Design Principles and Coupling Mechanisms in the 2D Quantum Well Topological Insulator HgTe/CdTe

Jun-Wei Luo and Alex Zunger
National Renewable Energy Laboratory, Golden, Colorado 80401, USA
(Received 13 July 2010; published 21 October 2010)

We present atomistic band structure calculations revealing a different mechanism than recently surmised via $\mathbf{k}\cdot\mathbf{p}$ calculations about the evolution of the topological state (TS) in HgTe/CdTe. We show that 2D interface (not 1D edge) TSs are possible. We find that the transitions from a topological insulator at critical HgTe thickness of $n = 23$ ML (62.5 Å) to a normal insulator at smaller n is due to the crossing between two *interface-localized* states: one derived from the *S*-like Γ_{6c} and one derived from the *P*-like Γ_{8v} light hole, not because of the crossing of an interface state and an extended quantum well state. These atomistic calculations suggest that a 2D TS can exist in a 2D system, even without truncating its symmetry to 1D, thus explaining the otherwise surprising similarity between the 2D dispersion curves of the TS in HgTe/CdTe with those of the TS in 3D bulk materials such as Bi_2Se_3.

DOI: 10.1103/PhysRevLett.105.176805 PACS numbers: 73.43.−f, 72.25.Hg, 73.20.−r, 85.75.−d

The quantum spin-Hall (QSH) effect occurring in topological insulators consists of the counterpropagation of opposite spins in spatially distinct channels in the absence of a magnetic field [1], thus allowing for dissipationless transport. This is enabled by a special band structure effect, the search of which [1–7] can be summarized by the following steps: (i) Find a primary compound that has a large spin-orbit coupling, leading to spin-orbit induced band inversion (SOBI). In zinc-blende symmetry, e.g., HgTe [2,3,8] band inversion occurs when the *S* orbital Γ_{6c} conduction band lies below the *P* orbital Γ_{8v} valence band; since the Fermi energy crosses the twofold degenerate Γ_{8v} band, such a bulk system has a zero band gap. Other known examples of SOBI are Bi_2Se_3 [5], Bi_2Te_3 [5], Sb_2Te_3 [5], α-Sn [3], and some Heusler compounds such as LuPtSb and YPtSb proposed recently [6,7]. If the primary material has zero gap (e.g., HgTe), then in the second step one needs to (ii) find a way to open a band gap in this framework system, e.g., by using quantum confinement [2,9–11], or strain [6,7,12], while maintaining the inverted order of bands (''topologically nontrivial''). The question that arises here is, what is the nature [light hole (LH); heavy hole (HH); electron (*E*)] of the extended valence band maximum (VBM) and conduction band minimum (CBM) states of the HgTe/CdTe quantum well (QW) that form the framework band gap? Finally, (iii) within this framework band gap, one or more pairs of linear-dispersed (''massless'') ''topological states'' (TSs) crossing each other (''gapless'') will emerge upon truncating the framework material (form an interface between band-inverted material and normal insulator). It is expected [1,2,8] that an N-dimensional TS will exist at the boundary of an $(N + 1)$-dimensional framework material. Thus, a band-inverted 3D material (e.g., bulk Bi_2Se_3) will give rise to a 2D TS if the former is layered with a normal-band material (including vacuum) [1,5,13]. Likewise, it was argued that a 2D band-inverted material (e.g, HgTe/CdTe QW) will give rise to a 1D TS if the former is physically truncated in-plane, reducing its 2D symmetry to 1D [2,8]. In this Letter we demonstrate among others that a 2D TS can exist in a band-inverted 2D HgTe/CdTe QW without truncating its in-plane symmetry. Thus, the 3D Bi_2Se_3 and 2D HgTe/CdTe have generically similar dispersion curves for the 2D TS.

