3-D Quantum Transport Solver Based on the Perfectly Matched Layer and Spectral Element Methods for the Simulation of Semiconductor Nanodevices

Candong Cheng, Joon-Ho Lee, Kim Hwa Lim, Hisham Z. Massoud, and Qing Huo Liu
Duke University Department of Electrical and Computer Engineering Box 90291, 129 Hudson Hall Durham, NC 27708-0291, USA

Abstract

A 3-D quantum transport solver based on the spectral element method (SEM) and perfectly matched layer (PML) is introduced to solve the 3-D Schrödinger equation with a tensor effective mass. In this solver, the influence of the environment is replaced with the artificial PML open boundary extended beyond the contact regions of the device. These contact regions are treated as waveguides with known incident waves from waveguide mode solutions. As the transmitted wave function is treated as a total wave, there is no need to decompose it into waveguide modes, thus significantly simplifying the problem in comparison with conventional open boundary conditions. The spectral element method leads to an exponentially improving accuracy with the increase in the polynomial order and sampling points. The PML region can be designed such that less than −100 dB outgoing waves are reflected by this artificial material. The computational efficiency of the SEM solver is demonstrated by comparing the numerical and analytical results from waveguide and plane-wave examples, and its utility is illustrated by multiple-terminal devices and semiconductor nanotube devices.

Keywords

Perfectly matched layer (PML); spectral element method (SEM); open boundary condition; quantum transport; nanodevice simulation; Schrödinger equation; tensor effective mass

1 Introduction

Numerical solution of Schrödinger equation has become increasingly important because of the increasing demand for design optimization of nanodevices where quantum effects are significant. Various numerical methods have been developed for this purpose, for examples, [1]. These numerical methods have been developed for one, two, and three dimensions. It has been noted that some quantum effects can only be explained if a three-dimensional model is used [2]. 3D numerical solutions of Schrödinger equation and self-consistent Schrödinger–Poisson system have attracted much attention recently (see, for examples, [2,3,4,5,6,7]). The objective of this work is to develop an alternative 3D numerical method that can effectively account for open boundaries in Schrödinger equation.
In nanoelectronic devices, states carrying current are important in considering current flowing through a device region [8]. The calculation of these states are, however, complicated by the influence of the external environment. The most popular numerical method to calculate these states is to apply an open boundary condition to replace the influence of the environment. This boundary condition enables the electron to transmit out of the devices without any reflection from the environment.

Several open boundary conditions have been previously proposed [8,9,10]. The easiest way is to apply an infinite barrier at the boundary, i.e. enforcing the total wave function to be zero [9]. However, this method does not enforce the reflected wave to be zero at the boundary, thus is not a true open boundary condition. In a more accurate quantum transmitting boundary method (QTBM)[8,10], the contact region along the exterior environment is replaced by outgoing waveguide modes with known transmitting characteristics [8,10,11]. This method needs, however, to independently consider each mode of the incident, reflected and transmitted waves from each terminal. The summation up to a large number of modes makes the implementation, therefore, more complex. In addition, one PML absorber method was introduced to solve the time-dependent Schrödinger equation [12]. It was also implemented in a 2-D finite-element method with a diagonal effective mass matrix. On the other hand, the SEM has been applied to fluid dynamics with the outflow boundary conditions [13]. However, to our knowledge, such a method has never been implemented in the semiconductor field for 3-D problems.

In this paper, an efficient 3-D quantum transport solver is introduced. It is based on the spectral element method (SEM) and perfectly matched layer method to calculate the current-carrying states in devices with arbitrary geometry. In this accurate open boundary based on the PML, the contact regions of the device are extended into artificial PML media. The PML-modified Schrödinger equation with complex stretched coordinates is formulated and solved by the spectral element method. In the continuous limit, the interfaces between the PML and device region have zero reflection, and outgoing waves attenuate rapidly into the PML region before being terminated by an outer boundary. Consequently, the solution of Schrödinger equation and thus the current in the original device region do not deviate from the correct solution in the unbounded domain. Moreover, in this solver, the SEM is applied to achieve an error that decreases exponentially with the increase in the polynomial order and sampling points, thus significantly reducing the CPU time and memory requirement compared to conventional finite-element and finite-difference methods [14,15]. In addition, the effective mass is implemented as a full anisotropic mass tensor, which provides an excellent tool to study anisotropic effects along any arbitrary orientation.

The organization of this paper is as follows. In Section 2, we introduce the problem and the PML-modified Schrödinger equation, and the SEM applied to solve this PML-modified Schrödinger equation. In Section 3, we determine this solver’s accuracy and efficiency by waveguides and spherical quantum-dot examples. Moreover, the utility of this solver will be illustrated by multiple-terminal devices and nanotube examples.

2 Formulation

2.1 Problem Statement and Overview of the PML Method

2.1.1 Problem Statement—The objective of this work is to solve 3-D Schrödinger equation with a tensor effective mass in an unbounded domain. Although the solver and devices considered are all three-dimensional, for the sake of presentation we use a 2-D geometry to represent a cross section of a 3-D problem, as shown in Fig. 1. The entire device region is partitioned into one (or more) semiconductor region(s) \( \Omega_{\text{d}} \) and several contact regions \( \Omega_{\text{d}1}, \, \Omega_{\text{d}2}, \, \ldots, \, \Omega_{\text{d}k} \), where \( K \) is the number of the contacts connecting the device to the
external environment. We extend each contact region $\Omega_{d0}$ into an artificial PML region $\Omega_{\rho_n}$ with exactly the same material properties (such as the effective mass and potential energy) as its original contact region at the interface, except that it satisfies a PML-modified Schrödinger equation. The outer boundaries of the semiconductor region(s) and the contact regions are denoted as $\Gamma_{d0}$ ($n$ from 0 to $K$), and the outer boundaries of the PML regions are denoted as $\Gamma_{\rho_n}$ ($n$ from 1 to $K$). The interfaces between the contact regions and PML media are denoted as $\Gamma_{n}$ ($n$ from 1 to $K$), which are assumed as planes without loss of generality. Again, each region of $\Omega_{n}$ is a 3-D region and each boundary of $\Gamma_{n}$ is a 2-D surface.

