

ARTICLE OPEN

Observation of a topological insulator Dirac cone reshaped by non-magnetic impurity resonance

Lin Miao^{1,2}, Yishuai Xu¹, Wenhan Zhang³, Daniel Older¹, S. Alexander Breitweiser¹, Erica Kotta¹, Haowei He¹, Takehito Suzuki⁴, Jonathan D. Denlinger², Rudro R. Biswas⁵, Joseph G. Checkelsky⁴, Weida Wu³ and L. Andrew Wray^{1,6}

The massless Dirac electrons found at topological insulator surfaces are thought to be influenced very little by weak, non-magnetic disorder. However, a resonance effect of strongly perturbing non-magnetic impurities has been theoretically predicted to change the dispersion and physical nature of low-energy quasiparticles, resulting in unique particle-like states that lack microscopic translational symmetry. Here we report the direct observation of impurities reshaping the surface Dirac cone of the model three-dimensional topological insulator bismuth selenide. A pronounced kink-like dispersion feature is observed in disorder-enriched samples, and found to be closely associated with the anomaly caused by impurity resonance in the surface state density of states, as observed by dichroic angle-resolved photoemission spectroscopy. The experimental observation of these features, which closely resemble theoretical predictions, has significant implications for the properties of topological Dirac cones in applied scenarios that commonly feature point-defect disorder at surfaces or interfaces.

npj Quantum Materials (2018)3:29; doi:10.1038/s41535-018-0101-8

INTRODUCTION

Three-dimensional topological insulators are bulk semiconductors with spin-helical Dirac cone surface states that span the bulk band gap.^{1,2} Since their discovery around 2007,^{3–6} the topological Dirac cone has appeared at the heart of a wide range of proposals for novel emergent quasiparticles and next-generation electronics, such as for the realization of exotic dyon-based, axion-based, or Majorana fermion-based physics.^{1,2,7–10} Moreover, the spin-helical Dirac surface states are very robust against dilute non-magnetic impurities due to intrinsic immunities to backscattering and Anderson localization,^{11–16} and numerous studies have shown the surface Dirac cone remains qualitatively intact in the presence of weak non-magnetic disorder.^{17–19} However, local probe studies have found that the effect of disorder on the *real space* electronic structure can be remarkably strong. Scanning tunneling microscopy (STM) experiments and complementary theoretical works have identified that even the simplest form of the non-magnetic impurity, a crystallographic point defect, will tend to give rise to resonance states very close to the Dirac point in two-dimensional (2D) Dirac fermion systems such as topological insulators and graphene.^{20–27} In systems like bismuth selenide (Bi₂Se₃), these surface defects are associated with distinctive patterns in STM topography images, and can redistribute local density of states (DOS) by building up a new peak tens of millielectronvolts above the surface Dirac point.

The relative significance of impurity resonances for real space electronic structure as opposed to momentum space is easy to understand if the impurities are low in density, because the associated resonance states will adhere locally to the sparse

defects. However, if there is a sufficiently high density of non-magnetic impurities, the dispersion of the topological surface Dirac cone is expected to change significantly. Recent models have shown that when the density of impurities reaches an experimentally achievable high level, the impurity resonance states will behave collectively much like a flat band that hybridizes coherently with the upper Dirac cone, breaking the Dirac band into upper and lower branches.^{28,29}

Measurements and theory suggest that electrons near the resonance adhere spatially to the disordered impurity lattice, which lacks translational symmetry. This scenario is at odds with the standard definition of a quasiparticle as a near-eigenstate of momentum. The momentum operator (K) is the generator of translations, and the spatial translation operator is defined as $T(x) = \exp(-ix \cdot K)$, implying that (near-) eigenstates of one operator should be (near-) eigenstates of the other. When an electron system is so disordered that translational symmetry in single-particle wavefunctions is absent on the length scale of one de Broglie wavelength, the result is termed a “bad metal,” and it is assumed that the quasiparticle picture no longer applies. However, the resistance of topological surface electrons to Anderson localization and backscattering is thought to result in a unique quasiparticle-like character for electrons at the impurity resonance. Though these electrons profoundly lack translational symmetry, simulations suggest that their width cross-section in momentum space is narrower than one inverse wavelength,^{28,29} and meets the nominal definition for a “good” quasiparticle. This is quite surprising, and means that inducing disorder at a topological insulator surface may enable the first experimental realization of

¹Department of Physics, New York University, New York, NY 10003, USA; ²Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; ³Rutgers Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854, USA; ⁴Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; ⁵Department of Physics and Astronomy, Purdue University, West Lafayette, IN 47907, USA and ⁶NYU-ECNU Institute of Physics at NYU Shanghai, 3663 Zhongshan Road North, Shanghai 200062, China

Correspondence: L. Andrew Wray (lawray@nyu.edu)

These authors contributed equally: Lin Miao, Yishuai Xu, Wenhan Zhang.

Received: 22 November 2017 Revised: 1 May 2018 Accepted: 4 May 2018

Published online: 04 June 2018

an itinerant quasiparticle that propagates with a well-defined momentum, but occupies a spatial basis that lacks even near-neighbor translational symmetry. A quantitative correspondence with STM measurements of the spatially resolved DOS distribution has been cited as experimental evidence for this novel quasiparticle character.²⁸ However, the larger scale morphology of such highly disordered samples has been problematic for high-resolution angle-resolved photoemission (ARPES), and a topological Dirac cone reshaped by non-magnetic disorder has not been previously reported.

