P H Y S I C A L   R E V I E W   L E T T E R S

Quantum Confinement and Electronic Properties of Tapered Silicon Nanowires

Zhigang Wu,[1,2] J. B. Neaton,[2,*] and Jeffrey C. Grossman[1,†]

[1]*Berkeley Nanotechnology and Nanoscience Institute (BNNI), University of California at Berkeley, Berkeley, California 94720, USA*
[2]*Molecular Foundry, Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
(Received 26 August 2007; published 20 June 2008)

Using *ab initio* calculations, structural tapering of silicon nanowires is shown to have a profound effect on their electronic properties. In particular, the electronic structure of small-diameter tapered silicon nanowires is found to have a strong axial dependence, with unoccupied eigenstates being substantially more sensitive to diameter. Moreover, the states corresponding to the highest occupied and the lowest unoccupied states are spatially separated along the wire axis by the tapering-induced charge transfer and a strong electrostatic potential gradient, due to an appreciable variation in quantum confinement strength with diameter.

In recent years semiconductor nanowires (NWs) have attracted tremendous attention due to their remarkable thermal, electronic, photonic, chemical, and mechanical properties [1–4]. Advances in synthesis will likely enable potential applications in electronics, spectroscopic sensing, biological sciences, and photovoltaics [5]. Silicon nanowires (Si NW) are particularly appealing because they can be fabricated routinely (with diameters d as small as ~ 1 nm [2,4]), and they are readily compatible with current Si-based technology [6–8]. With a reduction in d, Si NWs exhibit quantum confinement perpendicular to the growth direction, which, as in the case of quantum dots, results in unique electronic and optical properties [9]. For example, as d decreases, a strong blueshift of photon emission has been observed in photoluminescence studies [10,11]. Experiments report the optical gaps of Si NWs with $d =$ 1.3 and 2.0 nm to be 3.5 and 2.3 eV [4], respectively, compared to the bulk silicon optical absorption band edge of about 1.2 eV.

Previous theoretical studies have examined the electronic properties of straight-edged nanowires (SNWs) [12–18]; however, under typical growth conditions, NWs are often observed to be tapered [19,20] by as much as 1 nm in diameter for every 10 nm in length (L) [21,22]. Significant efforts have been expended in experiments to prevent tapering [22], even though the possible advantages or disadvantages are currently unknown. Tapering increases the surface to volume ratio (relative to SNWs with the same volume) and the variation of diameter along the wire alters quantum confinement in a manner analogous to wires with different surface reconstructions [15,16] or terminating chemistry groups [17,18,23]. For small-diameter SNWs, previous *ab initio* calculations predicted that surface morphology can change the optical gap by as much as 1 eV [18]. To our knowledge, however, theoretical studies have yet to examine the effects of a tapered morphology on confinement, despite its fundamental and practical importance.

In this Letter, calculations based on density functional theory (DFT) are employed to elucidate the effects of tapering on the electronic properties of Si NWs. Tapered nanowires (TNWs) possess axially dependent electronic properties. In contrast to SNWs, where confinement tends to shift both valence bands and conduction bands equally away from the bulk gap, the *unoccupied* states in TNWs are much more sensitive to diameter than the occupied states. Furthermore, our calculations show that in TNWs the highest-occupied state is *spatially isolated* in the narrowest parts, while the lowest-unoccupied state resides principally in the widest parts. Thus, tapering causes these two states to be separated along the wire axis, a consequence of its nanoscale morphology which may be exploited in electronic and optical applications.

Structural and electronic properties are computed for both TNWs and SNWs, passivated with H atoms, with wire axes oriented along the [011] direction, as is common experimentally [4]. These NWs have a direct gap at the Γ point [24]. All NWs are constructed without SiH_3 present on the surface. Calculations are carried out with the SIESTA package [25], using the generalized gradient approximation (GGA) [26] and norm-conserving pseudopotentials. A double-ζ polarization basis is used for both Si and H, and is found to reproduce well the DFT bulk Si lattice constant and band structure. The distance between NWs is set so that the vacuum length is at least 1.5 nm. An energy cutoff of 100 Ry is sufficient to converge the grid integration of the charge density, and all structures are relaxed until the forces are less than 0.02 eV/Å.