Our objective is to solve for the wave function in the device regions $\psi_n \in C^3(\Omega_{d0}), n = 0, \ldots, K$ that satisfies the following 3-D normalized Schrödinger equation with a mass tensor [14,16]

$$
-\nabla \cdot \eta \nabla + iU(\mathbf{r}) - \epsilon(\mathbf{r}) \psi = 0,
$$

(1)

where the differentiation operator $\nabla$ and position coordinates have been normalized with $\mathbf{r} = r/d_0$, $U(\mathbf{r}) = U(r)/E_1$ is the normalized potential energy, $U(r) = -qV(r)$ the potential energy, $V(r)$ the potential, $\epsilon = E/E_1$, $E_1 = h^2/2m_0d_0^2$ (corresponding to the eigen energy of the ground state in a rectangular infinite quantum well with well width $d_0$), $m_0$ the free-electron mass, and $\eta$ is defined as

$$
\eta = \frac{1}{\pi^2m} \begin{bmatrix}
m_{xx} & m_{xy} & m_{xz} \\
m_{yx} & m_{yy} & m_{yz} \\
m_{zx} & m_{zy} & m_{zz}
\end{bmatrix}^{-1},
$$

where $m = m/m_0$ is the normalized effective-mass tensor, and $m$ is the effective-mass tensor. The reason to normalize $E$ by $E_1$ is to obtain a well conditioned matrix. For simplicity and without confusion, in the rest of this paper, we will choose $d_0 = 1$ nm and use $r$ to represent the normalized spatial position.

2.1.2 Overview of the PML Method—We apply the scatter-field/total-field formulation [17] to solve the scattered field in the PML region and the total field in the device region. The PML method results in zero reflection coefficients at interfaces $\Gamma_n$, and outgoing waves that attenuate rapidly into the PML region before reaching the PML outer boundary, which will be proved both analytically and numerically in the next subsections. On all of the outer surfaces of the simulation regions (all $\Gamma_{d0}$ and $\Gamma_{\rho_n}$), the zero Dirichlet boundary conditions or zero Neumann boundary conditions will be applied for the unknowns. On the other hand, on those interfaces between PML and device regions $\Gamma_n$, incident waves exist and the total field is chosen as the unknown variable, and the wave function continuity and current continuity are applied at these interfaces.

2.2 Principle of the PML Method

In this subsection, we will analytically prove that the reflection coefficients at interfaces $\Gamma_n$ are zero.

2.2.1 Modified Schrödinger Equation in PML—We have previously shown that the PML as an open boundary condition for Schrödinger equation is identical to that for acoustic waves in [18]. In this paper, we follow the complex coordinate stretching technique [19,20], and apply the following PML transformation: $\partial \xi \Rightarrow \epsilon_{s} \partial \xi$, where $\epsilon_{s} = a_{s} + i\omega_{s}/\omega$, and $s = x, y, \text{or} z$. In the above, the scaling coefficient $a_{s}$ is usually chosen as 1, because it will transform the modified
Schrödinger equation to the regular Schrödinger equation at the boundary nodes $\Gamma_n$. \( \omega = \frac{\xi}{\hbar} \) the frequency in the solution of the time-dependent Schrödinger equation given by
\[
\psi(x,t) = \psi(x) \exp \left( \frac{-iE t}{\hbar} \right) = \psi(x) \exp (-i\omega t),
\]
and $\omega_s$, the attenuation coefficient along $s$ direction, that is nonzero only along the propagating direction inside the PML region, expressed as
\[
\omega_s(s) = \alpha \max \omega_0 \delta \epsilon,
\]
where $\delta = |s - s_0|/L$ and $s_0$ is the location of the interface $\Gamma_n$, and $L$ is the PML thickness. By experience, we found that $L = 1.5/\lambda_{\min}$ and $\omega_0 = \omega$ could achieve good enough performance and require small computational size, where $\lambda_{\min}$ is the wavelength corresponding to the maximum energy of interest. The PML material parameters $\alpha_{\max}$ and $\epsilon$ determine the wave function attenuation rate in the PML region. Their optimization will be explained in Section 3.

Applying the above transformation into (1), we obtain
\[
-\nabla \cdot (\sigma \nabla \psi) + \epsilon_s e_s (\mu - \epsilon) \psi = 0
\]
as the PML-modified Schrödinger equation, where
\[
\sigma \equiv \begin{pmatrix}
\eta_{xx} & \eta_{xy} & \eta_{xz} \\
\eta_{yx} & \eta_{yy} & \eta_{yz} \\
\eta_{zx} & \eta_{zy} & \eta_{zz}
\end{pmatrix},
\]
Schrödinger equation in the device regions, given in (1), is a special form of (4) with $\omega_s(s) = 0$ and $\epsilon_s = 1$ ($s = x, y, \text{ or } z$).

