Observation of strong reflection of electron waves exiting a ballistic channel at low energy

Canute I. Vaz, Changze Liu, Jason P. Campbell, Jason T. Ryan, Richard G. Southwick III, David Gundlach, Anthony S. Oates, Ru Huang, and Kin. P. Cheung

Citation: AIP Advances **6**, 065212 (2016); doi: 10.1063/1.4954083
View online: http://dx.doi.org/10.1063/1.4954083
View Table of Contents: http://scitation.aip.org/content/aip/journal/adva/6/6?ver=pdfcov
Published by the AIP Publishing

Observation of strong reflection of electron waves exiting a ballistic channel at low energy

Canute I. Vaz,[1] Changze Liu,[1,2] Jason P. Campbell,[1] Jason T. Ryan,[1] Richard G. Southwick III,[1,3] David Gundlach,[1] Anthony S. Oates,[4] Ru Huang,[2] and Kin. P. Cheung[1,a]

[1]*National Institute of Standards and Technology, Gaithersburg, MD 20899-8120, USA*
[2]*Institute of Microelectronics, Peking University, Beijing 100871, China*
[3]*IBM Research, Albany, NY 12205, USA*
[4]*Taiwan Semiconductor Manufacturing Corporation, Hsinchu 30844, Taiwan*

(Received 21 October 2015; accepted 5 June 2016; published online 10 June 2016)

Wave scattering by a potential step is a ubiquitous concept. Thus, it is surprising that theoretical treatments of ballistic transport in nanoscale devices, from quantum point contacts to ballistic transistors, assume no reflection even when the potential step is encountered upon exiting the device. Experiments so far seem to support this even if it is not clear why. Here we report clear evidence of coherent reflection when electron wave exits the channel of a nanoscale transistor and when the electron energy is low. The observed behavior is well described by a simple rectangular potential barrier model which the Schrodinger's equation can be solved exactly. We can explain why reflection is not observed in most situations but cannot be ignored in some important situations. Our experiment also represents a direct measurement of electron injection velocity - a critical quantity in nanoscale transistors that is widely considered not measurable. [http://dx.doi.org/10.1063/1.4954083]

One of the celebrated advancements in solid-state physics in the last 40 years is the quantized conductance and the minimum resistance in perfect conductors.[1–10] This minimum resistance is caused by the partial reflection of incoming electron flux at the *entrance* of the channel.[13] Consequently, the conductance per mode (quantum state) is independent of channel length as long as it remains ballistic.[11,12] For this to be true electrons exiting the channel must suffer no reflection.[14] Otherwise, a resonance that modulates the conductance will occur. This reflectionless assumption is ubiquitous in models that describe ballistic and quasi-ballistic transport in nanoscale devices.[15–17]

Experimentally, resonance due to reflection has not been observed in nanoscale devices, even in devices where the potential change is relatively sharp.[18,19] While reflection can be negligible if potential change is gradual, it was predicted that even for abrupt potentials, resonance can only be observed when the channel length is equal or larger than the channel width at temperatures well below liquid helium.[14]

Models that describe transport in nano-scale Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs)[15] also posit that only electron scattering at the source-side (entrance) of the channel is important.

On the other hand, exiting from a highly constricted region to a less constricted region was predicted to cause reflection when the electron energy is small[19,20] because a small potential step results from the change in quantum confinement. Since a MOSFET uses a potential barrier to modulate conduction, the potential step is much larger. If coherent reflection at the end of the channel exists, a nanoscale MOSFET is where it may be found.

[a]Corresponding author, Kin.Cheung@NIST.gov

We note that even though the question of reflection for ballistic transport is intimately related to the quantized conductance phenomenon, quantized conductance is not a prerequisite experimentally to address the question. Indeed, failure to observe interference in quantized conductance experiments so far strongly indicates a different approach is needed.

Here we present experimental evidence of strong drain-side reflection upon exiting a 2-D ballistic channel, leading to self-interference of the de Broglie waves of the electrons. The experimental condition is such that sub-bands due to lateral confinement (300 nm) are washed out (unresolved) while sub-bands due to vertical confinement (~2 nm) are well resolved. The experimental data are restricted to within the first sub-band of the vertical confinement. This combination removed the conductance modulation resulting from crossing sub-bands, allowing the interpretation of the observed current modulations (fig. 1(c)) to be focused on reflection.

