Alloy theory with atomic resolution for Rashba or topological systems

Zhi Wang\(^1\) and Alex Zunger\(^1\), *

\(^1\) Renewable and Sustainable Energy Institute, University of Colorado, Boulder, Colorado 80309, USA

*alex.zunger@colorado.edu

Interest in substitutional disordered alloys has recently reemerged with focus on the symmetry-sensitive properties in the alloy such as topological insulation and Rashba effect. A substitutional random alloy (AX)\(_{x}\)(BX)\(_{1-x}\) of components AX and BX generally manifests a distribution of local environments, whereby each X site for example can be locally decorated by different substitutional arrangements of \(\{A, B, X\}\) atoms, thus creating an inherently polymorphous network. Electrons will then respond to the existence of different local environments and site symmetries, creating local charge transfer and atomic displacements patterns observed in experiments. While the macroscopic average structure \(S_0\), as seen by probes with long coherence length, may have the original high symmetry of the constituent compounds, many observable physical properties are sensitive to local symmetry, and are hence the average \(<P(S_i)>\) of the properties \(\{P(S_i); i=1...N\}\) of the individual microscopic configurations \(\{S_i; i=1,N\}\) rather than the property \(P(<S>)=P(S_0)\) of the macroscopically averaged high-symmetry (monomorphous) configuration \(S_0\). The fundamental difference between the polymorphous representation \(<P(S_i)>\) vs the monomorphous \(P(S_0)\) in modeling substitutionally disordered alloys led to the often diverging results between methods that ‘see’ atomic details and those that see only the high symmetry of the constituents, while missing the atomic-scale resolution needed in many cases to discern local symmetry-related physics. A natural approach that captures the polymorphous aspect of random alloys is the well-known supercell approach where lattice sites are occupied by the alloyed elements with a particular form of disorder and solved via periodic electronic structure methods for sufficiently large supercells. However, such approaches tend to produce complex \(E\) vs \(k\) dispersion relations (‘spaghetti bands’), rendering the wavevector \(k\) information needed in theories of topology and Rashba physics and seen in angular resolved experiments, practically inaccessible. The results of such calculations have consequently been displayed as density of states. A solution that retains the polymorphous nature of the random alloy but reinstates the \(E\) vs \(k\) relation in the base Brillouin zone is to unfold the supercell bands. This yields alloy “Effective Band Structure” (EBS), providing a three-dimensional picture of the distribution of spectral density in the whole Brillouin zone. It consists of \(E\)- and \(k\)-dependent spectral weight with coherent and incoherent features, all created naturally by the specific nature of the chemical bonding underlying the polymorphous distribution of many local environments. We illustrate this EBS approach for CdTe-HgTe, PbSe-SnSe and PbS-PbTe alloys, showing atomic-scale effects such as formation of a distribution of A-X and B-X bond lengths, local charge transfer, and the creation and destruction of valley degeneracies. In CdTe-HgTe the disorder effect is so weak that the incoherent term is negligible, and the monomorphous approaches are still feasible in this alloy. In PbSe-SnSe the stronger disorder effect introduces significant (~150 meV) band splitting of the topological band inversion, forming a sequential inversion of multiple bands which is important for
the topological transition but absent in monomorphous methods. In PbS-PbTe there is a strong disorder effect, revealing the emergence of ferroelectricity from the polymorphous network in this alloy.

I. Introduction

**Alloy theories with or without atomic resolution:** Many target properties of materials are not available in currently known individual components AX or BX but do exist in alloys of such components (AX)\(_x\)(BX)\(_{1-x}\). Examples include band gap and effective mass tuning in semiconductors, ductility, brittleness and a given degree of short-range ordering (SRO) in intermetallics, and topological properties existing only after alloying\(^1\). Inevitably, disorder effects due to the substitutional occupations of A and B atoms in (AX)\(_x\)(BX)\(_{1-x}\) is the key to understand alloy properties such as mobility, conductivity, electronic structure and localization. Of particular recent interest are *alloy properties that depend on local symmetry*, such as the emergence of Rashba effect, predicated on absence of inversion symmetry, in alloys of centrosymmetric components\(^2\), or the appearance of topological properties in alloys e.g. (PbSe)\(_x\)(SnSe)\(_{1-x}\)\(^3\) and Mo\(_x\)W\(_{1-x}\)Te\(_2\)\(^4\) at specific, time reversal invariant wavevectors.

