PHYSICAL REVIEW B **81**, 041307(R) (2010)

Oscillatory crossover from two-dimensional to three-dimensional topological insulators

Chao-Xing Liu,[1,2] HaiJun Zhang,[3] Binghai Yan,[4] Xiao-Liang Qi,[5] Thomas Frauenheim,[4] Xi Dai,[3] Zhong Fang,[3] and Shou-Cheng Zhang[5]

[1]*Physikalisches Institut (EP3) and Institute for Theoretical Physics and Astrophysics, University of Würzburg, 97074 Würzburg, Germany*

[2]*Center for Advanced Study, Tsinghua University, Beijing 100084, China*

[3]*Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*

[4]*Bremen Center for Computational Materials Science, Universität Bremen, Am Fallturm 1, 28359 Bremen, Germany*

[5]*Department of Physics, McCullough Building, Stanford University, Stanford, California 94305-4045, USA*

(Received 7 December 2009; published 19 January 2010)

We investigate the crossover regime from three-dimensional topological insulators Bi_2Te_3 and Bi_2Se_3 to two-dimensional topological insulators with quantum spin Hall effect when the layer thickness is reduced. Using both analytical models and first-principles calculations, we find that the crossover occurs in an oscillatory fashion as a function of the layer thickness, alternating between topologically trivial and nontrivial two-dimensional behavior.

DOI: 10.1103/PhysRevB.81.041307 PACS number(s): 73.20.−r, 72.25.Dc, 73.50.−h

Introduction. Recent discovery of the two-dimensional (2D) and three-dimensional (3D) topological insulator (TI) state has generated great interests in this new state of topological quantum matter.[1–7] In particular, Bi_2Te_3 and Bi_2Se_3 are predicted to have bulk energy gaps as large as 0.3 *eV*, and gapless surface states consisting of a single Dirac cone.[5] Angle-resolved-photo-emission spectroscopy on both of these materials observed the single Dirac cone linearly dispersing from the Γ point.[6,7] These materials have a layered structure consisting of stacked quintuple layers (QL), with relatively weak coupling between the QLs. Therefore, it should be relatively easy to prepare these materials in the form of thin films, either by nanoribbon growth method,[8] or by molecular beam epitaxy.[9] In the limit when the thickness d of the thin film are much smaller than the lateral dimensions of the device, it is natural to ask whether or not the resulting 2D system is a 2D TI similar to the HgTe quantum wells.[1,2] In this Rapid Communication, we investigate this question, and find a surprising result that the crossover from the 3D to the 2D TI occurs in a oscillatory fashion as a function of the layer thickness d.

Effective model analysis. We begin by recalling the four-band effective model of the 3D TI introduced by Zhang *et al.*[5] with the following Hamiltonian:

$$H_{3D}(\mathbf{k})=\begin{pmatrix} \mathcal{M}(\mathbf{k}) & A_1k_z & 0 & A_2k_- \\ A_1k_z & -\mathcal{M}(\mathbf{k}) & A_2k_- & 0 \\ 0 & A_2k_+ & \mathcal{M}(\mathbf{k}) & -A_1k_z \\ A_2k_+ & 0 & -A_1k_z & -\mathcal{M}(\mathbf{k}) \end{pmatrix}+\epsilon_0(\mathbf{k}), \tag{1}$$

with $k_\pm=k_x\pm ik_y$, $\epsilon_0(\mathbf{k})=C+D_1k_z^2+D_2k_\perp^2$, and $\mathcal{M}(\mathbf{k})=M_0+B_1k_z^2+B_2k_\perp^2$. This Hamiltonian has been successfully used to discuss the property of V-VI semiconductor, such as Bi_2Se_3 and Bi_2Te_3.[5] The four basis of the above effective Hamiltonian are denoted as $|P1_z^+,\uparrow\rangle,|P2_z^-,\uparrow\rangle,|P1_z^+,\downarrow\rangle,|P2_z^-,\downarrow\rangle$ with the superscript $\pm$ standing for even and odd parity and $\uparrow(\downarrow)$ for spin up (down). An important feature is that the two orbitals $P1_z^+$ and $P2_z^-$ have the opposite parities, so that the off-diagonal terms are linear in k_z and $k_\pm$. Another important model is the four-band effective model proposed by Bernevig, Hughes, and Zhang (BHZ)[1] for 2D quantum spin Hall (QSH) insulator, given by the effective Hamiltonian,

