Energy Gaps in Graphene Nanoribbons

Young-Woo Son,[1,2] Marvin L. Cohen,[1,2] and Steven G. Louie[1,2,*]

[1]*Department of Physics, University of California at Berkeley, Berkeley, California 94720, USA*
[2]*Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

(Received 29 June 2006; published 22 November 2006)

Based on a first-principles approach, we present scaling rules for the band gaps of graphene nanoribbons (GNRs) as a function of their widths. The GNRs considered have either armchair or zigzag shaped edges on both sides with hydrogen passivation. Both varieties of ribbons are shown to have band gaps. This differs from the results of simple tight-binding calculations or solutions of the Dirac's equation based on them. Our *ab initio* calculations show that the origin of energy gaps for GNRs with armchair shaped edges arises from both quantum confinement and the crucial effect of the edges. For GNRs with zigzag shaped edges, gaps appear because of a staggered sublattice potential on the hexagonal lattice due to edge magnetization. The rich gap structure for ribbons with armchair shaped edges is further obtained analytically including edge effects. These results reproduce our *ab initio* calculation results very well.

DOI: 10.1103/PhysRevLett.97.216803 PACS numbers: 73.22.−f, 72.80.Rj, 75.70.Ak

The electronic structure of nanoscale carbon materials such as fullerenes and carbon nanotubes has been the subject of intensive research during the past two decades [1] because of fundamental scientific interest in nanomaterials and because of their versatile electronic properties that are expected to be important for future nanoelectronics [2,3]. Among the carbon nanostructures, a simple variation of graphene, ribbons with nanometer sized widths, has been studied extensively [4–15]. Because of recent progress in preparing single graphite layers on conventional device setups, the graphene nanoribbons (GNRs) with varying widths can be realized either by cutting [16] mechanically exfoliated graphenes [17,18], or by patterning epitaxially grown graphenes [19,20].

Since GNRs are just geometrically terminated single graphite layers, their electronic structures have been modeled by imposing appropriate boundary conditions on Schrödinger's equation with simple tight-binding (TB) approximations based on π-states of carbon [4–7] or on a two-dimensional free massless particle Dirac's equation with an effective speed of light ($\sim 10^6$ m/s) [8–10]. Within these models, it is predicted that GNRs with armchair shaped edges can be either metallic or semiconducting depending on their widths [4–10], and that GNRs with zigzag shaped edges are metallic with peculiar edge states on both sides of the ribbon regardless of its widths [4–14].

Although the aforementioned models are known to describe the low energy properties of graphene very well [17–22], a careful consideration of edge effects in nanometer sized ribbons are required to determine their band gaps because, unlike the situation in graphene, the bonding characteristics between atoms change abruptly at the edges [7,15]. Moreover, the spin degree of freedom is also important because the GNRs with zigzag shaped edges have narrow-band edge states at the Fermi energy (E_F) implying possible magnetization at the edges [4,11–13]. Motivated by the recent experimental progress in this area, we have carried out first-principles calculation and theoretical analysis to explore the relation between the band gap and the geometries of GNRs.

In this Letter, we show that GNRs with hydrogen passivated armchair or zigzag shaped edges both always have nonzero and direct band gaps. The origins of the band gaps for the different types of homogenous edges vary. The band gaps of GNRs with armchair shaped edges originate from quantum confinement, and edge effects play a crucial role. For GNRs with zigzag shaped edges, the band gaps arise from a staggered sublattice potential due to spin ordered states at the edges [4,11–13]. Although the ribbon widths and energy band gaps of the GNRs are related to each other primarily in inverse proportion, there is a rich structure in the ratio of the proportionalities as in the behavior of carbon nanotubes [1]. For GNRs with armchair edges, analytic scaling rules for the size of the band gaps are obtained as a function of width including the effect of the edges and give a good agreement with our first-principles calculations.

Our electronic structure calculation employs the first-principles self-consistent pseudopotential method [23] using the local (spin) density approximation [L(S)DA] [23,24]. An energy cutoff (for a real space mesh size) of 400 Ry is employed and a double-ζ plus polarization basis set [23] is used for the localized basis orbitals to deal with the many atoms in a unit cell of the GNRs of various widths. We obtained the electron density by integrating the density matrix with a Fermi-Dirac distribution [23,25]. The geometry for each GNR studied is fully relaxed until the force on each of the atoms is less than 16 pN. A k-point sampling of 32 (96) k points that are uniformly positioned along the 1D Brillouin zone is employed for GNRs with armchair (zigzag) shaped edges. We have tested the change of gap size by increasing the vacuum between edges from 20 to 40 Å and between planes from 16 to 25 Å and found no changes.

