

SCIENTIFIC REPORTS



OPEN

Data Mining for Three-Dimensional Organic Dirac Materials: Focus on Space Group 19

R. Matthias Geilhufe¹, Stanislav S. Borysov¹, Adrien Bouhon² & Alexander V. Balatsky^{1,3,4}

We combined the group theory and data mining approach within the Organic Materials Database that leads to the prediction of stable Dirac-point nodes within the electronic band structure of three-dimensional organic crystals. We find a particular space group $P2_12_12_1$ (#19) that is conducive to the Dirac nodes formation. We prove that nodes are a consequence of the orthorhombic crystal structure. Within the electronic band structure, two different kinds of nodes can be distinguished: 8-fold degenerate Dirac nodes protected by the crystalline symmetry and 4-fold degenerate Dirac nodes protected by band topology. Mining the Organic Materials Database, we present band structure calculations and symmetry analysis for 6 previously synthesized organic materials. In all these materials, the Dirac nodes are well separated within the energy and located near the Fermi surface, which opens up a possibility for their direct experimental observation.

Recently, we have witnessed growing interest in the research community in the Dirac materials where the low-energy excitations behave as massless Dirac fermions¹. Among the most prominent examples are the two-dimensional material graphene², the surface of bulk topological insulators³ like $\text{Pb}_x\text{Sn}_{1-x}\text{Te}$ ^{4,5}, Dirac-line materials^{6–8} and Weyl semimetals like TaAs⁹. To date, the strong focus within electronic Dirac materials lies in the inorganic crystals. The class of organic crystals remains rather unexplored and only a few organic Dirac materials are known. One prominent example is the quasi two-dimensional charge transfer salt α -(BEDT-TTF)₂I₃ which shows a tilted Dirac cone located at the Fermi energy under high pressure¹⁰. At the same time, organic crystals offer a high potential for technological applications due to low production costs, elastic properties (flexible electronics) and the opportunity to build large-area devices¹¹. We therefore will focus on exploring the space of organic materials with the specific goal of identifying Dirac materials.

Since the crystal structure plays a crucial role for hosting Dirac nodes, attempts of identifying organic Dirac materials so far are mainly based on variations of already known Dirac materials. In two dimensions, this can be done by starting with the graphene structure and replacing the carbon atoms by more complex organic molecules¹². A similar strategy was also discussed with respect to α -(BEDT-TTF)₂I₃^{13,14}. To go beyond this approach, we adopt a new strategy for the search for Dirac nodes in the class of three-dimensional organic crystals. We performed a data mining study on the basis of 5217 electronic Kohn-Sham band structures calculated using Density Functional Theory (DFT)^{15,16} and stored within the Organic Materials Database (OMDB)¹⁷. We looked for *isolated* linear crossings, i.e., where no other bands can be found besides the crossing within the corresponding energy range. Although all considered organic materials were previously synthesized, little attention has been paid to the electronic structure for most of them so far.

To achieve stable Dirac points within the electronic structure, symmetry or topological protection needs to be present. In this connection, crystals with non-symmorphic space groups have been widely discussed^{18–20} where the key role is played by high-dimensional irreducible representations at the Brillouin zone boundary. As a consequence of the Wigner-Eckart theorem, the degeneracy of an electronic state is equal to the dimension of an irreducible representation. In the context of crystals, these degeneracies were discussed in great detail during the 1960s²¹. Recently, with a reinterpretation, degenerate electronic states in crystals again attracted attention as a host to unconventional fermions as low-energy excitation²². There is also a second opportunity of hosting Dirac

¹Nordita, Center for Quantum Materials, KTH Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, SE, 106 91, Stockholm, Sweden. ²Department of Physics and Astronomy, Uppsala University, Box 516, SE, 751 20, Uppsala, Sweden. ³Institute for Materials Science, Los Alamos National Laboratory, Los Alamos, NM, 87545, USA. ⁴ETH Institute for Theoretical Studies, ETH Zürich, 8092, Zürich, Switzerland. Correspondence and requests for materials should be addressed to R.M.G. (email: geilhufe@kth.se)

