Article

Tunable non-integer high-harmonic generation in a topological insulator

C. P. Schmid[1], L. Weigl[1], P. Grössing[2], V. Junk[2], C. Gorini[2,7], S. Schlauderer[1], S. Ito[3], M. Meierhofer[1], N. Hofmann[1], D. Afanasiev[1], J. Crewse[2], K. A. Kokh[4,5], O. E. Tereshchenko[5,6], J. Güdde[3], F. Evers[2], J. Wilhelm[2✉], K. Richter[2✉], U. Höfer[3] & R. Huber[1✉]

When intense lightwaves accelerate electrons through a solid, the emerging high-order harmonic (HH) radiation offers key insights into the material[1-11]. Sub-optical-cycle dynamics—such as dynamical Bloch oscillations[2-5], quasiparticle collisions[6,12], valley pseudospin switching[13] and heating of Dirac gases[10]—leave fingerprints in the HH spectra of conventional solids. Topologically non-trivial matter[14,15] with invariants that are robust against imperfections has been predicted to support unconventional HH generation[16-20]. Here we experimentally demonstrate HH generation in a three-dimensional topological insulator—bismuth telluride. The frequency of the terahertz driving field sharply discriminates between HH generation from the bulk and from the topological surface, where the unique combination of long scattering times owing to spin–momentum locking[17] and the quasi-relativistic dispersion enables unusually efficient HH generation. Intriguingly, all observed orders can be continuously shifted to arbitrary non-integer multiples of the driving frequency by varying the carrier-envelope phase of the driving field—in line with quantum theory. The anomalous Berry curvature warranted by the non-trivial topology enforces meandering ballistic trajectories of the Dirac fermions, causing a hallmark polarization pattern of the HH emission. Our study provides a platform to explore topology and relativistic quantum physics in strong-field control, and could lead to non-dissipative topological electronics at infrared frequencies.

High-order harmonic generation (HHG) in solids[1-11] is a dramatic manifestation of strong-field light–matter interaction: an atomically strong lightwave with a well defined carrier-envelope phase (CEP) acts as an a.c. bias, which can drive non-perturbative transitions between different electronic bands[3-9] and charge currents within the bands[2-13,17]. Symmetries have a key role. In inversion-symmetric media, positive and negative half cycles of the driving field induce ultrashort bursts of electromagnetic radiation, resulting in a comb of odd multiples of the driving frequency[1,2,8-10]. In contrast, even and odd orders emerge from solids with broken inversion symmetry[3,4]. For conventional semiconductors[3,4,6,9], dielectrics[2,5] and graphene[8,10,21,22], efficient scattering and dephasing destroys the electronic quantum memory on extremely short timescales[1-10], setting narrow limits to strong-field quantum control.

Non-trivial topology promises a paradigm change[15-20,23-29]. In three-dimensional topological insulators (TIs), spin–orbit interaction and time-reversal symmetry lead to a unique scenario, where an insulating bulk with an inverted band order is inseparably connected with conducting surface states[14-17,23]. The latter feature an odd number of gapless Dirac cones with helical spin texture[14] that are topologically protected against non-magnetic perturbations. Spin–momentum locking suppresses electronic backscattering and facilitates scattering times of more than 1 ps (ref. [17]). Moreover, acceleration through the Dirac point

should abruptly reverse the group velocity[8,10,21,22,26,30], causing dramatic optical nonlinearities[15] and making TIs a prime medium for HHG. Topological invariants have been predicted to imprint a characteristic polarization on the HH emission[18,19,22,28,29]. Despite these exciting prospects, HH studies of three-dimensional TIs have been hampered by the challenge to selectively investigate the surface topological Dirac fermions, the response of which is often overwhelmed by the bulk contribution.

Here we explore HHG from the prototypical three-dimensional TI bismuth telluride (Bi_2Te_3). By tuning the driving frequency, we select between HHG dominated by the bulk or the surface state and discover a novel mechanism of HHG: ballistic acceleration of Dirac currents in the topologically protected surface state (TSS). These dynamics manifest in a smooth continuous shift of the HH spectra with the CEP of the driving field, allowing for purely non-integer HHG. A characteristic alternating polarization of adjacent harmonic orders reflects the unusual shape of the Berry curvature in the vicinity of the Dirac point warranted by the topological invariant.

Separating bulk and surface HHG

The low photon energy of driving fields in the multi-terahertz (multi-THz) spectral window has proven particularly suitable to

[1]Institute of Experimental and Applied Physics, University of Regensburg, Regensburg, Germany. [2]Institute of Theoretical Physics, University of Regensburg, Regensburg, Germany. [3]Department of Physics, Philipps-University of Marburg, Marburg, Germany. [4]V.S. Sobolev Institute of Geology and Mineralogy SB RAS, Novosibirsk, Russia. [5]Novosibirsk State University, Novosibirsk, Russia. [6]A.V. Rzhanov Institute of Semiconductor Physics SB RAS, Novosibirsk, Russia. [7]Present address: Université Paris-Saclay, CEA, CNRS, SPEC, Gif-sur-Yvette, France. ✉e-mail: jan.wilhelm@ur.de; klaus.richter@ur.de; rupert.huber@ur.de
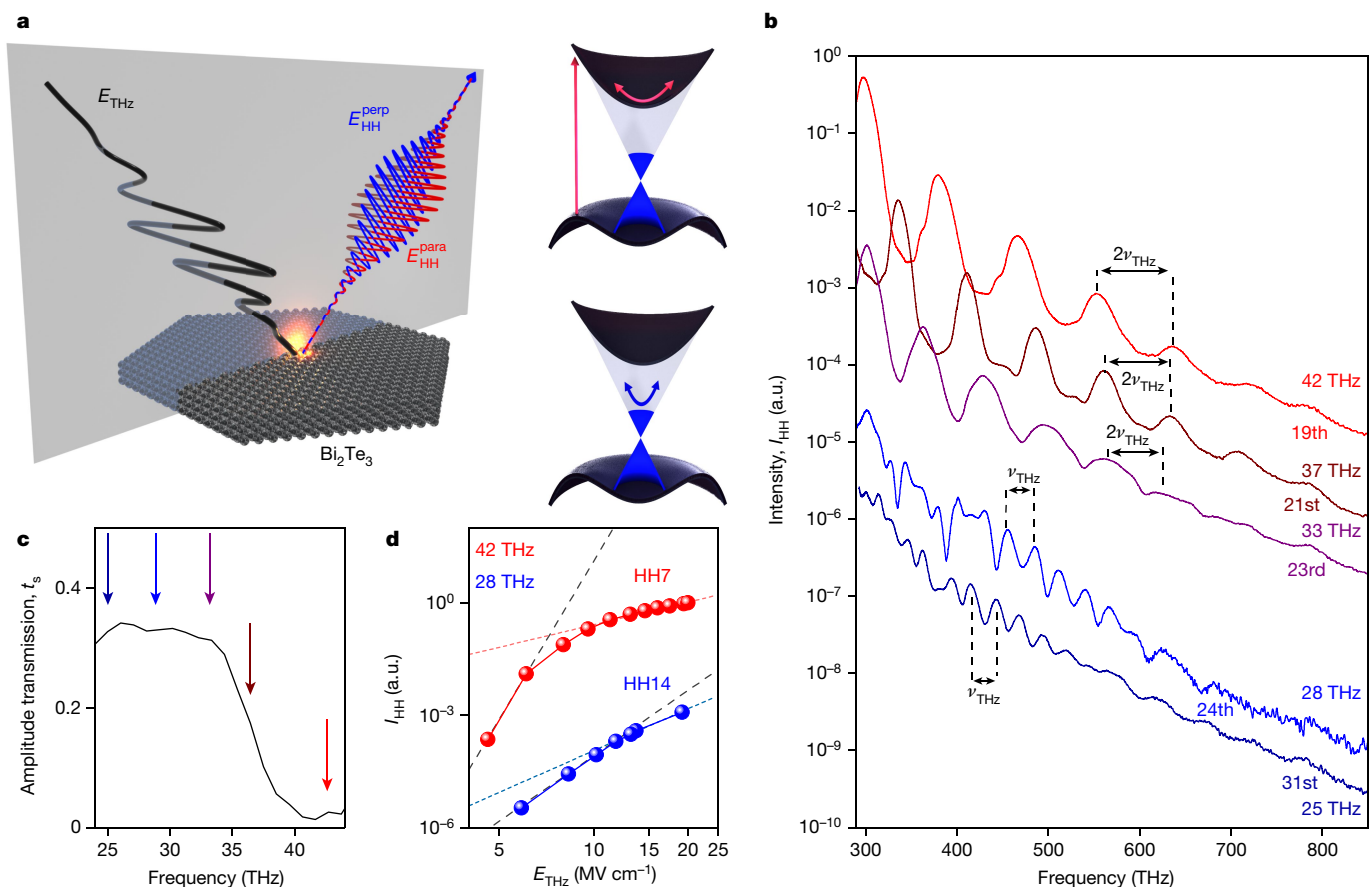
Fig. 1 | HH emission from a TI. a, Experimental scheme. An intense s-polarized THz pulse (black curve) is focused onto single-crystalline Bi_2Te_3 kept at room temperature; emerging HH radiation polarized parallel (E_{HH}^{para}, red waveform) and perpendicular (E_{HH}^{perp}, blue waveform) to the driving field is collected in reflection geometry. Inset: schematic band structure, featuring the Dirac-like dispersion of the TSS (light grey) and parabolic bulk bands (dark grey). The top and bottom sketches represent HHG from the bulk and the surface, respectively. **b**, HH spectral intensity, I_{HH}, for five driving frequencies between 25 THz and 42 THz ($E_{THz} = 20$ MV cm^{-1}). Neighbouring spectra are vertically offset by a factor of 10 for clarity. **c**, Amplitude transmission, t_s, of a 6.5-μm-thick Bi_2Te_3 crystal, obtained by THz time-domain spectroscopy. The edge at about 37 THz corresponds to the bulk bandgap. Coloured arrows mark the frequencies of the THz waveforms used for HHG in **b**. **d**, Intensity of harmonic order $n = 7$ (HH7, red, $\nu_{THz} = 42$ THz) and $n = 14$ (HH14, blue, $\nu_{THz} = 28$ THz) as a function of the peak THz electric field, E_{THz}. The black dashed line indicates a perturbative scaling law $I_{HH} \propto E^{2n} = E^{14}$. At higher field strengths, the intensity of the seventh harmonic scales much weaker with $I_{HH} \propto E^2$ (red broken line). For $\nu_{THz} = 28$ THz, the intensity of the emitted harmonics follows the scaling law $I_{HH} \propto E^{5.8\pm0.1}$ (black dashed line), for intermediate field strength, before it starts to scale as $I_{HH} \propto E^{3.7\pm0.3}$ (blue broken line), for the highest field strengths.

investigate semiconductors and Dirac materials[1,3,4,8,10,12,13,17,22–26,29,30]. We focus CEP-stable, few-cycle pulses featuring a tunable centre frequency of 25 THz $\leq \nu_{THz} \leq$ 42 THz onto a freshly cleaved (0001) surface of Bi_2Te_3 kept at room temperature (Fig. 1a). The bulk single crystal hosts a TSS with a Fermi level about 200 meV above the Dirac point (Methods). The incident peak THz electric field can be tuned up to $E_{THz} = 29$ MV cm^{-1} without damaging the sample. Owing to the large mid-infrared dielectric function of Bi_2Te_3, these fields are reduced inside the material to peak values below 3 MV cm^{-1} (Methods), which is smaller than typical peak fields used for HHG in conventional dielectrics by approximately one order of magnitude.

