Ballistic to Diffusive Crossover in III–V Nanowire Transistors

M. J. Gilbert, Member, IEEE, and Sanjay K. Banerjee, Fellow, IEEE

Abstract—In this paper, we examine the crossover between ballistic and diffusive transport in III–V nanowire transistors. We find that at lower drain voltages the ballistic-to-diffusive crossover occurs at channel lengths of approximately 2.3 nm at room temperature. However, when we increase the drain voltage, we find that the ballistic-to-diffusive crossover can be more than nine times as long at room temperature. As the temperature is decreased, we find that the initially the performance is not significantly improved. However, as the temperature approaches 100 K, the ballistic-to-diffusive crossover increases to longer channel lengths quite dramatically. When InAs, InSb, and InP nanowires are compared at room temperature, we find that InAs and InSb perform in a similar fashion each with ballistic regions in excess of 10 nm, but that InP has no significant ballistic regime. Finally, we simulate several 10-nm InAs trigate transistors and show that for dopants deeply buried in the source and drain the devices appear ballistic, but when dopants appear near the source-channel interface, significant reductions in performance occur.

Index Terms—Nanowire transistors (NWTs), phonon interactions, quantum simulation, III–V.

I. INTRODUCTION

The nanowire transistor (NWT) [1] may be a viable device to replace bulk Si MOSFETs due to its ability to reduce short-channel effects and give greater control over the channel electron density. There has also been a resurgence in the exploration of III–V materials for next generation CMOS technology. While silicon is the unquestioned material of choice in CMOS technology, III–V materials have some advantages over silicon such as higher mobility, which could potentially lead to very high-speed low-power transistors. Of course, we recognize the tradeoff in terms of low density-of-states, i.e., Landauer channels. InSb has measured electron mobilities of 7.8 × 10⁴ cm² V⁻¹ s⁻¹ [2] and has long been touted as a good candidate for high-frequency devices. Recently, high-mobility wrap-gated p-channel InAs NWTs have been experimentally demonstrated which also have exceptionally high mobilities and solid device performance [3]. With the advent of technology capable of producing high quality III–V nanowires for transistor applications, it is natural to question the limits of device performance in III–V materials.

Of particular interest for future generation CMOS applications is the location of the ballistic-to-diffusive crossover in III–V nanowire systems.

There have been many previous studies demonstrating the performance of silicon NWT [4]–[6] and InAs NWT [7]. However, since we are interested in operating these devices at or near room temperature, phonon processes are very important. A recent study has placed the ballistic-to-diffusive crossover for silicon nanowire systems at 1.42 nm [8]. This places serious limits on the effectiveness of silicon in next generation CMOS, but it also shows the importance of phonon modes in the determination of the performance characteristics. In this paper, we examine the dependence of the ballistic-to-diffusive (mobility dominated transport) crossover point in III–V NWT as a function of temperature, drain voltage, and semiconductor material covering three of the most popular III–V materials: InAs, InSb, and InP using a fully 3-D self-consistent quantum transport formalism. The effects of phonon scattering are included in the simulations as separable self-energy terms in the Hamiltonian. To make the simulations as complete as possible, we include interactions with the dopant ions, and inelastic effects (the effects of acoustic deformation potential, intervalley phonons [Γ → X and Γ → L], and polar optical phonons [POPs]).

II. DEVICE STRUCTURE AND SIMULATION METHOD

In Fig. 1, we display a schematic of the system under consideration in the x–y plane of the z-axis is normal to the 69 plane shown). The device dimensions (in multiples of the lattice constant) use in the simulations were chosen to aid in the consideration of the discrete dopants. The thickness of the InAs layer is 9.09 nm and the width of the channel is 8.48 nm. The source and drain of the device are n-type with a doping density of 1.3 × 10¹⁸ cm⁻³, while the channel of the device is considered to be p-type, but undoped. The gate material is assumed to be platinum and the gate oxide on each side is 1 nm of hafnium oxide (HfO₂). Underneath the device, we 78
have assumed a thick silicon dioxide substrate. The length of
the device is set to 50.89 nm. While the channel of the device
is typically much smaller than the overall length of the device
geometry, a large domain is set so that as the length of the
channel is increased, the bandstructure for the overall device
remains constant. The length of the channel starts at a length
of 6.0583 Å and is increased by this length until the diffusive
regime is obtained. During this increase of the length, the width,
and depth of the nanowire channel are kept constant. The length
of the gate that surrounds the nanowire on three sides is set to
be the same length as the channel of the device. The source and
drain of the device are divided equally depending on the length
of the channel; however, they never are reduced below 10 nm
in the x-direction.

Once the device geometry is defined, the InAs lattice is
scanned and the dopants are distributed according to the method
presented in [9]. Following the distribution of the dopants, they
are then mapped back onto the grid of the simulation mesh and
the initial self-consistent Poisson solution is obtained. In this
case, the full Coulomb potential of the dopants is incorporated.
Then, the solution of Poisson’s equation for the local potential
is no longer smoothly varying in the source and drain of the
device. The inclusion of discrete dopants causes the formation
of potential variations in the source and drain. The density
throughout the device is calculated using a variant of the re-
cursive scattering matrix method which solves the Schrödinger
equation in the effective mass approximation, as described in
[4]. In short, this method provides us with a 3-D, fully quantum
mechanical treatment. Since the transport calculation is per-
formed in real space, the different excited modes in the system
are automatically coupled unlike in other simulations [6]. In
order to achieve self-consistency, the density obtained from the
transport calculation is then updated using Broyden’s method
[10] and a new guess for the potential is obtained through
the solution of Poisson’s equation. The process is repeated
until a desired level of convergence is obtained. Exchange and
correlations terms are included in each simulation through a

In this paper, we seek the ballistic-to-diffusive crossover in
these nanowire systems. Our procedure for determining this
quantity is given below. We know that even when the transport
is ballistic, the lower limit of the resistance of the channel is
determined by the inverse of the Landauer conductance

\[ R_{\text{ballistic}} = \frac{2e^2}{\hbar N} \]  

where \( N \) is the number of transverse modes flowing in the
quantum wire. Clearly, the resistance in (1) has no dependence
on the length of the quantum wire. Conversely, when the
transport is determined by the carrier mobility and the carrier
density, then the resistance becomes

\[ R_{\text{diffusive}} = \frac{1}{n e \mu} \frac{L}{A} \]  

where \( L \) is the channel length and \( A \) is the cross-sectional
area of the inversion layer. We have been careful to use the
area of the inversion layer rather than the width because we are
dealing with a 3-D quantum wire with quantization in the transverse directions. In this regime, there is a clear linear de-
pendence of the resistance with the length of the quantum wire.

\[ R_{\text{diffusive}} = \frac{1}{n e \mu} \frac{L}{A} \]  

III. POP SCATTERING

The inclusion of separable scattering mechanisms is not a
new concept. In previous work, acoustic and optical phonon
processes have been derived [8] and so the work will not be
repeated here. However, we are now interested in including the
nonlocal effects of POPs which, to this point, have not been
treated. This is a crucial difference between silicon and III–V
materials. In silicon, the local phonon processes dominate, but
in III–V materials nonlocal optical processes dominate the
room temperature transport. We begin by noting that to use the
method presented in [8], we are assuming that the scattering
is weak relative to the energies in the system, and that we are
only interested in the steady-state limit. Therefore, we can use
the Fermi golden rule expression, equivalent to a first-
order nonself-consistent Born approximation, for each of the
scattering processes and generate a real space self-energy from
g it. In particular, we note that the imaginary part of the self-
energy term is related to the scattering rate and it is the nearest
scattering rate that we wish to calculate [8]. This will result in
an x-directed momentum which is related to the carrier energy
in the quantum wire. While the local scattering rates have been
previously derived [8], we will present the important changes
in the general derivation approach to take into account the
nonlocal nature of POP scattering. In this paper, we use the
Fermi golden rule to calculate scattering rates. This is treated in
many other places, however, we must account for the transverse
modes of the quantum wire. Therefore, we begin with the general form