#	OMDB ID	COD ID	sum formula	Refs
1	3407	1504134	C ₆ H ₇ ClO ₃	25
2	4071	1503207	C ₁₀ H ₁₀ Br ₂ Cl ₃ NO ₂	26
3	3180	1506714	C ₁₂ H ₁₃ NO ₂	27
4	5010	7153350	C ₁₃ H ₁₂ N ₂ O	28
5	3617	1506293	C ₉ H ₁₀ F ₃ NO	29
6	690	4066669	C ₁₀ H ₁₂ BrNO	30

Table 1. Data-mined organic Dirac materials with the space group 19. The first four materials show crossings in the valence band, the latter two have crossings in the conduction band. References to the papers where the synthesis of the materials was reported are given in the last column.

crossing as accidental crossings²³, which are protected by band topology. Such crossings can be found for instance in crystals with the monoclinic space group $P2_1/c$ (#14)^{6,24}. In crystals with this space group, electronic energy bands are sticking together in groups of four bands. As reported in ref. 6, for each of these groups, three topologically different orderings of electronic states can be found at the Γ -point within the Brillouin zone—a trivial phase and two different line-node phases. Within the space group $P2_12_12_1$ (#19), an almost similar situation is present, with the difference that at least one crossing has to occur along one of the paths $\Gamma\bar{X}$, $\Gamma\bar{Y}$ or $\Gamma\bar{Z}$ within the Brillouin zone. This group will be discussed below in Results.

In this paper, we report results of a combined study using abstract group theory and data mining within the Organic Materials Database (OMDB)¹⁷. We point out the first real material examples in the class of three-dimensional organic crystals hosting isolated Dirac nodes in the electronic structure: C₆H₇ClO₃²⁵, C₁₀H₁₀Br₂Cl₃NO₂²⁶, C₁₂H₁₃NO₂²⁷, C₁₃H₁₂N₂O²⁸, C₉H₁₀F₃NO²⁹, and C₁₀H₁₂BrNO³⁰. It will be shown that the found Dirac nodes are a consequence of the orthorhombic crystal structure of the space group $P2_12_12_1$ (#19). Within the band structure of the materials, two different kinds of nodes can be distinguished: 8-fold degenerate Dirac nodes protected by the crystalline symmetry and 4-fold degenerate tilted Dirac nodes protected by the band topology.

Results

Data mining and electronic structure calculations. As reported in ref. 17, most three-dimensional organic crystals are insulating. However, the doping of organic materials is extensively studied, opening the opportunity of shifting the Fermi level into the valence band (p -doping) or conduction band (n -doping)^{11,31}. Even though it is chemically more difficult to achieve in comparison to inorganics, for example, due to purification effects of organics, it was successfully implemented within OLEDs^{32,33}, solar cells^{34,35} or thermoelectric materials³⁶. Therefore, we searched for isolated linear crossings in a neighborhood of 0.5 eV above the lowest unoccupied electronic state and below the highest occupied electronic state within the Kohn-Sham band structures stored in the OMDB by explicitly focusing our attention to materials with the space group $P2_12_12_1$ (#19). For the 6 most promising structures tabulated in Table 1, additional refined DFT calculations were performed (see Methods for more details).

In the following, we concentrate on the discussion of the material C₆H₇ClO₃ (the electronic and molecular structures for the other 5 materials can be found in the supplementary material). Its orthorhombic crystallographic unit cell is build up of four copies of C₆H₇ClO₃ molecules as shown in Fig. 1(a). The calculated band structure along several high symmetry paths within the Brillouin zone (Fig. 1(b)) is plotted in Fig. 1(d). The isolated linear crossings occur at the high-symmetry point R (green dashed circles in Fig. 1(d)) as well as along the high-symmetry path $\Gamma\bar{X}$ (red dotted circles in Fig. 1(d)). A picture of the linear energy dispersion within the \vec{b}_2 - \vec{b}_3 -plane within reciprocal space in the vicinity of the R -point can be seen in Fig. 1(c). The linear crossings are well separated within the energy and as a consequence the electronic density of states grows quadratically with the energy in the vicinity of the three-dimensional Dirac crossing ($n(E) \sim E^{d-1}$, where $d=3$) as can be seen in Fig. 1(d).

