Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Journal of Computational Physics 209 (2005) 421–447

JOURNAL OF COMPUTATIONAL PHYSICS

www.elsevier.com/locate/jcp

Numerical simulation of two-dimensional electron transport in cylindrical nanostructures using Wigner function methods

Greg Recine *, Bernard Rosen, Hong-Liang Cui

Applied Electronics Laboratory, Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken, NJ 07030, United States

Received 6 December 2004; received in revised form 8 March 2005; accepted 10 March 2005
Available online 2 June 2005

Abstract

We have constructed a lattice Wigner–Weyl code to expand the Buot–Jensen algorithm to calculation of electron transport in two-dimensional cylindrically symmetric structures. Almost all of the numerical simulations to date have dealt with the restricted problem of one-dimensional transport. In real devices, electrons are not confined to a single transport dimension and the coulombic potential is fully present and felt in three dimensions. We show the derivation of the 2D equation in cylindrical coordinates as well as approximations employed in the calculation of the four-dimensional convolution integral of the Wigner function and the potential. We work under the assumption that longitudinal transport is more dominant than radial transport and employ parallel processing techniques. The total transport is calculated in two steps: (1) transport the particles in the longitudinal direction in each shell separately, then (2) each shell exchanges particles with its nearest neighbor. Most of this work is concerned with the former step: A 1D space and 2D momentum transport problem. Time evolution simulations based on these method are presented for three different cases. Each case lead to numerical results consistent with expectations. Discussions of future improvements are discussed.

1. Introduction

Electron transport in a resonant tunneling structure (RTS) has been studied in detail over the past decade [3–6,9,11]. Almost all of the numerical simulations have restricted the problem to one-dimensional transport. Much progress has been made using the 1D theory, however, in real devices electrons are not confined to transport in a single dimension and the coulombic potential is fully present and felt in three

* Corresponding author. Tel.: +1 201 736 1111.
E-mail address: gjr5y@virginia.edu (G. Recine).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2005.03.009

dimensions. Here, we present a method for numerical simulation of electronic transport through a cylindrical device that possesses azimuthal symmetry. Fig. 1(a) shows a schematic representation of such a device.

We work under the assumption that longitudinal transport dominants over the radial transport. The total transport is calculated in two steps: (1) transport the particles in the longitudinal direction in each shell separately, then (2) each shell exchanges particles with its nearest neighbor. During a given time step the particles are advanced longitudinally through the device, as in a 1D problem, but with the inclusion of radial momentum. This changes the form of the potential and interaction terms of the familiar 1D Wigner function (transport) equation (WFE). Since the latter step is computationally simple, most of this work is concerned with the former step: A 1D space and 2D momentum transport problem ($1x + 2k$). In this paper, the WFE is solved self-consistently with the Poisson equation.

In order to perform the numerical simulation, parallel programming techniques are used. A simplest way to attack this problem to slice up the device into cylindrically concentric shells as shown in Fig. 1(b). The two above steps now become: (1) each processor (shell) calculates the ($1x + 2k$) transport problem, then (2) each processor exchanges particle information with its nearest neighbor.

This papers is organized as follows: The first three sections detail the derivations: Section 2 the 2D WFE in cylindrical coordinates (assuming azimuthal symmetry), and Sections 3 and 4 the discretization. Section 5 discusses different methods of solving the $1x + 2k$ problem in regard to the limitations of today's computational resources. This includes a re-derivation of the potential term in the WFE and a splitting of the potential into static (conduction band edge) and changing (self-consistent) potentials. Finally, our concluding remarks are in Section 7.

2. Derivation

The 3D form of the Wigner function equation (WFE) can be written as (scattering will be added in later)

$$\frac{\mathrm{d}f(\mathbf{q},\mathbf{k})}{\mathrm{d}t} = -\frac{\hbar\mathbf{k}}{m^*}\cdot\nabla_{\mathbf{q}}f(\mathbf{q},\mathbf{k}) - \frac{\mathrm{i}}{(2\pi)^3\hbar}\int \mathrm{d}\mathbf{k}'\int 2\,\mathrm{d}\mathbf{y}\mathrm{e}^{-2\mathrm{i}(\mathbf{k}-\mathbf{k}')\cdot\mathbf{y}}\{V(\mathbf{q}+\mathbf{y}) - V(\mathbf{q}-\mathbf{y})\}f(\mathbf{q},\mathbf{k}'). \tag{1}$$

The derivation of the WFE was in the same vein as Frensley [5] (but kept in full 3D form here), and, as he has noted, has been derived while assuming no boundaries exist. Because of this, the integral limits have been omitted in the above and will be discussed later.

First, let us rewrite this equation in cylindrical coordinates by making the substitutions: $\mathbf{q}\to r,z,\phi$; $\mathbf{k}\to k_r,k_z,\chi_\phi$; $\mathbf{y}\to\rho,\zeta,\theta$. It is important to note that while $\mathbf{q}$ represents the spatial (i.e., center of mass)

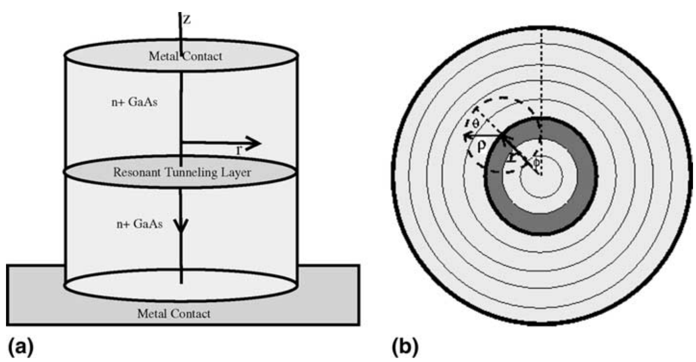

Fig. 1. Cylindrical RTD: (a) side view and (b) top view.

coordinate with respect to the origin on the cylindrical axis, **y** represents the spatial coordinate with respect to the origin at an arbitrary point within the cylinder, not necessarily on the axis. Fig. 1(b) illustrates the geometry of these two variables. Eq. (1) now becomes:

$$\begin{aligned}\frac{\mathrm{d}f(r,z,\phi,k_z,k_r,\chi_\phi)}{\mathrm{d}t} &= -\frac{\hbar}{m^*}\left[k_r\frac{\partial}{\partial r}+\frac{\chi_\phi}{r}\frac{\partial}{\partial\phi}+k_z\frac{\partial}{\partial z}\right]f(r,z,\phi,k_z,k_r,\chi_\phi)-\frac{2\mathrm{i}}{(2\pi)^3\hbar}\int\mathrm{d}k'_z\int\mathrm{d}k'_r\\ &\quad\times\int_0^{2\pi}|k'_r|\mathrm{d}\chi'_\phi\int\mathrm{d}\zeta\int\mathrm{d}\rho\int_0^{2\pi}|\rho|\,\mathrm{d}\theta\,\mathrm{e}^{-2\mathrm{i}[(k_z-k'_z)\zeta+(k_r\cos\chi_\phi-k'_r\cos\chi'_\phi)\rho]}\\ &\quad\times\{V(r+\rho,z+\zeta,\phi+\theta)-V(r-\rho,z-\zeta,\phi-\theta)\}f(r,z,\phi,k'_z,k'_r,\chi'_\phi).\end{aligned} \tag{2}$$

Considering a 2D problem in cylindrical coordinates with azimuthal symmetry, the first thing to notice is that even though the potential is symmetric in ϕ ($V(r,z,\phi)\to V(r,z)$), the potential difference *is dependent on angle*, $V(r\pm\rho,z\pm\zeta,\phi\pm\theta)\to V(r\pm\rho,z\pm\zeta,\pm\theta)$. Next, integrate out the remaining azimuthal spatial and momentum components. Using

$$\int_0^{2\pi}\int_0^{2\pi}f(r,z,\phi,k_z,k_r,\chi_\phi)|r|\,\mathrm{d}\phi|k_r|\,\mathrm{d}\chi_\phi=(2\pi)^2|r||k_r|f(r,z,k_z,k_r), \tag{3}$$

one obtains

$$\begin{aligned}\frac{\mathrm{d}f(r,z,k_z,k_r)}{\mathrm{d}t} &= -\frac{\hbar}{m^*}\left[k_r\frac{\partial}{\partial r}+k_z\frac{\partial}{\partial z}\right]f(r,z,k_z,k_r)-\frac{2\mathrm{i}}{(2\pi)^3\hbar}\int\mathrm{d}k'_z\int\mathrm{d}k'_r\int_0^{2\pi}|k'_r|\,\mathrm{d}\chi'_\phi\\ &\quad\times\int\mathrm{d}\zeta\int\mathrm{d}\rho\int_0^{2\pi}|\rho|\,\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\chi_\phi\mathrm{e}^{-2\mathrm{i}[(k_z-k'_z)\zeta+(k_r\cos\chi_\phi-k'_r\cos\chi'_\phi)\rho]}\\ &\quad\times\{V(z+\zeta,r+\rho,\theta)-V(z-\zeta,r-\rho,-\theta)\}f(r,z,k'_z,k'_r,\chi'_\phi).\end{aligned} \tag{4}$$

The drift term is relatively simple, but the potential term is a bit completed. Rewriting the potential term as

$$-\frac{2\mathrm{i}}{(2\pi)^3\hbar}\int|k'_r|\,\mathrm{d}k'_r\int\mathrm{d}k'_z\int\mathrm{d}\zeta\mathrm{e}^{-2\mathrm{i}(k_z-k'_z)\zeta}\int|\rho|\,\mathrm{d}\rho\mathscr{J}(\rho,k_r)\mathscr{V}(z,\zeta,r,\rho)\mathscr{F}(z,r,\rho,k'_z,k'_r), \tag{5}$$

and using the following definitions

$$\mathscr{J}(\rho,k_r)=\int_0^{2\pi}\mathrm{d}\chi_\phi\mathrm{e}^{-2\mathrm{i}k_r\cos\chi_\phi\rho}, \tag{6}$$

$$\mathscr{V}(z,\zeta,r,\rho)=\int_0^{2\pi}\mathrm{d}\theta\{V(z+\zeta,r+\rho,\theta)-V(z-\zeta,r-\rho,-\theta)\}, \tag{7}$$

$$\mathscr{F}(z,r,\rho,k'_z,k'_r)=\int_0^{2\pi}\mathrm{d}\chi'_\phi\mathrm{e}^{+2\mathrm{i}k'_r\cos\chi'_\phi\rho}f(r,z,k'_z,k'_r,\chi'_\phi), \tag{8}$$

reduces the complexity. Eq. (6) is just the definition of the Bessel function $2\pi J_0(2k_r\rho)$. Eq. (7) is evaluated at a given ρ,θ by noting that $\rho(\theta)=\rho\cos\theta$, leaving Eq. (8) to be dealt with. By expanding the exponential in terms of Bessel functions of χ'_ϕ, the integral becomes

$$\mathscr{F}(z,r,\rho,k'_z,k'_r)=\int_0^{2\pi}\mathrm{d}\chi'_\phi 2\pi\sum_{m,m'}J_{m'}(2k'_r\rho)f_m(r,z,k'_z,k'_r)\mathrm{e}^{\mathrm{i}(m+m')\chi'_\phi}=2\pi\sum_m J_m(2k'_r\rho)f_m(r,z,k'_z,k'_r). \tag{9}$$

Now there is an infinite series in m, but only the term $m = 0$ must be counted. This is because it represents the azimuthally independent functions, which, as dictated by the current problem, is the form that the Wigner distribution function, f, should take.

So, finally, the complete 2D form of the WFE in azimuthally independent cylindrical coordinates is

$$\frac{\mathrm{d}f(r,z,k_z,k_r)}{\mathrm{d}} = -\frac{\hbar}{m^*}\left[k_r\frac{\partial}{\partial r}+k_z\frac{\partial}{\partial z}\right]f(r,z,k_z,k_r)+\frac{1}{\pi\hbar}\int \mathrm{d}k_z'\int|\rho|\,\mathrm{d}\rho\,\mathscr{U}(r,\rho,z,k_r,k_z-k_z')\mathscr{F}(r,\rho,z,k_z'), \tag{10}$$

where (notice the terms $\mathscr{F}$, $\mathscr{U}$ and $\mathscr{V}$ have been redefined from what was written above)

$$\mathscr{F}(r,\rho,z,k_z) = \int |k_r'|\,\mathrm{d}k_r' J_0(2k_r'\rho)f(r,z,k_z,k_r\prime), \tag{11}$$

$$\mathscr{U}(r,\rho,z,k_z,k_r) = \int \mathrm{d}\zeta\,\sin(2k_z\zeta)J_0(2k_r\rho)\mathscr{V}(z,\zeta,r,\rho), \tag{12}$$

$$\mathscr{V}(z,\zeta,r,\rho) = \int_0^{2\pi}\mathrm{d}\theta\{V(z+\zeta,r+\rho\cos\theta)-V(z-\zeta,r-\rho\cos\theta)\}. \tag{13}$$

The integral limits are $\int_{-k_z^{\max}}^{+k_z^{\max}}\mathrm{d}k_z'$, $\int_{-k_r^{\max}}^{+k_r^{\max}}\mathrm{d}k_r'$, $\int_0^{L/2}\mathrm{d}\zeta$, and $\int_0^{R/2}\rho\,\mathrm{d}\rho$ for a cylindrical system of length L and radius R, recognizing that $r > 0$ always. It is worthwhile to note here that since the momentum variable comes from the Fourier transform $\int \mathrm{d}\mathbf{r}e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}}$, the value of $k^{\max}$ is determined by the spatial length of the box. This will become important when the problem is discretized.