Despite these moderate THz field strengths, distinct HH radiation is detected in a 45° reflection geometry. Figure 1b depicts the HH spectral intensity, I_{HH}, for five THz driving frequencies, ν_{THz}. All spectra cover the infrared, visible and near-ultraviolet window, with harmonic orders up to $n = 31$. Interestingly, for $\nu_{THz} = 42$ THz, 37 THz and 33 THz, the emission is dominated by odd-order harmonics. With decreasing ν_{THz}, the overall HH intensity drops and faint even orders emerge; even-order and odd-order harmonics are emitted with comparable strength, for $\nu_{THz} = 28$ THz and $\nu_{THz} = 25$ THz. As the inversion symmetry of Bi_2Te_3 is broken at the surface only, the emergence of even orders indicates

that the spectra recorded for $\nu_{THz} \leq 28$ THz predominantly stem from the TI surface. Conversely, the odd-order HH radiation obtained for $\nu_{THz} \geq 35$ THz originates from the inversion-symmetric bulk.

The linear THz transmission spectrum of our Bi_2Te_3 sample (Fig. 1c) reveals how the choice of ν_{THz} selects between bulk and surface HHG. The drop in transmission for frequencies above 35 THz coincides with the bulk bandgap of Bi_2Te_3, E_g (Methods). Hence, THz pulses with $h\nu_{THz} \geq E_g$ (where h is Planck's constant) can photoexcite electron–hole pairs in the bulk material by resonant interband transitions. These charge carriers may subsequently be accelerated by the THz carrier wave for efficient HHG (Fig. 1a, top right inset). Conversely, for $h\nu_{THz} < E_g$, the photon energies are too small for one-photon interband transitions in the bulk and for the moderate THz fields used here, non-resonant interband transitions are not efficient enough to allow for strong HHG from the bulk either (see Methods for a quantitative estimate). Hence, the odd-order harmonic generation from the bulk becomes too weak to dominate our spectra. Yet the surface band contains a sizeable density (about 4.5×10^{12} cm^{-2}) of Dirac fermions that can be directly accelerated by the THz wave (Fig. 1a, bottom right inset). This suggests that lightwave acceleration of Dirac fermions in the TSS governs HHG
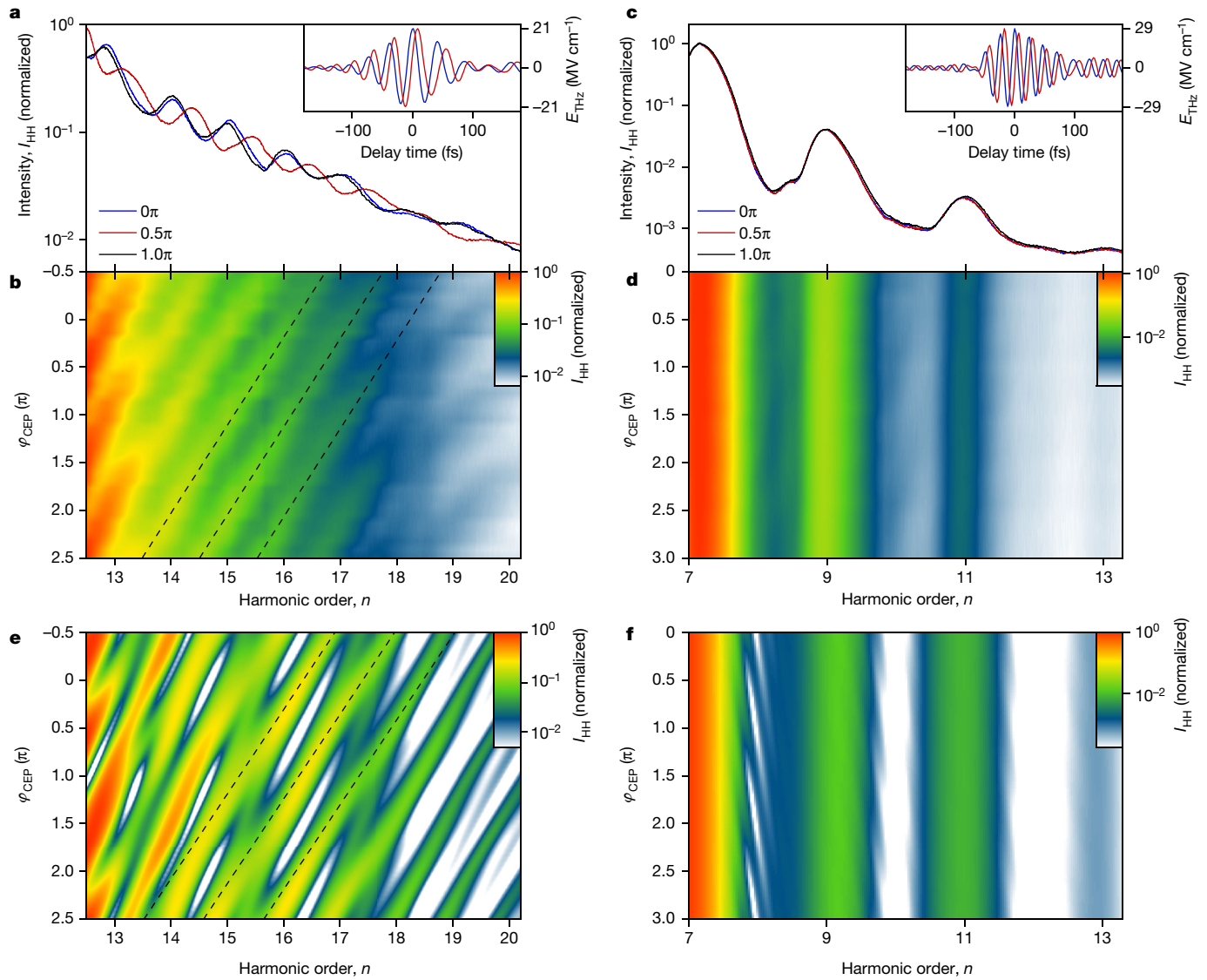
Fig. 2 | Dependence of the HH emission on the CEP of the driving field. a, Surface HH spectra for three different CEPs, $\varphi_{CEP} = 0$ (blue curve), $\varphi_{CEP} = 0.5\pi$ (red curve) and $\varphi_{CEP} = \pi$ (black curve) of the driving THz waveform ($\nu_{THz} = 25$ THz, $E_{THz} = 21$ MV cm^{-1}). Inset: electro-optically detected THz waveform for $\varphi_{CEP} = 0$ (blue curve) and $\varphi_{CEP} = 0.5\pi$ (red curve). **b**, Full CEP dependence of the HH spectra (colour map) from the TSS. The spectral position of the individual HH orders continuously changes while varying the CEP. The black dashed lines indicate the slope of the shift of I_{HH} with the CEP. **c**, Bulk-generated HH spectra, for three different CEPs of the driving waveform ($\nu_{THz} = 42$ THz,

$E_{THz} = 29$ MV cm^{-1}). The observed odd-order harmonics $n = 7, 9, 11$ and 13 show virtually no CEP dependence. **d**, A full scan reveals no strong influence of the CEP on I_{HH}, (colour scale). **e**, A simulation of I_{HH} from the TSS with a full quantum mechanical model based on the SBEs reproduces the strong CEP dependence, including the quantitative slope of the frequency shift (dashed lines). The waveform parameters, including pulse duration, field strength and spectral phase, have been adapted from the experiment. **f**, A simulation of the CEP dependence of bulk HHG confirms a negligible shift of the harmonic orders.

for $\nu_{THz} < 35$ THz, despite the low thickness of the TSS (about 1 nm) compared with typical optical absorption lengths in the bulk (about 30 to 100 nm, corresponding to a thickness ratio of about 10^{-2}). A detailed comparison of I_{HH} obtained for above- and below-bandgap driving is shown in Extended Data Fig. 3.

This scenario is further supported by the power scaling of HHG. In bulk semiconductors, HHG has been shown to scale perturbatively[2–7], for intermediate field strengths, such that the intensity of the nth harmonic grows as $I_n \approx E_{THz}^{2n}$. For yet stronger fields, non-perturbative nonlinearities set in and I_n increases more slowly, owing to subcycle interband Rabi oscillations[4]. I_{HH} follows this pattern very well for $h\nu_{THz} \geq E_g$ (Fig. 1d, red data points), corroborating that bulk HHG dominates for large ν_{THz}. In contrast, surface-dominated HHG ($h\nu_{THz} < E_g$) scales differently (Fig. 1d, blue data points), reflecting the different nature of this process.