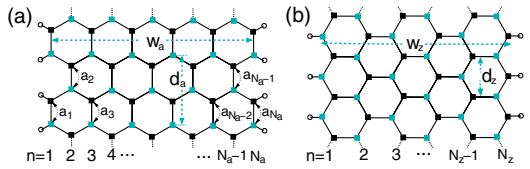

FIG. 1 (color online). (a) Schematic of a 11-AGNR. The empty circles denote hydrogen atoms passivating the edge carbon atoms, and the black and gray rectangles represent atomic sites belonging to different sublattice in the graphene structure. The 1D unit cell distance and ribbon width are represented by d_a and w_a, respectively. The carbon-carbon distances on the nth dimer line is denoted by a_n. (b) Schematic of a 6-ZGNR. The empty circles and rectangles follow the same convention described in (a). The 1D unit cell distance and the ribbon width are denoted by d_z and w_z, respectively.

Following previous convention [4–15], the GNRs with armchair shaped edges on both sides are classified by the number of dimer lines (N_a) across the ribbon width as shown in Fig. 1(a). Likewise, ribbons with zigzag shaped edges on both sides are classified by the number of the zigzag chains (N_z) across the ribbon width [Fig. 1(b)]. We refer to a GNR with N_a dimer lines as a N_a-AGNR and a GNR with N_z zigzag chains as a N_z-ZGNR.

Our calculations show that the N_a-AGNRs are semiconductors with energy gaps which decrease as a function of increasing ribbon widths (w_a). The variations in energy gap however exhibit three distinct family behaviors [Fig. 2]. Moreover, the energy gaps obtained by a simple TB model are quite different from those by first-principles calculations. The TB results using a constant nearest neighbor hopping integral, $t = 2.7$ eV [22] between π-electrons are summarized as function of width in Fig. 2(a). It shows that a N_a-AGNR is metallic if $N_a = 3p + 2$ (where p is a positive integer) or otherwise, it is semiconducting, in agreement with previous calculations [4–10]. The gap of a N_a-AGNR (Δ_{N_a}) is inversely proportional to its width,

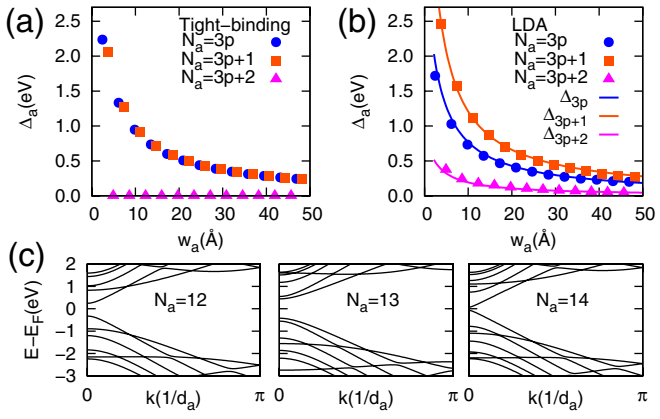

FIG. 2 (color online). The variation of band gaps of N_a-AGNRs as a function of width (w_a) obtained (a) from TB calculations with $t = 2.70$ (eV) and (b) from first-principles calculations (symbols). The solid lines in (b) are from Eq. (1). (c) First-principles band structures of N_a-AGNRs with $N_a = 12$, 13, and 14, respectively.

separated into basically two groups with a hierarchy of gap size given by $\Delta_{3p} \gtrsim \Delta_{3p+1} > \Delta_{3p+2}(= 0)$ for all p's. For the first-principles calculations, however, there are *no* metallic nanoribbons. The gaps as a function of ribbon width are now well separated into three different categories (or family structures) as shown in Fig. 2(b). Moreover, the gap size hierarchy is also changed to $\Delta_{3p+1} > \Delta_{3p} > \Delta_{3p+2}(\neq 0)$. For example, in the first-principles calculation for $p = 13$, the lowest energy gap is $\Delta_{38} = 45$ meV and $\Delta_{40} - \Delta_{39} = 68$ meV, all of which are quite larger values compared to those (0 and -2 meV, respectively) obtained from TB approximations. The first-principles band structures of N_a-AGNRs are shown in Fig. 2(c) for $N_a = 12$, 13, and 14. They exhibit direct band gaps at $kd_a = 0$ for all cases.

