Why does silicon have an indirect band gap?

Emily Oliphant[1], Veda Mantena,[1] Madison Brod[2], G. Jeffrey Snyder[2], Wenhao Sun[1*]
[1] Department of Materials Science, University of Michigan, Ann Arbor, Michigan 48109, United States
[2] Northwestern University, Materials Science and Engineering Evanston IL, 60208 USA

*Corresponding Author: whsun@umich.edu

Significance Statement

Density functional theory (DFT) is a powerful tool for predicting the electronic band structures of functional materials. However, it is often difficult to intuit how major band structure features—such as band gap magnitude, location of band extrema, effective masses, *etc.*—arise from the underlying crystal chemistry of a material. Here, we present a conceptual and computable framework to extract the chemical bonding origins of DFT-calculated band structures. As a key example, we explain the indirect band gap of silicon, revealing new insights to understand and engineer the position of its conduction band minimum. These calculation techniques can be broadly applied to reveal the crystal chemistry origins of electronic structure features in other optical, electronic, and magnetic materials.

Abstract

Chemical bonds give rise to electronic structure, but the complex bonding environments of crystals make it difficult to isolate which specific orbitals help to shape the overall electronic band dispersion. Here we present a systematic three-step approach to interpret how band dispersion arises from multiple orbital interactions in a 3D crystal. This process proceeds by determining 1) which orbitals contribute to a band, 2) how these orbitals bond across *k*-space, and 3) the strength of the orbital bonds. Applying this approach to silicon, we provide a revised interpretation for the origin of the silicon conduction band minimum at the low-symmetry Δ point along the Γ-X line. Specifically, we find that the dip in the silicon conduction band near X originates from a cosine shape along Γ-X arising from second nearest neighbor p_x–p_x bonding, combined with a positive linear slope due to first nearest neighbor s-s, s-p_x, and p_x–p_x interactions. Based on these insights, we illustrate how the bonding interactions can be tuned to morph the silicon band structure into the germanium band structure, and how short-range ordering can produce a direct band gap in a specific sequence of Si-Ge layers. Our 3-step process offers a general framework to extract the crystal chemistry origins of electronic structure features from DFT calculations, enabling a new paradigm of *bonding-by-design*.

Introduction

Silicon has an indirect band gap, with the valence band maximum (VBM) at the Γ point and the conduction band minimum (CBM) at a low-symmetry point ~85% of the way between the Γ and X points. This indirect band gap determines the essential electronic and optical properties of silicon, and thereby its performance in photovoltaic and electronic devices.[1–3] Although the low-symmetry CBM of silicon is a basic fact of semiconductor physics, it is not so simple to answer *why* silicon has an indirect band gap. Similar 'why' questions can be generally raised about the electronic structures of materials. Why does germanium have a CBM at the L point, despite also being a Group IV semiconductor in the diamond structure? Why does zincblende GaAs have a direct band gap, with such a light electron effective mass? Understanding how crystal chemistry influences band structure would enable the rational design of band structure for thermoelectrics, *p*-type transparent conducting oxides, topological insulators, and other advanced electronic materials.[4–6] However, without a conceptual framework to approach *why* questions, one must rely on simple heuristics to post-rationalize the chemical origins of band structure. The computational search of next-generation optoelectronic materials can therefore only proceed by brute-force screening via a "needle-in-a-haystack" approach [7–9], rather than by rational and intuitive design.

Roald Hoffman presented a beautiful theoretical framework to examine how physics and chemistry meet in the solid-state (10, 11), arguing that chemists approach electronic structure from a bottom-up molecular hybridization approach, whereas physicists adopt a top-down planewave interpretation of electronic structure. Although Hoffman provided instructive tight-binding interpretations for how 1D periodic bonding interactions lead to *k*-dependent band structures, his insights are difficult to extrapolate to crystals in three dimensions, as 3D atomic environments offers many more potential orbital interactions than 1D crystals. In particular, in three dimensions there are more geometric degrees-of-freedom for bonding, there are more competing nearest-neighbor environments, and the phase-factor modulation becomes 3-dimensional in *k*-space—making it difficult to deconvolute the contributions of various orbital interactions to the overall band structure.

Given this complexity, creating tight-binding models in 3D structures requires a fundamental trade-off between simplicity and interpretability. Simple tight-binding models are usually constrained to first nearest neighbors (1NN), where one first asserts *a priori* which bonds are considered. While the resulting models are usually easy to interpret, they may not be physically robust. In the case of silicon, achieving a CBM near the X point using 1NN tight-binding has relied on including an $s*$ and/or d states beyond the standard sp atomic orbitals, which are arguably dubious in their physical relevance, as they do not reproduce or describe the precise physics of the actual excited states.[12–19] On the other hand, tight-binding models that consider further atomic neighbors, such as by interpolation from DFT band structures or fitting with many-NN, increases the accuracy of a tight-binding model but the model combinatorically explodes in the number of terms—precluding chemical interpretability[20–31]. This simplicity vs. interpretability trade-off is also a central question in the machine-learning community, where powerful black-box models fail to deliver physical insights on their predictions.[32, 33]

This paper aims to bridge the chemical intuition of Hoffman and the practical toolkit of DFT, so that we can better interpret the chemical origins of electronic structure in real materials. To do so, we start from a tight-binding interpolation of the DFT-calculated electronic structure using maximally localized Wannier functions (MLWFs).[34] From this tight-binding interpolation, we chemically interpret how band shapes in *k*-space derive from orbital interactions in real space. Our three-step process proceeds by determining: **1)** *Which orbitals contribute to a band*—specifically, what are the orbital characters

(eigenvectors) that contribute to the wavefunction. **2)** *How orbitals bond across* k-*space*—where the real or imaginary nature of the phase factor ($e^{ik \cdot r}$) dictates the allowed bonding/antibonding/non-bonding interactions of orbitals in real space. **3)** *How strongly the orbitals bond*—determined by the magnitudes of the tight-binding hopping parameters. These three steps systematically sieve through hundreds of TB parameters to build a sparse and physically robust model to interpret chemical bonding contributions to the electronic structure.

Here, we begin with an illustrative 1D model system to emphasize the impact of multi-orbital and >1NN interactions on band structure. Then we apply our three-step process to build a concise and chemically robust tight-binding interpretation for how multiple orbital interactions combine to form a low-symmetry CBM in Silicon. Specifically, we find that this low-symmetry CBM along the Γ-X line derives primarily from a cosine shape from 2NN p_x-p_x bonds, combined with a linear shape near X from 1NN orbital bonds. Finally, we present a new computational tool (https://viz.whsunresearch.group/tb/) to interactively execute our three-step process for any band and *k*-point in an electronic band structure. This tool enables us to precisely identify which bonds affect which segments of the silicon band structure. We then rationally tune these orbital interactions to morph the silicon band structure towards the germanium band structure, in accordance with the actual chemical differences between Si and Ge; and also rationalize how a specific Si-Ge short range ordering can promote a direct band gap. Altogether, our work serves as a general blueprint to extract the crystal chemistry origins of electronic band structure, and provides a pathway for rational band-structure engineering by chemical and structural design.

Uncovering the chemistry in band structure

All electronic properties of a material—including band gap, effective mass, band extrema location, *etc.*—are characterized by its band structure. The chemical origin of these electronic properties can be elucidated from accurate tight-binding interpolations of DFT-calculated electronic structure. Tight-binding decomposes a band structure into a summation of terms, similar to a Fourier decomposition, but with basis functions that correspond to physically-relevant orbital interactions.(35) This decomposition enables one to trace back which bonds manifest which specific band features of interest.

In tight-binding, each band dispersion (shape in *E-k* space) is shaped from numerous pairwise bonds between atomic orbitals. In **Equation 1**, the overall shape of the band, $E_n(\boldsymbol{k})$ for band *n*, results from the sum of bond energies $E_{n,\alpha\beta}(\boldsymbol{k})$—which is the *E-k* shape of the bond between two Bloch orbitals α and β.

$$E_n(\boldsymbol{k}) = \langle \Psi_{nk}|\widehat{H}|\Psi_{nk}\rangle = \sum_\alpha \sum_\beta c_{n\alpha}^{k}{}^{\dagger} c_{n\beta}^{k} \langle \Phi_\alpha(\boldsymbol{k})|\widehat{H}|\Phi_\beta(\boldsymbol{k})\rangle = \sum_\alpha \sum_\beta E_{n,\alpha\beta}(\boldsymbol{k}) \qquad \textbf{Eqn. 1}$$

This equation substitutes the wavefunction, Ψ_{nk}, as the sum of atomic Bloch orbitals, Φ_α, multiplied by their eigenvector weights, $c_{n\alpha}^{k}$. From this, the bond energy, $E_{n,\alpha\beta}(\boldsymbol{k})$, is the bond weight, $c_{n\alpha}^{k}{}^{\dagger} c_{n\beta}^{k}$, multiplied by the *bond run*, which we name the Hamiltonian matrix element between Bloch orbitals α and β, $\langle \Phi_\alpha(\boldsymbol{k})|\widehat{H}|\Phi_\beta(\boldsymbol{k})\rangle$. The name 'bond run' is inspired by Hoffman's discussion(10) that bands made from *s* orbital bonds 'run up' in energy from Γ to the Brillouin zone edge (from bonding to antibonding), while bands from *p* orbital bonds 'run down' from Γ to the Brillouin zone edge (from antibonding to bonding).

To isolate the impact of atomic orbital bonds, the bond run—written as $H_{\alpha\beta}(\boldsymbol{k})$ for brevity—can be reduced by substituting the Bloch orbitals as the sum of k-modulated atomic orbitals. This results in **Equation 2**, which is a sum over bonds between an atomic orbital α at $\boldsymbol{r}_\alpha$ and atomic orbital β at $\boldsymbol{r}_\beta + \boldsymbol{R}$, where $\boldsymbol{R}$ is a translation vector. Each bond is the tight-binding hopping parameter $V_{\alpha\beta}(\boldsymbol{R})$—the bond strength at $k=0$, multiplied by a phase factor, which modulates the bond strength for given k.

$$\langle \Phi_\alpha(\boldsymbol{k})|\widehat{H}|\Phi_\beta(\boldsymbol{k})\rangle \equiv H_{\alpha\beta}(\boldsymbol{k}) = \sum_{\boldsymbol{R}} V_{\alpha\beta}(\boldsymbol{R})\, e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_\alpha - \boldsymbol{r}_\beta - \boldsymbol{R})} \qquad \textbf{Eqn. 2}$$

Here, the k-dependence of a bond run arises from the sum of phase factors $e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_\alpha - \boldsymbol{r}_\beta - \boldsymbol{R})}$, where $\boldsymbol{r}_\alpha$ and $\boldsymbol{r}_\beta$ are atomic orbital centers in the primitive unit cell. The sum of these weighted phase factors over translation vectors $\boldsymbol{R}$ dictate the sign and shape of the bond run—where negative sums indicate bonding at that k-point, and positive indicates anti-bonding. In general, the bond run can be of any complex phase, but when multiplied with the eigenvectors $c_{n\alpha}^{k}{}^{\dagger} c_{n\beta}^{k}$, must be real.

Our three-step process is grounded in these two fundamental equations. **Step 1** is bond weight, which quantifies the orbitals that contribute to a band. **Step 2** is the shape of the bond run, which describes how the bond-type (bonding, antibonding, or non-bonding) changes across k-space. While **Step 3** is the amplitude of the bond run, dictated by the strength of the bond between two atomic orbitals α and β, $V_{\alpha\beta}(\boldsymbol{R})$.