However, to justify the claim of the found materials being Dirac materials, a protection of the crossings by crystalline symmetry is necessary. Indeed, we find that the nature of the crossings forming Dirac nodes in the spectra can be explained within the framework of group theory below.

Group theory analysis. The space group $P2_12_12_1$ (#19) itself (hereinafter denoted by \mathcal{G}) is an infinite group having the group of pure translations \mathcal{T} as an infinite, normal and Abelian subgroup. The point group of the lattice \mathcal{G}_0 , i.e., the group of all rotational parts of the space group elements is given by 222 (D_2). The factor group \mathcal{G}/\mathcal{T} is isomorphic to 222 and the coset representatives are given by

$$T_1 = (E, 0, 0, 0), \quad (1)$$

$$T_2 = (C_{2x}, 1/2, 1/2, 0), \quad (2)$$

$$T_3 = (C_{2y}, 0, 1/2, 1/2), \quad (3)$$

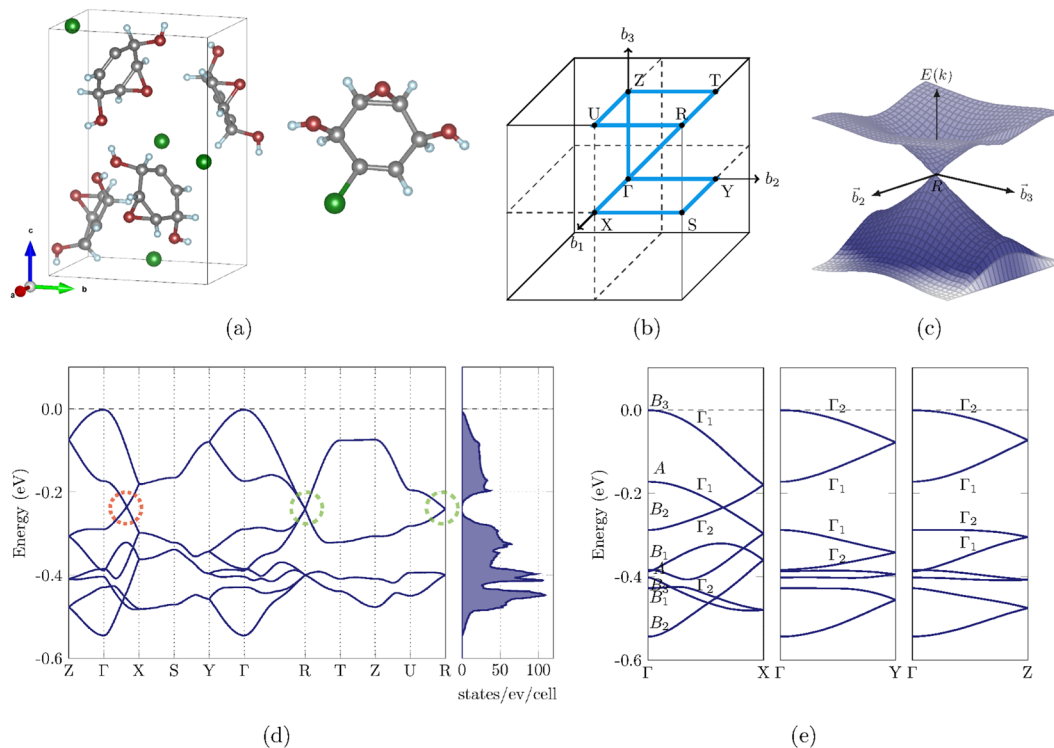


Figure 1. Crystal structure and electronic structure of the orthorhombic organic crystal $C_6H_7ClO_3$. **(a)** Orthorhombic unit cell and single molecule of $C_6H_7ClO_3$. Grey denotes carbon, red oxygen, blue hydrogen and green chlorine. **(b)** Brillouin zone path for band structure calculation. **(c)** Energy dispersion in b_2 - b_3 -plane around the R point. **(d)** Band structure and density of states. Crossings highlighted in green circles are 8-fold degenerate Dirac-points protected by crystalline symmetry and crossings in red circles are 4-fold degenerate Dirac points protected by band topology. **(e)** Energy dispersion along $\Gamma\bar{X}$, $\Gamma\bar{Y}$, $\Gamma\bar{Z}$ and irreducible representations.

	E	C_{2x}	C_{2y}	C_{2z}	$\{E, C_{2x}\}$	$\{E, C_{2y}\}$	$\{E, C_{2z}\}$
A	1	1	1	1	Γ_1	Γ_1	Γ_1
B_1	1	-1	-1	1	Γ_2	Γ_2	Γ_1
B_2	1	-1	1	-1	Γ_2	Γ_1	Γ_2
B_3	1	1	-1	-1	Γ_1	Γ_2	Γ_2

Table 2. Character table and compatibility relations of the point group 222 (D_2).

$$T_4 = (C_{2z}, 1/2, 0, 1/2). \quad (4)$$