Since the field-theoretic development of the special properties of the QSH effect is predicated on rather specific band structure properties [items (i)–(iii) above], a proper band-theoretic description of the latter is needed. Recent theoretical studies of the formation of the band gap in the framework system (ii) and the establishment of the TS within that gap (iii) in the paradigm SOBI material HgTe have used $\mathbf{k}\cdot\mathbf{p}$ approaches [1,2,8,14,15]. In this approach it is necessary to establish at the outset which energy bands in the primary 3D bulk material will couple most prominently in forming the extended 2D VBM and CBM [item (ii) above] and the TS within the framework gap [item (iii) above]. Given the difficulties of a small basis-set $\mathbf{k}\cdot\mathbf{p}$ to properly recognize the atomistic symmetry of interfaces (''farsightedness of $\mathbf{k}\cdot\mathbf{p}$'' [16,17]), the proper rendering of the electronic structure can be challenging. We use instead an atomistic band theory approach that treats both the primary bulk material and its ensuing 2D superstructures as independent systems in their own right. The identity of the primary states that couple most prominently emerges as ''output,'' not as an input hypothesis. We find a different picture for (ii) and (iii) than in the recent $\mathbf{k}\cdot\mathbf{p}$ literature [2,8,14,15]. These differences are important because they suggest a 2D TS can exist in a 2D system, without truncating its symmetry to 1D, thus explaining the otherwise surprising similarity with the dispersion curves of 2D TSs in 3D bulk materials.

The elements of the electronic structure of band-inverted bulk HgTe.—Whereas ordinary binary cubic zinc-blende

semiconductors such as GaAs or CdTe have a finite band gap separating the S-type Γ_{6c} conduction band from the P-type Γ_{8v} valence band below it (see central part of Fig. 1), the larger spin-orbit splitting (~ 1 eV) of the heavier HgTe material has long been known [12,18] to have the Γ_{8v} band move up in energy above the Γ_{6c} band (by about 300 meV). However, this spin-orbit induced S-P inversion in HgTe does not persist throughout the Brillouin zone, but is restricted to a region around the zone center [6]. This is shown in Fig. 2 which highlights the P-orbital character by green (medium gray) and the S-orbital character by red (dark gray), as obtained by projecting the pseudopotential band structure wave functions on a spherical harmonics basis. We first inquire, what will the levels of a 2D HgTe film of thickness L be if the barrier around it was infinitely high? According to the "Truncated Crystal Approximation" [19] the 2D levels of a (001) oriented film are approximately the bulk bands $E_n(\mathbf{k})$ evaluated at the discretized wave vectors $k = i(\pi/L)$, with $i = 1, 2, \ldots$. The left-hand side of Fig. 2 shows that a *thick* film will have a band gap between P-type LH1 and HH1, while still maintaining the band inversion (the S-type $E1$ below the P-type LH1). The right-hand side of Fig. 2 shows that a *sufficiently thin* film has a gap between the now S-type LH1 and the P-type HH1, thus losing the band inversion [S-type state (LH1) is above the P-type state ($E1$)]. Thus, analysis of the orbital makeup of the bands of bulk HgTe away from the zone center suggests that the quantization of bulk LH, HH, and E bands in a 2D system leads to opening of a gap primarily due to the coupling between the LH1 and HH1 subbands and to a transition from band inversion to normal-band order due to the coupling between LH1 and $E1$. We next examine this qualitative predication quantitatively.

