Topological electronic properties of silicon

A. Shtyk\textsuperscript{1} and C. Chamon\textsuperscript{2}

\textsuperscript{1}Department of Physics, Harvard University, Cambridge, MA 02138, USA
\textsuperscript{2}Department of Physics, Boston University, Boston, MA, 02215, USA

The central role that materials play in human history is exemplified by the three-age division of prehistory into the stone, bronze, and iron ages. References to our present time as the information age or silicon age epitomizes the important role that this semiconducting material came to play in the development of computers and devices that permeate our daily lives. Here we show that the electronic states in silicon have nontrivial topological structures that are captured by a network of Berry flux lines that link at points of high symmetry in the Brillouin zone. This complex network has ice-nodal points where fluxes satisfy ice rules, making silicon a “nodal-chain insulator”. Fixing the longitudinal momentum parallel to such flux lines yields a two-dimensional Dirac Hamiltonian for the transverse degrees of freedom. This complex Berry-flux network implies a topologically stable two-fold degeneracy along the X-W direction in all of silicon bands, a fact that is supported by crystal symmetry arguments as well as direct inspection of the vast literature on silicon band structures. Similarly to nodal-chain semimetals, we find drumhead-like states in the regions that are delimited by the projections of the bulk Berry flux network on the surface Brillouin zone.

I. INTRODUCTION

The experimental discovery of the integer quantum Hall effect\textsuperscript{1}, where the Hall resistance is quantized to the extraordinary precision of one part in a billion, led to the new standard of resistance for the international system of units. A degree of precision such as this has its roots in a fruitful confluence of physics and mathematics, which ties the Hall resistance to a topological quantity. In the case of the integer Hall effect, this quantity is the first Chern number associated to each filled Landau level\textsuperscript{2}. The integer Hall effect was the first example of a system with topological electronic properties; the number of systems in which topology plays a prominent role has grown explosively in the recent past, fueled by the discovery of a new class of topological band insulators occurring in semiconductors with strong spin-orbit coupling, in which gapless surface states exist\textsuperscript{3–16}. (For reviews, see Refs. 17,18.) After the discovery of topological insulators, many examples of topological semimetals were identified, such as Weyl metals\textsuperscript{19–25} and systems with Weyl nodal lines\textsuperscript{19,26}, and nodal chains\textsuperscript{27}.

Spin-orbit interactions play a prominent role in the topological insulators and Weyl systems, but spin-orbit coupling is not central to the understanding of the basic electronic properties of silicon. Indeed, in the standard classification of topological insulators, silicon is not classified\textsuperscript{28} as one with protected Dirac surface states. Therefore, silicon thus far has sat on the sidelines. Here we show that the sublattice structure of crystals such as silicon is responsible for a network of Berry flux lines in the Brillouin zone that link at points of high symmetry in momentum space. This Berry flux network is topologically stable, obeys ice rules (two in, two out) at the X points, and is responsible for topological protection of degeneracies along the X-W direction. The nontrivial topological structure of the Berry flux network in silicon shares the same physical origin as the Berry flux in graphene: the fact that there are two atoms in the unit cell gives rise to a spinor structure with associated Berry phases. The existence of the Berry flux network opens a novel “topological knob” to manipulate electrons in silicon, especially in light of progress made in the past decade from studying the effects of Berry phases in the electronic properties in graphene.

II. BERRY FLUX NETWORK

We uncover the Berry flux network in two steps. First we identify the spinor structure and the Berry fluxes within a general tight-binding approximation (valid for an arbitrary number of orbitals). Second we argue based on topological and symmetry arguments that the Berry flux network is robust and remains beyond the tight-binding description of silicon.

A tight-binding Hamiltonian for a system on a bipartite lattice comprised of sublattices A and B can be represented as

\[ H(k) = \begin{bmatrix} H_{AA}(k) & H_{AB}(k) \\ H_{BA}(k) & H_{BB}(k) \end{bmatrix}, \tag{1} \]

where the blocks \( H_{AA} \) and \( H_{BB} \) contain hoppings between sites in the same sublattice, and the blocks \( H_{AB} \) and \( H_{BA} = H_{AB}^\dagger \) contain hoppings between sites in different sublattices. The size of these blocks depend on the number of orbitals included in the tight-binding model. For example, in graphene the blocks are \( 1 \times 1 \) if one considers only the \( \pi \)-orbital, and in silicon the blocks are \( 4 \times 4 \) if one takes account of only the \( s, p_x, p_y, p_z \) orbitals (or \( 5 \times 5 \) in the \( sp^3s^* \) model\textsuperscript{29}, with the inclusion of the \( s^* \) orbital). The diagonal blocks are periodic in \( k \)-space: \( H_{AA}(k) = H_{AA}(k + G_i) \) and \( H_{BB}(k) = H_{BB}(k + G_i) \), where \( G_i \) \( (i = 1, \ldots, d) \) is a reciprocal lattice basis vector \( (d = 2 \) for graphene, \( d = 3 \) for silicon). The matrix
Consider the walk in $k$-space that visits, in order, the points $\mathbf{k}, \mathbf{k} + \mathbf{G}_1, \mathbf{k} + \mathbf{G}_1 + \mathbf{G}_2, \mathbf{k} + \mathbf{G}_1 + \mathbf{G}_2 + \mathbf{G}_3$ and back to $\mathbf{k}$. This walk passes through all four of the tetrahedron formed by the four points in $k$-space, closing a loop. We choose the initial point not to be one of high symmetry, to avoid that the edges pass through band crossings, thus avoiding degeneracies along the walk. For example, one may choose to start close to but not at the $\Gamma$ point, say at $\mathbf{k} = (\delta_x, \delta_y, \delta_z)$, with infinitesimal $\delta_x, \delta_y, \delta_z$. At the end of the walk, the Hamiltonian returns to $H(\mathbf{k})$, but the eigenvector is rotated by the sequence of unitary operations