Here, E denotes the identity element and C_{2x} , C_{2y} , and C_{2z} denote two-fold rotations (rotations by 180°) about the Cartesian x -, y - and z -axis, respectively. In general, since \mathcal{G} is an infinite group, it has infinitely many irreducible representations. However, due to the special structure of space groups they can be indexed by the combined index (\vec{k}, p) , where \vec{k} denotes a vector in reciprocal space and p denotes an additional index running over all the allowed representations at \vec{k} . A degeneracy of a state with energy $E(\vec{k})$ can be expected when the associated irreducible representation $\Gamma_{\vec{k}}^p$ has dimension $d > 1$ or when pairs of complex conjugate representations are present. Such a degeneracy is denoted as “protected by the crystalline symmetry”. At the R point within the Brillouin zone, the physically-irreducible representation is four dimensional³⁷. Additional spin-degeneracy leads to an 8-fold degenerate linear crossing. Such crossings were recently referred to as double Dirac crossings³⁸. Within the band structure in Fig. 1(d), these crossings are highlighted by green dashed circles.

At the Γ point, the group of the \vec{k} -vector is given by the whole space group. Each of the eigenstates is one-fold degenerate and belongs to one of the irreducible representations listed in Table 2 (spin degeneracy is omitted for the moment). As soon as one moves slightly away from the Γ point, for instance on the path $\Gamma\bar{X}$, the little group of the \vec{k} -vector only contains the elements T_1 and T_2 . Hence, according to Table 3, the bands will be classified by their transformation behavior with respect to the two-fold rotation, namely even (Γ_1) or odd (Γ_2). Every state at

	E	C_2
Γ_1	1	1
Γ_2	1	-1

Table 3. Character table of the point group 2 (C_2).

	T_1	T_2	T_3	T_4
E	2	0	0	0

Table 4. Character table of the space group $P2_12_12_1$ at X , Y and Z .