As others have previously done [3,5], scattering is included simply by the addition of a relaxation time approximation, given as

$$\left.\frac{\mathrm{d}f(r,z,k_z,k_r)}{\mathrm{d}t}\right|_{\text{coll}} = \frac{1}{\tau}\left(f_0(r,z,k_z,k_r)\frac{\int|k_r'|\,\mathrm{d}k_r'\int\mathrm{d}k_z' f(r,z,k_z',k_r')}{\int|k_r'|\,\mathrm{d}k_r'\int\mathrm{d}k_z' f_0(r,z,k_z',k_r')}-f(r,z,k_z,k_r)\right), \tag{14}$$

where the relaxation time, τ, is computed from the material parameters describing scattering due to: ionized impurities and longitudinal, piezoelectric, acoustic and optical phonons.

3. Discretization

The WFE in discretized (matrix) form is written as $\frac{\mathrm{d}\mathbf{f}}{\mathrm{d}t} = (\mathbf{T}+\mathbf{U}+\mathbf{S})\mathbf{f}-\mathbf{B}$, where $\mathbf{T}$ represents the drift (kinetic) operator, $\mathbf{U}$ the potential operator, $\mathbf{S}$ the scattering (interaction) operator, $\mathbf{f}$ the Wigner function and $\mathbf{B}$ the boundary conditions arising from the drift term. In this section, we will present the details of our discretization of the equations derived above.

3.1. Variables, operators and functions

The space and momentum variables are discretized as

$$z(i) = \frac{1}{2}(2i-1)\Delta z, \quad i = 1..N_z, \;\; \Delta z = L/N_z, \tag{15}$$

$$r(n) = \frac{1}{2}(2n-1)\Delta r, \quad n = 1..N_r, \;\; \Delta r = R/N_r, \tag{16}$$

$$\phi(m) = (m-1)\Delta\phi, \quad m = 1..N_\phi, \;\; \Delta\phi = 2\pi/N_\phi, \tag{17}$$

$$k_z(j) = \frac{1}{2}(2j-N_{k_z}-1)\Delta k_z, \quad j = 1..N_{k_z}, \;\; \Delta k_z = \pi/\Delta z N_{k_z}, \tag{18}$$

$$k_r(l) = \frac{1}{2}(2k-N_{k_r}-1t)\Delta k_r, \quad k = 1..N_{k_r}, \;\; \Delta k_r = \pi/\Delta r N_{k_r} \tag{19}$$

and the functions f, $\mathscr{V}$, and $\mathscr{U}$ are discretized as

$$f(r,z,k_z,k_r) \to f(n,i,j,l), \tag{20}$$
$$V(r-r',z-z') \to V(n-n',i-i'), \tag{21}$$

$$\mathscr{U}(r,r',z,k_z,k_r) \to \mathscr{U}(n,n',i,j,l), \tag{22}$$
$$\mathscr{V}(z,z',r,r') \to \mathscr{V}(i,i',n,n'), \tag{23}$$
$$\mathscr{F}(r,r',z,k_z) \to \mathscr{F}(n,n',i,j). \tag{24}$$

We note here that since ρ and ζ are on the same grid as r and z, they will be denoted as r' and z', respectively.

3.2. 2D Poisson equation

We use a Fourier transform method to solve the 2D Poisson equation. The Fourier transform of the charge density with respect to z, $\bar{\rho}_e$, is calculated using a fast Fourier sin transform (sinFFT). Each shell exchanges $\bar{\rho}_e$ to every other shell so that each shell knows $\bar{\rho}_e(r)$, the electron density of the entire cylinder. Using a standard tridiagonal solver, the sinFFT of the 2D potential, $\bar{\phi}(r)$, is calculated, then transformed back (via a sinFFT) to the full 2D potential, $\phi(r,z)$.

3.3. Drift and boundary conditions terms

We will separate the drift (or kinetic) term into longitudinal and radial components and treat each one slightly differently. We do this by breaking up

$$[\mathbf{T}\cdot\mathbf{f}](r,z,k_z,k_r) = -\frac{\hbar}{m^*}\left[k_r\frac{\partial}{\partial r} + k_z\frac{\partial}{\partial z}\right]f(r,z,k_z,k_r) \tag{25}$$

into

$$[\mathbf{T_z}\cdot\mathbf{f}](r,z,k_z,k_r) = -\frac{\hbar k_z}{m^*}\frac{\partial}{\partial z}f(r,z,k_z,k_r), \tag{26}$$
$$[\mathbf{T_r}\cdot\mathbf{f}](r,z,k_z,k_r) = -\frac{\hbar k_r}{m^*}\frac{\partial}{\partial r}f(r,z,k_z,k_r). \tag{27}$$

First the longitudinal drift term, Eq. (26), is computed using a second order "upwind/downwind" differencing scheme:

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} \simeq \mp\frac{1}{2\Delta x}[3f(x) - 4f(x\pm\Delta x) + f(x\pm 2\Delta x)]. \tag{28}$$

For $k_z < 0$, the upwind scheme is used, and for $k_z > 0$, the downwind scheme is used, giving

$$k_z \lessgtr 0 : \frac{\partial}{\partial z}f(r,z,k_z,k_r) \to \mp\frac{1}{2\Delta z}[3f(r,z,k_z,k_r) - 4f(r,z\pm\Delta z,k_z,k_r) + f(r,z\pm 2\Delta z,k_z,k_r)]. \tag{29}$$

Which, when discretized, gives $(C_j = \frac{\hbar\Delta k_z}{4m^*\Delta z}(2j - N_{k_z} - 1))$

$$j \lessgtr \frac{N_{k_z}}{2} : [\mathbf{T_z}\cdot\mathbf{f}](n,i,j,l) \to \pm\frac{1}{2}C_j[3f(n,i,j,l) - 4f(n,i\pm 1,j,l) + f(n,i\pm 2,j,l)]. \tag{30}$$

When the second order differencing scheme gets to the boundary, it is advantageous to have the function extend only one unit distance into the boundary. For the upwind scheme, this occurs at $i = Nz - 1$ and $i = Nz$, and for the downwind scheme at $i = 1$ and $i = 2$. (the n and l indexes will be dropped since there is no dependence on them). When $i = 1, N_z$ a first order upwind/downwind differencing scheme was employed ($\frac{\mathrm{d}f(x)}{\mathrm{d}x} \simeq \mp\frac{f(x)-f(x\pm 1)}{\Delta x}$) in order to preserve the continuity of the derivative. By saying that the distribution function past the boundaries have a constant value, $f(i = N_z + 1, j) = f(i = 0, j) = f_{\text{Fermi}}(j)$ (the two-dimensional Fermi distribution), these positions become constant longitudinal boundary conditions, $\mathbf{B_z}(i,j)$, defined as:

$$j > \frac{N_{k_z}}{2}:$$

$$\mathbf{B}_z(i = 1, j) = +2C_j f_{\text{Fermi}}(j), \tag{31}$$

$$\mathbf{B_z}(2, j) = -C_j f_{\text{Fermi}}(j), \tag{32}$$

$$\vdots$$

$$j \leqslant \frac{N_{k_z}}{2}:$$

$$\mathbf{B_z}(N_z - 1, j) = +C_j f_{\text{Fermi}}(j), \tag{33}$$

$$\mathbf{B_z}(N_z, j) = -2C_j f_{\text{Fermi}}(j). \tag{34}$$

This makes the final form of the longitudinal drift term, $[\mathbf{T_z} \cdot \mathbf{f}](i,j)$, as:

$$j > \frac{N_{k_z}}{2}:$$

$$[\mathbf{T_z} \cdot \mathbf{f}](i = 1, j) = -2C_j[2f(i = 1, j)], \tag{35}$$

$$[\mathbf{T_z} \cdot \mathbf{f}](2, j) = -C_j[3f(i = 2, j) - 4f(i = 1, j)], \tag{36}$$

$$[\mathbf{T_z} \cdot \mathbf{f}](i, j) = -C_j[3f(i, j) - 4f(i, j) + f(i, j)], \tag{37}$$

$$\vdots$$

$$j \leqslant \frac{N_{k_z}}{2}:$$

$$[\mathbf{T_z} \cdot \mathbf{f}](i, j) = C_j[3f(i, j) - 4f(i, j) + f(i, j)], \tag{38}$$

$$[\mathbf{T_z} \cdot \mathbf{f}](N_z - 1, j) = C_j[3f(i = N_z - 1, j) - 4f(i = N_z, j)], \tag{39}$$

$$[\mathbf{T_z} \cdot \mathbf{f}](N_z, j) = 2C_j[2f(i = N_z, j)]. \tag{40}$$

Eqs. (31)–(34) completely define the discretized longitudinal boundary conditions, while Eqs. (35)–(37) completely define the longitudinal drift for positive momenta and Eqs. (38)–(40) completely define the longitudinal drift for negative momenta.

Next, the radial drift term, $\mathbf{T_r} \cdot \mathbf{f}$, is computed using a first order differencing scheme since having each radial shell talking only to its nearest neighbor is needed when this algorithm is parallelized. For a shell that has neighbors on both sides, a central differencing scheme (CDS) is used. For the innermost and outermost shell, a forward/backwards differencing scheme (FBDS) is employed (the indexes i,j are omitted since there is no dependence on them) ($C = (\frac{\hbar k_r}{m^* \Delta r})$):

$$[\mathbf{T_r} \cdot \mathbf{f}](n = 1, l) \to -C[f(n = 1, l) - f(n = 2, l)], \tag{41}$$

$$[\mathbf{T_r} \cdot \mathbf{f}](n, l) \to -\frac{C}{2}[f(n + 1, l) - f(n - 1, l)], \tag{42}$$

$$[\mathbf{T_r} \cdot \mathbf{f}](n = N_r, l) \to C[f(n = N_r, l) - f(n = N_r - 1, l)]. \tag{43}$$

This form dictates that some tricks must be used at the innermost and outermost rings. For the innermost shell, it can be imagined that any particle possessing negative momenta (traveling inwards) will pass through the middle and then posses positive momenta (traveling outwards). This demands a ''particle mirror'' at the origin by, basically, saying that all values of $f(1, l \leqslant \frac{N_{k_r}}{2})$ at time t will be added to $f(1, l > \frac{N_{k_r}}{2})$ at time $t + \Delta t$. For the outermost shell, all values of $f(N_r, l > \frac{N_{k_r}}{2})$ at time t will be subtracted from $f(1, l > \frac{N_{k_r}}{2})$ at time $t + \Delta t$ and, depending on the chosen external conditions, given values of $f(N_r, l \leqslant \frac{N_{k_r}}{2})$ will be added at each time step. This is expressed as radial boundary conditions as ($C_l = \frac{\hbar \Delta k_r}{m^* \Delta r}(2l - N_{k_r} - 1)$):

$$\mathbf{B_r}\left(n = 1, l \leqslant \frac{N_{k_r}}{2}\right) = -C_l\left[-f\left(n = 1, l \leqslant \frac{N_{k_r}}{2}\right)\right], \tag{44}$$

$$\mathbf{B_r}\left(n = 1, l > \frac{N_{k_r}}{2}\right) = -C_l\left[+f\left(n = 1, l \leqslant \frac{N_{k_r}}{2}\right)\right], \tag{45}$$

$$\mathbf{B_r}\left(n = N_r, l > \frac{N_{k_r}}{2}\right) = +C_l\left[-f\left(n = N_r, l > \frac{N_{k_r}}{2}\right)\right], \tag{46}$$

$$\mathbf{B_r}\left(n = N_r, l \leqslant \frac{N_{k_r}}{2}\right) = +C_l[+f_{\text{Given}}(l)]. \tag{47}$$

3.4. Potential term

The potential term in Eq. (10) is written in operator form as:

$$[\mathbf{U} \cdot \mathbf{f}](r, z, k_z, k_r) \equiv +\frac{1}{\pi \hbar} \int_{-k_z^{\max}}^{+k_z^{\max}} \mathrm{d}k_z' \int_0^{R/2} |r'| \, \mathrm{d}r' \mathscr{U}(r, r', z, k_z - k_z', k_r) \mathscr{F}(r, r', z, k_z'), \tag{48}$$

where

$$\mathscr{V}(z, z', r, r') = \int_0^{2\pi} \mathrm{d}\phi' \{V(z + z', r + r' \cos \phi') - V(z - z', r - r' \cos \phi')\}, \tag{49}$$

$$\mathscr{F}(r, r', z, k_z) = \int_{-k_r^{\max}}^{+k_r^{\max}} |k_r'| \, \mathrm{d}k_r' J_0(2k_r' r') f(r, z, k_z, k_r'), \tag{50}$$

$$\mathscr{U}(r, r', z, k_z, k_r) = \int_0^{L/2} \mathrm{d}z' \sin(2k_z z') J_0(2k_r r') \mathscr{V}(z, z', r, r'). \tag{51}$$

For a given longitudinal position and radius, (z, r), one sweeps through disks of constant z'. For each disk of constant z', the potential difference contribution, $V(r + r'\cos\phi') - V(r - r'\cos\phi')$, between all points on the disk on (z, r) is calculated (see Fig. 1(b)). The result is that the effect of the potential of the entire cylinder on a given point (z, r) operates on the distribution function for that point. The terms $\mathscr{U}, \mathscr{F}$ and $\mathscr{V}$ each present some difficulties, but they only involve one integral each. It is best to go through each one separately.

The $\mathscr{V}$ term, when discretized, becomes

$$\mathscr{V}(i, i', n, n') = \Delta\phi \sum_{m'=1}^{N_\phi} \{V(i + i', n + n'') - V(i - i', n - n'')\}, \tag{52}$$

where

$$n'' = n' \mathrm{INT}[\cos \phi(m')] = n' \mathrm{INT}[\cos([m' - 1]\Delta\phi)], \tag{53}$$

INT[x] being a function returning the nearest integer to the real value x.