$$U(–\mathbf{G}_1 – \mathbf{G}_2 – \mathbf{G}_3) U(\mathbf{G}_3) U(\mathbf{G}_2) U(\mathbf{G}_1) = e^{i \pi \sigma_z} = -1,$$

which amounts to a rotation by $\pi$. This geometric phase implies the existence of $\pi$-flux lines in $k$-space, which pierce the loop we described above. An example of a $\pi$-flux network that threads the 4-edged loop in $k$-space is shown in Fig. 1, which, as we show below, corresponds to the cases of the lowest conduction and valence bands in silicon.

These singular flux lines are stable, and cannot be removed by small deformations. Silicon is inversion symmetric, which implies that the Berry curvature $\mathcal{B}(\mathbf{k}) = \mathcal{B}(–\mathbf{k})$; in the absence of time-reversal breaking perturbations, $\mathcal{B}(\mathbf{k}) = – \mathcal{B}(–\mathbf{k})$. These two symmetries, together, imply that the Berry curvature vanishes everywhere with the possible exception of singular lines carrying flux multiple of $\pi$, like those we identified above. These two symmetries thus ensure that the $\pi$-flux cannot spread over a finite region, and thus remains singular and contained within a network of flux lines circulating around the Brillouin zone. The number of orbits in the description of the system does not alter our conclusions based on the topological constraints imposed by Eq. (5).

To visualize the network of fluxes in silicon, we consider explicitly the 5-orbital $sp^3d^5$ nearest-neighbor tight-binding model[29]. This model captures essential features of silicon’s band structure; in particular, it reproduces the conduction band minimum along the $\Delta$ line connect-

**FIG. 1:** A loop in $k$-space encircling a single $\pi$-flux line. A Berry flux network in $k$-space is schematically shown in red color with the grey surface guiding the visualization of the flux flow. Blue color represents a $0 - \mathbf{G}_1 - (\mathbf{G}_1 + \mathbf{G}_2)^- \mathbf{(G}_1 + \mathbf{G}_2 + \mathbf{G}_3)^- 0$ walk that we use in the main text to argue the existence of the Berry $\pi$-flux line pincer this loop.

The elements of the off-diagonal blocks are written as

$$[H_{AB}(\mathbf{k})]_{\alpha\beta} = -\sum_{\mu} t_{\alpha\beta}(\mathbf{d}_\mu) e^{i\mathbf{k} \cdot \mathbf{d}_\mu},$$

where the vectors $\mathbf{d}_\mu$ connect the atoms from $A$ to $B$, and the $t_{\alpha\beta}(\mathbf{d}_\mu)$ contain the overlap of the orbitals $\alpha$ and $\beta$ separated by $\mathbf{d}_\mu$. Because the vectors $\mathbf{d}_\mu$ are not Bravais lattice vectors, $\mathbf{d}_\mu \cdot \mathbf{G}_i$ is generically not a multiple of $2\pi$, and therefore the off-diagonal blocks $H_{AB}$ and $H_{BA}$ are not periodic in $k$-space. An attempt to gauge out these non-periodicities comes with the price of adding singularities (branch cuts) to the phase of the wave function in momentum space.

In graphene the vectors $\mathbf{d}_\mu$, $\mu = 0, 1, 2$, point to the vertices of a triangle, while in silicon the $\mathbf{d}_\mu$, $\mu = 0, 1, 2, 3$, point to the vertices of a tetrahedron. In these lattices (see Supplementary Online Information), it follows that $H_{AB}(\mathbf{k} + \mathbf{G}_i) = e^{i \Phi(\mathbf{G}_i)} H_{AB}(\mathbf{k})$, where $\Phi(\mathbf{G}_i) = 2\pi/N$, with $N = 3$ and $N = 4$ for the graphene and silicon lattices, respectively. The Hamiltonian $H(\mathbf{k})$ is not periodic in $k$-space; however, it is periodic up to a unitary transformation that rotates the amplitudes on the two sublattices by opposite phases:

$$H(\mathbf{k} + \mathbf{G}_i) = U(\mathbf{G}_i) H(\mathbf{k}) U^\dagger(\mathbf{G}_i),$$

with

$$U(\mathbf{G}_i) = e^{i \frac{1}{2} \Phi(\mathbf{G}_i) \sigma_z},$$

and $\sigma_z$ a Pauli matrix acting on the $A/B$ sublattice grading. It follows from Eq. (3) that the eigenenergies $\epsilon(\mathbf{k}) = \epsilon(\mathbf{k} + \mathbf{G}_i)$, as expected. However, the spinor structure and the lack of periodicity of the matrix Hamiltonian $H(\mathbf{k})$ (not just its eigenvalues) is what leads to the Berry $\pi$-vortices at the $K$ points in graphene, and the Berry $\pi$-flux network that we uncover in this work. (In the Supplementary Online Information we illustrate how the generic framework above recovers the familiar results in graphene as a warm up for the calculations in silicon.)