We first consider a periodic Si TNW with d varying from 1.2 to 2.2 nm over a length of 10 nm. The tapering divides the unit cell into 9 connected segments with 5 distinct average diameters. The projected electron densities of states (PDOS) of Si atoms at the largest (solid curves) and smallest (dashed curves) diameters of the TNW are shown in Fig. 1(a), together with those PDOS for the SNWs of the same d. The average PDOS (per Si atom) at a given d in the TNW is quite similar to that of the corresponding SNW. For small d (1.2 nm), the PDOS curves almost lie on top of one another, with only a slight shift to higher energy visible; however as d increases, the
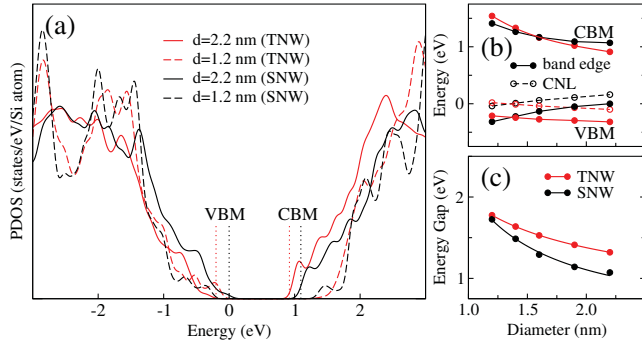
FIG. 1 (color). (a) Si atoms PDOS at the widest (2.2 nm) and the narrowest (1.2 nm) parts in the (relaxed) TNW (red curves), compared with those for the corresponding SNWs (black curves). The energy zero is set at the VBM of the SNW with $d = 2.2$ nm. The vertical dotted lines show the VBM and CBM energies for the TNW and the widest SNW. (b) Solid curves show the VBM and CBM energies as a function of d in the TNW (red) and the SNWs (black); and dashed curves show the charge neutrality levels (CNL). (c) The energy gaps between the local VBM and CBM as a function of d in the TNW and SNWs. The solid lines are fitted to the data points (circles) with the function $E_g^{bulk} + C(1/d)^\alpha$. For SNWs, $\alpha = 1.6$, in good agreement with the previous DFT result of 1.7 [12], but less than the experimental [4] and GW [12] result of 2.0. For the TNW, $\alpha = 1.1$.

TNW PDOS shifts to lower energy relative to the SNW, and the magnitude of this shift increases with d. The result is that the "local" valence band edge of each segment is at roughly the same energy over the length of the TNW, while the local conduction band edge varies considerably in energy, depending on d.

It is useful to compare the local valence and conduction band edges of TNW segments with those of SNWs with the same d. As summarized in Fig. 1(b), both the local valence band maximum (VBM) and conduction band minimum (CBM) in the TNW decrease with growing d in the TNW; in contrast, in SNWs the CBM decreases and the VBM increases. As noted above, the local VBM of the TNW has a weak dependence on d, while the CBM exhibits considerable variation. As d increases from 1.2 to 2.2 nm, the shifts of VBM and CBM in the TNW are about -0.1 and -0.6 eV, respectively; however, for SNWs, the corresponding shifts are 0.31 and -0.34 eV. Thus, in contrast to SNW, the widening energy gap with tapering originates from a systematic shift in the energetic position of the unoccupied states. Figure 1(c) shows that the local energy gap in a given TNW segment is larger than the gap of a SNW with corresponding d, and for our TNW, the difference increases with d. The bigger local gap in TNW is due to additional confinement along the z direction: segments with larger d are not as long as those with smaller d within our unit cell, resulting in greater confinement. For these reasons, the overall band gap of this TNW within DFT-GGA is 1.18 eV, slightly larger than 1.08 eV, the value computed for a SNW with the maximum d of the TNW.

To understand the diameter dependence of the local band edges in the TNW, we compute the local charge neutrality level (CNL) of each segment of the TNW. The CNL is defined as the branch point of the complex band structure [27], the point at which the gap states cross over from valence band- to conduction band-type character, and can be computed as $G(E) = \mathrm{Tr}(\frac{1}{E-\hat{H}}) = \sum_{n\mathbf{k}} \frac{1}{E-\varepsilon_{n\mathbf{k}}} = \int \frac{g(E')}{E-E'} dE' = 0$, where $g(E)$ is the DOS. For bulk Si, we obtain $E_{CNL} \sim 0.06$ eV above its VBM (DFT-GGA), in excellent agreement with previous results [28]. The CNLs of Si NWs are also all significantly closer to the VBM than to the CBM [Fig. 1(b)], and the E_{CNL} increases with d in straight NWs. This results in a tendency for charge to transfer from wider to narrower segments in the TNW in an effort to align the local CNLs. To explore this further, we plot the layer-by-layer integral of the difference in electron density between tapered and straight NWs of a given d in Fig. 2. Most importantly, we find a general trend that more electrons accumulate in narrow regions relative to wide regions, as indicated by the red arrows in Fig. 2; in addition, we observe appreciable changes in local electron distribution at interface regions where the diameter changes, and at surfaces where Si-Si bond lengths and angles deviate from those in SNWs. We also find that structural relaxation in the TNW reinforces the trend observed in the unrelaxed case. Therefore, electrostatic equilibrium results in a local CNL (and VBM) that is higher in energy at the narrowest diameters than at the widest diameters.