In this study, we present a high-resolution linear-dichroic ARPES (LD-ARPES) investigation of defect-enriched Bi_2Se_3 , and report the experimental observation of a topological surface Dirac cone reshaped by non-magnetic impurity resonance. The LD-ARPES spectra reveal an anomalous kink-like feature in the Dirac cone dispersion 40 meV above the Dirac point, which matches the predicted signature of coherent hybridization with an impurity resonance. Integrating over momentum shows that the dispersion anomaly is associated with a DOS peak, and that both features are progressively attenuated when successive intervals of low temperature (LT) annealing are applied to reduce disorder. All of these experimental results are highly consistent with theoretical predictions and with previous STM findings.

Modeling impurity resonances in a topological surface Dirac cone Bi_2Se_3 is widely seen as a model topological insulator system, with a single Dirac cone surface state connecting across a bulk band gap of roughly 0.3 eV. An idealized model of the ARPES spectral function of the linearly dispersive Bi_2Se_3 surface state is shown in Fig. 1a. When a surface is simulated with randomly distributed non-magnetic impurities represented by scalar delta-function potentials, the Dirac bands are broadened and their dispersion is changed (Fig. 1b). The bands develop an apparent kink at $E \sim 40$ meV above the Dirac point, corresponding with a blurred region in the simulated spectrum (dashed lines in Fig. 1b). Weighting the spectra with the participation ratio (PR) of each eigenstate, a method to reveal spatially inhomogeneous states,³⁰ highlights the impurity resonant states as they are more concentrated around the impurity sites (Fig. 1c). In the weighted spectrum, the blurred feature then can be distinguished as being composed of two bands (white dashed lines) that are broken by the impurity resonance. The two disconnected dispersions are easier to distinguish at higher defect densities,²⁸ and the corresponding features broaden as they approach the resonance energy of $E \sim 40$ meV. Integrating the simulation over the 2D momentum space extracts the DOS, which shows a hump at the

resonance energy (Fig. 1d) where a local DOS peak has been observed for point defects by STM.^{20–22}

RESULTS

Crystallographic point defects occur in typical bulk-grown Bi_2Se_3 samples with relatively low densities of $\sim 0.05\%$ per 5-atom formula unit. In this study, defect-enriched Bi_2Se_3 bulk samples are synthesized by abbreviating the final annealing stage of sample growth (see Methods), following the procedure in ref.³¹. X-ray diffraction (XRD) data (Fig. 2a) and STM topography (Fig. 2b) indicate that the resulting Bi_2Se_3 samples realize a large density of point defects while maintaining good single-phase crystallinity. The resulting defect species are labeled on the STM topography map, based on previously identified correspondences.³² The primary type of impurity is interstitial Se atoms residing on the outer surface or between the first and second Bi_2Se_3 quintuple layers, with a combined density of $\rho \sim 0.08\%$ per surface unit cell (i.e., per 0.15 nm^2 surface area). A small portion of anti-site Bi_{Se} and Se vacancy defects are also observed with densities of $\rho \sim 0.03\%$ and $\rho \sim 0.01\%$, respectively. This defect density has a good correspondence with the bulk Hall carrier density, but the precise numbers seen by STM fluctuate considerably from region to region (see Methods). The theoretical precondition for impurities significantly reconstructing the electronic structure is that there must be a defect population with local resonances at approximately the same energy E_R relative to the Dirac point, and a density that exceeds a cutoff proportional to $(E_R)^2$.²⁸ In the sparse-defect limit, E_R is proportional to the negative inverse of the effective interaction strength between defects and the surface state (U_{eff} , defined in Methods),²⁷ suggesting that defects can be neglected if the ratio $\rho/(U_{\text{eff}})^2$ is significantly smaller than a critical threshold. Moreover, the surface state skin depth in Bi_2Se_3 and ARPES measurement depth are both limited to $< \sim 1 \text{ nm}$,³³ meaning that the DOS contribution from deeper-lying resonances will not be strongly observed. Based on these considerations, the 2D density adopted for simulations is $\rho = 0.06\%$, representing the assumption that roughly half of the impurities observed in the top nanometer of the crystal will effectively share a common resonance energy, and collectively exceed the critical threshold.

High-resolution LD-ARPES was used to perform a targeted study of defect-derived changes in the surface electronic structure. Measurements were performed at the Advanced Light Source MERLIN beamline (BL4.0.3), which also provides a small $50 \mu\text{m}$ beam profile to minimize feature broadening from macroscopic inhomogeneity. Two linearly polarized photoemission experimental geometries were selected as shown in the Fig. 2c, d. The σ -polarization condition places the electric field parallel to the

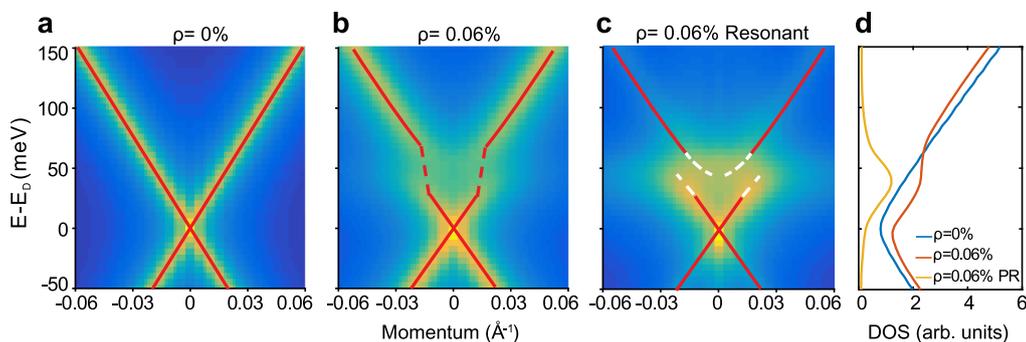


Fig. 1 Impurity resonance in momentum space. **a** A simulated Bi_2Se_3 surface Dirac cone without impurities. Red lines trace the linear band dispersion. **b** A simulated Dirac cone reshaped by scalar impurities ($U_{\text{eff}} = 1 \text{ eV}$, density $\rho = 0.06\%$ per 2D surface unit cell). A kink-like dispersive feature occurs at the impurity resonance energy, and is traced with dashed lines. **c** The simulated Dirac cone is weighted by participation ratio to highlight impurity-resonant states. The kink-like feature is resolved more clearly as containing the split dispersions predicted in ref.²⁸ (white dashed lines). **d** The corresponding simulated DOS curves of (blue) a pristine Dirac cone, (red) a Dirac cone with impurities, and (yellow) a participation ratio (PR) weighted simulation with impurities