\[ \left( \frac{1}{\tau_{\text{polar}}(\text{ij})} \right)_{\text{ij}}' = \frac{\pi e^2}{\gamma V \omega_0} \sum_{q_x} \sum_{q_y,q_z} \int dy \int dz \int dy' \int dz' \times \varphi_{i,j}(y,z) \varphi_{i',j}(y',z') \varphi_{i,j}(y',z') \times \frac{e^{i q_y (y-y')} + i q_z (z-z')}{q^2} \times \{ (N_q + 1) \delta(E_k - E_{k-q} - i \omega_0) + N_q \delta(E_k - E_{k-q} + i \omega_0) \} \]

where \( V \) is the volume, \( \omega_0 \) is the phonon frequency and \( \gamma \) is
the effective interaction parameter, \( i \) is the initial mode index
163 in the width direction, \( j \) is the initial mode index in the depth
164 direction, \( i' \) is the final mode index in the width direction, 165
and \( j' \) is the final mode index in depth direction. The delta
166 functions in (3) serve to conserve the energy in the process
167 of the interactions of the carriers with the POPs. \( \varphi_{i,j}(y,z) \) is
168 the transverse 2-D wave function in the particular slice under
169 consideration. We assume in this treatment that we are dealing
170 with parabolic bandstructure. Taking a closer look at the terms
171
172 contained within the delta functions, we realize that based on
173 this assumption
\[ E_k - E_{k+q} = \hbar \omega_0 \]
\[ = E_{i,j} - E_{i',j'} \mp \hbar \omega_0 + \frac{\hbar^2 k_x^2}{2m_x^*} - \frac{\hbar^2 (k_x + q_x)^2}{2m_x^*} \]
\[ = E_{i,j} - E_{i',j'} \pm \hbar \omega_0 - \frac{\hbar^2 k_x q_x}{m_x^*} \pm \frac{\hbar^2 q_x^2}{2m_x^*}. \] (4)

174 In (4), \( E_{i,j} \) and \( E_{i',j'} \) are the initial and final energies at the
175 bottom of their respective bands. We begin to simplify (3), by
176 first examining the longitudinal terms
\[ \sum \frac{1}{2\pi^2} \delta(E_k - E_{k+q} \pm \hbar \omega_0). \] (5)
We now expand (5) using the results from (4)
\[ \sum_{q_x} \frac{1}{q_x^2 + q_y^2 + q_z^2} \left( \frac{\hbar^2 q_x^2}{2m_x^*} + \frac{\hbar^2 k_x q_x}{m_x^*} + E_{i,j} - E_{i',j'} \pm \hbar \omega_0 \right). \] (6)
This sum may be simplified further by taking the Fourier
179 transform with respect to \( q_x \)
\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} dq_x e^{iq_x(x-x')} \left( \frac{\hbar^2 q_x^2}{2m_x^*} + \frac{\hbar^2 k_x q_x}{m_x^*} + \Delta_{i,j} \right) \] (7)
where we denote
\[ \Delta_{i,j} = E_{i,j} - E_{i',j'} \pm \hbar \omega_0. \] (8)
Representing (7) as a summation, we have
\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} dq_x e^{iq_x(x-x')} \frac{1}{q_x^2 + q_y^2 + q_z^2} \frac{m_x^*}{2\pi \hbar^2} \int_{-\infty}^{\infty} dq_x e^{iq_x(x-x')} \left( \frac{\hbar^2 k_x q_x}{m_x^*} + \frac{\hbar^2 q_x^2}{2m_x^*} \right). \] (9)
At this point, we examine the relationship for \( E_{k+q} \)
\[ E_{k+q} = \frac{\hbar^2 k_x^2}{2m_x^*} + \frac{\hbar^2 k_x q_x}{m_x^*} + \frac{\hbar^2 q_x^2}{2m_x^*}. \] (10)
Equation (11) is most easily evaluated by making a transformation
to cylindrical coordinates to obtain
\[ \frac{m_x^*}{4\pi \hbar^2} \int dq_x e^{iq_x(r-r')} \frac{\rho_{1D}(E') dE'}{q_x^2 + a^2}. \] (12)
We evaluate the integral in (12) using contour integrations to
arrive at a final, simplified expression for the sum over the
longitudinal wavevectors
\[ \sum_{q_x} \frac{1}{2q_x^2} \delta(E_k). \] (13)
The expression in (13) is substituted back into (3) to find an
overlap integral that is similar to both the one found in [8] and
in most textbooks
\[ = \sum_{q_x} \sum_{q_y \neq q_x} \int dy \int dz \int dy' \int dz' \phi_{i,j}(y, z) \phi_{i',j'}(y, z) \]
\[ \times \phi_{i,j}(y', z') e^{iq_y(y-y')} + i q_z(z - z') \delta(E_k) \]
\[ = A \frac{e}{2V} \sum_{q_x} \int dz dy \left[ \phi_{i,j}^*(y, z) \phi_{i',j'}(y, z) \right] \frac{1}{q_x} \delta(E_k) \] (14)
Where in (14), we have used the following definition
\[ I_{i,j}^{\prime,\prime} = \frac{A}{4\pi e} \int dz dy \left[ \phi_{i,j}^*(y, z) \phi_{i',j'}(y, z) \right]. \] (15)
Following the usual procedure, we replace the summation over
the final momentum states with an integration
\[ \sum_{k'} \to \frac{L}{2\pi} \int_{-\infty}^{\infty} dk' = \int_0^{\rho_{1D}(E')} dE' \] (16)
to obtain the final result
\[ \left( \frac{1}{\tau_{polar}} \right)_{i,j}^{\prime,\prime} = \frac{m_x^* L e^2}{32\pi^3 \rho_{1D}} I_{i,j}^{\prime,\prime} \int_{-\infty}^{\infty} \frac{dk_x}{k_x} \delta(E_k). \] (17)
However, we are still in the mode space representation and
the longitudinal momentum is not yet a viable operator. From
many-body physics, the momentum dependence arises from the
Fourier transform of the differences in the two coordinates in
the self-energy. Therefore, to include the proper momentum
dependence, we take the inverse transform. This essentially
involves solving a contour integral for both the emission and
absorption cases which results from (17). This gives us our
final, simplified and summarized results for absorption and
emission of POPs
\[ \left( \frac{1}{\tau_{polar}} \right)_{i,j}^{\prime,\prime} = \frac{m_x^* L e^2}{32\pi^3 \rho_{1D}} I_{i,j}^{\prime,\prime} \int_{-\infty}^{\infty} \frac{dk_x}{k_x} \delta(E_k). \] (18)
This result is reasonable as we now see the manifestation of the nonlocal nature of POP scattering as a clear dependence on distance in (18). To utilize this in the transport calculation, we must use the following unitary transformation to convert this form to the site representation for inclusion in the Hamiltonian

$$\Gamma_{\text{polar}} = \text{Im}\{\Sigma\} = U^+ \left( \frac{\hbar}{\tau_{\text{polar}}} \right)^{ij'} U$$

where $U$ is a unitary mode-to-site transformation matrix. The unitary matrix $U^+$ results from the eigenvalue solutions in the transverse slice and are composed of the various eigenfunctions in the site basis. Hence, it represents a mode-to-slice transformation.

IV. RESULTS

A. Drain-Voltage Dependence

In Fig. 2, we plot the resistance of an InAs quantum wire as a function of the length of the channel. The gate voltage has been set to $V_g = 0.5$ V and the drain voltage has been set to $V_d = 10$ mV. The temperature of the device is 300 K. The red lines signify the maximum and minimum ranges for the ballistic to diffusive crossover, and the blue line signifies the median value. These lines are used to determine the value quoted for the ballistic to diffusive crossover.

![Fig. 2. Resistance of an InAs quantum wire as a function of the length of the channel. The gate voltage has been set to $V_g = 0.5$ V and the drain voltage has been set to $V_d = 10$ mV. The temperature of the device is 300 K. The red lines signify the maximum and minimum ranges for the ballistic to diffusive crossover, and the blue line signifies the median value. These lines are used to determine the value quoted for the ballistic to diffusive crossover.](Image)

In Fig. 3, we now increase the drain voltage to $V_d = 0.5$ V. The gate voltage has been set to $V_g = 0.5$ V and the drain voltage has been set to $V_d = 0.5$ V. The temperature of the device is 300 K. The red lines once again signify the maximum and minimum ranges for the ballistic to diffusive crossover, and the blue line signifies the median value.

