Formation of a large gap quantum spin Hall phase in a 2D trigonal lattice with three p-orbitals

Chong Li,a Kyung-Hwan Jin,b Shuai Zhang,a Fei Wang,a Yu Jia*a and Feng Liu*b,c

The quantum spin Hall (QSH) phase in a trigonal lattice requires typically a minimal basis of three orbitals with one even parity s and two odd parity p orbitals. Here, based on first-principles calculations combined with tight-binding model analyses and calculations, we demonstrate that depositing 1/3 monolayer Bi or Te atom layers on an existing experimental Ag/Si(111) surface can produce a QSH phase readily but with three p-orbitals (p_x, p_y, and p_z). The essential mechanism can be understood by the fact while in 3D, the p_x orbital has an odd parity, its parity becomes even when it is projected onto a 2D surface so as to act in place of the s orbital in the original minimum basis. Furthermore, non-trivial large gaps, i.e., 275.0 meV for Bi and 162.5 meV for Te systems, arise from a spin–orbit coupling induced quadratic p_x–p_y band opening at the Γ point. Our findings will significantly expand the search for a substrate supported QSH phase with a large gap, especially in the Si surface, to new orbital combinations and hence new elements.

Introduction

The quantum spin Hall (QSH) phase is an exotic phenomenon in condensed-matter physics and has been attracting significant attention in recent years due to its gapless one-dimensional (1D) (2D) edge (surface) states existing inside the bulk band gap of an insulator.1–6 Great effort has been made recently in the intensive search for QSH materials in both theory and experiment,7–24 in the spirit of two well-known physical mechanisms of realization of a QSH phase in 2D materials. One is any finite spin–orbital coupling (SOC) interaction induced band gap opening at the Γ point, stemming from the Kane–Mele model2 or the Haldane model5 in the honeycomb lattice. The other is the thickness of quantum wells (QWs), strain or SOC interaction induced band inversion between two bands with different parities, originating from the Bernevig–Hughes–Zhang (BHZ) model in the square lattice.3

However, up to now, in contrast to a few 3D topological insulators (TIs) confirmed in experiments7–9 for 2D TIs only HgTe/CdTe10,11 and InAs/GaSb12 QWs have been confirmed experimentally for the BHZ model, but they have too small a gap. It is, therefore, highly desirable to extend the search beyond the original hexagonal and square lattices to increase the feasibility of experimental realization. On the other hand, beyond the freestanding 2D layers, the search for 2D topological phases has been extended to surface-based or substrate-supported 2D overlayers grown on conventional semiconductor and insulator substrates.25–29 It was shown that “orbital engineering” plays a key role in designing topological phases, e.g., through the so-called substrate orbital filtering effect,25 in addition to lattice symmetry and SOC26 mentioned above. For example, it has been shown that achieving a topological insulating phase in a trigonal lattice requires either a minimal basis of one even s and two odd p_x/p_y orbitals as exemplified in the Au/GaAs(111) surface,31 or interface orbital engineering of the p–d band inversion mechanism as exemplified in the Bi/Au/Si(111) surface.32 Excitingly, a recent experiment has indirectly confirmed such surface-based large-gap QSH states33 using spectroscopy techniques. However, the direct measurement of quantum transport properties is still lacking, partly due to the too small domain size of the samples that can be made. Therefore, predicting surface-based QSH states in more material systems with different lattice symmetries and orbital types remains a highly interesting topic to facilitate experimental measurements of QSH states at elevated temperature and especially on a Si substrate with a potential impact on device applications.

In this work, we have investigated the deposition of 1/3 monolayer (ML) heavy metal Bi and non-metal Te epitaxial overlayers on a Ag/Si(111) surface. Remarkably, we discovered that a non-trivial large gap originates from the SOC-induced quadratic p_x–p_y band gap opening at the Γ point. Three p-orbital tight-binding (TB) model calculations and analyses illustrate that the odd p_x orbital in 3D plays the role of even parity in the substrate surface, thus allowing a surface-
mediated interaction between \( p_x(p_y) \) and \( p_z \), which is usually absent otherwise. Thus, we demonstrate the possibility of realizing a surface-based large-gap QSH phase via three p orbitals in a trigonal lattice.

