Oscillatory crossover from two dimensional to three dimensional topological insulators

Chao-Xing Liu1,5, HaiJun Zhang2, Binghai Yan3, Xiao-Liang Qi4, Thomas Frauenheim3, Xi Dai2, Zhong Fang2 and Shou-Cheng Zhang4
1 Physikalisches Institut (EPI3) and Institute for Theoretical Physics and Astrophysics, University of Würzburg, 97074 Würzburg, Germany;
2 Beijing National Laboratory for Condensed Matter Physics, and Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China;
3 Bremen Center for Computational Materials Science, Universität Bremen, Am Fallturm 1, 28359 Bremen, Germany;
4 Department of Physics, McCullough Building, Stanford University, Stanford, CA 94305-4045; and Center for Advanced Study, Tsinghua University, Beijing, 100084, China;

(Dated: September 16, 2009)

We investigate the crossover regime from three dimensional topological insulators Bi2Te3 and Bi2Se3 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 non-trivial two dimensional behavior.

PACS numbers: 73.20.-r, 73.50.-h, 72.25.Dc

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, 2, 3, 4, 5, 6, 7]. In particular, Bi2Te3 and Bi2Se3 are predicted to have bulk energy gaps as large as 0.3eV, and gapless surface states consisting of a single Dirac cone[8]. 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 nano-ribbon growth method[5], or by molecular beam epitaxy[9]. In the limit when the thickness d of the thin film is 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 well[1, 2]. In this letter, we investigate this question, and find a surprising result that the crossover from the 3D to the 2D TI occurs in an oscillatory fashion as a function of the layer thickness d.

Effective model analysis- We begin by recalling the 4-band effective model of the 3D TI introduced by Zhang et al[5] with the following Hamiltonian:

\[ H_{3D}(k) = \begin{pmatrix} \mathcal{M}(k) & A_1 k_z & 0 & A_2 k_- \\ -A_1 k_z & -\mathcal{M}(k) & A_2 k_- & 0 \\ 0 & A_2 k_+ & \mathcal{M}(k) & -A_1 k_z \\ A_2 k_+ & -A_1 k_z & -\mathcal{M}(k) \end{pmatrix} + \epsilon_0(k) \]

with \[ k_\pm = k_x \pm i k_y \] and \[ \epsilon_0(k) = C + D_1 k_x^2 + D_2 k_y^2 \] and \[ \mathcal{M}(k) = M_0 + B_1 k_x^2 + B_2 k_y^2 \]. This Hamiltonian has been successfully used to discuss the property of V-VI semiconductor, such as Bi2Se3 and Bi2Te3[8, 9]. The four basis for the above effective Hamiltonian are denoted as \[ |P1_+^z, \uparrow\rangle, |P2_+^z, \downarrow\rangle, |P1_+^z, \downarrow\rangle, |P2_+^z, \downarrow\rangle \] 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 \[ P_1 \] and \[ P_2 \] have the opposite parities, so that the off-diagonal term are linear in \[ k_x \] and \[ k_y \]. Another important model is the 4-band effective model proposed by Bernevig, Hughes and Zhang (BHZ)[10] for 2D quantum spin Hall (QSH) insulator, given by the effective Hamiltonian

\[ H_{2D}(k) = \begin{pmatrix} \mathcal{M}(k) & 0 & 0 & \tilde{A} k_- \\ 0 & -\mathcal{M}(k) & \tilde{A} k_- & 0 \\ 0 & \tilde{A} k_+ & \mathcal{M}(k) & 0 \\ \tilde{A} k_+ & 0 & 0 & -\mathcal{M}(k) \end{pmatrix} + \epsilon_0(k) \]