The electronic structure of the TNW implies that tapering has a strong impact on the spatial location of the highest-occupied and lowest-unoccupied states. As indicated by the vertical red dotted lines in Fig. 1(a), the highest-occupied state of the TNW is mainly derived from electrons associated with atoms in the narrowest segments, and the lowest-unoccupied band is mainly associated with atoms in the widest segments. It is clear from Figs. 3(a) and 3(b) that these two states in the TNW are spatially separated and mainly located in the narrowest and the widest parts, respectively. This phenomenon is further
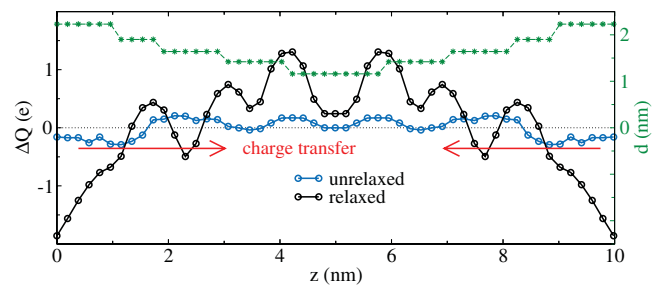

FIG. 2 (color). The layer-by-layer integral [35] of the difference in electron density ΔQ (circles) between the tapered and straight NWs as a function of d and layer. The green curve displays d values (right axis label) of the TNW along the wire axis.
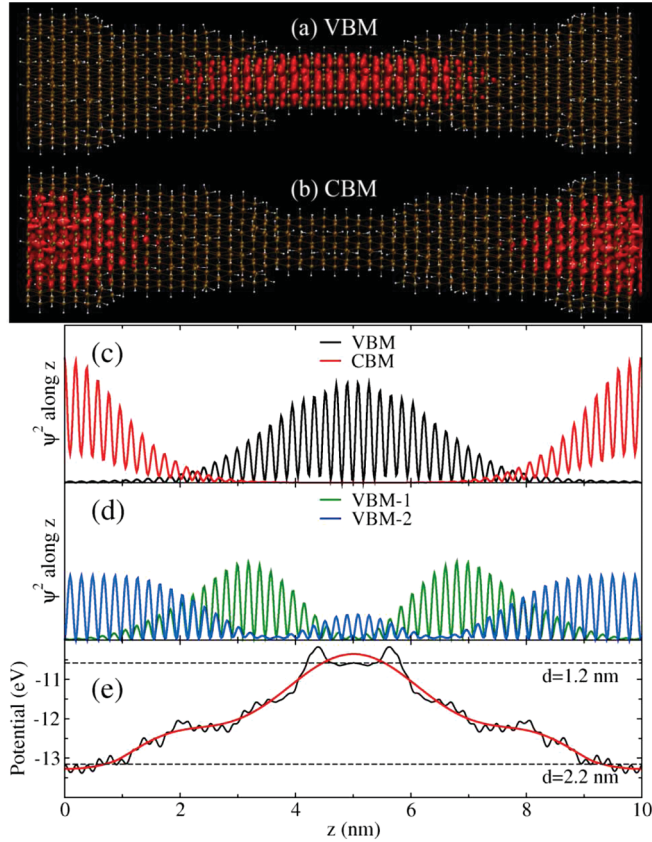
FIG. 3 (color). Panels (a–b) show isosurfaces (red) of the square of the wave functions of the VBM and CBM in the TNW at values of 10% of their respective maximum. Panels (c–d) show the charge density distribution of VBM, the CBM, VBM-1, and VBM-2 states along the wire axis (integrated over the *xy* plane). Panel (e) shows the planar-averaged total potential $V(z)$ along the wire axis for the TNW (solid black) and SNWs (dashed black). The thick red curve is a polynomial fit to the black solid curve.

illustrated by the planar-averaged axial charge distribution, shown in Fig. 3(c). These figures also show the difference in character between the highest occupied and lowest unoccupied states, revealing their respective bonding and antibonding nature. Figure 3(d) indicates that the VBM-1 and VBM-2 states primarily occupy the middle and widest segments; on the other hand, there are several states close in energy just above the CBM, and they are all spatially confined to the widest regions of the TNW. Note that these observations are all consistent with the PDOS curves [Fig. 1(b)].