**Methods**

All the first-principles density functional theory (DFT) calculations were performed using the Heyd–Scuseria–Ernzerhof (HSE06) hybrid functional\(^{34} \) as implemented in the Vienna Ab initio Simulation Package.\(^{35} \) The interactions between electrons and cores were treated by the projector-augmented-wave method.\(^{36} \) The cutoff energy of the plane-wave basis set was set at 250 eV. \( 5 \times 5 \times 1 \) Monkhorst–Pack \( k \)-point meshes were used for Brillouin zone integration. Seven Si layers were chosen as substrate, and the bottom layers were fixed and saturated with H atoms. The internal structures were fully relaxed until the Hellman–Feynman forces on all atoms were less than 0.01 eV Å\(^{-1} \). TB calculations and analyses are also presented in comparison with HSE06 results to reveal the detailed underlying mechanism of three p orbitals to form a QSH phase.

**Results and discussion**

As a benchmark, we first calculated the structural and electronic properties of a Ag/Si(111) surface as confirmed in experiment, where its surface is appropriately characterized by the inequivalent triangle (IET) model\(^{37} \) with small triangles (ST) and large triangles (LT) formed by Ag, and the Si trimer (SiT) formed by the topmost Si atoms with a single dangling bond, as shown in Fig. 1(a). Additionally, we reproduced a featured free-electron-like parabolic band present near the Fermi level,\(^{37,38} \) as shown in Fig. 1(d).

The pristine Ag/Si(111) surface has been used as a surface template for epitaxial growth of metal-atom overlayers (i.e., Na, K, Ag and Cu), as shown by both theory and experiment.\(^{38} \) These early studies suggest that it might be feasible to grow other metal (e.g. Bi) or non-metal (i.e. Te) elements as epitaxial overlayers in the Ag/Si(111) surface. We first consider the case of Bi, and our calculations show that the Bi atom on the ST site (the energetically most favorable) is about 5 meV lower than the ST site, which is lower than that on the SiT site by 1.16 eV, see Fig. 1(b) (top view) and 1(c) (side view), respectively. In addition, Bi atoms are slightly charged so that their Coulomb repulsion will suppress the aggregation of Bi atoms on the substrate. The small energy difference between the ST and LT sites might cause mixing of two site occupations, making the experiment more challenging. On the other hand, we carried out MD simulations to further confirm the stability of the system. We used a \( 2 \times 2 \) supercell and a MD run up to 2 ps. We found that the Bi trigonal lattice is stable up to 150 K, a typical growth temperature used in previous experiments on growth of Au epitaxial overlayers on the same Ag/Si(111) template surface.\(^{40} \) We note that due to a relatively low surface diffusion barrier,\(^{41} \) metastable epitaxial film structures are often grown at relatively low temperatures under far from equilibrium conditions. In general, our results agree well with those of Na, K, Ag and Cu,\(^{42} \) but in sharp contrast to the case of the Bi/Au/Si(111) surface,\(^{43,44} \) where SiT is the energetically most favorable site. This difference might be attributed to the fact that the bond lengths of the Si trimer shrink much more in the Ag/Si(111) than in the Au/Si(111) surface.

Next we studied the electronic properties of the Bi/Ag/Si(111) system. Our calculated band structures show evidently that there are three bands, contributed by Bi \( p_x, p_y \) and \( p_z \) orbitals (see Fig. 2(c)), residing in the bulk gap of Si, as seen in Fig. 2(a) and (b) without and with the SOC effect, respectively, similar to the case of the Bi/Au/Si(111) surface,\(^{43,44} \) where SiT is the energetically most favorable site. This difference might be attributed to the fact that the bond lengths of the Si trimer shrink much more in the Ag/Si(111) than in the Au/Si(111) surface.

Among these three bands, the upper two bands touch together at the Brillouin zone centre \( \Gamma \) point. Remarkably, upon including SOC, the touched two bands lift their degeneracy and open a quite large gap of \( \sim 275 \text{ meV} \). Additionally, since the inversion symmetry is broken in the Bi/Ag/Si(111) surface, the surface Rashba SOC effect splits the spin degeneracy of these three p bands. This can be seen clearly along the high symmetry line of the Brillouin zone (Fig. 2(b)). We further use DFT results to construct the maximally localized Wannier functions (MLWFs) in the Wannier90 package\(^{45} \) to fit the first-principles bands, as shown in Fig. 2(a) and (b). Apparently, the Wannier band shows good agreement with the HSE06 band.