**FIG. 2:** Berry curvature field. Left: a single first Brillouin zone. Right: a doubled Brillouin zone. The colored spheres represent three inequivalent X points. The grey surface guides the visualization of the flux flow and the ice-rule links at the X points.
ing the Γ and X points in the Brillouin zone. In this approximation, sites within the same sublattice are not connected, so that diagonal blocks simply contain the on-site potential energies \( H_{AA/BB} = \text{diag}(E_s, E_p, E_p, E_{s'}) \), while the inter-sublattice hoppings contain the non-trivial momentum dependence:

\[
H_{AB}(k) = \begin{pmatrix}
V_{ss} g_0(k) & V_{sp} g_1(k) & V_{sp} g_2(k) & V_{sp} g_3(k) \\
-V_{sp} g_1(k) & V_{xx} g_0(k) & V_{xy} g_3(k) & V_{xy} g_4(k) \\
-V_{sp} g_2(k) & V_{xy} g_3(k) & V_{xx} g_0(k) & V_{xy} g_2(k) \\
0 & V_{sp} g_3(k) & V_{xy} g_4(k) & V_{sp} g_3(k)
\end{pmatrix}
\]

where momentum functions

\[
\begin{align*}
g_0(k) &= \frac{1}{4} (e^{i d_0 k} + e^{i d_1 k} + e^{i d_2 k} + e^{i d_3 k}), \\
g_1(k) &= \frac{1}{4} (e^{i d_0 k} + e^{i d_1 k} - e^{i d_2 k} - e^{i d_3 k}), \\
g_2(k) &= \frac{1}{4} (e^{i d_0 k} - e^{i d_1 k} + e^{i d_2 k} - e^{i d_3 k}), \\
g_3(k) &= \frac{1}{4} (e^{i d_0 k} - e^{i d_1 k} - e^{i d_2 k} + e^{i d_3 k}),
\end{align*}
\]

and \( d_0 = \frac{a}{4} (1 1 1), \quad d_1 = \frac{a}{4} (1 \bar{1} \bar{1}), \quad d_2 = \frac{a}{4} (\bar{1} 1 \bar{1}), \quad d_3 = \frac{a}{4} (\bar{1} 1 1) \), with \( a = 5.4310 \text{Å} \).

This spectrum has an intricate set of lines along which the conduction band can be studied within a simple toy model. Left: all bands exhibit a two-fold degeneracy at the ice-nodal point of the lowest conduction band (see Supplementary Online Information for details)

\[
H_X = \varepsilon_0(p) \hat{s}_0 + v_x p_x \hat{s}_1 + \kappa_{yz} p_y p_z \hat{s}_2,
\]

where \( v_x = 0.51, \kappa_{yz} = 0.18, \varepsilon_0(\bar{p}) = 1.63 + 0.11 p_y^2 \), and \( \hat{s}_i \) are Pauli matrices. (Energies are measured in eV and momentum in units of \( 1/a \).) This expression explicitly shows the ice-nodal nature of the X point. For example, in the vicinity of the points \( 2\pi (1,0,\pm 0.1) \) the Hamiltonian takes the form

\[
H_{X \pm \delta} = \tilde{v}_x p_x \hat{s}_1 \pm \tilde{v}_y p_y \hat{s}_2,
\]

where \( \tilde{v}_x = 0.51 \) and \( \tilde{v}_y = 0.02 \) are electron velocities at the \( 2\pi (1,0,\pm 0.1) \) points. This Hamiltonian structure indicates two Dirac lines with opposite chiralities approaching the X point along the \( z \)-axis from opposite directions. Similarly, there are two more such lines along the \( y \)-axis (see Fig. 4).

It turns out that all crucial topological properties of the conduction band can be studied within a simple toy

**III. EFFECTIVE HAMILTONIAN NEAR THE X POINT OF THE LOWEST CONDUCTION BAND**

The effective two-band Hamiltonian in the vicinity of the ice-nodal point of the first conduction band is obtained by expanding \( k = (2\pi,0,0) + p \), yielding (see Supplementary Online Information for details)

\[
H_X = \varepsilon_0(p) \hat{s}_0 + v_x p_x \hat{s}_1 + \kappa_{yz} p_y p_z \hat{s}_2,
\]

where \( v_x = 0.51, \kappa_{yz} = 0.18, \varepsilon_0(\bar{p}) = 1.63 + 0.11 p_y^2 \), and \( \hat{s}_i \) are Pauli matrices. (Energies are measured in eV and momentum in units of \( 1/a \)). This expression explicitly shows the ice-nodal nature of the X point. For example, in the vicinity of the points \( 2\pi (1,0,\pm 0.1) \) the Hamiltonian takes the form

\[
H_{X \pm \delta} = \tilde{v}_x p_x \hat{s}_1 \pm \tilde{v}_y p_y \hat{s}_2,
\]

where \( \tilde{v}_x = 0.51 \) and \( \tilde{v}_y = 0.02 \) are electron velocities at the \( 2\pi (1,0,\pm 0.1) \) points. This Hamiltonian structure indicates two Dirac lines with opposite chiralities approaching the X point along the \( z \)-axis from opposite directions. Similarly, there are two more such lines along the \( y \)-axis (see Fig. 4).

It turns out that all crucial topological properties of the conduction band can be studied within a simple toy model. Left: all bands exhibit a two-fold degeneracy at the ice-nodal point of the lowest conduction band. Right: first Brillouin zone with points of high symmetry.