X transforms as the 2-dimensional irreducible representation E illustrated in Table 4. Moving towards X and coming from $\overline{\Gamma X}$ bands of character Γ_1 and Γ_2 have to merge pairwise. In general, bands can only cross (accidental crossings) when they belong to different irreducible representations²³. Otherwise they will hybridize and form a spectral gap. Clearly, a similar consideration holds along $\overline{\Gamma Y}$ and $\overline{\Gamma Z}$. Taking into account all possible permutations of the four irreducible representations at Γ as well as the possible connections to X , Y and Z , it can be verified that at least one crossing can be found along one of the 3 paths. Taking into account the C_2 rotation symmetry, a second copy of the crossing along the path $\overline{\Gamma(-X)}$ can be also found within the Brillouin zone. Furthermore, as soon as one slightly departs from one of the 3 paths towards the interior of the Brillouin zone, none of the symmetries is kept and only the identity element T_1 is present. Thus, there is no reason to protect the crossing at any point that is not lying on one of the 3 paths and the crossing itself is a Dirac point. Hence, by including spin-degeneracy, this crossing is 4-fold degenerate.

In the present case of $C_6H_7ClO_3$, the topologically protected crossing can be found along the path $\overline{\Gamma X}$, as can be verified from Fig. 1(d,e). In Fig. 1(d), the topologically protected crossing is highlighted by a red dotted circle. As can be seen, the four interesting bands below the Fermi level have the ordering B_3, A, B_2, B_1 . Bands originating from A and B_1 and from B_3 and B_2 are merging pairwise at X for the above discussed reason. As a consequence, a crossing of bands originating from A and B_2 can be observed (see Fig. 1(e)).

Without spin-orbit coupling the system can be split into two identical spin-polarized components. One simple (spin-polarized) point-node is naturally characterized through $\pi_2(Gr_2(\mathbb{C}^4)) \cong \pi_1(U(2)) = \mathbb{Z}$ with the topological charge given by the Chern number, $|C_1| = 1$. By C_2 symmetry, the two simple point-nodes (located here on the $\overline{\Gamma X}$ -axis) must carry the same charge, say $C_1 = -1$. According to the Nielsen-Ninomiya theorem^{39, 40}, the total charge over the whole Brillouin zone must be zero. We then deduce that the double point-node at R must carry a charge $C_1 = +2$. This can be confirmed both numerically and algebraically⁴¹.

Interestingly, $C_6H_7ClO_3$ is close to a Lifschitz transition between two inequivalent global band topologies, controlled by a band inversion at Γ . Here, A and B_1 are almost degenerate having an energy of ≈ -0.39 eV (see Fig. 1(e)). Indeed, the energy ordering of the irreducible representations at Γ , i.e., $\{B_2, B_1, B_3, A, B_1, B_2, A, B_3\}$, guarantees to have two disconnected blocks of four bands (without counting the spin degree of freedom) over the whole Brillouin zone. Alternatively, if we exchange A and B_1 , the global band topology guarantees to have a single block of eight bands being nontrivially connected⁴¹. Supporting this, our data mining has revealed a clear candidate of a nontrivially connected eight-band subspace which is characterized by two pairs of topologically stable simple point-nodes (one pair along $\overline{\Gamma X}$ and one pair along $\overline{\Gamma Y}$) shown in the supplementary material (see Fig. 2(b)). This is the first realistic band structure reported with this new type of global band topology.

Discussion

We presented the first 6 predicted compounds for three-dimensional organic crystals hosting isolated Dirac crossings. All 6 materials were synthesized before, so we encourage direct experimental verification of the results presented here. Within the respective band structures, we identified 8-fold degenerate Dirac nodes at the R point in the Brillouin zone together with 4-fold degenerate topologically protected Dirac points along the high-symmetry path $\overline{\Gamma X}$. The crossings are well separated in energy due to the flat electronic bands of organic crystals and potentially accessible via doping or gating. In comparison to inorganic materials, the spatially sparse unit cells of organic materials characterized by van der Waals bonding between large molecules leads to an electronic structure characterized by blocks of well separated flat bands. This particular property also increases chances of finding topologically protected crossings to be isolated within the energy. We expect that more organic Dirac materials will be reported elsewhere with growth of the OMDb database. Furthermore, the slope of the crossings is usually much smaller than for similar inorganic crystals leading to potential applications as slow Dirac materials⁴².