Scattering can be a source of reflection. Defects in the contact beyond the channel can cause reflections in 1-D devices[21,22] but the width of our 2-D device and the strength of the interference make it unlikely. So we assume the reflection is from the drain end of the channel and treat the self-interference of the de Broglie wave of the electron with a simple classical model. Comparing the extracted injection energy, along with the observed modulation depth to the solutions of the Schrodinger's equation for a one dimensional rectangular barrier, we found very good agreement to the physical parameters of the transistor, providing strong support for the drain end reflection assumption. The success of such a simple text book level model allows an intuitive discussion of why reflection is so hard to observe in ballistic transport.

Incidentally, this represents a direct measurement of the injection velocity of the electron coming into the channel - a task commonly thought impossible.[23] Unexpectedly, the injection energy is found to be much higher than the thermal energy that is commonly assumed in transport models.[15–17]

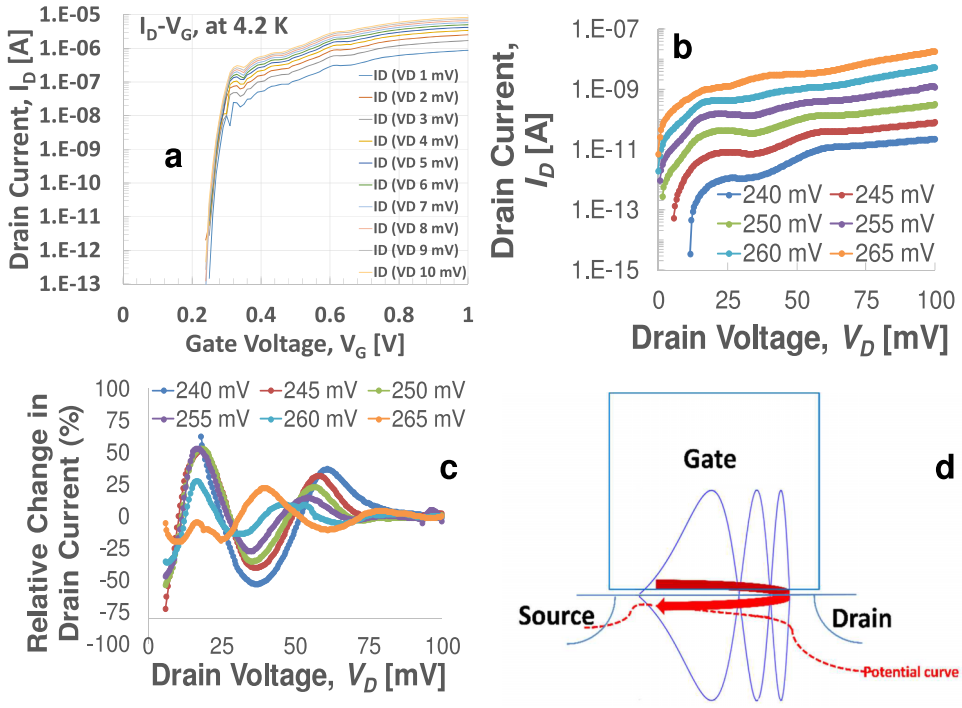

FIG. 1. a, I_D-V_G characteristics of the MOSFET at 4.2 K. b, I_D-V_D characteristics of the same MOSFET at 4.2 K, for several sub-threshold gate bias, V_G. c, The change in drain current, I_D, normalized to the background current as a function of drain voltage, V_D, for various V_G. d, The cross-section of a MOSFET and the potential barrier between the source and drain regions (broken red line). The arrow indicates the propagation, incidence and reflection, of electron de Broglie waves. The wave as depicted is at destructive resonance. (The potential barrier and de Broglie wave are not drawn to scale.)

The transistor used in this work has larger width (300 nm) than length (40 nm physical; 7 nm effective). The potential transition region at source or drain is roughly 20 nm for this device. At liquid helium temperature and low drain bias, this is comparable to the de Broglie wavelength which ranges from 52 nm (0.36 meV) to 10 nm (10 meV). According to some theory, no resonance should be observed.[14] Classical wave theory, on the other hand, predicts significant reflection if the step height is large. Consider a step height of >100 meV (as shown later), the transition is >5 meV/nm – a steep transition for 10 meV particle wave with 10 nm wavelength.

