration. Chem. Phys. **336**(2–3), 127–135 (2007). doi:10.1016/j.chemphys.2007.06.011. ISSN 0301–0104.
10. Frensley, W.R.: Numerical evaluation of resonant states. Superlattices Microstr. **11**(3), 347–350 (1992)
11. Bowen, R.C., Frensley, W.R., Klimeck, G., Lake, R.K.: Transmission resonances and zeros in multiband models. Phys. Rev. B **52**(4), 2754 (1995)
12. Klimeck, G., Lake, R., Chris Bowen, R., Frensley, W.R., Moise, T.S.: Quantum device simulation with a generalized tunneling formula. Appl. Phys. Lett. **67**(17), 2539–2541 (1995)
13. Karner, M., Gehring, A., Kosina, H.: Efficient calculation of lifetime based direct tunneling through stacked dielectrics. J. Comput. Electron. **5**(2–3), 161–165 (2006)

14. Blanchard, P., Stubbe, J.: Bound states for schrödinger Hamiltonians: phase space methods and applications. Rev. Math. Phys. **08**(04), 503–547 (1996). doi:10.1142/S0129055X96000172
15. Gustafson, S.J., Sigal, I.M.: Mathematical Concepts of Quantum Mechanics. Universitext (En ligne). Springer, Berlin (2011). ISBN 9783642218668
16. Hasan, M.Z., Kane, C.L.: Colloquium: topological insulators. Rev. Mod. Phys **82**, 3045–3067 (2010). doi:10.1103/RevModPhys.82.3045
17. Güttel, S., Van Beeumen, R., Meerbergen, K., Michiels, W.: NLEIGS: a class of robust fully rational Krylov methods for nonlinear eigenvalue problems. Technical Report TW633, Department of Computer Science, KU Leuven, Leuven (2013)
18. Effenberger, C.: Robust successive computation of eigenpairs for nonlinear eigenvalue problems. SIAM J. Matrix Anal. Appl. **34**(3), 1231–1256 (2013)
19. Van Beeumen, R., Meerbergen, K., Michiels, W.: A rational Krylov method based on Hermite interpolation for nonlinear Eigenvalue problems. SIAM J. Sci. Comput. **35**(1), A327–A350 (2013)
20. Jarlebring, E., Michiels, W., Meerbergen, K.: A linear eigenvalue algorithm for the nonlinear eigenvalue problem. Numerische Mathematik **122**(1), 169–195 (2012). doi:10.1007/s00211-012-0453-0. ISSN 0029–599X.
21. Flügge, S.: Practical Quantum Mechanics: Reprint of the Classics in Mathematics Series, 1994th edn. Springer-Verlag GmbH, Heidelberg (1994)
22. Andrei Bernevig, B., Hughes, T.L., Zhang, S.-C.: Quantum spin hall effect and topological phase transition in HgTe quantum wells. Science **314**(5806), 1757–1761 (2006). doi:10.1126/science.1133734
23. König, M., Wiedmann, S., Brüne, C., Roth, A., Buhmann, H., Molenkamp, L.W., Qi, X.-L., Zhang, S.-C.: Quantum spin hall insulator state in HgTe quantum wells. Science **318**(5851), 766–770 (2007). doi:10.1126/science.1148047
24. Novik, E.G., Pfeuffer-Jeschke, A., Jungwirth, T., Latussek, V., Becker, C.R., Landwehr, G., Buhmann, H., Molenkamp, L.W.: Band structure of semimagnetic hg_{1-y} mn_y quantum wells. Phys. Rev. B, **72**: 035321 (2005). doi:10.1103/PhysRevB.72.035321