Methods

Materials data and data mining. We analyzed 5217 Kohn-Sham band structures of the three-dimensional organic crystals stored within the Organic Materials Database (OMDb)¹⁷ (<http://omdb.diracmaterials.org>). The search algorithm for isolated Dirac crossings consisted of two major steps. On the first step, illustrated in Fig. 2, the algorithm selected all materials with either zero or tiny direct energy gaps (less than 1 meV) located up to 0.5 eV above the minimum energy of the lowest conductance band or 0.5 eV below the maximum energy of the highest valence band (for metals, the reference point was the calculated Fermi energy). The purpose of this gap is to introduce numerical tolerance since the band structure calculations were performed along a discrete mesh and the algorithm could miss the crossing point if it occurs between two mesh points. The algorithm also checked

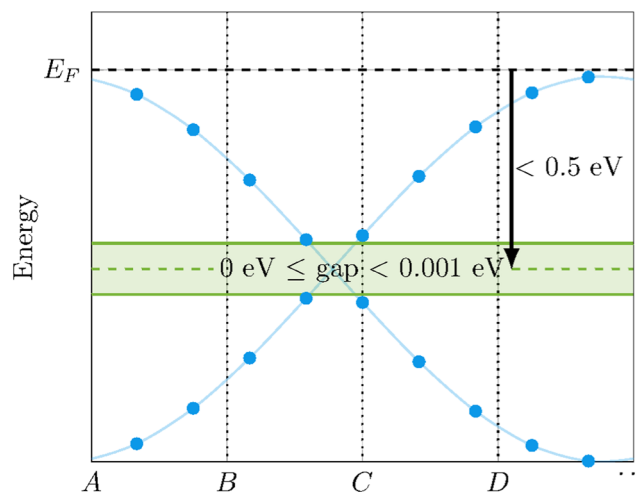


Figure 2. Illustration of the search criteria for isolated Dirac crossings near the Fermi level. The algorithm selected all materials with either zero or tiny direct energy gaps (less than 1 meV) located up to 0.5 eV above the minimum energy of the lowest conductance band or 0.5 eV below the maximum energy of the highest valence band, where no other bands can be found within the corresponding energy range of the gap. The purpose of this gap is to introduce numerical tolerance since the band structure calculations were performed along a discrete mesh.

that no other bands can be found within the corresponding energy range of the gap, which is a necessary (but not sufficient) criterion to find an isolated Dirac crossing. The first step allowed us to significantly reduce the search space as it found only 45 gaps (45 materials or 0.9% of the initial dataset) near the lowest conductance band and 92 gaps (91 materials or 1.8% of the initial dataset) near the highest valence band, making manual inspection of the search results feasible. The statistics suggest that this simple criterion performs well in filtering out irrelevant materials and can be used for the automated search for isolated Dirac crossings.

In the second step, we applied a nearest neighbor search algorithm to arrange the selected materials according to their similarity to a pattern of two crossing straight lines. With this aim, we considered the two nearest bands within a momentum window around a direct gap detected in the previous step. We empirically set the size of the momentum window to be 0.4 units wide as an expected characteristic scale of the Dirac crossing. Finally, we used the average Euclidean distance (the root mean square error) between the pattern and these two bands (within the momentum window) linearly scaled to the same bounding box to arrange the selected materials. The second step becomes important to prioritize search results when the number of materials is large.

Electronic structure calculations. Having found a subset of perspective materials, we performed refined electronic structure calculations in the framework of the density functional theory¹⁵ by applying a pseudopotential projector augmented-wave method^{43–45}, as implemented in the Vienna Ab initio Simulation Package (VASP)^{46–48} and the Quantum ESPRESSO code⁴⁹. The exchange–correlation functional was approximated by the generalized gradient approximation according to Perdew, Burke and Ernzerhof⁵⁰. The structural information were taken from the Crystallography Open Database (COD)^{51–53} and transformed into DFT input files by applying the Pymatgen package⁵⁴.