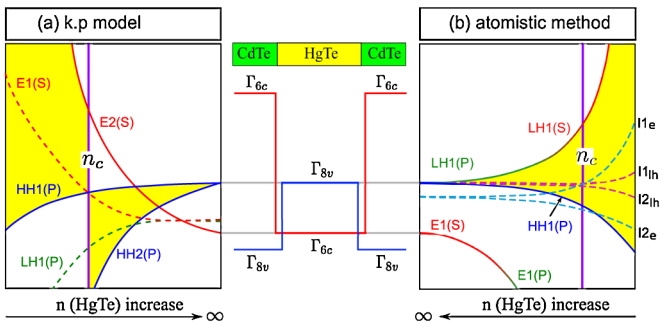

FIG. 1 (color online). Schematic description of the different dependence of the extended band edge states (solid lines) and localized interface states (dashed lines) in $(HgTe)_n/(CdTe)_\infty$ QWs on the HgTe layer thickness n according to (a) continuum $k \cdot p$ method following Refs. [2,8] and (b) atomistic calculation (Notations defined in text). The yellow (light gray) areas denote the confinement-induced framework band gap opening in the QW. The gradual color changes in (b) from red to green for $E1$ and green to red for LH1 represent the changes in orbital character. The central part of the figure denotes the band alignment of the QW at the asymptotic ($n \to \infty$) limit where $E1$ reaches Γ_{6c} and LH1, HH1 reach Γ_{8v}. The solid vertical line denotes the demarcation between topological-trivial and topological-nontrivial cases.

Formation of framework VBM and CBM in the $(HgTe)_n/(CdTe)_\infty$ *QW as a function of HgTe thickness* n.—We use the atomistic plane-wave pseudopotential approach [16] to calculate the band structure of (001) oriented $(HgTe)_n/(CdTe)_m$ superlattices. The crystal potential $V(\mathbf{r}) = \sum_{n,\alpha} \hat{v}_\alpha(\mathbf{r} - \mathbf{R}_{n,\alpha})$ is a superposition of screened atomic potentials $\hat{v}_\alpha$ of atom type α located at atomic site $\mathbf{R}_{n,\alpha}$ which contains a local part v_α^L and a nonlocal spin-orbit interaction part $\hat{v}_\alpha^{NL}$. The pseudopotentials $\hat{v}_\alpha$ are fitted [20] to experimental transition energies, effective masses, and deformation potentials of the bulk material. The HgTe/CdTe valence band offset is 400 meV [21] in agreement with the experimental measured value of 375 meV [22]. The results for QWs are shown first qualitatively in Fig. 1(b) compared with $\mathbf{k} \cdot \mathbf{p}$ [Fig. 1(a)]. The actual numerical results from our pseudopotential calculation and the corresponding QW wave functions are shown in Fig. 3. Regarding item (ii), recent $\mathbf{k} \cdot \mathbf{p}$ calculation [Fig. 1(a)] predicated [2,8,14,15] that the 3D bulk LH (Γ_{8v}) band generates in the HgTe/CdTe QW a series of extended 2D subbands LH2, LH3, ... whose energies shift down (up in present atomistic calculation) as the HgTe QW thickness n(HgTe) is reduced. Likewise, this literature predicted that the 3D bulk Γ_{6c} bulk band forms in the QW a series of extended 2D subbands $E2$, $E3$, ... whose energies shift up (down in present atomistic calculation) as n(HgTe) is reduced. The 2D subbands HH1, HH2 ... derived from 3D bulk HH(Γ_{8v}) band also move down as n(HgTe) is reduced. This $\mathbf{k} \cdot \mathbf{p}$ calculation [2,8] predicated that the extended framework band edges of the QW [solid lines in Fig. 1(a)] are HH1 (CBM) and HH2 (VBM) when $n > n_c$, changing to $E2$ (CBM) and HH1 (VBM) as n(HgTe) is reduced.