Eqn. 1 is the same theoretical starting point as the crystal orbital Hamiltonian population (36, 37) method, but the implementation of an analytical representation in **Eqn. 2** (instead of a numerical calculation, for example as done through LOBSTER(38)) requires a tight-binding interpolation, which enables us to then break down the contributions from individual bonds. These analytical representations of each orbital wavefunction lets us visualize and further separate into NN or long-range interactions.

To examine how multiple orbitals and further nearest-neighbor interactions manifest in both real space wavefunction and the reciprocal space bands, here we present an illustrative example on a one-dimensional monatomic chain, with two orbitals (s and p) per atom. Traditionally, 2NN and further terms are neglected in simple TB models, but here we show how they could induce major qualitative changes to the band structure. The details of the derivation are in **SI.1**.

Plotted in **Figure 1a**, the 1NN bond runs H_{ss} and H_{pp} are cosine curves with extrema at the high symmetry points, but all other bond runs (H_{sp} and 2NN bonds) have extrema at low symmetry points. Using **Eqn. 1**, we sum together the bond runs multiplied by their bond weight (orbital character) to find the band energy. For only 1NN, $E_2(k)$ is dominated by $H_{pp}(k)$, which has extrema at high symmetry points. We next add 2NN interactions, with an exaggerated bond strength of half the 1NN, which results in a qualitative change in the overall band shape. This additional strong 2NN interaction pulls the bands together near $X/2$ while pushing them apart near Γ and X, creating a low-symmetry band extremum reminiscent of silicon.

When we ground our tight-binding intuition in 1D models, the 2NNs are so far away that their contributions tend to be small. However, in 3D crystals, atoms have much higher coordination numbers. These 2NNs are also much closer in 3D crystals than they are in 1D systems, meaning the 2NN contribution to the tight-binding interactions can be substantial. In the case of silicon, the twelve 2NNs are only 1.6× further than the four 1NNs, with a geometry that allows for strong overlap between p orbitals. By studying silicon, we will show how bonds of similar strength with different frequency of bond runs leads to band extrema away from high symmetry points. Thus, if there is a low symmetry band extremum on a band that does not have a band inversion, long-range interactions beyond 1NN are a likely culprit.

Figure 1: Multiple bonds combine to form band energy with a monoatomic s+p 1D tight-binding model. When adding 2NNs, band extrema at low-symmetry k-points manifest. *In (a), the bond runs, weights, and energies are plotted for onsite (orbital energy), 1NN, and 2NN bond contributions to the second band, E_2. The bond energies sum to create the band energy, where the orange/blue projection is the orbital character, $|c_{s/p}|$. To make them real, $iH_{sp}(k)$ and $-ic_{ns}^{k\dagger}c_{np}^k$ are plotted for the sp bonds. On the right, 2NN terms are added, perturbing the bond weights and band energies to create a low symmetry extremum off-X. In (b), the real part of the $\Psi_2^{\frac{X}{2}}$ wavefunction is plotted to highlight the real-space bonding implicit in band structure. The orbital at each atom is determined by the phase factor, $e^{i\frac{X}{2}x}$, multiplied by the orbital coefficients, $c_s = 0.49$ and $c_p = -0.87i$. The nonzero bond energies are written and circled in red on the reciprocal-space plots.*

In our ambition to eventually design band structure from the underlying bonds, we need to first elucidate the orbital and bonding nature at the specific k-point of a band. The key term linking real space chemical bonding with reciprocal space bandstructure is the phase factor $e^{i\mathbf{k}\cdot\mathbf{R}}$, manifesting in the wavefunction as $\Psi_n^k(\mathbf{r}) = \sum_{\alpha,R} e^{i\mathbf{k}\cdot\mathbf{R}_\alpha} c_{n\alpha}^k \phi_\alpha(r - R_\alpha)$. For a given k-vector, the combined coefficient $e^{i\mathbf{k}\cdot\mathbf{R}_\alpha} c_{n\alpha}^k$ for each atomic orbital α dictates the bond type (bonding, antibonding, or non-bonding) between atomic orbitals in real space **(Step 2)**. If the complex phases of neighboring orbitals are orthogonal (e.g. real and imaginary), they do not interact and are non-bonding. If they are nonorthogonal (e.g. real and real, or imaginary and imaginary), they are bonding or antibonding, depending on the signs of the wavefunction.

In Figure 1b, we illustrate the relationship between the phase factor and the real-space $\Psi_2^{\frac{X}{2}}$ wavefunction at $k = X/2$, which corresponds to a 4-unit cell superstructure in real space. Across the 4 atoms in **Figure 1b**, the phase factor $e^{i\frac{X}{2}\cdot\mathbf{R}}$ modulates as +1, +i, –1, and –i. Because c_2^k coefficients are real and positive for s orbitals, whereas they are imaginary and negative for p orbitals, we see that the real part of $\Psi_2^{\frac{X}{2}}$ has s orbitals on atoms 1 and 3; and p orbitals on atoms 2 and 4. Therefore, the only 1NN interaction is *s-p* antibonding. The sign switching between atoms 1 and 3 gives 2NN *s-s* antibonding, and the sign switching between atoms 2 and 4 give 2NN *p-p* bonding. This example visualizes how the chemical bonding in real space implicitly derives from each k-point of a band in reciprocal space.

Detangling the Silicon band structure

Although silicon has been studied for decades, the crystal chemistry origins of its low-symmetry conduction band minimum still lack satisfactory explanation. Tight-binding (TB) models fitted with only 1NN incorrectly produce a conduction band minimum at Γ. (12, 13) While Vogl produced a CBM off-X with an additional excited s^* state, he acknowledges that "The inclusion of some such excited states in any minimal basis set is physically important–although the precise physics of the actual excited states need not be faithfully and quantitatively reproduced."(14) Indeed, while the sps^* model fits the Γ-X line, it sacrifices the accuracy of the rest of the conduction band structure along nearly every other k-path—especially the Λ_2, Σ_3, and Σ_1 bands and the W_1 conduction band points, as detailed further in **SI.2**.

Since band structure arises from the complex interactions between multiple orbitals, it is often possible to have multiple non-unique solutions that fit a singular band feature. Following Vogl, others have included additional orbital states to their tight-binding models—for example Jancu et. al. and others added d states, producing a reasonable fit of the lowest conduction bands but again with little physical insight regarding the additional parameters.(15–19) Tight-binding models with >1NN were also fit (often with s^* states) for silicon and zinc-blende semiconductors with varying degrees of 2NN contribution.(20–28, 31)

From a model-building perspective, it is not satisfactory to include terms *ad hoc* just to match a single band structure feature—rather, a term that is physically valid should improve the fit of all band energies throughout the entire Brillouin Zone. This is especially important if one aims to later engineer and design the band structure by modifying chemical interactions, which requires one to accurately identify the true chemical origin of band features.

At the other end of the spectrum, one can perform a tight-binding interpolation directly from DFT—which obtains the tight-binding parameters from a Fourier transformation of the k-dependent orbital Hamiltonian. This tight-binding interpolation circumvents the need to assume which interactions are present, but the resulting many-NN TB models can have hundreds of non-trivial interactions, which is too complex to interpret chemically. Sanchez-Portal, and later Qian et al, applied a TB interpolation which includes many NNs to silicon finding a low symmetry minimum along the Γ-X line with only a sp basis, indicating that s^* and d states are not required to reproduce the minimum off X.(29, 39) Since then, a reasonable TB interpolation of silicon has been achieved many times using MLWFs and other methods, but a simple chemical understanding has still not yet been detangled from the hundreds of hopping parameters found from the interpolation.

Here, we apply our three-step process to build a chemical interpretation for the low symmetry conduction band minimum in silicon along the Δ_1 band. Details of our MLWF parameters and process are discussed in **SI.3**. Our 3-step process is implemented as follows: **Step 1**) We determine which orbitals contribute to the Δ_1 band—finding that it is >80% p_x orbital character, with the remaining character being s orbitals. **Step 2**), we determine how the orbitals bond across k-space—finding from the bond runs that the second nearest neighbor p_x-p_x is the only interaction that decreases the band energy at the X point. **Step 3**) we determine how strongly the orbitals bond—showing that the 2NN p_x–p_x bond has large hopping parameters and high coordination which makes it a significant influence on the band structure.

Finally, individual chemical bonds are assessed for their contribution to the total shape of the Δ_1 band. From this, we determine that the low-symmetry conduction band minimum of silicon manifests from a linear shape of the 1NNs near X, combined with the cosine shape of the 2NN p_x–p_x bond. Crucially, including the 2NN p_x–p_x bond not only improves the Δ_1 band, but the band structure across all other high-symmetry lines (details in **SI.2**), validating its physical significance in creating the low-symmetry Si CBM.

1) *Orbital character of bands*

First, we determine which orbitals in the Δ_1 band are present to bond. In a *sp* model, silicon in the diamond structure has eight orbitals, four for each of the two atoms in the primitive cell. This amounts to 72 Hamiltonian matrix elements—8 onsite, 32 1NN, and 32 2NN interactions. After symmetry and group theory considerations, a wavefunction along the Γ–X line will have either $s + p_x$ orbitals or $p_y + p_z$ orbitals. With only s and p_x orbitals are on the Δ_1 band, the 72 matrix elements can be reduced to 8 unique elements.

To separate the character of these bands, it is important to use atomic orbitals. If sp^3 hybridized orbitals were used, it would be difficult to deconvolute the mixing of all four orbitals from each atom. While hybridized orbitals can serve well in simple molecular systems, momentum-dependent crystal wavefunctions rarely reduce to the hybridized symmetry of simple molecular wavefunctions and require the generalized framework of atomic orbitals.

Figure 2 plots positive (red) and negative (green) isosurfaces for the real part of the complex wavefunctions in one of each of the four doubly-degenerate bands at X. For all the X_1 bands, the first atom in the basis set has only p_x orbitals and the second atom has only (distorted) s orbitals. Away from X, the X_1 degeneracy splits into Δ_1 (CBM band) and Δ_2 bands with s and p_x orbital character, while the X_4 bands remain degenerate. In Silicon, tight-binding analysis of the Δ_1 conduction band character shows it is predominantly (>80%) p_x orbital character. The X_1 antibonding wavefunction is mainly 1NN s-p_x antibonding and 2NN p_x–p_x bonding. Importantly, the 2NN p_x–p_x interaction is the only one that is bonding along the x-direction and contributes to lowering the energy at X.

Figure 2: The crystal wavefunctions at the X point in silicon with the X_1 conduction band highlighted to show orbitals and bond type. Each of the four doubly degenerate bands is accompanied by the present 1NN bond and the plotted real part of a wavefunction. The atom sites are spheres colored to indicate the z-coordinate. The red and green show the positive and negative isosurfaces of the real wavefunction, where neighboring same color lobes are bonding and different color lobes are antibonding. Bonding lobes often mesh together while antibonding lobes are distorted apart. By looking closer at the X_1 band, we determine 1NN and 2NN bond-types, where the 2NN p_x – p_x interaction is the only bonding along the x-direction.

Although these arguments explain the lowering of the Δ_1 band energy *at* X, the actual CBM is at a low-symmetry point ~85% of the way from Γ to X. The corresponding real-space wavefunction for at this k-point in fact spans multiple unit cells. We next visualize the real-space bonding across this long-range wavefunction, so that we can directly observe these critical 2NN interactions.

