Simulation of Electron Transport in Quantum Well Devices

D. R. Miller, K. K. Gullapalli, V. R. Reddy, and D. P. Neikirk

Department of Electrical and Computer Engineering
Microelectronics Research Center
The University of Texas at Austin

1.0 Introduction

Double barrier resonant tunneling diodes (DBRTD) have received much attention as possible terahertz devices. Experimentally, DBRTD's have shown detection capabilities at sub-millimeter wavelengths\textsuperscript{1}. When used as oscillators, small amounts of power have also been measured in sub-millimeter range\textsuperscript{2}. Despite these impressive experimental results, the specific of the device physics (i.e., how the electrons propagate through the structure) are only qualitatively understood. Therefore, better transport models are warranted if this technology is to mature.

Near the heterostructure double barrier region, it is generally accepted that quantum mechanical transport, via tunneling and reflections, dominate the electron dynamics. However, most DBRTDs in use today are designed with extended spacer regions. These spacer regions serve the function of increasing the real part of the overall device impedance while simultaneously reducing the imaginary part, thereby incorporating DBRTDs in millimeter wave circuits a far easier task. Since the spacer regions are sufficiently removed from the heterojunctions, electron propagation should be govern by semiclassical and not quantum mechanical considerations. Here, semiclassical refers to transport which is adequately describe using some form of the semiclassical Boltzmann equation.
Past simulation models of DBRTD structures have evolved from simple Schrödinger equation solutions of a free electron in a double barrier potential to more complicated methods involving multiband, multivalley Schrödinger solutions\textsuperscript{3,4} or single valley kinetic equations that utilizes quantum Wigner functions\textsuperscript{5,6,7}. These methods are expected to work reasonably well for ideal DBRTD structures with parabolic bands in which only quantum mechanical reflections and tunneling are important. However, because the Schrödinger or single valley Wigner models do not include realistic phonon scattering or band structure effects (i.e., multiple valleys, non-parabolicity, multiple bands, etc.) these models do not adequately address the transport through the semiclassical region. Alternatively, the semiclassical Boltzmann equation provides an adequate description of the semiclassical region, but fails completely near the DBRTD heterostructure region.

To model a DBRTD structure with two distinct transport regions, two options are available. The first option incorporates a composite scheme by which each region is modeled with an equation suitable for that region. The two solutions are then matched at a quantum/classical interface to obtain a self consistent solution throughout the device. We find combinations of a free particle Schrödinger equation for the quantum region coupled with either the drift/diffusion\textsuperscript{8} or Monte Carlo\textsuperscript{9} formalism for the semiclassical region quite useful. However, agreement between experiment and theory is still lacking since the simple Schrödinger equation is only an approximate solution of the electron transport within the heterostructure region.

The second option is to model the entire device with one self consistent formalism that, in principle, can account for all the important device physics for each region. The kinetic equation based upon the Lattice Wigner function is a promising candidate for such a task. In this paper, we will use the Lattice Wigner function to explain important transport issues associated with DBRTD device behavior.

2.0 The Lattice Wigner Function

The lattice Wigner function\textsuperscript{10,11} we employ is based on the discrete spectrum composed of Wannier and Bloch crystal representations, making it different from other Wigner function methods. Because of the choice of representations, band structure effects are explicitly included in the kinetic equation. Thus, issues such as $\Gamma$ to $X$ tunneling, non-parabolicity of the conduction bands, or effective mass variations across the heterojunction,
can be examined in detail. Phonon scattering is also included through the standard semiclassical Boltzmann collision term.