**FIG. 3: ELECTRONIC BAND STRUCTURE OF SILICON WITHIN THE SP^3S' MODEL.** Left: all bands exhibit a two-fold degeneracy along the X-W line (highlighted in red). This degeneracy may be used as evidence for the existence of the Dirac lines along the X-W direction. Right: first Brillouin zone with points of high symmetry.
model, with a single orbital per site and nearest-neighbor tight-binding Hamiltonian

\[ H_{\text{toy}}(k) = t_{nn} \begin{bmatrix} 0 & g_0(k) \\ g_0^*(k) & 0 \end{bmatrix} \]

\[ = t_{nn} \left( \hat{\sigma}_1 \text{Re} g_0(k) + \hat{\sigma}_2 \text{Im} g_0(k) \right). \tag{10} \]

Hopping matrix elements to farther neighbors can be included, especially between the sites of the same sublattice, to reproduce features of silicon's band structure, such as a conduction band minimum along the \( \Delta \) line. Yet, the topological features are captured by the off-diagonal terms alone: the Hamiltonian is degenerate along the \( g_0(k) = 0 \) manifold, yielding essentially the same cage-like net of Dirac lines connecting at X points, just as in the full \( sp^3s^* \) model for silicon.

The real and imaginary parts of \( g_0(k) \) are

\[
\text{Re} \ g_0(k) = \cos \frac{k_x}{4} \cos \frac{k_y}{4} \cos \frac{k_z}{4}, \\
\text{Im} \ g_0(k) = -\sin \frac{k_x}{4} \sin \frac{k_y}{4} \sin \frac{k_z}{4}. \tag{11} 
\]

from which we identify

\[
(2\pi, 0, s), \ (2\pi, s, 0), \ (0, 2\pi, s), \ (s, 2\pi, 0), \\
(0, s, 2\pi), \ \text{and} \ (s, 0, 2\pi), \ \text{for} \ s \in [0, 2\pi), \tag{12} 
\]

as the set of nodal lines, crossing at the three ice-nodal X points,

\[
(2\pi, 0, 0), \ (0, 2\pi, 0), \ (0, 0, 2\pi). \tag{13} 
\]

Expanding around one of the ice-nodal points, \( k = (2\pi, 0, 0) + p \), we obtain

\[
\text{Re} \ g_0(k) = -\sin \frac{p_x}{4} \cos \frac{p_y}{4} \cos \frac{p_z}{4} \simeq -\frac{1}{4} p_x, \\
\text{Im} \ g_0(k) = -\cos \frac{p_x}{4} \sin \frac{p_y}{4} \sin \frac{p_z}{4} \simeq -\frac{p_y p_z}{16}, \tag{14} 
\]

reproducing the same structure of the X-point effective Hamiltonian derived using the \( sp^3s^* \) model:

\[
H_{\text{toy}, X} = -V_{AB} \left( \frac{p_x}{4} \hat{\sigma}_1 + \frac{p_y p_z}{16} \hat{\sigma}_2 \right). \tag{15} 
\]

We note that perturbing this toy model Hamiltonian with a \( \hat{\sigma}_3 \) term gaps out the Dirac lines (it breaks sublattice symmetry), while a \( \hat{\sigma}_2 \) perturbation separates the Dirac lines in different ways depending on the sign of the \( \hat{\sigma}_2 \) term, indicating the critical character of the ice-nodal point.

Finally, we gather all results above on the ice-nodal points and combine them with other information inferred from results on the band structure of silicon that is obtained from methods other than tight-binding. We condense this combined information into an effective Hamiltonian near the \( (2\pi, 0, 0) \) nodal point:

\[
H_{\text{eff}, X} = \frac{p_x^2}{2m_\ell} + \frac{p_y^2}{2m_\ell} + \frac{p_z^2}{2m_\ell} + v_x p_z \hat{\sigma}_1 + \kappa p_y p_z \hat{\sigma}_2, \tag{16} 
\]

where \( m_\ell = 0.98 m_e \) and \( m_1 = 0.19 m_e \) coincide with the masses from the standard low-energy description of the conduction band (with \( m_e \) the electron mass); the velocity \( v_x = 0.15 \, (2\pi/a)^2 (h/m_e) \) is estimated from the location of the conduction band minimum; and \( \kappa \) follows from the dispersion along the X-U direction. This effective Hamiltonian, we claim, captures the topological properties and hence gives a more accurate description of the low energy states in silicon’s conduction band than simply expanding the dispersion to quadratic order near the band minimum. That the minimum occurs close to the X point is captured by the interplay between the Dirac dispersion along the \( x \)-direction and the parabolic term, which places the minimum close to the X point (notice that the energy difference between the minimum and the X point is only 0.08eV).

## IV. Surface States and Bulk-Boundary Correspondence

Probably the most important theoretical insight in the field of topological insulators is the fact that non-trivial bulk topology leads to the existence of robust surface states. Recently this idea was extended to topological semimetals, where the Fermi surface consists of a nodal loop. In such materials surface states exist only in parts of the surface Brillouin zone that are determined by projecting the nodal loop from the bulk Brillouin zone onto the surface Brillouin zone. Here we argue that the same bulk-boundary correspondence applies to silicon and the existence of the Berry flux “wire frame” leads to the existence of non-trivial topological drumhead surface states.

The two-band toy model Hamiltonian (10) is of an archetypal semimetallic and the manifold \( g_0(k) = 0 \) defines a nodal loop. Moreover, this Hamiltonian describes a nodal chain semimetal, a topologically critical system with crossing nodal lines yielding ice-nodal X
points\textsuperscript{27,33}. According to the intuition from semimetallic systems\textsuperscript{34–37}, the projection of the bulk wire frame onto the surface Brillouin zone breaks it into segments. These segments can be colored with two colors, for example red and blue, in such way that adjacent parts are always of the opposite color. Surface drumhead states then exist in all segments colored either in blue or in red, depending whether the crystal surface is terminated at A or B sublattice. In the case of silicon that we consider, the projection of the wire frame breaks down surface Brillouin zone into quadrants. Surface states exist either in first and third or second and fourth quadrants, see Fig. 5.