The effective-mass approximation has proven useful for understanding electronic confinement in SNWs [12]. Along the *z* direction, the envelope function φ_f^z can be approximated by the wave function of a particle within the potential well $V(z) = \frac{1}{2a} \int_{z-a}^{z+a} \bar{V}(z') dz'$, with *a* the distance between layers along [011] and $\bar{V}(z)$ the planar average of the potential within NWs. Physically, these valence (conduction) band edge states can be approximately regarded as bound hole (electron) states within the finite potential

well of Fig. 3(e). For a harmonic potential well, $\varphi_f^z \propto \exp(-C_1 z^2) H_j(C_2 z)$, where C_1 and C_2 are positive constants, and H_j is the *j*-th order Hermite polynomial. For the lowest-lying states, $H_0(z) = 1$, resulting in valence and conduction band edge wave functions centered in the narrowest and widest parts, respectively. For the potential profile Fig. 3(e), the number of bound states can be estimated from the condition $\Delta V > (j + \frac{1}{2})\hbar\sqrt{\frac{k}{m^*}}$, where *k* is the potential well curvature, and m^* is the effective mass. Using previously computed values of effective masses of bulk Si, $m_h^* = 0.39 m_e$ and $m_e^* = 0.2 m_e$ [29], and $\Delta V \sim$ 2.5 eV, we find 4 and 7 bound states for holes and electrons, respectively. This simple estimate is consistent with our *ab initio* calculations, which show more bound electron states in the widest regions of the TNW, and fewer bound hole states in the narrowest regions.

An important question is whether many-electron interactions, which have been shown to be important for small-diameter straight Si NWs [12,30,31], might affect the spatial separation of the highest-occupied and lowest-occupied states. Recent first-principles quasiparticle calculations [32] for straight Si nanowires [12,30,31] reported quasiparticle gaps more than a factor of 2 larger than those obtained with DFT-GGA. In addition, previous theoretical studies demonstrate large exciton binding energies, of ~1.0 eV, were reported for 1.2 nm-diameter NWs using a Bethe-Salpeter approach [33]. Further consideration of our results, however, suggests that many-electron effects are not likely to alter our main conclusions for TNWs. First, we expect self-energy corrections will open the quasiparticle gap relative to DFT-GGA without affecting the order of the band edge states, preserving the spatial separation of low-energy quasiparticles we observe. Although the states near the VBM experience additional confinement along the wire axis as depicted in Fig. 3(e), Fig. 3(c) indicates that these VBM states are still spread over ~5 nm. Thus the bulk of the confinement energy associated with these VBM states comes from the 1.2 nm local diameter of the TNW, and their self-energy corrections would therefore be expected to be similar to an extended SNW state in a 1 nm diameter SNW. (Similar arguments should apply to the CBM states, where the confinement energy is likely even smaller, given that these states reside in regions of the wire with larger diameter.) Second, although excitonic effects are important for straight Si NWs, binding energies of ~1 eV are small compared to the potential difference of 2.5 eV that develops along the wire. A 1 eV exciton, consequently, would not be expected to remain bound in the presence of such a large potential difference, and the spatially separated electron and hole states we observe near the band edges should be only weakly correlated.

Given the difficulty of synthesizing periodically repeating TNWs, we also examine a finite tapered nanorod with $L = 10$ nm and *d* varying from 1.2 to 2.2 nm. Figs. 4(a) and 4(b) illustrate that the highest-occupied molecular
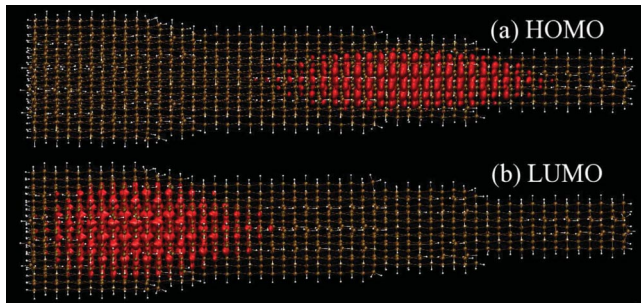
FIG. 4 (color). Isosurfaces (red) of (a) HOMO and (b) LUMO orbital densities of the tapered nanorod, plotted at 10% of their respective maximum.

orbital (HOMO) is located on the narrow side and the lowest-unoccupied molecular orbital (LUMO) the wide side, well-separated along the rod axis just as in the periodic wires. However, due to H-termination, the HOMO and LUMO must vanish close to the edges, leading to a shift of the center of the HOMO (LUMO) in the tapered rod away from the center of the narrowest (widest) segment. This shift can be altered by modifying the termination chemistry on the rod ends. Based on this and other calculations, we expect that the electronic structures of periodic TNWs are qualitatively similar to those of the corresponding monotonically tapered nanorods. In fact, we find the charge transfer and potential distribution in the tapered rod are nearly identical to those in the TNW.