We rewrite the potential term, by introducing new functions, as

$$[\mathbf{U}\cdot\mathbf{f}](r,z,k_z,k_r)=\frac{1}{\pi\hbar}\int_{-k_z^{\max}}^{+k_z^{\max}}\mathrm{d}k_z'\int_{-k_r^{\max}}^{+k_r^{\max}}\mathrm{d}k_r'U(r,z,k_z-k_z',k_r,k_r')f(r,z,k_z',k_r'), \tag{54}$$

$$U(r,z,k_z,k_r,k_r')=|k_r'|\int_0^{R/2}\mathrm{d}r'|r'|J_0(2k_r'r')J_0(2k_rr')\mathscr{P}(r,r',z,k_z), \tag{55}$$

$$\mathscr{P}(r,r',z,k_z)=\int_0^{L/2}\mathrm{d}z'\sin(2k_zz')\mathscr{V}(z,z',r,r'), \tag{56}$$

where $\mathscr{V}(z,z',r,r')$ is defined as above. In discretized form, these equations become

$$[\mathbf{U}\cdot\mathbf{f}](n,i,j,l)=\frac{\pi^2}{\hbar N_{k_z}N_{k_r}^2}\sum_{j'}\sum_{l'}U(n,i,j-j',l,l')f(n,i,j',l'), \tag{57}$$

$$U(n,i,j,l,l')=|(2l'-N_{k_r}-1)|\sum_{n'}\mathscr{J}(n',l,l')\mathscr{P}(n,n',i,j), \tag{58}$$

$$\mathscr{J}(n,l,l')=|(2n'-1)|J_0\left(\frac{\pi}{N_{k_r}}(2l'-N_{k_r}-1)(2n-1)\right)J_0\left(\frac{\pi}{N_{k_r}}(2l-N_{k_r}-1)(2n-1)\right), \tag{59}$$

$$\mathscr{P}(n,n',i,j)=\sum_{i'}\sigma(i,j)\mathscr{V}(i,i',n,n'), \tag{60}$$

$$\sigma(i,j)=\sin\left(\frac{2\pi}{N_{k_z}}j(2i-1)\right). \tag{61}$$

In performing all these calculations, it is important to remember that in current computing platforms, memory is both cheap and nicely managed so any number of the terms in these arrays can be calculated once or once per cycle and stored in advance. By employing distributed computing techniques, a dedicated CPU can spend its time calculating these arrays while other parts of the program are running. Parallelization of this algorithm will be discussed in detail below.

3.5. Interaction term

The scattering term is written using the relaxation time approximation

$$\left.\frac{\mathrm{d}f(r,z,k_z,k_r)}{\mathrm{d}t}\right|_{\mathrm{coll}}=\frac{1}{\tau}\left(\frac{f_0(r,z,k_z,k_r)}{\int|k_r'|\,\mathrm{d}k_r'\int\mathrm{d}k_z'f_0(r,z,k_z',k_r')}\int|k_r'|\,\mathrm{d}k_r'\int\mathrm{d}k_z'f(r,z,k_z',k_r')-f(r,z,k_z,k_r)\right), \tag{62}$$

where $f_0(r,z,k_z',k_r')$ is the equilibrium WDF. This term, discretized, becomes

$$[\mathbf{S}\cdot\vec{f}](n,i,j,l)=\frac{\beta(n,i,j,l)}{\tau}\sum_{j'=1}^{N_{k_z}}\sum_{l'=1}^{N_{k_r}}|(2l'-N_{k_r}-1)|f(n,i,j',l')-\frac{1}{\tau}f(n,i,j,l),$$

$$\beta(n,i,j,l)=\frac{f_0(n,i,j,l)}{\sum_{j'=1}^{N_{k_z}}\sum_{l'=1}^{N_{k_r}}|(2l'-N_{k_r}-1)|f_0(n,i,j',l')}. \tag{63}$$

4. 2D matrix setup

It is important that the matrix $\Omega=\mathbf{T}+\mathbf{U}+\mathbf{S}$ be such that the limitations of modern computer systems are able to handle the above equations in a reasonable amount of time (days and weeks as opposed to months). We know of two general methods that are able to efficiently solve the system of equations generated by the discretization of the WFE: Matrix inversion techniques and direct integration techniques (explicit Runge–Kutta-like, implicit BDF/Adams, etc). For the matrix inversion methods, a necessary condition is that either the whole matrix can to be stored in RAM or can be broken up into parts that can be separately stored in RAM and solved for sequentially (i.e., block diagonal matrix). The necessary condition for direct integration techniques is simply that the technique be stable and fast enough.

Each processor is handling the $1x + 2k$ problem, which implies a large number or matrix elements/equations. This number is too large to satisfy the above conditions. In order to remedy this, we try to find a form of the matrix Ω that is block diagonal in k_r, similar to the 1D problem, but with slightly different terms. We find that the potential term $\mathbf{U}$ cannot be expressed in such a way, and present an alternative method of solution, using approximation techniques, in Section 5.

4.1. Kinetic matrix (longitudinal)

Following the procedure well defined in previous literature [3,5,11], we say $[f(i,j)]_l$ is written in vector form as $[f_{1,1}f_{1,2}f_{1,3}\cdots f_{1,N_{k_z}}f_{2,1}\cdots f_{i,j-1}f_{i,j}f_{i,j+1}\cdots f_{N_z,N_{k_z}}]_l^{\mathrm{T}}$ so we can write that, for a given i, the longitudinal drift term can be written in matrix form as (< and > denotes downwind and upwind differentiation, respectively.)

$$\mathcal{T}_n^{>}\vec{f_{i,j,l}} = C\begin{bmatrix} 2T_n & & & & \\ & \ddots & & & \\ & & T_n & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}\begin{bmatrix} f_{i,1,l} \\ \vdots \\ f_{i,\frac{N_{k_z}}{2},l} \\ f_{i,\frac{N_{k_z}}{2}+1,l} \\ \vdots \\ f_{i,N_{k_z},l} \end{bmatrix}_l \quad \text{and} \quad \mathcal{T}_n^{<}\vec{f_{i,j,l}} = C\begin{bmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & T_n & \\ & & & & \ddots \\ & & & & & 2T_n \end{bmatrix}\begin{bmatrix} f_{i,1,l} \\ \vdots \\ f_{i,\frac{N_{k_z}}{2},l} \\ f_{i,\frac{N_{k_z}}{2}+1,l} \\ \vdots \\ f_{i,N_{k_z},l} \end{bmatrix}_l, \tag{64}$$

$\mathcal{T}_n^{\lessgtr}$ being a diagonal square matrix of size N_k, and $C = \frac{\hbar\Delta k_z}{4m^*\Delta z}$. The values of $T_0^{\lessgtr} = \pm 3$, $T_1^{\lessgtr} = \mp 4$, $T_2^{\lessgtr} = \pm 1$ are defined by their position in the complete matrix, as stated by the second order differencing scheme and the boundaries. By denoting $[f]_i$ as a vector of length N_k holding all the momentum (j) values of f_{ij} for a given i, the entire $\mathbf{T}\cdot\overrightarrow{f}$ term is written as

$$[\mathbf{T}\cdot\vec{f}]_l = C_j\left(\begin{bmatrix} \mathcal{T}_0^{<} & \mathcal{T}_1^{<} & \mathcal{T}_2^{<} & & \\ & \ddots & \ddots & \ddots & \\ & & \mathcal{T}_0^{<} & \mathcal{T}_1^{<} & \mathcal{T}_2^{<} \\ & & & \mathcal{T}_0^{<} & \mathcal{T}_1^{<} \\ & & & & \mathcal{T}_0^{<} \end{bmatrix} - \begin{bmatrix} \mathcal{T}_0^{>} & & & & \\ \mathcal{T}_1^{>} & \mathcal{T}_0^{>} & & & \\ \mathcal{T}_2^{>} & \mathcal{T}_1^{>} & \mathcal{T}_0^{>} & & \\ & \ddots & \ddots & \ddots & \\ & & \mathcal{T}_2^{>} & \mathcal{T}_1^{>} & \mathcal{T}_0^{>} \end{bmatrix}\right)\begin{bmatrix} [f]_{1,l} \\ \vdots \\ \vdots \\ \vdots \\ [f]_{1,N_{k_r}} \end{bmatrix}_l, \tag{65}$$

where $C_j = (2j - N_{k_z} - 1)C$, and $\mathbf{T}_l$ is a block tri-diagonal square matrix of rank $N_zN_{k_z}$.

Concerning the radial terms, no matrix operations are needed. Each ring receives the Wigner function distribution from its nearest inner and outer neighbor and the first order differencing scheme (described above) is used.

4.2. Potential matrix

Now, when $[f(i,j)]_{l'}$ is expressed in vector form (for a given r and k_r) as $[f_{1,1}f_{1,2}f_{1,3}\cdots f_{1,N_{k_z}}f_{2,1}\cdots f_{i,j-1}f_{i,j}f_{i,j+1}\cdots f_{N_z,N_{k_z}}]^{\mathrm{T}}_{l'}$, we can write Eq. (57) for a given j as

$$[\mathbf{V}(i,l')\cdot\vec{f_{i,j}}]_{n,l'} = C\begin{bmatrix} V_{n,l}(i,1-1,l') & V_{n,l}(i,1-2,l') & \cdots & V_{n,l}(i,1-[N_{k_z}-1],l') & V_{n,l}(i,1-N_{k_z},l') \\ V_{n,l}(i,2-1,l') & V_{n,l}(i,2-2,l') & \cdots & V_{n,l}(i,2-[N_{k_z}-1],l') & V_{n,l}(i,2-N_{k_z},l') \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ V_{n,l}(N_{k_z}-1-1,l') & V_{n,l}(i,N_{k_z}-1-2,l') & \cdots & V_{n,l}(i,N_k-1-[N_{k_z}-1],l') & V_{n,l}(i,N_{k_z}-1-N_{k_z},l') \\ V_{n,l}(i,N_{k_z}-1,l') & V_{n,l}(i,N_{k_z}-2,l') & \cdots & V_{n,l}(i,N_k-[N_{k_z}-1],l') & V_{n,l}(i,N_{k_z}-N_{k_z},l') \end{bmatrix}\begin{bmatrix} f_{i,1} \\ f_{i,2} \\ \vdots \\ f_{i,N_{k_z-1}} \\ f_{i,N_{k_z}} \end{bmatrix}_{l'},$$

where $V_{n,l}(i,j-j',l')=\sum_{n'}\mathscr{I}(n',l,l')\sum_{i'}\sigma(i',j-j')\mathscr{V}(i,i',n,n')$ and $C=\frac{\pi^2}{\hbar N_{k_r}^2 N_{k_z}}$. Since $V_{n,i}(i,j,l')\propto\sin j$, $V_{n,l}(i,-j,l')=-V_{n,l}(i,j,l')$ and $V_{n,l}(i,0,l')=0$, the above becomes

$$[\mathbf{V}(i,l')\cdot\vec{f_{i,j}}]_{n,l'} = C\begin{bmatrix} 0 & -V_{n,l}(i,1,l') & \cdots & -V_{n,l}(i,N_{k_z}-2,l') & -V_{n,l}(i,N_{k_z}-1,l') \\ V_{n,l}(i,1,l') & 0 & \cdots & -V_{n,l}(i,N_{k_z}-3,l') & -V_{n,l}(i,N_{k_z}-2,l') \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ V_{n,l}(i,N_{k_z}-2,l') & V_{n,l}(i,N_{k_z}-3,l') & \cdots & 0 & -V_{n,l}(i,1,l') \\ V_{n,l}(i,N_{k_z}-1,l') & V_{n,l}(i,N_{k_z}-2,l') & \cdots & V_{n,l}(i,1,l') & 0 \end{bmatrix}\begin{bmatrix} f_{i,1} \\ f_{i,2} \\ \vdots \\ f_{i,N_{k_z-1}} \\ f_{i,N_{k_z}} \end{bmatrix}_{l'}. \tag{66}$$

Note that $\mathbf{V}(i,l')$ is a $N_{k_z}\times N_{k_z}$ anti-symmetric matrix.

By denoting $[f]_{i,l}$ as a vector of length $N_{k_z}N_{k_r}$ holding all the momentum (j) values of f_{ijl} for a given i,l, the entire $\mathbf{U}\cdot\vec{f}$ term is written as

$$[\mathbf{U}\cdot\vec{f}]_n = \begin{bmatrix} \mathbf{V}(1,1) & \mathbf{V}(1,2) & \cdots & \mathbf{V}(1,N_{k_r}) \\ \mathbf{V}(2,1) & \mathbf{V}(2,2) & \cdots & \mathbf{V}(2,N_{k_r}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{V}(N_z,1) & \mathbf{V}(N_z,2) & \cdots & \mathbf{V}(N_z,N_{k_r}) \end{bmatrix}\begin{bmatrix} [f]_{1,1} \\ [f]_{1,2} \\ \vdots \\ [f]_{1,N_{k_r}} \end{bmatrix}, \tag{67}$$

so that $\mathbf{U}$ is a square matrix of rank $N_zN_{k_z}N_{k_r}$.

4.3. Interaction matrix

Whereas we previously wrote the discrete interaction term in Eq. (63), we rewrite it as

$$\begin{aligned}[\mathbf{S}\cdot\vec{f}](n,i,j,l) &\equiv \frac{\beta(n,i,j,l)}{\tau}\rho(n,i)-\frac{1}{\tau}f(n,i,j,l), \quad \rho(n,i) \\ &\equiv \sum_{j'=1}^{N_{k_z}}\sum_{l'=1}^{N_{k_r}}|(2l'-N_{k_r}-1)|f(n,i,j',l'), \quad \beta(n,i,j,k) \\ &= \frac{f_0(n,i,j,l)}{\sum_{j'=1}^{N_{k_z}}\sum_{l'=1}^{N_{k_r}}|(2l'-N_{k_r}-1)|f_0(n,i,j',l')},\end{aligned} \tag{68}$$

where $\rho(n,i)$ is the density of the previous time step.