For the quantum interference experiment, a few conditions must be met. First, the channel length must be short enough to ensure ballistic transport. Second, the electrons must have a tight energy distribution to avoid smearing the interference. Third, the device must be free from phase interrupting scattering. This last requirement is very difficult to meet with research devices. Even for our production quality (40 nm technology node) MOSFETs very few devices satisfy these conditions. Upon surveying many devices, we found one that shows clean and strong interference and is reported here. After a complete set of data is collected, this device was subjected to an electrical stress that is designed to create defects at the gate oxide interface. As expected, the interference was destroyed, presumably due to defect scattering.

At 4.2 K, the energy spread of electrons, k_BT, is 0.36 meV and phonon scattering is greatly suppressed. The energy spread of the electrons is large enough that sub-bands due to lateral confinement of the channel ($\ll$1 meV) are not resolved and narrow enough that the sub-bands due to vertical confinement (~100 meV) are fully resolved. As long as we keep the energies within the first sub-band of vertical confinement, we can rule out the observed modulation of current as a result of crossing sub-bands.[24]

Figure 1(a) shows the drain current characteristic of our transistor as a function of gate bias at 4.2 K. The modulation above threshold is well-known – a result of crossing sub-bands produced by vertical field quantum confinement.[25] The first peak at 0.33 V gate bias has a resistance of 33 kΩ. For ballistic transport involving many states in the first sub-band, the channel resistance should be much less than 12.9 kΩ - the ideal value for single state ballistic transport. This means the series resistance of the source and drain is greater than 10 kΩ each, a reasonable value at such low gate bias.[26] As gate bias drops below threshold, the channel is still ballistic, but the injection electron flux drops rapidly, leading to a rapid decrease in current.

Figure 1(b) shows the drain current at 4.2 K as a function of drain bias for a few gate voltages below threshold. A clear drain current modulation is evident. We note that our current level is so low that Joule heating is negligible. The observed current is a superposition of the modulation and transistor "on-current" which can be fitted to a second-order polynomial and removed to facilitate the study of the modulation.[27]

One way to explain the background "on-current" is not all electrons are incident into the channel parallel the direction of source to drain. Some truly oblique incident electrons will suffer scattering and behave as normal non-ballistic current that follows non-ballistic transport equation. At low drain bias and near threshold, this should behave like a "long-channel" transistor and have a square dependent on drain bias.

Given the large width and short channel length, many electrons with less oblique incident angles are ballistic. However, when the angle of incident to the reflection boundary is large, the self-interference is weak and do not contribute to the interference. This contribution to background current cannot be modeled by the "long-channel" equation. Empirically, we found the second order polynomial fits the background current rather well given the short range available for fitting.

Figure 1(c) shows the percent modulation of the background drain current. Over 50% modulation is observed at low gate voltage and low drain bias. Such deep modulation clearly requires a reflector that spans the entire width of the transistor, ruling out defect scattering. The only candidate is the potential drop at the end of the channel.

Even with the strong drain-end reflection picture, it is still a surprise at first glance for the deep modulation. If electrons were entering the channel at random direction, then the fraction that is close to normal incident and produce the strong interference should be relatively small. In reality electrons near the entrance are accelerated by a "drain bias penetration field" beyond the entrance. This boundary condition was discussed by Rahman *et al*. in terms of electrostatic requirement.[28]