In the nearest neighbor toy model, chiral symmetry protects the dispersionless nature of the wire frame and the Fermi Surface at this energy coincides with the wire frame. The same happens at the touching of the two lowest valence bands in the $sp^3s^*$ model, see Fig. 3. In real silicon, however, $X - W$ lines possess nontrivial dispersion. Hopping matrix elements to farther neighbors can be included in order to break sublattice symmetry and reproduce this feature:

$$H_{\text{toy-2}}(\mathbf{k}) = \tilde{\sigma}_0 \left( t_{n^3} v_{n^3}(\mathbf{k}) + t_{n^3} v_{n^3}(\mathbf{k}) \right) + t_{nn} \left( \tilde{\sigma}_1 \Re g_0(\mathbf{k}) + \tilde{\sigma}_2 \Im g_0(\mathbf{k}) \right),$$  

where

$$v_{n^3}(\mathbf{k}) = \frac{1}{2} \sum_{i \neq j} \cos \frac{k_i a}{2} \cos \frac{k_j a}{2},$$  

$$v_{n^3}(\mathbf{k}) = \sum_{i} \cos (k_i a).$$

We allow for both next-nearest-neighbor ($n^3$) and next-next-nearest-neighbor ($n^4$) hopping. It turns out that $n^3$ hopping alone is not sufficient to destroy the flatness of dispersion along X-W lines. We fit parameters of the toy model to reproduce $\Gamma, X, W$ point energies in the lowest valence band of real silicon. The resulting band structure and Fermi surface in the vicinity of the X point are shown in Fig. 6.

V. SILICON QUANTUM WELL

The clean and straightforward connection between bulk topology and surface states relies on the sublattice symmetry, which, unfortunately, is destroyed even within the $sp^3s^*$ model. While $sp^3s^*$ model has matrix elements only between A and B sublattices, only two of the ten states are involved in the formation of each wire frame. When the full Hamiltonian is projected on the relevant two dimensional subspace for each wire frame, longer range hopping elements are induced via transitions to the states that were projected out. For example, if we focus on the lowest energy wire frame that is formed by the first two bands, next-nearest neighbor element between an A site at the $i$-th and $(i+1)$-th cells is formed as

$$- \sum_{\alpha=2\ldots10} \frac{\langle A_i, 1 | H_{\text{eff}} | B_i, \alpha \rangle \langle B_i, \alpha | H_{\text{eff}} | A_{i+1}, 1 \rangle}{(E_{\alpha} - E_1)^2}. \quad (20)$$

Despite this unfortunate fact, the bulk-boundary correspondence and surface states in the toy model are of topological nature. Hence, we expect that breaking respective symmetries retains qualitatively bulk-boundary correspondence and surface states for reasonably large perturbations of the toy model.

To explore drumhead states within the $sp^3s^*$ model, we perform a numerical calculation of the electron dispersion in a silicon slab. We consider a slab of the material with [100] surface and the thickness approximately

![FIG. 5: Bulk-boundary correspondence within the toy model.](image)

![FIG. 6: Electronic band structure and Fermi surface of the toy model.](image)
20 nm (40 unit cells). The sample is infinite in the \( y, z \) directions parallel to the surface, for which we perform the Fourier transformation, retaining only the spatial dimension \( x \) perpendicular to the surface, so that the electron wavefunction can be factorized as

\[
\Psi^{\alpha \mu}(x_i, k_y, k_z) = e^{i(k_y y + k_z z)} \psi^{\alpha \mu}_{k_y, k_z}(x_i),
\]

where \( x_i \) is the position of the \( i \)-th unit cell \((i = 1 \ldots 40)\), \( \alpha = 1, 2 \) labels \( A/B \) sublattices and \( \mu = 1 \ldots 5 \) enumerates each of the five orbitals. We find all eigenvalues of the Hamiltonian \( E_i(k_{||}) \) for each value of the momentum \( k_{||} = (k_y, k_z) \) parallel to the surface and repeat this procedure along the path \( X_s - X_s1 - X_s2 \) within the surface Brillouin zone. The result is shown on the Fig. 7. While the exact nature of the bulk-boundary correspondence does not survive in the five band model, the qualitative connection remains clear.

As we have argued in the previous section, the energy dispersion of silicon in the vicinity of the touching of the lowest valence bands is fairly close to that of the toy model. Largely, this is the case due to other bands being well separated in energy from the first two valence bands. To further clarify the connection between the surface states and the bulk topology, we consider the following modification of the \( sp^3s^* \) model. We slowly tune interband hopping parameters from 0% to 100%,

\[
(V'_{sp}(\alpha), V'_{s*p}(\alpha), V'_{xy}(\alpha)) = \alpha \cdot (V_{sp}, V_{s*p}, V_{xy}),
\]

keeping other parameters intact. At \( \alpha = 0 \) we have five exactly solvable copies of the toy model and for each of them we know that the drumhead states exist and they are of topological nature. This way, by slowly tuning \( \alpha \) from 0 to 1, we can track the evolution of the surface state arising from the lowest energy wire frame.

In Fig. 8 we zoom into the region between points \( X_{s2} \) and \( X_{s1} \), to show a striking robustness of the drumhead state in the full \( sp^3s^* \) model. On the top panels, we show the evolution of the spectrum for \( \alpha = 0, 0.5, \) and 1.0, and on the bottom panels we show the spectrum resulting from further addition of a term that breaks inversion symmetry and opens a gap. Notice that the drumhead states acquire a dispersion as \( \alpha \) is turned on, and that the corresponding bandwidth is large. Thus, if one opens a gap that is smaller than that bandwidth (as in the bottom panels), a stable drumhead state remains at the surface, crossing the bulk gap. We remark that the scale of the bandwidth is rather large in the \( sp^3s^* \) model, of the order of several \( eV \). In other words, breaking of the chiral symmetry protects the drumhead state against other perturbations that weakly break any other symmetry.