$$H_{2D}(\mathbf{k})=\begin{pmatrix} \tilde{\mathcal{M}}(\mathbf{k}) & 0 & 0 & \tilde{A}k_- \\ 0 & -\tilde{\mathcal{M}}(\mathbf{k}) & \tilde{A}k_- & 0 \\ 0 & \tilde{A}k_+ & \tilde{\mathcal{M}}(\mathbf{k}) & 0 \\ \tilde{A}k_+ & 0 & 0 & -\tilde{\mathcal{M}}(\mathbf{k}) \end{pmatrix}+\tilde{\epsilon}_0(\mathbf{k}), \tag{2}$$

with $\tilde{\epsilon}_0(\mathbf{k})=\tilde{C}+\tilde{D}_2k_\perp^2$ and $\tilde{\mathcal{M}}(\mathbf{k})=\tilde{M}_0+\tilde{B}_2k_\perp^2$. The four basis for the HgTe system are taken as $|E1,\frac{1}{2}\rangle,|H1,-\frac{3}{2}\rangle,|E1,-\frac{1}{2}\rangle,|H1,\frac{3}{2}\rangle$. According to Ref. 1, the orbitals $E1$ and $H1$ also have the opposite parities, similar to the 3D TI model. The finite size effect has been studied previously by B. Zhou *et al.* within the BHZ model.[10] Another similarity between these two models is that in order to describe 2D or 3D TI, we need the condition $M_0B_{1,2}<0$ or $\tilde{M}_0\tilde{B}_2<0$, so that the system is in the inverted regime.[1] In fact if we simply take $k_z=0$, the A_1, B_1 and D_1 terms vanish and 3D TI model Eq. (1) will reduce exactly to the BHZ model for the 2D TI Eq. (2). These similarities suggest that it is possible to give a unified description of both the 2D and 3D TI, and it is helpful to investigate the crossover between them when the dimension is reduced by quantum confinement. Therefore, in the following, we would like to consider the 3D TI model Eq. (1) in a quantum well or a thin film configuration. For simplicity, here we assume the film or well thickness to be d and use an infinite barrier to represent the vacuum.

To establish the connection between 2D BHZ model and 3D TI model, we start from the special case of $A_1=0$ and turn on A_1 later. For $A_1=0$ the eigenvalue problem of the infinite quantum well can be easily solved at Γ point ($k_x=k_y=0$). The eigenstate are simply given by $|E_n(H_n)\rangle=\sqrt{\frac{2}{d}}\sin(\frac{n\pi z}{d}+\frac{n\pi}{2})|\Lambda\rangle$, with $|\Lambda\rangle=|P1_z^+,\uparrow(\downarrow)\rangle$ for