However, accounting for local symmetry effects in random disordered alloys is never a simple task. Common models of disorder have considered single-site models and continuum model that account for the changes of lattice vectors and volume (e.g., via Vegard’s law) but retain the macroscopic symmetry rather than the local symmetry. The Virtual Crystal Approximation\(^5\) (VCA) and the Single-site Coherent Potential Approximation\(^6\) (S-CPA) rely on the assumption that alloy short-range (atomic-size) disorder has negligible effects and can be averaged out, thus largely restoring in the alloy the symmetry (hence band structure shapes and degeneracies) of the parent compounds, with account k-dependent broadening of the band structure\(^7\). Nevertheless, the VCA and S-CPA generally lack atomic-scale resolution of disorder that should be visible when the alloyed elements differ sufficiently on some scale of atomic sizes, bonding characteristics, or charge transfer.

The insufficiency of the monomorphous alloy description has been shown in many previous works. Examples include (i) the extended X-ray absorption fine structure (EXAFS) and atomic pair distribution function (PDF) observation of the existence in random alloy of a distribution of A-X and B-X bond lengths\(^8\)–\(^13\); and (ii) the observation that atomic site charges \(\{Q_i\}\) in an alloy depend on the local neighborhood environment of site \(i\), which results in a non-vanishing, large electrostatic alloy (Madelung) energy\(^14\), contradicting the common assumption underlying S-CPA, of the independence of charges on local environment, leading to \(<Q_iQ_j>=<Q_i><Q_j>=0\ i.e.\ vanishing\ electrostatic\ energy. To achieve an atomic resolution of disorder one needs theory that recognizes local symmetries, yet informs about the extent to which the long-range translational symmetry underlying the concept of wavevector is retained in alloys.

Theories of topological effects in random alloys\(^15\)–\(^17\) argue that in an infinite sample of random alloy all symmetry elements (e.g. inversion center) of the constituent solids being mixed are preserved on average, so the latter configuration can be used to evaluate topological characteristics. However, even if this proposition were correct, the properties of the alloy \(<P>\) (such as band structure and band inversion) do not reflect the property of the macroscopically averaged
configuration \(<P>=P(S_0)\) but rather the average \(P_{\text{obs}}=\sum P(S_i)\) of the properties \(\{P(S_i)\}\) of the individual microscopic configurations \(\{S_i; i=1, N\}\).

**The atomically resolved perturbations induced by A-on-B substitution in alloys:** As is well known, disorder models with atomic resolution can be built by solving the band structure problem of supercell whose \(N\times N\times N\) primitive cells contain \(N^3\) sites are occupied randomly by the constituent atoms \(A\) and \(B\). The alloyed atoms can naturally have different electronic structures, atomic sizes, and tendencies for charge transfer, thereby creating a polymorphous representation where (unlike VCA or S-CPA) the common atom \(X\) is ‘seeing’ a variety of local environments, depending on the number of \(A\) and \(B\) atoms locally coordinating different \(X\) sites. In this representation, \(A-X\), \(B-X\) and \(A-B\) charge transfer, as well as the existence of a distribution of \(A-X\) and \(B-X\) bond lengths is allowed, in addition to the trivial variation of volume with composition. The spectra can be converged with respect to the size of the supercell and by averaging over a representative number of random realizations. More effectively, one can construct from the outset special supercells, ‘Special Quasirandom Structures’ (SQS) that are guaranteed to reproduce pair and many body correlation functions in the best way possible for a given supercell size \(N\).

**The limitation of supercells and band unfolding:** As the size of supercells increase, the \(E\) vs \(k\) dispersion relation also becomes more complex because of band folding. This leads to the difficulty of interpreting alloy effects that depend on wavevector \(k\) information, e.g. topological materials, Rashba physics and Angle-Resolved Photoemission Spectroscopy (ARPES) analyses. Such effects would be concealed inside the ‘spaghetti-like’ supercell bands. This is perhaps the primary reason density of states, rather than \(E\) vs \(k\) dispersion, is usually shown in such supercell calculations. This difficulty can be solved by the “Effective Band Structure” (EBS) method\(^\text{18}\), which unfolds the supercell band structures into the primitive Brillouin zone, same as the Brillouin zone used in the theoretical study of pure compounds and the experimental ARPES study of alloy. Similar to ARPES, the EBS method also provides a three-dimensional picture of the spectral function with \(E\)- and \(k\)-dependent features including coherent (dispersive term, or ‘sharpness’) and incoherent (band non-dispersive broadening, or ‘fuzziness’) spectral weights\(^\text{19}\), all naturally produced by the polymorphous nature of the many local environment in alloys. Depending on the electronic structure method used to solve the supercell Hamiltonian (mean-field like approaches, or explicitly correlated approaches), additional coherent or incoherent effects originating from many-body effects can come in. Here we emphasize that even a single deterministic electronic structure method such as Density Functional theory (DFT) will already produce three-dimensional spectral functions with coherent and incoherent features just because of allowing a polymorphous representation of the structure.