A determining factor in the semiconducting behavior of N_a-AGNR is quantum confinement which can be characterized by $\Delta_{N_a} \sim w_a^{-1}$ [4–10]. In addition, as will be discussed below, the edge effects play a crucial role and force the $(3p + 2)$-AGNRs (predicted to be metallic by TB model) to be semiconductors. The edge carbon atoms of our AGNRs are passivated by hydrogen atoms (by some foreign atoms or molecules in general) so that the σ bonds between hydrogen and carbon and the on-site energies of the carbons at the edges would be different from those in the middle of the AGNRs. The bonding distances between carbon atoms at the edges are also expected to change accordingly. Such effects have been observed in large aromatic molecules, e.g., ovalene ($C_{32}H_{14}$) [26]. In Fig. 3(a), we show that the bond lengths parallel to dimer lines at edges [a_1 and a_{N_a} for N_a-AGNR in Fig. 1(a)] are shortened by 3.3 ∼ 3.5% for the 12-, 13-, and 14-AGNR as compared to those in the middle of the ribbon. From the analytic expressions for TB matrix elements between carbon atoms in Ref. [27], a 3.5% decrease in interatomic distance from 1.422 Å would induce a 12% increase in the hopping integral between π-orbitals.

To see the consequence of such effects more clearly, we introduce a lattice model [Fig. 3(b)] which is equivalent to

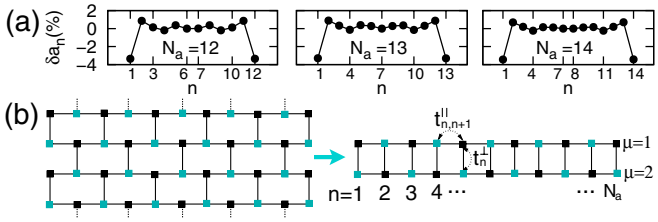

FIG. 3 (color online). (a) The ratio of the calculated change in the carbon-carbon distance (δa_n) [see Fig. 1(a)] to the carbon-carbon distance in the middle of the N_a-AGNRs, i.e., $\delta a_n \equiv 100 \times \frac{a_n - a_c}{a_c}$, where $a_c = a_6 = a_7 = 1.424$ Å for $N_a = 12$, $a_c = a_7 = 1.422$ Å for $N_a = 13$, and $a_c = a_7 = a_8 = 1.423$ Å for $N_a = 14$, respectively. (b) Topologically equivalent structure to the N_a-AGNR shown in Fig. 1(a). For the special case of $k = 0$, a lattice with periodic ladders (left) can be folded into a two-leg ladder with N_a rungs (right).

the AGNRs within the TB approximation [4–7]. The set of eigenvalues of a brick type lattice shown in Fig. 3(b) at $kd_a = 0$ is further equivalent to that of a two-leg ladder system with N_a rungs [4–7]. The Hamiltonian of this simpler model can be written as $\mathcal{H} = \sum_{n=1}^{N_a} \sum_{\mu=1}^{2} \varepsilon_{\mu,n} a^{\dagger}_{\mu,n} a_{\mu,n} - \sum_{n=1}^{N_a} t^{\perp}_{n} (a^{\dagger}_{1,n} a_{2,n} + \text{H.c.}) - \sum_{n=1}^{N_a-1} \sum_{\mu=1}^{2} t^{\parallel}_{n,n+1} (a^{\dagger}_{\mu,n} a_{\mu,n+1} + \text{H.c.})$, where $\{n, \mu\}$ denote a site, $\varepsilon_{\mu,n}$ site energies, $t^{\parallel}_{n,n+1}$ and $t^{\perp}_{n}$ the nearest neighbor hopping integrals within each leg and between the legs, respectively, and $a_{\mu,n}$ the annihilation operator of π-electrons on the nth site of the μth leg. As discussed above and shown in Fig. 3(a), $t^{\perp}_{n}$ and $\varepsilon_{\mu,n}$ at the edges would differ from those in the middle of GNRs. Hence, considering the simplest but essential variation from the exact solvable model to approximate the realistic situations, we assume that $t^{\perp}_{1} = t^{\perp}_{N_a} \equiv (1 + \delta)t$ and $t^{\perp}_{n} \equiv t$ for $n = 2, \ldots, N_a - 1$, and $t^{\parallel}_{n,n+1} \equiv t$ for all n's. The site energies are set at $\varepsilon_{\mu,n} \equiv \varepsilon_0$ for $n = 1$ and N_a and 0 otherwise regardless of μ. This model Hamiltonian is solved pertubatively and the resulting energy gaps to the first order in δ and ε_0 are as follows,