![Fig. 3. Resistance of an InAs gated quantum wire as a function of the length of the channel. The gate voltage has been set to $V_g = 0.5$ V and the drain voltage has been set to $V_d = 0.5$ V. The temperature of the device is 300 K. The red lines once again signify the maximum and minimum ranges for the ballistic to diffusive crossover, and the blue line signifies the median value.](Image)

The existence of the tunneling regime in Fig. 3 is an important point. If indeed, the physics is as we state then the 278 same tunneling regime should show up in simple ballistic 279 simulations of the same device structure. In Fig. 4, we plot 280 the resistance of the quantum wire channel as a function of its 281 length with no scattering included in the simulation. We find a 282 similar trend as in the case with the inelastic processes included. 283 In this simulation, we initially begin with three propagating 284 modes in the system. The initial resistances do not match the 285
other quantum wires as doping is slightly lower than in other devices. After a short period of nearly constant and quantized resistance, we see a similar rise in the resistance of the quantum wire as in the case where phonon interactions were included. Since no phonon modes are available to promote or demote the electron energies, we conclude that the tunneling regime is due to large energetic spacing of longitudinal states in the channel.

Therefore, while there is an initial ballistic region at high drain voltages, it is interrupted by a tunneling regime. After the tunneling regime, we see a dramatic reduction to the proper ballistic quantum resistance for several propagating modes. This ballistic region then persists, in the case of this specific device, for 12.33 ± 1.41 nm, after subtracting off the length of the tunneling region, before it begins to increase linearly as expected. This is almost 10 times the length found in a similar silicon NWT at room temperature.

In Fig. 5, we plot the ballistic-to-diffusive crossover as a function of the drain voltage applied to the device. There is an interesting trend present in this figure. We find that the ballistic-to-diffusive crossover seems to show a nonlinear increase with increasing drain voltage. This is quite different from the linearly reducing dependence on drain voltage found in silicon. To explain how this is possible, we reexamine the forms of the POP emission and absorption terms shown in (18). For low-energy carriers, the scattering rate is appreciable over a range of grid points in the longitudinal direction. This leads to large scattering rates which significantly affects the incident carriers. However, as the energy of the carriers is increased, the phonon scattering rates become local. This localization causes a significant reduction in the strength of the electron–phonon interaction contribution to the Hamiltonian. Therefore, we expect to see small ballistic-to-diffusive lengths at low drain energy but which increase as the carriers gain energy from the applied field. Perhaps the more unexpected part of this result is not that we obtain longer ballistic-to-diffusive lengths, but that the connection between the quasi-1-D electrons in the channel and the reservoirs produces a shift in the location of the ballistic regime. This is an important feature to note for future ultrasmall device fabrication as it imposes serious limits on the scalability of NWTs.

Put in more physical terms, we should expect to find the ballistic length to roughly follow

$$L_{\text{ballistic}} = \frac{e V t^2}{2 m L^2}.$$  (20)

Where in (20), $V$ is the applied voltage, $t$ is the mean scattering time, $L$ is the length of the channel, and $m$ is the effective mass.

Hence, by increasing the drain bias from 10 to 500 mV, we should find a corresponding increase of the ballistic length of 50 times. The fact that we do not see such a large increase is due to the fact that the mean scattering time is reduced significantly by hot carriers for the higher drain bias. This results in the value of about 12 nm for the ballistic length which is only an increase of 3.3 times compared to that of the low-voltage result. The reduction in the expected ballistic length corresponds to a decrease in the mean scattering time from 18 fs to about 8.5 fs. While these times are very short, this is indicative of the fact that the transport is not being conducted in a bulk sample. Using the bulk value of the mobility of InAs, we find that the mean scattering time is around 180 fs (assuming a mobility of 14 000 cm$^2$/V·s). If the carriers reach the saturation velocity, and drop the entire bias energy into longitudinal optical (LO) phonons, then the estimated mean scattering time would only be about 3.6 fs. Thus, the carriers are dropping a great deal of energy in the drain rather than in the channel of the device and our scattering time estimates seem to be in order.

### B. Temperature Dependence

In Fig. 6, we once again plot the resistance of an InAs quantum wire device as a function of the channel length. As expected, as the device is cooled, the interactions with the phonon modes in the device should be reduced. We see that the initial 12 nm of the device exhibits the same type of behavior as in the higher temperature cases, however, now the tunneling regime now takes on a more pronounced appearance. This is due to the fact that there is no longer enough energy to broaden the levels in the channel and the energetic distribution of...
Fig. 6. Resistance of an InAs quantum wire as a function of the length of the channel. The gate voltage has been set to $V_g = 0.5$ V and the drain voltage has been set to $V_d = 0.5$ V. The temperature of the device is 100 K.

Fig. 7. Summary of the ballistic to diffusive crossover in InAs NWTs for temperatures at a constant voltages of $V_g = V_d = 0.5$ V. We see a decreasing nonlinear trend, within error bars, forming in the ballistic to diffusive crossover length. There are no error bars on the 100 K case as it is an estimate.

Fig. 8. Plot of the resistance of an InSb NWT as a function of the channel length with $V_g = V_d = 0.5$ V. We show the maximum and median lines used to estimate the ballistic to diffusive crossover.

C. Material Dependence

As mentioned in the introduction, InAs is not the only III–V semiconductor currently under investigation for future use in CMOS architectures; InSb and InP are also being considered. In Fig. 8, we plot the resistance of an InSb NWT as a function of channel length with the gate and drain voltages affixed at $0.5$ V. The thickness of the InSb film is $9.07$ nm and the width of the channel is $8.42$ nm. The source and drain are $n$-type with $391$ doping set to $1.7 \times 10^{18}$ cm$^{-3}$ while the channel is undoped $p$-type. While the dimensions have again been chosen to facilitate the inclusion of the discrete doping in the source and the drain of the device, we have chosen very similar dimensions to those of the InAs devices to facilitate comparison. The overall width of $396$ of the source and drain are $27.21$ nm and the overall length of $397$ of the device is $50.54$ nm. In the InSb device, we see very similar trends as to those seen in the InAs devices. This is to be expected as their POP energies are quite similar. Nevertheless, while the overall shape of the curve remains similar, there are some notable differences. The location of the tunneling regime has shifted its position in channel length. This shift is attributed to the readjustment of the states in the channel due to the changes in effective mass between InAs and InSb. Therefore, the initial ballistic region associated with direct tunneling is estimated to be $4.73 \pm 0.82$ nm. After the tunneling region, we find a smooth ballistic region which extends $8.7 \pm 0.60$ nm, after subtracting off the tunneling region length. While the overall ballistic length in InSb and InAs NWTs are quite similar, the InAs device has a longer continuous region of ballistic transport. Based on these results, it seems that InAs and InSb are rather interchangeable.

Both InAs and InSb are starkly contrasted by the behavior of InP. In Fig. 9, we plot the resistance of an InP NWT as a function of the channel length for $V_g = V_d = 0.5$ V at $300$ K. The dimensions of the device are slightly different due to the change in lattice constant. Here, we use a overall source and drain width of $27.58$ nm with a total device length of $50.47$ nm. The InP film thickness is $9.39$ nm with a corresponding channel width of $8.22$ nm. The channel is assumed to be undoped $p$-type and the source and drain are doped $n$-type $2 \times 10^{18}$ cm$^{-3}$. In InP, we find that there is no significant ballistic transport regime and even very short NWTs show diffusive transport.
be explained in two ways. First, the energy of the POPs in InP is much larger than in InAs or InSb. When this is combined with the larger effective mass of InP, the POP scattering rate is quite large and ballistic transport cannot be obtained. More importantly is the fact that InP demonstrates that the channel of these NWT devices are strongly coupled to the source and the drain. This affects transport in that the electrons are not injected directly into the channel of the device. Before they reach the channel, they must first traverse the source. Therefore, by the time the electrons have reached the channel of the device, they have already undergone diffuse transport and this has affected the mean free path in InP. Almost certainly, it is possible for InP to exhibit ballistic characteristics in other situations, but not in the context of the device presented here.