Non-integer HHG from the surface state

Yet more intriguingly, HHG in the bulk and the TSS show fundamentally different dependences on the CEP of the THz waveform, φ_{CEP}. Whereas the HH spectra from the bulk crystal (Fig. 2c, d) show virtually no dependence on the CEP, acceleration of Dirac electrons in the TSS (Fig. 2a, b) reveals a pronounced shift of the harmonic orders with φ_{CEP}. In the TSS (Fig. 2a), the spectra for $\varphi_{CEP} = 0$ and $\varphi_{CEP} = 0.5\pi$ (see THz waveforms in the inset) feature maxima and minima at complementary positions, whereas the spectra for $\varphi_{CEP} = 0$ and $\varphi_{CEP} = \pi$ are practically identical. More systematically, all harmonic orders can be continuously shifted with the CEP (Fig. 2b) by $\Delta\nu = \nu_{THz} \times (\varphi_{CEP}/\pi)$ (black dashed lines). Notably, φ_{CEP} can be set such that the local maxima of I_{HH} are not related to any integer multiple of ν_{THz}. Also, any observed harmonic of order n can be continuously shifted to order $n + s$ by offsetting φ_{CEP} by
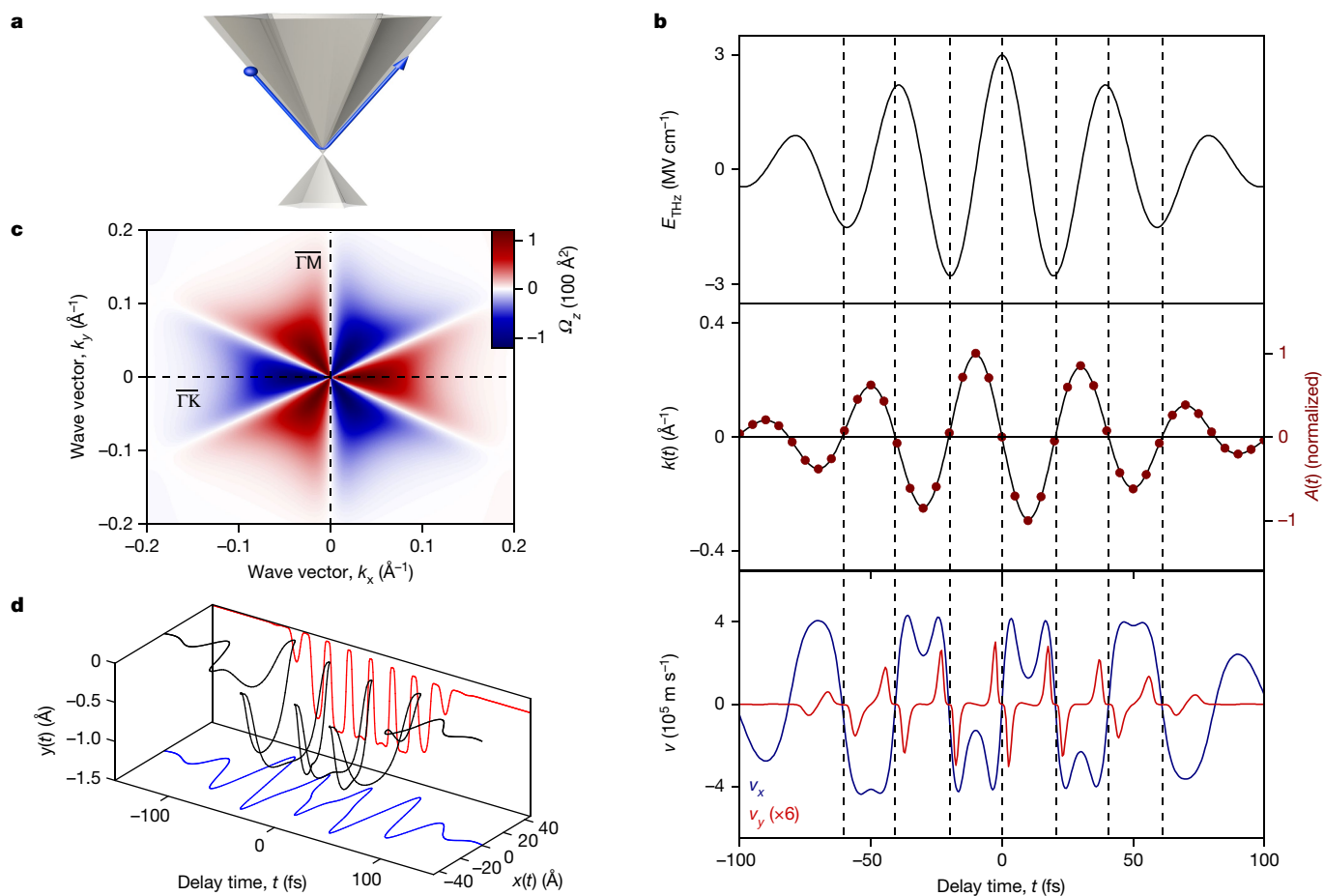
Fig. 3 | Microscopic origin of high-efficiency HHG from the TSS. a, The Dirac-like electron dispersion of the TSS leads to an abrupt change of the group velocity (slope of the band) when electrons are accelerated through the vicinity of the $\overline{\Gamma}$ point (blue trajectory). **b**, Lightwave-driven currents in the TSS. Top: THz driving waveform (frequency $\nu_{THz} = 25$ THz, internal peak THz field, $E_{THz} = 3$ MV cm^{-1}) used for the simulations. Middle: the semiclassically computed average momentum excursion, $k(t)$, of the Fermi surface of Dirac fermions (black curve, left axis) follows the vector potential, $A(t)$ (red data points, right axis), of the driving field. Bottom: corresponding group velocity components of the Dirac electrons in the TSS parallel (v_x, blue) and perpendicular (v_y, red) to the THz driving field. Both components abruptly reverse sign during zero crossings of the momentum space trajectories, that is, when the electrons traverse the Dirac point (dashed vertical lines), efficiently emitting broadband HH radiation. **c**, Out-of-plane component of the Berry curvature, Ω_z, of the TSS of Bi$_2$Te$_3$. Accelerating Dirac electrons along the $\overline{\Gamma K}$ direction can lead to an anomalous velocity contribution by the Berry curvature. **d**, Real-space trajectory of lightwave-driven Dirac electrons computed semiclassically (black curve). The blue and red lines represent projections of the trajectory onto the x and y directions.

$s \times \pi$, where s can be any real number. This CEP effect is qualitatively different from HH spectra reported for conventional insulators and semiconductors[3,5], where CEP-dependent spectral features are superimposed on CEP-independent integer-order harmonic radiation in the vicinity of the cut-off region. In contrast, all harmonic orders observed in Fig. 2a smoothly and continuously shift with φ_{CEP} without a trace of CEP-independent integer-order harmonics. This unique situation manifests in a featureless decrease of the intensity of each spectral maximum with φ_{CEP} (Extended Data Fig. 4). Importantly, this strong CEP dependence occurs only for HHG from the TSS ($h\nu_{THz} < E_g$), and is entirely absent for bulk HHG ($h\nu_{THz} > E_g$).

To pinpoint the unique nature of the underlying microscopic lightwave-driven dynamics in the TI, we simulate the experimental findings with a full quantum theory. We first compute the electronic band structure of bulk and surface states of Bi$_2$Te$_3$ and describe the TSS with an effective model Hamiltonian that reproduces the symmetry of our system, its band structure, the hexagonal warping of the spin texture, the interband dipole moments and the Berry curvature (Methods). To describe HHG from the TSS and the bulk on equal footing, we account for the possibility of both interband and intraband dynamics by solving the semiconductor Bloch equations[20] (SBEs, see Methods). The computed HH spectra attest to the fundamental difference between the TSS (Fig. 2e) and the bulk (Fig. 2f). The pronounced CEP dependence of surface-generated HH spectra, including the quantitative slope of the CEP-dependent frequency shift (black dashed lines), is faithfully confirmed. Conversely, CEP effects are negligible in the bulk HH spectra (Fig. 2f) for the orders depicted, in line with our experiment (Fig. 2d) and literature on conventional semiconductors[3].

As in other undoped semiconductors, HHG in bulk Bi$_2$Te$_3$ requires interband transitions to promote charge carriers into incompletely filled bands by subcycle Rabi oscillations, which tie the HH emission events temporally to the field crests of the carrier wave[4,7]. For $E_{THz} < 3$ MV cm^{-1}, non-resonant interband tunnelling is extremely unlikely (Extended Data Fig. 2), which explains why bulk contributions to HHG are weak for $h\nu_{THz} < E_g$. In contrast, a sizeable interband polarization is directly created by resonant excitation for $h\nu_{THz} > E_g$. The polarization then undergoes intraband acceleration or other anharmonic dynamics[1–6,8–10]. The intraband current becomes particularly anharmonic away from the parabolic band extrema where the electron–hole polarization initially emerges. Subcycle electron scattering and averaging over the spatially inhomogeneous transverse beam profile[11] suppress the effect of the driving waveform on intraband contributions.
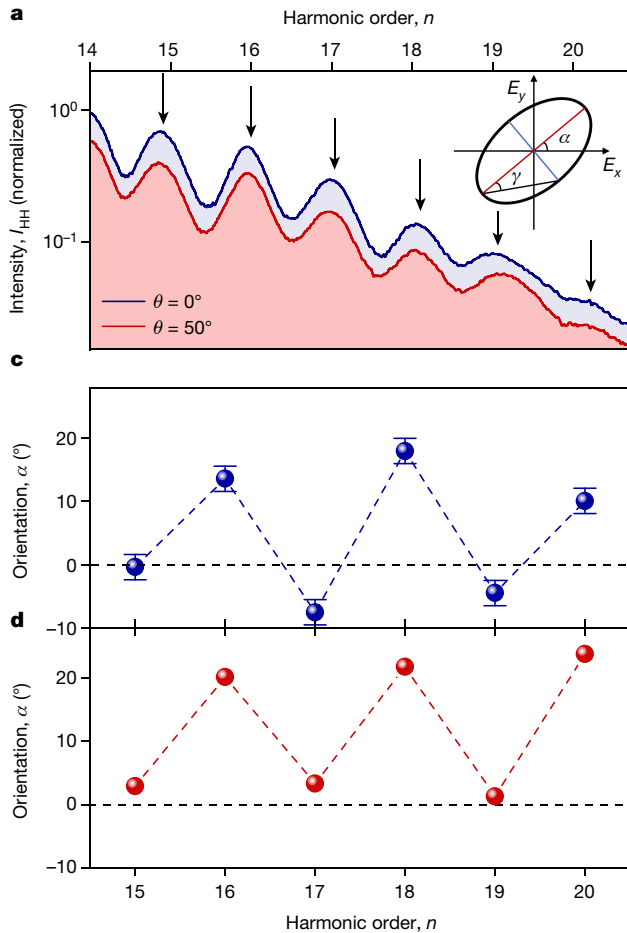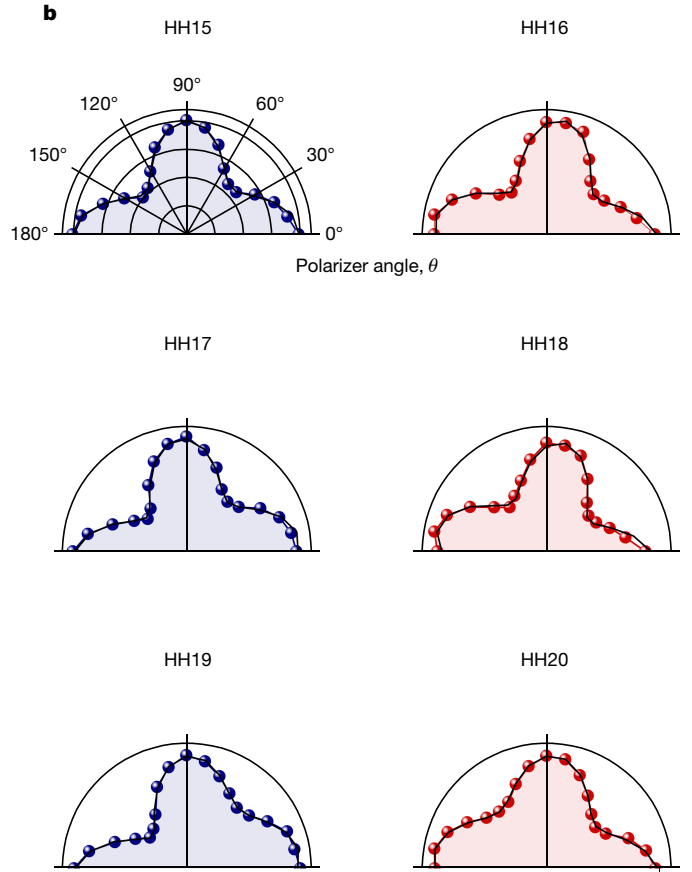
Fig. 4 | Tracing geometric phase effects in HH radiation from the TSS.
a, Polarization-resolved intensity, I_{HH}, ($v_{THz} = 25$ THz, $E_{THz} = 24$ MV cm^{-1}) along the $\overline{\Gamma K}$ direction, as measured with a polarimetry setup consisting of a rotatable quarter-wave plate and a fixed Glan–Thompson polarizer. Representative spectra for two angles of the quarter-wave plate, $\theta = 0°$ (blue curve) and $\theta = 50°$ (red curve). Inset: polarization ellipse defining the orientation angle, α, and the ellipticity angle, γ. **b**, Spectrally integrated intensity of harmonic orders $n = 15$ to $n = 20$ as a function of θ. Black curves represent a fit to the experimental data employing the Stokes formalism (Methods). **c**, The orientation angle, α, alternates between neighbouring harmonic orders, characteristic of the effect of the Berry curvature. Error bars indicate the average 95% confidence interval of the numerical extraction of α from the experimental data. **d**, The orientation angle, α, extracted from a solution of the SBEs[20] reproduces the alternating polarization pattern observed in **c**.