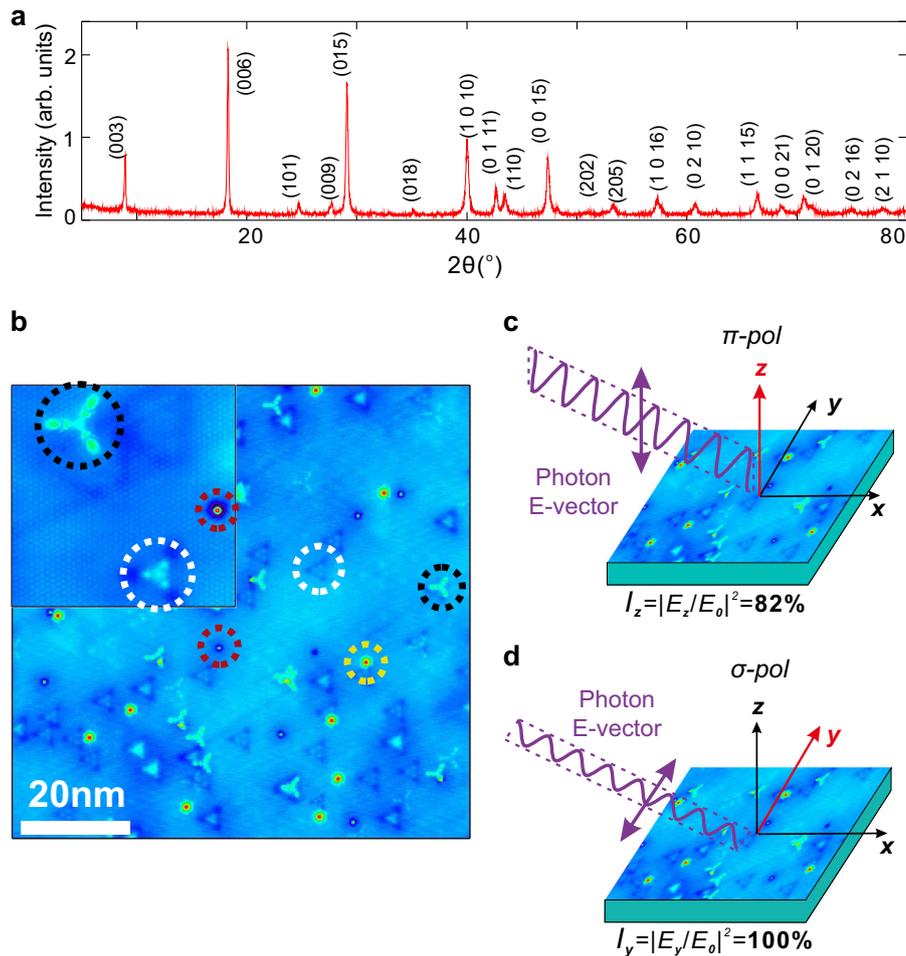


Fig. 2 Defect-enriched Bi_2Se_3 . **a** XRD from the defect-enriched Bi_2Se_3 sample, showing no impurity phase features. **b** STM topography of an $80 \times 80 \text{ nm}^2$ cleaved Bi_2Se_3 surface with lattice defects/impurities. The defects are predominantly excess Se (red circles) on the top surface and (white circles) between first and second quintuple layers. (Yellow and black circles) Anti-site Bi_{Se} defects are also common. An expanded inset shows distinctive defect profiles used for characterization. **c** The π -polarization ARPES measurement geometry, with the electric field of incident photons mostly normal to the sample surface (projecting 82% onto the z -axis). **d** The σ -polarization ARPES measurement geometry gives an electric field that projects 100% onto the in-plane y -axis (the ARPES analyzer slit axis)

sample surface, and the π measurement condition polarization is inside the scattering plane, with a primarily out-of-plane (crystal-line c -axis) orientation. Incident photon polarization was switched between π and σ geometries via control of the synchrotron beamline elliptically polarized undulator, to keep exactly the same beam spot on the sample. Previous ARPES measurements on the topological Dirac cone mainly used the π -pol experimental geometry, which gives strong emission from the primary P_z orbital component of the Dirac cone. The σ -pol condition is rarely used due to its low efficiency. However, the weakness of emission from the electronic structure of a *pristine Dirac cone* is also a positive feature of this polarization condition, as the derivative term in the electron-photon interaction couples in-plane polarization directly to nanoscale in-plane inhomogeneity, which is greatest within defect resonance states and defect-rich regions of the sample surface. The asymmetric properties of these polarization matrix elements are reviewed in the Online Supplementary Material, and qualitatively associate the π and σ geometries with selective sensitivity to defect-poor and defect-rich surface regions, respectively.