D. Effect of Discrete Doping

In Fig. 10(a), we plot the $I_d-V_g$ curve for three discretely doped trigate InAs quantum wire devices with only elastic (boundary and impurity) scattering considered (solid) and with phonon processes (dotted) while in Fig. 10(b), we repeat the plots again on a logarithmic scale. For device simulation, we have increased the source and drain doping to $6 \times 10^{18}$ cm$^{-3}$ and set the channel length to 9.69 nm. The device temperature is 300 K and the drain voltage is held at $V_d = 0.6$ V. In terms of simple device performance, averaging over four different devices, we find that the threshold voltage is found to be $0.373$ V ± $11$ mV. These figures are slightly different from those quoted elsewhere [9], but it should be noted that we have used a smaller sampling of devices in this paper. While in Fig. 10, the majority of the dopant atoms were located near the bottom of the InAs layer allowing the device to recover quasi-ballistic behavior, one device shows a very distinct difference between the ballistic and quasi-ballistic cases. Here, the locations of the dopant atoms are such that in the source of the device there are two dopant atoms near the entrance to the channel in the middle of the semiconductor layer. These dopant atoms cause significant modifications to the energy of the incident electrons. The lower energy electrons now see more POP scattering which, when combined with the typical reflections from the channel entrance, gives rise to significant reductions in the amount of charge in the channel. With less conduction in the channel, the device sees significantly degraded performance.

In Fig. 11, we plot the electron density taken at a depth of approximately 5 nm into the InAs device layer at $V_g = 0.6$ V for the elastic case. The black dots in the figure represent the locations of the dopant atoms in the system. Dots that are larger in size are closer to the surface of the device, while dots that are smaller are buried farther down in the device. At this voltage, we are above the threshold voltage for the device; the channel is now heavily populated with carriers. In the source and drain, the location of the electron density depends rather weakly on the locations of the dopants. This is mainly due to the scarcity of the dopants and that they are buried deeply in the InAs substrate which reduces their effect on the propagating electrons. Nevertheless, the density tends to roughly follow the path of the dopants as it makes its way to the channel. In the channel, we find that the some of the electrons have been trapped by multiple sequential reflections off of the...
The great difference in device performance, with a ballistic performance greatly due to weak phonon interactions. InSb devices have similar device performance, with a ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly with the placement of dopants in the source and channel interface. Further, we find that reductions in the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian.

In this paper, we have presented a method for including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the nonlocal effects of POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly with the voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in temperature will not improve the ballistic-to-diffusive crossover length, but cooling the device to around 100 K improves the performance greatly due to weak phonon interactions. InSb devices have similar device performance, with a ballistic-to-diffusive crossover of about 8.7 nm. The great difference between InAs and InSb is the locations of longitudinal states in the channel which shifts a highly resistive region responsible for poor device performance in both materials. This is not the case with InP where we find that the POP scattering is too strong and diffusive transport is achieved in the source of the device. We simulated several 10-nm InAs devices to find that at high gate voltages they behave as ballistic devices, but for the placement of dopants in the source and drain is crucial to device performance. These results demonstrate the potential and limitations of III–V NWTs for use in highly scaled CMOS architectures. With ballistic lengths nine times longer than silicon, they are certainly worthy of continued exploration for novel electron quantum interference and waveguide devices.

ACKNOWLEDGMENT

M. J. Gilbert, one of the authors, would like to thank D. K Ferry and L. F. Register for insightful conversations.

REFERENCES


M. J. Gilbert (M’06) received the B.S. (with honors), M.S., and Ph.D. degrees in electrical engineering, all from Arizona State University, Tempe, in 2000, 2003, and 2005, respectively.

He is the Assistant Director of the South West Academy of Nanoelectronics at the University of Texas at Austin. He has contributed to research on nanostructures, many-body theory, and quantum transport. He has authored more than 30 refereed publications and has given presentations at over 30 international conferences. His current research centers on spintronics, strongly correlated electron systems, and quantum transport theory.

Sanjay K. Banerjee (S’80–M’83–SM’89–F’96) received the B.Tech. degree from the Indian Institute of Technology (IIT), Kharagpur, India, in 1979, and the M.S. and Ph.D. degrees from the University of Illinois at Urbana–Champaign, in 1981 and 1983, respectively, all in electrical engineering.

He is the Cockrell Family Regents Chair Professor of Electrical and Computer Engineering and Director of the Microelectronics Research Center at the University of Texas at Austin. As a Member of Technical Staff, Corporate Research, Development 265 and Engineering of Texas Instruments, Inc., from 1983–1987, he worked on polysilicon transistors and dynamic random access memory cells 627 used by Texas Instruments in the world’s first 4-Mb dynamic random access memory. He has been Assistant Professor (1987–1990), Associate Professor 629 (1990–1993), and Professor (since 1993) with The University of Texas at Austin. He is currently active in the areas of ultrahigh vacuum and remote 631 plasma-enhanced chemical vapor deposition for silicon–germanium–carbon 632 heterostructure MOSFETs and nanostructures. He is also interested in the areas of ultrashallow junction technology and semiconductor device modeling. He has over 580 archival refereed publications/talks, seven books/chapters, and 635 26 U.S. patents. He has supervised over 40 Ph.D. and 50 M.S. students.

Ballistic to Diffusive Crossover in III–V Nanowire Transistors

M. J. Gilbert, Member, IEEE, and Sanjay K. Banerjee, Fellow, IEEE

Abstract—In this paper, we examine the crossover between ballistic and diffusive transport in III–V nanowire transistors. We find that at lower drain voltages the ballistic-to-diffusive crossover occurs at channel lengths of approximately 2.3 nm at room temperature. However, when we increase the drain voltage, we find that the ballistic-to-diffusive crossover can be more than nine times as long at room temperature. As the temperature is decreased, we find that the initially the performance is not significantly improved. However, as the temperature approaches 100 K, the ballistic-to-diffusive crossover increases to longer channel lengths quite dramatically. When InAs, InSb, and InP nanowires are compared at room temperature, we find that InAs and InSb perform in a similar fashion each with ballistic regions in excess of 10 nm, but that InP has no significant ballistic regime. Finally, we simulate several 10-nm InAs trigate transistors and show that for dopants deeply buried in the source and drain the devices appear ballistic, but when dopants appear near the source-channel interface, significant reductions in performance occur.

Index Terms—Nanowire transistors (NWTs), phonon interactions, quantum simulation, III–V.

I. INTRODUCTION

THE NANOWIRE transistor (NWT) [1] may be a viable device to replace bulk Si MOSFETs due to its ability to reduce short-channel effects and give greater control over the channel electron density. There has also been a resurgence in the exploration of III–V materials for next generation CMOS technology. While silicon is the unquestioned material of choice in CMOS technology, III–V materials have some advantages over silicon such as higher mobility, which could potentially lead to very high-speed low-power transistors. Of course, we recognize the tradeoff in terms of low density-of-states, i.e., Landauer channels. InSb has measured electron mobilities of 7.8 × 10^4 cm^2 V^−1 s^−1 [2] and has long been touted as a good candidate for high-frequency devices. Recently, high-mobility wrap-gated p-channel InAs NWts have been experimentally demonstrated which also have exceptionally high mobilities and solid device performance [3]. With the advent of technology capable of producing high quality III–V nanowires for transistor applications, it is natural to question the limits of device performance in III–V materials.

Of particular interest for future generation CMOS applications is the location of the ballistic-to-diffusive crossover in III–V nanowire systems.