We applied our supercell model with DFT to several substitutionally random alloys CdTe-HgTe (topological alloy), PbSe-SnSe (topological alloy) and PbS-PbTe (bulk Rashba alloy). By using the EBS, we restored the \(E\) vs \(k\) band dispersion for alloy into the primitive Brillouin zone. We found that:

1. In **CdTe-HgTe** is a weakly perturbed alloy made of nearly size-matched components of similar electronic and bonding properties, with band gap occurring at nondegenerate state at the \(\Gamma\) point. The polymorphous theory gives band structure and topological band inversion point that are rather similar to those found previously in the monomorphous theory.
(2) In PbSe-SnSe is an alloy with moderate chemical disparity in the alloyed elements Pb vs Sn but with degenerate band edge states at L point. Here, the monomorphous theory fails to describe the disorder-induced band edge splitting, whereas the polymorphous representation shows that the states split and invert \textit{sequentially}. This is because the monomorphous approaches do not consider the symmetry breaking induced by charge exchange and bond relaxation, masking such events by a sweeping band broadening parameter. Such approaches are inadequacy for prediction of topological properties in this system.

(3) PbS-PbTe represents a strongly scattering alloy (8\% lattice mismatch). We find in the high-resolution picture a coherent, Rashba-like band splitting (revealing the ferroelectricity) emerging from the incoherent band broadening (revealing the alloy disorder). The mixture of coherent and incoherent features in this alloy is absent in the monomorphous approaches.

\section*{II. Modeling the physical changes in the constituent compounds upon forming an alloy.}

To analyze the specific physical effects contributing to alloy formation we will decompose the alloy formation into three physical steps illustrated in Fig. 1:

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{The physical steps from pure compounds to alloy. (a) Pure compounds AX and BX at their equilibrium lattice vectors \{\(\vec{a}_0(AX)\)\} and \{\(\vec{a}_0(BX)\)\}; (b) deform AX and BX lattice vectors, and let them both have the alloy lattice vectors \{\(\vec{a}_0(x)\)\}; (c) form a supercell of AX-BX with lattice vectors \{\(\vec{a}_0(x)\)\} while keeping all bond lengths equal to each other; (d) relax all bonds in the supercell with lattice vectors \{\(\vec{a}_0(x)\)\}. Step a→b is the lattice deformation (LD); step b→c is charge exchange (CE), whereas step c→d is the bond relaxation (BR) step.}
\end{figure}

(1) \textbf{Lattice deformation (LD) step} (from Fig. 1(a) to Fig. 1(b)): Here we prepare the two constituent compounds so they could form a common alloy lattice in the next step. To do so we distort the lattice vectors for both components, so both have the same lattice vectors \{\(\vec{a}_0(x)\)\} appropriate to this alloy of composition \(x\), where the subscript 0 means the equilibrium lattice vector. For example, pure PbS and PbTe are both face-center cubic (FCC) structure, with \(a_0(PbS) = 6.03\ \text{Å}\), \(a_0(PbTe) = 6.55\ \text{Å}\) and \(\alpha = \beta = \gamma = 90\) degree; while in PbSTe alloy the lattice constant has changed \((a_0(PbS) < a_0(x) < a_0(PbTe))\), and the crystalline structure has transformed into a distorted rhombohedral structure. We then in this step expand the smaller component (here, PbS) and
compress the larger component (here, PbTe), and distort the two compounds into the same, distorted rhombohedral structure as in alloy. Note that the alloy lattice constants as well as the cell distortions are calculated from DFT and validated with experiments. The change in extensive property \( P(x) \) (total energy etc.) in this step can be modelled by the formal reaction

\[
AX|\vec{a}_0(AX) + BX|\vec{a}_0(BX) \rightarrow AX|\vec{a}_0(x) + BX|\vec{a}_0(x)
\]

This step reveals the contribution of the deformation of lattice on the alloy formation.