2.2.2 The Solution of the Modified Schrödinger Equation—For a homogeneous PML medium, the modified Schrödinger equation in (4) has a plane-wave solution of the form
\[
\psi = A \exp \left[ i \left( k \cdot r \right) \right] = A \exp \left[ i \left( \epsilon_s (k x + e_s k_y + e_s k_z) \right) \right],
\]
where $k = (\epsilon_s k_x, \epsilon_s k_y, \epsilon_s k_z)$ and $A$ a constant used to normalize the wave function. Substituting (5) into (4) yields the dispersion relation
\[
\kappa = \sqrt{\frac{\epsilon_s^2}{\pi^2} - \kappa^2},
\]
where $\kappa^2 = \kappa^2 (\epsilon - \mu)$. For the special case of a diagonal mass tensor, (6) reduces to
\[
\frac{k_1^2}{m_{xx}} + \frac{k_2^2}{m_{yy}} + \frac{k_2^2}{m_{zz}} = \kappa^2.
\]
Equation (7) is the equation of an ellipsoid in 3D and is satisfied by
\[
k_1 = \kappa \sqrt{m_{xx}} \sin \theta \cos \phi,
\]
\[
k_2 = \kappa \sqrt{m_{yy}} \sin \theta \sin \phi,
\]
\[
k_3 = \kappa \sqrt{m_{zz}} \cos \theta,
\]
where $\theta$ is the polar angle from the $z$-axis in a spherical coordinate, and $\phi$ the azimuthal angle in the $xy$-plane from the $x$-axis.
2.2.3 Reflection Coefficient of an Oblique Incident Wave at a PML Interface—
Zero reflection coefficients at the interfaces between PML and regular materials with full tensor anisotropic material have been analytically proved for Maxwell’s equations [21] and widely applied in the field of electromagnetics [22, 23]. As the properties of Maxwell’s equations and Schrödinger equation are both second order partial differential equations, the above analytical proofs are also applicable for the solution of the Schrödinger equation in the semiconductor field. For simplicity, here we only consider a diagonal mass tensor in the contact and PML regions. Assume that a group of electrons with wave function \( \psi \) is obliquely incident on the interface at \( z = 0 \), as shown in Fig. 2. According to the solution of the modified Schrödinger equation, the incident waves \( \psi_i \), the reflected waves \( \psi_r \), and the transmitted waves \( \psi_t \) can be expressed as

\[
\begin{align*}
\psi_i &= A \exp (i k_{1z} z), \\
\psi_r &= R A \exp (i k_{2z} z), \\
\psi_t &= T A \exp (i k_{2z} z),
\end{align*}
\]

where \( k_{2z} = -k_{1z} \), and the subscripts 1 and 2 represent regions 1 and 2, respectively.

The continuity of the wave function at the interface \( z = 0 \) requires that

\[
1 + R = T,
\]

where \( R \) is the reflection coefficient and \( T \) is the transmission coefficient.

The continuity of the current at the interface at \( z = 0 \) requires the continuity of \( \vec{n} \cdot (\vec{\sigma} \nabla \psi) \), i.e.

\[
\frac{e_{1x} e_{1y} k_{1z}}{m_{1zz}} (1 - R) = \frac{e_{2x} e_{2y} k_{2z}}{m_{2zz}} T.
\]

Thus, from (11) and (12), we can write that

\[
R = \frac{e_{1x} e_{1y} m_{2zz} - e_{2x} e_{2y} m_{1zz}}{e_{1x} e_{1y} m_{2zz} + e_{2x} e_{2y} m_{1zz}}.
\]

2.2.4 Zero Reflection at a Perfectly Matched Interface—The phase matching condition requires that \( e_{1x} k_{1x} = e_{2x} k_{2x} \) and \( e_{1y} k_{1y} = e_{2y} k_{2y} \). Substituting these conditions into (8) and (9), we obtain that

\[
\begin{align*}
e_{1x} k_{1x} \sqrt{m_{1zz}} \sin \theta_1 \cos \phi_1 &= e_{2x} k_{2x} \sqrt{m_{2zz}} \sin \theta_2 \cos \phi_2, \\
e_{1y} k_{1y} \sqrt{m_{1zz}} \sin \theta_1 \sin \phi_1 &= e_{2y} k_{2y} \sqrt{m_{2zz}} \sin \theta_2 \sin \phi_2.
\end{align*}
\]

If we choose \( m_{1zz} = m_{2zz}, m_{1yy} = m_{2yy}, m_{1zz} = m_{2zz}, e_{1x} = e_{2x}, e_{1y} = e_{2y}, V_1 = V_2 \) (i.e. \( \kappa_1 = \kappa_2 \)), we can then obtain that \( \theta_1 = \theta_2 \), and \( \phi_1 = \phi_2 \), and therefore \( k_{1z} = k_{2z} \). Thus, from (13), the reflection coefficient \( R = 0 \) is always satisfied for any angle of incidence.

2.3 PML Implementation in the Spectral Element Method

In order to solve for the wave function using the spectral element method, we use the weak form of (4). By multiplying both sides of (4) with a testing function \( \psi_i \), and integrating by parts over the computational domain, the weak form of (4) can be written as

\[
\int_\Omega \left( \overline{\nabla \Psi_i} \cdot \nabla \psi_i + q \psi_i \Psi_i \right) dV = \int_{\partial \Omega} \overline{\nabla \Psi_i} \cdot (\sigma \nabla \psi_i) \Psi_i dS,
\]

\( \Omega \) being the computational domain, \( \partial \Omega \) being the boundary of \( \Omega \).
where $q = e_x e_y e_z (u - e)$. Within the framework of the spectral element method introduced previously [14,15], both the testing and basis functions are chosen as $\psi_i$, which we have selected the Gauss-Lobatto-Legendre (GLL) polynomials defined on a cubic reference element. The physical domain is divided into non-overlapping curved hexahedron elements conforming to the problem geometry; each curved hexahedron element is mapped into the cubic reference element. The unknown wave function is expanded in terms of the basis functions as

$$\phi = \sum_{j=1}^{N} \phi_j \psi_j(r)$$

where $\{ \phi_j \}$ are the values of the unknown wave functions at the SEM nodal points, and $N$ is the number of the total grid points in the curved hexahedron elements. Next, the SEM is applied to the device and PML regions, respectively.