There have been many previous studies demonstrating the performance of silicon NWT [4]–[6] and InAs NWT [7]. However, since we are interested in operating these devices at 49 or near room temperature, phonon processes are very important. A recent study has placed the ballistic-to-diffusive crossover for silicon nanowire systems at 1.42 nm [8]. This places serious limits on the effectiveness of silicon in next generation CMOS, but it also shows the importance of phonon modes in the determination of the performance characteristics. In this paper, we examine the dependence of the ballistic-to-diffusive (mobility dominated transport) crossover point in III–V NWT as a function of temperature, drain voltage, and semiconductor material covering three of the most popular III–V materials: InAs, InSb, and InP using a fully 3-D self-consistent quantum transport formalism. The effects of phonon scattering are included in the simulations as separable self-energy terms in the Hamiltonian. To make the simulations as complete as possible, we include interactions with the dopant ions, and inelastic effects (the 64 effects of acoustic deformation potential, intervalley phonons [Γ → X and Γ → L], and polar optical phonons (POPs)).

II. DEVICE STRUCTURE AND SIMULATION METHOD

In Fig. 1, we display a schematic of the system under 68 consideration in the x–y plane (the z-axis is normal to the 69 plane shown). The device dimensions (in multiples of the lattice constant) used in the simulations were chosen to aid in the 71 inclusion of the discrete dopants. The thickness of the InAs 72 layer is 9.09 nm and the width of the channel is 8.48 nm. The source and drain of the device are n-type with a doping 74 density of 1.3 × 10^18 cm^−3, while the channel of the device 75 is considered to be p-type, but undoped. The gate material is assumed to be platinum and the gate oxide on each side is 77 1 nm of hafnium oxide (HfO2). Underneath the device, we 78
have assumed a thick silicon dioxide substrate. The length of
the device is set to 50.89 nm. While the channel of the device
is typically much smaller than the overall length of the device
geometry, a large domain is set so that as the length of the
channel is increased, the bandstructure for the overall device
remains constant. The length of the channel starts at a length
of 6.0583 Å and is increased by this length until the diffusive
regime is obtained. During this increase of the length, the width,
and depth of the nanowire channel are kept constant. The length
of the gate that surrounds the nanowire on three sides is set to
be the same length as the channel of the device. The source and
drain of the device are divided equally depending on the length
of the channel; however, they never are reduced below 10 nm
in the x-direction.

Once the device geometry is defined, the InAs lattice is
scanned and the dopants are distributed according to the method
presented in [9]. Following the distribution of the dopants, they
are then mapped back onto the grid of the simulation mesh and
the initial self-consistent Poisson solution is obtained. In this
case, the full Coulomb potential of the dopants is incorporated.

Then, the solution of Poisson’s equation for the local potential
is no longer smoothly varying in the source and drain of the
device. The inclusion of discrete dopants causes the formation
of potential variations in the source and drain. The density
thoughout the device is calculated using a variant of the re-
cursive scattering matrix method which solves the Schrödinger
equation in the effective mass approximation, as described in
short, this method provides us with a 3-D, fully quantum
mechanical treatment. Since the transport calculation is per-
formed in real space, the different excited modes in the system
are automatically coupled unlike in other simulations [6]. In
order to achieve self-consistency, the density obtained from the
transport calculation is then updated using Broyden’s method
[10] and a new guess for the potential is obtained through
the solution of Poisson’s equation. The process is repeated
until a desired level of convergence is obtained. Exchange and
correlations terms are included in each simulation through a

In this paper, we seek the ballistic-to-diffusive crossover in
these nanowire systems. Our procedure for determining this
quantity is given below. We know that even when the transport
is ballistic, the lower limit of the resistance of the channel is
determined by the inverse of the Landauer conductance

\[ R_{\text{ballistic}} = \left[ \frac{2e^2}{hN} \right]^{-1} \]  

where \( N \) is the number of transverse modes flowing in the
quantum wire. Clearly, the resistance in (1) has no dependence
on the length of the quantum wire. Conversely, when the
transport is determined by the carrier mobility and the carrier
density, then the resistance becomes

\[ R_{\text{diffusive}} = \frac{1}{ne\mu} \frac{L}{A} \]  

where \( L \) is the channel length and \( A \) is the cross-sectional
area of the inversion layer. We have been careful to use the
area of the inversion layer rather than the width because we are
dealing with a 3-D quantum wire with quantization in the transverse directions. In this regime, there is a clear linear de-
pendence of the resistance with the length of the quantum wire.

III. POP SCATTERING

The inclusion of separable scattering mechanisms is not a new concept. In previous work, acoustic and optical phonon processes have been derived [8] and so the work will not be repeated here. However, we are now interested in including the nonlocal effects of POPs which, to this point, have not been treated. This is a crucial difference between silicon and III–V materials. In silicon, the local phonon processes dominate, but in III–V materials nonlocal optical processes dominate the room temperature transport. We begin by noting that to use the method presented in [8], we are assuming that the scattering is weak relative to the energies in the system, and that we are only interested in the steady-state limit. Therefore, we can use the Fermi golden rule expression, equivalent to a first-
order nonself-consistent Born approximation, for each of the scattering processes and generate a real space self-energy from it. In particular, we note that the imaginary part of the self- energy term is related to the scattering rate and it is the latter scattering rate that we wish to calculate [8]. This will result in an x-directed momentum which is related to the carrier energy in the quantum wire. While the local scattering rates have been previously derived [8], we will present the important changes in the general derivation approach to take into account the nature of POP scattering. In this paper, we use the Fermi golden rule to calculate scattering rates. This is treated in many other places, however, we must account for the transverse modes of the quantum wire. Therefore, we begin with the general form

\[ \frac{1}{\tau_{\text{polar}}} = \sum_{q} \sum_{i,j} \int dy \int dz \int dy' \int dz' \times \phi_{i,j}(y,z)\delta(E_k - E_{k-q} - \hbar\omega_0) \times \frac{\{N_q + 1\} \delta(E_k - E_{k-q} + \hbar\omega_0)}{\pi e^2 \gamma V \omega_0} \]  

where \( V \) is the volume, \( \omega_0 \) is the phonon frequency and \( \gamma \) is the effective interaction parameter, \( i \) is the initial mode index in the width direction, \( j \) is the initial mode index in the depth direction, \( i' \) is the final mode index in the width direction, \( j' \) is the final mode index in the depth direction. The delta functions in (3) serve to conserve the energy in the process of the interactions of the carriers with the POPs. \( \phi_{i,j}(y,z) \) is the transverse 2-D wave function in the particular slice under consideration. We assume in this treatment that we are dealing with parabolic bandstructure. Taking a closer look at the terms
172 contained within the delta functions, we realize that based on 173 this assumption

\[ E_k - E_{k\pm q} \pm \hbar \omega_0 \]

\[ = E_{i,j} - E_{i',j'} \pm \hbar \omega_0 + \frac{\hbar^2 k_x^2}{2m_x^*} - \frac{\hbar^2(k_x \pm q_x)^2}{2m_x^*} \]

\[ = E_{i,j} - E_{i',j'} \pm \hbar \omega_0 - \frac{\hbar^2 k_x q_x}{m_x^*} \pm \frac{\hbar^2 q_x^2}{2m_x^*.} \]  \hspace{1cm} (4)

174 In (4), \( E_{i,j} \) and \( E_{i',j'} \) are the initial and final energies at the 175 bottom of their respective bands. We begin to simplify (3), by 176 first examining the longitudinal terms

\[ \sum_{q_x} \frac{1}{q_x^2} \delta(E_k - E_{k\pm q} \pm \hbar \omega_0). \]  \hspace{1cm} (5)

177 We now expand (5) using the results from (4)

\[ \sum_{q_x} \frac{1}{q_x^2 + q_y^2 + q_z^2} \left( \frac{\hbar^2 q_x^2}{2m_x^*} + \frac{\hbar^2 k_x q_x}{m_x^*} + E_{i,j} - E_{i',j'} \pm \hbar \omega_0 \right). \]  \hspace{1cm} (6)