(2) Charge exchange (CE) step (from Fig. 1(b) to Fig. 1(c)): Here we mix the structures prepared in the previous step to form the random alloy supercell at lattice vectors \( \{\vec{a}_0(x)\} \). At this step the A-X and B-X bonds are allowed to coexist in the alloy so charge exchange can occur among different atomic sites, but all bonds are still constrained to equal to each other. Each atomic site of a given chemical identity (such as X of AX) will have in principle, a different charge distribution around it, generally reflecting the number of A and B atoms around it. CE step is a polymorphous effect thus not captured by monomorphous approaches. The change in extensive property \( P(x) \) in this step can be modelled by the formal reaction

\[
AX|\vec{a}_0(x) + BX|\vec{a}_0(x) \rightarrow A_{1-x}B_xX|\vec{a}_0(x)
\]

representing charge exchange at constant volume and ideal bond geometry.

(3) Bond relaxation (BR) step (from Fig. 1(c) to Fig. 1(d)): Here we take the previous step where a supercell with its attendant charge transfer was already formed and now allow the relaxation for all internal atomic positions at the fixed alloy lattice vectors \( \{\vec{a}_0(x)\} \). Note that for each composition the bond lengths are not single-valued but have distributions due to the polymorphous local environment effect, i.e., bond lengths \( R_{A-X}^{(n)}(x) \) and \( R_{B-X}^{(n)}(x) \) are neighborhood-configuration-dependent \(((n)\)-dependent). The BR step is a polymorphous effect thus not captured by monomorphous approaches. The change in extensive property \( P(x) \) in this step can be modelled by the formal reaction

\[
A_{1-x}B_xX|\vec{a}_0(x) \rightarrow \text{relaxed } A_{1-x}B_xX|\vec{a}_0(x)
\]

The total change in extensive property \( P(x) \) of alloy relative to the linearly weighted average of the constituents can be written as

\[
\Delta P_{\text{tot}}(x) = P(x) - \left[ xP(AX) + (1 - x)P(BX) \right] = \Delta P_{\text{LD}}(x) + \Delta P_{\text{CE}}(x) + \Delta P_{\text{BR}}(x)
\]

which will assist us in analyzing physical alloy effect.

III. Computational details

This work used the computational resources of the Extreme Science and Engineering Discovery Environment (XSEDE)\textsuperscript{20}. We have performed DFT calculations as applied within the Vienna \textit{ab initio} simulation package (VASP)\textsuperscript{21} using the projector-augmented wave (PAW)\textsuperscript{22} pseudopotentials. Cd 4d, 5s, Hg 5d, 6s, Pb 5d, 5s, 5p, Sn 4d, 5s, 5p, S 3s, 3p, Se 4s, 4p and Te 5s, 5p have been treated as valence electrons. For all primitive cells of pure compounds, we used an 8×8×8 Γ-center \( k \) mesh in the electronic self-consistent iterations and in the atomic relaxations. Table I shows the space group, energy cutoff and exchange correlation terms, and the comparison of relaxed lattice constant and band gap for all pure compounds between DFT and experimental results. All alloy supercells have been constructed using the SQS method as implemented in the Alloy Theoretic Automated
Alloy supercell sizes are 32 formula unit (f.u.) (CdTe-HgTe), 128 f.u. (PbSe-SnSe) and 32 f.u. (PbS-PbTe), while the \( k \) meshes are \( 4 \times 4 \times 4 \) (CdTe-HgTe), \( 3 \times 3 \times 2 \) (PbSe-SnSe) and \( 4 \times 4 \times 4 \) (PbS-PbTe). We calculated all alloy supercells using the same parameters as in their constituent compounds (e.g., for all CdTe-HgTe alloy supercells we used the same parameters as in CdTe and HgTe). Note that the space groups of alloys are always different from constituent compounds, because all atomic positions as well as the lattice vectors in alloys have been determined by fully relaxing (atomic positions, cell shape and cell volume) the alloy supercells. CdTe-HgTe and PbSe-SnSe alloy supercells are still in cubic phase after relaxation, however they are no longer F-43m or Fm-3m because of the polymorphous network (different atomic sites have different element occupations and different atomic displacements). This makes it completely different with the monomorphous approaches. Meanwhile PbS-PbTe alloy supercells become distorted rhombohedral after relaxation. EBS calculations have been done by a modified version of BandUP code\(^{25}\).