By writing $f(i,j)$ as a vector, $[f_{1,1}f_{1,2}f_{1,3}\cdots f_{1,N_k}f_{2,1}\cdots f_{i,j-1}f_{i,j}f_{i,j+1}\cdots f_{N_x,N_k}]^{\mathrm{T}}$ and by denoting $[f]_i$ as a vector of length N_k holding all the momentum (j) values of f_{ij} for a given i, the entire $\mathbf{S}\cdot\vec{f}$ term is written as

$$\left[\vec{\Sigma}\right]_{n,l}-\left[\mathbf{S}\cdot\vec{f}\right]_{n,l}=\frac{1}{\tau}\begin{bmatrix}[\rho]_1[\beta]_1\\ [\rho]_2[\beta]_2\\ \vdots\\ [\rho]_{N_z}[\beta]_{N_z}\end{bmatrix}-\begin{bmatrix}S(1)&0&\cdots&0\\ 0&S(2)&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\cdots&S(N_z)\end{bmatrix}\begin{bmatrix}[f]_1\\ [f]_2\\ \vdots\\ [f]_{N_z}\end{bmatrix}, \tag{69}$$

where $\mathbf{S}$ is a square block diagonal matrix of rank $N_Z N_{k_z}$ and $\vec{\Sigma}$ is a $N_z N_{k_z}$ vector.

4.4. Boundary conditions

By denoting $[f]_i$ as a vector of length N_k holding all the momentum (j) values of the Fermi distribution $f_{\mathrm{Fermi}}(j)$ for a given i, the longitudinal boundary equations (31)–(34) can be written as

$$\vec{\mathbf{B}}=C_j\left[B_1^{>}[f]_1B_2^{>}[f]_2\cdots B_2^{<}[f]_{N_z-1}B_1^{<}[f]_{N_z}\right]^{\mathrm{T}}, \tag{70}$$

$B_n^{\lessgtr}[f]_i$ being a vector of size N_{k_z}, $\vec{\mathbf{B}}$ a vector of size $N_z N_{k_z}$, and $C_j=\frac{\hbar\Delta k_z}{4m^*\Delta z}(2j-N_{k_z}-1)$. The values of $B_1^{\lessgtr}=\pm 2, B_2^{\lessgtr}=\mp 1$ are defined by Eqs. (31)–(34), as stated by the second order differencing scheme at the boundaries.

5. Methods of solution

5.1. Parallelization

So far, we have dealt almost exclusively with the 1D space, 2D momentum transport problem ($1x$ + $2k$). As stated above, we are working under the assumption that *longitudinal transport is more dominant than radial transport*. This allows the total transport to be calculated in two steps: (1) transport the particles in the longitudinal direction in each shell separately ($1x$ + $2k$), then (2) each shell exchanges particles with its nearest neighbor. This latter step is where we employ parallel processing techniques. The $1x$ + $2k$ problem is performed on P processors, where P is the number of cylindrical shells into which we have divided up the RTS. Once each shell has advanced a given amount of time, then communication between shells (radial drift) can commence. As described in Section 3.3, a central differencing scheme (CDS) is used and shown in Eqs. (41)–(43). The boundaries consist of the material external to the shell and the innermost shell. As per Eqs. (44)–(47) and explained in Section 3.3, the exterior boundary defines the device. For example, if the device is a mesa RTS, and there is nothing but vacuum outside the shell, one should choose a boundary shell that injects into the outermost shell the same particles that the outermost shell ejected (keeping the momenta the same). If the device is a slab with a circular contact for an emitter, then the material outside the cylinder is in equilibrium. The boundary shell would be chosen to reflect this.

5.2. Potential transform

As explained is Section 4 and seen in Eq. (67), the Wigner integral (potential term) is the only term that cannot be made diagonal in k_r, leading to a full matrix that is too big to store in memory. If one uses direct integration methods [1], the number of terms becomes large enough to make the problem intractable. We now describe a method of using the Fourier transform property of the Wigner integral to eliminate the k_r dependence inherent in Eq. (11).

We need to have either the matrix $\Omega = \mathbf{T}_z + \mathbf{U} + \mathbf{S}$ be able to be stored in the RAM of present day computers (for the matrix methods), or limit the number of simultaneous equations to solve (the direct integration methods). The idea behind our method of solution of the 2D transport equation,

$$\frac{\mathrm{d}\mathbf{f}}{\mathrm{d}t} = \Omega\mathbf{f} - \mathbf{B}, \tag{71}$$

is that if the matrix Ω is block diagonal in k_r then one can progress through all the values of k_r solving the matrix equation at fixed values of k_r each time, effectively reducing the simulation to a series of 1D problems. Unfortunately, the equations for some of the matrix operators are not block diagonal in k_r in their present form. Some manipulation will be needed to obtain block diagonal terms. The drift term is already independent of k_r, and the scattering term is coupled to k_r via off-diagonal elements due to the integral $\int |k_r'|\, \mathrm{d}k_r' \int \mathrm{d}k_z' f(r,z,k_z',k_r')$. This integral is just the 1D density in the shell $\rho_r(z)$, which can be calculated at the beginning of the time step. This approximation turns the scattering term into the desired diagonal matrix without losing much detail. The potential term, however, is not so simple. In Eq. (48), the term $\mathscr{F}(r,z,\rho,\zeta)$, as defined in Eq. (11), makes the matrix operator $\mathbf{U}$ a full matrix. We will now outline a method to circumvent this problem below.

In order to solve the transport equation, we use an implicit method (first proposed in [3]) by rewriting Eq. (71) as (dropping the z index from the potential matrix)

$$\frac{\bar{\mathbf{f}} - \mathbf{f}}{\Delta t} = (\mathbf{T} + \mathbf{U} + \mathbf{S})\frac{\vec{\mathbf{f}} + \mathbf{f}}{2} - \mathbf{B}, \tag{72}$$

where $\vec{\mathbf{f}}$ means the new (next) value of $\mathbf{f}$ in time. Rewriting it as

$$\left[1 - \frac{\Delta t}{2}(\mathbf{T} + \mathbf{U} + \mathbf{S})\right] \cdot (\bar{\mathbf{f}} + \mathbf{f}) = 2\mathbf{f} + \mathbf{B}\Delta t, \tag{73}$$

and making the approximation (accepting error in terms of order Δt^2)

$$1 - \frac{\Delta t}{2}(\mathbf{T} + \mathbf{U} + \mathbf{S}) \simeq \left[1 - \frac{\Delta t}{2}(\mathbf{T} + \mathbf{S})\right]\left[1 - \frac{\Delta t}{2}\mathbf{U}\right] \tag{74}$$

allows us to write

$$\left[1 - \frac{\Delta t}{2}(\mathbf{T} + \mathbf{S})\right]\left[1 - \frac{\Delta t}{2}\mathbf{U}\right] \cdot (\bar{\mathbf{f}} + \mathbf{f}) = 2\mathbf{f} + \mathbf{B}\Delta t. \tag{75}$$

For convenience, we will rewrite this as

$$\Omega\mathbf{f}' = \Omega_0\Omega_U\mathbf{f}' = \mathbf{b}, \tag{76}$$

where we have defined

$$\mathbf{b} = 2(\mathbf{f} + \mathbf{B}\tau), \tag{77}$$

$$\tau = \frac{\Delta t}{2}, \tag{78}$$

$$\Omega_U = (1 - \tau\mathbf{U}), \tag{79}$$

$$\Omega_0 = (1 - \tau[\mathbf{T} + \mathbf{S}]), \tag{80}$$

$$\mathbf{f}' = (\bar{\mathbf{f}} + \mathbf{f}). \tag{81}$$

The solution to this equation involves solving two matrix equations.

(1) Solve $\Omega_0\Gamma = \mathbf{b}$ for Γ (quick).
(2) Solve $\Omega_U\mathbf{f}' = \Gamma$ for $\mathbf{f}'$ (impractical).

The matrix Ω_U is still too big to store in memory and, consequently, solving $\Omega_U\mathbf{f}' = \Gamma$ for $\mathbf{f}$ is not practical for modern computers. By making the approximation that led to Eq. (75), we are able to solve $\Omega_U\mathbf{f}' = \Gamma$ separately, allowing us to reformulate it in a form more suitable for computation.

Recall that the Wigner integral, $\mathbf{U}\cdot\mathbf{f}$, was derived by taking the Fourier transforms of the Greens function, $G^<$ [7, Section 6.4]. From the last term in Eq. (10) can be written out in full as

$$\int \mathrm{d}z'\mathrm{e}^{-\mathrm{i}2k_z z'}\int |r'|\,\mathrm{d}r' J_0(2k_r r')\mathscr{V}(z,z',r,r')\int \mathrm{d}k_z'\mathrm{e}^{+\mathrm{i}2k_z' z'}\int |k_r'|\,\mathrm{d}k_r' J_0(2k_r' r')f(r,z,k_z',k_r') = \mathbf{F}\mathscr{V}\mathbf{F}^{-1}\mathbf{f}, \tag{82}$$

$$\mathscr{V}(z,z',r,r') = \int_0^{2\pi}\mathrm{d}\theta\{V(z+z',r+r'\cos\theta) - V(z-z',r-z'\cos\theta)\}. \tag{83}$$

The Fourier Bessel transform, $\mathbf{F}$, and its inverse, F^{-1}, are given by

$$\mathbf{F} \equiv \mathbf{F}(k_r,k_z;r',z') = \int \mathrm{d}z'\mathrm{e}^{-\mathrm{i}2k_z z'}\int \mathrm{d}r'|r'|J_0(2k_r r'), \tag{84}$$

$$\mathbf{F}^{-1} \equiv \mathbf{F}^{-1}(k_r,k_z;r',z') = \frac{1}{(2\pi)^2}\int \mathrm{d}k_z\mathrm{e}^{+\mathrm{i}2k_z' z'}\int \mathrm{d}k_r|k_r|J_0(2k_r r'). \tag{85}$$

When we define

$$g(\mathbf{x},\mathbf{y}) = \frac{1}{(2\pi)^3}\int \mathrm{d}^3\mathbf{k}\mathrm{e}^{2\mathrm{i}\mathbf{k}\cdot\mathbf{y}}f(\mathbf{x},\mathbf{k}) = \mathbf{F}^{-1}\mathbf{f}, \tag{86}$$

we can rewrite $\Omega_U\mathbf{f}' = \Gamma$ as

$$(1-\tau\mathbf{F}\mathscr{V}\mathbf{F}^{-1})(\mathbf{F}\mathbf{g}') = \Gamma, \tag{87}$$

$\mathbf{g}$ being the Fourier transform of $\mathbf{f}$ and $\mathbf{g}' = \bar{\mathbf{g}} + \mathbf{g}$. Some manipulation gives

$$\Omega_U'\mathbf{g}' = \gamma, \tag{88}$$

where

$$\Omega_U' = (1-\tau\mathscr{V}) \quad \text{and} \quad \gamma = \mathbf{F}^{-1}\Gamma. \tag{89}$$

The procedure is now reduced to solving the equation, $\Omega_U'\mathbf{g}' = \gamma$.

The term $\mathscr{F}(r,z,\rho,\zeta)$, as defined in Eq. (11), can now be written as

$$\mathscr{F}(r,z,\rho,\zeta) = \int |k_r'|\,\mathrm{d}k_r'\,\mathrm{d}k_z'\int_0^{2\pi}\mathrm{d}\chi_\phi'\mathrm{e}^{+2\mathrm{i}(k_z'\zeta+k_r'\cos\chi_\phi'\rho)}f(r,z,k_z',k_r',\chi_\phi') = \int \mathrm{d}^3\mathbf{k}/\mathrm{e}^{2\mathrm{i}\mathbf{k}\cdot\mathbf{y}}f(r,z,\mathbf{k})$$
$$= (2\pi)^3 g(r,z,\rho,\zeta,\theta). \tag{90}$$

This lets us define $\Omega_U' = [1-\frac{\Delta t}{2}\mathbf{U}']$ in terms of

$$\mathbf{U}'\mathbf{g} = \frac{1}{2\pi^2\hbar}\int \mathrm{d}\rho\,\mathrm{d}\zeta|\rho|\sin(2k_z\zeta)J_0(2k_r\rho)\mathscr{V}(r,z,\rho,\zeta)(2\pi)^3 g(r,z,\rho,\zeta,\theta), \tag{91}$$

or, in discrete form, as (completing Eq. (57))

$$[\mathbf{U}'\mathbf{g}](n,i,j,k) = +\frac{16\pi}{\hbar}\Delta r^2\Delta z\sum_{n',i'}\mathscr{U}(n,i,n',i',k,j)g(n,i,n',i',m'). \tag{92}$$

The WDF is recovered by

$$f(r,z,k_r,k_z) = \int_0^{2\pi} |k_r|\,\mathrm{d}\chi_\phi f(r,z,k_r,k_z,\chi_\phi) = \int_0^{2\pi} |k_r|\,\mathrm{d}\chi_\phi \int \mathrm{d}^3\mathbf{y}\mathrm{e}^{2\mathrm{i}\mathbf{k}\cdot\mathbf{y}} g(r,z,\mathbf{y})$$
$$= |k_r| \int \mathrm{d}\rho|\rho| J_0(2k_r\rho) \int \mathrm{d}\zeta \mathrm{e}^{-2\mathrm{i}k_z\zeta} g(r,z,\rho,\zeta). \tag{93}$$

The complete process to solve the WFE equation is given in the following steps

(1) Solve $\Omega_0\Gamma = \mathbf{b}$ for Γ.
(2) Define $\gamma = \mathbf{F}^{-1}\Gamma$.
(3) Solve $\Omega'_U\mathbf{g}' = \gamma$ for $\mathbf{g}'$.
(4) Recover $\mathbf{f}'$ from $\mathbf{f}' = \mathbf{F}\mathbf{g}'$.