2) *Visualizing the k-dependence of chemical bonds*

At a low symmetry k-point, the long-range wavefunction in real space modulates between bonding, non-bonding, and anti-bonding. The bond-type depends on the k-point modulation of the atomic orbitals by the phase factor $e^{ik \cdot R_i}$. When constructing the electron density by $\psi^*\psi$, this phase factor is lost, meaning that chemical bonding (which results when atomic wavefunctions of the same phase overlap) is not a quantum mechanical observable. Hence, methods to evaluate chemical bonds using the relative energies and slopes of the charge density are necessarily indirect.

On the other hand, by constructing the supercell corresponding to a low-symmetry k-point, and then plotting the long-range wavefunction, we visualize chemical bonding directly. Bonding is then determined by the phase and sign of the orbitals. That is, if neighboring orbitals have orthogonal phase (e.g. real and imaginary), they do not interact and are nonbonding. Whereas when neighboring orbitals have nonorthogonal phase (e.g. real and real), they will be bonding for same-sign isosurface lobes or antibonding for opposite-sign lobes. This analysis is the same as we used for the model system in **Figure 1b**, but can now be applied to DFT calculations of materials. This brings the crystal bonding schematics used by Woodward (40), Snyder (41), and others (10, 42, 43) to a computable and generalizable scale.

This visualization is applied to Silicon in **Figure 3**, where the smallest supercell of the real crystal wavefunctions for three different k-points are plotted. Along the Δ_1 band the s orbital coefficient is real and positive, while the p_x orbital coefficient is imaginary and negative. To create the full wavefunction, the orbital coefficient is multiplied by the phase factor to determine the phase of the orbital on atom at center $\boldsymbol{R}$ as $c_\alpha^k e^{ik \cdot \boldsymbol{R}}$ which is expanded for the Δ_1 band as $\left(\left|c_s^k\right| + i\left|c_p^k\right|\right)(\cos(\boldsymbol{k} \cdot \boldsymbol{R}) + i\sin(\boldsymbol{k} \cdot \boldsymbol{R}))$. Now, the oscillation of orbital character between s and p_x in the real wavefunction is seen as the s orbital changes as $\left|c_s^k\right|\cos(\boldsymbol{k} \cdot \boldsymbol{R})$ and the p_x orbital as $-\left|c_p^k\right|\sin(\boldsymbol{k} \cdot \boldsymbol{R})$. We plot this in the 'Total' column of **Fig 3** with vertical lines, where the color represents the s vs. p_x character on the atom.

Figure 3: Crystal wavefunctions describe variation in 2NN bond-types along the Δ_1 band. The orbital on an atom is colored to indicate s orbitals (yellow), sp_x hybridized orbitals (blue), and p_x orbitals (purple). Atom 1 of the two-atom basis is color white while Atom 2 is gray. Positive (green) and negative (red) isosurfaces of the real crystal wavefunction in their smallest repeatable supercell are plotted to inspect bond-type. The total column shows how the phase factor dictates orbital character and the other columns allow for visualizing the 2NN bond type.

To isolate the 2NN bonds, Figure 3 shows the orbitals for Atom 1 (white) and Atom 2 (gray) of the silicon primitive cell separately. The real-space oscillation in the x-direction of the conventional unit cell (corresponding to the {110} direction of the primitive cell, visualized here) reveals that only bonds with an interatomic displacement along the x-direction change bond-type along the $\Gamma - X$ line. All four 1NN around an atom have the same x-translation of $|0.25a|$, while the twelve 2NNs split into eight with an x-translation of $|0.5a|$ and four with no x-translation.

The plotted wavefunctions can be used to visually determine bond-type at specific k-points. As an example, we describe some interactions for the 2NN with an x-translation. At $k = X$, the alternating sign for the p_x orbitals on Atom 2 allows [red + red] or [green + green] lobes to be nearest, resulting in bonding. At $k = 0.87X$, the supercell size increases, each atom has a complicated mixture of orbitals, the wavelike oscillation between s and p_x orbital character is most apparent (caused by the orbital coefficients being different phases), and the p_x orbitals are bonding on average. At $k = \frac{1}{2}X$, the Atom 1 switches between s and p_x orbitals, while Atom 2 maintains sp_x orbitals. While the s orbitals on Atom 1 do not appear at this isosurface value, their sign yields antibonding. For Atom 2, the red and green sp_x lobes are always closest, also yielding antibonding. While not pictured, at $k = \Gamma$, s-s is bonding, s-p is non-bonding, and p-p is antibonding, for both 1NN and 2NN. All the s-s and s-p interactions for both 1NN and 2NN lead to a higher energy at the X point than at Γ, with the major exception of the 2NN p_x–p_x interaction—which sweeps from most antibonding at Γ to most bonding at X, resulting in a lower band energy at X.

3) *Strength of hopping parameters*

Based on intuition built from 1D models, we would anticipate the hopping parameters for 1NNs to generally be much larger than for 2NNs. However, here we find that the 2NN bonds are very important in silicon, as shown in **Table 1** by the hopping parameters from our atomic-like Wannier TB interpolation. When including bond multiplicity, the twelve 2NN p_x–p_x parameters sum as $8V_{xx}(110) + 4V_{xx}(011) = 2.00\text{eV}$, which is 4× larger than the four 1NN p_x–p_x parameters $V_{xx}(\frac{1}{2}\frac{1}{2}\frac{1}{2})$. Combined with the dominant orbital character being p_x, the 2NN p_x–p_x contributes significantly to the energy of the Δ_1 wavefunctions. Other tight-binding models which have included 2NN parameters either did not include a $V_{xx}(110)$ term, or they were ~10× smaller than our MLWF-derived result.(20–28) An exception to this is Grosso and Piermarocchi who fit a $V_{xx}(110)$ about 2× larger than our result in **Table 1**.(31) In all these cases, the 2NN contributions have not been individually analyzed for their role in shaping the CBM.

Table I. Silicon hopping parameters (eV) from MLWF used to reconstruct the Δ_1 band.

ε_s	ε_p	$V_{ss}(\frac{1}{2}\frac{1}{2}\frac{1}{2})$	$V_{sx}(\frac{1}{2}\frac{1}{2}\frac{1}{2})$	$V_{xx}(\frac{1}{2}\frac{1}{2}\frac{1}{2})$	$V_{xx}(110)$	$V_{xx}(011)$
-5.467	1.650	-1.639	1.075	0.126	0.117	-0.267

The small hopping parameter between 1NN p_x–p_x orbitals can be understood from the geometry as discussed by Slater.(44) The hopping parameter between two p_x orbitals is given by $l^2 V_{pp\sigma} + (1 - l^2)V_{pp\pi}$, where l is the direction cosine in the x direction. If the p_x-p_x lobes are facing each other (like ∞-∞) then there is perfect σ bonding, where $l = 1$, whereas if p_x-p_x lobes are parallel (like 8-8) then $l = 0$ and there is perfect π bonding. Because $V_{pp\sigma}$ and $V_{pp\pi}$ have opposite sign, an intermediate orientation between perfectly aligned (∞-∞) and perfectly parallel (8-8), will result in the hopping parameter canceling to zero.

In the tetrahedral coordination environment of the diamond structure, 1NN have $l^2 = \frac{1}{3}$, such that the $V_{pp\sigma}$ and $V_{pp\pi}$ components nearly cancel. The 2NN have four neighbors with $l^2 = 0$, meaning these $V_{xx}(011)$ parameters are entirely π bonding, while the other eight neighbors have $l^2 = \frac{1}{2}$, allowing the stronger σ antibonding to compensate for the weaker π bonding in the $V_{xx}(110)$ parameters. In the next section we put together how the orbital character, bond run, and bond strength combine to form the total band dispersion.

Steps 1+2+3: The low-symmetry Δ_1 minimum in silicon

Finally, we rationalize the total shape of the silicon Δ_1 band dispersion from individual bonds and determine which bonds are necessary to capture the correct qualitative band shape. Each bond contributes a distinct shape over some high-symmetry line of the band structure. To obtain the correct band dispersion that matches a DFT band structure, all significant bonds must be included. As Vogl showed, a 1NN tight-binding model with an sp basis cannot produce a low-symmetry conduction band minimum(14), whereas an many-NN sp basis can (29, 39). This indicates that physically significant bonds are missing from the 1NN sp model, which as we have argued, are the 2NN interactions.

To better account for how a bond contributes to the energy of a given band, we introduce a band-dependent bond run $H_{n,\alpha\beta}(k)$, **Eqn. 4**. This multiplies the Hamiltonian element given by **Eqn. 3** by the phase of the orbital coefficients, ensuring that $H_{n,\alpha\beta}$ is real and the correct sign for band n. Finally, we write the energy that a bond between orbitals α and β contributes to band n as $H_{n,\alpha\beta}(k)$ multiplied by the absolute value of the orbital coefficients, **Eqn. 5**, which follows clearly from **Eqns. 1** and **4**.

$$H_{n,\alpha\beta}(k) \equiv \frac{{c_{n\alpha}^{k}}^{\dagger} c_{n\beta}^{k}}{|c_{n\alpha}^{k}||c_{n\beta}^{k}|} H_{\alpha\beta}(k) \qquad\qquad \text{Eqn. 4}$$

$$E_{n,\alpha\beta}(k) = |c_{n\alpha}^{k}||c_{n\beta}^{k}| H_{n,\alpha\beta}(k) \qquad\qquad \text{Eqn. 5}$$

The Δ_1 band energy as a function of the bonds can be simplified from the sum over each matrix element, **Eqn. 1**, to a sum over unique elements for the s and p_x (written as x for simplicity) orbitals, **Eqn. 6**. As discussed earlier, this reduces the 72 parameters for the $\Gamma - X$ line to only 8: two onsite (orbital energy) terms, three 1NN terms, and three 2NN terms.

$$E_{\Delta_1}(k) = 2E_{\Delta_1,ss}^{0} + 2E_{\Delta_1,xx}^{0} + 2E_{\Delta_1,ss}^{1NN} + 4E_{\Delta_1,sx}^{1NN} + 2E_{\Delta_1,xx}^{1NN} + 2E_{\Delta_1,ss}^{2NN} + 4E_{\Delta_1,sx}^{2NN} + 2E_{\Delta_1,xx}^{2NN} \qquad \text{Eqn. 6}$$

Each term has an analytical expression as seen from **Eqn. 5**, which is the product of relevant orbital coefficients (a multiplicative factor) with the band-dependent bond run (a cosine or sine shape). For example, the onsite p_x orbital term is $E_{\Delta_1,xx}^{0} = \varepsilon_0^x |c_{\Delta_1,x}^{k}||c_{\Delta_1,x}^{k}|$ and the important 2NN p_x-p_x orbital bonding term is $E_{\Delta_1,xx}^{2NN} = |c_{\Delta_1,x}^{k}||c_{\Delta_1,x}^{k}| \cdot [8V_{xx}(110)\cos(k_x\pi) + 4V_{xx}(011)]$, where ε_0^x is the atomic p_x-like orbital energy, $c_{\Delta_1,x}^{k}$ is the p_x orbital coefficient **(Step 1)**, and $8V_{xx}(110)\cos(k_x\pi) - 4V_{xx}(011)$ is the bond run $H_{\Delta_1,xx}^{2NN}(k)$ which shows the 2NN p_x-p_x bond-type (energy) for the Δ_1 band **(Step 2 and 3)**. Crucially, the bond run for the $V_{xx}(110)$ hopping parameter is a positive cosine, lowering the energy at X. For full derivation and decomposition of **Eqn. 6**, see **SI.4**.