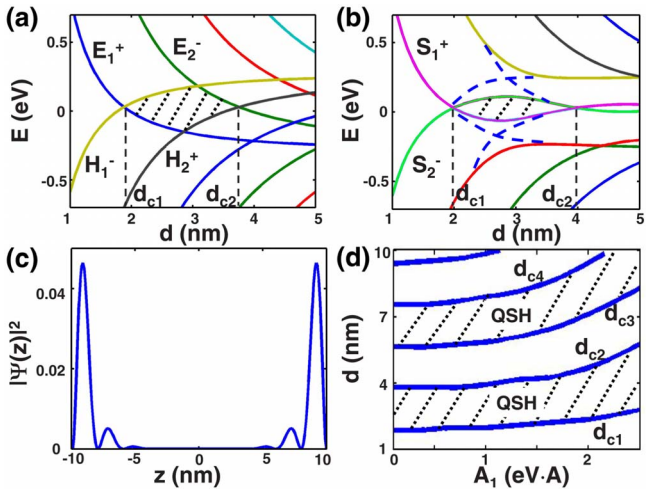


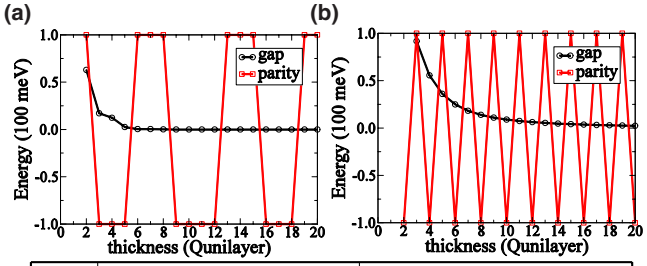
FIG. 1. (Color online) The energy level versus the thickness of the quantum well is shown for (a) $A_1=0$ eV·Å, (b) $A_1=1.1$ eV·Å. Other parameters are taken from Ref. 5. The shaded region indicates the regime for QSH states. The blue dashed line in (b) shows how the crossing between $|E_1(H_1)\rangle$ and $|H_2(E_2)\rangle$ is changed to anticrossing when A_1 is nonzero. In (c), the density of $|S_1^+\rangle$ ($|S_2^-\rangle$ has the same density) is plotted for $A_1=1.1$ eV·Å. In (d), the critical thickness $d_{cn}(n=1,2,\cdots)$ is plotted as a function of A_1. QSH states appear in the shaded region.

electron sub-bands and $|\Lambda\rangle=|P2_z^-,\uparrow(\downarrow)\rangle$ for hole sub-bands, and the corresponding energy spectrum are $E_e(n)=C+M_0+(D_1+B_1)(\frac{n\pi}{d})^2$ and $E_h(n)=C-M_0+(D_1-B_1)(\frac{n\pi}{d})^2$, respectively. Here, we assume $M_0<0$ and $B_1>0$ so that the system stays in the inverted regime.[1] The energy spectrum is shown in Fig. 1(a). When the width d is small enough, electron sub-bands E_n have higher energy than the hole sub-bands H_n due to the quantum confinement effect. Because the bulk band is inverted at Γ point ($M_0<0$), with increasing d the energy of the electron sub-bands will decrease toward their bulk value $M_0<0$, while that of the hole sub-bands will increase toward $-M_0>0$. Thus, there must be crossing points between the electron and hole sub-bands.