Usually, an SOC induced gap opening may result in a nontrivial band topology.\(^{46} \) To confirm this, we calculated the \( Z_2 \) invariant and found that the system is indeed a nontrivial topological insulator (\( Z_2 = 1 \)). This means that three p orbitals...
can also produce a QSH state. Here we stress that the SOC gap at the $\Gamma$ point is purely via quadratic $p_x$–$p_y$ band opening, which is significantly different from the Bi/Au/Si(111) system whose mechanism is via $p$–$d$ band inversion. Among these three bands, the fundamental distinctions between them lie in $p_x$–$p_y$ quadratic bands opening from the two upper bands, while the $p$–$d$ band inversion originates from the two lower bands.

However, the Bi/Ag/Si(111) surface is not an intrinsic topological insulator because the Fermi level is located below the $\Gamma$ point where the gap opens. To overcome this deficiency, one needs to add one electron into the system to shift the Fermi level up to the touching points of $p_x$–$p_y$ bands without SOC. In general, $n$-doping can be more easily achieved in surface systems experimentally by the deposition of alkali-metal atoms. We therefore adopt this strategy by employing the K atom as a dopant in the Bi/Ag/Si(111) surface (here the K adsorption is one K atom per surface unit cell or $\sim 0.022$ electron per cm$^2$. The lattice constant of the surface unit cell is 6.68 Å of Si substrate). We found that the doped K atom prefers to adsorb on the SiT site, distributing uniformly due to the Coulomb repulsion between the charged K atoms. Remarkably, in the calculated band structures, as shown in Fig. 3(a) and (b), the doped Bi/Ag/Si(111) system becomes an intrinsic topological insulator. On the other hand, the SOC gap of the K doped Bi/Ag/Si(111) surface decreases noticeably to 143.7 meV, which, we suspect, may be attributed to the level repulsion between the filled sub-bands and band splitting induced by SOC. Namely, upon electron doping, the three $p$ bands shift downward simultaneously, leading to repulsion between the lower filled $p$ orbitals and those of substrates. Importantly, the topological features formed by the three $p$ orbitals remain intact upon electron doping (e.g. still $Z_2 = 1$), as presented in Fig. 3(c), except that the SOC gap is reduced. This is also confirmed by the Wannier band, which exhibits very good agreement with the HSE06 band, as seen in Fig. 3(a) and (b), respectively.

One may raise the question: is it possible to form an intrinsic QSH phase without extrinsic doping by growing a different epitaxial layer? To answer this question, we selected the non-metallic element Te as the epitaxial layer because it has one more electron than the Bi atom. We found that the Te atom also prefers to adsorb on the LT site just as the Bi atom. The band structure shows that the Fermi level lies exactly at the touching point of the two upper $p$ bands, as seen in Fig. 4(a). When SOC is considered, it opens a sizable
global gap of 162.5 meV, as presented in Fig. 4(b). The Wannier band (fitted by the three \( p_x, p_y \) and \( p_z \) orbitals) shows also good agreement with the HSE06 band. \( Z_2 = 1 \) as calculated, confirming that the Te/Ag/Si(111) surface is indeed an intrinsic QSH phase, as shown in Fig. 4(c). Mechanistically, this system is the same as the Bi/Ag/Si(111) surface with a nontrivial gap opened by p-p inversion with a basis of three p orbitals, and is different from the Bi/Au/Si(111) surface where the nontrivial gap is opened by the p-d inversion mechanism.\(^{32}\)