Figure 4: Deconstructing how each bond contributes to the Δ_1 conduction band in silicon by plotting the band-dependent bond runs $H_{\Delta_1,ij}$ and corresponding energies $E_{\Delta_1,ij}$ for the onsite (atomic orbital energy), 1NN, and 2NN bonds. *Interactions between p_x-p_x, s-p_x, and s-s orbitals are colored purple, teal, and yellow, respectively, which corresponds with the orbital character color bar used in the total energy plots E_{Δ_1}. Altogether, the linear behavior near X achieved with onsite + 1NN bonds and the cosine shape of the 2NN p_x-p_x bond combine to form the minimum near X. The gray lines in the bottom right plot show the Silicon band structure with all onsite, 1NN, and 2NN bonds, where the small error between the colored and gray Δ_1 band results from including the 2NN s-p_x and s-s bonds.*

In **Figure 4**, we plot the band-dependent bond runs $H_{\Delta_1,\alpha\beta}$ and the energy from each bond $E_{\Delta_1,\alpha\beta}$ for the onsite, 1NN, and 2NN interactions. The bond run shapes are similar to the illustrative 1D example from **Figure 1b**, but with half the length in reciprocal space since silicon has a two-atom primitive cell. The bond run magnitudes are dictated by the hopping parameter.

When looking only at the bond runs, the onsite and 1NN terms dominate the 2NN, with the 1NN s–p_x and s–s spanning ~8 eV each, while the 2NN p_x–p_x only reaches 2 eV. But once the strong p_x orbital character is included with $E_{\Delta_1,\alpha\beta}$, the p_x–p_x interactions are nearly unchanged, while the s–p_x is decreased significantly, and the s–s drops nearly to zero. This puts the [onsite + 1NN] energy magnitude in the same range as the 2NN p_x– p_x, where each span ~1.5 eV. Now considering the shape, we find the onsite + 1NN combine to form a pseudo-linear increase near X. This results primarily from the $E_{\Delta_1,ss}^{1NN}$ shape which arises from the sine curve of $H_{\Delta_1,ss}^{1NN}$ (flat at X) which is heavily distorted (almost to a cosine-like curve) by the coefficient weight $\left|c_{\Delta_1 s}\right|\left|c_{\Delta_1 p}\right|$ increasing from Γ to X. The nonzero slope of the Δ_1 band at X is then transferred to the degenerate Δ_2 band.

Finally, the pseudo-linear shape of the onsite + 1NN near X plus the cosine curve of the 2NN p_x–p_x combine to form the conduction band minimum away from the high-symmetry X point in silicon. Importantly, including the 2NN p_x– p_x bond also provides a good band structure fit on all other k-paths (see **SI.2** for details), compared to the s^* state from Vogl *et al.*(14), which validates the physical importance of 2NN bonding in governing the low-symmetry conduction band minimum of silicon.

Towards Bonding-by-Design

Band engineering for solar cells, semiconductors, and thermoelectrics frequently requires control over the energy level of bands at specific *k*-points. Because we now have a theoretical pathway to connect the bonding interactions to the band structure, we can examine the *inverse* electronic structure design problem—*How can I modify chemical interactions to morph an existing band structure to a new band structure with more desirable features*? As a representative example, here we will modify the bonding interactions to shift the CBM from the Γ-X line in silicon to the L point as it is in germanium. We then rationalize the short range ordering in d'Azerac's 'magic sequence' of Si-Ge layers that led to a direct band gap.(45) This illustration paves the way to a vision of *bonding-by-design*, where instead of searching for pristine materials with a given band structure feature, we can rationally tune the chemistry (by substitutional doping or alloying) to morph a given band structure towards a desired one.

Thoroughly analyzing a band structure feature is arduous, which motivated us to create a computational analysis package (https://viz.whsunresearch.group/tb/) which systematically executes our three-step process. Our package features an interactive interface that populates tables with the orbital character and important bonds for any selected point of the $E(k)$ diagram. In addition, the band-dependent bond runs and bond energies for any of the important bonds can be plotted upon selection, allowing a user to rapidly discern how each bond contributes to the band shape. A detailed explanation and tutorial are provided in **SI.5** and **SI.6**. Here we use our package to demonstrate how different segments of a band can be selectively raised or lowered towards a desired shape by modifying a single bond.

As illustrated in **Fig 5a** and discussed previously, the 2NN p_x-p_x bond contributes a shifted cosine shape to the Δ_1 band. Thus, strengthening the 2NN p_x-p_x bond lowers the corresponding off–X minimum. **Figure 5b** shows that by increasing the 1NN p–p bond strength, the L point moves down. Increasing the 1NN p-p bond affects the L point because the L point in reciprocal space corresponds to the [111] direction in real space, and in Silicon this (111) direction corresponds to a nearest-neighbor atom from the tetrahedral coordination environment. The band character of the $L - \Gamma$ line is split ~55% *s* orbital and ~45% evenly amongst the *p* orbitals, a combination which corresponds to the sp^3 hybridized orbital in the (111) direction with variable *s* character. Using our analysis tool, we found that increasing 1NN p–p bond strength moves the L point down in energy below the Γ and X points, which can make the CBM at the L point.

In **Figure 5c**, we examine the conduction band at the Γ point, which is entirely *s* orbital character. The *s* character increases from ~55% to 100% in the last tenth of the line from L to Γ—this dramatic change in orbital character contributes to the sharp curvature near Γ, which leads to low effective masses in direct-gap tetrahedral semiconductors like GaAs. Consequently, reducing the 1NN s–s bond strength lowers the energy of the band at Γ, which is the *s* orbital antibonding state.

Altogether, it is possible to morph the Si band structure towards the Ge band structure by increasing bonding between *p* orbitals while decreasing bonding of *s* orbitals. This effect is qualitatively consistent with changing chemistry from Si to Ge. The occupied *d* shell in Ge allows more partially screened nuclear charge to attract the valence shell—an effect called scandide contraction—which disproportionately impacts the *s* orbital near the nucleus, thereby reducing the *s* orbital radius in Ge compared to the *p* orbital radii. Yuan *et al*. also found that the *d* orbitals are important in changing CBM location (46), and here we see that this is due to the indirect screening for the *s* and *p* orbitals. Thus, when augmenting Si by increasing all 1NN p–p interactions by 30% and decreasing all 1NN s–p and s–s bonds by 15%, we reproduce the characteristic band structure of Ge with a CBM at the L-point. Further discussion of the Si and Ge MLWF band structure can be found in **SI.7**.

Figure 5: The role of chemical bonding in band extrema. In (a-d) the original Si band structure (gray) is plotted against augmented band structures (red), which changes one bond to selectively lower X, L, or Γ (plots a, b, and c, respectively) or changes bonds based on the chemistry of Ge (plot d). In e, Zunger's magic sequence of Si (blue) and Ge (gray) layers is plotted (unrelaxed) with the decomposition of Si-Si and Si-Ge bonds for 1NN, 2NN, 3NN, and 4NN. The table reveals how the strength of each NN bond should change to increase directness and is highlighted for whether Si-Ge or Si-Si would achieve the desired qualitative bond strength based on their different chemistry.

These conceptual insights into how different orbitals affect the bandstructure can also qualitatively guide the design of short-range alloy order to promote the frequency of certain bond-types over others (47). In a random solid-solution of $Si_{0.5}Ge_{0.5}$, the frequency of A-A bonds (e.g., Si-Si or Ge-Ge) and A-B bonds (Si-Ge) are equally 50% likely for all nearest-neighbor environments. But because each bond type and nearest-neighbor number can affect band extrema differently, the band structure can be selectively tuned by designing short-range order that preferences different fractions of nearest neighbor bond types.

For example, d'Avezac *et al.* used genetic algorithms to identify a 'magic sequence' of Si-Ge layers that results in a direct band gap Si_xGe_{1-x} alloy(45). Qualitatively, to build a direct band gap structure, the CBM has to be lowered at Γ while being raised at X and L. In the magic sequence structure, Si is bonded to Si 25% of the time for 1NNs, 33% for 2NNs and 3NNs, and 83% for 4NNs (**Fig 5e**). To understand how this short-range ordering leads to a direct band gap, we show in the table of **Fig 5e** whether a weaker or stronger orbital interaction for each nearest neighbor would improve directness (lower Γ while raising X

and L). For example, a weaker 1NN *s-s* interaction results in diminished 1NN *s-s* antibonding, lowering the energy of the CBM at Γ. On the other hand, stronger 2NN and 4NN *s-s* interactions lower the energy at Γ because their interactions exhibit more bonding in the CBM wavefunction. Based on our chemical intuition that Si-Si bonds have stronger *s-s* bonds and weaker *s-p* and *p-p* bonds relative to Si-Ge, we anticipate that to engineer a direct band gap, we should increase the relative fraction of Si-Ge bonds for 1NN and 3NN bonds, while increasing the fraction of Si-Si bonds for 2NN and 4NN—which is consistent with the magic sequence short-range ordering in d'Avezac *et al*. While the d'Avezac structure was difficult to realize experimentally due to strain between layers,(48) it is possible that this ratio of nearest-neighbor bonds can be found in a more energetically-favorable ordering that also yields a direct band gap.

Outlook

Here we presented a computable and chemically motivated framework that considers **1)** *Which orbitals are in a band*, **2)** *How are they allowed to bond*, and **3)** *How strongly do they bond?* This framework produces a sparse and therefore interpretable tight-binding model that can help us intuitively understand the crystal chemistry origins of band structure. When we applied our approach to silicon, we found that the low-symmetry conduction band minimum of silicon originates primarily from 2NN p_x–p_x bonds, which significantly lowers the energy at X. The significance of the 2NN p_x-p_x orbital bond compared to the 1NN is explained from the geometry of the bonding angles, in addition to there being 3× as many 2NN atoms than 1NN. This explanation is a revision on Vogl's *sps** model, which captures the CBM position in Silicon but at the expense of other conduction bands in the Brillouin zone. Our interpretation provides a clear physical mechanism compared to previous *sp* models with multiple NN.

Figure 6: In materials-by-design, the electronic properties of pristine materials are calculated from DFT. By inverting this paradigm to bonding-by-design, one starts with the desired band feature for a given application, and rationally tunes the crystal chemistry to achieve this band feature.

Broadly speaking, our approach allows us to clearly and robustly pinpoint the physical origin of electronic structure features in complex 3D crystals. This framework is general and can be applied to any tight-binding interpolation of a DFT-calculated band structure. By better understanding how crystal chemistry translates to major electronic structure features, we can more intuitively design chemistries and bonding environments to yield a desired band structure feature. A major advantage of this approach opportunity to search within the 'perturbation space' of a given material, allowing us to find best-in-class semiconductors which are often minor perturbations (strain, doping, alloying, *etc.*) from their pristine forms. This approach would invert the design paradigm from electronic 'materials-by-design' to the inverse approach of *bonding-by-design* (**Figure 6**)—where instead of searching for materials with specific properties, we can chemically or structurally modify the band structure of a given material to tune it towards next-generation electronic, optic, thermoelectric, and correlated quantum materials.