In a multiple barrier heterostructure device with multiple non-parabolic conduction band valleys and no interband coupling, it is possible to write separate, but coupled, kinetic equations for each valley. The Wigner function for the \(i\)th valley, \(f^W_i\), is found from the solution of

\[
0 = \frac{\partial f^W_i(R,k)}{\partial t} + \frac{1}{\hbar} \frac{\partial e_n(k)}{\partial k} \frac{\partial f^W_i(R,k)}{\partial R} + \frac{eE}{\hbar} \frac{\partial f^W_i(R,k)}{\partial k} - \left( \frac{\partial f^W_i(R,k)}{\partial t} \right)_{\text{col}}
- \sum_{j=1}^{\#\text{Valleys}} \sum_{m=1}^{\#\text{Barriers}} B^{i,j,m}_0 + B^{i,j,m}_1 + B^{i,j,m}_2
\]

where the barrier scattering matrices, \(B_0\), \(B_1\), and \(B_2\) are given by

\[
B^{i,j,m}_0 = \frac{\Delta E^{(m,j)}(k)}{i\hbar} \sum_k f^W_j(R,k') \left( \frac{1}{N_L} \sum_{R'} e^{i(k' - k)} \left( V^{mn}_{\text{Pulse}(m)} \left( R + \frac{R'}{2} \right) - V^{nm}_{\text{Pulse}(m)} \left( R - \frac{R'}{2} \right) \right) \right)
\]

\[
B^{i,j,m}_1 = -\frac{\Delta v^i_\Delta}{2} \sum_k \frac{\partial f^W_j(R,k')}{\partial R} \left( \frac{1}{N_L} \sum_{R'} e^{i(k' - k)} \left( V^{mn}_{\text{Pulse}(m)} \left( R + \frac{R'}{2} \right) + V^{nm}_{\text{Pulse}(m)} \left( R - \frac{R'}{2} \right) \right) \right)
\]

\[
B^{i,j,m}_2 = -\frac{\hbar}{8im_\Delta^*} \sum_k \frac{\partial^2 f^W_j(R,k')}{\partial R^2} \left( \frac{1}{N_L} \sum_{R'} e^{i(k' - k)} \left( V^{mn}_{\text{Pulse}(m)} \left( R + \frac{R'}{2} \right) - V^{nm}_{\text{Pulse}(m)} \left( R - \frac{R'}{2} \right) \right) \right)
\]

The indices, \(j\) and \(q\), indicate the \(j\)th conduction band valley and the \(q\)th barrier. The coefficients, \(B^{i,j,q}\), are the potential energy terms that account for one heterostructure barrier. In equations 1 through 4, \(n\) is the band index, \(k\) and \(k'\) are crystal momenta, \(R\) and \(R'\) are lattice vectors, \(N_L\) is the number of lattice sites, and \(E\) is the electric field. \(V^{mn}_{\text{Pulse}}\) is the potential energy diagonal matrix element of the \(m\)th pulse function that localizes the barrier. The change in the \(i\)th valley effective mass \((m^{*i}_\Delta)\), velocity \((v^{*i}_\Delta)\), and the offset energy \((\Delta E^i)\) are given by

\[
\frac{1}{m^{*i}_\Delta} = \frac{1}{m^{*i}_{\text{Barrier}}} - \frac{1}{m^{*i}_{\text{Bulk}}}
\]

\[
v^{*i}_\Delta = \frac{1}{\hbar} \frac{\partial e^{\text{Barrier}}(k)}{\partial k} - \frac{1}{\hbar} \frac{\partial e^{\text{Bulk}}(k)}{\partial k}
\]

\[
\Delta E^i(k) = e^{\text{Barrier}}(k) - e^{\text{Bulk}}(k)
\]
where $\varepsilon_n^{\text{barrier}}(k_i)$ and $\varepsilon_n^{\text{bulk}}(k_i)$ are the minimum conduction band energies of the $i^{\text{th}}$ valley for the barrier and bulk materials, respectively.

As noted earlier, equation 1 accounts for both effective mass variations across the heterostructure interface and intervalley coupling. The intervalley coupling is possible since the $i^{\text{th}}$ valley distribution function $f_i^{W}$ is coupled to the $j^{\text{th}}$ valley distribution function $f_j^{W}$ through the barrier terms. Effective mass variation effects are due to the crystal momentum dependence of the barrier, and result in the spatial derivatives of the distribution function in the $B_1^q$ and $B_2^q$ terms. Mass variations are also responsible for raising and/or lowering the barrier height, as seen in the expression for $B_0^q$. This is because the energy band of the bulk material may rise at a different rate than the energy band of the barrier material for a given momentum change. For semiconductors with non-parabolic valleys, additional barrier scattering terms arise due to the higher derivatives associated with the Wigner-Moyal expansion. However, in this paper we will assume that the effective masses and the non-parabolicity factors are approximately the same for the bulk and barrier material. Therefore, $B_1$, $B_2$, and all other higher barrier scattering terms will be set to zero.