High-resolution ARPES measurements of the Dirac cone at a freshly cleaved sample surface are shown for the π -pol and σ -pol geometries in Fig. 3a. The surface Dirac bands cross at the Dirac point and the ARPES matrix element effect gives the bands

different intensity profiles under different polarizations. The lower Dirac cone is more difficult to trace due to the close proximity of the bulk valence band, and both the lower Dirac cone and the valence band show much higher intensity in σ -pol measurements. Momentum distribution curves (MDCs) of the Dirac bands are shown in Fig. 3e, and show a large discrepancy between σ and π polarizations. At energies greater than $E \sim > 30 \text{ meV}$ relative to the Dirac point, the σ -pol ARPES-imaged Dirac bands always appear to be centered at smaller momenta, indicating a dispersion different from the states seen under π -pol. The MDCs of the simulated impurity-rich surface show a spectrum very similar to the Dirac cone imaged under σ -pol. In the simulation (Fig. 3d), the Dirac bands have smaller momenta above the resonance energy ($E \sim > 30 \text{ meV}$) within the kinked Dirac cone. This polarization dependence between the dispersion of ARPES-imaged Dirac bands is not possible for a perfectly homogenous sample, but the σ -pol dispersion closely resembles the emergent “kink-like” feature associated with higher impurity densities. To reveal the quasi-particle dispersions more clearly, the MDCs of simulated and dichroic ARPES-imaged Dirac cones (Fig. 3d, e) are fitted with Voigt functions to track their dispersions (Fig. 3h, i). For simplicity, the “kink-like” feature in the Dirac cone with resonance states was also treated as being composed of just two peaks for the purposes of the fitting procedure (see Supplementary Fig. 1 for the curve-

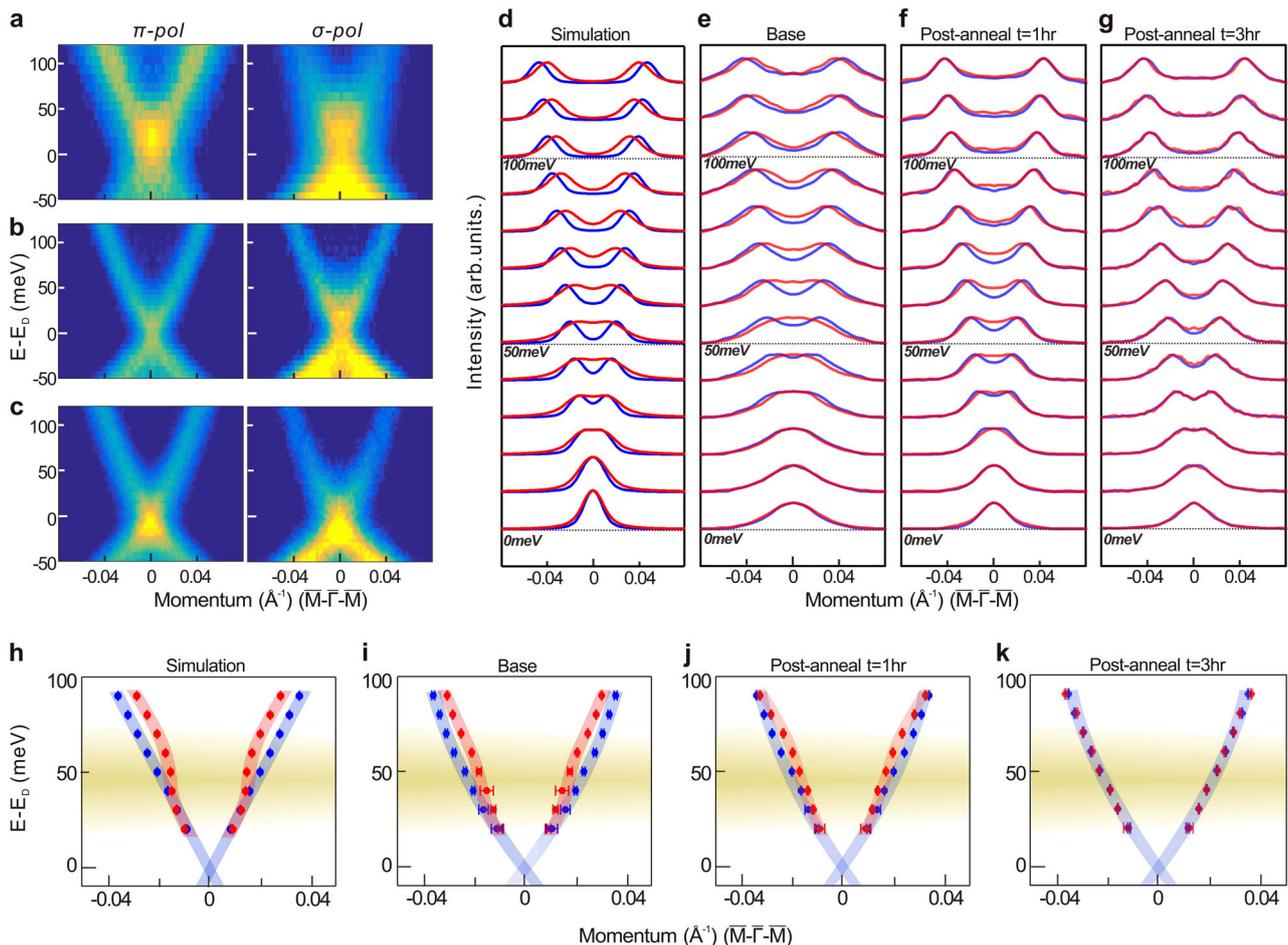


Fig. 3 Measuring the defect-derived dispersion anomaly. Symmetrized π -pol and σ -pol ARPES images of Dirac cone band structure are shown for the same defect-enriched sample (a) before LT annealing and after (b) 1 h and (c) 3 h of LT annealing. c The MDCs of a simulated Dirac cone (red) with and (blue) without impurities are extracted from Fig. 1a, b, and shown with a 10 meV step, starting from the Dirac point (set to 0 energy). To facilitate comparison, all MDCs are normalized to the same amplitude. e, f The corresponding MDCs of panels (a–c) are shown with (red) σ -pol and (blue) π -pol. h–k Dispersions are obtained from fitting the red-coded and blue-coded MDCs in d–g with two Voigt functions. A horizontal shaded region indicates the expected defect resonance energy.^{20,28}

by-curve fitting of MDCs). The σ -pol and π -pol dispersions have a relatively constant momentum offset at energies high above the Dirac point. At lower energies approaching the Dirac point, the σ -pol group velocity appears to become very large, as is common on the low-energy side of dispersion kinks, and the two dispersions merge rapidly beneath $E \sim 40$ –50 meV. These anomalous features of the σ -pol dispersion can be closely reproduced by the simulation of a Dirac cone with dense impurities (Fig. 3h).