### 2.3.1 System Equations in the Device Region—
The weak form equation for the device region (including all contacts) can be obtained simply from (14) with $e_x = e_y = e_z = 1$. To facilitate the interface between the PML and the contact regions within the total-field/scattered-field formulation, we also expand the normal flux term in terms of the basis functions at this interface. Assuming that there are $N_b$ SEM boundary nodes at the interfaces between the PML and the contacts, we expand the normal flux term in the contacts at the PML interface as

$$n \cdot (n \nabla \phi) = \sum_{j=1}^{N_b} c_j \psi_j(r)$$

where $c_j$ denotes the value of the normal flux at the $j$-th nodal point. The system equation in the device region can be written as

$$\sum_{e=1}^{K_D} \sum_{j=1}^{N} \int_{-1}^{1} \int_{-1}^{1} (\nabla \psi_j \cdot J_e^T n \nabla \phi_j) \sigma J_e |J_e| d\xi d\eta d\zeta$$

$$+ \sum_{e=1}^{K_D} \sum_{j=1}^{N} \int_{-1}^{1} \int_{-1}^{1} (\psi_j \cdot J_e^T n \nabla \phi_j) \sigma J_e |J_e| d\xi d\eta d\zeta$$

$$= \sum_{b=1}^{K_D} \sum_{j=1}^{N_b} \int_{-1}^{1} \int_{-1}^{1} (\psi_j \cdot J_b) d\xi d\eta d\zeta + \sum_{b=1}^{K_D} \sum_{j=1}^{N_b} \int_{-1}^{1} \int_{-1}^{1} (\psi_j \cdot J_b) d\xi d\eta d\zeta$$

(15)

where $\{ \phi_j \}$ are the unknown values of the wave functions at the nodal points defined on curved hexahedron elements in the device region, $K_D$ the total number of volume hexahedron elements partitioning the whole device region represented by $N_D$ unknown variables, $K_b$ the total number of surface elements on the interfaces $\{ \Gamma_n \}$ represented by $N_b$ unknown variables. Again, each of the curved volume hexahedron elements is mapped into a cubic reference element, and $J_e$ and $J_b$ are the Jacobian matrices in the volume element and in the boundary surface element, respectively, and $|J_e|$ and $|J_b|$ mean their corresponding determinants.

After applying the zero Dirichlet boundary conditions or zero Neumann boundary conditions on all $\Gamma_n$ in Fig. 1, the last term in (15) becomes zero and only the first term in the right-hand side of (15) remains and is denoted as $S_1$.

### 2.3.2 System Equations in the PML Region—
In the PML regions, we use the scattered field as the unknown variable in order to avoid the attenuation of the incident field from outside the device. Similar to the formulation for the device region, by considering the zero Dirichlet boundary condition or zero Neumann boundary condition for the scattered field at the outer surfaces of the PML regions, then we can obtain the system equation for the PML regions as

$$\sum_{e=1}^{K_D} \sum_{j=1}^{N} \int_{-1}^{1} \int_{-1}^{1} (\nabla \psi_j \cdot J_e^T n \nabla \phi_j) \sigma J_e |J_e| d\xi d\eta d\zeta$$

$$+ \sum_{e=1}^{K_D} \sum_{j=1}^{N} \int_{-1}^{1} \int_{-1}^{1} (\psi_j \cdot J_e^T n \nabla \phi_j) \sigma J_e |J_e| d\xi d\eta d\zeta = S_2$$

(16)
where $\phi_j^f$ is the scattered wave function at the $j$-th node, and $S_2$ is equal to

$$
S_2 = - \sum_{b=1}^{K_b} \sum_{i=1}^{N_b} \int_{I_i^b} \int_{I_i^b} c_j^f \overline{\psi_j^f} |\mathbf{J}_b| dudv - \sum_{b=1}^{K_b} \sum_{i=1}^{N_b} \int_{I_i^b} \int_{I_i^b} c_j^i \overline{\psi_j^i} (c_j^i - c_j^f) |\mathbf{J}_b| dudv,
$$

(17)

and $c_j^i$ and $c_j^f$ are the values of the normal flux at the $j$-th node corresponding to the scattered and incident wave functions, respectively. The relation $c_j^i = c_j^i - c_j^f$ is used in (17) since at the interfaces between the PML and contacts we have

$$
\vec{n} \cdot (\eta \nabla \psi) = \vec{n} \cdot (\sigma \nabla \psi).
$$

The definitions of $K_b$, $N_b$, and $N_b$ in the PML regions are similar to those in the device regions. The minus signs in those terms on the right-hand of (17) result from the normal directions of the interface for the PML regions opposite to the normal directions of the device regions.

Adding $S_1$ and $S_2$ yields to the source term $S$ at the internal interfaces between the PML and device regions

$$
S = S_1 + S_2 = \sum_{b=1}^{K_b} \sum_{i=1}^{N_b} \int_{I_i^b} \int_{I_i^b} c_j^i \overline{\psi_j^i} |\mathbf{J}_b| dudv,
$$

where $c_j^i$ is $\vec{n} \cdot (\eta \nabla \psi_j)$ evaluated at the $j$-th nodal point.

### 2.3.3 System Equation in a Matrix Form

Now we are in a position to combine the system equations for the device and PML regions by considering the boundary conditions at the PML-device interfaces. To this end, we separate the unknowns in the interior nodes and those at the PML-device interfaces. The system equation in the device regions can be rewritten in a matrix form as

$$
\begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix}
\begin{bmatrix}
\phi \\
\phi_b
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
S_1
\end{bmatrix},
$$

(18)

where $\phi$ and $\phi_b$ represent the wave functions on the inner and boundary nodal points in the device regions.

Similarly, the system equation in the PML region can be rewritten in a matrix form as

$$
\begin{bmatrix}
B_{11} & B_{12} \\
B_{21} & B_{22}
\end{bmatrix}
\begin{bmatrix}
\phi^f \\
\phi^b
\end{bmatrix}
= 
\begin{bmatrix}
S_2 \\
0
\end{bmatrix},
$$

(19)

where $\phi^f$ and $\phi^b$ represent the scattered wave functions on the inner and boundary nodal points in the PML regions.