TABLE I. The calculation details for pure compounds, and the comparison between DFT and experimental results.

| Compounds (space group) | Exchange correlation term | Cutoff energy | Lattice constant (DFT) | Lattice constant (Exp.) | Band gap (DFT) | Band gap (Exp.) |
|-------------------------|---------------------------|---------------|------------------------|------------------------|----------------|----------------|
| CdTe (F-43m)            | LDA+U (U\(_d\)=10 eV)    | 400 eV        | 6.410 Å                | 6.48 Å                | 0.86 eV        | 1.65 eV        |
| HgTe (F-43m)            | LDA+U (U\(_d\)=10 eV)    | 400 eV        | 6.436 Å                | 6.46 Å                | -0.26 eV       | -0.3 eV        |
| PbSe (Fm-3m)            | GGA+U (U\(_{Pb_s}\)=2 eV) | 360 eV        | 6.22 Å                | 6.12 Å                | 0.23 eV        | 0.17 eV        |
| SnSe (Fm-3m)            | GGA+U (U\(_{Pb_s}\)=2 eV) | 360 eV        | 5.99 Å                | 6.00 Å                | 0.72 eV        | 0.62~0.72 eV   |
| PbS (Fm-3m)             | GGA+U (U\(_{Pb_s}\)=2 eV) | 360 eV        | 6.03 Å                | 5.93 Å                | 0.3 V          | 0.29 eV        |
| PbTe (Fm-3m)            | GGA+U (U\(_{Pb_s}\)=2 eV) | 360 eV        | 6.55 Å                | 6.44 Å                | 0.2 eV         | 0.19 eV        |

**IV. Unfolding the supercell energy bands and recovering \( E \) vs \( k \) alloy EBS**

Here we briefly summarize the basic equations of EBS. The band folding mechanism in a supercell can be expressed as

\[
|Km\rangle = \sum_{i=1}^{N_K} \sum_{n} F(k_i, n; K, m)|k_i n\rangle
\]  

(5)
where $|Km\rangle$ is the $m$-th electronic state at $K$ in supercell Brillouin zone, $|k_in\rangle$ is the $n$-th electronic state at $k_i$ in primitive Brillouin zone. One can then unfold the supercell band structure by calculating the spectral weight $P_{Km}(k_i)$ from

$$P_{Km}(k_i) = \sum_n |\langle Km | k_in \rangle|^2$$  \hspace{1cm} (6)

which is the Bloch ‘preservation’ of Bloch wavevector $k_i$ in $|Km\rangle$ when $E_n = E_m$. Finally, the EBS can be obtained using the spectral function $A(k_i, E)$,

$$A(k_i, E) = \sum_m P_{Km}(k_i) \delta(E_m - E)$$  \hspace{1cm} (7)

As an example, Fig. 2 shows the comparison among pure PbTe band structure, supercell PbSTe band structure and supercell PbSTe EBS, all plotted in the PbTe primitive Brillouin zone. The spectral function can be sharply dispersive (coherent) e.g. conduction band minimum (CBM) along $\Gamma$-$Z$ direction, or become completely non-dispersive (incoherent) e.g. valence band maximum (VBM) at $\Gamma$, or be a mixture of both e.g. VBM along $\Gamma$-Z direction.

![Diagram](image)

**FIG. 2.** (a) Conventional band structure of the distorted R3m PbTe (2-atom primitive cell, plotted in primitive Brillouin zone). (b) Supercell band structure (64-atom supercell PbS$_x$Te$_{1-x}$ at $x=9.375\%$, plotted in the same primitive Brillouin zone) and (c) EBS (unfolded from 64-atom supercell for PbS$_x$Te$_{1-x}$ alloy at $x=9.375\%$ into the same primitive Brillouin zone). The primitive lattice vectors have been distorted according to the relaxed supercell structure in order to make a direct comparison. (a) (b) and (c) are all plotted along the same $\Gamma$-$Z$-$U$ direction in the primitive Brillouin zone [$Z=(\pi/a_1, \pi/a_2, \pi/a_3)$, $U=(\pi/2a_1, 2\pi/a_2, \pi/2a_3)$].

**V. Results and discussion**

**A. Decomposition of alloy effects into physical terms**

We study alloys having different scales of disorder: HgTe-CdTe, PbSe-SnSe and PbS-PbTe. We will see that the scale of disorder is system-dependent, from weak (HgTe-CdTe) to intermediate (PbSe-SnSe), and to strong (PbS-PbTe). It is an interesting question how different scales of disorder in different materials can affect $E$ vs $k$ structure.