178 This sum may be simplified further by taking the Fourier 179 transform with respect to \( q_x \)

\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} dq_x e^{iq_x(x-x')} \frac{1}{q_x^2 + q_y^2 + q_z^2} \left( \frac{\hbar^2 q_x^2}{2m_x^*} + \frac{\hbar^2 k_x q_x}{m_x^*} + \Delta_{i,j}^{i',j'} \right) \]  \hspace{1cm} (7)

180 where we denote

\[ \Delta_{i,j}^{i',j'} = E_{i,j} - E_{i',j'} \pm \hbar \omega_0. \]  \hspace{1cm} (8)

181 Representing (7) as a summation, we have

\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} dq_x e^{iq_x(x-x')} \frac{1}{q_x^2 + q_y^2 + q_z^2} \left[ \delta(E_{k\pm q} - \hbar \omega_0) \right] \]

\[ = \frac{m_x^*}{2\pi\hbar^2} \int_{-\infty}^{\infty} dq_x e^{iq_x(x-x')} \frac{1}{q_x^2 + q_y^2 + q_z^2} \left[ \frac{\hbar^2 k_x q_x}{m_x^*} \right]. \]  \hspace{1cm} (9)

182 At this point, we examine the relationship for \( E_{k\pm q} \)

\[ E_{k\pm q} = \frac{\hbar^2 k_x^2}{2m_x^*} \pm \frac{\hbar^2 k_x q_x}{m_x^*} \pm \frac{\hbar^2 q_x^2}{2m_x^*}. \]  \hspace{1cm} (10)

183 Equation (10) is quadratic in \( q_x \) and may then be solved to yield 184 a solution for \( q_x \), which is then substituted back into (9) making 185 the term in the absolute value independent of \( q_x \) as shown 186 below regardless of if the phonon process under consideration 187 is emission or absorption. Therefore, we may group all of 188 the variables that are independent of \( q_x \) into a simple dummy 189 variable \( a \). With this pairing, we may reduce our summation in 200 (9) to reach

\[ \frac{m_x^*}{2\pi\hbar^2} \int_{-\infty}^{\infty} dq_x e^{iq_x(x-x')} \frac{1}{q_x^2 + a^2}. \]  \hspace{1cm} (11)

Equation (11) is most easily evaluated by making a transformation 202 to cylindrical coordinates to obtain

\[ \frac{m_x^*}{4\pi\hbar^2} \int dq_x e^{iq_x(r-r')} \frac{1}{q_x^2 + a^2} \]  \hspace{1cm} (12)

We evaluate the integral in (12) using contour integrations to 203 arrive at a final, simplified expression for the sum over the 204 longitudinal wavevectors

\[ \sum_{q_x} \frac{1}{2q_x} \delta(E_{k}). \]  \hspace{1cm} (13)

The expression in (13) is substituted back into (3) to find an 206 overlap integral that is similar to both the one found in [8] and 207 in most textbooks

\[ = \sum_{q_x} \sum_{q_y,q_z} \int dq_y \int dq_z \int dq_0 \int dq_1 \sum_{i,j} \delta(E_{k}). \]

\[ \times \varphi_{i,j}(y, z) \varphi_{i',j'}(y, z) e^{iq_y(y-y')} + iq_z(z-z') \frac{\delta(E_{k})}{2q_x}. \]

\[ = \frac{A}{2V} \sum_{q_x} \int dq x [\varphi_{i,j}(y, z) \varphi_{i',j'}(y, z)]^2 \frac{1}{q_x} \delta(E_{k}). \]  \hspace{1cm} (14)

Where in (14), we have used the following definition

\[ \varphi_{i,j}^{i',j'} = \frac{A}{4\pi V} \int dq x [\varphi_{i,j}^*(y, z) \varphi_{i',j'}(y, z)]^2. \]  \hspace{1cm} (15)

Following the usual procedure, we replace the summation over 200 the final momentum states with an integration

\[ \sum_{k'} \rightarrow \frac{L}{2\pi} \int_{-\infty}^{\infty} dk' = \frac{L}{2\pi} \int_{0}^{\infty} \rho_{1D}(E') dE' \]  \hspace{1cm} (16)

to obtain the final result

\[ \left( \frac{1}{t_{polar}} \right)_{i,j}^{i',j'} = \frac{m_x^*Lc^2}{32\pi^3\gamma \omega_0} \int_{-\infty}^{\infty} \frac{dk_x}{k_x} \delta(E_{k}). \]  \hspace{1cm} (17)

However, we are still in the mode space representation and 203 the longitudinal momentum is not yet a viable operator. From 204 many-body physics, the momentum dependence arises from the 205 Fourier transform of the differences in the two coordinates in 206 the self-energy. Therefore, to include the proper momentum 207 dependence, we take the inverse transform. This essentially 208 involves solving a contour integral for both the emission and 209 absorption cases which results from (17). This gives us our 210 final, simplified and summarized results for absorption and 211 emission of POPs

\[ \left( \frac{1}{t_{polar}} \right)_{i,j}^{i',j'} = \frac{m_x^*Lc^2}{32\pi^3\gamma \omega_0} \int_{-\infty}^{\infty} \frac{dk_x}{k_x} \delta(E_{k}) \left\{ \frac{N_q e^{-k_0|x-x'|}}{(N_q + 1) \sin(k_0|x-x'|)} \right\}. \]  \hspace{1cm} (18)
This result is reasonable as we now see the manifestation of the nonlocal nature of POP scattering as a clear dependence on distance in (18). To utilize this in the transport calculation, we must use the following unitary transformation to convert this form to the site representation for inclusion in the Hamiltonian

$$\Gamma_{\text{polar}} = \text{Im}\{\Sigma\} = U^+ \left( \frac{\hbar}{\tau_{\text{polar}}} \right)_{i,j}^\prime U$$

where $U$ is a unitary mode-to-site transformation matrix. The unitary matrix $U^+$ results from the eigenvalue solutions in the transverse slice and are composed of the various eigenfunctions in the site basis. Hence, it represents a mode-to-slice transformation.

IV. RESULTS

A. Drain-Voltage Dependence

In Fig. 2, we plot the resistance of an InAs quantum wire as a function of the length of the channel. The gate voltage has been set to $V_g = 0.5$ V and the drain voltage has been set to $V_d = 10$ mV. The temperature of the device is 300 K. The red lines signify the maximum and minimum ranges for the ballistic to diffusive crossover, and the blue line signifies the median value. These lines are used to determine the value quoted for the ballistic to diffusive crossover.

![Fig. 2. Resistance of an InAs quantum wire as a function of the length of the channel. The gate voltage has been set to $V_g = 0.5$ V and the drain voltage has been set to $V_d = 10$ mV. The temperature of the device is 300 K. The red lines signify the maximum and minimum ranges for the ballistic to diffusive crossover, and the blue line signifies the median value. These lines are used to determine the value quoted for the ballistic to diffusive crossover.](image_url)

In Fig. 3, we now increase the drain voltage to $V_d = 0.5$ V. We find that this drain voltage consists of much more structure than the device held at reduced drain voltages. We begin our examination of Fig. 3, by noticing that there is an initial period of slightly more than 3 nm where the resistance of the quantum wire channel does not change. In this region, the length of the channel is such that the incident electrons see almost no tunneling potential and simply tunnel from the source to the drain of the device. As the length of the channel is increased beyond 4 nm, we see that there is a large spike in the resistance of the channel. This behavior is attributed to the fact that the locations of the longitudinal states in the channel are too high in energy to support any modes and the channel becomes a large tunnel barrier. This argument is supported by the fact that the form of the drain voltage versus channel length in the regime of 4 to 8 nm has an exponential form which is indicative of tunneling transport. We do not see this in the lower drain voltage sweeps due to the fact that scattering has blurred energy spectrum of the electrons before they reach the channel. This widening of the range of energies allows electrons not energetically capable of accessing the states to scatter into them. Thus, we see no rise in resistance in Fig. 2. While this type of interaction is fairly uncommon in semiconductor devices due to their relatively large sizes, this is common in systems of quasi-1-D electrons connected to external reservoirs [13].