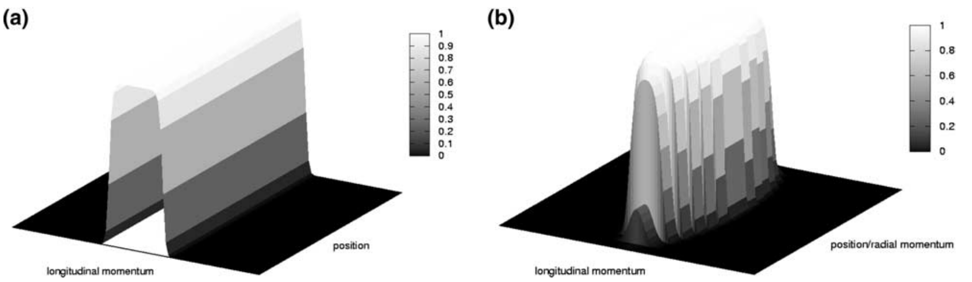

Fig. 2. Initial WDF as boundary value. (a) Longitudinal phase space for one k_r slice. (b) All k_r slices put together.

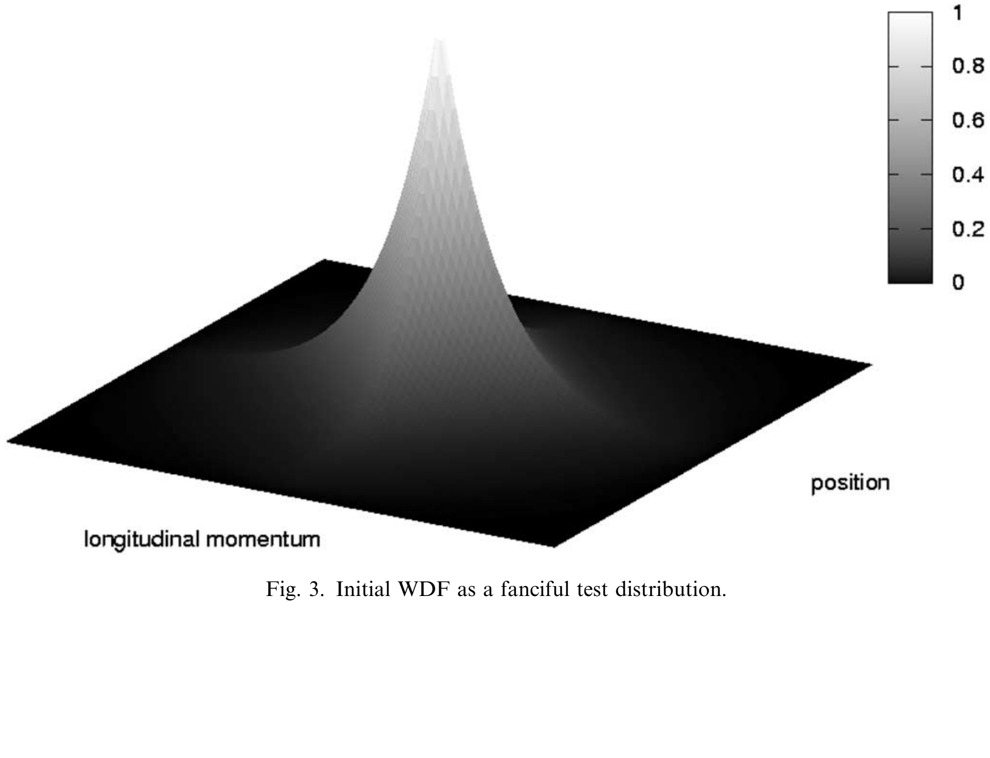

Fig. 3. Initial WDF as a fanciful test distribution.

5.3. Aim and shoot

We find that using direct integration for a system of linear equations grew slower and less stable as the number of simultaneous equations grew. For example, a 1 fs time step involved hundreds of iterations were involved, each one taking about 90 s. When using an implicit matrix method, as is standard in the 1D codes, the matrices were too big to store. Because we have now, using the approximation in Eq. (74), separated the operator Ω into the product of a drift/scattering operator, Ω_0, and a potential operator, Ω_U, a combination of these two methods can be used. As we shown, by itself, Ω_0 can be block diagonal and sparse enough to be easily solved by either of the two methods. Also, by itself, Ω_U can be rewritten in a form that also can be easily solved by either of the two methods.

The following method allows us to take advantage of this split is a way that is analogous to the accelerated convergence method used to obtain steady-state solutions in the 1D problem [3]. We have dubbed this a "aim and shoot" approach. The static potential term is calculated on a time scale of $\delta t = 0.01$ fs for only one step, then held constant while the system evolves for $\Delta t = 1$ fs. This is the aim part. The shoot part involves the drift and scattering matrices being solved implicitly by matrix inversion, with the potential term

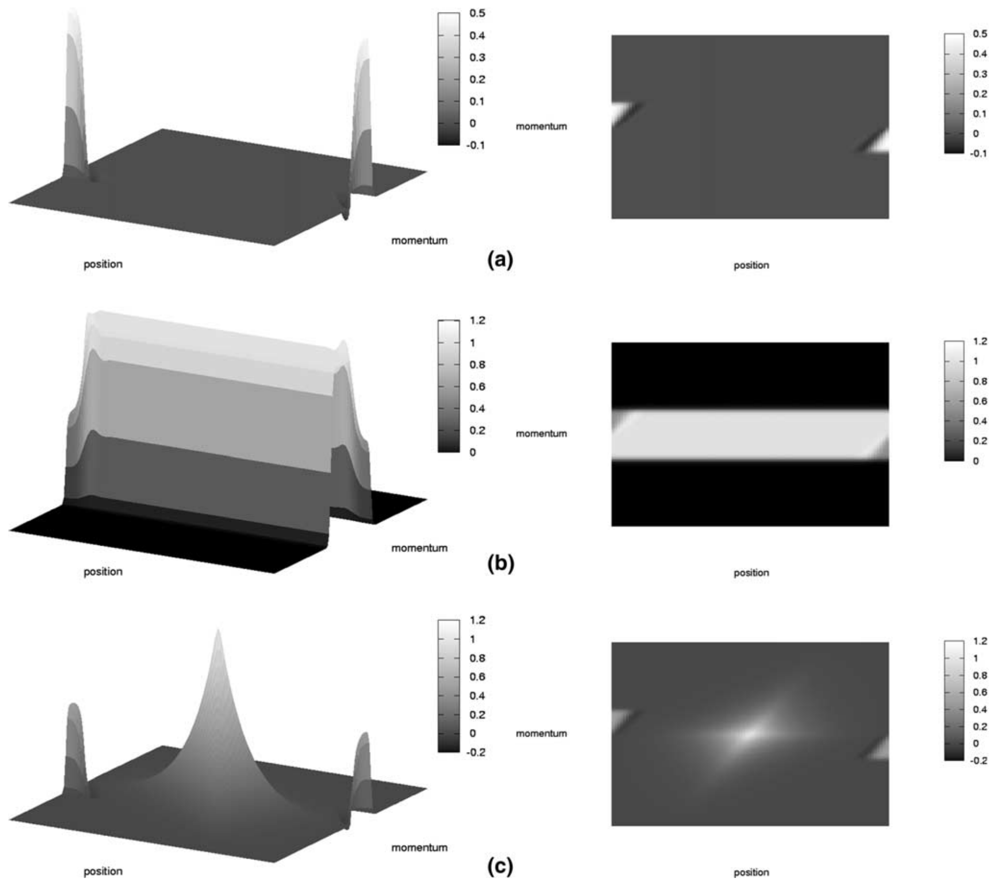

Fig. 4. Neglecting coulomb interactions: surface and contour plots of the WDF at 10 fs for the Initial WDF of (a) zero, (b) pseudoFermi, (c) central.

static. This is fine for a steady-state solution, but it is still uncertain if this method will give the proper transient behavior of the system.

6. Implementation and simulation results

In the preceding sections a computational method of solving for both the time dependent and steady state two-dimensional Wigner function transport equation was presented. The 2D equations and computational method were derived for the case of longitudinal transport through a cylinder while taking account of the effects of radial momentum in addition to the longitudinal momentum. As previously stated, the numerical solution is broken into two parts: (1) transport in the longitudinal direction then (2) transport in the radial direction. The cylinder is divided in to a number of concentric cylindrical shells in which the longitudinal transport takes place as in the 1D problem, but with the inclusion of radial momentum. The radial transport involves a simple exchange of particles (dictated by the newly calculated radial momentum). Sine this step is computationally trivial, this work was concerned with the former step: A 1D space and 2D momentum transport problem ($1x + 2k$). Below we present some proof-of-principle

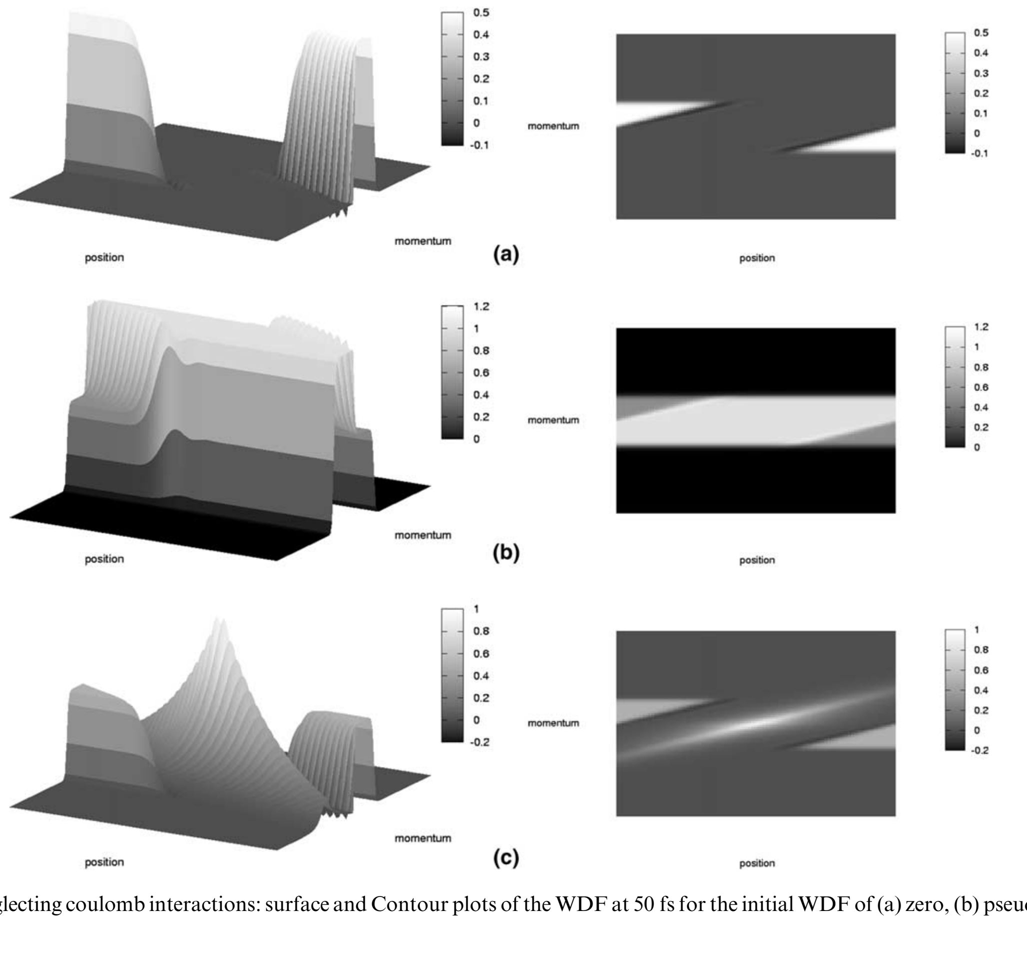

Fig. 5. Neglecting coulomb interactions: surface and Contour plots of the WDF at 50 fs for the initial WDF of (a) zero, (b) pseudoFermi, (c) central.

simulation results obtained using the methods developed in Section 5. From these results, the future usefulness of each of these methods, in light of current computing trends, will be discussed.

This simulations were performed on Linux workstations (2 GHz Pentium 4, 1.7 GHz Pentium 4s and 1.2 GHz AMD AthlonMPs). For phase space, our solution method is most easily formulated when $N_z = N_{k_z}$ and $N_r = N_{k_r}$ since the Fourier transforms between momentum space and displacement space must be on the same lattice. Simulations were performed on longitudinal phase space grid sizes of $N_z = N_{k_z} = 96$ and $N_z = N_{k_z} = 48$ with radial phase space grid sizes of $N_r = N_{k_r} = 2, 4, 8, 16, 20$. In all of the simulations presented, the device is constructed of bulk *n*-doped GaAs with a RTS of undoped GaAs/$Al_{0.3}Ga_{0.7}As$ with a barrier potential is 0.3 eV. The device temperature is 77 K, the electron effective mass is $0.0667m_0$ and the donor density is 2×10^{18} cm^{-3}. The cylinder has dimensions of 1000 Å in both length (z) and diameter (r). For most of the simulations the active RTS region is approximately 170 Å, the well, spacer and barrier lengths being 50, 30 and 30 Å, respectively.