Fig. 3 shows the polymorphous local environment effects in the CE and BR terms in the alloy forming reactions. The effects of CE step have been shown in Fig 3(a)(c)(e) by plotting the contours...
of the charge density nearby one common atom in the three alloy systems. We see that the charge density around the common atom (Te in CdHgTe; Se in PbSnSe and Pb in PbSTe) has different shapes when considering the bonds formed with the dissimilar alloyed atoms. The effects in BR step have been shown in Fig 3(b)(d)(f) by the bond length distribution profiles. The asymmetry of charge density along different bonds, as well as the spread of bond length variations of different bonds, becomes more significant as one progresses from weak to strong alloying. In CdTe-HgTe, the electron density distributions along Cd-Te and Hg-Te bonds show only small differences (Fig. 2(a)), and the bond lengths of Cd-Te (as well as Hg-Te) are virtually equal with negligible distribution of values ($\sigma<0.002$ Å as shown in Fig. 2(b)). In the strongly perturbed alloy PbS-PbTe, (i) the Pb-S and Pb-Te bonds have distinct charge densities, (ii) the bond lengths show a significant statistical spread for the same chemical bond (Fig. 2(e)(f), e.g., the Pb-S bonds vary in a range of ~1 Å, and (iii) chemically different bonds Pb-S and Pb-Te have different lengths away from the macroscopic lattice constant. Clearly, the high symmetric alloy model assumed in monomorphous theories ignore the local atomic environment effects for both electron density and geometric bond structures.

FIG. 3. The charge density and bond length profiles in three alloy systems. (a) (c) (e) show the logarithmically spaced contours for the charge density nearby one common atom for each of the alloys in CE step. (b) (d) (f) show the bond length distributions for the different types of bond (red and blue solid lines) with means (circles) and standard deviations (bars). The uniform bond lengths $R_0(x)$ in the unrelaxed lattice (i.e. before BR step) are shown as the black dash lines in (b)(d)(f). $R^{(0)}$ shown in red and blue dash lines are the bond lengths in pure compounds.
The disorder effects can also be seen from the decomposition of the alloy mixing energy (total energy of alloy with respect to equivalent amounts of its constituents) from Eq. (4), shown in Fig. 4. Recall that the LD step introduces only monomorphous effects while the CE and BR steps result in polymorphous local environments. The scale of the total mixing energies in parts (a) (b) and (c) already disclose the scale of disorder, CdTe-HgTe having 7 meV (x=50%), PbSe-SnSe having 3 meV (x=31.25%), whereas PbS-PbTe having 70 meV (x=50%) which agrees with previous works26,27. In the weak alloy CdTe-HgTe where the lattice mismatch is tiny (~0.3%), the lattice deformation and bond relaxation energies are negligible, and the charge exchange contributes most to the mixing enthalpy (ΔH). As the lattice mismatch increases (PbSe-SnSe and PbS-PbTe), the bond relaxation BR step does not release as much energy as the lattice deformation costs in the first place, and the charge exchange energy is small, so the total mixing energy is positive and non-negligible. It is hence inadequate to include only the trivial lattice deformation effect as in simple monomorphous models, since the polymorphous terms, CE and BR can be important, neglecting which would lead incorrectly to high mixing enthalpies.

FIG. 4. Alloy mixing energies in each physical step and the total formation energies, for (a) CdTe-HgTe, (b) PbSe-SnSe and (c) PbS-PbTe.

B. Effective Band Structure of the three alloys:
Here we show EBS pictures for each of the 3 alloys. We found that the spectral functions show a clear trend with respect to the scale of disorder: from weak to strong alloy, the spectral weights lose the coherent dispersion more quickly when leaving the reciprocal high-symmetric $k$ points. Moreover, each alloy shows some unique features:

(1) In pure CdTe and HgTe, each band at $\Gamma$ point is a twofold degenerate. CdTe-HgTe EBS shows very sharp band structure near the time reversal invariant momentum (TRIM) $\Gamma$ point and no band splitting at $\Gamma$ point (Fig. 5), which can be attributed to the very weak alloy disorder effect as shown in Fig 3(a)(b) and Fig 4(a). All bands near the Fermi level, including the light electron, light hole and heavy hole states (corresponding to $\Gamma_6$ and $\Gamma_8$ states in pure HgTe and CdTe), are sharply dispersive and almost 100% coherent. Note that there is a tiny splitting (<25 meV) on the heavy hole state along $\Gamma$-L direction, which is attributed to the small atomic displacement (see Fig. 3(b)) and agrees with previous work. Therefore, in this alloy, the monomorphous theories might be adequate, e.g., for predicting the topological band inversion between $\Gamma_6$ and $\Gamma_8$ states at $\Gamma$ point as Cd composition increases.