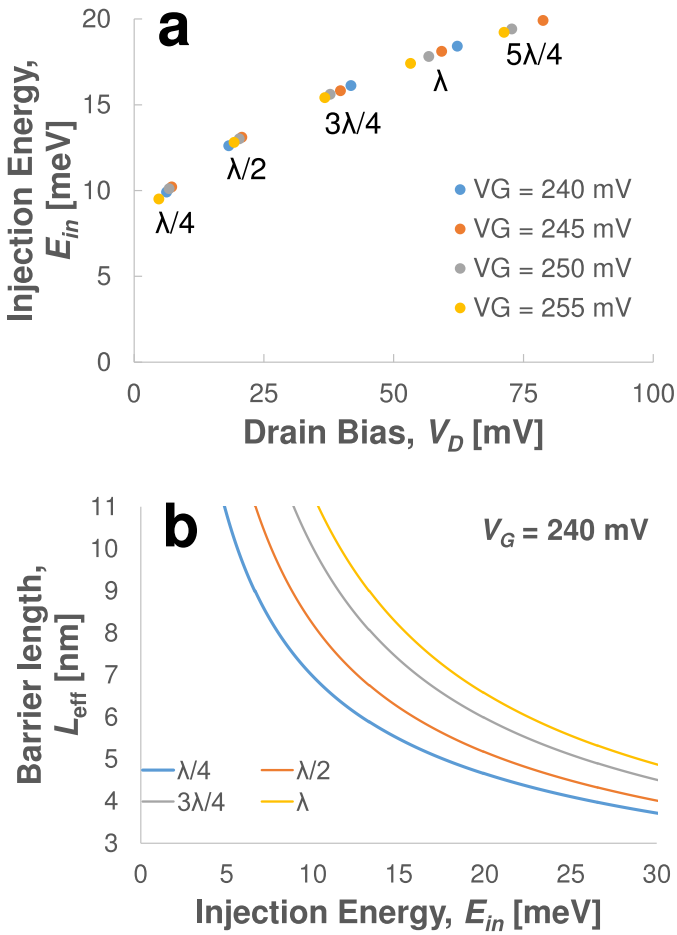

FIG. 2. Electron injection energies. a, Extracted electron injection energies (for 7 nm barrier length) at various resonance conditions versus the drain bias of each resonance, for four different gate bias. Error bars are comparable to the size of the data point and are not shown. b, the variation of barrier length with electron injection energy is calculated for 4 resonance conditions based on the measured I_D-V_G characteristics at a gate voltage, V_G, of 240 mV.

Simply put, as electrons near the entrance of the channel inject into the channel, the local neutrality is disturbed. An acceleration field must develop to attract electrons further away in the source region. The magnitude of the acceleration field is current dependent. Figure 2(a) shows the extracted injection energy as a function of drain bias. This provides a way to estimate the acceleration field. The slope of the first group of injection energy is 0.3, meaning 3 meV for every 10 mV of drain bias. Some of this is due to gate voltage change. Accounting for the clearly steeper slope at lower drain bias, it is safe to say that the acceleration field is roughly 3 mV at 10 mV drain bias. This field is 10 times larger than the thermal energy and has a strong collimating effect on the injecting electrons, leading to most electrons are injecting at close to normal incident angle.

To sum up the experimental evidence: at liquid Helium temperature strong drain current modulation is observed as the drain bias sweep across a region that should not produce sub-band crossing and the modulation depth precludes defect scattering as possible cause. While it is possible alternative explanation other than drain end reflection can produce such result, it is not obvious. In what follows, we explore the theoretical explanation of the observed modulation.

The reflected electron wave interfering with the same electron's incoming wave is illustrated in Figure 1(d). Also shown is the cross-section of the MOSFET and the electrostatic potential barrier from source to drain (dotted line, not to scale). At high enough gate bias, electrons with energy higher than the barrier enter the channel with finite efficiency from the source (drain side injection is neglected because of the drain bias). Once injected into the channel, electrons accelerate toward the drain. In the absence of scattering the electron waves remain coherent.

First, we use a classical wave analysis. At the source end, the de Broglie wavelength is given by the energy of the injecting electrons, which is the Fermi energy of electrons in the source *minus the source side barrier height*. The wavelength shrinks as the electrons accelerate across the channel. At the drain end, the electron waves encounter a change in potential and a partial reflection is assumed. As the reflected waves traverse back toward the source, their wavelength increases. When the reflected wave is exactly 180° out-of-phase with the incident wave at the injection point, as depicted in Figure 1(d), a transmission minimum occurs.

The first self-interference maximum occurs when the one-way transit time equals a half period of the incident wave,

$$t_T = \frac{\lambda_{in}/2}{v_{in}} = \frac{\lambda_{in}/2}{\sqrt{\frac{2E_{in}}{m}}} \tag{1}$$

where λ_{in}, v_{in}, and E_{in} are de Broglie wavelength, velocity and energy of the incoming electron at the source-end, respectively, m is the electron effective mass, and t_T is the transit time through the channel. To handle the position-dependent velocity, we divide the channel length, L_{eff}, into N equal segments of length ξ and each has a transit time, t_i, and injection velocity, v_i. The total transit time is:

$$t_T = \sum t_i = \sum \frac{\xi}{v_i} = \frac{L_{eff}}{N} \sum \frac{1}{v_i} = \frac{L_{eff}\sqrt{m}}{N\sqrt{2}} \sum \frac{1}{\sqrt{E_{in} - E_i}} \tag{2}$$

where E_i is the energy of the i^{th} segment in excess of the incident energy E_{in}, and L_{eff} is the effective channel length. Combining (1) and (2), we get

$$\frac{Nh}{2L_{eff}\sqrt{2m}} = E_{in} \sum \frac{1}{\sqrt{E_{in} - E_i}} \tag{3}$$

where h is Planck's constant. One can similarly obtain expressions for resonance conditions of $\lambda/4$, $3\lambda/4$, λ, etc. Equation (3) links the injection energy of the electron for each resonance condition to the measured drain bias of the resonance, if L_{eff} is known.

As figure 1(d) clearly depicts, L_{eff}, is much shorter than the physical channel length. A method to estimate the barrier length is to assume that the transistor has been engineered to achieve the shortest effective channel length without short-channel effects,[29] a reasonable assumption for our production quality transistor. That means the barrier length goes to zero at full drain bias (V_{DD}). The effective barrier length at any drain bias (V_D) is proportionally given by

$$L_{eff} = L_{drawn}\left(1 - \frac{\sqrt{V_{bi}} + \sqrt{V_{bi} + V_D}}{\sqrt{V_{bi}} + \sqrt{V_{bi} + V_{DD}}}\right) \tag{4}$$

where L_{drawn}, is the drawn channel length and the built-in potential across the source (drain) to substrate junction, V_{bi}, is taken as 0.9V. This results in an estimated effective barrier length (L_{eff}) of 7 nm for our 40 nm drawn transistor. This method of estimating L_{eff} was proven very successful in similar transistors.[29]

Figure 2(a) shows the extracted injection energies for the measured resonances. The first minimum ($\lambda/4$) for each gate bias has injection energy around 10 meV. At first glance, this value seems too large. Room temperature transistor models[15–17] generally assume that the injecting electrons to have thermal energy which is 0.36 meV at 4.2 K, almost 30 times smaller.

Since the extracted injection energy is dependent on the barrier length (fig. 2(b)). One may argue that the effective channel length may be under estimated. However, the curve trend in figure 2(b) suggests that even if the full 40 nm physical gate length is the channel length (not possible), the injection velocity will still be higher than thermal energy. In the next section we will use a quantum model and known physical parameters of the transistor to corroborate the combination of 7 nm barrier length and 10 meV extracted injection energies.

We now examine the one dimensional problem of electron injection over a barrier by solving the Schrodinger equation, as illustrated in figure 3(a). The rectangular barrier has a length of L_{eff},