Furthermore, to understand better the band inversion associated with the gap opening process, we next investigated the evolution of a SOC induced band gap as a function of SOC strength (\( \lambda_{so} \)). We have artificially increased \( \lambda_{so} \) from 0.0 to 1.0, as shown in Fig. 5(a); for the K doped Bi/Ag/Si(111) surface, the gap first increases with increasing \( \lambda_{so} \). When \( \lambda_{so} \) exceeds 0.4, the gap is almost constant, while for the Te/Ag/Si(111) surface it monotonically increases with increasing \( \lambda_{so} \) in the whole range of \( \lambda_{so} \). This difference could be attributed to the electronic doping resulted effect as mentioned above, which is similar to band features of other QSH insulators in a trigonal lattice such as Au/GaAs(111).\(^{31}\) For the normal band order, the \( p_x \) band of a Bi or Te atom is above the \( p_{x,y} \) bands without SOC, as shown in Fig. 5(b). It is noted that there is one Bi or Te atom in a unit cell and the electronic states around the Fermi level are mainly contributed by the p orbitals. One can also induce the inverted band order with proper on-site energies for \( p_x \), \( p_y \) and \( p_z \) orbitals. The band inversion occurs between the \( p_x \) and \( p_y \) orbitals, which have the opposite-parities in 2D. Without SOC, the \( p_x \), \( p_y \) states are doubly degenerate at the \( \Gamma \) point; accordingly the system is a semimetal. With SOC, the \( p_x \), \( p_y \) states split into two non-degenerate states, opening a nontrivial gap.

Another hallmark of a 2D QSH phase is that there exist gapless helical edge states inside the band gap. We have also calculated the topological edge states by constructing the edge Green’s function of semi-infinite K/Bi/Ag/Si(111) and Te/Ag/Si(111) surfaces from the MLWFs. The local density of states of the zigzag edge is shown in Fig. 5(c) and (d), respectively. Both of them indicate that there are odd gapless topological nontrivial edge states, connecting the upper and lower band edges of bulk, and crossings over the Fermi level. This is a convincing feature of the QSH state.

Lastly, to further our understanding of the QSH phase formed by three \( p_x, p_y \) and \( p_z \) orbitals, we constructed a three-orbital tight-binding (TB) model. In the basis of \( \{p_x, p_y, p_z\} \) the corresponding spinless TB Hamiltonian for the triangular lattice, as seen in Fig. 6(a), can be written as

\[
H_0 = \sum_{\alpha} \epsilon_{\alpha} c_{\alpha}^{\dagger} c_{\alpha} + \sum_{i} \left( t_{0a,i} c_{0a}^{\dagger} c_{i0} + t_{i0,a} c_{i0}^{\dagger} c_{0a} \right),
\]

where \( a = p_x p_y p_z \) are the orbital indices, \( \epsilon_{\alpha} \) is the on-site energy and \( t_{0a,i} \) is the nearest-neighbor (NN) hopping parameter. The on-site SOC term can be written as

\[
H_{SOC} = -i\lambda (c_{\alpha}^{\dagger} c_{\beta} - c_{\beta}^{\dagger} c_{\alpha}) \cdot s_z
\]

where \( \lambda \) is the atomic SOC strength, \( i.e. \), the Te atom, and \( s_z \) is the Pauli matrix. Since spin-up and spin-down Hamiltonians are decoupled, without losing generality, we will focus on the spin-up part (one just needs to change \( \lambda \) to \( -\lambda \) for the spin-down Hamiltonian). The spin-up Hamiltonian can be represented as

\[
H = H_0 + H_{SOC} = \begin{pmatrix}
H_{11} & H_{12} & H_{13} \\
H_{12} & H_{22} & H_{23} \\
H_{13} & H_{23} & H_{33}
\end{pmatrix}
+ \begin{pmatrix}
0 & -\lambda i & 0 \\
\lambda i & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}
\]

Fig. 5 (a) The SOC gap of electronic doping K atom in Bi/Ag/Si(111) and Te/Ag/Si(111) surfaces as a function of \( \lambda_{so} \). (b) Schematic illustration of three \( p \) orbitals splitting in the crystal field and the evolution of the SOC band gap opening. (c) The energy and momentum-dependent local density of state of the zigzag edge of the semi-infinite K/Bi/Ag/Si(111) or Te/Ag/Si(111) surface.
with the expressions