References

1. M. A. Green, Silicon photovoltaic modules: a brief history of the first 50 years. *Prog. Photovolt. Res. Appl.* **13**, 447–455 (2005).

2. A. I. Boukai, *et al.*, Silicon nanowires as efficient thermoelectric materials. *Nature* **451**, 168–171 (2008).

3. Y. Cui, Z. Zhong, D. Wang, W. U. Wang, C. M. Lieber, High Performance Silicon Nanowire Field Effect Transistors. *Nano Lett.* **3**, 149–152 (2003).

4. G. J. Snyder, E. S. Toberer, Complex thermoelectric materials. *Nat. Mater.* **7**, 105–114 (2008).

5. G. Brunin, F. Ricci, V.-A. Ha, G.-M. Rignanese, G. Hautier, Transparent conducting materials discovery using high-throughput computing. *Npj Comput. Mater.* **5**, 1–13 (2019).

6. M. Toriyama, G. J. Snyder, "Are Topological Insulators Promising Thermoelectrics?" (Chemistry, 2023).

7. G. Hautier, Finding the needle in the haystack: Materials discovery and design through computational ab initio high-throughput screening. *Comput. Mater. Sci.* **163**, 108–116 (2019).

8. S. Curtarolo, *et al.*, The high-throughput highway to computational materials design. *Nat. Mater.* **12**, 191–201 (2013).

9. W. Chen, *et al.*, Understanding thermoelectric properties from high-throughput calculations: trends, insights, and comparisons with experiment. *J. Mater. Chem. C* **4**, 4414–4426 (2016).

10. R. Hoffmann, How Chemistry and Physics Meet in the Solid State. *Angew. Chem. Int. Ed. Engl.* **26**, 846–878 (1987).

11. R. Hoffmann, *Solids and Surfaces: A Chemist's View of Bonding in Extended Structures* (John Wiley & Sons, 2021).

12. W. A. Harrison, Bond-Orbital Model and the Properties of Tetrahedrally Coordinated Solids. *Phys. Rev. B* **8**, 4487–4498 (1973).

13. D. J. Chadi, M. L. Cohen, Tight-binding calculations of the valence bands of diamond and zincblende crystals. *Phys. Status Solidi B* **68**, 405–419 (1975).

14. P. Vogl, H. P. Hjalmarson, J. D. Dow, A Semi-empirical tight-binding theory of the electronic structure of semiconductors†. *J. Phys. Chem. Solids* **44**, 365–378 (1983).

15. J.-M. Jancu, R. Scholz, F. Beltram, F. Bassani, Empirical spds * tight-binding calculation for cubic semiconductors: General method and material parameters. *Phys. Rev. B* **57**, 6493–6507 (1998).

16. S. Sapra, N. Shanthi, D. D. Sarma, Realistic tight-binding model for the electronic structure of II-VI semiconductors. *Phys. Rev. B* **66**, 205202 (2002).

17. T. B. Boykin, G. Klimeck, F. Oyafuso, Valence band effective-mass expressions in the sp 3 d 5 s * empirical tight-binding model applied to a Si and Ge parametrization. *Phys. Rev. B* **69**, 115201 (2004).

18. A. S. Martins, T. B. Boykin, G. Klimeck, B. Koiller, Conduction-band tight-binding description for Si applied to P donors. *Phys. Rev. B* **72**, 193204 (2005).

19. Y. M. Niquet, D. Rideau, C. Tavernier, H. Jaouen, X. Blase, Onsite matrix elements of the tight-binding Hamiltonian of a strained crystal: Application to silicon, germanium, and their alloys. *Phys. Rev. B* **79**, 245201 (2009).

20. G. Klimeck, *et al.*, Si tight-binding parameters from genetic algorithm fitting. *Superlattices Microstruct.* **27**, 77–88 (2000).

21. D. N. Talwar, C. S. Ting, Tight-binding calculations for the electronic structure of isolated vacancies and impurities in III-V compound semiconductors. *Phys. Rev. B* **25**, 2660–2680 (1982).

22. T. B. Boykin, Improved fits of the effective masses at Γ in the spin-orbit, second-nearest-neighbor sp 3 s * model: Results from analytic expressions. *Phys. Rev. B* **56**, 9613–9618 (1997).

23. J. P. Loehr, D. N. Talwar, Exact parameter relations and effective masses within sp 3 szinc-blende tight-binding models. *Phys. Rev. B* **55**, 4353–4359 (1997).

24. K. C. Hass, H. Ehrenreich, B. Velický, Electronic structure of Hg 1 − x Cd x Te. *Phys. Rev. B* **27**, 1088–1100 (1983).

25. W. A. Harrison, New tight-binding parameters for covalent solids obtained using Louie peripheral states. *Phys. Rev. B* **24**, 5835–5843 (1981).

26. K. E. Newman, J. D. Dow, Theory of deep impurities in silicon-germanium alloys. *Phys. Rev. B* **30**, 1929–1936 (1984).

27. T. J. Lenosky, *et al.*, Highly optimized tight-binding model of silicon. *Phys. Rev. B* **55**, 1528–1544 (1997).

28. Y. M. Niquet, C. Delerue, G. Allan, M. Lannoo, Method for tight-binding parametrization: Application to silicon nanostructures. *Phys. Rev. B* **62**, 5109–5116 (2000).

29. D. Sanchez-Portal, E. Artacho, J. M. Soler, Projection of plane-wave calculations into atomic orbitals. *Solid State Commun.* **95**, 685–690 (1995).

30. S. Ciraci, I. P. Batra, Electronic-energy-structure calculations of silicon and silicon dioxide using the extended tight-binding method. *Phys. Rev. B* **15**, 4923–4934 (1977).

31. G. Grosso, C. Piermarocchi, Tight-binding model and interactions scaling laws for silicon and germanium. *Phys. Rev. B* **51**, 16772–16777 (1995).

32. S. M. Moosavi, K. M. Jablonka, B. Smit, The Role of Machine Learning in the Understanding and Design of Materials. *J. Am. Chem. Soc.* **142**, 20273–20287 (2020).

33. C. Rudin, Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **1**, 206–215 (2019).

34. N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, D. Vanderbilt, Maximally localized Wannier functions: Theory and applications. *Rev. Mod. Phys.* **84**, 1419–1475 (2012).

35. G. Dresselhaus, M. S. Dresselhaus, Fourier Expansion for the Electronic Energy Bands in Silicon and Germanium. *Phys. Rev.* **160**, 649–679 (1967).

36. R. Dronskowski, P. E. Bloechl, Crystal orbital Hamilton populations (COHP): energy-resolved visualization of chemical bonding in solids based on density-functional calculations. *J. Phys. Chem.* **97**, 8617–8624 (1993).

37. V. L. Deringer, A. L. Tchougréeff, R. Dronskowski, Crystal Orbital Hamilton Population (COHP) Analysis As Projected from Plane-Wave Basis Sets. *J. Phys. Chem. A* **115**, 5461–5466 (2011).

38. S. Maintz, V. L. Deringer, A. L. Tchougréeff, R. Dronskowski, LOBSTER: A tool to extract chemical bonding from plane-wave based DFT: Tool to Extract Chemical Bonding. *J. Comput. Chem.* **37**, 1030–1035 (2016).

39. X. Qian, *et al.*, Quasiatomic orbitals for *ab initio* tight-binding analysis. *Phys. Rev. B* **78**, 245112 (2008).

40. H. Mizoguchi, H. W. Eng, P. M. Woodward, Probing the Electronic Structures of Ternary Perovskite and Pyrochlore Oxides Containing Sn4+ or Sb5+. *Inorg. Chem.* **43**, 1667–1680 (2004).

41. M. K. Brod, M. Y. Toriyama, G. J. Snyder, Orbital Chemistry That Leads to High Valley Degeneracy in PbTe. *Chem. Mater.* **32**, 9771–9779 (2020).

42. W. G. Zeier, *et al.*, Thinking Like a Chemist: Intuition in Thermoelectric Materials. *Angew. Chem. Int. Ed.* **55**, 6826–6841 (2016).

43. H. Mizoguchi, P. M. Woodward, Electronic Structure Studies of Main Group Oxides Possessing Edge-Sharing Octahedra: Implications for the Design of Transparent Conducting Oxides. *Chem. Mater.* **16**, 5233–5248 (2004).

44. J. C. Slater, G. F. Koster, Simplified LCAO Method for the Periodic Potential Problem. *Phys. Rev.* **94**, 1498–1524 (1954).

45. M. d'Avezac, J.-W. Luo, T. Chanier, A. Zunger, Genetic-Algorithm Discovery of a Direct-Gap and Optically Allowed Superstructure from Indirect-Gap Si and Ge Semiconductors. *Phys. Rev. Lett.* **108**, 027401 (2012).

46. L.-D. Yuan, H.-X. Deng, S.-S. Li, S.-H. Wei, J.-W. Luo, Unified theory of direct or indirect band-gap nature of conventional semiconductors. *Phys. Rev. B* **98**, 245203 (2018).

47. X. Jin, S. Chen, C. Lemkan, T. Li, Role of local atomic short-range order distribution in alloys: Why it matters in Si-Ge-Sn alloys. *Phys. Rev. Mater.* **7**, L111601 (2023).

48. B. Huang, *et al.*, Exceptional Optoelectronic Properties of Hydrogenated Bilayer Silicene. *Phys. Rev. X* **4**, 021029 (2014).

Why does silicon have an indirect band gap?

Emily Oliphant[1], Veda Mantena,[1] Madison Brod[2], G. Jeffrey Snyder[2], Wenhao Sun[1*]
[1] Department of Materials Science, University of Michigan, Ann Arbor, Michigan 48109, United States
[2] Northwestern University, Materials Science and Engineering Evanston IL, 60208 USA

*Corresponding Author: whsun@umich.edu

Supplementary information

SI.1: Details of 1D monatomic *s* and *p* tight binding model
SI.2: Comparing our 2NN model to Vogl's *sps** model
SI.3: Obtaining the best chemistry from MLWF
SI.4: Derivation of final Δ_1 band dispersion
SI.5: Details of computational tool
SI.6: Tutorial of computational tool
SI.7: Details of Si tight binding parameters compared to Ge

SI.1: Details of 1D monatomic *s* and *p* tight binding model

In this section we review the simple tight binding theory behind our model in Figure 2 of the main text. The tight binding wavefunction are the sum of *s* and *p* Bloch orbitals with their respective contribution $c_{n\alpha}^k$, shown in Eqn. 1, where α indicates the orbital, n the band, and k the k-point.

$$\Psi_n^k(r) = c_{ns}^k \, \Phi_s(r) + c_{np}^k \, \Phi_p(r) \tag{1.1}$$

The Hamiltonian for our system including first and second nearest neighbors is shown in Eqn. 2.

$$H = \begin{bmatrix} E_s + H_{ss}^{1NN} + H_{ss}^{2NN} & H_{sp}^{1NN} + H_{sp}^{2NN} \\ -H_{sp}^{1NN} - H_{sp}^{2NN} & E_p + H_{pp}^{1NN} + H_{pp}^{2NN} \end{bmatrix} \tag{1.2}$$