Thus, shifting φ_{CEP} typically affects only the highest harmonic orders while leaving the spectral shape and position of intermediate orders untouched (Fig. 2c, d, f).

HHG in the TSS is qualitatively different. (1) The pre-existing Fermi surface can be accelerated without the need for strong-field interband transitions. Indeed, switching off interband transitions in the SBEs yields qualitatively identical HH spectra and CEP dependence (Extended Data Fig. 5). (2) As spin–momentum locking suppresses scattering, Dirac fermions in Bi$_2$Te$_3$ can be accelerated ballistically over timescales exceeding 1 ps (ref. [17]). (3) The Dirac-like dispersion entails an extremely anharmonic response to even weak driving because the largest non-linearity occurs in the centre of the surface Brillouin zone ($\overline{\Gamma}$ point, see Extended Data Fig. 6), where the slope of the conduction band abruptly flips sign (Fig. 3a). According to Bloch's acceleration theorem, the THz driving field (Fig. 3b, top) shifts the centre of the Fermi circle in momentum space (Fig. 3b, centre, black curve) and the excursion follows the vector potential of the driving field (Fig. 3b, centre, red spheres). When the Dirac fermions are accelerated through the vicinity of the $\overline{\Gamma}$ point, their group velocity abruptly flips (Fig. 3b, bottom), generating high-frequency radiation at the Brillouin zone centre (see also Extended Data Fig. 6). For illustration, these curves were obtained with the semiclassical Boltzmann equation, but quantum mechanical wave-packet dynamics yield comparable results (Extended Data Fig. 7). If the Dirac

fermions move ballistically between successive cycles of the driving field, the instance when they pass the $\overline{\Gamma}$ point depends on their acceleration history. Thus, the action of the THz wave accumulates over multiple oscillation periods and the exact timing of the HH train changes with the THz waveform, chirp and CEP, causing the shift of the HH spectra observed in Fig. 2a, b, e.

Tracing geometric phase effects

These ballistic dynamics open a very rare view onto coherent transport in topological matter. As a first exciting example, we investigate a finite geometric phase effect on HHG. Figure 3c depicts the Berry curvature[27] $\mathbf{\Omega} = \nabla_k \times i\langle \Psi_+ | \nabla_k | \Psi_+ \rangle$, along different crystal directions, with Ψ_+ being the positive-energy eigenstate of the surface state Hamiltonian (Methods) and k denoting the crystal momentum . $\mathbf{\Omega}$ vanishes along the high-symmetry direction $\overline{\Gamma M}$, where net effects on HHG cancel. In contrast, for an acceleration in the $\overline{\Gamma K}$ direction, the Berry curvature can induce an average perpendicular velocity component, v_y (Fig. 3b, red curve), which abruptly changes upon acceleration through the centre of the Brillouin zone. The resulting meandering trajectory (Fig. 3d) should manifest in a perpendicular polarization component of the HH spectra. Note that v_x flips sign with alternating positive and negative slopes in successive THz half cycles, whereas v_y shows the same

Article

negative slope for all zero crossings (Fig. 3b). Therefore, v_x and v_y cause odd-order and even-order harmonics, respectively, and odd and even orders should feature different polarizations[19].

To test this conjecture, we performed systematic polarimetry[29] by recording the HH intensity after a fixed polarizer, while a quarter-wave plate placed before the polarizer was continuously rotated by an angle θ (see exemplary spectra in Fig. 4a). The resulting polar diagram retrieved for each harmonic order n encodes the polarization state (Fig. 4b). For $n = 15$, a regular four-lobe pattern aligned with the driving field is obtained; the patterns of all other orders are rotated. From these data, we can uniquely extract the degree of polarization, σ, the orientation angle, α, and the ellipticity angle, γ (Fig. 4a, inset, Extended Data Fig. 8), which perfectly describe the polarimetry data (Fig. 4b, black lines). Strikingly, α alternates between adjacent orders (Fig. 4c); it remains small for odd orders and reaches values of up to 20° for even orders. Because the ellipticity angle γ remains small throughout (Extended Data Fig. 8), these modulations describe a polarization rotation of adjacent HH orders—in qualitative agreement with the above prediction and a quantum simulation (Fig. 4d). Our results corroborate that the surface-based HH emission originates from the ballistic motion of quasi-relativistic fermions through the Dirac point while the electron wavefunctions accumulate a geometric phase.

Concluding remarks

We have demonstrated HHG in a TI. By choosing a driving frequency above the bulk bandgap, bulk-dominated odd-order HHG featuring no CEP dependence is observed. If v_{THz} is tuned below E_g/h, interband excitation across the bulk bandgap is strongly suppressed and the HH spectra are dominated by contributions from the metallic surface state, resulting in even-order and odd-order harmonics. An unprecedented dependence of the emitted spectra on the CEP of the driving field reveals that HHG in the surface state is directly linked to the presence of a Dirac point. Moreover, an anomalous polarization component unveils the influence of the Berry curvature in the TSS. Strong-field ballistic control of topologically protected surface states opens a unique laboratory to test fundamental predictions of relativistic quantum physics. Specifically, the abrupt spin flip occurring directly at the Dirac point deserves additional experimental and theoretical attention. Going beyond the dominant intraband currents studied here, intriguing signatures of Zitterbewegung (see quantum simulations in Extended Data Fig. 9), Landau–Zener tunnelling and strong-field axion dynamics in a topologically non-trivial environment may soon become accessible to experimental observation.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03466-7.

1. Chin, A. H., Calderón, O. G. & Kono, J. Extreme midinfrared nonlinear optics in semiconductors. *Phys. Rev. Lett.* **86**, 3292–3295 (2001).
2. Ghimire, S. et al. Observation of high-order harmonic generation in a bulk crystal. *Nat. Phys.* **7**, 138–141 (2011).
3. Schubert, O. et al. Sub-cycle control of terahertz high-harmonic generation by dynamical Bloch oscillations. *Nat. Photon.* **8**, 119–123 (2014).
4. Hohenleutner, M. et al. Real-time observation of interfering crystal electrons in high-harmonic generation. *Nature* **523**, 572–575 (2015).
5. Luu, T. T. et al. Extreme ultraviolet high-harmonic spectroscopy of solids. *Nature* **521**, 498–502 (2015).
6. Vampa, G. et al. Linking high harmonics from gases and solids. *Nature* **522**, 462–464 (2015); corrigendum **542**, 260 (2017).
7. Garg, M. et al. Multi-petahertz electronic metrology. *Nature* **538**, 359–363 (2016).
8. Yoshikawa, N., Tamaya, T. & Tanaka, K. High-harmonic generation in graphene enhanced by elliptically polarized light excitation. *Science* **356**, 736–738 (2017).
9. Sivis, M. et al. Tailored semiconductors for high-harmonic optoelectronics. *Science* **357**, 303–306 (2017).
10. Hafez, H. A. et al. Extremely efficient terahertz high-harmonic generation in graphene by hot Dirac fermions. *Nature* **561**, 507–511 (2018).
11. Floss, I. et al. Ab initio multiscale simulation of high-order harmonic generation in solids. *Phys. Rev A.* **97**, 011401(R) (2018).
12. Langer, F. et al. Lightwave-driven quasiparticle collisions on a subcycle timescale. *Nature* **533**, 225–229 (2016).
13. Langer, F. et al. Lightwave valleytronics in a monolayer of tungsten diselenide. *Nature* **557**, 76–80 (2018).
14. Chen, Y. L. et al. Experimental realization of a three-dimensional topological insulator, Bi_2Te_3. *Science* **325**, 178–181 (2009).
15. Giorgianni, F. et al. Strong nonlinear terahertz response induced by Dirac surface states in Bi_2Se_3 topological insulator. *Nat. Commun.* **7**, 11421 (2016).
16. Bauer, D. & Hansen, K. K. High-harmonic generation in solids with and without topological edge states. *Phys. Rev. Lett.* **120**, 177401 (2018).
17. Reimann, J. et al. Subcycle observation of lightwave-driven Dirac currents in a topological surface band. *Nature* **562**, 396–400 (2018).
18. Silva, R. E. F., Jiménez-Galán, A., Amorim, B., Smirnova, O. & Ivanov, M. Topological strong field physics on sub-laser cycle timescale. *Nat. Photon.* **13**, 849–854 (2019).
19. Baykusheva, D. et al. Strong-field physics in three-dimensional topological insulators. *Phys. Rev. A* **103**, 023101 (2021).
20. Wilhelm, J. et al. Semiconductor-Bloch formalism: derivation and application to high-harmonic generation from Dirac fermions. *Phys. Rev. B* **103**, 125419 (2021).
21. Higuchi, T., Heide, C., Ullmann, K., Weber, H. B. & Hommelhoff, P. Light-field-driven currents in graphene. *Nature* **550**, 224–228 (2017).
22. McIver, J. W. et al. Light-induced anomalous Hall effect in graphene. *Nat. Phys.* **16**, 38–41 (2020).
23. Kuroda, K., Reimann, J., Güdde, J. & Höfer, U. Generation of transient photocurrents in the topological surface state of Sb_2Te_3 by direct optical excitation with midinfrared pulses. *Phys. Rev. Lett.* **116**, 076801 (2016).
24. Wu, L. et al. Quantized Faraday and Kerr rotation and axion electrodynamics of a 3D topological insulator. *Science* **354**, 1124–1127 (2016).
25. Mahmood, F. et al. Selective scattering between Floquet–Bloch and Volkov states in a topological insulator. *Nat. Phys.* **12**, 306–310 (2016).
26. Braun, L. et al. Ultrafast photocurrents at the surface of the three-dimensional topological insulator Bi_2Se_3. *Nat. Commun.* **7**, 13259 (2016).
27. Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. Lond. A* **392**, 45–57 (1984).
28. Luu, T. T. & Wörner, H. J. Measurement of the Berry curvature of solids using high-harmonic spectroscopy. *Nat. Commun.* **9**, 916 (2018).
29. Banks, H. B. et al. Dynamical birefringence: electron–hole recollisions as probes of Berry curvature. *Phys. Rev. X* **7**, 041042 (2017).
30. Cheng, B. et al. Efficient terahertz harmonic generation with coherent acceleration of electrons in the Dirac semimetal Cd_3As_2. *Phys. Rev. Lett.* **124**, 117402 (2020).