Each numerical experiment has been carried out in the following way. First, the cylinder is populated with electrons according to one of three specific WDF: (a) $f(z, k_Z, k_r) = 0$, corresponding to no excess electrons in the cylinder, (b) $f(z, k_Z, k_r) = f_{\text{Fermi}}$, corresponding to the same distribution of the metallic leads (Fig. 2), and (c) a fanciful distribution corresponding to electrons mostly at the center of phase space (Fig. 3). To be exact, an error was made in preparing for case (b). We meant to set the WDF

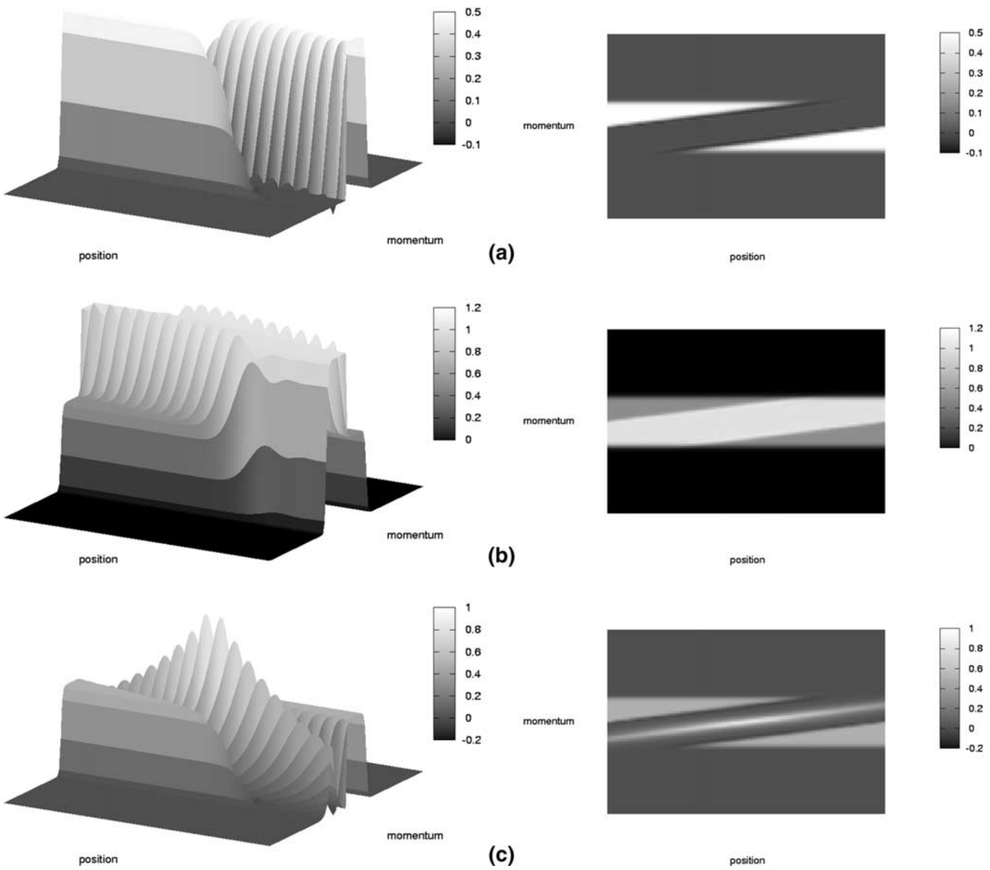

Fig. 6. Neglecting coulomb interactions: surface and contour plots of the WDF at 100 fs for the initial WDF of (a) zero, (b) pseudoFermi, (c) central.

each k_r slice to the Fermi distribution of the boundary of that specific k_r slice. This way the integral of the total WDF would be unity, as expected. Instead, we accidentally normalized the WDF of each k_r slice such that the integral of the WDF in each slice in unity. We have kept this error here for the reason that the simulations illustrates that the system will adjust itself and still tend toward the expected result.

Next, at zero bias, the system is allowed to evolve with scattering turned off. The system is allowed to evolve for a suitable time, until it settles into a steady state, and then scattering is turned on. During this time, the previous step's value of the WDF is used as the "equilibrium value" needed in the relaxation time approximation of the present step. Once again, the system is allowed to evolve for a suitable time, and at this time, the current WDF is set to be the "equilibrium value" for the rest of the evolution, which continues until the system settles into a steady state. By observing this time evolution of each of the three cases, we can determine how well a given method behaves as well as gather timing and other information. Although measurable quantities, such current and carrier densities, are calculated from the WDF by our simulation, due to space constraints, we will restrict our discussion to the evolution of the WDF only. For the test problems we examine, such quantities other than the WDF will yield little useful information into the correctness of our algorithm.

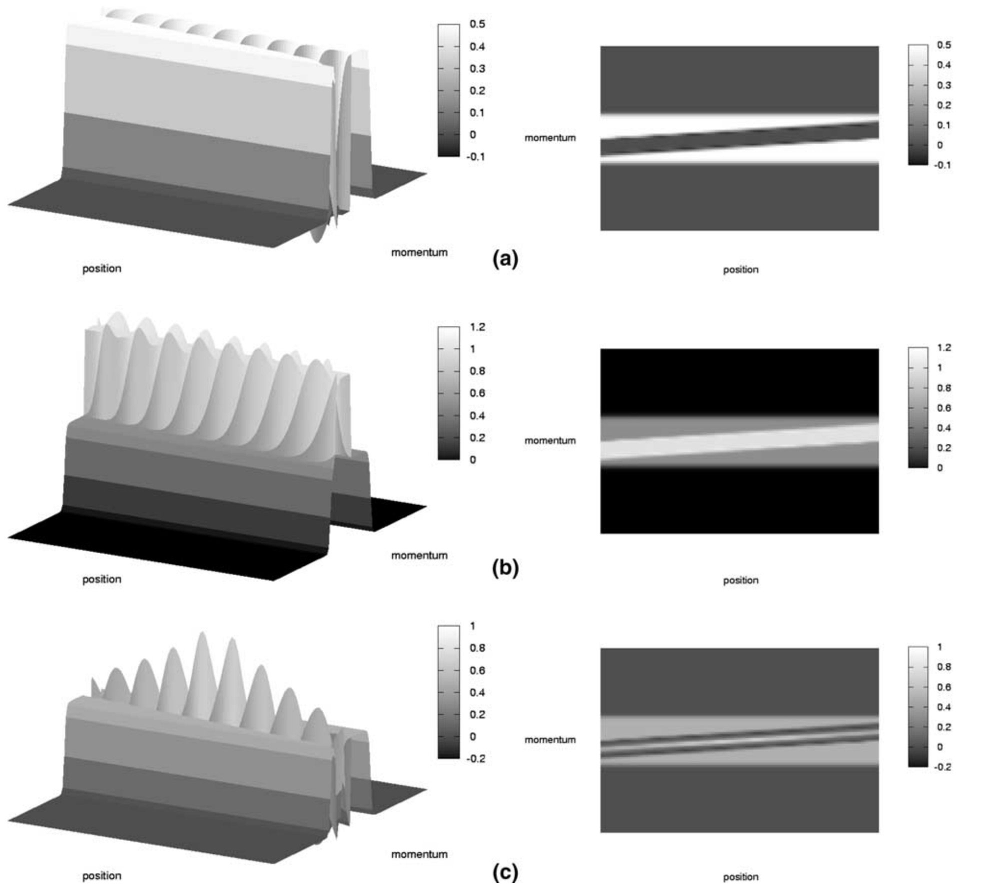

Fig. 7. Neglecting coulomb interactions: surface and contour plots of the WDF at 200 fs for the initial WDF of (a) zero, (b) pseudoFermi, (c) central.

One important item of note is that the examples below are performed with $N_z = N_{k_z} = 96$ and $N_r = N_{k_r} = 2$. In effect we are including only one positive and one negative radial momentum value. While this is fine for testing purposes where we are basically reducing the simulation to a 1D problem, for a real 2D problem N_{k_r} and N_r should be large enough ($\sim$16) to encompass a phase space greater than the radial Fermi momentum of the material outside the cylinder. The computational issue is that while $N_r = N_{k_r} = 2$ can be solved in under 20 min for a 2000 fs run at time steps of 1 fs, the addition of more radial points increases the time dramatically (this will be mentioned below). As a result, the phase space plots given are showing only one value of the radial momentum (the positive momentum) since they are symmetric in this case.

In order to follow the time evolution of this system, we have employed two different types of integration methods: Direct integration (explicit) and Matrix solvers (implicit). Direct integration is done using a prepackaged integrator called ROCK4 [1], which is a fourth order Runge–Kutta like integrator for a system of equations. With ROCK4, there is no need to compute and store the right hand side of the discretized equation as a general band matrix, and consequently, no need for a matrix inversion of the time evolution operator which have been used in previous simulations. Instead, all that is needed is to calculate the right hand side on the fly. Implicit integration is done by using LAPACK [2] to solve the matrices.

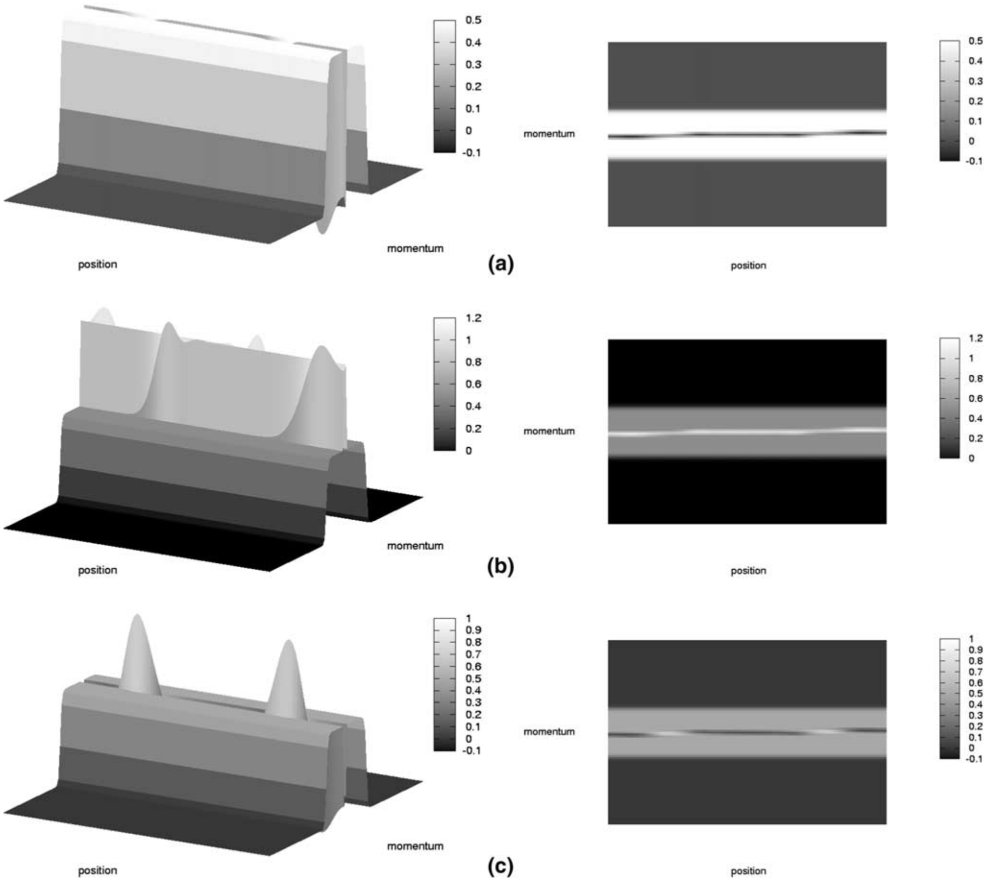

Fig. 8. Neglecting coulomb interactions: surface and contour plots of the WDF at 1000 fs for the initial WDF of (a) zero, (b) pseudoFermi, (c) central.

6.1. Direct integration methods

Not much will be said for the explicit method since, on the same computer as the runs below, after 36 min the method progressed only to a time of 14 fs. When the number of radial grid points is increased from 2 up to 8, the number of equations increases 256 times the original number. This fact renders ROCK4 useless for any future 2D simulations. Recently, we have been introduced [8] to implicit direct integration methods (BDF/Adams) and Newton solvers that, so far, outperform the ROCK4 method for the 1D simulations. We have yet to include this method in our 2D simulations. The next phase of the ongoing research is to see how such methods compare to what we will present below.

6.2. Matrix split

The so-called matrix split refers to the method described in Section 5.2, which is when taking the WFE, apply the approximation of Eq. (74) and split the matrix Ω into a product of a drift/scattering matrix and a potential matrix (Eq. (75)). This can be written as

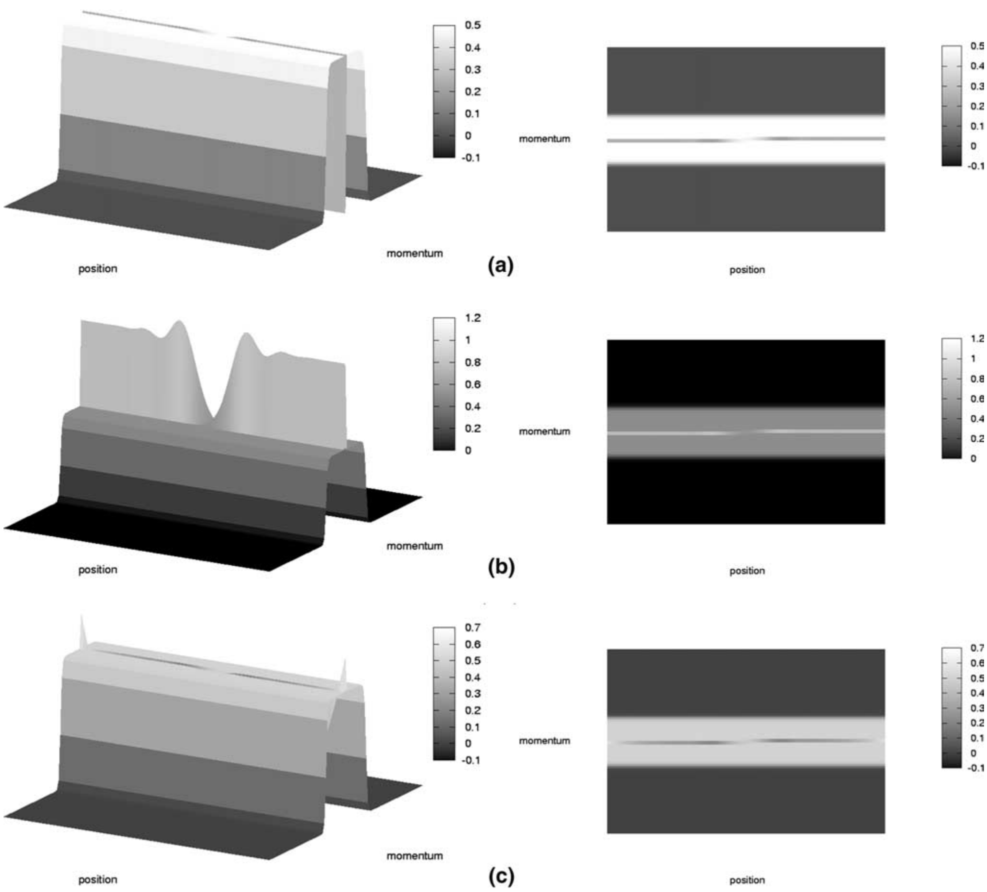

Fig. 9. Neglecting coulomb interactions: surface and contour plots of the WDF at 2000 fs for the initial WDF of (a) zero, (b) pseudoFermi, (c) central.

$$\Omega \mathbf{f}' = \Omega_0 \Omega_U \mathbf{f}' = \mathbf{b}, \tag{94}$$

The solution to this equation involves solving two matrix equations, one for the drift/scattering

$$\Omega_0 \Gamma = \mathbf{b}, \tag{95}$$

and one for the potential

$$\Omega'_U \mathbf{g}' = \gamma. \tag{96}$$

In solving the potential equation, we must perform an inverse transform, $\gamma = \mathbf{F}^{-1}\Gamma$, and then a transform, $f' = \mathbf{F}\mathbf{g}'$.