![EBS of CdTe-HgTe supercell (32 f.u.) at (a) Cd=12.5%, (b) Cd=18.75%, (c) Cd=25% and (d) Cd=31.25%, unfolded to the zinc-blende primitive Brillouin zone. All EBS are plotted along the same X-(\Delta)-(\Gamma)-(\Lambda)-L direction in the primitive Brillouin zone.](image)

FIG. 5. EBS of CdTe-HgTe supercell (32 f.u.) at (a) Cd=12.5%, (b) Cd=18.75%, (c) Cd=25% and (d) Cd=31.25%, unfolded to the zinc-blende primitive Brillouin zone. All EBS are plotted along the same X-(\Delta)-(\Gamma)-(\Lambda)-L direction in the primitive Brillouin zone.

(2) As the scale of disorder increasing, in PbSe-SnSe EBS, the degeneracy of band valleys at the TRIM L point is no longer preserved: there is a significant band splitting (~150 meV) on the band valleys at L point (marked by red circles in Fig. 6). Both valence band and conduction bands are split but sharply dispersive near $\Gamma$; moving along the $\Gamma$-L-W lines, one can see that the VBM quickly loses its coherent feature, first in the middle of $\Gamma$-L line then near the W point. More importantly, the
band valley splitting at TRIM L further indicates that the ‘sharp’, concurrent NI-TCI transition at specific composition, which has been predicted earlier\textsuperscript{3,7} to be due to the band inversion between highly degenerated bands, is not what a theory with atomic resolution finds: the alloy system actually experiences a regime having sequential inversions of multiple bands nearby Fermi level at L point. The sequential band inversion regime is visible only within the polymorphous model e.g. supercell but not in the monomorphous models.

![Diagram](image)

**FIG. 6.** EBS of PbSe-SnSe supercell (128 f.u.) at Sn=25% unfolded to the FCC PbSe primitive Brillouin zone, plotted (a) along Γ-(Λ)-L-(Q)-W direction and then (b) zoomed-in around L point. The red circles mark the sequential band inversion in this alloy attributed to band edge splitting at L point.

(3) In the strong alloy PbS-PbTe, the introduction of S atom leads to a ferroelectric (FE) sublattice displacement at low T, making the alloy a famous candidate of bulk Rashba and FE materials\textsuperscript{10}. In VCA and S-CPA the ferroelectricity was mimic monomorphously by using a uniform displacement between cation and anion sublattice, while in our supercell, ferroelectricity is polymorphous. We found that (Fig. 7) in this alloy most bands suffer splitting and broadening, but the CBM and VBM at Z point is relatively sharp and dispersive, which agrees well with previous work\textsuperscript{7}. Moreover, we see that each conduction bands clearly split into two bands along Γ-Z and Z-U directions, while VBM, although mixed with incoherent broadening, also shows such two-band splitting along Z-U direction. Note that this type of band splitting is coherent because each split branch shows an individual dispersion. We found in previous S-CPA work a similar result\textsuperscript{7}, where however it was suggested to be an incoherent band broadening. We suggest that this band splitting of band edge states can be Rashba-like and reveal the ferroelectricity of this alloy system.
FIG. 7. EBS of PbS-PbTe supercell (32 f.u.) at (a) S=9.375%, (b) S=18.75%, (c) S=25% and (d) S=31.25%, unfolded to the distorted R3m PbTe primitive Brillouin zone. All EBS are plotted along the same Γ-Z-U direction in the primitive Brillouin zone \([Z=(\pi/a_1, \pi/a_2, \pi/a_3), U=(\pi/2a_1, 2\pi/a_2, \pi/2a_3)]\). The white dot lines are only for eye-guiding to show the coherent splitting on VBM along Z-U.
VI. Conclusions

With the aid of polymorphous supercell approach and band unfolding, we restore the all-important $E$ vs $k$ dispersion relation to alloy theory in CdTe-HgTe, PbSe-SnSe and PbS-PbTe alloys, revealing various sources of alloy formation, such as lattice deformation, charge exchange and bond relaxation. This allows one to define a scale of disorder, by the deviations that these effects create relative to the monomorphous level. We find that the spectral weights unfolded to primitive Brillouin zone shows: (1) both coherent, dispersive splitting of band degeneracies and incoherent band broadening that depends on the wavevectors and on the scale of alloy disorder; (2) coherent-incoherent transition on different bands along different $k$ space directions; and (3) Rashba-like band splitting consisting of both coherent and incoherent features. We expect that such effects—notably the splitting of band degeneracies—could be observed by ARPES.