The existence of the tunneling regime in Fig. 3 is an important point. If indeed, the physics is as we state then the same tunneling regime should show up in simple ballistic simulations of the same device structure. In Fig. 4, we plot the resistance of the quantum wire channel as a function of its length with no scattering included in the simulation. We find a similar trend as in the case with the inelastic processes included. In this simulation, we initially begin with three propagating modes in the system. The initial resistances do not match the...
other quantum wires as doping is slightly lower than in other devices. After a short period of nearly constant and quantized resistance, we see a similar rise in the resistance of the quantum wire as in the case where phonon interactions were included. Since no phonon modes are available to promote or demote the electron energies, we conclude that the tunneling regime is due to large energetic spacing of longitudinal states in the channel.

Therefore, while there is an initial ballistic region at high drain voltages, it is interrupted by a tunneling regime. After the tunneling regime, we see a dramatic reduction to the proper ballistic quantum resistance for several propagating modes. This ballistic region then persists, in the case of this specific device, for 12.33 ± 1.41 nm, after subtracting off the length of the tunneling region, before it begins to increase linearly as expected. This is almost 10 times the length found in a similar silicon NWT at room temperature.

In Fig. 5, we plot the ballistic-to-diffusive crossover as a function of the drain voltage applied to the device. There is an interesting trend present in this figure. We find that the ballistic-to-diffusive crossover seems to show a nonlinear increase with increasing drain voltage. This is quite different from the linearly reducing dependence on drain voltage found in silicon. To explain how this is possible, we reexamine the forms of the POP emission and absorption terms shown in (18).

For low-energy carriers, the scattering rate is appreciable over a range of grid points in the longitudinal direction. This leads to large scattering rates which significantly affects the incident carriers. However, as the energy of the carriers is increased, the phonon scattering rates become local. This localization causes a significant reduction in the strength of the electron–phonon interaction contribution to the Hamiltonian. Therefore, we expect to see small ballistic-to-diffusive lengths at low drain energy but which increase as the carriers gain energy from the applied field. Perhaps the more unexpected part of this result is not that we obtain longer ballistic-to-diffusive lengths, but that the connection between the quasi-1-D electrons in the channel and the reservoirs produces a shift in the location of the ballistic regime. This is an important feature to note for future ultrasmall device fabrication as it imposes serious limits on the scalability of NWTs.

Put in more physical terms, we should expect to find the ballistic length to roughly follow

$$L_{\text{ballistic}} = \frac{eVt^2}{2mL}.$$  \hspace{1cm} (20)$$

Where in (20), $V$ is the applied voltage, $t$ is the mean scattering time, $L$ is the length of the channel, and $m$ is the effective mass.

Hence, by increasing the drain bias from 10 to 500 mV, we should find a corresponding increase of the ballistic length of 50 times. The fact that we do not see such a large increase is due to the fact that the mean scattering time is reduced significantly by hot carriers for the higher drain bias. This results in the 335 value of about 12 nm for the ballistic length which is only an increase of 3.3 times compared to that of the low-voltage result.

The reduction in the expected ballistic length corresponds to a decrease in the mean scattering time from 18 fs to about 8.5 fs. While these times are very short, this is indicative of the fact that the transport is not being conducted in a bulk sample.

Using the bulk value of the mobility of InAs, we find that the 340 mean scattering time is around 180 fs (assuming a mobility of 1 4 000 cm^2/V·s). If the carriers reach the saturation velocity, 344 and drop the entire bias energy into longitudinal optical (LO) phonons, then the estimated mean scattering time would only be about 3.6 fs. Thus, the carriers are dropping a great deal of energy in the drain rather than in the channel of the device and our scattering time estimates seem to be in order.

### B. Temperature Dependence

In Fig. 6, we once again plot the resistance of an InAs quantum wire device as a function of the channel length. As expected, as the device is cooled, the interactions with the phonon modes in the device should be reduced. We see that the initial 12 nm of the device exhibits the same type of behavior as in the higher temperature cases, however, now the tunneling regime now takes on a more pronounced appearance. This is due to the fact that there is no longer enough energy to broaden the levels in the channel and the energetic distribution of...
In Fig. 7, we summarize the performance of InAs NWTs at several different temperatures. We find that the trend in the plot shows a quickly decaying dependence. Clearly, the mobility increases as the temperature is reduced. However, even at reduced temperatures the electron–phonon interactions are crucial to understanding the performance limits of III–V NWTs. It should be noted that these simulations do not account for the effect of carrier heating near the drain of the device. With a relatively large drain voltage of $V_d = 0.5 \text{ V}$ applied to the NWT, carrier heating will undoubtedly play a nontrivial role in determining the principles of thermal operation.

C. Material Dependence

As mentioned in the introduction, InAs is not the only III–V semiconductor currently under investigation for future use in CMOS architectures; InSb and InP are also being considered. In Fig. 8, we plot the resistance of an InSb NWT as a function of channel length with the gate and drain voltages affixed at $V_g = V_d = 0.5 \text{ V}$. The thickness of the InSb film is 9.07 nm and the width of the channel is 8.42 nm. The source and drain are $n$-type with 391 doping set to $1.7 \times 10^{18} \text{ cm}^{-3}$ while the channel is undoped $p$-type. While the dimensions have again been chosen to facilitate the inclusion of the discrete doping in the source and the drain of the device, we have chosen very similar dimensions to the InAs devices to facilitate comparison. The overall width of 396 of the source and drain are 27.21 nm and the overall length of 397 of the device is 50.54 nm. In the InSb device, we see very similar trends as to those seen in the InAs devices. This is to be expected as their POP energies are quite similar. Nevertheless, while the overall shape of the curve remains similar, there are some notable differences. The location of the tunneling regime has shifted its position in channel length. This shift is attributed to the readjustment of the states in the channel due to changes in effective mass between InAs and InSb. Therefore, the initial ballistic region associated with direct tunneling is estimated to be 4.73 ± 0.82 nm. After the tunneling region, we find a smooth ballistic region which extends 8.7 ± 0.60 nm, after subtracting off the tunneling region length. While the 409 ballistic length in InSb and InAs NWTs are quite similar, the InAs device has a longer continuous region of ballistic transport. Based on these results, it seems that InAs and InSb are rather interchangeable.

Both InAs and InSb are starkly contrasted by the behavior of InP. In Fig. 9, we plot the resistance of an InP NWT as a function of channel length with $V_g = V_d = 0.5 \text{ V}$ at 300 K. The dimensions of the device are slightly different due to the change in lattice constant. Here, we use a overall source and drain width of 27.58 nm with a total device length of 50.47 nm. The InP film thickness is 9.39 nm with a corresponding channel width of 8.22 nm. The channel is assumed to be undoped $p$-type and the source and drain are doped $n$-type $2 \times 10^{18} \text{ cm}^{-3}$. In 422 InP, we find that there is no significant ballistic transport regime and even very short NWTs show diffusive transport. This can be attributed to the much faster scattering associated with POP phonons increases at 100 K, POP scattering still dominates the transport. When 423 at 100 K, POP scattering still dominates the transport. When 424 at several different temperatures. We find that the trend in the plot shows a quickly decaying dependence. Clearly, the mobility increases as the temperature is reduced. However, even at reduced temperatures the electron–phonon interactions are crucial to understanding the performance limits of III–V NWTs. It should be noted that these simulations do not account for the effect of carrier heating near the drain of the device. With a relatively large drain voltage of $V_d = 0.5 \text{ V}$ applied to the NWT, carrier heating will undoubtedly play a nontrivial role in determining the principles of thermal operation.

6 IEEE TRANSACTIONS ON ELECTRON DEVICES

![Fig. 6. Resistance of an InAs quantum wire as a function of the length of the channel. The gate voltage has been set to $V_g = 0.5 \text{ V}$ and the drain voltage has been set to $V_d = 0.5 \text{ V}$ The temperature of the device is 100 K.

![Fig. 7. Summary of the ballistic to diffusive crossover in InAs NWTs for temperatures at a constant voltages of $V_g = V_d = 0.5 \text{ V}$. We see a decreasing nonlinear trend, within error bars, forming in the ballistic to diffusive crossover length. There are no error bars on the 100 K case as it is an estimate.