With the boundary conditions at the interface $\Gamma_n$: $\psi^f = \psi - \psi^i$, (19) can be rewritten as

$$
\begin{bmatrix}
B_{11} & B_{12} \\
B_{21} & B_{22}
\end{bmatrix}
\begin{bmatrix}
\phi_b \\
\phi^f
\end{bmatrix}
= 
\begin{bmatrix}
S_2 + B_{11} \phi^f \\
B_{21} \phi^f
\end{bmatrix},
$$

(20)

Combination of (18) and (20) results to the global matrix equation

$$
\begin{bmatrix}
A_{11} & A_{12} & 0 \\
A_{21} & A_{22} + B_{11} & B_{12} \\
0 & B_{21} & B_{22}
\end{bmatrix}
\begin{bmatrix}
\phi \\
\phi_b \\
\phi^f
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
S + B_{11} \phi^f \\
B_{21} \phi^f
\end{bmatrix},
$$

(21)
The specific expressions of the above matrix elements are given in the Appendix. This system equation is solved by the conjugate-gradient method to obtain the unknown values of the wave function at nodal points.

3 Numerical Results and Discussions

Below we first use numerical results of a waveguide and a spherical quantum dot to show the validity and the exponential convergence of the SEM solver. We then apply this solver to a multiple-terminal device and a carbon nanotube.

3.1 A Waveguide example

First, we consider a rectangular waveguide with an electron wave propagating from a PML region into a device region and exiting via another PML region. The analytical solution available for this problem will be used to test the validity of the PML method and the convergence of the SEM method, and to optimize the PML parameters. The rectangular waveguide shown schematically in Fig. 3 is divided into three SEM elements, each having a size of 8 nm × 8 nm × 8 nm. The incident wave \( \phi_i = \sin(k_y y) \sin(k_z z) \exp(ik_x x) \) with its wavelength \( \lambda = 10 \) nm propagates along the +x direction from the left PML box, through the device region, and then transmits into a right PML box. The zero Dirichlet boundary conditions have been applied on all the outer surfaces.

3.1.1 Comparison with the Analytical Solution—We first apply a quadratic PML profile (i.e. \( c = 2 \)) with \( \alpha_{\text{max}} = 6.5 \) in (3). The wave function along the x direction in the device region is plotted in Fig. 4. The \( L_2 \) error for all nodes in the device region at the sampling density \( SD = 10 \) PPWs (points per wavelength) reaches 0.02\%, which is within the acceptable error range.

3.1.2 Optimization of PML Profile Parameters and the SEM Convergence—We use this example to optimize the PML profile parameters \( c \) and \( \alpha_{\text{max}} \), and to test the error convergence of the SEM.

We choose \( c = 0, 1, 2 \) and \( 3 \), and apply different values of \( \alpha_{\text{max}} \) to compare the error performance for \( SD=14 \) PPWs, as shown in Fig. 5. This error is the difference between the calculated wave function and the analytical solution within the central device region. It is observed that the curve with \( c = 0 \) and \( \alpha_{\text{max}} = 4.0 \) has the minimum error (\( L_2 \) error = \( 3.16 \times 10^{-8} = -150 \) dB).

From the above study, we obtain the optimal value of \( \alpha_{\text{max}} \) for different profiles of \( c \) that give the minimum error: \( \alpha_{\text{max}} = 4.0, 6.5, 8.3 \) and \( 11.0 \) for \( c = 0 \) to \( 3 \) respectively. Using these optimal values of \( \alpha_{\text{max}} \), we now test the error convergence with the increasing sampling density for \( c = 0 \) to \( 3 \), as shown in Fig. 6. The error decreases exponentially with the increase of the sampling density for all these profiles. Moreover, as expected, the smallest reachable error increases with the increase of \( c \).

Overall, the PML profile with \( c = 0 \) and \( \alpha_{\text{max}} = 4.0 \) has the minimum error in the above waveguide device simulations. The value \( c = 0 \) yields the smallest error because it does not increase the polynomial order of the integrand of \( B_{ij} \) in (A.1). The Gauss-Lobatto-Legendre quadrature order to evaluate this integrand is, therefore, not required to be increased in the SEM. On the other hand, \( \alpha_{\text{max}} = 4.0 \) yields the smallest error because an increase in \( \alpha_{\text{max}} \) will speed up the decay rate inside the PML region, while large values of \( \alpha_{\text{max}} \) will introduce reflection because of the significant difference between the device material and PML material. This optimal value of \( \alpha_{\text{max}} \) may be problem-dependent, but the error achieved by this choice is in general much smaller than that required by most engineering applications. In the remaining examples, the values of \( c = 0 \) and \( \alpha_{\text{max}} = 4.0 \) will be used.
3.2 Plane-Wave Incidence on a Spherical Quantum Dot

Next, to test the SEM solver on curved structures, we simulate a spherical quantum dot with plane-wave incidence.

The incident plane wave $\phi^i = \exp(ikz)$ propagates along the $+z$ direction from a high potential energy region (region 2, $V_2 = 0.014$ V), and reaches a low potential energy quantum dot sphere (Region 1, $V_1 = 0$ V, radius $r = 8$ nm). Region 2 is unbounded, and thus should not produce reflections. We truncated Region 2 with a cube ($24$ nm $\times$ $24$ nm $\times$ $24$ nm) and applied a PML region with $8$ nm thick outside the cube. The zero Neumann boundary conditions have been applied on all outer surfaces. The energy for the incident wave is $E = 0.166$ eV and the effective mass is $m = m_0$. The geometry is shown in Fig. 7. The sphere being modeled by a cubic element at the center and six curved hexahedron elements conforming to the spherical surface. The sampling densities are approximately even in $x$, $y$, and $z$ directions.

The differences of the real (a) and imaginary (b) parts of the wave function between numerical results and analytical solutions at central $y$ surface are shown in Fig. 8. In this example, basis order $N = 8$, i.e. $SD = 9$ PPW. We found that the error is within acceptable range.

The error of the wave function difference between the analytical solution and numerical results in the device region is shown in Fig. 9. The error decreases exponentially with the increase of basis function order $N$. When $N = 8$ (i.e. $SD = 8$ PPWs), the error reaches $0.047\%$.