In addition to tracking the \( \alpha \)-dependence of the energy spectrum, for each value of \( \alpha \) we also calculate the inverse participation ratio defined as

\[
\text{ipr}(k_y, k_z) = \sum_i \left| \psi_{k_y, k_z}(x_i) \right|^4.
\]

For localized states the ipr is inversely proportional to the localization length \( \propto 1/L \), while for bulk states it vanishes as \( \propto 1/L \), where \( L \) is the size of the system. The resulting

---

**FIG. 7:** Energy dispersion \( E(k_y, k_z) \) within a silicon slab (valence bands on the left and conduction bands on the right). Figure in the center shows surface Brillouin zone with shaded areas (2nd and 4th quadrants) revealing region where a surface state exists within a toy model.

**FIG. 8:** Drumhead state dispersion for different interband interaction strengths \( \alpha = 0.0, 0.5, 1.0 \) (columns). Bottom row shows dispersion with an additional \( A/B \) staggered chemical potential corresponding to opening of 1 \( eV \) bulk gap between the first and second valence bands.
Inverse participation ratios are shown in the Fig. 9. Aside from the emergence of new surface states as we tune $\alpha$, the topological nature this surface state is unambiguous.

VI. CONCLUSIONS AND OUTLOOK

In this work we identified a network of $\pi$ Berry flux lines in reciprocal space, for silicon. We first constructed a rather general argument, based solely on the existence of two sublattices, to argue that there must be $\pi$ flux lines independently of how many bands there are. The $\pi$ flux lines are tied to the spinor structure due to the two sublattices, and the flux cannot spread out because of time-reversal and inversion symmetry. The situation is analogous to what happens in graphene, where the Dirac points carry $\pi$ flux and cannot be removed perturbatively.

We then identified these lines in a tight-binding model containing 5-orbitals per sublattice (a $10 \times 10$ matrix Hamiltonian). We showed explicitly that the $\pi$ flux lines appear, and identified the $X$ point as a location where flux lines meet. The electronic dispersion near the $X$ point can be described in terms of the Dirac lines analyzed in this paper.

Flux lines inside the bulk Brillouin zone imply the existence of drumhead surface states, which are confined within the projection of the flux lines onto the surface Brillouin zone. We discussed how the breaking of sublattice symmetry makes the identification of the drumhead states less obvious, but that nevertheless the qualitative connection between bulk-boundary remains.

There are two sets of questions that our work suggests:

- Can one observe all or some of these features experimentally?
- Can the knowledge that these topological features exist in silicon lead to novel electronics in this "old" material?

Regarding the first question, one of the possible tools to probe these flux lines and the dispersion of wire frames is angle resolved photoemission spectroscopy (ARPES). Since the features we identified are present in all bands of silicon, one can study them in the valence bands, which are accessible in ARPES. If high energy photons are used, not only can one probe electronic states deep inside the valence bands but also use the high penetration depth (here the $x$-direction) of those photons to map constant energy surfaces as function of $k_y$ and $k_z$ for different $k_x$. Such techniques are used in Ref. 38, for example. These types of scans would be able to identify the wire frames, providing evidence for the lines of nodes where the flux runs through.

As for addressing the second question, one must find ways to pull the physics of these nodal lines or surface states to the Fermi level. While the Fermi level lies in the gap for intrinsic silicon, one can reach the regime where the Fermi level crosses the conduction band in inversion layers in doped silicon field effect transistors (FETs).

Another possibility is to use undoped silicon, and pull down the surface states by electric fields. States that are already localized at the boundary are more sensitive to potentials caused by an electric field (a linear potential has its largest and smallest values at the boundary). Intrinsic silicon cannot screen the electric field. Once the surface states are pulled to the Fermi level, there should be metallic boundary states. The absence of disorder should lead to high mobilities at these surfaces. While the mechanism described above does not require that the surface states be drumhead ones, we already observe from Fig. 7 that the lowest surface band at positive energies is of the drumhead type.

The findings presented in this work reveal novel topological electronic features in the band structure of silicon, one of the best known and most studied materials. That these features had been missed does not signal an accident, but rather suggests that there are a number of topological properties occurring in many, if not most, other materials. The topological features of silicon that we expose provide new impetus to revisit the physics of bulk silicon and two-dimensional electron gases in silicon FETs, particularly in light of what is now known from recent studies of both graphene and topological insulators.
Acknowledgments

We thank Michael El-Batanouny, Shyam Erramilli, Thomas Iadecola, Christopher Mudry, Andrei Rucken-stein, and Alex Sushkov for insightful discussions. We acknowledge financial support from the U.S. Dept. of Energy Grant DE-FG02-06ER46316 (C.C.).
Let us first apply the formalism presented in the main text to derive the familiar result that in graphene there is a \(\pi\)-vortex in the Brillouin zone at the K point. The carbon atoms form a honeycomb lattice with interpenetrating triangular sublattices \(A\) and \(B\). The \(A\) sites sit on lattice sites spanned by the basis vectors \(\mathbf{R}_1 = \sqrt{3}/2 \ a (1/2, \sqrt{3}/2)\) and \(\mathbf{R}_2 = \sqrt{3}/2 \ a (-1/2, \sqrt{3}/2)\), where \(a = 1.42\,\text{Å}\). The reciprocal lattice vectors are \(\mathbf{G}_1 = \frac{4\pi}{3a} (\sqrt{3}/2, 1/2)\) and \(\mathbf{G}_2 = \frac{4\pi}{3a} (-\sqrt{3}/2, 1/2)\). The three vectors connecting sublattice \(A\) to \(B\) are \(\mathbf{d}_0 = a (0,1)\), \(\mathbf{d}_1 = a (-\sqrt{3}/2, -1/2)\), and \(\mathbf{d}_2 = a (\sqrt{3}/2, -1/2)\).