Now let us focus on the crossing points between the nth electron sub-band $|E_n\rangle$ and the nth hole sub-band $|H_n\rangle$, which occurs at the critical thickness $d_{c1},d_{c2},\cdots$ in Fig. 1(a). Near one of these critical values of d, the low-energy physics can be obtained by projecting the 3D TI model Eq. (1) into the basis $|E_n,\uparrow(\downarrow)\rangle$ and $|H_n,\uparrow(\downarrow)\rangle$. The resulting effective Hamiltonian is nothing but BHZ model Eq. (2) with $\tilde{C}=C+D_1(n\pi/d)^2$, $\tilde{M}_0=M_0+B_1(n\pi/d)^2$, $\tilde{A}=A_2$, $\tilde{B}_2=B_2$ and $\tilde{D}_2=D_2$. Following the argument of BHZ,[1] we know that there is a topological phase transition between QSH state and ordinary insulator across the critical point where the gap is closed. The above result can also be obtained from the parities of the sub-bands $|E_n\rangle$ and $|H_n\rangle$ at Γ point. Since now, the wave function has four components, the parity should be determined by both the function $\sin(\frac{n\pi z}{d}+\frac{n\pi}{2})$ and the basis $|\Lambda\rangle$. Define $\hat{P}$ as the inversion operator and we find that $\hat{P}|E_n\rangle=(-1)^{n+1}|E_n\rangle$ and $\hat{P}|H_n\rangle=(-1)^n|H_n\rangle$, hence $|E_n\rangle$ and $|H_n\rangle$ always have the opposite parities. According to the parity criterion of Fu and Kane,[3] the topological property of the system can be determined by the product of the parities of all the occupied energy levels at all time-reversal invariant momenta, which is denoted as ν. Near the critical thickness d_{cn}, the band gap is determined by $|E_n\rangle$ and $|H_n\rangle$ at Γ point and all the other energy levels are far from the Fermi surface. Therefore, due to their opposite parities, the crossing of $|E_n\rangle$ and $|H_n\rangle$ will change the total parity ν and correspondingly the topological property of the system. The crossing points can be determined by the condition $E_e(n)=E_h(n)$, which leads to $d_{cn}=n\pi\sqrt{\frac{B_1}{|M_0|}}$. A topological phase transition occurs at d_{cn} for each n, so that the system oscillates between QSH phase and ordinary insulator phase with the period of $l_c=\pi\sqrt{\frac{B_1}{|M_0|}}$, if the Fermi energy always stays within the gap.

Next, we consider the effect of turning on A_1 term. As shown in Fig. 1(b), A_1 term induces the coupling between $|E_n(H_n)\rangle$ and $|H_{n\pm1}(E_{n\pm1})\rangle$, which results in an avoid crossing between these sub-bands. However, a finite A_1 does not break parity, so that the crossing at d_{cn} between $|E_n\rangle$ and $|H_n\rangle$ with opposite parity remains robust. Since the topological phase transition is mainly determined by the crossing between $|E_n\rangle$ and $|H_n\rangle$, the system still oscillates between QSH insulator and ordinary insulator when the thickness d is tuned. Due to the hybridization induced by A_1 term, two special states (each doubly degenerate due to spin), denoted as $|S_1^+\rangle$ and $|S_2^-\rangle$ in Fig. 1(b), are formed within the bulk gap and well-separated in energy from other sub-bands when the thickness d is large. $|S_1^+\rangle$ is in fact the superposition of $|E_{2n-1}\rangle$ and $|H_{2n}\rangle(n=1,2,\cdots)$ while $|S_1^-\rangle$ is the superposition of $|H_{2n-1}\rangle$ and $|E_{2n}\rangle(n=1,2,\cdots)$. In large d limit, the gap between $|S_1^+\rangle$ and $|S_2^-\rangle$ goes to zero and the two states become nearly degenerate. In Fig. 1(c), the density $|\Psi(z)|^2$ for $|S_1^+\rangle$ or $|S_2^-\rangle$ is plotted as a function of the position z for a quantum well with the width $d=20$ nm, which indicates that $|S_1^+\rangle$ and $|S_2^-\rangle$ are localized on the two surfaces of the thin film. One should recall that the bulk material is a 3D TI with surface states on each surface,[5] so that in the large d limit $|S_1^+\rangle$ and $|S_2^-\rangle$ are nothing but the bonding and the antibonding state of the two surface states at the two opposite surfaces. Since $|S_1^+\rangle$ and $|S_2^-\rangle$ have opposite parities, the topological phase transition can also be understood as the change of sequence between these two states. On top of the exponential decay, the wave function shown in Fig. 1(c) also oscillates with a period $\sim\pi\sqrt{\frac{B_1}{|M_0|}}$, which coincides with the oscillation period of the system between QSH insulator and ordinary insulator. Although A_1 term cannot eliminate the crossing between $|E_n\rangle$ and $|H_n\rangle$, it can shift the critical thickness d_{cn}. The dependence for the critical thickness on the parameter A_1 is shown in Fig. 1(d), with the QSH regime labeled by the shading regions. One can see that when A_1 is increased, all d_{cn} ($n=1,2,\cdots$) are shifted to the large values, but the oscillation length does not change too much.