In 1D system with interatomic spacing of a, there are two 1NN at locations $+a$ and $-a$ and two 2NN at $+2a$ and $-2a$. The Brillouin zone is also limited to 1D, from Γ ($k = 0$) to X ($k = \frac{\pi}{a}$). Thus, the $\mathbf{k} \cdot \mathbf{R}$ term is $\pm ak$ for the 1NN and $\pm 2ak$ for the 2NN. When defining k_x from 0 at Γ to 1 at X the phase becomes $\pm \pi k_x$ for the 1NN and $\pm 2\pi k_x$ for the 1NN. Because each term is the sum of two opposite sign imaginary exponentials, the functional form reduces to sine for s-p bonds where the hopping parameter changes sign, and cosine for s-s and p-p bonds. The full bond runs are written below with the sign of the hopping parameter carried out to the front.

$$H_{sp}^{1NN}(k) = 2i \left| V_{sp}^{1NN} \right| \sin(\pi k_x)$$

$$H_{ss}^{1NN}(k) = -2|V_{ss}^{1NN}| \cos(\pi k_x)$$

$$H_{pp}^{1NN}(k) = 2\left| V_{pp}^{1NN} \right| \cos(\pi k_x)$$

$$H_{sp}^{2NN}(k) = 2i \left| V_{sp}^{2NN} \right| \sin(2\pi k_x)$$

$$H_{ss}^{2NN}(k) = -2|V_{ss}^{2NN}| \cos(2\pi k_x)$$

$$H_{pp}^{2NN}(k) = 2\left| V_{pp}^{2NN} \right| \cos(2\pi k_x)$$

From these, we see that H_{sp}^{1NN}, H_{ss}^{2NN}, H_{pp}^{2NN}, and H_{sp}^{2NN} bond runs all have extrema away from high symmetry points. Then the Hamiltonian is diagonalized with and without the 2NN terms. While the energy can be directly obtained in the diagonalization it can be equivalently written as the sum of the bonds as described below. The band dispersions for the 1NN and 2NN are than found as the equations below. Note that $E_n^{on+1NN}(k)$ is slightly different when the eigenvectors are calculated with the 2NN terms than without.

$$E_n^{on+1NN}(k) = c_{ns}^{k}{}^* c_{ns}^k \left(E_s + H_{ss}^{1NN}(k)\right) + c_{np}^{k}{}^* c_{np}^k \left(E_p + H_{pp}^{1NN}(k)\right) + c_{ns}^{k}{}^* c_{np}^k \left(2H_{sp}^{1NN}(k)\right)$$

$$E_n^{on+1NN+2NN}(k) = E_n^{on+1NN}(k) + c_{ns}^{k}{}^* c_{ns}^k \left(H_{ss}^{2NN}(k)\right) + c_{np}^{k}{}^* c_{np}^k \left(H_{pp}^{2NN}(k)\right) + c_{ns}^{k}{}^* c_{np}^k \left(2H_{sp}^{2NN}(k)\right)$$

SI.2: Comparing our 2NN model to Vogl's *sps** model

Here we compare our 1NN+2NN model, which relies crucially on the 2NN same *p* orbital bond (e.g. p_x-p_x), to the Vogl *sps** model, which also reproduces a minimum off-X for silicon. While the motivation for considering additional parameters from 1NNs is to capture the CBM off-X, the impact that the added parameter has on the rest of the bandstructure reveals how physical the parameter is. For physical accuracy, including an additional term into a model should qualitatively improve your fit across the entire Brillouin Zone or *k*-path. Because our model includes only 1NN and some 2NN terms from a TB interpolation of DFT bandstructure, it will have some quantitative error due to limiting the number of terms. While Vogl's model (which is fit only using 1NN *sps** terms) may have less quantitative error along the Γ-X line, it produces qualitatively incorrect band shapes along various *k*-paths.

To assess the physical accuracy of our model compared to *sps**, we look at changes in qualitative band shape on other *k*-paths. First examining the Γ-L line, a clear physical problem for the *sps** model, is the significant lowering of the Λ_2 band. While the DFT bandstructure exhibits a slight lowering close to L, the L_2 point is still significantly higher than the L_1 and L_3. But in the *sps** model, the L_2 band is lower than L_3, sitting around L_1. Because the 2NN *p-p* interaction have almost no effect on any of the bands along the Γ-L line, it can lower the X_1 point without affecting the L_2 point. Additionally, while the run of the Λ_1 band with the *sps** model is lowered near the L point, consistent with the DFT bandstructure, the band minimum is significantly off-L, which is not reproduced from DFT.

Figure SI.1: The band structures from DFT, the limited Wannier TB model, and Vogl's sps* TB model are plotted. *Bands and high symmetry point are labeled as needed. The TB models are plotted with (solid red) and without (dashed gray) the interaction important to the CBM near X to analyze the affect the additional term has on the bandstructure as a whole. Ultimately, we see that the 2NN same* p-p *bond adds changes consistent with the DFT bandstructure, while the* s*-p *interaction creates unphysical changes.*

The other band shapes that are qualitatively changed by the *s*-p* interaction are the bottom two conduction bands along the Γ-K line, Σ_3 and Σ_1. Both of these bands with the *s*-p* interaction run down towards the K point, where the 1NN interactions mostly run up towards K (with a slight turn downwards in the Σ_1 band near K). In the DFT bandstructure, both bands move up from Γ to ~0.5K (Σ_3) or ~0.8K (Σ_1) than turn down sharply towards K for the remainder of the *k*-path. This matches up poorly with the *sps** model, where they run down starting at Γ (Σ_3) or ~0.3K (Σ_1) and the K_1 point is brought much lower in

energy than in DFT. Including the 2NN same p-p interaction decreases the band energies near K while leaving the bands largely unchanged from Γ-0.5K. While the exact shape does still not match perfectly, the decrease in the band energies only near K is seen in the DFT bandstructure. Thus, our 2NN same p-p interaction is deemed a physical correction to a 1NN model, while the extent of the s^*-p appears unphysical.

Furthermore, there are several features in the valence bands and conduction bands of silicon which are not reproduced with a 1NN model. The 2NN same p-p bond address these exceptionally well, while the s^*-p interactions have no effect. Firstly, the 1NN TB bandstructure is incorrectly flat along the X-W. Previously documented by Chadi and Cohen, including the 2NN same p-p interaction creates a dispersion along X-W which aligns well with DFT valence bands Z_1 and Z_4.(13) The added s^*-p interaction cannot produce the dispersion from X-W, which becomes especially problematic for the conduction Z_1 band. Because the sps^* model goal is to lower the X point, it inevitably lowers the W by an equal amount. This is highly unphysical as in the DFT bandstructure the W_1 point is ~3 eV higher than the X_1 point. Whereas the 2NN same p-p orbital bond only slightly lowers the W_1 point when lowering the X_1 point. The 1NN model also poorly predicts the top two VBs along the Γ-K, Σ_1 and Σ_2, as being too close together and does not capture the low symmetry minimum off-K in the Σ_1 band. Including the 2NN same p-p bond fixes this, separating the K_1 and K_2 points and creating the minimum near K in the Σ_1 band, while the s^*-p interaction has no affect.

Overall, including Vogl's s^*-p interaction induces unphysical changes along every other high symmetry line other than Γ-X line. This reveals that within Vogl's sps^* model, the s^*-p interaction is merely a fitting term which solely reproduces the CBM while negatively impacting the rest of the silicon bandstructure. On the other hand, the 2NN same p-p interactions leave other high symmetry lines either unchanged or makes changes consistent with the DFT bandstructure. Thus, the inclusion of the 2NN same p-p interaction and its effect on the CBM in silicon is deemed physical.

Ultimately, we believe that while physical s^* states do impact the bandstructure, it is in a much less significant role than suggested by Vogl, and that the essential physics of the low symmetry CBM in silicon is reproduced without relying on excited states. The impact of excited states on the conduction bands is likely well estimated as the difference between the DFT bandstructure and TB models reproduced using solely a sp^3 orbital basis. This is well illustrated in Figure 1a of Ref [2] by Sanchez-Portal et. al.(29)

SI.3: Obtaining the best chemistry from Maximally Localized Wannier Functions

This section outlines the essential attributes of a "good" Wannierisation such that the resulting TB model can be used for physical intuition in the third step. For additional Wannier90 formalism and code examples the reader is deferred to the Wannier90 papers and user guide.

To achieve the correct generalizable chemical intuition, the Wannierization must be done with projections of all the outer shell atomic orbitals that are included in the pseudo-potential files. For example, Si requires s and p orbitals, while atoms like Ga or Pb require s, p, and d orbitals. In the projections block of the seedname.win file this translates to having '$l = 0$; $l = 1$' for Si and '$l = 0$; $l = 1$; $l = 2$' for Ga or Pb. Specifying the zaxis or xaxis is not necessary with atomic orbitals and setting the radial part is not necessary with projection on the pseudo-wavefunctions, which have no radial nodes for the outer shell orbitals which are not excited states. Setting zona to set orbital size may be necessary if the default projected atomic orbitals are much smaller or larger than the PAW pseudo-orbitals. This is increasingly important for compounds with atoms of significantly different sizes but not strictly necessary for silicon.

Projecting all the outer shell orbitals will include some of the conduction bands in the Wannierisation. This often requires a detangling of the desired bands. In the case of silicon, the bottom four conduction bands are nearly separated from the higher conduction bands (no overt band crossings) but of course still hybridizes with higher bands. Because the TB parameters have nontrivial variation ($\pm 20\%$) depending on the window selection for disentangling, it is preferable to do no disentanglement for a clear comparison to Germanium later. When comparing the TB interpolations of silicon with and without the detangling procedure, we find similar TB parameters and bandstructures for distances < 8 Angstroms. Only above this does the poor localization of the Wannier orbitals without disentanglement manifest with higher TB parameters and a bandstructure with incorrect high frequency wiggles. The disentangled bandstructure uses parameters of num_wann = 8, num_bands = 15, dis_win_max = 21.5, dis_froz_max = 7.6 . The paper analysis is completed using the TB model without disentanglement because it produces the same qualitative dependencies on NN as the disentangled TB model and has reasonable quantitative accuracy up to hopping parameters with a distance of < 8 Angstrom.

Finally, check the results of the Wannierisation to determine whether it can be used to gain chemical intuition. Common metrics for a 'good' Wannierisation are that the change in spread of the Wannier functions is <20% and that the tight binding parameters exponentially decay with distance. While these are good indicators, a change in spread of <20% is impossible with the initial projection of atomic orbitals for

the Wannierisation of Si (and other covalent compounds). Atomic orbital ought to be used despite this because hybridized orbitals only work well in molecular chemistry because they are the final combination of the atomic orbitals. But in crystal, the localized orbitals combine to form a nearly infinite number of crystal orbitals across momentum space (k-space) producing a variety of hybridizations which only reduces to what human intuition might expect at limited high symmetry points. Thus, atomic orbitals must be used to clearly show the natural hybridization at any band and point in k-space. Fortunately, a precise tight binding model that maintains crystal symmetry and has near zero imaginary components can still be interpolated.

The high-quality tight binding model will be obtained when two things are true throughout the Wannierisation: the centers of the Wannier orbitals are that of the host atom in all symmetrically constrained directions, and the spread of orbitals for atoms of the same type are identical. These conditions generally need to be met immediately in the initial projection or Wannierisation tends to indirectly optimize one of the constraints at the cost of the other. For the initial projection to be good enough, I have found an odd Γ-centered k-point grid of the DFT calculation to be necessary. Our code requires Wannier90 output files _hr.dat, _center.xyz, and .win to generate the interactive TB band structure using a tight binding package. Then, the corresponding eigenvectors are used to determine the orbital contributions and plot the crystal wavefunctions with the e^{ikR} term using the .xsf Wannier orbital files.