\[ H_{11} = \epsilon_p + (3t_{pp} + \epsilon_{pp}) \cos \left( \frac{\sqrt{3}}{2} k_x \right) \cos \left( \frac{1}{2} k_y \right) + 2t_{pzx} \cos(k_y) \]
\[ H_{12} = (t_{pp} - \epsilon_{pp}) \sqrt{3} \sin \left( \frac{\sqrt{3}}{2} k_x \right) \sin \left( \frac{1}{2} k_y \right) \]
\[ H_{13} = t_{pp} \left[ 2\sqrt{3}\lambda \sin \left( \frac{\sqrt{3}}{2} k_x \right) \cos \left( \frac{1}{2} k_y \right) \right] \]
\[ H_{22} = \epsilon_p + (3t_{pp} + \epsilon_{pp}) \cos \left( \frac{\sqrt{3}}{2} k_x \right) \cos \left( \frac{1}{2} k_y \right) + 2t_{pzy} \cos(k_y) \]
\[ H_{23} = t_{pp} \left[ 2\sqrt{2} \lambda \sin \left( \frac{\sqrt{3}}{2} k_x \right) \cos \left( \frac{1}{2} k_y \right) \right] \]
\[ H_{33} = \epsilon_p + t_{pp} \left[ 4 \cos \left( \frac{\sqrt{3}}{2} k_x \right) \cos \left( \frac{1}{2} k_y \right) + 2 \cos(k_y) \right] \]

(4)

where \( \epsilon_p, \epsilon_{pp}, \) and \( \epsilon_{pp} \) are the on-site energies of p orbitals, and \( t_{pp}, t_{pp}, t_{pp}, \) and \( t_{pzz} \) are NN hopping parameters. The band structures of the Hamiltonian are presented in Fig. 6(b), which exhibits three typical bands without and with SOC. The different colours distinguish the components of pₓy and p₂ orbitals of opposite parity. Without SOC, there is no gap with valence and conduction bands degenerate at the Γ point. In the 3D space, the p₂ orbital has negative parity; when it is projected into a 2D plane, its parity becomes positive, and this 2D feature of the p₂ orbital is taken into account (here we neglect the Rashba effect). Such a general trend is well consistent with the SOC induced parity change of pₓy orbitals: the 3D space, the p₂ orbital has negative parity; when it is projected into a 2D plane, its parity becomes positive, and this 2D feature of the p₂ orbital is taken into account (here we neglect the Rashba effect). Such a general trend is well consistent with the SOC induced parity change of pₓy orbitals. Thus, we add an additional \( t_{pp} \) term to our hopping parameters, which are usually present in a 3D system. This is consistent with the requirement for a QSH phase in a trivalent lattice of a minimal basis of three orbitals with p₂ acting like s orbital in 2D. For the trivial case, the p₂ orbital is above the pₓy orbital, and there is no band inversion between orbitals. For the nontrivial case, the p₂ orbital shifts down to below the pₓy energy level, i.e., in the C₃v crystal field (trigonal lattice), the energy level of the p₂ orbital is below that of pₓy, as seen in Fig. 5(b), and the band order has been inverted even without SOC. Moreover, near the Fermi level, pₓy bands touch each other at the Γ point with quadratic band touching. This is a general feature of a trigonal lattice with multiple orbitals. When SOC is taken into account (here we neglect the Rashba effect because it does not change the topological order) the SOC gap and bands can be well reproduced by including an on-site SOC term in our three p-orbital model. Such a general trend is well consistent with the SOC induced pₓy > p₁ > p₀ orbital band gap opening as shown in both Fig. 5(b) and 6(b). The degenerate point at the Γ point splits into two single states; any finite SOC will open a nontrivial gap to turn the system into a QSH phase.

Moreover, we introduce an effective model at the continuum limit, which is adiabatically connected to the original lattice model. If we assume that the energy of the p₂ orbital is far away from the pₓ and pᵧ orbitals, we can neglect this high-energy degree of freedom and focus only on the low-energy states from pₓ and pᵧ orbitals. Consequently, the three-orbital Hamiltonian is reduced to an effective two-band model with

\[ H_{\text{eff}} = \sum_k \left( p_{x,k} \right) H_k \left( p_{x,k} \right) \]

(5)