Methods

Experimental setup

A femtosecond titanium-sapphire laser amplifier (repetition rate, 3 kHz; pulse energy, 5.5 mJ; pulse duration, 33 fs; centre wavelength, 808 nm) is used to pump two parallel dual-stage optical parametric amplifiers, which deliver two signal pulses with tunable output wavelengths between 1.1 μm and 1.7 μm (pulse energy 0.5 mJ). Difference frequency generation between the two spectrally detuned phase-correlated near-infrared pulses from the optical parametric amplifiers generates intense, phase-locked waveforms in the multi-THz spectral region. The subcycle time delay between the near-infrared generation pulses sets the CEP of the THz pulses. The centre frequency of the few-cycle THz pulses is varied between 25 THz and 42 THz, while the electric peak fields reach values of multiple 10 MV cm^{-1} (ref. [31]). The field transients are focused onto the Bi$_2$Te$_3$ sample under an angle of incidence of 45°. The emitted HH radiation is collimated and the corresponding spectra are recorded with a spectrograph coupled to a thermoelectrically cooled silicon charge-coupled device camera, setting a natural long-wavelength limit of detection of about 1.2 μm. The diffraction efficiency of the grating and the quantum efficiency of the detector are corrected for in all spectra. The THz waveforms are electro-optically sampled in a 6.5-μm-thin zinc telluride crystal. For the spectrum shown in Fig. 1c, an ultrabroadband THz transient was transmitted through a thin Bi$_2$Te$_3$ window (thickness, 6.5 μm) under normal incidence and electro-optically detected. Recording a reference waveform transmitted through air allowed for the extraction of the transmission coefficient (Fig. 1c), which indicates a bandgap of about 150 meV, in good agreement with literature values[32]. Moreover, the transmission coefficient allows us to calculate the refractive index of the Bi$_2$Te$_3$ crystal, which was found to be $n_r = 10$, in line with literature values of $n_r = 9.2$ (ref. [33]). Using this value of the refractive index as well as the experimental angle of incidence of 45°, we retrieve a peak field strength in the bismuth telluride crystal of roughly $E_{THz} = 3$ MV cm^{-1}.

Sample preparation

The Bi$_2$Te$_3$ sample was grown by a floating zone Bridgman technique. By choosing an appropriate solidification condition, a gradation in the carrier concentration was realized in an ingot naturally forming a p–n junction[34]. A clean and well-ordered surface was obtained by cleaving the sample using the Scotch tape method. The samples were slightly p-doped and the Dirac point was found to be located 200 meV below the Fermi energy, by angle-resolved photoemission spectroscopy. The sample orientation was determined by low-energy electron diffraction (Extended Data Fig. 1).

Model Hamiltonian for Bi$_2$Te$_3$

A realistic description of the energy landscape of charge carriers in the TSS is the starting point for the simulation of the lightwave-driven electron dynamics. To this end, we set up a model $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian that features the correct symmetry properties[35]:

$$H_{TSS}(\mathbf{k}) = C_0 + C(\mathbf{k})\mathbf{k}^2 + A\left(\sigma_x k_y - \sigma_y k_x\right) + 2R(\mathbf{k})\left(k_x^3 - 3k_y^2 k_x\right)\sigma_z. \quad (1)$$

Here σ_x, σ_y and σ_z describe the Pauli matrices and k_x and k_y represent the wave vectors along the x ($\overline{\Gamma K}$) and y ($\overline{\Gamma M}$) directions, respectively. The coefficients C_0, $C(\mathbf{k})$, A and $R(\mathbf{k})$ were determined by fitting the corresponding eigenenergies $\epsilon(\mathbf{k})$ to a band structure calculated by density functional theory as follows:

$$C(\mathbf{k}) = \frac{c_2}{k_{sym}^2} \frac{1}{1 + \left(\frac{k}{k_{sym}}\right)^2}, \quad (2)$$

$$R(\mathbf{k}) = \frac{r_2}{k_{asym}^2} \frac{1}{1 + \left(\frac{k}{k_{asym}}\right)^4}, \quad (3)$$

where $c_2 = 0.32$ eV, $k = |\mathbf{k}|$, $k_{sym} = 0.12$ Å^{-1}, $r_2 = 1.57$ eV Å and $k_{asym} = 0.215$ Å^{-1}. The values for the remaining coefficients are given by $C_0 = 0.176$ eV and $A = 0.609$ eV Å. Using the Hamiltonian in equation (1), we also calculated the Berry curvature[27,36,37] $\boldsymbol{\Omega}(\mathbf{k})$ in reciprocal space (Fig. 3c).

Carrier-injection into bulk states and comparison of bulk- and surface-state HHG

Although resonant injection of charge carriers into the bulk conduction band is not possible, for $v_{THz} < E_g$, electrons can still undergo Zener tunnelling. With an internal peak THz electric field of about 3 MV cm^{-1} and a carrier frequency of 25 THz, we obtain a Keldysh parameter of $\gamma < 1$, where interband excitations are well described by a tunnelling integral. Assuming a realistic value for the reduced effective band mass in the bulk $m_r = m_e$ (m_e, bare electron mass), we can easily estimate[38,39] the density of electrons tunnelling from the valence into the conduction band within one half cycle of the driving field. Inserting the electric field strength $E_{THz} = 3$ MV cm^{-1} applied in the experiment, we calculated a three-dimensional electron–hole pair density of the order of 5×10^{19} cm^{-3} (Extended Data Fig. 2a). This value corresponds to a sheet density of 5×10^{12} cm^{-2} in a 1-nm-thick layer, comparable to the density of Dirac fermions (4.5×10^{12} cm^{-2}) in our topological surface layer. Yet, massive band electrons and holes in the bulk are much less efficient in HHG than Dirac fermions: for intraband currents of massive electron–hole pairs to contribute to HHG in the visible spectral range, the charge carriers would have to be accelerated far out of the parabolic region of their band structure. In sharp contrast, Dirac fermions feature by far the strongest nonlinearity in the vicinity of the centre of the Brillouin zone. At the moderate THz fields used in the experiment, the efficiency of HHG with massive electron–hole pairs in bulk is therefore strongly suppressed compared with the case of the Dirac fermions in the topological surface state (compare Extended Data Fig. 2b).

Conversely, if resonant one-photon absorption becomes possible, the electron–hole pair density strongly rises. For excitation at 40 THz and a THz absorption length of 1 μm, one obtains a three-dimensional electron–hole pair density of as much as 5×10^{21} cm^{-3}. At these densities, the transition should be strongly driven into a nonlinear response regime[40] — all of which boosts the efficiency of HHG. This explains why the efficiency of HHG from the bulk abruptly increases for THz driving frequencies exceeding the bulk bandgap.

Simulation of HHG in bulk and TSS with the SBEs

The CEP dependence of the HH spectra was modelled by SBEs[20,41] using the two-band model (equation (1)) of the Bi$_2$Te$_3$ surface state. For a realistic description of the experiments, we drove the dynamics by an ultrashort Gaussian pulse

$$E(t) = E_{THz} e^{-\frac{t^2}{\tau^2}} \sin(2\pi v_{THz} t(1 + f_{chirp} t) - \varphi_{CEP}), \quad (4)$$

with time, t, pulse duration, τ, a variable CEP, φ_{CEP}, and a fixed chirp, f_{chirp}, adapted to fit the experimentally measured waveform. For instance, the below-bandgap simulations of Fig. 2 were carried out for $E_{THz} = 3$ MV cm^{-1}, $v_{THz} = 25$ THz, $\tau = 45$ fs and $f_{chirp} = -1.25$ THz. The Fermi level was set to 0.2 eV above the band crossing at the $\overline{\Gamma}$ point confirmed with angle-resolved photoemission spectroscopy of the sample used in the experiment. The decay of off-diagonal density matrix elements towards equilibrium was modelled with a time constant of 10 fs, which represents an accepted simulation value[11]. Intraband currents (diagonal density matrix elements) decay towards their equilibrium distribution with a scattering time $T_1 = 1$ ps as

Article

demonstrated in ref. [17]. The dephasing thus introduced should be considered as an approximative means to account for experimental effects that cause decoherence and that are not explicitly included in our modelling. As an example, we mention the spatial inhomogeneity of the laser intensity, which leads to an incoherent superposition of signals irradiated from different sample regions. We employed a 420×42 k-point grid for sampling the hexagonal two-dimensional Brillouin zone. Although the shape of the emission spectrum is sensitive to the specific parameter choice, the CEP dependence of the harmonic orders (Fig. 2e, f) and the polarization pattern (Fig. 4d) remain robust under parameter variation.

For the CEP dependence of the HH spectra in the bulk Bi_2Te_3, we employed a two-band model with the band structure

$$H_{\text{bulk}}(\mathbf{k}) = 0.05\,\text{eV}\,\sigma_x + 2.0\,\text{eV}\left[2.033 - \cos(1.5ak_x) - \cos(1.5ak_y)\right]\sigma_z, \quad (5)$$

where σ_x and σ_z are Pauli matrices and $a = 4.38$ Å is the lattice constant of Bi_2Te_3. The parameters are chosen such that the maximum dipole moment computed from the Hamiltonian is 2 eÅ—a typical value for semiconductors[3,4]. We drove the dynamics by an ultrashort Gaussian pulse (equation (4)) with a peak field strength inside the medium, $E_{\text{THz}} = 3\,\text{MV cm}^{-1}$ and a frequency $\nu_{\text{THz}} = 40$ THz. The pulse duration and chirp of the pulses was adapted to match the experimentally measured transient (see inset Fig. 2c). For simulating the bulk emission, we started the dynamics with a filled valence band and an empty conduction band. Here, diagonal density matrix elements decay with a time constant of 10 fs, whereas off-diagonal terms are damped with a decay time of 1 fs. For the SBE simulations, we have used our in-house program package CUED (see 'Code availability').

HHG in the semiclassical limit

By solving the SBEs[20], we fully account for intraband and interband dynamics on equal footing, which is a prerequisite for HHG in intrinsic semiconductors featuring a finite bandgap. In these semiconductors, there would be no HHG in the absence of interband excitations—a process not covered by semiclassical intraband currents. In contrast, HHG from the topological surface state can occur even without interband transitions, merely by ballistically accelerating the pre-existing Fermi circle, which can be consistently modelled by the Boltzmann equation. Indeed, by switching off all dipole moments in the SBEs[20], they morph into the Boltzmann equation, which yield qualitatively identical HH spectra and CEP dependence (Extended Data Fig. 5) as the full SBE approach.