Finally, our framework requires a *real*-space tight binding description of the *reciprocal*-space band structure. In order to apply this framework generally to materials an automated tight binding interpolation of nonorthogonal atomic orbitals from DFT needs to be realized. While MLWFs often provide a good TB interpolation, the basis set of Wannier orbitals are distorted, often significantly, from the original atomic orbitals(34, 49), creating hopping parameters and coefficients which may be inconsistent with chemical intuition. On the other hand, methods that perform a TB interpolation of predefined (orthogonalized) atomic orbitals have significant errors, preventing generalization and confidence in such approaches.(50) Although the orbitals generated from MLWFs are orthogonal and vary from traditional atomic orbitals, they can still yield valuable insight when the hopping parameters produce the same sign and magnitude as from atomic orbitals.

SI.4: Derivation of final Δ_1 band dispersion

We show the full derivation and decompose of equation 6 of the main text.

The total energy of a TB wavefunction can be broken down into components from individual Bloch orbitals i and j, Eqn. 1.

$$E_n(k) = \langle \Psi_{nk}|H|\Psi_{nk}\rangle = \sum_{ij} c_{ni}^{k}{}^\dagger c_{nj}^{k} H_{ij}(k) = \sum_{ij} E_{n,ij}(k) \tag{4.1}$$

Along the Δ_1 band, there are only s and p_x orbitals for each atom, limiting i and j to four options. This expansion is shown by the Eqn. 2 where the superscript indicates the atom of the orbital and p_x is written as p for simplicity.

$$E_{\Delta_1}(k) = E_{n,s^1s^1} + E_{n,s^1p^1} + E_{n,s^1s^2} + E_{n,s^1p^2} + E_{n,p^1s^1} + E_{n,p^1p^1} + E_{n,p^1s^2} + E_{n,p^1p^2}$$

$$+E_{n,s^2s^1} + E_{n,s^2p^1} + E_{n,s^2s^2} + E_{n,s^2p^2} + E_{n,p^2s^1} + E_{n,p^2p^1} + E_{n,p^2s^2} + E_{n,p^2p^2} \tag{4.2}$$

These 16 terms are reduced considering that the coefficients for atom 1 are the same as atom 2 for the entire Δ_1 band and that atom 1 and atom 2 are the same element and thus have the same orbitals. Terms where $ji=ij$ are also grouped together.

$$E_{\Delta_1}(k) = 2E_{n,s^1s^1} + 2E_{n,p^1p^1} + 4E_{n,s^1p^1} + 2E_{n,s^1s^2} + 4E_{n,s^1p^2} + 2E_{n,p^1p^2} \tag{3.3}$$

The 1NN interactions are between atoms 1 and 2, being the last three terms of Eqn 3, while interactions between the same atom include both the onsite (orbital energy) terms and the 2NN interactions. This equation written as onsite, 1NN, and 2NN terms is Equation 6 of the main text and Eqn. 4 below.

$$E_{\Delta_1}(k) = 2E_{\Delta_1,ss}^0 + 2E_{\Delta_1,pp}^0 + 2E_{\Delta_1,ss}^{1NN} + 4E_{\Delta_1,sp}^{1NN} + 2E_{\Delta_1,pp}^{1NN} + 2E_{\Delta_1,ss}^{2NN} + 4E_{\Delta_1,sp}^{2NN} + 2E_{\Delta_1,pp}^{2NN} \tag{3.4}$$

The individual terms are expanded using the bond run, where $\langle i,j\rangle$ goes over the onsite, 1NN, or 2NN bonds and $V_{ij}(\boldsymbol{R})$ is the hopping parameter.

$$E_{n,ij}(k) = c_{ni}^{k}{}^\dagger c_{nj}^{k} H_{ij}(k) = c_{ni}^{k}{}^\dagger c_{nj}^{k} \sum_{\boldsymbol{R}} V_{ij}(\boldsymbol{R}) e^{ik\cdot(\boldsymbol{R}+r_j-r_i)} \tag{4.5}$$

The onsite terms are most obvious, where E_s^0 is the orbital energy and $c^2 = c^\dagger c$:

$$E_{\Delta_1,ss}^0 = c_{\Delta_1,s}^{k_x}{}^2 E_s^0 \tag{4.6}$$

$$E_{\Delta_1,pp}^0 = c_{\Delta_1,p}^{k_x}{}^2 E_p^0 \tag{4.7}$$

1st nearest neighbors

The 1NN terms have four bonds, with $R_j - R_i$ terms of $\left[\frac{1}{4},\frac{1}{4},\frac{1}{4}\right], \left[\frac{1}{4},-\frac{1}{4},-\frac{1}{4}\right], \left[-\frac{1}{4},\frac{1}{4},-\frac{1}{4}\right], \left[-\frac{1}{4},-\frac{1}{4},\frac{1}{4}\right]$ in units of the lattice vectors $[\overrightarrow{a_1},\overrightarrow{a_2},\overrightarrow{a_3}]$, or denoted as $\left(\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)$ in Cartesian coordinates. The Δ high symmetry line goes from $[0,0,0]$ to $[0.5,0,0.5]$ in units of the reciprocal lattice vectors $[\overrightarrow{b_1},\overrightarrow{b_2},\overrightarrow{b_3}]$ where $\overrightarrow{a_i}\cdot\overrightarrow{b_j} = 2\pi\,\delta_{ij}$. Writing $\vec{k}$ as k_x, which is 0 at Γ and 1 at X, we have $\sum_{1NN} V_{ij} e^{ik\cdot(\boldsymbol{R}+r_j-r_i)} = V_{ij}\left(\frac{1}{2},\frac{1}{2},\frac{1}{2}\right) e^{i\frac{k_x\pi}{2}} + V_{ij}\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2}\right) e^{-i\frac{k_x\pi}{2}} + V_{ij}\left(-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right) e^{i\frac{k_x\pi}{2}} + V_{ij}\left(-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right) e^{-i\frac{k_x\pi}{2}}$. Because s-s and p-p bonds have the same sign and magnitude for all V_{ij} while s-p has opposite sign but same magnitude we obtain Eqns 8-10 below.

$$E_{\Delta_1,ss}^{1NN} = 4c_{\Delta_1,s}^{k_x}{}^2 V_{ss}(\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2})\cos\left(\tfrac{k_x\pi}{2}\right) \tag{4.8}$$

$$E_{\Delta_1,pp}^{1NN} = 4c_{\Delta_1,p}^{k_x}{}^2 V_{pp}(\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2})\cos\left(\tfrac{k_x\pi}{2}\right) \tag{4.9}$$

$$E_{\Delta_1,sp}^{1NN} = 4c_{\Delta_1,s}^{k_x}{}^\dagger c_{\Delta_1,p}^{k_x} V_{sp}(\tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{2})\sin\left(\tfrac{k_x\pi}{2}\right)i \tag{4.10}$$

The $s - s$ ($p_x - p_x$) bond shape along the $\Gamma - X$ line is a negative (positive) cosine wave with only a quarter of the cosine wavelength shown, relating to bonding (antibonding) at Γ and nonbonding at X. The $s - p_x$ bond has the shape of a quarter of a positive sine wave, with nonbonding at Γ and antibonding at X. Including the TB parameters, and the phase of the coefficients (real for s-s and p-p but imaginary for s-p), we obtain all the 1NN contributions written as Eqn. 11 below.

$$E_{\Delta_1}^{1NN}(k_x) = 2E_{\Delta_1,pp}^{1NN} + 2E_{\Delta_1,ss}^{1NN} + 4E_{\Delta_1,sp}^{1NN} \tag{4.11}$$

2nd nearest neighbors

The 2NN terms have twelve bonds, with $R_j - R_i$ terms with variations of $[\pm 1,0,0]$, $[0,\pm 1,0]$, $[0,0,\pm 1]$, $[0,\pm 1,\mp 1]$, $[\pm 1,0,\mp 1]$ and $[\pm 1,\mp 1,0]$ in lattice vectors. Of the 6 groups listed 1, 3, 4, and 6 have nonzero projections along $\Gamma \to X$ line of $\pm k_x \pi$ with bond vector denoted as (110) in Cartesian coordinates, while groups 2 and 5 have zero projection with bond vector denoted as (011) in Cartesian coordinates. Similar to the 1NN, these breakdown to cosine for s-s and p-p bonds and sine for s-p as shown in Eqns 12-14, where the extra x in the subscript denotes the TB parameters between 2NNs along the x-direction.

$$E_{\Delta_1,ss}^{2NN} = c_{\Delta_1,s}^{k_x}{}^2 [8V_{ss}(110)\cos(k_x\pi) + 4V_{ss}(011)] \tag{4.12}$$

$$E_{\Delta_1,pp}^{2NN} = c_{\Delta_1,p}^{k_x}{}^2 \left[8V_{pp}(110)\cos(k_x\pi) + 4V_{pp}(011)\right] \tag{4.13}$$

$$E_{\Delta_1,sp}^{2NN} = c_{\Delta_1,s}^{k_x}{}^{\dagger} c_{\Delta_1,p}^{k_x} \left[8V_{sp}(110)\sin(k_x\pi)\,i + 4V_{sp}(011)\right] \tag{4.14}$$

Where V_{sp}^{2NN} should be zero by symmetry. The bonding for the variable eight 2NNs should have the same shapes as their 1NN counterparts but will now cover half, rather than a quarter, of the cosine or sine wave along the $\Gamma - X$ line. These are written with the TB parameters and the coefficients phase to give the contribution of 2NN, which is Eqn. 15 below. Here the $p_x - p_x$ bonds dominate with larger hopping parameters and coefficients.

$$E_{\Delta_1}^{2NN}(k_x) = 2E_{\Delta_1,pp}^{2NN} + 2E_{\Delta_1,ss}^{2NN} + 4E_{\Delta_1,sp}^{2NN} \tag{4.15}$$

SI.5: Details of computational tool

The code starts from a tight binding interpolation of DFT results, requiring basic information about the primitive cell, the number of orbitals, and the tight binding (or hopping) parameters. There will be a separate hopping parameter to represent a bond between each orbital set of i and j with all possible displaces for orbital j of $\mathbf{R} = R_1\widehat{a_1} + R_2\widehat{a_2} + R_3\widehat{a_3}$. If the interpolation is good, many of the hopping parameters will be equal or equal and opposite according to the symmetry of the orbitals and crystal. To obtain the initial band structure, the TB model is solved in the standard way using the hopping parameters to constructing k-dependent Hamiltonians which are diagonalized to obtain the energies and eigenvectors (orbital coefficients).