Realistic materials. The above discussion based on the analytical model gives us a clear physical picture of the crossover between 2D BHZ model and 3D TI model, and in the following, we would like to consider about the possible realization in the realistic TI materials Bi_2Se_3 and Bi_2Te_3. Here, we carried out the first-principles calculations with the BSTATE package[11] and Vienna *ab initio* simulation package (VASP).[12] The generalized gradient approximation of



(c)	Bi_2Se_3		Bi_2Te_3	
	Gap at Γ (eV)	Parity ν	Gap at Γ (eV)	Parity ν
1QL	0.710	+	0.391	+
2QLs	0.098	+	0.148	-
3QLs	0.004	-	0.036	+
4QLs	0.012	-	0.005	-
5QLs	0.004	-		

FIG. 2. (Color online) The band gap and the total parity are plotted as the function of the number of the QLs for (a) Bi_2Se_3 and (b) Bi_2Te_3. The calculation is based on TB model constructed by MLWF from first-principles calculation. The results from the fully first-principles calculation are shown in table (c) for $1\sim5$ QLs. Good agreement between the TB model calculation and first-principles calculation is found.

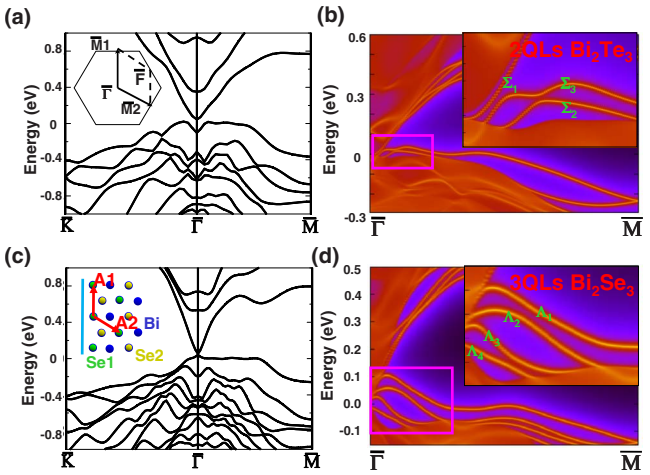


FIG. 3. (Color online) Left: the energy dispersion of 2D thin film is plotted for (a) 2QLs Bi_2Te_3 and (c) 3QLs Bi_2Se_3. Right: the energy dispersion of the semi-infinite film with the edge along $A1$ direction is plotted for (b) 2QLs Bi_2Te_3 and (d) 3QLs Bi_2Se_3. The inset of (a) shows the 2D BZ and that of (c) is the top view of 2D thin film with two in-plane lattice vectors $A1$ and $A2$. The 1D edge is indicated by the blue line along $A1$ direction. The insets of (b) and (d) are the zoom-in of the energy dispersion near the bulk gap.

Perdew-Burke-Ernzerhof (PBE)-type is used for exchange-correlation potential and the spin orbit coupling is taken into account. For 2D thin film, we construct the free-standing slab model with crystal parameters taken from the experimental data. For Bi_2Se_3 and Bi_2Te_3, five-atom layers, including two Bi layers and three Se or Te layers, are stacked along z direction, forming a QL.[5] There exists strong coupling between two atomic layers within one QL but much weaker coupling, predominantly of the van der Waals type, between two QLs. Therefore, it is natural to regard a QL as a unit for the thin film. We will first neglect the lattice relaxation effect, which depends on the detail of the lattice environment and address this problem in the end. As we have discussed above, the topological property can be determined by the total parity ν, which has been successfully utilized to predict the 3D TI, such as Bi_xSb_{1-x} (Ref. 3) and Bi_2Se_3.[5,6] Here, we apply this method to the thin film with four time-reversal invariant points in 2D Brillouin zone (BZ), namely, $\bar{\Gamma}(0,0)$, $\bar{M}1(\pi,0)$, $\bar{M}2(0,\pi)$, and $\bar{F}(\pi,\pi)$, as shown in the inset of Fig. 3(a). The parity of the wave function at these points is well-defined and can be easily calculated to determine the topological property.