Semiclassical simulation of HHG in the topological surface

For an intuitive first-order approximation, the microscopic mechanism of HHG in the TSS can be modelled starting with the semiclassical equations of motion for an electron in an electric field, $\mathbf{E}$ (ref. [37]):

$$\dot{\mathbf{r}} = \frac{1}{\hbar}\frac{\partial\epsilon(\mathbf{k})}{\partial\mathbf{k}} - \dot{\mathbf{k}} \times \mathbf{\Omega}(\mathbf{k}), \quad (6)$$

$$\hbar\dot{\mathbf{k}} = -e\mathbf{E}. \quad (7)$$

Here, $\dot{\mathbf{r}}$ describes the real space velocity and $\dot{\mathbf{k}}$ is the time derivative of the electrons' crystal momentum. Moreover, $\hbar$ is the reduced Planck's constant, e is the elementary charge, and ϵ and $\mathbf{\Omega}$ describe the band structure and the corresponding Berry curvature extracted from the Hamiltonian of equation (1). As we are interested in the acceleration within the two-dimensional surface state, we focus on the in-plane motion. Thus, the velocities in real and reciprocal space are given by $\dot{\mathbf{r}} = (\dot{r}_x, \dot{r}_y, 0)^T$ and $\dot{\mathbf{k}} = (\dot{k}_x, \dot{k}_y, 0)^T$, where the superscript T indicates transpose. To describe the motion of electrons within the topological surface state in reciprocal space, we employed the Boltzmann equation[42]:

$$\frac{\partial}{\partial t}f(\mathbf{k}, t) = -\frac{e}{\hbar}\mathbf{E}(t)\nabla_{\mathbf{k}}f(\mathbf{k}, t) + \left(\frac{\partial f(\mathbf{k}, t)}{\partial t}\right)_{\text{scattering}}, \quad (8)$$

with $f(\mathbf{k}, t)$ denoting the distribution of electrons in momentum space. The first term on the right-hand side describes the redistribution of electrons in the presence of an electromagnetic field (compare equation (7)), whereas the second term includes all changes of $f(\mathbf{k}, t)$ caused by scattering within the surface state. Owing to spin–momentum locking, which leads to extremely large scattering times of more than 1 ps (ref. [17]), this term is almost negligible. Solving equation (8) for a THz waveform similar to the experimental condition (peak field strength inside the medium, $E_{\text{THz}} = 3\,\text{MV cm}^{-1}$, $\nu_{\text{THz}} = 25$ THz), we obtained the full temporal evolution of the momentum distribution function in the TSS. From the velocity of the charge carriers averaged over momentum space::

$$\langle\dot{\mathbf{r}}\rangle = \langle f(\mathbf{k})\dot{\mathbf{r}}(\mathbf{k})\rangle, \quad (9)$$

and its Fourier transform, we can calculate the emitted HH spectra and their CEP dependence.

To demonstrate where an individual order is generated in momentum space, we solved equations (6) and (7) under the action of a transient electric field pointing along the x direction. We avoided trivial perpendicular velocity contributions by symmetrically placing two test charges with a separation of $k_y = \pm 0.001$ Å^{-1} from the Dirac point (perpendicular to the direction of the driving field) and calculated the emitted HH spectra (Extended Data Fig. 6a). The momentum space origin of HHG can be located by tracking the intensity of a given harmonic of order n as a function of the starting point of the momentum space trajectory (k_x, k_y). Extended Data Fig. 6b shows the intensity of harmonic order $n = 15$ (colour scale) as a function of the starting point in momentum space. Electrons placed close to the Dirac point perceive the strongest nonlinearity whereas electrons starting far away from the Dirac point do not significantly contribute to the HHG emission.

Stokes–Mueller formalism

To measure the polarization state of the emitted HH radiation, we used a polarimetry setup consisting of a quarter-wave plate on a rotating mount and a fixed Glan–Thompson polarizer oriented parallel to the THz driving field. Employing the Stokes formalism with the Stokes parameters S_0, S_1, S_2 and S_3 and the Mueller matrices for quarter-wave plate and polarizer, we find the following expression for the intensity, $I(\theta)$ (ref. [29,43]):

$$I(\theta) = S_0 - \frac{S_1}{2}(1 + \cos(\delta)) - S_3\sin(\delta)\sin(2\theta) - \frac{S_1}{2}(1 - \cos(\delta))$$
$$\cos(4\theta) - \frac{S_2}{2}(1 - \cos(\delta))\sin(4\theta). \quad (10)$$

Here θ represents the rotation angle and δ represents the retardance of the quarter-wave plate. The wavelength dependence of δ was measured and corrected for. Fitting equation (10) to the experimental data allowed us to retrieve the complete polarization state defined by the Stokes parameters. To identify possible anomalous polarization components from the Berry curvature, we extracted the spherical coordinates α and γ, which span the polarization ellipse, as well as the degree of polarization σ (ref. [29]):

$$2\alpha = \arctan\left(\frac{S_2}{S_1}\right), \quad (11)$$

$$2\gamma = \arctan\left(\frac{S_3}{\sqrt{S_1^2 + S_2^2}}\right), \quad (12)$$

$$\sigma = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0}. \tag{13}$$

In Extended Data Fig. 8, the extracted parameters are shown for the polarization-resolved measurement illustrated in Fig. 4a, b.

Quantum mechanical wave-packet dynamics in Bi_2Te_3

The fundamental HHG mechanism by ballistic Dirac currents computed semiclassically in Fig. 3b can be confirmed more rigorously by numerically propagating quantum mechanical wave packets of select test charges. To this end, we solved the full time-dependent Schrödinger equation[44] including the electric field of the THz pulse ($E_{THz} = 1\,MV\,cm^{-1}$, $v_{THz} = 25\,THz$), that is, $H(\mathbf{k}, t) = H_{TSS}(\mathbf{k}) + eE(t)x$. We employed the C++ library 'time-dependent quantum transport' by Krueckl[45] for our simulations. A more detailed description of the library is given in ref. [44]. Using the model Hamiltonian for Bi_2Te_3 from equation (1), we computed the total evolution of two Gaussian wave packets centred at $k_{x,0} = 0\,Å^{-1}$ and $k_{y,0} = \pm 0.04\,Å^{-1}$

$$\Psi(\mathbf{k}, t) = \sqrt{\frac{1}{\sqrt{\pi}\,\Delta k}} \exp\left(-\frac{1}{2\Delta k^2}\left[(k_x - k_{x,0})^2 + (k_y - k_{y,0})^2\right]\right) \cdot \phi(\mathbf{k}, 0), \tag{14}$$

where $\phi(\mathbf{k}, 0)$ is the positive eigenstate of $H_{TSS}(\mathbf{k})$ and $\Delta k = 0.002\,Å^{-1}$. Similar to the semiclassical model, v_x exhibits a strongly anharmonic evolution (Extended Data Fig. 7), explaining efficient HHG even in the presence of transitions between the two bands, as included in the wave-packet calculations. The qualitative features of v_y also resemble the semiclassical model, but are smoother and show an amplitude of similar order of magnitude as v_x. As we did not include Pauli blocking mechanisms and decoherence in the model, it might overestimate interband contributions but clearly confirms the influence of the Dirac-like band structure and the Berry curvature on carrier motion.

Potential signatures of Zitterbewegung

The influence of interband transitions on the wave-packet motion becomes evident when increasing the driving field above $E_{THz} = 3\,MV\,cm^{-1}$ and considering the same wave packets as in the last section. We obtain a similar temporal evolution of v_x (not shown) as in Fig. 3b with stronger anharmonic features than for $E_{THz} = 1\,MV\,cm^{-1}$. In v_y (Extended Data Fig. 9), however, additional oscillations in the frequency window between harmonic order $n = 2$ and $n = 5$ are expected to arise, which are not yet observed in our experiment. They are easy to distinguish from Berry curvature effects as they peak at zero crossings of the THz field and show two key characteristics of Zitterbewegung. (1) They only occur when states living in the different branches of a Dirac system interfere with each other. The high-frequency oscillations start to appear after the electric field becomes strong enough to drive Landau–Zener transitions between both bands. (2) The frequency of Zitterbewegung depends on the energy separation of the interfering states. As one would expect from this, the frequency of the oscillations increases as long as the amplitude of the driving THz field E_{THz} increases and decreases again when the amplitude of E_{THz} returns to zero. Observable signatures of Zitterbewegung are certainly overestimated in this model calculation because of the neglection of decoherence and the use of only two test charges instead of the full Fermi surface. Yet, these predictions forecast which signatures one

may have to search for in future experiments of ballistic strong-field Dirac control.

Data availability
The data supporting the findings of this study are available from the corresponding authors upon request. Source data are provided with this paper.

Code availability
The in-house program package CUED that was used to solve the SBE is freely available from GitHub (https://github.com/ccmt-regensburg/CUED).

31. Sell, A., Leitenstorfer, A. & Huber, R. Phase-locked generation and field-resolved detection of widely tunable terahertz pulses with amplitudes exceeding 100 MV/cm. *Opt. Lett.* **33**, 2767–2769 (2008).
32. Michiardi, M. et al. Bulk band structure of Bi_2Te_3. *Phys. Rev. B* **90**, 075105 (2014).
33. Austin. I. G. The optical properties of bismuth telluride. *Proc. Phys. Soc.* **72**, 545–552 (1958).
34. Kokh, K. A. et al. Melt growth of bulk Bi_2Te_3 crystals with a natural p–n junction. *CrystEngComm* **16**, 581–584 (2014).
35. Liu, C.-X. et al. Model Hamiltonian for topological insulators. *Phys. Rev. B* **82**, 045122 (2010).
36. Xiao, D., Chang, M.-C. & Niu, Q. Berry phase effects on electronic properties. *Rev. Mod. Phys.* **82**, 1959–2007 (2010).
37. Gradhand, M. et al. First-principle calculations of the Berry curvature of Bloch states for charge and spin transport of electrons. *J. Phys. Condens. Matter* **24**, 213202 (2012).
38. Kane, E. O. Zener tunneling in semiconductors. *J. Phys. Chem. Solids* **12**, 181–188 (1960).
39. Lange, C. et al. Extremely nonperturbative nonlinearities in GaAs driven by atomically strong terahertz fields in gold metamaterials. *Phys. Rev. Lett.* **113**, 227401 (2014).
40. Junginger, F. et al. Nonperturbative interband response of a bulk InSb semiconductor driven off resonantly by terahertz electromagnetic few-cycle pulses. *Phys. Rev. Lett.* **109**, 147403 (2012).
41. Kira, M. & Koch, S. W. *Semiconductor Quantum Optics* (Cambridge Univ. Press, 2012).
42. Mikhailov, S. A. Non-linear electromagnetic response of graphene. *Europhys. Lett.* **79**, 27002 (2007).
43. Huard, S. *Polarization of Light* (Wiley, 1997).
44. Junk, V., Reck, P., Gorini, C. & Richter, K. Floquet oscillations in periodically driven Dirac systems. *Phys. Rev. B* **101**, 134302 (2020).
45. Krückl, V. *Wave Packets in Mesoscopic Systems: From Time-dependent Dynamics to Transport Phenomena in Graphene and Topological Insulators*. PhD thesis, Univ. Regensburg (2013).