Including some sort of phonon collision processes into a Wigner function calculation of an DBRTD is not new. Most calculations have approximated the influence on carrier transport from phonon collisions using the relaxation time approximation\textsuperscript{7,12,13}. In addition, these calculations have restricted phonon scattering events to one valley. However, because of the high electric fields within the DBRTD structure, intervalley scattering is required. Thus, we will assume that the total scattering matrix, $S_{\text{Total}}$, is given by

$$S_{\text{Total}}^{ij}(k', k) = S_{\text{pop}}^{ij}(k', k) + S_{\text{Acoustic}}^{ij}(k', k) + S_{\text{Intervalley}}^{ij}(k', k)$$

(9)

where $S_{\text{pop}}$, $S_{\text{Acoustic}}$, and $S_{\text{Intervalley}}$ refer to the polar optical, acoustical, and intervalley phonon scattering matrices, respectively. Assuming non-degenerate statistics, we can write the collision term in equation 1 for the $i^{\text{th}}$ valley Wigner function as

$$\left( \frac{\partial f_i^{W}(R, k)}{\partial t} \right)_\text{Col} = \frac{1}{N_v} \sum_{j=1}^{N_v} \left\{ \sum_k^{\text{Valley}} \left( S_{\text{Total}}^{ij}(k, k') f_j^{W}(R, k') - S_{\text{Total}}^{ji}(k', k) f_i^{W}(R, k) \right) \right\}$$

(8)

where $N_v$ is the total number of valleys. The material parameters as well as the functional form for each scattering matrix are identical to those used in Monte Carlo device simulations and are given in reference 14.
The Wigner distribution function for each valley is defined in the three dimensional crystal momentum space. If the structure and the electric field are homogeneous in two of the spatial variables (independent of $R_y$ and $R_z$), equation 1 reduces to a four dimensional integral/differential equation. Since solutions to the four dimensional problems are exceedingly difficult, further simplification is required. We do this by characterizing the distribution function that is transverse to the electric fields by a Maxwellian defined at some transverse temperature $T_t$. Furthermore $T_t$ is assumed to vary with longitudinal position. Thus, for electric fields in the (100) direction with quantum well barriers grown in the (100) plane, the total distribution function is approximated by the product of a transverse and a longitudinal distribution functions

$$f^{W}_{i} (R_x, k_x, k_y, k_z) = f^{R}_{i} (R_x, k_x) \exp\left(-\frac{E^i_t}{k_b T_t (R_x)}\right)$$

where $k_x$, $k_y$, and $k_z$ are defined with respect to the valley minimum. The transverse energy for non-parabolic valleys, $E^i_t$, is assumed to be of the form

$$E^i_t = \frac{\left(\frac{\hbar^2 k_x^2}{2m_y^i} + \frac{\hbar^2 k_y^2}{2m_z^i}\right)}{1 + \sqrt{1 + 4\alpha_i \left(\frac{\hbar^2 k_x^2}{2m_y^i} + \frac{\hbar^2 k_y^2}{2m_z^i} + \frac{\hbar^2 k_z^2}{2m_x^i}\right)}}$$

where $\alpha_i$ is the non-parabolicity factor for the $i^{th}$ valley. Note that although the longitudinal and transverse distribution functions are not coupled through the electric field or the barrier scattering matrix, they are coupled by the phonon scattering matrix. It is because of this coupling that the transverse temperature significantly impacts the device physics, as discussed in the next section.