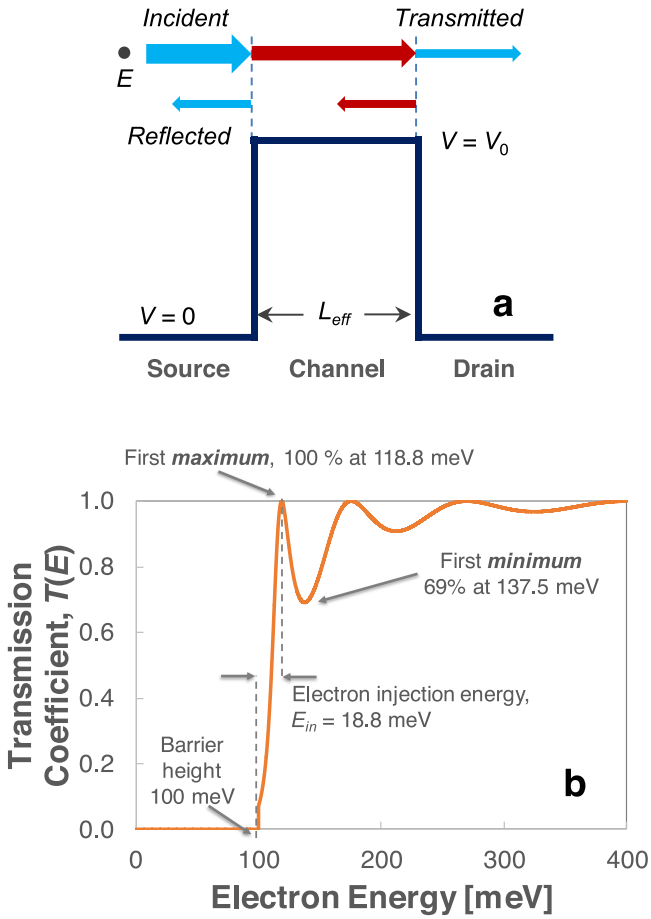

FIG. 3. a, A rectangular potential barrier is used to represent the channel region of the MOSFET without drain bias to examine the quantum transmission of electrons with energy higher than the barrier height. Reflection at entrance and at exit are indicated by arrows pointing backward. b, Calculated transmission coefficient, $T(E)$, of an electron for a barrier height of 100 meV and barrier length of 7 nm. The first maximum occurs at 118.8 meV, leading to electron injection energy of 18.8 meV at round trip = $\lambda/2$ resonance.

and a height of V_0. The incoming electrons are at energy above the barrier. This text book problem has a well-defined solution[30] and the results are shown in Figure 4.

Transmission coefficients as a function of incoming electron energy for four different barrier heights are plotted for each barrier length in figure 4. Notice that in each case (barrier length) the modulation depth increases with barrier height. Thus our observed 50% modulation can be linked to a specific barrier height for a given barrier length. The barrier heights that match experiment are approximately 233 meV, 176 meV, 133 meV, and 98 meV for barrier lengths of 6 nm, 7 nm, 8 nm, and 9 nm, respectively.

The barrier height is the difference between the injection energy and the source region Fermi energy E_F which can be found from the dopant concentration using the free-electron model[31]:

$$E_F = \frac{\hbar^2}{2m}\left(\frac{3\pi^2 N}{V}\right)^{2/3} \tag{5}$$

where N/V is electron density. We should mention that since we are using a MOSFET, the source region is not a 2-D electron gas as in many quantum point contact experiments. The 3-D equation (5) is appropriate for our purpose. A plot of Fermi energy as a function of electron density (assume full ionization- valid for degenerate doping even at 4.2 K) is shown in Figure 5.

For the 40 nm transistor technology, typical source doping concentration range is 2 to 3.5 x 10^{20}/cm^3, corresponding to an E_F range of 124 meV to 177 meV. With 10 meV injection energy,

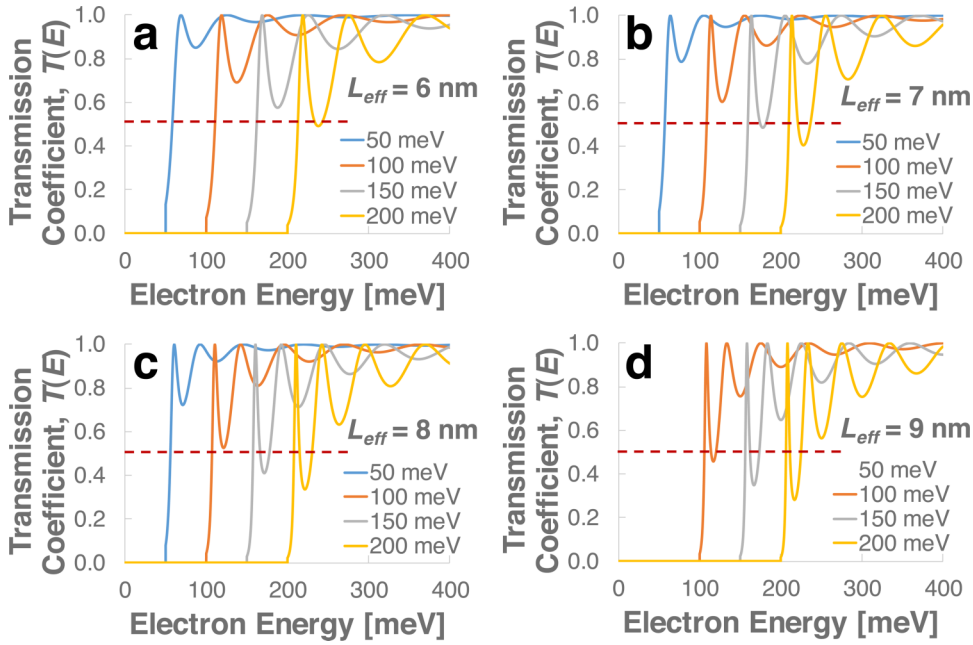

FIG. 4. a-d, Transmission coefficients for barrier lengths, L_{eff}, of 6, 7, 8, and 9 nm, respectively, and for 4 barrier heights. Broken horizontal line indicates the transmission coefficient, $T(E)$, at 0.5 (i.e., to 50%) – a point matched in the experiment.

the range of barrier height comes to 114 meV to 167 meV. Thus we can pin down the barrier length to within the range of roughly 7 nm to 8 nm.