The results reported in the above subsections show that the PML method has been applied effectively, as the reflection coefficient has been theoretically proved to be zero, and numerical results are within an acceptable error range. In addition, we have implemented the SEM in the waveguide and quantum-dot examples, yielding errors that decrease exponentially with the increase of the sampling density $SD$. In the next subsections, we will utilize this solver to calculate the electron concentration and transmission coefficient in semiconductor devices.

3.3 A Multiple-Terminal Device with Quantum Dots

In this example, we will simulate the electron concentration distribution in a multiple-terminal device that includes several quantum dots. There are five contacts connecting the device to the external environment, which can be treated as five terminals with each having an incident and a reflected wave. This example can be easily transformed into a quantum coupler, a quantum dot, or a MOSFET.

3.3.1 Geometry—The geometry of the device is shown in Fig. 10. Five cubic contact terminals (Regions 1, 4, 5, 6, and 7, with $V_1 = V_4 = V_6 = V_7 = -5.5$ V, $V_5 = 5.9$ V) carrying waveguide modes $\psi^i = \cos(k_ax) \cos(k_by) \exp(-ik_cz)$ with energy $E = 0.38$ eV propagate into a semiconductor region (Region 2, $V_2 = 0.38$ V, $72$ nm $\times$ $24$ nm $\times$ $24$ nm) with three inner quantum-dot spheres (Region 3, $V_3 = -0.153$ V, $r = 5$ nm), where $a$ and $b$ are the transverse variables and $c$ the longitudinal variable in the rectangular waveguides. These five contact terminals are also the exit terminals of the device. All terminals have the same size of $8$ nm $\times$ $8$ nm $\times$ $2$ nm, where $8$ nm is the transverse direction and $2$ nm is the propagating direction. The Fermi level in Regions 1 is chosen to be zero. Each contact region has been extended into a $2$ nm thick PML region. The zero Neumann boundary conditions have been applied on all outer surfaces.

The electron effective mass for all materials are assumed to be $m_e = 0.5m_0$. This example can be treated with terminal 1 as the gate region, terminals 4 and 5 as the source and drain regions, and terminals 6 and 7 as the bulk region. It is easy to change the bias applied at each terminal. Here, we assume that the potential distribution is known and is constant in each region, while
the real potential distribution will be determined from a Poisson solver, which will be investigated in a forthcoming paper.

### 3.3.2 Electron Concentration Distribution

The electron concentration distribution $| \psi^2 |$ along the central $z$ surface is shown in Fig. 11. As can be seen, the electrons are accumulated in the quantum-dot low-energy regions. Moreover, as the geometry of the device and the bias applied on this device is symmetric along $x = 0$, the electron distribution is also symmetric along $x = 0$, as expected.

This multiple-terminal device example has proved the solver's ability to calculate the wave function and electron concentration distribution in the quantum coupler, quantum dot and MOSFET devices.

### 3.4 A Carbon Nanotube

Finally, we will simulate the wave function in a carbon nanotube. We will compare our results with the transmission coefficient reported in [24].

We will only simulate the nanotube, source and drain regions in the geometry of [24], as shown in Fig. 12, i.e. $R_t = 0.63$ nm, $L_t = 20$ nm, work function $\Phi_{\text{metal}} = 4.5$ V, $\chi_{\text{metal}} = 10.0$ eV, $\chi_{\text{CN}} = 4.2$ eV. A bias voltage $V_{DS} = 0.4$ V is applied between the drain and source contacts. The potential energy was obtained from Fig. 4(b) in [24].

#### 3.4.1 Transmission Coefficients

Based on the geometry in Fig. 12, we simulated the transmission coefficients with two different incident electron waves: (a) a constant wave front, i.e., the first circular waveguide mode $\psi = AJ_0 \left( x_1 \rho / a \right) \exp \left( ik z \right)$ (where $J_m(x)$ is the Bessel function of the $m$-th order, $x_0 = 0$ is the $n$-th zero of the derivative of $J_0(x)$, and $A$ is the amplitude of the wave function calculated from Landauer’s equation [24]). This example is used to compare the 1-D transmission coefficient with the result reported in [24].

(b) the second circular waveguide mode $\psi = AJ_1 \left( x_1 \rho / a \right) \exp \left( ik z \right)$ $k_z = \sqrt{k^2 - \left( x_1 / a \right)^2}$. This example is used to calculate the 3-D transmission coefficient. These two examples are respectively denoted as $J_{01}$ and $J_{11}$. In the $J_{11}$ example, the incident wave does not propagate in the nanotube with $R_t = 0.63$ nm, which leads to negligible transmission coefficients. In order to compare the 1-D and 3-D transmission coefficients, in the $J_{11}$ mode, we have chosen the radius $R_t = 4.6$ nm whose propagating probability is large enough. To confirm the 1-D transmission coefficients results, we also simulated a third example with a constant wave front (i.e. a square waveguide mode $\psi = A \cos \left( m \pi x / a \right) \cos \left( n \pi y / b \right) \exp \left( ik z \right)$, denoted as “$S_{00}$” for $m = n = 0$; $k_z = \sqrt{k^2 - m^2 \pi^2 / a^2 - n^2 \pi^2 / b^2}$ on a square cylinder nanotube with the same cross section surface area as the case with $R_t = 0.63$ nm. All three incident waves satisfy the Neumann boundary conditions at the outer surface of the cylinders. The example we consider here only has a source injection of electrons, as the opposite direction of the injection is analogous. In order to obtain the same result as [24], we choose the effective mass for the carbon nanotube as $m_{\text{CN}} = 0.0615 m_0$. We calculated the transmission coefficient from both 1-D and 3-D models for the above three examples. The 1-D transmission coefficient is defined as [24]

$$T_{1D} = \frac{m_s^2 k_d |C|^2}{m_s^2 k_s |A|^2},$$

(22)