![Fig. 8. Plot of the resistance of an InSb NWT as a function of the channel length with $V_g = V_d = 0.5 \text{ V}$. The temperature of the device is 100 K.](image-url)
be explained in two ways. First, the energy of the POPs in InP is much larger than in InAs or InSb. When this is combined with the larger effective mass of InP, the POP scattering rate is quite large and ballistic transport cannot be obtained. More importantly is the fact that InP demonstrates that the channel of these NWT devices are strongly coupled to the source and the drain. This affects transport in that the electrons are not injected directly into the channel of the device. Before they reach the channel, they must first traverse the source. Therefore, by the time the electrons have reached the channel of the device, they have already undergone diffusive transport and this has affected the mean free path in InP. Almost certainly, it is possible for InP to exhibit ballistic characteristics in other situations, but not in the context of the device presented here.

D. Effect of Discrete Doping

In Fig. 10(a), we plot the $I_d-V_g$ curve for three discretely doped trigate InAs quantum wire devices with only elastic (boundary and impurity) scattering considered (solid) and with phonon processes (dotted) while in Fig. 10(b), we repeat the plots again on a logarithmic scale. For device simulation, we have increased the source and drain doping to $6 \times 10^{18}$ cm$^{-3}$ and set the channel length to 9.69 nm. The device temperature is 300 K and the drain voltage is held at $V_d = 0.6$ V. In terms of simple device performance, averaging over four different devices, we find that the threshold voltage is found to be 0.373 V ± 11 mV. These figures are slightly different from those quoted elsewhere [9], but it should be noted that we have used a smaller sampling of devices in this paper. While the threshold voltage is, at present, too high for next generation technology, with the use of gate stack engineering [7] this figure can be brought down to acceptable levels. Further, we see that the spread in the threshold voltage is quite small. This is due to the fact that fewer dopant atoms are required to reach these levels of performance. Thus, with fewer elastic scattering sites to induce quantum interference, the threshold voltage is expected to be more stable than in its silicon-based equivalents.

Upon further examination of the curves in Fig. 10(b), we find that in the low gate voltage regime POP scattering significantly reduces the current in all three cases. This shift in low gate voltage current leads to modifications to the performance. The threshold voltage averaged over four devices is 0.423 V ± 465 mV. Based on this result, we find that the POP scattering does indeed shift the threshold voltage, but it also reduces $I_d$ nonuniformities by reducing spread in carrier energy in the source. At high gate voltages, we find that the drain currents in both the elastic and combined (elastic scattering contribution + inelastic scattering contribution) cases begin to merge after $V_g = 0.5$ V. This can be most directly attributed to a high-energy localization of the POP scattering rate which limits the effect of the inelastic perturbations to current. The end result is that the elastic interactions dominate the current–voltage characteristics.

While in Fig. 10, the majority of the dopant atoms were located near the bottom of the InAs layer allowing the device to recover quasi-ballistic behavior, one device shows a very distinct difference between the ballistic and quasi-ballistic cases. Here, the locations of the dopant atoms are such that in the source of the device there are two dopant atoms near the entrance to the channel in the middle of the semiconductor layer. These dopant atoms cause significant modifications to the energy of the incident electrons. The lower energy electrons now see more POP scattering which, when combined with the typical reflections from the channel entrance, gives rise to significant reductions in the amount of charge in the channel. With less conduction in the channel, the device sees significantly degraded performance.

In Fig. 11, we plot the electron density taken at a depth of approximately 5 nm into the InAs device layer at $V_g = 0.6$ V for the elastic case. The black dots in the figure represent the locations of the dopant atoms in the system. Dots that are larger in size are closer to the surface of the device, while dots that are smaller are buried farther down in the device. At this voltage, we are above the threshold voltage for the device; the channel is now heavily populated with carriers. In the source and drain, the location of the electron density depends rather weakly on the locations of the dopants. This is mainly due to the scarcity of the dopants and that they are buried deeply in the InAs substrate which reduces their effect on the propagating electrons. Nevertheless, the density tends to roughly follow the path of the dopants as it makes its way to the channel. In the channel, we find that the some of the electrons have been trapped by multiple sequential reflections off of the boundary and impurity scattering considered (solid) and with phonon processes (dotted) cases begin to merge after $V_g = 0.5$ V. This can be most directly attributed to a high-energy localization of the POP scattering rate which limits the effect of the inelastic perturbations to current. The end result is that the elastic interactions dominate the current–voltage characteristics.

In Fig. 10(a), we plot the $I_d-V_g$ curve for three discretely doped trigate InAs quantum wire devices with only elastic (boundary and impurity) scattering considered (solid) and with phonon processes (dotted) while in Fig. 10(b), we repeat the plots again on a logarithmic scale. For device simulation, we have increased the source and drain doping to $6 \times 10^{18}$ cm$^{-3}$ and set the channel length to 9.69 nm. The device temperature is 300 K and the drain voltage is held at $V_d = 0.6$ V. In terms of simple device performance, averaging over four different devices, we find that the threshold voltage is found to be 0.373 V ± 11 mV. These figures are slightly different from those quoted elsewhere [9], but it should be noted that we have used a smaller sampling of devices in this paper. While the threshold voltage is, at present, too high for next generation technology, with the use of gate stack engineering [7] this figure can be brought down to acceptable levels. Further, we see that the spread in the threshold voltage is quite small. This is due to the fact that fewer dopant atoms are required to reach these levels of performance. Thus, with fewer elastic scattering sites to induce quantum interference, the threshold voltage is expected to be more stable than in its silicon-based equivalents. Upon further examination of the curves in Fig. 10(b), we find that in the low gate voltage regime POP scattering significantly reduces the current in all three cases. This shift in low gate voltage current leads to modifications to the performance.
In this paper, we have presented a method for including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering. In InAs devices, we find that NWTs have a significant ballistic length at room temperature of about 12.33 nm. This ballistic length increases nonlinearly as the drain voltage increases as the nonlocal POPs become localized at higher energy. Further, we find that reductions in this method to simulate III–V NWTs while including the nonlocal effects of POP scattering into simulation techniques relying on a propagating Hamiltonian. We then used this method to simulate III–V NWTs while including the effects of not only POP scattering, but also impurity, intervalley, and acoustic phonon scattering.


M. J. Gilbert (M’06) received the B.S. (with honors), M.S., and Ph.D. degrees in electrical engineering, all from Arizona State University, Tempe, in 2000, 2003, and 2005, respectively.

He is the Assistant Director of the South West Academy of Nanoelectronics at the University of Texas at Austin. He has contributed to research on nanostuctures, many-body theory, and quantum transport. He has authored more than 30 refereed publications and has given presentations at over 30 international conferences. His current research centers on spintronics, strongly correlated electron systems, and quantum transport theory.

Sanjay K. Banerjee (S’80–M’83–SM’89–F’96) received the B.Tech. degree from the Indian Institute of Technology, Kharagpur, India, in 1979, and the M.S. and Ph.D. degrees from the University of Illinois at Urbana–Champaign, in 1981 and 1983, respectively, all in electrical engineering.

He is the Cockrell Family Regents Chair Professor of Electrical and Computer Engineering and Director of the Microelectronics Research Center at the University of Texas at Austin. As a Member of Technical Staff, Corporate Research, Development 625 and Engineering of Texas Instruments, Inc., from 1983–1987, he worked on polysilicon transistors and dynamic random access trenched memory cells used by Texas Instruments in the world’s first 4-Mb dynamic random access 628 memory. He has been Assistant Professor (1987–1990), Associate Professor 629 (1990–1993), and Professor (since 1993) with The University of Texas at 630 Austin. He is currently active in the areas of ultrahigh vacuum and remote 631 plasma-enhanced chemical vapor deposition for silicon–germanium–carbon heterostructure MOSFETs and nanostructures. He is also interested in the areas of ultrashallow junction technology and semiconductor device modeling. He has over 580 archival refereed publications/talks, seven books/chapters, and 635 26 U.S. patents. He has supervised over 40 Ph.D. and 50 M.S. students.


467