### 3.0 Simulation Results

The first obvious question that can be addressed with the lattice Wigner function is how close to the quantum well an electron is when quantum mechanical reflections and/or tunneling affect the electron's behavior. One measure of this distance is obtained by comparing the size of the barrier matrix ($B_0$) with the size of the phonon matrix ($S_{\text{Total}}$) at each lattice site. However, such a direct comparison may provide an overestimate of the extent of the quantum region. During one phonon scattering interval, an electron may
propagate through many lattice sites due to either a high electric field or a high initial velocity. Therefore, the effective barrier scattering strength is an average over all lattice sites occupied during the mean free flight time between collisions. In other words, the barriers' influence on the electrons is best determined by the matrix

\[ B_{\text{Total}} = \frac{1}{t_f - t_i} \sum_{j=1}^{\# \text{Barriers}} \int_{t_i}^{t_f} B_0^{j,m}(R_x(t), k_x(t)) \, dt \]  \hspace{1cm} (11)

integrated over the classical trajectory. In equation 11, \( B_0 \) is the matrix defined by equation 2, and \( R_x(t) \) and \( k_x(t) \) are given by the semiclassical equations of motion.

![Graph showing quantum well scattering strength versus position](image)

**Figure 1.** Effective quantum well scattering strength versus position from the center of a 6/19/6 GaAs/AlAs quantum well.

It is difficult to evaluate the above equation for the general case. However, Figure 1 shows the effective quantum well scattering strengths determined by taking the matrix norm of \( B_{\text{Total}} \) for two simplified conditions. The structure is a 6/19/6 GaAs/AlAs quantum well. The first solid line assumes the case of an extremely slow electron in which the electron transverses only one lattice site between phonon collisions. The second line assumes that the mean free path was 32 monolayers, corresponding to a distance of 90.6 Å. In both cases the electric field was set to zero, which significantly simplified the integration along the classical trajectory. Also plotted are the momentum relaxation times corresponding to different electric fields. These values represent the size of the phonon
scattering matrix \( S_{\text{Total}} \) determined using standard Monte Carlo simulations of homogeneous material\(^9\). Thus, when comparing the relative strengths of the quantum well and the phonon scattering matrices, one should compare the fast electron case with the high field momentum relaxation time since these two conditions usually exist together. Similarly, the slow electron case corresponds to the zero electric field relaxation time.

Two points are worth emphasizing from Figure 1. The first is that slow electrons are influenced by the quantum well with greater strength and at greater distances when compared to fast electrons. The second point, however, is the one which we want to emphasize. DBRTD structures are usually operated at extremely high electric fields (200 kV/cm) and under ballistic conditions. Therefore, as can be seen from the fast electron curve of Figure 1, the region where the barrier potential has the greatest impact on electron dynamics is within 65 monolayers of the center of the well. Beyond that region, phonon scattering is more important. This fact supports the hypothesis that much of the DBRTD is governed by semiclassical considerations for devices that extend 400 to 800 monolayers from the quantum well.

The Wigner function has been used extensively to account for the quantum transport within the heterostructure region of a DBRTD device. However, the question arises as to its ability to properly account for electron transport within the semiclassical region. To show that this is possible, we have used the Lattice Wigner Function to calculate the velocity versus electric field characteristics for homogeneous GaAs. For homogeneous samples, equation 1 reduces to

\[
0 = -\frac{eE}{\hbar} \cdot \frac{\partial f_j^W(R, k)}{\partial k} - \frac{1}{N_{\text{valley}}} \sum_{j=1}^{N_{\text{valley}}} \left\{ \sum_{k'} \left( S_{\text{Total}}^{ij}(k, k')f_j^W(R, k') - S_{\text{Total}}^{ji}(k', k)f_i^W(R, k) \right) \right\} \tag{12}
\]

To solve this equation, the first step is to integrate over the transverse crystal momentum directions. The resulting one dimensional phonon scattering matrix is then combined with the matrix generated by the electric field term to yield the equation

\[
0 = T_{R, j, k_x} \cdot f_{R, j, k_x} \tag{13}
\]

where \( T \) is the combined phonon/electric field matrix and \( f_{R, j, k_x} \) is the vectorized longitudinal distribution function. The valid solution exists when the minimum eigenvalue of \( T \) is equal to zero\(^{15}\). If the minimum eigenvalue is not zero, the electric field and the valley transverse temperatures, \( T_t \), are adjusted so that homogeneous equation can be
satisfied. In affect, we are using the transverse valley temperatures as fitting parameters to achieve a reasonable velocity versus field curve. Once the zero eigenvalue is calculated, the eigenvector for that eigenvalue can be determined, from which the velocity is obtained. Note that pure state tunneling models using only the Schrödinger equation do not produce reasonable velocity-field curves for GaAs.