$$\begin{aligned} \Delta_{3p} &\simeq \Delta^{0}_{3p} - \frac{8\delta t}{3p+1} \sin^2 \frac{p\pi}{3p+1}, \\ \Delta_{3p+1} &\simeq \Delta^{0}_{3p+1} + \frac{8\delta t}{3p+2} \sin^2 \frac{(p+1)\pi}{3p+2}, \\ \Delta_{3p+2} &\simeq \Delta^{0}_{3p+2} + \frac{2|\delta|t}{p+1}, \end{aligned} \tag{1}$$

where Δ^{0}_{3p}, Δ^{0}_{3p+1}, and Δ^{0}_{3p+2} are the gaps of the ideally terminated ribbon when $\delta = \varepsilon_0 = 0$. They are given by $t[4\cos\frac{p\pi}{3p+1} - 2]$, $t[2 - 4\cos\frac{(p+1)\pi}{3p+2}]$ and 0, respectively. The zeroth-order gaps are identical to the values obtained from numerical calculations in Fig. 2(a) [4–7]. With $t = 2.7$ (eV) [22] and $\delta = 0.12$, the calculated gaps obtained using Eq. (1) are in good agreement with our LDA results [Fig. 2(b)]. This implies that the 12% increase of the hopping integrals between carbon atoms at the edges opens the gaps of the $(3p + 2)$-AGNRs and decreases (increases) the gaps of $3p$-AGNRs [$(3p + 1)$-AGNRs]. This analysis provides the physical explanation of the changes in the gap hierarchy discussed before. We note that there is no contribution from the variation in the site energies (ε_0) at the edges to first order.

Next, we find that nanoribbons with zigzag shaped edges also have direct band gaps which decrease with increasing width (w_z). The eigenstates of the ZGNRs near E_F, without considering spins, have a peculiar edge-state structure. As noted earlier within the tight-binding picture [4], there are two edge states decaying into the center of the ZGNR with a decay profile depending on their momentum as $\sim e^{-\alpha_k r}$, where $\alpha_k \equiv -\frac{2}{\sqrt{3}d_z} \ln|2\cos\frac{kd_z}{2}|$ [$\frac{2\pi}{3} \le kd_z \le \pi$, d_z = unit cell length shown in Fig. 1(b)]. Our first-principles calculation also predicts a set of doubly degenerate flat edge-state bands at E_F when not considering spins (not shown here). Since the edge-states around E_F form flat bands, they give rise to a very large density of states at E_F. Thus, infinitesimally small on-site repulsions could make the ZGNRs magnetic [4], unlike the case with two-dimensional graphene which has a zero density of states at E_F. As pointed out in a TB study earlier [4] and later confirmed by first-principles studies [11–13], our LSDA calculation also shows that the ground state of ZGNRs with hydrogen passivated zigzag edges indeed have finite magnetic moments on each edge with negligible change in atomic structure [11–14].

Upon inclusion of the spin degrees of freedom within LSDA, the ZGNR are predicted to have a magnetic insulating ground state with ferromagnetic ordering at each zigzag edge and antiparallel spin orientation between the two edges. The spatial spin distributions of the ground state in the case of 12-ZGNR is displayed in Fig. 4(a). The small spin-orbit coupling [28] in carbon atoms is neglected in the present study, and we label one spin orientation as α-spin (red) and the opposite as β-spin (blue) in Fig. 4(a). The total energy difference per edge atom between nonspin-polarized and spin-polarized edge states increases from 20 meV ($N_z = 8$) to 24 meV ($N_z = 16$). These energy differences are further stabilized by an antiferromagnetic coupling between the two edges. The total energy difference between ferromagnetic and antiferromagnetic couplings between edges, however, decreases as N_z increases and eventually becomes negligible if the width is significantly larger than the decay length of the spin-polarized edge states [12]. The ferromagnetic-antiferromagnetic energy differences per unit cell are 4.0, 1.8, and 0.4 meV for the 8-, 16-, and 32-ZGNR, respectively. Our LSDA results agree with previous studies [4,11,12] and are consistent with a theorem based on the Hubbard Hamiltonian on a bipartite lattice [29]. Though infinite range spontaneous