Then, once the user selects a point, the energy contributed by each individual bond is calculated by Eqn. 1 below, combining to Equation 1 and 3 of the main text, where the second sum is over all relevant displacements $\mathbf{R}$ between orbitals i and j, often written as $\langle i, j \rangle$.

$$E_n(k) = \sum_{ij} c_{ni}^{k}{}^{\dagger} c_{nj}^{k} \, H_{ij}(k) = \sum_{ij} c_{ni}^{k}{}^{\dagger} c_{nj}^{k} \sum_{\mathbf{R}} V_{ij}^{\mathbf{R}} \, e^{ik \cdot \mathbf{R}} \tag{5.1}$$

This creates an array with a flattened size of $i \cdot j \cdot R_1 \cdot R_2 \cdot R_3$, where each element represents a single bond in the crystal and its corresponding value is the energy that bond contributes. This energy per bond array is sorted and equal magnitude elements are grouped into bonding groups, which includes all bonds that contribute the same amount of energy to that point on the bandstructure. For bonds to be grouped, their displacement must have the same magnitude when projected on k-point, $|k \cdot \mathbf{R}|$, in addition to having the same hopping parameter. One group of bonds will contribute a characteristic real shape to the total energy dispersion. For example, Eqns. 3.12 describes the energy for all 12 of the 2NN s-s bonds, where 8 are grouped into a cosine shape and 4 are grouped into a constant.

Finally, the band-dependent bond runs and bond energies can be plotted for each bond group using Equations 4 and 5 of the main text. This is written using the hopping parameters in Eqns. 2 and 3, where $\langle i, j, \mathbf{R} \rangle_b$ includes all the hopping parameters in bond group number b. While the phase of the coefficients $\frac{c_{ni}^{k}{}^{\dagger} c_{nj}^{k}}{|c_{ni}^{k}||c_{nj}^{k}|}$ changes gradually, if at all, along high symmetry lines, degenerate bands (and band crossing) cause significant jumps. Because of this, the band-dependent bond run for degenerate bands is plotted using the coefficient phase of only the clicked point, rather than changing the phase based on k-point.

$$H_{n,b}(k) \equiv \sum_{\langle i,j,\mathbf{R} \rangle_b} \frac{c_{ni}^{k}{}^{\dagger} c_{nj}^{k}}{|c_{ni}^{k}||c_{nj}^{k}|} V_{ij}^{\mathbf{R}} \, e^{ik \cdot \mathbf{R}} \tag{5.2}$$

$$E_{n,b}(k) = |c_{ni}^{k}||c_{nj}^{k}| \, H_{n,b}(k) \tag{5.3}$$

Altogether, this process allows us to circumvent the long steps of calculating by hand the energies for a given band and high symmetry line as was done in SI3, which is even more challenging for less symmetric high symmetry paths.

SI.6: Tutorial of computational tool

General process

1. Open https://viz.whsunresearch.group/tb/, and click on point in bandstructure.
2. **Observe tables** with the orbital character and the most important bonds.
3. **Plot the bond run and energies** for the important bond groups by clicking on the row. This informs how a bond group effects the band dispersion along a given high symmetry line. The bond runs plotted will always be for the nth band based on the band index n that was selected in the first step.
4. **Edit hopping parameter** magnitudes to achieve a bandstructure with desired change.

Next, we analyze the Γ, X, and L points using our online tool. We show how to use our app to make figures similar to Figure 5(a-d) of the main text. Although, the band structures will not be exactly the same because the online app has a further NN cutoff than that plotted for our manuscript. Areas which are clicked or edited are highlighted in red, showing in the top left panel which point is clicked, in the bottom left panel which bond is selected to plot bond run and energy, and in the main panel which TB parameters are edited to generate the desired change in bandstructure.

Γ point

Observe tables: When the Γ point is clicked, the tables populate to reveal that the character is 50/50 Wannier orbital 1 and 5 which corresponds to the s orbital on atom 1 and 2. The bonds at play at this point are the 1NN, 2NN, and 3NN s-s interaction. The 1NN s-s TB parameter is –1.64 eV, in the antibonding state selected this increases the energy by 6.56 eV. The 2NN s-s TB parameter is 0.084, this is counterintuitive if the Wannier orbitals are perfect s orbitals as they should have bonding (negative) TB parameters at any distance. The positive TB parameter is reflective of the nodal behavior of Wannier orbital due to their orthogonality requirement.

Plot bond runs: In this case, the value strictly at Γ is what we are most interested in, so the bond runs play a less significant role. Particularly because the bond runs are for the highest band (rather than the lowest conduction band) along the L to Γ high symmetry line.

Edit hopping parameters: Here we want to lower the s antibonding state which was selected. We to this by reducing the TB parameter of the 1NN s-s bond by 15%. This successfully lowers s antibonding state while minimally effecting the energies of surrounding bands and the CBMs at the L and X points.

X point

Observe tables: When near the CBM near the X point is clicked, the orbital character is shown to be *s* and p_x orbitals on atom 1 and 2 equally. The p_x orbitals have ~80% total character while the *s* orbitals have the remaining ~20%. The important bonds include (1) all the 1NN s-p bonds, (2) the 2NN *p-p* bonds without a projection on x, (3) the 2NN *p-p* bonds with a projection on *x*.

Plot bond runs: The band-dependent bond runs and bond energies will be consistent with Figure 4 and equations 7 and 8 of the main paper. With the blue lines representing the band-dependent bond runs plotted in the top of Figure 4 and the pink lines representing how the bond contributes to the total energy by scaling the bond run with the orbital character as in the middle of Figure 4.

Edit hopping parameters: With the goal of selectively lowering the X point we increase 2NN *p-p* hopping parameters by 30%. This has the desired effect because the 2NN with and without an *x* projection destructively combine at Γ, but they constructively combine near X.

L point

The "Orbital character of selected point" and "Important bonds of selected point" tables shown in the figure:

Orbital character of selected point

Orbital Type	1 - Si s	2 - Si pz	3 - Si py	4 - Si px	5 - Si s
Percent Character	49.4	29.2	29.2	29.2	49.4

Important bonds of selected point

Orbitals	Total Energy	Degeneracy	TB parameters	Modified TB
15, 51	2.402	6	[-1.639]	1.639
16, 17, 18, 25, ...	1.86	36	[-1.075, 1.075]	1.075
16, 17, 18, 25, ...	1.354	12	[-1.075, 1.075]	1.075
27, 28, 36, 38,...	-1.154	36	[-1.13, 1.13]	1.469
27, 28, 36, 38,...	-0.818	12	[1.13]	1.469

Observe tables: This band along L to Γ has character from *s, p_x, p_y,* and *p_z* orbitals. Atoms 1 and 2 again contribute equally and the *p_x, p_y,* and *p_z* orbitals have equal character. Much of the band is split between ~50% character *s* orbitals and ~50% character combined *p* orbitals. The diversity of orbital character increases the number of active bonds. Along the L direction, the four 1NN split into one with a nonzero projection on k-point and three with zero projection. This is distinguished in their bond run, as the bonds with zero projection are flat. The bonds contributing most to the energy at this k-point are (1) the three *s-s* bonds with zero projection (x2 for orbital variations), (2) the three *s-p* bonds (x12 for orbital variation), (3) the one *s-p* bond with a projection (x12), (4) the three *p-p* bonds (x12), (5) the one *p-p* bond with projection (x12), (6) (shown online when by scrolling) the one s-s bond with projection (x2). The extra multiplicity for the *p-p* bonds comes for the six possible combinations of different *p_x-$p_{y/z}$* orbitals.

Plot bond runs: The bond runs break into ones with more # of params that are flat because the bonds have no projection along *k*-point (bonds 1, 2, and 4) and ones with less # of params which have a cosine shape (bonds 3, 5, and 6). While the first set have flat bond runs, they can have a shape in the bond energy if the orbital character changes significantly along the band. Because this band has minimal change in orbital character, the plotted bond energies remain mostly flat, except when very close to Γ.

Editing hopping parameters: With the goal of selectively lowering the L point, we target the 1NN different orbital p-p bonds by increasing them by 30%. This successfully lowers this point while have no impact on the CBM at Γ and X. This time it has no effect because there are no *p* orbitals in the antibonding *s* state at Γ and only one type of *p* orbital at the X point, restricting any different *p* orbital bonding.

SI.7: Details of Si tight binding parameters compared to Ge

The main chemical difference between silicon and germanium is the filled d shell in Ge. This introduces 10 electrons which do not fully screen the increased positive nuclear charge for the outer shell *s* and *p* orbitals. This especially impacts the *s* orbitals, which are smaller and lower in energy compared to the *p* orbitals than the *s* orbitals in silicon. This effect is often referred to as *d*-block or scandide contraction.

Because of this, we expect a rebalancing of the bond strengths to favor *p* orbitals. The *s-s* and possibly *s-p* bonds should get weaker while the *p-p* bonds should get stronger. Additionally, the *s* orbital energies should decrease as they are more bound. While the exact energy level cannot be compared without band alignment calculations, we can compare the energy difference between the *s* and *p* orbitals. A final note I is that the Wannier orbitals will not be exactly atomic orbitals. Because of this, the TB parameters following the developed chemical trend is obscured as the nodes near neighboring atom in the Wannier function has inconsistent effects. Still, we are able to observe the expected changes in the TB parameters. Although, this is less severe than I would anticipate for perfect atomic orbitals as the changes could be mapped onto multiple higher NN.

All of this considered, the 1NN TB parameters in eV for silicon are 1.64, 1.08, 1.13, and 0.13 for the *s-s*, *s-p*, different type *p-p*, and same type *p-p*, respectively. The *s* and *p* orbital energies are 7.12 eV apart. In germanium, the TB parameters are very similar with the largest difference occurring in the *s-s* hopping and *s* orbital energy difference from the *p* orbital. The parameters are 1.40, 1.03, 1.15, and 0.16 for the *s-s*, *s-p*, different type *p-p*, and same type *p-p*, respectively. The *s* and *p* orbital energies are 8.31 eV apart.

To test how these different parameters change the bandstructure, we can use our new computational tool. (Changing the orbital energy has to be done in code because it is excluded from the list of important bonds). Because we are more concerned with the accuracy of the CBM points, these bandstructures are recreated using TB parameters that are <8 Angstroms apart. Most of the change in bandstructure from silicon to germanium within the Wannierisation is caused by the *s* orbital energy that is 1.19 eV lower and the *s-s* bond that is 0.24 eV weaker.

Ultimately, the Γ point is exclusively lowered by changing either the s orbital energy or s-s hopping parameters but the L and X points can be lowered through a couple chemical mechanisms, including increased *p-p* hopping, decreased s-p hopping, or decreased *s-s* hopping. While the Wannier90 results point to changes in the *s* orbital being the main cause for the onsite and 1NN terms, it is not conclusive due to the non-atomic nature of the Wannier orbitals. In other words, changes in the *p* orbitals could be mapped to the *s* orbitals as the orbitals hybridize to maintain orthogonality.

Supplementary Information References

1. D. J. Chadi, M. L. Cohen, Tight-binding calculations of the valence bands of diamond and zincblende crystals. *Phys. Status Solidi B* **68**, 405–419 (1975).

2. D. Sanchez-Portal, E. Artacho, J. M. Soler, Projection of plane-wave calculations into atomic orbitals. *Solid State Commun.* **95**, 685–690 (1995).

3. N. Marzari, A. A. Mostofi, J. R. Yates, I. Souza, D. Vanderbilt, Maximally localized Wannier functions: Theory and applications. *Rev. Mod. Phys.* **84**, 1419–1475 (2012).

4. J. Qiao, G. Pizzi, N. Marzari, Projectability disentanglement for accurate and automated electronic-structure Hamiltonians. *Npj Comput. Mater.* **9**, 1–14 (2023).

5. D. Sánchez-Portal, E. Artacho, J. M. Soler, Analysis of atomic orbital basis sets from the projection of plane-wave results. *J. Phys. Condens. Matter* **8**, 3859–3880 (1996).