Author contributions F.E., J.W., K.R., U.H. and R.H. conceived the study. K.A.K. and O.E.T. provided the high-quality Bi_2Te_3 samples and performed transport measurements. C.P.S., L.W., S.S., S.I., M.M., N.H., D.A. and J.G. set up and carried out the optical experiments and characterized the sample orientation. C.P.S., L.W., P.G., J.C., F.E. and J.W. developed and carried out the semiconductor Bloch equations simulations. V.J., C.G. and K.R. set up and carried out the quantum mechanical wave-packet simulation. C.P.S., L.W. and P.G. implemented the semiclassical calculations. All authors analysed the data and discussed the results. C.P.S. and R.H. wrote the manuscript with contributions from all authors.
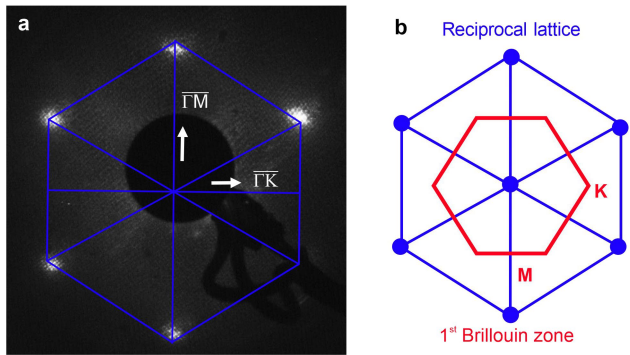
Article

Extended Data Fig. 1 | Crystallographic orientation of the Bi$_2$Te$_3$ sample.
**a**, Low-energy electron diffraction of the Bi$_2$Te$_3$ sample measured with an electron energy of 78 eV. The white spots mark the reciprocal lattice vectors.
b, Schematic of the reciprocal lattice vectors (blue) overlaid with the surface Brillouin zone (red) and the corresponding high-symmetry directions.
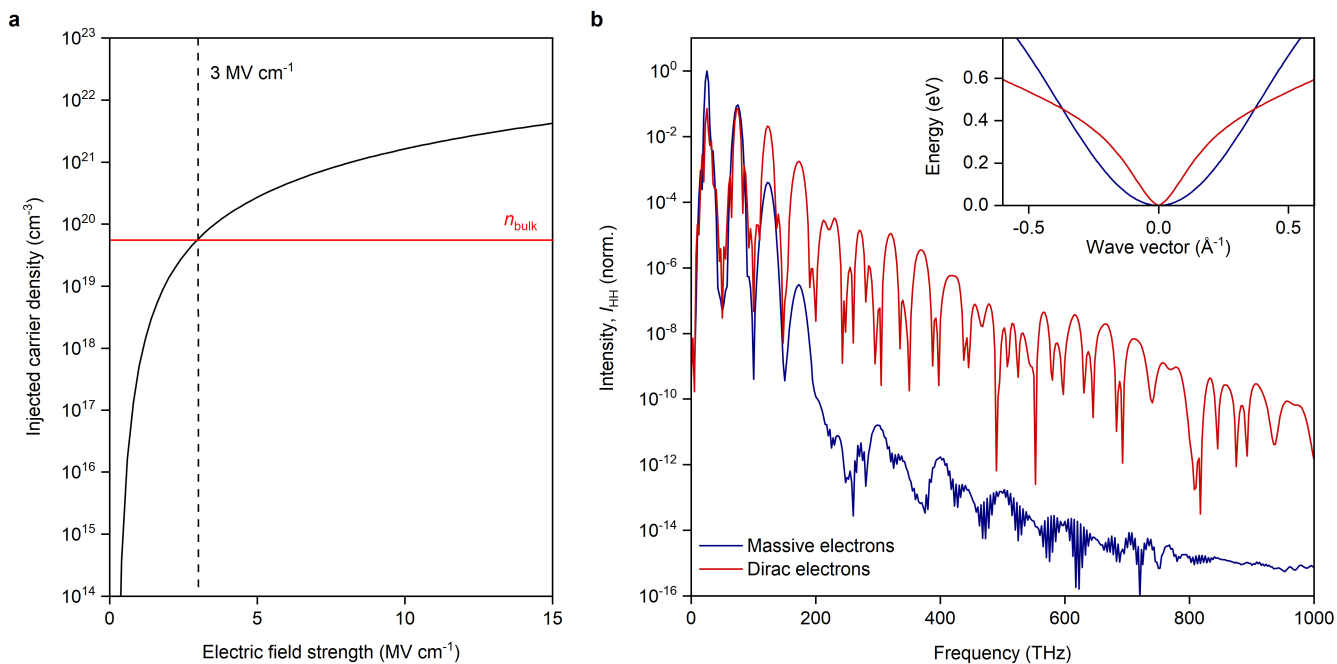
Extended Data Fig. 2 | Carrier-injection into bulk states and comparison of bulk- and surface-state HHG. a, Calculated carrier density injected, within one THz half cycle, into the bulk conduction band by Zener tunnelling, as a function of the peak THz electric field. Red horizontal line: carrier density, n_{bulk}, injected for $E_{THz} = 3 \, MV \, cm^{-1}$. **b**, HH spectra, I_{HH}, calculated by Boltzmann equations for massive electrons in the bulk crystal described by a cosine-shaped band structure (blue curve), and by Boltzmann equation for Dirac electrons in the surface state of a topological insulator. Inset: corresponding band structures for bulk (blue curve) and surface state electrons (red curve).

Article
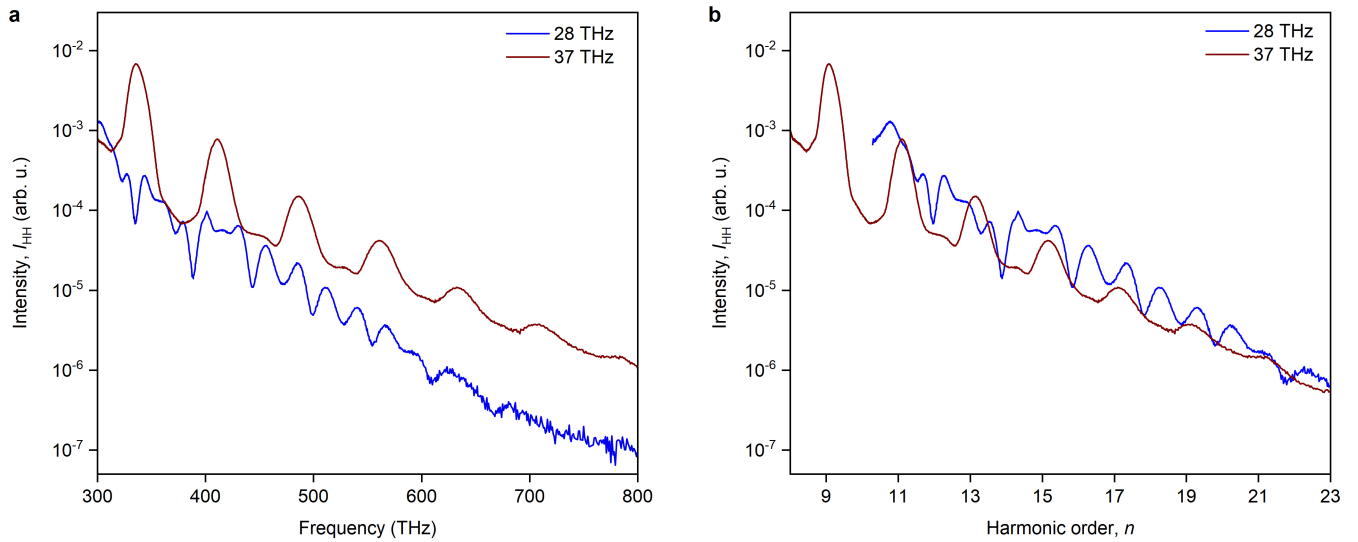
Extended Data Fig. 3 | Comparison of the intensity of bulk and surface HHG. a, HH intensity, I_{HH}, for two select driving frequencies. Above-bandgap excitation at $v_{THz} = 37$ THz (red curve) allows for strong bulk contributions. For $v_{THz} = 28$ THz (blue curve) resonant interband transitions in the bulk are not possible and the peak electric THz field (about 3 MV cm^{-1}) is too low for efficient non-resonant interband excitation. Therefore, the bulk contribution drops by orders of magnitude and the observed spectrum is dominated by HHG from the metallic TSS. In a direct comparison of the two spectra, this contribution is reduced with respect to the above-bandgap bulk HH intensity by only one order of magnitude. **b**, A direct comparison of the same spectra I_{HH} as a function of the harmonic order, n, instead of the harmonic frequency, even reveals a slight enhancement of HHG in the TSS ($v_{THz} = 28$ THz) with respect to the above-bandgap bulk HHG ($v_{THz} = 37$ THz). Considering the low effective thickness of the TSS of about 1 nm compared with the optical penetration depth of about 30 nm to 100 nm over which bulk HHG is collected, this comparison attests to the strong nonlinearity of Dirac electrons.
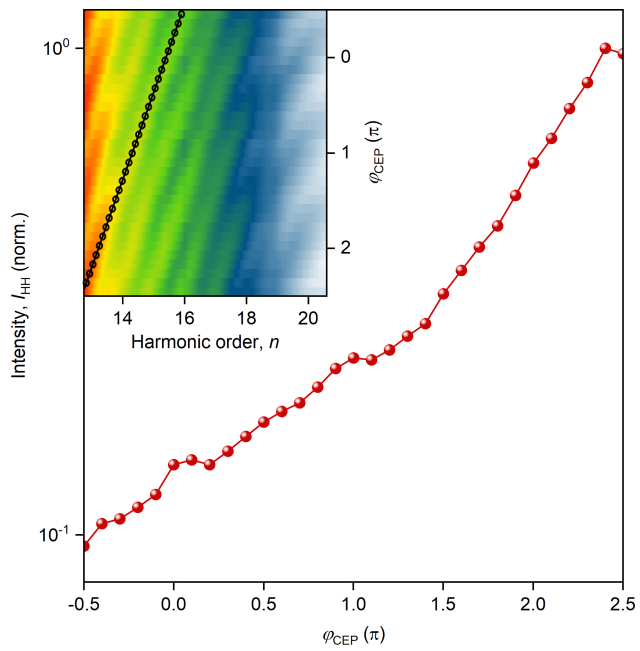
Extended Data Fig. 4 | CEP dependence for HHG in the TSS. HH intensity, I_{HH}, generated in the TSS as a function of the CEP, φ_{CEP}, along the black dotted line in the inset. The intensity of the emitted HHs monotonically increases with increasing φ_{CEP}.
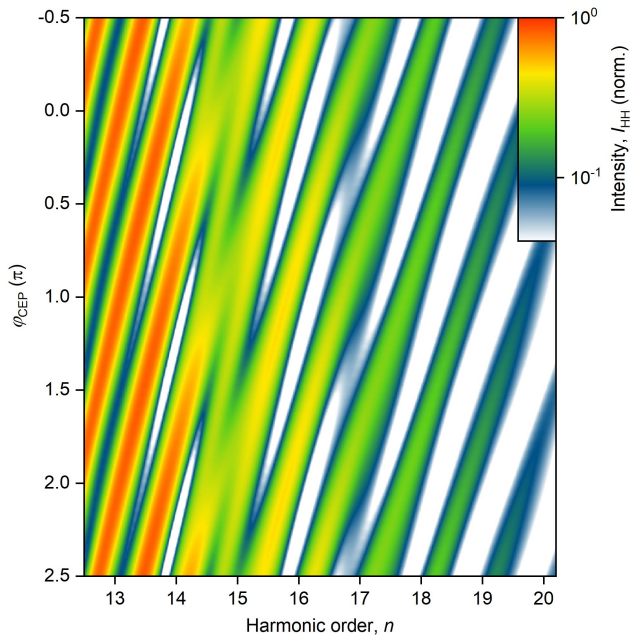
Extended Data Fig. 5 | SBE simulation of CEP-dependent HHG from TSS without interband transitions. Numerical simulation of I_{HH} from the TSS with the SBEs, as in Fig. 2e, but deactivated interband transitions. This calculation is equivalent to the semiclassical solution using the Boltzmann equation, which accounts only for intraband dynamics. The results reproduce both the CEP dependence observed in the experiment of Fig. 2b and the full SBE results of Fig. 2e.