In figure 4 the first transmission peak is located just above the barrier at energies (E_{peak}-$E_{barrier}$) 18.9 meV, 13.7 meV, 10.6 meV, 8.33 meV for barrier lengths of 6 nm, 7 nm, 8 nm, and 9 nm, respectively. They correspond to our injection energies in figure 2(b). However, the quantum calculation did not include drain bias while the extracted value is from experiment under ~10 mV bias. One expects the quantum calculated value to be larger than the extracted value and not by just 0.6 meV. Thus we can reasonably argue that the first peak location points to a barrier length closer to 7 nm than 8 nm.

In other words, the quantum calculation confirms that the combination of 7 nm barrier length and 10 meV extracted injection energy is correct. The much higher than thermal energy of the injecting electron is a necessary consequence of the wave nature of the electrons.

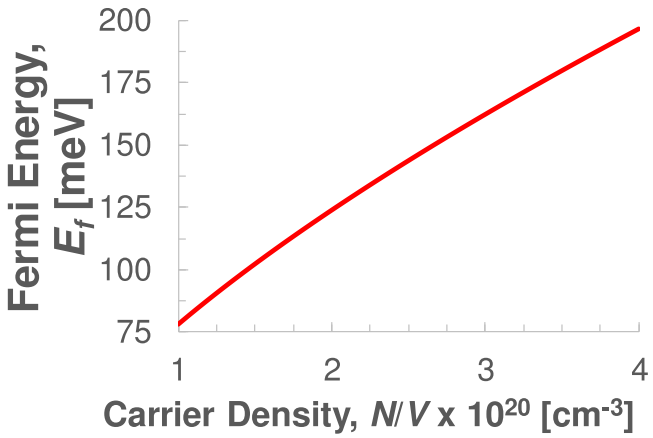

FIG. 5. Fermi energy in the source region versus carrier densities (dopant concentrations, full ionization assumed) according to equation (5). Typical source doping concentrations in 40 nm transistor technologies are in the 2 – 3.5 x 10^{20} cm^{-3} range, corresponding to Fermi energy from 124 meV to 177 meV.

With the classical and quantum model in full agreement, and all the parameters extracted in full agreement with the physical property of the device, our drain end reflection interference model is fully versified.

It is instructive to clarify the picture of carrier transport in ballistic MOSFET at liquid Helium temperature: As the barrier is driven by gate bias to below the Fermi energy of the source and drain, carrier from both side can flow across. Since the thermal energy is only 0.36 meV, a small drain bias of 1 mV will stop the drain side injection (no empty state available). The source side injection, while allowed, is inefficient due to the wave nature of the electrons. As the barrier is further lowered, the source side injection efficiency rises quickly as indicated by the transmission probabilities shown in figures 3 and 4. This rise is not due to change in the fraction of electrons with energy over the barrier, as is expect for room temperature operation. It is entirely a result of the wave nature of electrons. At this point the barrier is well below the Fermi surface and there are many filled energy levels within the transistor channel. As only electrons in the narrow thermal energy range at the Fermi surface is mobile, the injection remains one sided as long as there is a drain bias.

Returning to the question of why is reflection generally not observed? The simple quantum model of the barrier provides a clear answer. As figure 4 shows, at high electron energy (relative to the barrier height), the resonance quickly diminishes. Most quantum point contact experiments involve metallic sample with high Fermi energy. At very low barriers, the resonance also diminishes. A pure constriction change-induced potential step is likely too low to produce an observable resonance. In nanoscale MOSFETs, the barrier height and the electron energy are both adjustable, giving the best chance to see the resonance. However, defect-free devices are rare. So even if one is looking for the interference, it is clearly elusive.