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. [11], 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 [11]. Another similarity between these two models is that in order to describe 2D or 3D TI, we need the condition \[ M_0 B_{1,2} < 0 \] or \[ M_0 B_2 < 0 \], so that the system is in the inverted regime [11]. In fact if we simply take \[ k_z = 0 \], the \[ A_1 \], \[ B_1 \] and \[ D_1 \] terms vanish and 3D TI model [11] will reduce exactly to the BHZ model for the 2D TI [12]. 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 [11] 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 $|E_n(H_n)| = \sqrt{\frac{2}{\pi}} \sin \left( \frac{n_x \pi}{d} + \frac{n_z \pi}{d} \right) |\Lambda\rangle$, with $|\Lambda\rangle = |P1\uparrow, \uparrow (1)\rangle$ for electron sub-bands and $|\Lambda\rangle = |P2\uparrow, \uparrow (1)\rangle$ for hole sub-bands, and the corresponding energy spectrum are $E_e(n) = C + M_0 + (D_1 + B_1) \left( \frac{n_x}{d} \right)^2$ and $E_h(n) = C - M_0 + (D_1 - B_1) \left( \frac{n_x}{d} \right)^2$ respectively. Here we assume $M_0 < 0$ and $B_1 > 0$ so that the system stays in the inverted regime. The energy spectrum is shown in Fig.1(a). When the width $d$ is small enough, electron sub-bands $E_e$ 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 towards their bulk value $M_0 < 0$, while that of the hole sub-bands will increase towards $-M_0 > 0$. Thus there must be crossing points between the electron and hole sub-bands.

Now let us focus on the crossings points between the $n$th electron sub-band $|E_n\rangle$ and the $n$th 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 into the basis $|E_n, \uparrow (1)\rangle$ and $|H_n, \uparrow (1)\rangle$. The resulting effective Hamiltonian is nothing but BHZ model with $C = C + D_1(n\pi/d)^2$, $M_0 = M_0 + B_1(n\pi/d)^2$, $\tilde{A} = A_2$, $\tilde{B}_2 = B_2$ and $\tilde{D}_2 = D_2$. Following the argument of BHZ, 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 $\left( \frac{n_x \pi}{d} + \frac{n_z \pi}{d} \right)$ and the basis $|\Lambda\rangle$. Define $\hat{P}$ as the inversion operator and we find that $P|E_n\rangle = (-1)^{n+1}|E_n\rangle$ and $P|H_n\rangle = (-1)^{n}|H_n\rangle$, hence $|E_n\rangle$ and $|H_n\rangle$ always have the opposite parities. According to the parity criterion of Fu and Kane, 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 $\nu$. 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 $\nu$ 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 $d_{cn} = \pi \sqrt{\frac{d}{M_0}}$. 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 $l_c = \pi \sqrt{\frac{d}{M_0}}$, 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 $|H_n\pm (E_n\pm 1)\rangle$, 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 $|S_2^-\rangle$ 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, so that in the large $d$ limit $|S_1^+\rangle$ and $|S_2^-\rangle$ are nothing but the bonding and the anti-bonding 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

FIG. 1: (Color online) The energy level versus the thickness of the quantum well is shown for (a) $A_1 = 0 eV\cdot A$, (b) $A_1 = 1.4 eV\cdot A$. Other parameters are taken from Ref. [2]. 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 anti-crossing when $A_1$ is non-zero. In (c), the density of $|S_1^+\rangle$ $|S_2^-\rangle$ has the same density) is plotted for $A_1 = 1.4 eV\cdot A$. 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.
of the exponential decay, the wave function shown in Fig. 1(c) also oscillates with a period \( \sim \pi \sqrt{\frac{E}{|V_0|}} \), which coincides with the oscillation period of the system between QSH insulator and ordinary insulator. Although \( A_1 \) term cannot eliminate the crossing between \( |E_n| \) and \( |H_n| \), it can shift the critical thickness \( d_\text{cr} \). 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 \( d_\text{cr} \) \((n = 1, 2, \cdots) \) are shifted to the large values, but the oscillation length doesn’t change too much. This indicates that the QSH regime will always exist, although the gap will be reduced when \( A_1 \) is large.

**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 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 \([2] \). 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 \( \nu \), which has been successfully utilized to predict the 3D TI such as \( Bi_2Se_3 \) and \( Bi_2Te_3 \). Here we apply this method to the thin film with four time-reversal invariant points in 2D Brillouin zone (BZ), namely \( \Gamma(0,0), M(1/2,0), M(1,0,0), \bar{F}(\pi, \pi) \), as shown in the inset of Fig.2(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 \( \Gamma \) point and the related total parity for different QLS are shown in the table of Fig.2(c). The first non-trivial QSH phase appears at 3QLs for \( Bi_2Se_3 \) and at 2QLs for \( Bi_2Te_3 \). The gap at \( \Gamma \) point of the non-trivial QSH phase for 2QLs \( Bi_2Te_3 \) is quite large, of about 148meV, 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 \( \sim 10 \text{meV} \), similar to the gap of 4QLs \( Bi_2Se_3 \). When the thickness of the film increases, the time cost for the \textit{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 \textit{ab initio} calculation, which has also successfully applied to calculate the surface states of \( Bi_2Se_3 \) type of materials \([5] \). The band gap and the total parity \( \nu \) obtained from the TB calculation are summarized in Fig.2(a) for \( Bi_2Se_3 \) and (b) \( Bi_2Te_3 \). The results for the \( 1 \sim 5 \) QLS. Good agreement between the TB model calculation and first-principles calculation is found.