The off-diagonal hopping matrix element for the \(\pi\)-orbital is

\[
H_{AB}(\mathbf{k}) = -\sum_{\mu=0}^{2} t(\mathbf{d}_\mu) e^{i\mathbf{k} \cdot \mathbf{d}_\mu},
\]

where we allowed generically for unequal values for the hoppings to the three nearest neighbors. The vectors \(\mathbf{G}_{1,2}\) and \(\mathbf{d}_{0,1,2}\) satisfy

\[
\mathbf{G}_i \cdot \mathbf{d}_\mu = \frac{2\pi}{3} \mod 2\pi \quad \text{for} \quad i = 1, 2 \quad \text{and} \quad \mu = 0, 1, 2.
\]

It then follows that the \(2 \times 2\) Hamiltonian for graphene is not periodic in \(\mathbf{k}\), but instead is periodic up to the unitary transformation in Eq. (3) of the main text (repeated here for convenience):

\[
H(\mathbf{k} + \mathbf{G}_i) = U(\mathbf{G}_i) \ H(\mathbf{k}) \ U^\dagger(\mathbf{G}_i), \quad \text{with} \quad U(\mathbf{G}_i) = e^{i\frac{\phi(\mathbf{G}_i)}{\mathbf{G}_z}} \sigma_z,
\]
This rotation changes the sign of the wavefunction, which is equivalent to an accumulation of a Berry phase of $\pi$.

We then consider the walk in $k$-space that visits, in order, the points $\mathbf{k}, \mathbf{k} + \mathbf{G}_1, \mathbf{k} + \mathbf{G}_1 + \mathbf{G}_2$ and back to $\mathbf{k}$. This walk visits three neighboring Brillouin zones, returning to the original point in $k$-space, closing a loop, shown in Fig. 10. We know that, in graphene, this loop will not visit a degeneracy point if we choose $\mathbf{k} = 0$ (the $\Gamma$ point). (In silicon, we shift the point $\mathbf{k}$ from the origin to avoid going through degeneracies.) At the end of the walk, the Hamiltonian returns to $H(\mathbf{k})$, but the eigenvector is rotated by the sequence of unitary operations

$$
U(-\mathbf{G}_1 - \mathbf{G}_2) U(\mathbf{G}_2) U(\mathbf{G}_1) = e^{i \frac{1}{2} [\Phi(-\mathbf{G}_1 - \mathbf{G}_2) + \Phi(\mathbf{G}_2) + \Phi(\mathbf{G}_1)]} \sigma_z = e^{i \pi} \sigma_z = -\mathbb{1}.
$$

This rotation changes the sign of the wavefunction, which is equivalent to an accumulation of a Berry phase of $\pi$.

Indeed, the Dirac node at the $K$ point contains the $\pi$-vortex when all the three hoppings to the nearest neighbors are equal. If these three hoppings are not equal, the vortex moves location in $k$-space, but cannot disappear; it must be contained within the triangle. To disappear, the vortex must come to the boundaries of the triangle to meet an anti-vortex, but when that happens our assumption that the path does not include a degenerate point no longer applies.

Having completed the warm up exercise of recovering known results for graphene in two spatial dimensions, we apply the same steps for silicon in three dimensions. In silicon the $\mathbf{A}$ applies.

In silicon there are four possible axial phases and hence four colors for the reciprocal lattice sites, indicating the quadrupling of the Brillouin zone.

FIG. 10: **Reciprocal lattice.** Left: reciprocal lattice of graphene. There are three possible axial phases at the center of the Brillouin zone, $0, 2\pi/3, 4\pi/3$, which are represented by three colors. This leads to tripling of the Brillouin zone, as shown by the dashed line enclosing three hexagons of different colors. The figure also shows a loop constructed out of $\mathbf{G}$ vectors with an overall $\pi$ Berry phase that establishes existence of the Dirac points at the $K/K'$ points.

Right: reciprocal lattice of silicon. In the case of silicon there are four possible axial phases and hence four colors for the reciprocal lattice sites, indicating the quadrupling of the Brillouin zone.

with $\Phi(\mathbf{G}_i) = 2\pi/3$. One can also check that $-(\mathbf{G}_1 + \mathbf{G}_2) \cdot \mathbf{d}_\mu = \frac{2\pi}{\pi} (\text{mod } 2\pi)$, for $\mu = 0, 1, 2$, or equivalently, $\Phi(-\mathbf{G}_1 - \mathbf{G}_2) = 2\pi/3$.