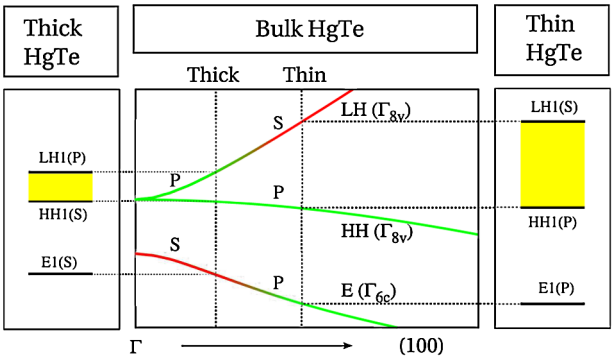

FIG. 2 (color online). Band structure of bulk zinc-blende HgTe. Near the Γ point the S-type Γ_{6c} band is below the P-type Γ_{8v} band, thus forming topological nontrivial band inversion. Away from the Γ point the upward parabolic band transitions from P-type near Γ to S-type, whereas the downward parabolic band transitions from S-type near Γ to P-type. The 2D levels of HgTe film with thickness L can be guessed from the bulk HgTe bands $E_{n,\mathbf{k}}$ at the discretized wave vectors $k = i(\pi/L)$ with ($i = 1, 2, \ldots$). For thin HgTe (right) the band inversion between S-type and P-type is lost whereas for thick HgTe (left) the inversion order is maintained. Hence, the dominant coupling forming a TS in a QW is between the P-type LH1 and the S-type E1.

In contrast, the present atomistic calculation finds that the extended framework band edges of the QW [solid lines in Fig. 1(b)] are always LH1 (CBM) and HH1 (VBM) moving monotonically up and down, respectively, as n(HgTe) is reduced. The band gap is never between HH1 and HH2. Figures 3(b) and 3(c) show the wave function of LH1 and HH1 clarifying that they are extended throughout the HgTe layer. We see [also in Fig. 1(b)] that as n(HgTe) is reduced from its asymptotic limit in infinite thick HgTe: (a) the S-type $E1$ subband moves down in energy and morphs into a P-type $E1$ state, staying all the time below the LH1/HH1 subbands; (b) the band that moves up in energy is the (light mass) P-type LH1 which morphs into a S-type LH1 state; in turn, (c) the band that moves down in energy is the P-type HH1 state. The character (LH or HH) of 2D QW states is revealed by projecting them onto bulk bands by the method described in [17]. The framework band gap of QWs, that is formed between the 2D LH1 and HH1, is simply due to the quantization of bulk LH and HH bands and is very different from that envisioned from the recent $\mathbf{k}\cdot\mathbf{p}$ method [2,8,14]. In this respect, our results are similar to the earlier results of Lin-Liu and Sham [11] who used the HH(Γ_{8v}) and LH(Γ_{8v}) states of bulk HgTe and CdTe as a basis.