Acknowledgments

The work at the University of Colorado at Boulder was supported by the National Science foundation NSF Grant NSF-DMR-CMMT No. DMR-1724791. We thank Qihang Liu for fruitful discussions on the subject. The ab-initio calculations were done using the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562.

References

1. D. Hsieh, D. Qian, L. Wray, Y. Xia, Y.S. Hor, R.J. Cava, and M.Z. Hasan, Nature 452, 970 (2008).
2. O. Rubel and A. Bokhanchuk, ArXiv150803612 Cond-Mat (2015).
3. P. Dziawa, B.J. Kowalski, K. Dybko, R. Buczko, A. Szczerback, M. Szot, E. Łusakowska, T. Balasubramanian, B.M. Wojek, M.H. Berntsen, O. Tjernberg, and T. Story, Nat. Mater. 11, 1023 (2012).
4. T.-R. Chang, S.-Y. Xu, G. Chang, C.-C. Lee, S.-M. Huang, B. Wang, G. Bian, H. Zheng, D.S. Sanchez, I. Belopolski, N. Alidoust, M. Neupane, A. Bansil, H.-T. Jeng, H. Lin, and M. Zahid Hasan, Nat. Commun. 7, 10639 (2016).
5. L. Nordheim, Ann Phys 9, 607 (1931).
6. P. Soven, Phys. Rev. 156, 809 (1967).
7. D. Di Sante, P. Barone, E. Plekhanov, S. Ciuchi, and S. Picozzi, Sci. Rep. 5, 11285 (2015).
8. J.C. Mikkelsen and J.B. Boyce, Phys. Rev. B 28, 7130 (1983).
9. Q.T. Islam and B.A. Bunker, Phys. Rev. Lett. 59, 2701 (1987).
10. B.A. Bunker, Z. Wang, and Q. Islam, Ferroelectrics 150, 171 (1993).
11. B. Ravel, E. Cockayne, M. Newville, and K.M. Rabe, Phys. Rev. B 60, 14632 (1999).
12. I.-K. Jeong, T.W. Darling, J.K. Lee, T. Proffen, R.H. Heffner, J.S. Park, K.S. Hong, W. Dmowski, and T. Egami, Phys. Rev. Lett. 94, 147602 (2005).
13. S. Kastbjerg, N. Bindzus, M. Søndergaard, S. Johnsen, N. Lock, M. Christensen, M. Takata, M.A. Spackman, and B.B. Iversen, Adv. Funct. Mater. 23, 5477 (2013).
14. R. Magri, S.-H. Wei, and A. Zunger, Phys. Rev. B 42, 11388 (1990).
15. L. Fu and C.L. Kane, Phys. Rev. B 76, 045302 (2007).
16. L. Fu and C.L. Kane, Phys. Rev. Lett. 109, 246605 (2012).
17. Y. Ando and L. Fu, Annu. Rev. Condens. Matter Phys. 6, 361 (2015).
18 V. Popescu and A. Zunger, Phys. Rev. Lett. **104**, 236403 (2010).
19 A. Damascelli, Z. Hussain, and Z.-X. Shen, Rev. Mod. Phys. **75**, 473 (2003).
20 J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G.D. Peterson, R. Roskies, J.R. Scott, and N. Wilkins-Diehr, Comput. Sci. Eng. **16**, 62 (2014).
21 G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
22 G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
23 A. van de Walle, M. Asta, and G. Ceder, Calphad **26**, 539 (2002).
24 A. van de Walle, P. Tiwary, M. de Jong, D.L. Olmsted, M. Asta, A. Dick, D. Shin, Y. Wang, L.-Q. Chen, and Z.-K. Liu, Calphad **42**, 13 (2013).
25 P.V.C. Medeiros, S. Stafström, and J. Björk, Phys. Rev. B **89**, 041407 (2014).
26 S.-H. Wei and A. Zunger, Phys. Rev. B **55**, 13605 (1997).
27 J.W. Doak and C. Wolverton, Phys. Rev. B **86**, (2012).
28 O. Rubel, A. Bokhanchuk, S.J. Ahmed, and E. Assmann, Phys. Rev. B **90**, 115202 (2014).