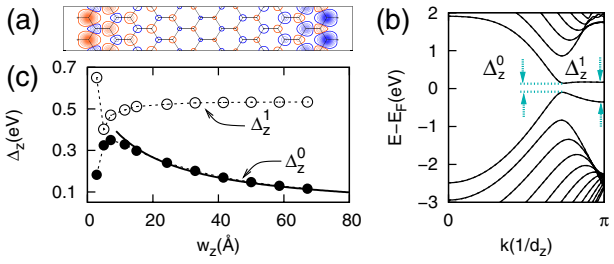

FIG. 4 (color online). (a) Contour graph for $\rho_\alpha(r) - \rho_\beta(r)$ of a 12-ZGNR (the density is integrated over the normal direction to the ribbon plane). The lowest (highest) contour of $\pm 0.4 \times 10^{-4} a_0^{-2}$ is drawn by a thick blue (red) line and the spacing for blue (red) lines is $1.0 \times 10^{-4} a_0^{-2}$ (a_0 = Bohr radius). (b) The band structure of a 12-ZGNR. The α- and β-spin states are degenerate in all energy bands. Δ^0_z and Δ^1_z denote the direct band gap and the energy splitting at $kd_z = \pi$, respectively. (c) The variation of Δ^0_z and Δ^1_z as function of the width (w_z) of N_z-ZGNRs. The solid line is a fit curve for Δ^0_z ($N_z \ge 8$) and the dotted lines are drawn to guide the eyes.

magnetic ordering in a one-dimensional Heisenberg model is ruled out [30], spin orderings in nanoscale system are realizable in practice [31–34] at finite temperature assisted by the enhanced anisotropy on substrates [33,34].

The energy gaps in ZGNRs originate from the staggered sublattice potentials resulting from the magnetic ordering, which introduce band gaps for electrons on a honeycomb hexagonal lattice [35]. This is realized because the opposite spin states on opposite edges occupy different sublattices, respectively, [black rectangles on the left side and gray ones on the right belong to different sublattice, respectively, in Fig. 1(b)]. So, the ZGNRs can be considered as the magnetic analog of a single BN sheet because the former has a band gap which originates from the exchange potential difference on the two sublattices while the band gap of the latter is from the ionic potential difference between boron and nitrogen atoms located on the different sublattices [36]. The Hamiltonian ($\mathcal{H}$) and Bloch wave functions ($\psi_{nk\alpha(\beta)}$) of the ground states satisfy $[\mathcal{H}, \mathcal{TM}] = 0$ and $\mathcal{TM}\psi_{nk\alpha} = \psi_{nk\beta}$, where $\mathcal{T}$ is the time-reversal symmetry operator and $\mathcal{M}$ a mirror symmetry operator interchanging sites on opposite sides. Hence, α- and β-spin states are degenerate in all bands and have the same gap as shown in Fig. 4(b).

Since the strength of the staggered potentials in the middle of the ribbon decreases as the width increases, the LSDA band gaps of ZGNRs are inversely proportional to the width. The calculated energy gaps (in eV) can be fit by $\Delta_z^0(w_z) = 9.33/(w_z + 15.0)$ with w_z in angstroms when $N_z \geq 8$ as shown in Fig. 4(c). It is also shown that the energy splitting at $ka = \pi$ [Δ_z^1 in Fig. 4(b)] converges to 0.53 eV (32-ZGNR) from 0.51 eV (8-ZGNR).

In summary, we have shown that graphene nanoribbons with homogeneous armchair or zigzag shaped edges all have energy gaps which decrease as the widths of the system increase [37]. The role of the edges are crucial for determining the values and scaling rule for the band gaps.

Y.-W. S. thanks F. Giustino, D. Prendergast, and Ç. Girit for discussions. This research was supported by NSF Grant No. DMR04-39768 and by the Director, Office of Science, Office of Basic Energy under Contract No. DE-AC02-05CH11231. Computational resources have been provided by NSF at NPACI and DOE at NERSC.