where $k_d$ and $k_s$ are the wave numbers along the propagation $z$ direction in the drain and source contacts, respectively, $m_d$ and $m_s$ the electron effective mass in the drain and source contacts, respectively, and $C$ and $A$ the amplitudes of the wave functions at the drain and source contacts.
respectively [24]. The 3-D transmission coefficient is defined in terms of the current density ratio between the outermost interfaces at the drain and source contacts as

$$T_{3D} = \frac{J_D}{J_S} = \frac{\int_{\partial \Omega_D} \left[ \psi_D^{-1} m_D \cdot \nabla \psi_D - \psi_D^{-1} m_D \cdot \nabla \psi_D \right] \cdot ds}{\int_{\partial \Omega_S} \left[ \psi_S^{-1} m_S \cdot \nabla \psi_S - \psi_S^{-1} m_S \cdot \nabla \psi_S \right] \cdot ds},$$

where $J_D$ and $\psi_D$ are the current density and transmitted wave at the drain-contact exit interface, respectively, $J_S$ and $\psi_S$ the current density and incident wave at the source contact interface, respectively, $m_D$ and $m_S$ the electron effective mass tensors in the drain and source contacts, respectively.

The transmission coefficients obtained in this work and those reported in [24] (denoted as “JCPP”) are compared in Fig. 13. The dotted line represents the $T_{3D}$ value using a constant incident wave ($S_{00}$ mode) passing through a square cylinder. The dash-dotted line represents the $T_{3D}$ value using a constant incident wave ($J_{01}$ mode, denoted as “CJ01”) passing through the circular cylinder. Our calculated 1-D transmission coefficients values $T_{1D}$ also agree with the results of John et al. [24].

As can be seen in Fig. 13, when we use a constant incident wave, (i.e. $J_{01}$ and $S_{00}$ modes, same as in [24]), the magnitudes of the $T_{3D}$, the constructive and the destructive energies are almost the same for the circular and the square nanotubes for all energies, and also agree well with the results reported in [24]. The transmission coefficient results for the $J_{11}$ incident wave on the circular cylinder (denoted as “CJ11”) with $R_t = 4.6$ nm shown in Fig. 13 have a little shift in energy from the $J_{01}$ and $S_{00}$ fundamental modes. This shift is due to the wave number $k_z$ for the $J_{11}$ mode being smaller than $k_z = k$ for the $J_{01}$ and $S_{00}$ modes. Furthermore, as can be seen in Fig. 13, small energy level ranges ($E \in [-0.2, -0.07]$ eV) lead to almost zero propagation probability, because $k_z$ in this energy range has an imaginary value, and the incident mode is an evanescent one. This result is based on the enlarged radius $R_t = 4.6$ nm for obtaining real value of $k_z$ for energies larger than $-0.07$ eV.

### 3.4.2 Electron Distribution at Different Interference Energies

From Fig. 13, we choose a constructive interference at energy $E = 0.1085$ eV, a destructive energy at $E = 0.1712$ eV, an intermediate one at energy $E = 0.14$ eV based on the $J_{01}$ example, and plot the normalized electron concentration distribution along the central axial $z$ direction of the coaxial CNFET in Fig. 14. As can be seen, the peak of the electron concentration for $E = 0.1085$ eV is largest among the three energies, which implies its constructive interference property. Moreover, the peak of the electron concentration for the destructive interference at energy $E = 0.1712$ eV is smallest among the three energies. Furthermore, at $E = 0.14$ eV, an intermediate electron concentration is obtained in the nanotube.

### 3.4.3 Effects of the Mass Tensor Anisotropy

Finally, the solver is used to model the effects of the mass tensor by comparing the transmission coefficient corresponding to a scalar mass, a diagonal mass tensor, and a full mass tensor for the above example. The incident wave is the fundamental $J_{01}$ mode along a circular cylinder with a radius $R_t = 0.63$ nm (same as above examples) for three different effective masses inside the CNFET region. In the first example, $m_{SM} = 0.0615 m_0$ is the same scalar mass as in the above example. In the second example, the diagonal mass tensor

$$m_{D} = \begin{bmatrix} 0.0523 & 0 & 0 \\ 0 & 0.0523 & 0 \\ 0 & 0 & 0.0851 \end{bmatrix} m_0$$

and
has the same determinant as that of the scalar mass, i.e., $|m_{DMT}| = 0.0615^3 m_0^3$, but the transverse component has a 15% perturbation.

In the third example, the full mass tensor
\[
\begin{bmatrix}
-0.0646 & -0.0071 & -0.0142 \\
-0.0071 & 0.0564 & 0.0082 \\
-0.0142 & 0.0082 & 0.0687
\end{bmatrix}
\]

is obtained by rotating the principal axis of the diagonal mass tensor by 30° along the $z$ axis then by 45° along the $y$ axis. It also has the same determinant as the scalar mass.

Meanwhile, the effective mass inside the lead region is still chosen to be the same as in the previous example, i.e., a scalar mass with value $m_{lead} = 0.0615 m_0$.

The values of the transmission coefficients $T_{3D}$ obtained for the three mass tensors are compared in Fig. 15. It is observed that the transmission coefficient peaks for the diagonal mass tensor have large negative energy shifts from the corresponding ones for the scalar mass. These shifts are caused by the increased $k_z$ value (thus decreased effective wavelength in the $z$ direction) because of the increased $m_{zz}$, as can be seen in equation (7). Note that the decreased values of $m_{xx} = m_{yy}$ have much less effect for this fundamental mode. For the case of a full mass tensor, since the principal axis has been rotated, this anisotropy effect is reduced because of the averaging effect caused by rotation, thus the negative shifts in the transmission coefficient are less than the case of the diagonal mass tensor.