The calculated band gap at Γ point and the related total parity ν for different QLs are shown in the table of Fig. 2(c). The first nontrivial QSH phase appears at 3QLs for Bi_2Se_3 and at 2QLs for Bi_2Te_3. The gap at Γ point of the nontrivial QSH phase for 2QLs Bi_2Te_3 is quite large, of about 148 meV, however as shown in Fig. 3(a), the Bi_2Te_3 is an indirect gap material, and the indirect gap is an order smaller, only ~10 meV, similar to the gap of 4QLs Bi_2Se_3. When the thickness of the film increases, the time cost for the *ab initio* calculation also increases rapidly, therefore, we perform the tight-binding (TB) model calculation based on the maximally localized Wannier function (MLWF)[13,14] from the *ab initio* calculation, which has also been successfully applied to calculate the surface states of Bi_2Se_3 type of materials.[5] The band gap and the total parity ν obtained from the TB calculation are summarized in Fig. 2(a) for Bi_2Se_3 and Fig. 2(b) for Bi_2Te_3. The results for the $1\sim5$ QLs fit with *ab initio* calculation, which confirms the validity of our method. From Figs. 2(a) and 2(b), it is clear that the system oscillates between QSH insulator and ordinary insulator, which verifies our previous analysis of oscillating crossover.

When the 2D system stays in the QSH phase, there are topologically protected helical edge states at the 1D edge.[15–17] To show the topological feature more explicitly, we calculate the dispersion spectra of the helical edge states directly. As examples, here we study the edge states of the 2QLs Bi_2Te_3 and 3QLs Bi_2Se_3 film along $A1$ direction, as shown in the inset of Fig. 3(c). For a semi-infinite system, combining the TB model from MLWF with the iterative method, we can calculate the Green's function[17] for the edge states directly. The local density of states is directly related to the imaginary part of Green's function, from which we can obtain the dispersion of the edge states. The topological nature of the edge states can be determined by the method suggested by Fu and Kane.[3] As shown in Fig. 3(b) for 2QLs Bi_2Te_3, there exist one edge state Σ_1, which stays in the valence band at $\bar{\Gamma}$ point and goes to conduction band at $\bar{M}$ point. Such edge states are in fact the helical edge states which cannot be eliminated by the local time-reversal invariant perturbation. For 3QLs Bi_2Se_3 in Fig. 3(d), there are three edge states $\Lambda_{1,2,3}$ connecting the conduction and valence band, which guarantee the system to be nontrivial. There are also other trivial edge states ($\Sigma_{2,3}$ and Λ_4) only connected to the conduction band or valence band, which do not change the topological property of the system.

In the discussion above, we have only considered the bulk parameters and have neglected the influence of the surface

lattice relaxation. Since for 2D thin film, the lattice relaxation always plays an important role in the electronic band structure, we would like to address this problem in the following. The largest nontrivial gap appears for 2QLs Bi_2Te_3 and 4QLs Bi_2Se_3, which is suitable for experiment to observe. If the thin film is relaxed in vacuum environment, the nontrivial QSH phase could be changed to ordinary insulator phase, due to the change of the distance between two QLs. The reason is that the van der Waals interaction between two QLs is very week and quite sensitive to the lattice environment. This also indicates that the lattice relaxation should depend strongly on the pressure and the substrate. Therefore, here we propose two different ways to overcome the negative influence from surface lattice relaxation. One simple way is to apply uniaxial compressive stress (along the growth direction) to the film. Based on first-principles calculation, it is found that only 0.1 GPa stress can recover the nontrivial gap, which can be easily achieved in experiment.[18] Another way is to fabricate the sample on the proper substrate. For the material with the positive Poisson ratio, when the lattice is stretched within the film plane, it tends to get thinner in the perpendicular direction. Therefore, we can use the materials with the larger in-plane lattice constant than that of Bi_2Se_3 or Bi_2Te_3 as the substrate. Since in the present paper, we focus on the principle of realizing QSH insulator from 3D TI, further study about the effect of pressure and substrate will be addressed elsewhere.