LT annealing provides a relatively safe way to mobilize the interstitial Se atoms in the van de Waals layer without creating new impurities of Se vacancy or Bi_{Se} anti-sites. The temperature $T \sim 120$ °C is chosen to remain well below the thermal activation energy for creating new point defects, and beneath the $T > 150$ °C low-energy cutoff for eliminating larger morphological defects.^{34,35} To reduce disorder, the same sample was treated with intervals of LT annealing between synchrotron beamtimes, and re-cleaved for each new ARPES measurement. In a second experiment following 1 h of LT annealing, ARPES data (Fig. 3b) show significantly sharper bands, which is consistent with a reduced defect density. From the dichroic MDC comparison (Fig. 3f), the discrepancy of dispersion between the π -pol and σ -pol imaged Dirac cone still exists, but is visibly smaller. The corresponding traced dispersion (Fig. 3j) shows the same kink-like feature around $E \sim 40$ meV in σ -pol imaged band structure, but with a smaller amplitude and an onset slightly closer to the Dirac

point. The momentum offset between σ -pol and π -pol dispersions is reduced to roughly 60% of the value before annealing, and is less prominent outside of a narrow energy window from $E \sim 50$ –70 meV above the Dirac point. A final experiment followed 2 more hours of LT annealing (3 h in total), and found almost no difference between the π -pol and σ -pol ARPES spectra (Fig. 3g). Unlike the first 1 h LT anneal, very little change is noted in the sharpness of the bands following this final 2 h LT anneal. The traced bands (Fig. 3h) show the same dispersion under both polarizations, and there is no such anomalous kink-like feature.

Though changes in band dispersion are of particular interest for physics and applications, a more basic property of impurity resonance known from STM and theory is the build-up of a DOS peak at the resonance energy.^{20–29} This DOS feature cannot be identified from the ARPES band dispersions, as Luttinger’s theorem does not apply to disordered systems.^{36,37} To evaluate DOS, we instead sum the ARPES spectral function over the 2D momentum space, making use of the continuous rotational symmetry near the Dirac point to define $\text{DOS}(E) = \sum_k 2\pi |k| * I(E, k)$. This symmetrization procedure is applied separately to the σ -pol and π -pol dichroic ARPES data to obtain differently weighted approximations for the surface state DOS (Fig. 4a). In all cases, the π -pol imaged DOS curve has a linear trend in the upper Dirac cone, matching expectations for a massless 2D

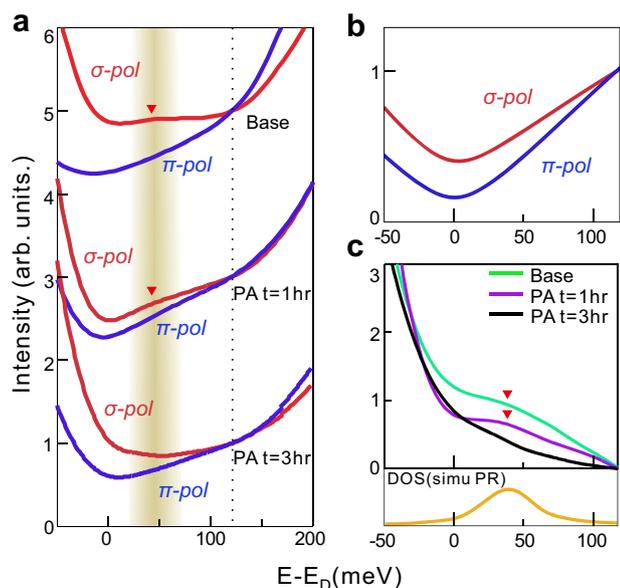


Fig. 4 The impurity resonance DOS peak. **a** ARPES DOS curves estimated from (red) σ -pol and (blue) π -pol are shown for different levels of LT post-annealing (PA). The curves are offset in increments of 2, and normalized at $E = 120$ meV above the Dirac point (dashed line), an energy that is higher than (shaded region) the expected impurity resonance and low enough to avoid the bulk conduction band. **b** Simulated ARPES DOS curves for a pristine Dirac cone, based on empirical photoemission matrix elements from ref. ³⁸. **c** Dichroic DOS curves the subtracting the π -pol ARPES DOS from σ -pol ARPES DOS. The dichroic DOS curves are compared with (bottom) the PR-weighted impurity resonance simulation from Fig. 1d. The anomalous peak-like feature that vanishes with annealing is indicated with filled triangles

Dirac fermion. However, the σ -pol data reveal an extra hump around $E \sim 40$ meV in the non-annealed base sample and after 1 h of annealing. To exclude the possibility that this feature is from an extraneous matrix element effect, a corresponding simulation for a pristine Dirac cone is shown in Fig. 4b based on the energy-resolved ARPES matrix elements identified in ref. ³⁸. The symmetrized DOS curves for both measurement conditions show a good linear character, and neither has an anomalous peak feature, suggesting that the extra hump in the symmetrized DOS curves represents a real anomaly within the Dirac surface state. A cleaner view of the dichroic feature is obtained by subtracting the π -pol curve from the σ -pol DOS (Fig. 4c), which reveals the anomalous hump as a peak at $E \sim 40$ meV in the dichroic DOS curve of the non-annealed sample. This peak continuously loses intensity with LT annealing, and is no longer identifiable after 3 h of annealing. The PR-weighted impurity resonance peak simulated in Fig. 1d is reproduced at the bottom of Fig. 4c, and is energetically consistent with the dichroic DOS feature. This correspondence can only be qualitatively meaningful as it is filtered through ARPES matrix elements; however, it indicates that the kinked band dispersion observed in σ -pol ARPES spectra is appropriately aligned with the impurity resonance in exactly the way that was theoretically predicted.