a

b
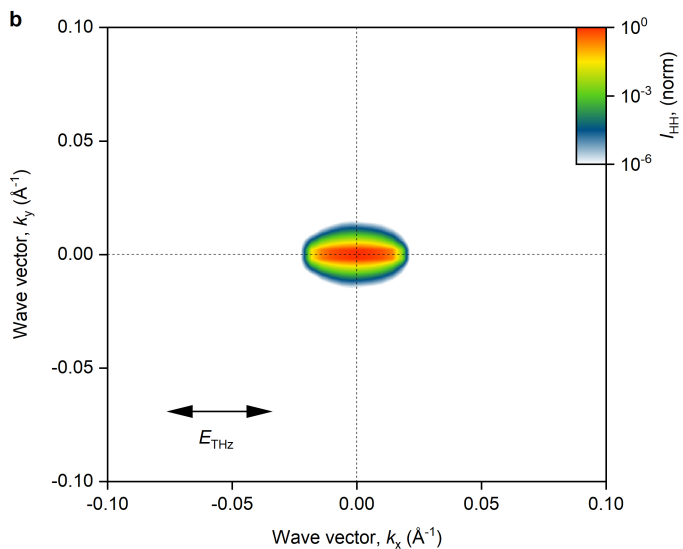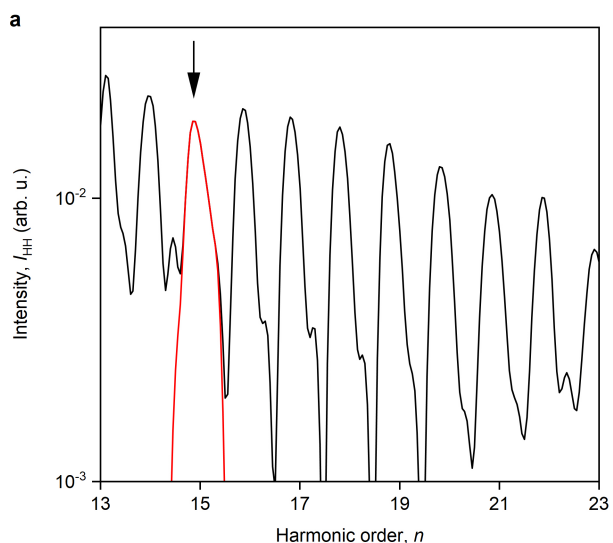
Extended Data Fig. 6 | Momentum-space origin of HHG in the TSS.
a, Calculated HH spectra, I_{HH}, (black curve) for two test charges placed at the wave vector $k_y = \pm 0.001\ \text{Å}^{-1}$ ($k_x = 0$), as obtained from a semiclassical solution of the equations of motion ($v_{THz} = 25\ \text{THz}$, $E_{THz} = 0.1\ \text{MV cm}^{-1}$). **b**, HH intensity (colour scale) of order $n = 15$ (see arrow and red dotted area in **a**) as a function of the starting point (k_x, k_y) of the test charges in momentum space.

a

b
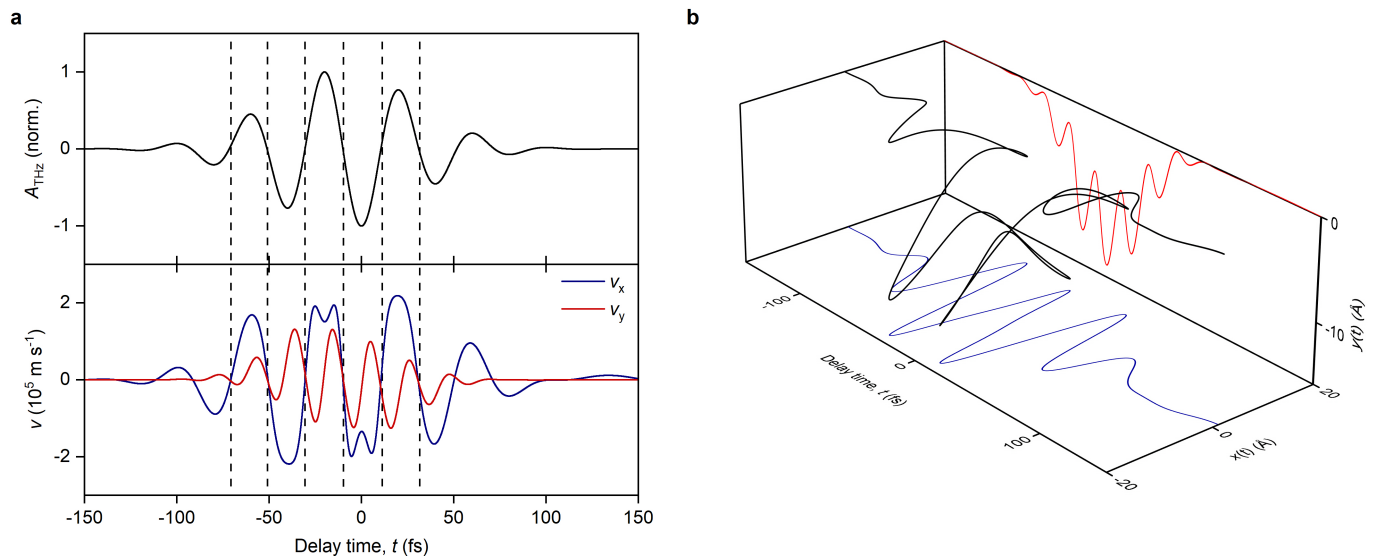
Extended Data Fig. 7 | Quantum mechanical wave-packet motion in the TSS. a, Top: normalized vector potential, A_{THz}, of the driving multi-THz waveform (frequency $v_{THz} = 25$ THz; peak electric field $E_{THz} = 1$ MV cm^{-1}). Dashed lines highlight the zero crossings of the vector potential and the momentum space trajectories. Bottom: group velocity components of the electrons in the TSS parallel (v_x, blue) and perpendicular (v_y, red) to the THz driving field calculated by solving the full time-dependent Schrödinger equation. Both components reverse sign during zero crossings of the momentum space trajectories. **b**, Real space trajectory of lightwave-driven Dirac electrons calculated by the velocities in **a**.
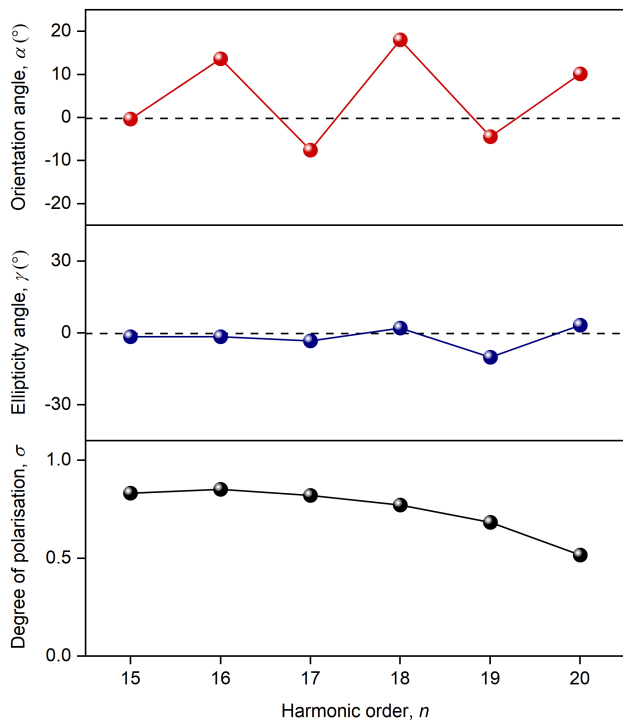
Extended Data Fig. 8 | Extended analysis of the polarimetry measurements.
Extracted orientation angle, α, ellipticity angle, γ, and degree of polarization, σ,
as a function of the harmonic order, n. Although α shows an alternating
behaviour for even and odd orders, the ellipticity remains relatively small for all
orders. The degree of polarization, σ, decreases with increasing order, but still
remains sufficiently high to guarantee a reliable extraction of α and γ.
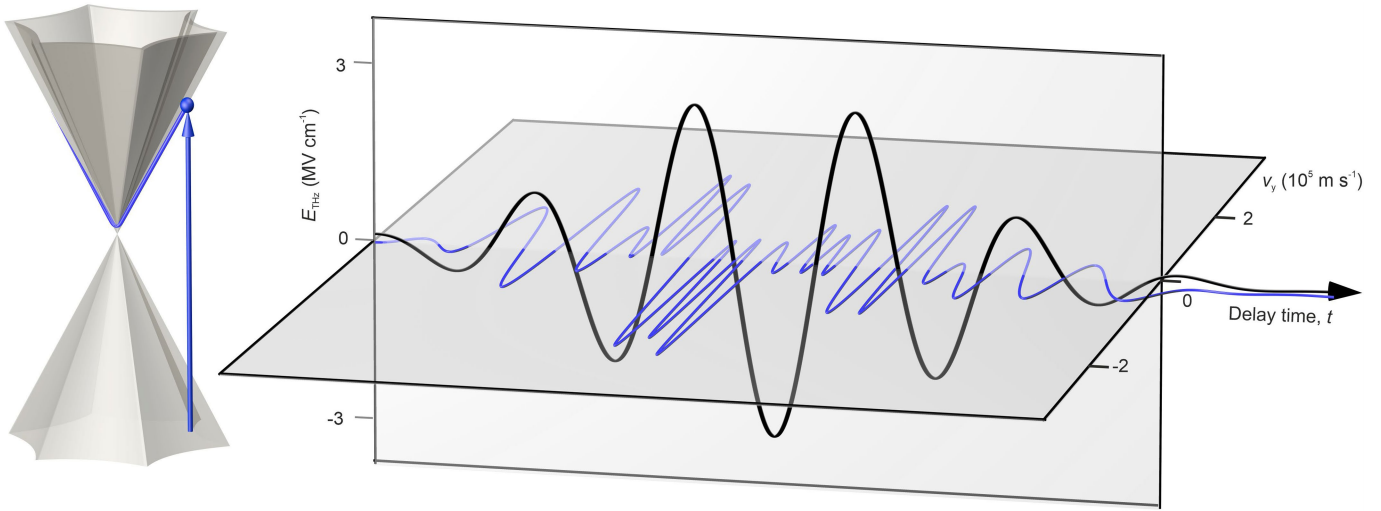
Extended Data Fig. 9 | Zitterbewegung in a topological surface state. Left: three-dimensional scheme of the Dirac-like electron dispersion of the TSS. The blue arrow highlights the quantum interference of different branches of the Dirac system. Right: high-frequency oscillations (blue waveform) indicative of Zitterbewegung depend on the energy separation of the interfering states residing at different energy branches in our quantum mechanical calculations. The black waveform represents the driving THz field E_{THz}.