Figs. 4–15 illustrate the time evolution of our device up to 2000 fs. Each figure is a snapshot in the evolution of three WDFs whose initial values are one of the three discussed in Section 6. They will be referred to as (a) *Zero* – $f(z,k_Z,k_r) = 0$, (b) pseudoFermi – $f(z,k_Z,k_r) = f_{\mathrm{Fermi}}$, (Fig. 2) and (c) central initial values – electrons mostly at the center of phase space (Fig. 3). In each figure, the left column shows the WDF plotted in 3D, while the right column shows a contour plot of the WDF. We will examine two distinct sets of cases to illustrate our method. The first will have the potential term set to zero, corresponding to no coulombic

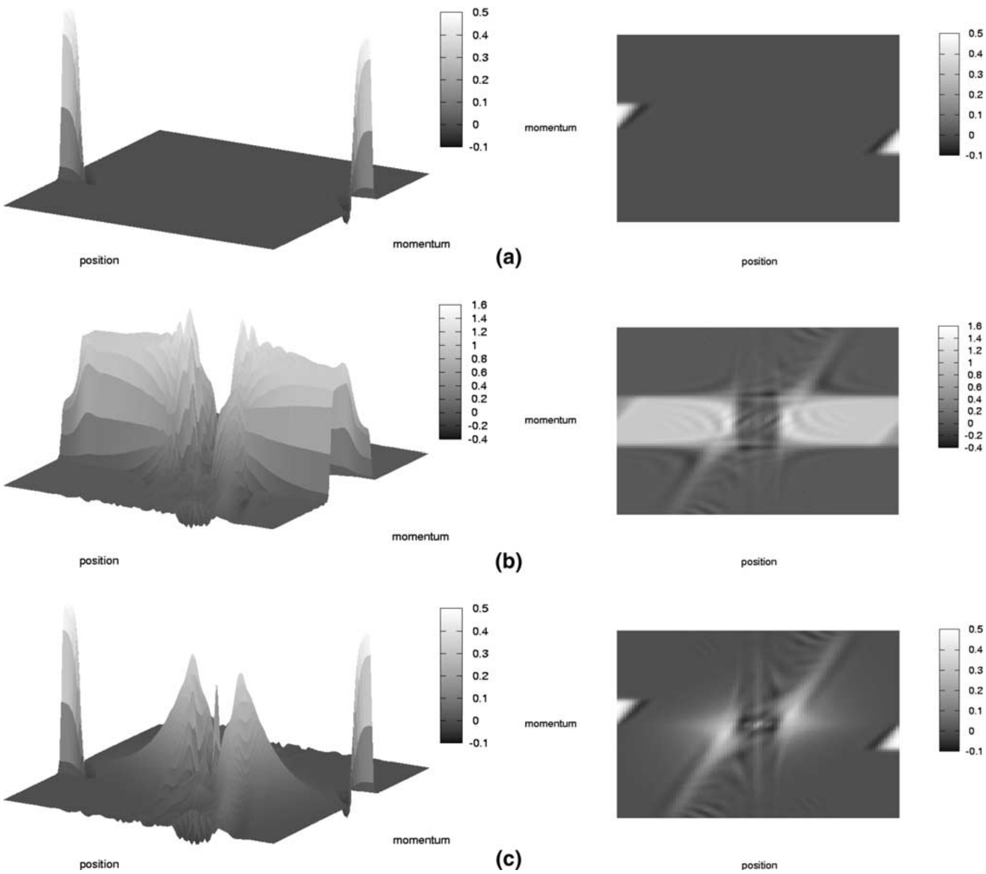

Fig. 10. Including coulomb interactions: surface and contour plots of the WDF at 10 fs for the initial WDF of (a) zero, (b) pseudoFermi, (c) central.

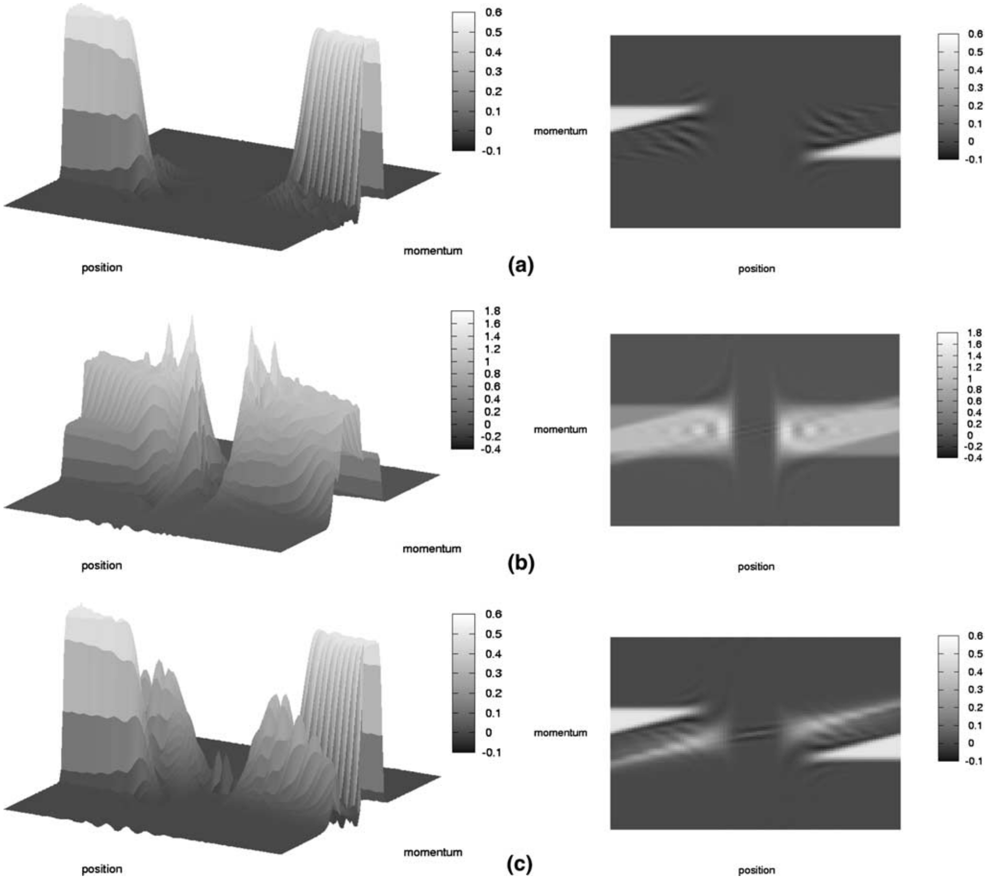

Fig. 11. Including coulomb interactions: surface and contour plots of the WDF at 50 fs for the initial WDF of (a) zero, (b) pseudoFermi, (c) central.

interaction (Note that this has the effect of rendering the barrier invisible in the constant effective mass approximation.) The second will include the coulombic effects.

6.2.1. Potential "Turned Off"

Figs. 4–9 show the evolution of the electron distribution in our device without any coulomb interactions. We see that for each of the three WDFs, the system behaves as expected. With a zero initial WDF, we see the electrons move in from either end of the device with the high momentum carriers further in than the low momentum carriers. As time progresses, the phase space fills with carriers and the WDF tends towards the distribution of the boundaries, namely, the Fermi distribution. The next case, that of the initial WDF set to the PseudoFermi distribution, shows the excess carriers leaving the system. Ultimately, the phase space once again tends towards the distribution of the boundaries. In the final case, the central distribution, we see the carriers moving in from the boundary, as in the first case. At the same time, the central distribution itself relaxes, with the positive momentum carriers moving one direction and the negative momentum carriers the other way. Eventually, the central distribution relaxes while the boundary carriers move in. Once again, the end result tends towards the boundary distribution.

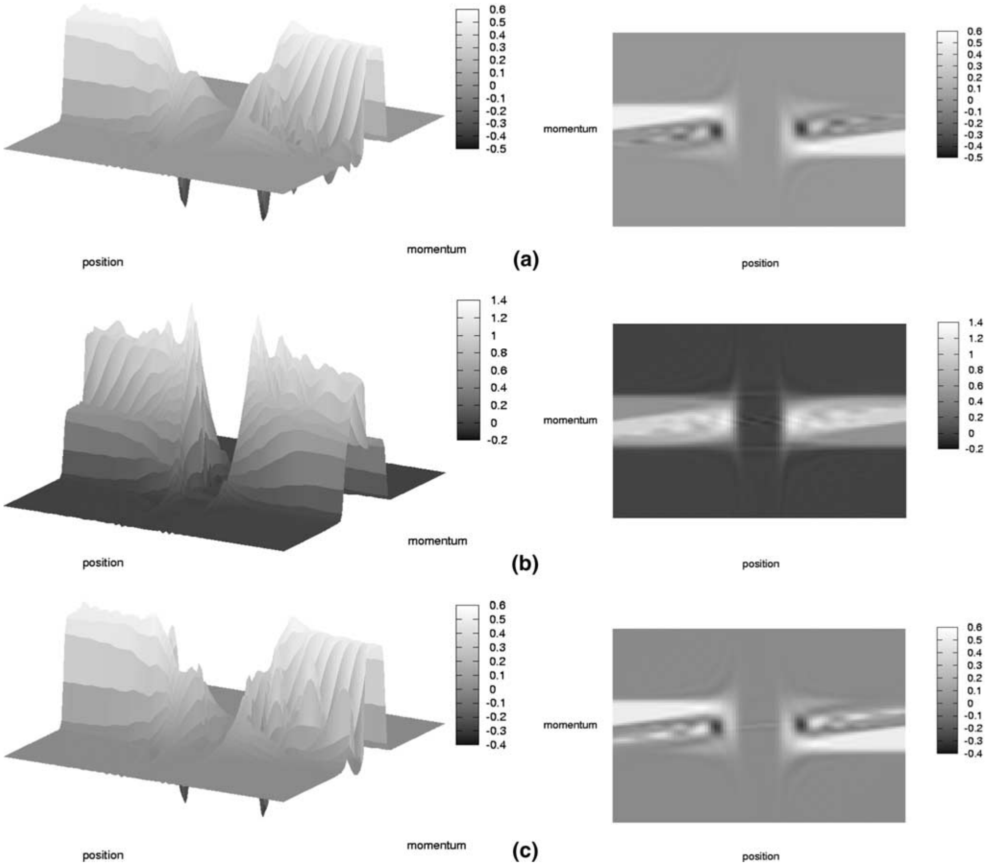

Fig. 12. Including coulomb interactions: surface and contour plots of the WDF at 100 fs for the Initial WDF of (a) zero, (b) pseudoFermi, (c) central.

All three of the different initial WDFs evolve towards the same, expected, final WDF. This simply shows that the drift/scattering part of the simulation works as expected, which is expected. What we can learn from this exercise is the CPU time to calculate the drift and scattering terms up to 200 and 2000 fs (Table 1). As stated above, this is the first of two matrix equations that must be solved. We see from the table that this is not where most of the CPU will spend its time. Rather, the potential term will take the bulk of the computing time. Next, we see how the simulation behaves when this term is turned on.