The above procedure yields good agreement with the Monte Carlo results, as indicated in Figure 2. Here, a velocity field curve generated by equations 12 and 13 is compared against a three valley model calculated using standard Monte Carlo techniques. Furthermore, the temperatures required to obtain this level of agreement are comparable to the transverse temperatures calculated from the homogeneous Monte Carlo simulations. However, at low electric fields there is a larger discrepancy between the two alternative approaches. This difference is probably due to the fact that the transverse distribution function is not precisely Maxwellian. However, this error is fairly small, demonstrating that the Lattice Wigner function is capable of simulating both the quantum transport region and the classical transport regions of a device.

![Graph](image)

Figure 2. Velocity field calculation for a 3 valley spherical non-parabolic Wigner function model. Also shown is the corresponding result from a Monte Carlo simulation. The scattering parameters used for each calculation are identical, and are found in reference 14.

Thus far, we have applied the Lattice Wigner function to determine generic properties of bulk materials and of resonant tunneling transport. This formalism’s real
usefulness can be demonstrated by applying it to an actual DBRTD device. Shown in
Figure 3 is a typical GaAs/AlAs DBRTD. The quantum well consists of 6 monolayer
AlAs barriers separated by 19 monolayer GaAs well. As can be seen in Figure 3, a
moderately doped, extended spacer region is added to the right hand side of the quantum
well. As will be shown latter, it is the semiclassical transport through the extended spacer
layer that can have a serious impact on device behavior.

![Graph](image)

Figure 3. 6/19/6 GaAs/AlAs DBRTD used for the device simulations. The x axis origin is defined as the center of the quantum well.

The equilibrium electron concentration, found by integrating the Lattice Wigner
function over momentum space, is shown in Figure 4. However, under bias conditions,
obtaining the electron concentration via the Lattice Wigner function is much more
complicated. The basic problem is in determining the transverse temperature profile of the
spacer region. We have previously shown using a composite Schrödinger/Monte Carlo
model, that for typical bias voltages, electrons are quickly scattered into the upper satellite
values once they emerge from the quantum well. This intervalley scattering between
equivalent and non-equivalent conduction band valleys quickly heats up the carrier
distribution functions in the transverse direction. The effect of the carrier heating on the
charge densities can be considerable.
Figure 4. Equilibrium electron concentration for the device shown in Figure 3.

Figure 5. Electron concentrations obtained from the Lattice Wigner formalism for two different transverse temperature profiles. The dc bias is at 0.7 V.
Figure 5 illustrates the effect of transverse temperature on the simulated electron concentration for this device biased at 0.7 V. The dashed curve shown in this figure was generated under the assumption of constant transverse temperature within the spacer region. As seen from this curve, the electron density actually decreases past the quantum well region, indicating an increase in the overall electron velocities. The increase in the electron velocities is a result of a non-physical assumption. Because the transverse temperature in this space charge region is kept artificially low, the electron population is also artificially cooled, keeping all electrons in the fast $\Gamma$ valley. Therefore, unrealistically high velocities would be predicted under this constant temperature assumption. This can lead to unrealistic high frequency performance predictions, since fast electrons generally result in improved frequency response.

Also shown in Figure 5 are the results of a Lattice Wigner function calculation using a transverse temperature profile obtained through the Schrödinger/Monte Carlo model. As seen from this figure, the electron concentration is increased significantly over the constant temperature model, indicating a slowing down of the electron population within the spacer region. The reduced velocity is a result of the increased phonon scattering within the $\Gamma$ valley, as well as some $\Gamma$ to $L$ intervalley transfer. This slowing of the electrons can have an significant impact on the device performance, as discussed below.