### 4 Conclusions

We have demonstrated both analytically and numerically that the reflection coefficients at the interfaces between device regions and perfectly matched layers can be made smaller than $-100$ dB when we appropriately choose the parameters of the perfectly matched layers in the spectral element method. Moreover, we have shown that this solver can achieve exponential convergence, reaching 0.1% accuracy when the sampling density exceeds eight points per wavelength, enabling us to significantly reduce the computational time and memory from conventional finite-element and finite-difference methods [14,16]. Although we focused on a Schrödinger solver in this article, it is natural for us to extend this solver to the self-consistent Schrödinger-Poisson system in 3D in our future work, as has been done in the 1-D case [14]. Such a self-consistent solver has been done in the first author’s thesis, and will be reported in the near future. In our self-consistent solver, Schrödinger and Poisson equations are solved iteratively by a nonlinear solver. Within each iteration of the nonlinear solution, both Schrödinger and Poisson equations are linear, and thus can be solved by the SEM [25].

### Acknowledgments

We are grateful to the helpful discussion with Dr. Eric Polizzi from University of Massachusetts and Dr. David John from University of Wales, Bangor.

### A Expression of Each Term in the Global Matrix Equation

The global matrix equation is expressed in (21) as:

\[
\begin{bmatrix}
A_{11} & A_{12} & 0 \\
A_{21} & A_{22} + B_{11} & B_{12} \\
0 & B_{21} & B_{22}
\end{bmatrix}
\begin{bmatrix}
\phi \\
\phi_b
\end{bmatrix}
= \begin{bmatrix}
0 \\
S + B_{11} \phi^t_b \\
B_{21} \phi^t_b
\end{bmatrix}.
\]

The expressions of the corresponding elemental matrices $A_{ij}^T$, $B_{ij}^T$ and $S_i^T$ are as follows:
\[ A_{ij} = \int_{-1}^{1} \int_{-1}^{1} (\nabla \Psi_i) \cdot J^T e^{-i \sum_{j} \Psi_j} J_0 \int_{D} d\xi d\eta d\zeta + \int_{-1}^{1} \int_{-1}^{1} (u - \varepsilon) \Psi_i \Psi_j J_0 \int_{D} d\xi d\eta d\zeta, \]  
(A.1)

where \( i, j = 1, 2, \ldots, N_D \), and \( N_D \) is the number of nodes in each element in the device region,

\[ B_{ij} = \int_{-1}^{1} \int_{-1}^{1} (\nabla \Psi_i) \cdot J^T \sigma J_0 (\nabla \Psi_j) J_0 \int_{D} d\xi d\eta d\zeta + \int_{-1}^{1} \int_{-1}^{1} e_\xi e_\eta e_\zeta (u - \varepsilon) \Psi_i \Psi_j J_0 \int_{D} d\xi d\eta d\zeta, \]  
(A.2)

where \( i, j = 1, 2, \ldots, N_P \), and \( N_P \) is the number of nodes in each PML element, and

\[ S_j = \sum_{i=1}^{N_b} \int_{-1}^{1} \int_{-1}^{1} e^j \Psi_i \Psi_j J_0 \int_{D} d\xi d\eta d\zeta, \]  
(A.3)

where \( i = 1, 2, \ldots, N_P \), and \( N_P \) is the number of nodes on each boundary element at the interfaces between the PML and contacts.

For the global matrices, the range of \( i \) and \( j \) in each term of the above equation is as follows:

- \((A_{11})_{ij} : i = 1, \ldots, N_D - N_b; \)
- \((A_{12})_{ij} : i = 1, \ldots, N_D - N_b; j = N_D - N_b + 1, \ldots, N_D; \)
- \((A_{21})_{ij} : i = N_D - N_b + 1, \ldots, N_D; j = 1, \ldots, N_D - N_b; \)
- \((A_{22})_{ij} : i = N_D - N_b + 1, \ldots, N_D; j = 1, \ldots, N_D; \)
- \((B_{11})_{ij} : i = 1, \ldots, N_P; j = 1, \ldots, N_P; \)
- \((B_{12})_{ij} : i = 1, \ldots, N_P; j = N_P + 1, \ldots, N_P; \)
- \((B_{21})_{ij} : i = N_P + 1, \ldots, N_P; j = 1, \ldots, N_P; \)
- \((B_{22})_{ij} : i = N_P + 1, \ldots, N_P; j = 1, \ldots, N_P; \)

where \( N_D \) is the total nodes in the device regions, \( N_P \) is the total nodes in the PML regions, and \( N_B \) is the total boundary nodes at the interfaces \( \Gamma_n \).

References


Fig. 1.
The problem geometry for a semiconductor device with $K$ contact regions truncated by the PML medium.
Fig. 2.
Arbitrary angle electron incident on the plane $z=0$. 
Fig. 3.
The geometry for three-element PML Model.
Fig. 4.
Comparison of SEM results and analytical solutions for the wave components versus $x$ at the center $y$ and center $z$ line in the device region.
Fig. 5.
The error convergence with the change of $\alpha_{max}$ for $c = 0$ to $c = 3$ in the PML region.
Fig. 6.
The error convergence with the increase of the sampling density.
Fig. 7. The geometry for a spherical quantum dot with plane wave incidence.
Fig. 8.
The differences of the real (a) and imaginary (b) parts of the wave function at central y surface between numerical results and analytical solutions.
Fig. 9.
The error of the wave function between numerical results and analytical solutions for a sphere quantum dot with plane wave incidence.
Fig. 10.  
The geometry for a device with three inner quantum dots, five incident terminals and five exit terminals.
Fig. 11.
The electron concentration distribution on the central $z$ surface for a device with three inner spherical quantum dots and five contact terminals. The unit for the electron concentration shown in this figure is cm$^{-3}$. 
Fig. 12.
The geometry for a coaxial Carbon nanotube FET device.
Fig. 13.
The transmission coefficient comparison for a CNFET with different input energies and different incident waves.
Fig. 14.
The electron concentration along $z$ direction of the coaxial CNFET with input energies $E = 0.1085$ eV, $0.1712$ eV and $0.14$ eV.
Fig. 15.
Comparison of the transmission coefficient for three CNFETs with a scalar mass, diagonal mass tensor and full mass tensor versus the input electron energy.