The establishment of topological interface states within the framework band gap.—Regarding item (iii), the $\mathbf{k}\cdot\mathbf{p}$ literature predicted a [2,14] single pair of 2D interface states [dashed lines in Fig. 1(a)] called $E1$ and LH1 with energies moving up and down, respectively, as n(HgTe) is reduced. Furthermore, physical truncation of the 2D in-plane symmetry to 1D is needed to create TSs. This TS is argued [1,2,8] to emerge for QW thickness n(HgTe) above n_c when $E1$ lies below HH1 [see Fig. 1(a)]. In contrast to $\mathbf{k}\cdot\mathbf{p}$, our atomistic calculation finds [Fig. 1(b)], without 1D truncation, two pairs of 2D interface-localized states whose crossing as a function of n(HgTe) signals the creation of TSs. Figures 3(b) and 3(c) show the wave function of these states demonstrating their interface localization. One pair ($I1_{\rm lh}$ + $I2_{\rm lh}$) is derived from the 3D LH. When n(HgTe) $> n_c$ it lies in energy above the second pair ($I1_e$ + $I2_e$) which is derived from the 3D Γ_{6c} [Fig. 1(b)]. The $I1_{\rm lh}$ + $I2_{\rm lh}$ pair appears inside the framework QW band gap, whereas the $I1_e$ + $I2_e$ pair below gap couple to the extended hole states [Fig. 3(a)] when $n \gg n_c$. As n(HgTe) decreases from infinity $I1_e$ moves up and $I2_e$ moves down. These shifts are more rapid than the corresponding upward movement of $I1_{\rm lh}$ and downward movement of $I2_{\rm lh}$. Eventually, $I1_e$ crosses $I1_{\rm lh}$ + $I2_{\rm lh}$ at a critical thickness n_c = 23 ML [see vertical dashed line in Fig. 3(a)]. The in-plane band dispersion is shown in Fig. 4(a) for a thin QW ("topologically trivial" $n < n_c$) and in Fig. 4(b) for a thick QW ("topologically nontrivial" $n > n_c$). For the thick HgTe case we find two pairs of massless (linearly dispersed) bands crossing at a Dirac point $k_\parallel = 0$ forming two Dirac cones as required by the time reversal symmetry [1]. These linearly dispersed bands (crossing at $k_\parallel = 0$) are interface-localized states. Although the pair lower in energy is HH1 at the Dirac point, they are interface-localized states away from $k_\parallel = 0$ (not shown). They exhibit a tiny gap (~ 0.5 meV) away from $k_\parallel = 0$ [Fig. 4(b)]. A very thick HgTe well must approximate two independent interfaces. Indeed our Fig. 4(b) resembles qualitatively the earlier $\mathbf{k}\cdot\mathbf{p}$ result of a single interface [23]. However, the recent $\mathbf{k}\cdot\mathbf{p}$ model [2] for two equivalent interfaces (2D QW) does not approach the limit of a single interface as depicted in the earlier $\mathbf{k}\cdot\mathbf{p}$ result [23]. For the thin HgTe case (n = 16 ML; 43.5 Å), as shown in Fig. 4(a), the Dirac cones disappear and the strong coupling between $I1_e$ and $I1_{\rm lh}$ interface bands opens a large

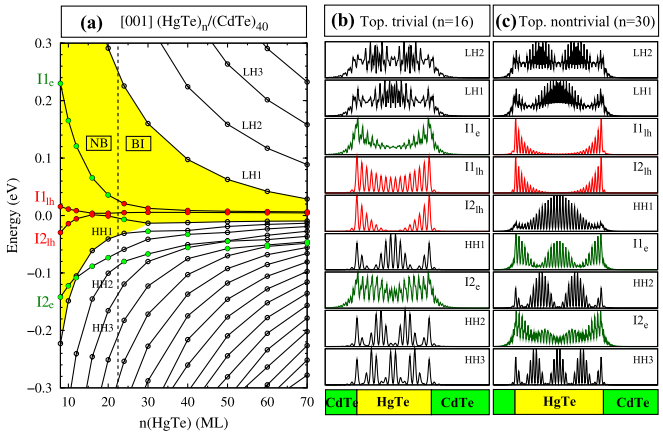

FIG. 3 (color online). (a) Presently calculated atomistic energy states $k_\parallel = 0$ as a function of HgTe layer thickness n (ML) with CdTe barrier thickness fixed to 40 ML. The yellow (light gray) shaded area indicates the framework band gap formed by intrinsic extended QW states LH1 and HH1. Interface states are denoted as $I1_e$, $I1_{\rm lh}$, $I2_{\rm lh}$, $I2_e$. The vertical dashed line denotes the demarcation between band-inverted (BI) and normal-band (NB) cases. Planer-averaged wave functions at $k_\parallel = 0$ within the energy range of interest for (b) the topological trivial case (n = 16 ML) and (c) the topological nontrivial case (n = 30 ML). The interface states ($I1_{\rm lh}$ and $I2_{\rm lh}$) are indicated in red (medium gray), whereas the interface states ($I1_e$ and $I2_e$) are in green (dark gray).