![Figure 6](image.png)

Figure 6. Transverse temperature profile for the device given in Figure 3. The dc bias is at 2.0 volts. The temperature was obtained from a Schrödinger/Monte Carlo transport model discussed in reference 9.
The bias of 0.7 V is well below the voltage at which peak current is expected in this device. However, even under these conditions, the Monte Carlo simulated $\Gamma$ transverse temperature peaks at 2500 Kelvin. At even higher biases, the transverse temperature becomes extremely high. Figure 6 shows the transverse temperature profiles extracted through a three valley Schrödinger/Monte Carlo model for a device biased at 2.0 V. As seen from this figure, the $\Gamma$ temperature can exceed 7000 degrees Kelvin. Even the upper satellite valleys are exceptionally hot. Thus, we would expect that most of the electrons for this bias voltage are in the upper satellite valley, which is indeed the case.

4.0 Impact on High Frequency Behavior

Much of the discussion concerning the ultimate frequency limitations of DBRTD devices has centered on the frequency limitations imposed by the quantum well itself. It has been projected that the quantum well is capable of operating at terahertz frequencies\textsuperscript{16}. However, the preceding discussion illustrates that the semiclassical spacer regions can have a dramatic impact on the behavior of DBRTD devices. Because of the efficient phonon scattering processes, the actual number of free carriers within the spacer layer is much higher than predicted by either a pure state Schrödinger solution, a one valley Wigner solution, or a multiple valley constant temperature Lattice Wigner solution. This free charge contributes to a positive resistance which is equivalent in every way to the space charge resistance found in transit time diodes \textsuperscript{17}. The magnitude of the space charge resistance is fundamentally determined by the total number of free carriers within the region. Therefore, any analysis which does not realistically describe electron densities cannot be used to project the high frequency performance of a real device.

Understanding space charge resistance is important because it is generally felt that in order to improve the output power of a DBRTD device, one must dramatically increase the current density. In reality, this procedure will be useful only up the point where deleterious space charge resistance effects become dominant. This can be seen even under static conditions. For example, considered the measured dc-$IV$ curve of an InGaAs/AlAs DBRTD structure, shown in Figure 7. The total spacer region for this device was 1250Å. With available simulation tools, it is possible to model comparable structures with artificially increased current densities. Thus, five static simulations where performed using a composite quantum injector/drift/diffusion model, as described in reference 18. For each
simulation, the DBRTD spacer regions where kept constant. The only difference between each simulation is an assumed increase in the current density of the device, starting with the dc-$IV$ curve given in Figure 7. The results of the simulations are shown in Figure 8. As the current density is magnified, the presence of the free carriers within the spacer regions can cause a major portion of the negative resistance regime to become positive. It is obvious that this positive resistance would prevent useful device operation, despite the fact that the difference between the peak and valley current is exceptionally high.

![Current Density vs Voltage](image)

Figure 7. The experimentally measured dc-$IV$ curve for an InGaAs/AlAs DBRTD. The structure is similar to Figure 3 except that the total spacer layer to the right of the quantum well is 1250Å.

The current density where space charge resistance becomes important depends on two factors. The first is the length of the spacer regions. As the length is decreased, the total number of carriers within the region is also reduced. The price paid for this reduction is a corresponding decrease in the overall device impedance. Furthermore, it is doubtful that the spacer region can be eliminated beyond a certain point, since there is always a depletion region formed due to the high electric fields near the quantum well. This is even true if the heavily doped contacts are immediately adjacent to the quantum well.
Figure 8. The simulated effect of the space charge resistance on the dc-$IV$ curve. These curves were generated using the Schrödinger/Drift Diffusion composite model in which the current given by Figure 7 was scaled upward 10, 20, 30, 40, and 50 times.

The second factor dictating when the space charge resistance becomes important is determined by the velocity of the carriers. Higher velocities result in lower electron concentrations, since total current must be conserved. It is in predicting these concentrations and velocities that previous quantum mechanical models have failed, thus failing to predict the importance of the spacer regions in overall low and high frequency device behavior. Thus, in order to accurately project the ultimate performance of DBRTDs, it is critical to use a quantum kinetic formalism such as the Lattice Wigner function. Future results using this model should lead to a determination of the behavior of DBRTDs at terahertz frequencies.

This work has been supported by the Texas Advance Research Program.
5.0 References