DISCUSSION

These results show that Bi_2Se_3 samples with a high point defect density can exhibit significant changes in surface state dispersion and DOS, both of which are reversible via an aging process that includes LT annealing. Numerical modeling closely reproduces the experimental features, and suggests that the dispersion kink and higher energy momentum shift in σ -pol ARPES band structure

originate from a small discontinuity in the Dirac bands due to hybridization with impurity resonance states. While this may seem like a disruption of the topological band connectivity, it does not involve time-reversal symmetry breaking, and should not necessarily be viewed as a band gap. Band topology constraints are absent in the case of strong disorder, and the DOS distribution is peaked, rather than gapped, at the resonance energy.

With respect to earlier studies, it is noteworthy that the π -pol data *do not* show a similar kink-like dispersion feature, suggesting that the measurement conditions used for most ARPES experiments will give little weight to impurity-rich surface regions (see Supplementary Note 1). Moreover, the in-plane σ -pol orientation chosen for this study was selected to achieve particularly favorable matrix elements for resolving impurity physics (see Supplementary Note 2).

Unlike the conventional manifestation of a non-dispersive impurity band, the impact of disorder is seen to play out over a wide energy scale of several hundred millielectronvolts. Together with the theoretical protection against Anderson localization,²⁸ this suggests that the Dirac cone electrons have a unique coherent relationship with the disordered impurity lattice, which preserves reasonably sharp quasiparticle-like character in spite of allowing large changes to the overall electronic structure. The DOS peak found at the dispersion kink is consistent with real space STM investigations of impurity resonance, and has been predicted to greatly enhance the susceptibility to magnetic order in mean-field modeling for an ordered impurity lattice.²⁷ Further theoretical work will be needed to more fully understand the many-body susceptibilities of a dispersive quasiparticle-like mode that is profoundly inhomogeneous on a <100 nm scale, and for which the number of states in a “band” cannot be treated as constant throughout momentum space (i.e., for which Luttinger’s theorem does not hold). These results and predictions suggest that the non-magnetic impurities found ubiquitously at the surfaces and buried interfaces of real-world samples and devices may provide a far-reaching mechanism for shaping the physical properties of a TI surface.

METHODS

High-resolution linear dichroic ARPES measurements

All ARPES measurements were performed at the BL4.0.3 MERLIN ARPES endstation at the Advanced Light Source, with a Scienta R8000 analyzer and base pressure better than 5×10^{-11} Torr. The sample was maintained at $T \sim 20$ K, the time between s-polarization and p-polarization measurements is roughly 4 h, and surface band structure features were observed to be stable within ~ 10 meV on the time scale of the each LD-ARPES experiment ($\sim <24$ h). The chemical potential for all samples was $E - E_D \sim 300$ meV above the Dirac point, which matches the expectation of $E - E_D = 325$ meV above the Dirac point for the Hall effect carrier density of $2.6 \times 10^{19} \text{ cm}^{-3}$. This surface potential is calculated using the band bending model in ref. ³⁹, which treats the surface state and bulk charge carriers on equal footing in the modified Thomas-Fermi approximation.

Surface doping of Bi_2Se_3 by adatoms and photon exposure is a natural concern in quantitative ARPES experiments, and multiple doping mechanisms can come into play.^{19,33,39–47} However, the observed stability of the measured surface in this particular case is consistent with strong bulk screening associated with the high bulk carrier density, and with the observation that aging effects from residual gas and photon exposure tend to saturate at lower surface potentials of $E - E_D \sim <290$ meV.⁴⁰ The overall energy resolution was between 15 and 20 meV, and photon flux was well below the regime on which photo-gating effects have been observed.⁴¹ Measurements on the base sample were performed at $h\nu = 30$ eV, and later post-annealing measurements were performed at $h\nu = 34$ eV to obtain a higher photon flux, due to the rapid loss of photon throughput on the high-energy grating beneath 50 eV.

Growth and STM characterization of defect-enriched Bi_2Se_3 samples

High-quality defect-enriched samples were grown following the procedure described in ref. ³¹. The resulting crystals show no sign of impurity phases.

The Hall effect carrier density was $2.6 \times 10^{19} \text{ cm}^{-3}$, corresponding to an expected defect density of $\sim 0.19\%$, which corresponds reasonably well with the roughly $\sim 0.12\%$ defect count in the top nanometer of the crystal seen by STM topography maps described in the main text. ARPES and STM measurements were performed close to the sample center, where the defect density is expected to be lower than on the periphery. The Hall carrier density and surface chemical potential are not greatly changed following the aging/annealing process. A slight reduction in the surface chemical potential suggests that the bulk carrier density is reduced by 10–20% in later measurements, relative to density beneath the beam spot in the initial measurement. Additionally, it is expected that the post-annealed defect distribution may be more homogeneous across micro-crystalline domains.