Here, while the p₂ orbital drops out of the Hamiltonian, the kernel of the Hamiltonian receives correction from a virtual hopping process between the pₓ, pᵧ and p₂ orbitals in which an electron from pₓ, pᵧ orbitals is transferred to the p₂ orbital and then back to pₓ, pᵧ orbitals and other higher-order virtual processes, which we omit here. Using perturbation theory, the matrix \( H_k \) can be determined order by order as

\[ H_k = \left( \begin{array}{c} h_{11} + \lambda i \hbar \left( h_{12} - \lambda i \right) \hbar \left( h_{21} - h_{12} \right) - \frac{1}{\mu} \left( h_{12}^* h_{13} + h_{13}^* h_{22} \right) \right) \]

(6)

where \( \mu \) is the chemical potential and \( h_{22} \) is the (i, j) component of the matrix shown in matrix form. Here, the first term on the right hand side is the zero-th order term in the perturbation expansion, generated by direct hopping between p orbitals, while the second term is from the second order perturbation, which describes the virtual hopping processes mentioned above.

By expanding k to the leading order around the Γ point, we obtain an effective Hamiltonian \( H_{\text{eff}} = d \hat{t} + \hat{d} \cdot \vec{\sigma} \), where I is the identity matrix and \( \vec{\sigma} \) is the Pauli matrix and

\[ d_0 = \epsilon_p + 3(t_{pp} + \epsilon_{pp}) - \frac{9t_{pzz}^2}{2(t_{pp} + \epsilon_{pp}) + \epsilon_p + 6t_{pzz} - \epsilon_p - 3(t_{pp} + \epsilon_{pp})} \]
\[ d_1 = \frac{1}{4} \left[ t_{pp} - \epsilon_{pp} + \frac{9t_{pzz}^2}{\epsilon_p + 6t_{pzz} - \epsilon_p - 3(t_{pp} + \epsilon_{pp})} \right] k_x k_y \]
\[ d_2 = \lambda \]
\[ d_3 = -\frac{1}{4} \left[ t_{pp} - \epsilon_{pp} + \frac{9t_{pzz}^2}{\epsilon_p + 6t_{pzz} - \epsilon_p - 3(t_{pp} + \epsilon_{pp})} \right] k_x^2 k_y^2 \]

(7)

For the nontrivial band topology, \( \hat{d} = \hat{d}/|\hat{d}| \) has a vortex structure around the Γ point, as presented in Fig. 6(c). At the Γ point, \( \hat{d} \) is along the South Pole. When \( \hat{d} \) goes away from the Γ point, \( \hat{d} \) changes direction gradually from out-of-plane to in-plane. This vortex is a topological defect described by the Chern number (C). Using the \( \hat{d} \) vector, the Chern number for the spin-up states can be further defined as

\[ C_{\uparrow} = \frac{1}{4\pi} \int \frac{dk_x}{dk_y} \frac{\partial \hat{d}}{\partial k_x} \times \frac{\partial \hat{d}}{\partial k_y} \]

(8)

We calculate the Berry curvature as shown in Fig. 6(d). The Berry curvature is nonzero around the Γ point and leads to \( C_{\uparrow} = 1 \left( C_{\downarrow} = -1 \right) \). So the spin Chern number, defined as \( C_s = 1/2 \left( C_{\uparrow} - C_{\downarrow} \right) \), equals 1, confirming a QSH phase.
Conclusions

In summary, we have studied 1/3 ML Bi and Te atoms grown on a Ag/Si(111) surface, and discovered that both systems can produce a QSH phase readily. The non-trivial large gap originates from an SOC-induced $p_x-p_y$ band gap opening at the $\Gamma$ point. Such a QSH phase can be well reproduced using three $p$ orbitals with a reversed parity of the $p_z$ orbital when projected from 3D into 2D in a trigonal lattice. Our findings are mechanistically different from those of the already reported QSH phase in a trigonal lattice, and thus will expand significantly the search for a large-gap substrate-supported QSH phase to new orbital combinations.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The work was financially supported by the National Natural Science Foundation of China (grant no. 11304288 and 11274280), the Postdoctoral Science Foundation of China and Henan Province (grant no. 2014M552009 and 2013001). KHZ and FL acknowledge support from DOE-BES (grant no. DE-FG02-04ER46148). The calculations were performed on the High Performance Clusters of Zhengzhou University.

References