**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. The results from the fully first-principles calculation are shown in table (c) for \( 1 \sim 5 \) QLS. Good agreement between the TB model calculation and first-principles calculation is found.

When the 2D system stays in the QSH phase, there are topologically protected helical edge states at the 1D edge \([13, 16, 17] \). 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 \( A_1 \) 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 \([8] \). As shown in Fig.3(b) for 2QLs \( Bi_2Te_3 \), there exist one edge state \( \Sigma_1 \) which stays in the valence band at \( \Gamma \) point and goes to conduction band at \( M \) 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 non-trivial. There are also other trivial
edge states ($\Sigma_{2,3}$ and $\Lambda_4$) always staying in the conduction band or valence band, which will 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 non-trivial 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 non-trivial 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 weak 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 non-trivial gap, which can be easily achieved in experiment. 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, and our work may open up a new way to search for new materials with 2D QSH effect. Recently, the high-quality samples of $Bi_2Se_3$ and $Bi_2Te_3$ thin films have already been fabricated in experiment, so that it is likely to observe the proposed phenomenon soon.

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 (No. 2007CB925000), the International Science and Technology Cooperation Program of China and by the US Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering, under contract DE-AC02-76SF00515. C.X. Liu and B.H. Yan acknowledge financial support from the Alexander von Humboldt Foundation of Germany.

*Note added:* During the course of this work, we have learned about the work of H.Z. Lu et al.[20] and J. Linder et al.[21], in which the similar finite size effect has been studied within the 3D TI model of Zhang et al.[2].

---

[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. 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 452, 970 (2008).
[5] H. Zhang, C. Liu, X. Qi, X. Dai, Z. Fung, and S. Zhang, Nat Phys 5, 438 (2009).
[6] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y. S. Hor, R. J. Cava, et al., Nat Phys 5, 398 (2009).
[7] Y. L. Chen, J. G. Analytis, J. Chu, Z. K. Liu, S. Mo, X. L. Qi, H. J. Zhang, D. H. Lu, X. Dai, Z. Fung, 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, to be published.
[9] Y.-Y. Li, G. Wang, X.-G. Zhu, X. Chen, Y. Wang, K. He, L. Wang, X. Ma, H. Zhang, X. Dai, et al., to be published.
[10] B. Zhou, H. Lu, R. Chu, S. Shen, and Q. Niu, Phys. Rev. Lett. 101, 246807 (2008).
[11] Z. Fang and K. Terakura, Journal of Physics: Condensed Matter 14, 3001 (2002).
[12] G. Kresse and J. Furthmueler, 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, Chemistry of Materials 13, 2068 (2001).
[19] G. Zhang, H. Qin, J. Teng, J. Guo, Q. Guo, X. Dai, Z. Fang, and K. Wu, arxiv: cond-mat/0906.5306 (2009).
[20] H. Lu, W. Shan, W. Yao, Q. Niu, and S. Shen, arxiv: cond-mat/0908.3120 (2009).
[21] J. Linder, T. Yokoyama, and A. Sudb, arxiv: cond-mat/0908.2992 (2009).
|        | $\text{Bi}_2\text{Se}_3$ |        | $\text{Bi}_2\text{Te}_3$ |        |
|--------|--------------------------|--------|---------------------------|--------|
|        | Gap at $\Gamma$ (eV)     | Parity $\nu$ | Gap at $\Gamma$ (eV)     | Parity $\nu$ |
| 1QL    | 0.710                    | +       | 0.391                     | +       |
| 2QLs   | 0.098                    | +       | 0.148                     | -       |
| 3QLs   | 0.004                    | -       | 0.036                     | +       |
| 4QLs   | 0.012                    | -       | 0.005                     | -       |
| 5QLs   | 0.004                    | -       |                           |         |