In summary, we have experimentally demonstrated strong quantum interference in a ballistic transistor. Although only evidence from one device is obtained, the signal to noise ratio and the strength of the modulation demands that it be taken seriously. The simple text book level analysis is completely consistent with the physical parameter of the device, providing strong support that the observed interference is due to reflection of coherently-transported electron waves at the drain end of the channel due to the change in potential. The conceptual implication of this work on ballistic transport in general and quantized conductance in particular is significant.

The authors like to acknowledge the helpful discussions with Professor S. Datta and Professor M. Lundstrom at Purdue University.

[1] E. N. Economou and C. M. Soukoulis, Phys. Rev. Lett. **46**, 618-621 (1981).
[2] D. S. Fisher and P. A. Lee, Phys. Rev. B. **23**, 6851-6854 (1981).
[3] D. J. Thouless, Phys. Rev. Lett. **47**, 972-972 (1981).
[4] D. C. Langreth and E. Abrahams, Phys. Rev. B. **24**, 2978-2984 (1981).
[5] C. P. Umbach, S. Washburn, R. B. Laibowitz, and R. A. Webb, Phys. Rev. B. **30**, 4048-4051 (1984).
[6] R. Webb, S. Washburn, C. P. Umbach, and R. B. Laibowitz, Phys. Rev. Lett. **54**, 2696-2699 (1985).
[7] V. Chandrasekhar, M. J. Rooks, S. Wind, and D. E. Prober, Phys. Rev. Lett. **55**, 1610-1613 (1985).
[8] S. Datta *et al.*, Phys. Rev. Lett. **55**, 2344-2347 (1985).
[9] A. Benoit, C. P. Umbach, R. B. Laibowitz, and R. A. Webb, Phys. Rev. Lett. **58**, 2343-2346 (1987).
[10] G. Timp *et al.*, Phys. Rev. Lett. **59**, 732-735 (1987).
[11] A. D. Stone and Szafer, IBM J. Res. Develop. **32**, 384-413 (1988).
[12] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, 1995).
[13] R. Z. Landauer, Phys. B - Condensed Matter. **68**, 217-228 (1987).
[14] A. Szafer and A. D. Stone, Phys. Rev. Lett. **62**, 300-303 (1989).
[15] M. S. Lundstrom, IEEE Electron Dev. Lett. **18**, 361-363 (1997).
[16] M. V. Fischetti, S. Jin *et al.*, in 13th Int. Workshop on Computational Electronics, IWCE '09, 2009.
[17] V. K. Arora, in *Proc. 19th Int. Conf. Mixed Design of Integrated Circuits and Systems (MIXDES), 2012*. pp. 17-24.
[18] B. J. van Wees *et al.*, Phys. Rev. Lett. **60**, 848-850 (1988).
[19] A. M. Kriman and P. P. Ruden, Phys. Rev. B **32**(12), 8013-8020 (1985).
[20] R. Frohne and S. Datta, J. Appl. Phys. **64**(8), 4086-4090 (1988).
[21] P. A. M. Holweg *et al.*, Phys. Rev. Lett. **67**, 2549-2552 (1991).
[22] C. Untiedt, G. R. Bollinger, S. Vieira, and N. Agrait, Phys. Rev. B. **62**, 9962-9965 (2000).
[23] L. Yang, M. Luisier *et al.*, IEEE Trans. Elect. Dev. **59**(4), 994-1001 (2012).
[24] G. Wirth, U. Hilleringmann, J. Horstmann, and K. Goser, in Proc. 27th European.Solid-State Device Research Conf., 1997.
[25] F. Stern and W. E. Howard, Phys. Rev. **163**, 816-835 (1967).

[26] J. P. Campbell, K. P. Cheung, J. S. Suehle, and A. Oates, IEEE Electron Device Letters **32**(8), 1047-1049 (2011).
[27] See supplementary material at http://dx.doi.org/10.1063/1.4954083 for the background removal procedure.
[28] A. Rahman *et al.*, IEEE Trans. Electron Dev. **50**, 1853-1864 (2003).
[29] R. G. Southwick *et al.*, in *Proc. 2012 IEEE Silicon Nanoelectron. Workshop*, *Honolulu, HI*, *June 10-11* (2012).
[30] See, for example, R. M. Eisberg, *Fundamentals of Modern Physics* (John Wiley & Sons, Hoboken, USA, 1967).
[31] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Cengage Learning, Belmont, USA, 1976).