Conclusion. We have studied the crossover between 3D and 2D TI in Bi_2Se_3 and Bi_2Te_3 systems. Based on both effective model analysis and *ab initio* calculation, we found that it is possible to obtain 2D QSH state by confining 3D TI in a quantum well or thin film configuration with the proper thickness. Up to now, 2D QSH effect has only been observed in HgTe quantum wells,[2] and our work may open up a new way to search for new materials with 2D QSH effect. Recently, the high-quality samples of Bi_2Se_3 [19] and Bi_2Te_3 [9] thin films have already been fabricated in experiment, so that it is likely to observe the proposed phenomenon soon. Recently, we learned about the work of H. Z. Lu *et al.*[20] and J. Linder *et al.*,[21] in which similar finite size effect has been studied within the 3D TI model of Zhang *et al.*[5]

We would like to thank L. W. Molenkamp for the helpful discussion. This work is supported by the NSF of China, the National Basic Research Program of China (Grant No. 2007CB925000), the International Science and Technology Cooperation Program of China and by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Contract No. DE-AC02–76SF00515. C.X.L. and B.H.Y. acknowledge financial support by the Alexander von Humboldt Foundation of Germany.

[1] B. A. Bernevig, T. L. Hughes, and S. C. Zhang, Science **314**, 1757 (2006).
[2] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science **318**, 766 (2007).
[3] L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).
[4] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Nature (London) **452**, 970 (2008).
[5] H. Zhang, C. Liu, X. Qi, X. Dai, Z. Fang, and S. Zhang, Nat. Phys. **5**, 438 (2009).
[6] Y. Xia *et al.*, Nat. Phys. **5**, 398 (2009).
[7] Y. L. Chen *et al.*, Science **325**, 178 (2009).
[8] H. Peng, K. Lai, D. Kong, S. Meister, Y. Chen, X.-L. Qi, S.-C. Zhang, Z.-X. Shen, and Y. Cui, Nature Mater., doi:10.1038/nmat2609 (2009).
[9] Y.-Y. Li *et al.*, arXiv:0912.5054 (unpublished).
[10] B. Zhou, H.-Z. Lu, R.-L. Chu, S.-Q. Shen, and Q. Niu, Phys. Rev. Lett. **101**, 246807 (2008).
[11] Z. Fang and K. Terakura, J. Phys.: Condens. Matter **14**, 3001 (2002).
[12] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
[13] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).
[14] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2001).
[15] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
[16] C. Wu, B. A. Bernevig, and S. C. Zhang, Phys. Rev. Lett. **96**, 106401 (2006).
[17] X. Dai, T. L. Hughes, X.-L. Qi, Z. Fang, and S.-C. Zhang, Phys. Rev. B **77**, 125319 (2008).
[18] D. A. Polvani, J. F. Meng, N. V. C. Shekar, J. Sharp, and J. V. Badding, Chem. Mater. **13**, 2068 (2001).
[19] G. Zhang, H. Qin, J. Teng, J. Guo, Q. Guo, X. Dai, Z. Fang, and K. Wu, Appl. Phys. Lett. **95**, 053114 (2009).
[20] H. Lu, W. Shan, W. Yao, Q. Niu, and S. Shen, arXiv:0908.3120 (unpublished).
[21] J. Linder, T. Yokoyama, and A. Sudbø, Phys. Rev. B **80**, 205401 (2009).