STM data were obtained at LT ($T < 50 \text{ K}$) with tips calibrated on an Au (111) surface. The topographic map in Fig. 2b was measured using a bias of -0.7 V and a tunneling current of 100 pA . Rapid topographical maps of multiple $\sim 100 \text{ nm}$ regions were sampled, and were consistent with previous analyses that suggest a stochastically random placement of impurities within local regions (see the Supplemental material in ref. 28). However, the standard deviation in defect density for different surface regions separated by several microns was much higher than the variation expected from the Poisson distribution. This variation reveals large fractional differences (up to a factor of ~ 2) in the local impurity densities that are averaged over in the ARPES spectral function.

Modeling the TI surface

The surface state is modeled as a spin-helical 2D Dirac cone on a hexagonal lattice, perturbed by scalar delta-function-like impurities, as described in ref. 28. The impurity potential is effectively reduced by the fact that the delta-function-like potential is not well resolved in the model, which imposes a high-energy cutoff on the kinetic basis for state diagonalization. An effective value for the potential is calculated as the trace of the potential Hamiltonian for a single impurity ($U_{\text{eff}} = \text{Tr}(H_{ij}) = -1 \text{ eV}$), whereas the uncorrected potential would have a value of $U = -35 \text{ eV}$, as defined in ref. 28. The impurities are randomly distributed with a density of $\rho = 0.06\%$. Spectral features were convoluted by a Lorentzian function with 30 meV peak width at half-maximum, except where otherwise noted.

Intensities in Fig. 1c are weighted by the PR P of each single-particle eigenstate. The PR gives higher intensity for states that are less evenly distributed throughout space, and is used to highlight emission from defect resonance states. It is defined as:

$$P_{\alpha} = \sum_i n_{\alpha,i}^2, \quad (1)$$

where the sum is overall sites in the system, and $n_{\alpha,i}$ is the local DOS on site i of $|\alpha\rangle$, which is an eigenstate of the full Hamiltonian.

Data availability

The data and source code that support the findings of this study are available from the corresponding authors on reasonable request.

ACKNOWLEDGEMENTS

We are grateful for discussions with P. Chaikin, Y.-D. Chuang, D. Huse, and P. Moon. This research used resources of the Advanced Light Source, which is a DOE Office of Science User Facility under contract no. DE-AC02-05CH11231. R.R.B. was supported by Purdue University startup funds. Synthesis and analysis instrumentation at NYU is supported by NSF under MRI-1531664, and from the Gordon and Betty Moore Foundation's EPIQS Initiative through Grant GBMF4838. Work at NYU was supported by the MRSEC Program of the National Science Foundation under Award Number DMR-1420073. The STM work at Rutgers is supported by NSF under grant DMR-1506618.

AUTHOR CONTRIBUTIONS

L.M. and Y.X. carried out the ARPES experiments with support from D.O., S.A.B., E.K., H. H., and J.D.D.; STM measurements were performed by W.Z., with guidance from W.W.; high-quality defect-enriched samples were developed by T.S. and J.C.; simulations were performed by Y.X. with guidance from R.R.B.; L.M., Y.X., W.Z., and L.A.W. participated in the analysis, figure planning, and draft preparation; L.A.W. was responsible for the conception and the overall direction, planning, and integration among different research units.

ADDITIONAL INFORMATION

Supplementary Information accompanies the paper on the *npj Quantum Materials* website (<https://doi.org/10.1038/s41535-018-0101-8>).