6.2.2. Potential "Turned On"

Figs. 10–15 show the evolution of the electron distribution in our device including coulomb interactions. We can compare the simulations of these three cases with coulombic interactions to those above, where coulombic interactions were ignored. By following the evolution of the first case (zero initial WDF), we see how the carriers interact with the barriers as they move towards the center of the device. We also see (more clearly in the contour plots) how the carriers interact with each other, spreading out slightly as the progress inwards. In Figs. 12 and 13 we begin to see the interaction of the reflected carriers with the incoming carriers. This effect grows as the system evolves, which is evident in the dark/light patterns along the momentum axis. The same effects are seen in the other two cases, with the exception that they begin with the unlikely distribution having carriers in the barriers. We see these carriers begin ejected from

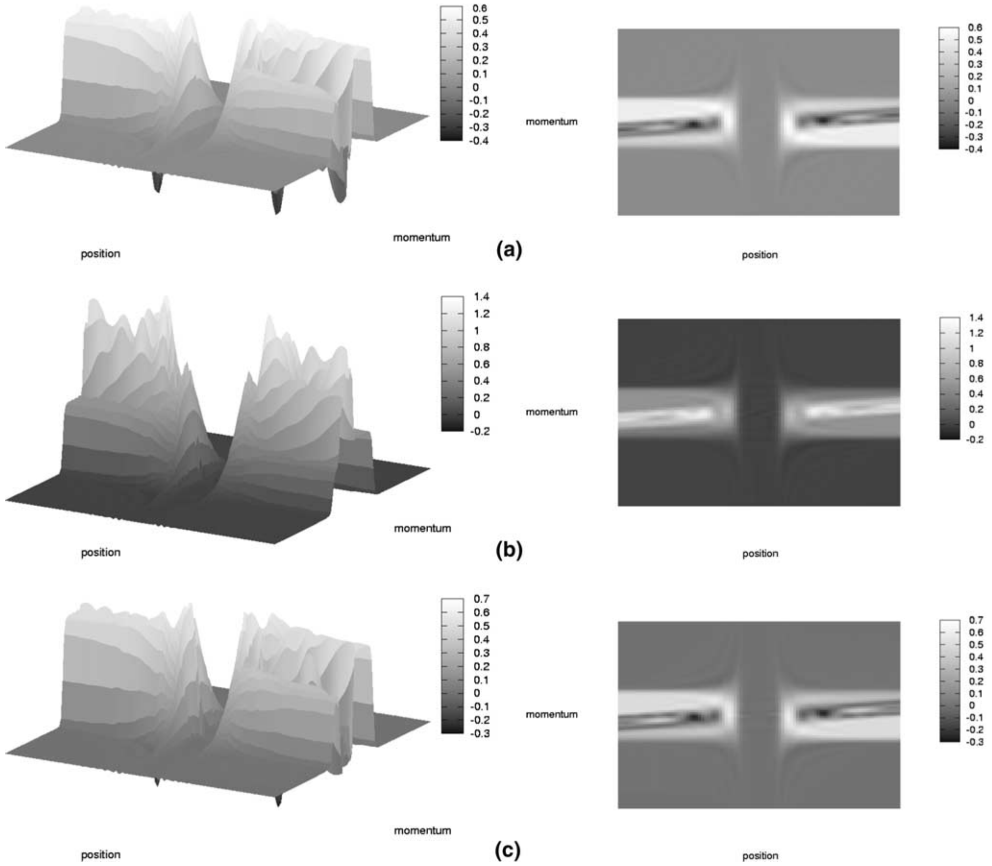

Fig. 13. Including coulomb interactions: surface and contour plots of the WDF at 200 fs for the initial WDF of (a) zero, (b) pseudoFermi, (c) central.

the barrier region at high momentum at first, then at later times these cases evolve to same result as the first case, where we see the expected WDF.

As we increase $N_r = N_{k_r}$ from 2 to 4, a 200 fs run increases from about 1 min 40 s to 5 min (a factor of 3). Projecting this to 2000 fs, we see the simulation would take approximately 47 min to complete. So far, this is not unreasonable, since for a given bias point, a 2000 fs run is enough to assure convergence. A 90 point IV curve (including the reverse sweep) would take about 70 h (about 3 days). Increase $N_r = N_{k_r}$ to 8 and a 200 fs run takes 20 min ($4 \times N_r = 4$, or $12 \times N_r = 2$). A 2000 fs run will take $3\frac{1}{3}$ h, which means a 90 point IV curve will take 12.5 days. A beginning run at $N_r = N_{k_r} = 16$ returns a time rate of 38 s/fs. At that rate, a 2000 fs run would take 21.1 h and a 90 point IV curve almost 80 days.

7. Conclusions

We have shown here a 2D drift/scattering and 3D potential form of the Wigner function transport equation for the case of a cylindrical device. This is an important step in Wigner function simulations of electronic transport since previous simulations have been restricted to 1D drift/scattering and 1D potential. Any comparison to a real device must answer the question of how important both radial drift/scattering

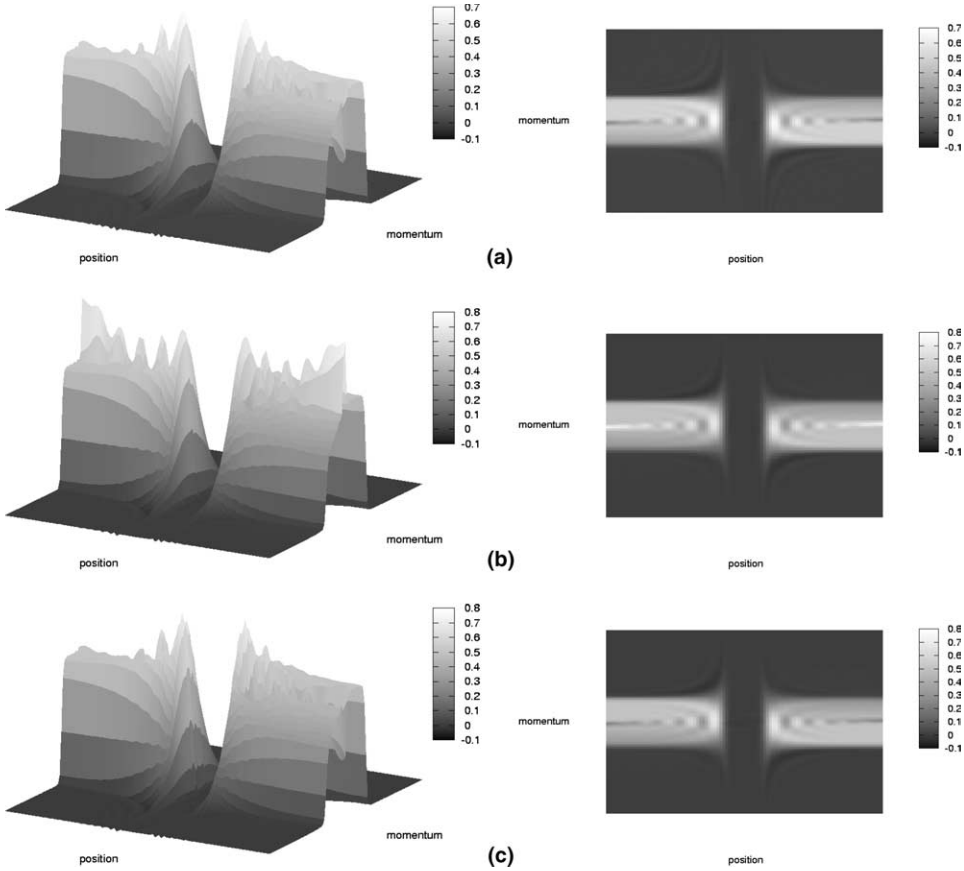

Fig. 14. Including coulomb interactions: surface and contour plots of the WDF at 1000 fs for the initial WDF of (a) zero, (b) pseudoFermi, (c) central.

and a non-1D coulombic charge density are. We now have a framework developed to examine these questions.

The exact effect of the radial drift/scattering will depend on the device parameters. Specifically, the WDF radial boundary condition at the outermost shell (Eqs. (46) and (47)) can be set up to describe a planar device (allow electrons to flow out) or a ''quantum mesa'' device (vacuum outside, therefore the electrons cannot exit). In the former case, we would simply see a decrease in the current density with radial distance (the magnitude of which is dependent on the device geometry). We are currently working on implementing a method where the radius of the ohmic contact at the emitter can be smaller than the device radius in order to perform a more realistic simulation. In the case of a quantum mesa, their should be carrier build up at the radial boundary thereby confining all radial transport, changing the quantum well into quantum dot [10]. It is this device that we ultimately hope to examine in detail.

The computational hurdles of solving a 2D WFE have been identified: (1) not all the matrices are sparse enough to fit into the RAM of present day computers, and (2) direct integration becomes more time consuming as the number of simultaneous equations to solve grows large. Some solutions of these hurdles have been described and a workable way of numerically simulating a RTS that exhibits cylindrical symmetry has been given. If one can separate the operator $\Omega = \mathbf{T}_z + \mathbf{U} + \mathbf{S}$ into the product of a drift/scattering operator

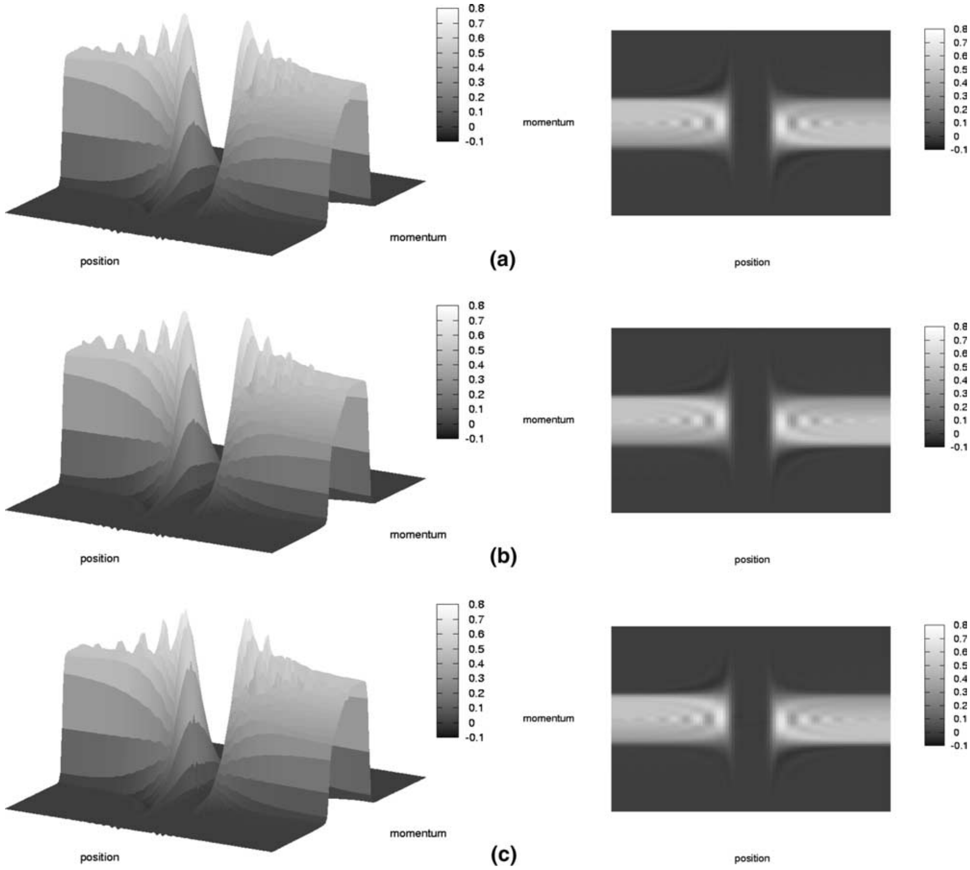

Fig. 15. Including coulomb interactions: surface and contour plots of the WDF at 2000 fs for the Initial WDF of (a) zero, (b) pseudoFermi, (c) central.

and a potential operator, then each matrix can be made sparse (overcoming the first hurdle) and/or block diagonal (the second hurdle). This also allows the separation of the transport problem into a two step problem, drift/scatter and potential computations. An additional time saving method, based on this split, is introduced. Finally, but taking advantage of the Fourier transform nature of the potential term, one can decrease the operator size substantially further.

Time evolution simulations based on these method were then presented for three different cases. Each case lead to numerical results consistent with expectations. To the author's knowledge, this is the first proof-of-principle of a 1D space, 2D momentum simulation. At the present time, a full transient treatment

Table 1
Timings for the matrix split runs

	Potential	$U \neq 0$		$U = 0$	
	# Time steps:	$200\Delta t$	$2000\Delta t$	$200\Delta t$	$2000\Delta t$
Initial	Zero	0 min 7.428 s	0 min 51.376 s	1 min 38.178 s	16 min 37.437 s
WDF	PseudoFermi	0 min 7.349 s	0 min 51.762 s	1 min 36.676 s	15 min 45.484 s
Value	Central	0 min 7.396 s	0 min 52.361 s	1 min 35.993 s	15 min 38.711 s

of a forward/backwards bias sweep would take upwards of 3 months. The work shown here still must be numerically scrutinized in order to shorten the computer times involved, our goal being the enhancement of the PDE solver. One important item to remember is that the computational hardware is still following Moore's law, that is, approximately doubling in speed every 18 months. In a few years, the techniques shown here, which push current technology to the limit, will prove to be even more feasible in the foreseeable future.

It is our belief that the methods outlined in this paper will finally allow a full treatment of the RTS transport problem.

Acknowledgments

The authors thank Peiji Zhao and Harold Grubin for many insightful discussions regarding RTS Wigner function simulations, C.T. Kelley for showing us different numerical techniques for solving the WFE, and M. Burcin Unlu for discussions on the 2D derivation. This work has been supported but the Defense University Research Initiative on Nanotechnology (DURINT) project grant from the Army Research Office.

References

[1] A. Abdulle, Fourth order Chebyshev methods with recurrence relation, SIAM J. Sci. Comput. 23 (6) (2002) 2042–2055.
[2] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen, LAPACK Users' Guide, third ed., Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999.
[3] F.A. Buot, K.L. Jensen, Lattice Weyl–Wigner formation of exact many-body quantum transport theory and applications to novel solid state quantum-based devices, Phys. Rev. B 42 (15) (1990) 9429.
[4] W.R. Frensley, Wigner-function model of a resonant-tunneling semiconductor device, Phys. Rev. B 36 (3) (1987) 1570.
[5] W.R. Frensley, Boundary conditions for open quantum systems driven far from equilibrium, Rev. Mod. Phys. 62 (3) (1990) 745.
[6] V.J. Goldman, D.C. Tsui, J.E. Cunningham, Observation of intrinsic bistability in resonant tunneling structures, Phys. Rev. Lett. 58 (1987) 1256.
[7] L.P. Kadanoff, G. Baym, Quantum Statistical Mechanics, Benjamin, New York, 1962.
[8] C.T. Kelley, Private Communication.
[9] H. Mizuta, T. Tanoue, The Physics and Applications of Resonant Tunnelling Diodes, Cambridge Studies in Semiconductor Physics and Microelectronic Engineering, Cambridge University Press, Cambridge, 1995.
[10] M.A. Reed, J.N. Randal, R.J. Aggarwak, R.J. Matyi, T.M. Moore, A.E. Wetsel, Observation of discrete electronic states in a zero-dimensional semiconductor nanostructure, Phys. Rev. Lett. 60 (6) (1988) 535.
[11] P. Zhao, H.L. Cui, D. Woolard, K.L. Jensen, F.A. Buot, Simulation of resonant tunneling structures: origin of the i–v hysteresis and plateau-like structure, J. Appl. Phys. 87 (3) (2000) 1337.