The four vectors connecting sublattice $A$ to $B$ are

$$
\mathbf{d}_A = \frac{a}{4} (1, 1, 1), \quad \mathbf{d}_1 = \frac{a}{4} (1, -1, -1), \quad \mathbf{d}_2 = \frac{a}{4} (-1, -1, 1), \quad \text{and } \mathbf{d}_3 = \frac{a}{4} (-1, 1, 1).
$$

The off-diagonal block matrix, whose dimension depends on how many orbitals we consider, is given by Eq. (2) of the main text with the index $\mu = 0, 1, 2, 4$:

$$
[H_{AB}(\mathbf{k})]_{\alpha\beta} = -\sum_{\mu=0}^{4} t_{\alpha\beta}(\mathbf{d}_\mu) e^{i \mathbf{k} \cdot \mathbf{d}_\mu}, \tag{S5}
$$

where again we allowed generically for unequal values for the hoppings to the three nearest neighbors. The vectors $\mathbf{G}_{1,2,3}$ and $\mathbf{d}_{0,1,2,3}$ satisfy

$$
\mathbf{G}_i \cdot \mathbf{d}_\mu = \frac{2\pi}{4} (\text{mod } 2\pi) \quad \text{for } i = 1, 2, 3 \text{ and } \mu = 0, 1, 2, 4. \tag{S6}
$$
The Hamiltonian for silicon is therefore not periodic in \( k \), but instead is periodic up to the unitary transformation in Eq. (3) of the main text [repeated in Eq. (S3) for convenience], with \( \Phi(G_1) = 2\pi/4 \). One can also check that 

\[-(G_1 + G_2 + G_3) \cdot d_\mu = \frac{2\pi}{\pi} \] (mod 2\( \pi \)), for \( \mu = 0, 1, 2, 3 \), or equivalently, \( \Phi(-G_1 - G_2 - G_3) = 2\pi/4 \).

We then consider the walk in \( k \)-space that was described in the main text, one that visits, in order, the points \( k, k + G_1, k + G_1 + G_2, k + G_1 + G_2 + G_3 \) and back to \( k \). As explained in the main text, we start at a \( k \) near but not at the \( \Gamma \) point, to avoid passing through lines of degeneracy. At the end of the walk, the Hamiltonian returns to \( H(k) \), but the eigenvector is rotated by the sequence of unitary operations

\[
U(-G_1 - G_2 - G_3) U(G_3) U(G_2) U(G_1) = \\
e^{i\frac{\pi}{2}[\Phi(-G_1-G_2-G_3)+\Phi(G_3)+\Phi(G_2)+\Phi(G_1)]} \sigma_z \\
e^{i\pi \sigma_z} = -\mathbb{I}.
\]

The wavefunction changes sign upon returning from the walk, which requires that a \( \pi \)-flux lines pierce the region enclosed by the walk. The flux lines required by the argument above are precisely those described in the main text, obtained via direct calculation of the Berry curvature in the \( sp^3s^* \) model.

**Appendix B: Derivation of the effective Hamiltonian at the X points for the conduction band of silicon**

Since electron energies at the X point come in degenerate pairs, the relevant physics of the two lowest conduction bands (that together are degenerate along the X-W direction) is to be described by an effective \( 2 \times 2 \) Hamiltonian. This can be done within the framework of degenerate perturbation theory\(^{32}\).

In order to derive such an effective Hamiltonian we first diagonalize the full \( 10 \times 10 \) Hamiltonian of the \( sp^3s^* \) model at the X point to identify pairs of eigenstates corresponding to five degenerate eigenvalues. Then we expand the Hamiltonian in the newfound basis to the second order in momentum (around the X point) to get a matrix of the type

\[
H = \begin{pmatrix}
\hat{H}_{I_1} & \hat{H}_{I_1I_2} & \cdots & \hat{H}_{I_1I_5} \\
\hat{H}_{I_2I_1} & \hat{H}_{I_2} & \cdots & \hat{H}_{I_2I_5} \\
\vdots & \vdots & \ddots & \vdots \\
\hat{H}_{I_5I_1} & \hat{H}_{I_5I_2} & \cdots & \hat{H}_{I_5}
\end{pmatrix},
\]

where \( I_\alpha \) label groups of degenerate sublevels and \( \hat{H}_{I_\alpha} \) are \( \mathcal{D}_\alpha \) by \( \mathcal{D}_\alpha \) matrices with \( \mathcal{D}_\alpha \) being the degeneracy of a given level group. \( \hat{H}_{I_\alpha I_\beta} \) represent matrix elements between two such groups of levels. In the case of the X point in silicon all degeneracies are twofold and \( \{I_1, I_2, I_3, I_4, I_5\} = \{(1, 2), (3, 4), (5, 6), (7, 8), (9, 10)\} \), the enumeration starting from the lowest valence band. Then within second-order perturbation theory the effective Hamiltonian describing level group \( I_\alpha \) is given by

\[
\hat{H}^\text{eff}_{I_\alpha} = \hat{H}_{I_\alpha} - \sum_{\beta \neq \alpha} \frac{\hat{H}_{I_\alpha I_\beta} \hat{H}_{I_\beta}}{E_{I_\alpha} - E_{I_\beta}}.
\]

Since \( I_\alpha \) and \( I_\beta \) are groups of indices, we can also clarify the equation above and expand in a more explicit form using ordinary level indices \( m, n, l \) and the full Hamiltonian \( H \) as

\[
\left(\hat{H}^\text{eff}_{I_\alpha}\right)_{m,n} = H_{m,n} - \sum_{\beta \neq \alpha} \frac{1}{E_{I_\alpha} - E_{I_\beta}} \sum_{l \in I_\beta} H_{m,l} H_{l,n}, \quad m, n \in I_\alpha.
\]

Performing this procedure for silicon numerically, we obtain the effective Hamiltonian for bands 5 and 6 that was presented in the main text,

\[
H^\text{eff}_{(5,6)} = \begin{bmatrix}
1.63 + 0.02p_x^2 & 0.29pp_z + 0.02ip_x \\
0.29pp_z - 0.02ip_x & 1.63 + 0.02p_x^2
\end{bmatrix}.
\]