Our results point to the possibility of a new mechanism to separate electrons and holes in nanowire devices. Although the HOMO and LUMO are mostly localized in the narrowest and widest parts separately, the orbitals at higher excitation energies extend along the wire axis, as shown in Fig. 1(a), with a large overlap in real space. One example of a way to take advantage of this unique effect would be in a solar cell, where thermalization and internal electric fields could separate spatially the electron and hole into different ends of the TNW, as apposed to the use of dopants to form a *p-n* junction. In another example, TNWs could be useful in Si-based laser devices because the exciton carriers would have longer lifetime without introducing interface states [34].

*jbneaton@lbl.gov
†jgrossman@berkeley.edu

[1] A. M. Morales and C. M. Lieber, Science **279**, 208 (1998).
[2] J. D. Holmes *et al.*, Science **287**, 1471 (2000).
[3] Y. Cui *et al.*, Appl. Phys. Lett. **78**, 2214 (2001).
[4] D. D. D. Ma *et al.*, Science **299**, 1874 (2003).
[5] C. M. Lieber and Z. L. Wang, MRS Bull. **32**, 99 (2007), and references therein.
[6] Y. Cui and C. M. Lieber, Science **291**, 851 (2001).
[7] L. J. Lauhon, M. S. Gudiksen, D. Wang, and C. M. Lieber, Nature (London) **420**, 57 (2002).
[8] Y. Huang *et al.*, Science **294**, 1313 (2001).
[9] G. Hodes, Adv. Mater. **19**, 639 (2007).
[10] L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
[11] X. Duan, J. Wang, and C. M. Lieber, Appl. Phys. Lett. **76**, 1116 (2000).
[12] X. Zhao, C. M. Wei, L. Yang, and M. Y. Chou, Phys. Rev. Lett. **92**, 236805 (2004).
[13] J. Li and A. J. Freeman, Phys. Rev. B **74**, 075333 (2006).
[14] Y. M. Niquet *et al.*, Phys. Rev. B **73**, 165319 (2006).
[15] R. Rurali, A. Poissier, and N. Lorente, Phys. Rev. B **74**, 165324 (2006).
[16] T. Vo, A. J. Williamson, and G. Galli, Phys. Rev. B **74**, 045116 (2006).
[17] P. W. Leu, B. Shan, and K. Cho, Phys. Rev. B **73**, 195320 (2006).
[18] M. Nolan, S. O'Callaghan, G. Fagas, and J. C. Greer, Nano Lett. **7**, 34 (2007).
[19] Y. Kim *et al.*, Nano Lett. **6**, 599 (2006).
[20] Y. Wang, V. Schmidt, S. Senz, and U. Gösele, Nature Nanotechnology **1**, 186 (2006).
[21] E. Garnett and P. Yang (private communication).
[22] L. Tong *et al.*, Nanotechnology **16**, 1445 (2005).
[23] A. Puzder, A. J. Williamson, J. C. Grossman, and G. Galli, Phys. Rev. Lett. **88**, 097401 (2002).
[24] The electronic properties of Si NWs of other growth directions have been shown to be very similar [4,12].
[25] J. M. Soler *et al.*, J. Phys. Condens. Matter **14**, 2745 (2002).
[26] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[27] J. Tersoff, Phys. Rev. Lett. **52**, 465 (1984).
[28] J. Robertson, J. Vac. Sci. Technol. B **18**, 1785 (2000).
[29] M. A. Green, J. Appl. Phys. **67**, 2944 (1990).
[30] J.-A. Yan, L. Yang, and M. Y. Chou, Phys. Rev. B **76**, 115319 (2007).
[31] L. Yang, C. D. Spataru, S. G. Louie, and M. Y. Chou, Phys. Rev. B **75**, 201304(R) (2007).
[32] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[33] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000).
[34] L. Pavesi *et al.*, Nature (London) **408**, 440 (2000).
[35] To reduce spurious contributions from differences in atomic positions between SNWs and TNWs, we average the density difference over an interval of 0.5 nm along the wire axis, then perform the layer-by-layer integral to obtain the charge density.