*Electronic address: sglouie@berkeley.edu

[1] G. Dresselhaus, M. S. Dresselhaus, and P. C. Eklund, *Science of Fullerenes and Carbon Nanotubes: Their Properties and Applications* (Academic, New York, 1996).
[2] L. Chico *et al.*, Phys. Rev. Lett. **76**, 971 (1996).
[3] P. L. McEuen, M. S. Fuhrer, and H. Park, IEEE Trans. Nanotechnol. **1**, 78 (2002).
[4] M. Fujita, K. Wakabayashi, K. Nakada, and K. Kusakabe, J. Phys. Soc. Jpn. **65**, 1920 (1996).
[5] K. Nakada, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, Phys. Rev. B **54**, 17 954 (1996).
[6] K. Wakabayashi, M. Fujita, H. Ajiki, and M. Sigrist, Phys. Rev. B **59**, 8271 (1999).
[7] M. Ezawa, Phys. Rev. B **73**, 045432 (2006).
[8] L. Brey and H. A. Fertig, Phys. Rev. B **73**, 235411 (2006).
[9] K.-I. Sasaki, S. Murakami, and R. Saito, J. Phys. Soc. Jpn. **75**, 074713 (2006).
[10] D. A. Abanin, P. A. Lee, and L. S. Levitov, Phys. Rev. Lett. **96**, 176803 (2006).
[11] S. Okada and A. Oshiyama, Phys. Rev. Lett. **87**, 146803 (2001).
[12] H. Lee *et al.*, Phys. Rev. B **72**, 174431 (2005).
[13] Y.-W. Son, M. L. Cohen, and S. G. Louie, Nature (London) **444**, 347 (2006).
[14] Y. Miyamoto, K. Nakada, and M. Fujita, Phys. Rev. B **59**, 9858 (1999).
[15] T. Kawai, Y. Miyamoto, O. Sugino, and Y. Koga, Phys. Rev. B **62**, R16 349 (2000).
[16] H. Hiura, Appl. Surf. Sci. **222**, 374 (2004).
[17] K. S. Novoselov *et al.*, Nature (London) **438**, 197 (2005).
[18] Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, Nature (London) **438**, 201 (2005).
[19] C. Berger *et al.*, Science **312**, 1191 (2006).
[20] C. Berger *et al.*, J. Phys. Chem. B **108**, 19 912 (2004).
[21] M. L. Sadowski *et al.*, cond-mat/0605739.
[22] S. Reich, J. Maultzsch, C. Thomsen, and P. Ordejón, Phys. Rev. B **66**, 035412 (2002).
[23] J. M. Soler *et al.*, J. Phys. Condens. Matter **14**, 2745 (2002), and references therein.
[24] We expect that the generalized gradient approximation would not change the main results.
[25] We have tested the various values for electronic temperature from 12 to 580 K but found no variation of the size of the gap.
[26] C. A. Coulson, J. Phys. Chem. **56**, 311 (1952).
[27] D. Porezag *et al.*, Phys. Rev. B **51**, 12 947 (1995).
[28] G. Dresselhaus, M. Dresselhaus, and J. G. Mavrodies, Carbon **4**, 433 (1966).
[29] E. H. Lieb, Phys. Rev. Lett. **62**, 1201 (1989).
[30] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).
[31] P. Gambardella *et al.*, Nature (London) **416**, 301 (2002).
[32] A. Delin, E. Tosatti, and R. Weht, Phys. Rev. Lett. **92**, 057201 (2004).
[33] J. Dorantes-Dávila and G. M. Pastor, Phys. Rev. Lett. **81**, 208 (1998).
[34] P. Gambardella *et al.*, Phys. Rev. Lett. **93**, 077203 (2004).
[35] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
[36] X. Blase, A. Rubio, S. G. Louie, and M. L. Cohen, Phys. Rev. B **51**, 6868 (1995).
[37] LDA Kohn-Sham gaps in general underestimate the quasiparticle band gaps of semiconductors. [See, e.g., M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986)]; We expect an overall increase in the value of the gap calculated here when considering quasiparticle corrections using the GW approximation as discussed in T. Miyake and S. Saito, Phys. Rev. B **68**, 155424 (2003); **72**, 073404 (2005).