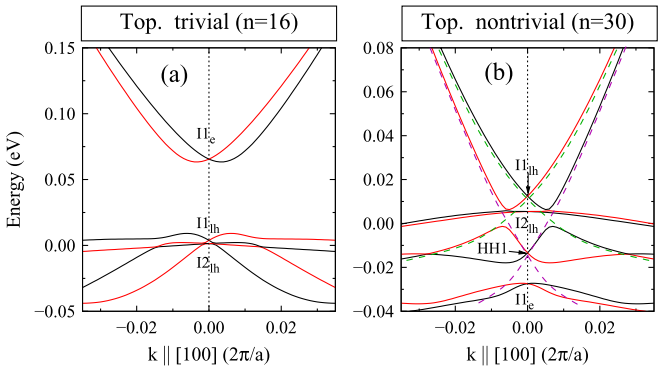

FIG. 4 (color online). The in-plane band dispersion in (a) thin (n = 16 ML) (b) thick (n = 30 ML) (HgTe)$_n$/(CdTe)$_{40}$ QWs within the energy range of interest. The red and black lines represent the two spin component of each subband, respectively. The dashed lines in (b) are guides for the eye to highlight the two pairs of linear-k dispersion bands (Dirac cones).

"QSH gap" [13] [instead of a tiny gap (~ 0.5 meV) away from $k_{\|}$ in Fig. 4(b)]. Thus, we conclude that the topological phase transition in the HgTe/CdTe QW [2,23,24] (represented by the transition of the superpotential W in the effective supersymmetry Hamiltonian [23,24], or of the gap parameter M in the effective Dirac Hamiltonian [2]) from a topological insulator to a normal insulator is due to the crossing between interface states $I1_e(\Gamma_{6c})$ and $I2_{\mathrm{lh}}(\mathrm{LH})$ and not because of the crossing of the extended QW HH1 and interface state $E1$ as previously thought [1,2,8,25]. This distinction is important because it changes the design principles (no 1D truncation is needed to create TSs in QWs).

The in-plane dispersion of the interface bands of 2D HgTe/CdTe QWs in both topological trivial [Fig. 4(a); $n < n_c$] and topological nontrivial [Fig. 4(b); $n > n_c$] greatly resembles the recent measured dispersion for a very different system: Bi_2Se_3 thin films (Figs. 2(c) and 2(d) in Ref. [13]). This resemblance despite the striking differences in the underlying 3D symmetries (T_d for bulk HgTe and D_{3d} for bulk Bi_2Se_3) suggests that there is a universal picture of interface band dispersions for all 3D framework band inversion materials in their 2D secondary systems. The only distinguishing feature is the critical thickness at which the topological transition occurs. This is a material-dependent quantity. The calculated critical thickness of HgTe layer is of $n_c = 23$ ML (62.5 Å) in agreement with the experimentally measured value of 63 Å [8]. It should be noted that the spin splitting of the interface bands shown in Fig. 4(a) is induced here by the Dresselhaus effect (the lack of bulk inversion symmetry [17]) rather than by the Rashba effect (the lack of structure inversion symmetry). The latter was recently suggested as an explanation for the observed large spin splitting in ultrathin 2D Bi_2Se_3 thin film [13].

Design principles for topological insulators.—(1) The extended VBM and CBM framework states of the HgTe/CdTe QW are LH1 (CBM) and HH1 (VBM), never HH1 (CBM) and HH2 (VBM). These identities should be distinguishable in IR absorption. (2) Two-dimensional topological states can exist in 2D material with band inversion, not requiring 1D truncation. (3) The topological phase transition of the HgTe/CdTe QW is due to the crossing between interface-localized 2D states $I1_e(\Gamma_{6c})$ and $I2_{\mathrm{lh}}(\mathrm{LH})$ and not because of the crossing of the framework QW HH1 and interface state $E1$ as previously thought [1,2,8,25]. (4) All 2D thin films made of framework 3D materials with band inversion share the same interface band dispersion pictures below and above the topological transition. (5) Robust QSH effect requires an odd number of Dirac cones [1,3]. Two Dirac cones in a symmetric 2D system can be spatially separated to two opposite interfaces [1] by breaking the structural inversion symmetry.