Competing interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Hasan, M. Z. & Kane, C. L. Colloquium: topological insulators. *Rev. Mod. Phys.* **82**, 3045 (2010).
- Qi, X.-L. & Zhang, S.-C. Topological insulators and superconductors. *Rev. Mod. Phys.* **83**, 1057 (2011).
- Fu, L., Kane, C. L. & Mele, E. J. Topological insulators in three dimensions. *Phys. Rev. Lett.* **98**, 106803 (2007).
- Zhang, H. et al. Topological insulators in Bi_2Se_3 , Bi_2Te_3 and Sb_2Te_3 with a single Dirac cone on the surface. *Nat. Phys.* **5**, 438–442 (2009).
- Xia, Y. et al. Observation of a large-gap topological-insulator class with a single Dirac cone on the surface. *Nat. Phys.* **5**, 398–402 (2009).
- Chen, Y. L. et al. Experimental realization of a three-dimensional topological insulator, Bi_2Te_3 . *Science* **325**, 178–181 (2009).
- Qi, X.-L., Li, R., Zang, J. & Zhang, S.-C. Inducing a magnetic monopole with topological surface states. *Science* **323**, 1184–1187 (2009).
- Yu, R. et al. Quantized anomalous Hall effect in magnetic topological insulators. *Science* **329**, 61–64 (2010).
- Chang, C.-Z. et al. Experimental observation of the quantum anomalous Hall effect in a magnetic topological insulator. *Science* **340**, 167–170 (2013).
- Fu, L. & Kane, C. L. Superconducting proximity effect and Majorana fermions at the surface of a topological insulator. *Phys. Rev. Lett.* **100**, 096407 (2008).
- Hsieh, D. et al. Observation of unconventional quantum spin textures in topological insulators. *Science* **323**, 919–922 (2009).
- Fu, L. & Kane, C. L. Topology, delocalization via average symmetry and the symplectic Anderson transition. *Phys. Rev. Lett.* **109**, 246605 (2012).
- Roushan, P. et al. Topological surface states protected from backscattering by chiral spin texture. *Nature* **460**, 1106–1109 (2009).
- Zhang, T. et al. Experimental demonstration of topological surface states protected by time-reversal symmetry. *Phys. Rev. Lett.* **103**, 266803 (2009).
- Alpichshev, Z. et al. STM imaging of electronic waves on the surface of Bi_2Te_3 : topologically protected surface states and hexagonal warping effects. *Phys. Rev. Lett.* **104**, 016401 (2010).
- Beidenkopf, H. et al. Spatial fluctuations of helical Dirac fermions on the surface of topological insulators. *Nat. Phys.* **7**, 939–943 (2011).
- Hsieh, D. et al. A tunable topological insulator in the spin helical Dirac transport regime. *Nature* **460**, 1101–1105 (2009).
- Wray, L. A. et al. Observation of topological order in a superconducting doped topological insulator. *Nat. Phys.* **6**, 855859 (2010).
- Valla, T., Pan, Z.-H., Gardner, D., Lee, Y. S. & Chu, S. Photoemission spectroscopy of magnetic and non-magnetic impurities on the surface of the Bi_2Se_3 topological insulator. *Phys. Rev. Lett.* **108**, 117601 (2012).
- Alpichshev, Z. et al. STM imaging of impurity resonances on Bi_2Se_3 . *Phys. Rev. Lett.* **108**, 206402 (2012).
- Teague, M. L. et al. Observation of Fermi-energy dependent unitary impurity resonances in a strong topological insulator Bi_2Se_3 with scanning tunneling spectroscopy. *Solid State Commun.* **152**, 747–751 (2012).
- Ugeda, M. M., Brihuega, I., Guinea, F. & GomezRodriguez, J. M. Missing atom as a source of carbon magnetism. *Phys. Rev. Lett.* **104**, 096804 (2010).
- Wehling, T. O. et al. Local electronic signatures of impurity states in graphene. *Phys. Rev. B* **75**, 125425 (2007).
- Biswas, R. R. & Balatsky, A. V. Impurity-induced states on the surface of three-dimensional topological insulators. *Phys. Rev. B* **81**, 233405 (2010).
- Black-Schaffer, A. M. & Balatsky, A. V. Subsurface impurities and vacancies in a three-dimensional topological insulator. *Phys. Rev. B* **86**, 115433 (2012).
- Black-Schaffer, A. M. & Balatsky, A. V. Strong potential impurities on the surface of a topological insulator. *Phys. Rev. B* **85**, 121103(R) (2012).
- Black-Schaffer, A. M. & Yudin, D. Spontaneous gap generation on the surface of weakly interacting topological insulators using non-magnetic impurities. *Phys. Rev. B* **90**, 161413(R) (2014).
- Xu, Y. et al. Disorder enabled band structure engineering of a topological insulator surface. *Nat. Commun.* **8**, 14801 (2017).
- Zhong, M. et al. Effect of impurity resonant states on optical and thermoelectric properties on the surface of a topological insulator. *Sci. Rep.* **7**, 3971 (2017).

30. Thouless, D. J. Electrons in disordered systems and the theory of localization. *Phys. Rep.* **13C**, 93–142 (1974).
31. Petrushevsky, M. et al. Probing the surface states in Bi_2Se_3 using the Shubnikov–de Haas effect. *Phys. Rev. B* **86**, 045131(R) (2012).
32. Dai, J. et al. Toward the intrinsic limit of the topological insulator Bi_2Se_3 . *Phys. Rev. Lett.* **117**, 106401 (2016).
33. Wray, L. A. et al. Spin-orbital ground states of superconducting doped topological insulators: a Majorana platform. *Phys. Rev. B* **83**, 224516 (2011).
34. Xue, L. et al. First-principles study of native point defects in Bi_2Se_3 . *AIP Adv.* **3**, 052105 (2013).
35. Li, H. D. et al. The van der Waals epitaxy of Bi_2Se_3 on the vicinal Si(111) surface: an approach for preparing high-quality thin films of a topological insulator. *N. J. Phys.* **12**, 103038 (2010).
36. Luttinger, J. M. & Ward, J. C. Ground-state energy of a many-fermion system. II. *Phys. Rev.* **118**, 1417–1427 (1960).
37. Luttinger, J. M. Fermi surface and some simple equilibrium properties of a system of interacting fermions. *Phys. Rev.* **119**, 1153–1163 (1960).
38. Cao, Y. et al. Mapping the orbital wavefunction of the surface states in three-dimensional topological insulators. *Nat. Phys.* **9**, 499–504 (2013).
39. Wray, L. A. et al. Chemically gated electronic structure of a superconducting doped topological insulator system. *J. Phys. Conf. Ser.* **449**, 012037 (2013).
40. Frantzeskakis, E. et al. Trigger of the ubiquitous surface band bending in 3D topological insulators. *Phys. Rev. X* **7**, 041041 (2017).
41. Kordyuk, A. A. et al. Photoemission-induced gating of topological insulators. *Phys. Rev. B* **83**, 081303 (2011).
42. Wray, L. A. et al. A topological insulator surface under strong Coulomb, magnetic and disorder perturbations. *Nat. Phys.* **7**, 32–37 (2011).
43. Bianchi, M. et al. Coexistence of the topological state and a two-dimensional electron gas on the surface of Bi_2Se_3 . *Nat. Commun.* **1**, 128 (2010).
44. Benia, H. M., Lin, C., Kern, K. & Ast, C. R. Reactive chemical doping of the Bi_2Se_3 topological insulator. *Phys. Rev. Lett.* **107**, 177602 (2011).
45. Bahramy, M. S. et al. Emergent quantum confinement at topological insulator surfaces. *Nat. Commun.* **3**, 1159 (2012).
46. Chen, C. et al. Robustness of topological order and formation of quantum well states in topological insulators exposed to ambient environment. *Proc. Natl. Acad. Sci. USA* **109**, 3694 (2012).
47. Jiang, R. et al. Reversible tuning of the surface state in a pseudobinary $\text{Bi}_2(\text{Te-Se})_3$ topological insulator. *Phys. Rev. B* **86**, 085112 (2012).



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018