We acknowledge helpful discussions with O. A. Pankratov, J. Kübler, S. C. Zhang, X. L. Qi, and S. H. Wei. This work was funded by the U.S. Department of Energy, Office of Science, Basic Energy Science, Materials Sciences and Engineering, under Contract No. DE-AC36-08GO28308 to NREL.

[1] See review X. L. Qi and S. C. Zhang, Phys. Today **63**, No. 1, 33 (2010).
[2] B. A. Bernevig, T. L. Hughes, and S. C. Zhang, Science **314**, 1757 (2006).
[3] L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).
[4] D. Hsieh *et al.*, Nature (London) **452**, 970 (2008).
[5] H. Zhang, C. X. Liu, X. L. Qi, X. Dai, Z. Fang, and S. C. Zhang, Nature Phys. **5**, 438 (2009).
[6] H. Lin *et al.*, Nature Mater. **9**, 546 (2010).
[7] S. Chadov, X. L. Qi, J. Kübler, G. H. Fecher, C. Felser, and S. C. Zhang, Nature Mater. **9**, 541 (2010).
[8] M. König *et al.*, Science **318**, 766 (2007).
[9] M. Jaros, A. Zoryk, and D. Ninno, Phys. Rev. B **35**, 8277 (1987).
[10] Y. C. Chang, J. N. Schulman, G. Bastard, Y. Guldner, and M. Voos, Phys. Rev. B **31**, 2557 (1985).
[11] Y. R. Lin-Liu and L. J. Sham, Phys. Rev. B **32**, 5561 (1985). These authors found just one pair of interface states since they neglected the Γ_{6c} conduction band in their 4×4 Luttinger Hamiltonian.
[12] C. Y. Moon and S. H. Wei, Phys. Rev. B **74**, 045205 (2006).
[13] Y. Zhang *et al.*, Nature Phys. **6**, 584 (2010).
[14] E. G. Novik *et al.*, Phys. Rev. B **72**, 035321 (2005).
[15] C. Brüne *et al.*, Nature Phys. **6**, 448 (2010).
[16] A. Zunger, in *Quantum Theory of Real Materials*, edited by J. R. Chelikowsky and S. G. Louie (Kluwer, Boston, 1996), p. 173.
[17] J. W. Luo, A. N. Chantis, M. van Schilfgaarde, G. Bester, and A. Zunger, Phys. Rev. Lett. **104**, 066405 (2010).
[18] *Semiconductors: II-VI and I-VII Compounds, Semimagnetic Compounds*, edited by U. Rössler, Landolt-Börnstein, New Series, Group III, Vol. 41B, (Springer-Verlag, Berlin, 1999).
[19] S. B. Zhang, C. Y. Yeh, and A. Zunger, Phys. Rev. B **48**, 11 204 (1993) and references therein.
[20] J. W. Luo, G. Bester, and A. Zunger, Phys. Rev. Lett. **102**, 056405 (2009).
[21] Y. H. Li *et al.*, Appl. Phys. Lett. **94**, 212109 (2009).
[22] S. P. Kowalczyk, J. T. Cheung, E. A. Kraut, and R. W. Grant, Phys. Rev. Lett. **56**, 1605 (1986).
[23] O. A. Pankratov, S. V. Pakhomov, and B. A. Volkov, Solid State Commun. **61**, 93 (1987). These authors used a too small band offset (40 rather than 400 meV) so the results are not quantitatively close to the atomistic results.
[24] E. Witten, Nucl. Phys. B **188**, 513 (1981).
[25] Xi Dai, T. L. Hughes, X. L. Qi, Z. Fang, and S. C. Zhang, Phys. Rev. B **77**, 125319 (2008). In this tight-binding work the authors incorrectly identified the band-edge states to be the same as those found in the $\mathbf{k} \cdot \mathbf{p}$ calculation of Ref. [2]. As shown in Fig. 1 here atomistic and $\mathbf{k} \cdot \mathbf{p}$ states are different.