Within VASP, the precision flag was set to “normal” meaning that the energy cut-off is given by the maximum of the specified maxima for the cut-off energies within the POTCAR files (for example, for carbon this value is given by 400 eV). The calculations were performed spin-polarized but without spin-orbit coupling. For the integration in \vec{k} -space, a $6 \times 6 \times 6$ Γ -centred mesh according to Monkhorst and Pack⁵⁵ was chosen during the self-consistent cycle. A structural optimization was performed by allowing the ionic positions, the cell shape and the cell volume to change (ISIF = 3). Optimized structures from the VASP calculations were used. Quantum ESPRESSO was applied to estimate the associated irreducible representations of the energy levels within the band structure. The cut-off energy for the wave function was chosen to be 48 Ry and the cut-off energy for the charge density and the potentials was chosen to be 316 Ry. The calculated band structures using VASP and Quantum ESPRESSO are in perfect agreement.

References

1. Wehling, T., Black-Schaffer, A. M. & Balatsky, A. V. Dirac materials. *Adv. Phys.* **63**, 1–76 (2014).
2. Abergel, D., Apalkov, V., Berashevich, J., Ziegler, K. & Chakraborty, T. Properties of graphene: a theoretical perspective. *Adv. Phys.* **59**, 261–482 (2010).
3. Fu, L., Kane, C. L. & Mele, E. J. Topological insulators in three dimensions. *Phys. Rev. Lett.* **98**, 106803 (2007).
4. Tanaka, Y. *et al.* Experimental realization of a topological crystalline insulator in SnTe. *Nat. Phys.* **8**, 800–803 (2012).
5. Geilhufe, M. *et al.* Effect of hydrostatic pressure and uniaxial strain on the electronic structure of $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$. *Phys. Rev. B* **92**, 235203 (2015).
6. Geilhufe, R. M., Bouhon, A., Borysov, S. S. & Balatsky, A. V. Three-dimensional organic dirac-line materials due to nonsymmorphic symmetry: A data mining approach. *Phys. Rev. B* **95**, 041103 (2017).

7. Kim, Y., Wieder, B. J., Kane, C. L. & Rappe, A. M. Dirac line nodes in inversion-symmetric crystals. *Phys. Rev. Lett.* **115**, 036806, doi:10.1103/PhysRevLett.115.036806 (2015).
8. Yamakage, A., Yamakawa, Y., Tanaka, Y. & Okamoto, Y. Line-node dirac semimetal and topological insulating phase in noncentrosymmetric pnictides CaAgX (X = P, As). *J. Phys. Soc. Jpn.* **85**, 013708, doi:10.7566/JPSJ.85.013708 (2016).
9. Lv, B. *et al.* Experimental discovery of weyl semimetal TaAs. *Phys. Rev. X* **5**, 031013 (2015).
10. Katayama, S., Kobayashi, A. & Suzumura, Y. Pressure-induced zero-gap semiconducting state in organic conductor α -(BEDT-TTF)₃I₃ salt. *J. Phys. Soc. Jpn.* **75**, 054705, doi:10.1143/JPSJ.75.054705 (2006).
11. Lüssem, B., Riede, M. & Leo, K. Doping of organic semiconductors. *physica status solidi (a)* **210**, 9–43 (2013).
12. Wang, Z., Su, N. & Liu, F. Prediction of a two-dimensional organic topological insulator. *Nano letters* **13**, 2842–2845 (2013).
13. Choji, T., Kobayashi, A. & Suzumura, Y. Zero-gap state in organic conductor α -(BEDT-TTF)₂NH₄Hg(SCN)₄. *J. Phys. Soc. Jpn.* **80**, 074712 (2011).
14. Morinari, T. & Suzumura, Y. On the possible zero-gap state in organic conductor α -(BEDT-TSF)₃I₃ under pressure. *J. Phys. Soc. Jpn.* **83**, 094701 (2014).
15. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871, doi:10.1103/PhysRev.136.B864 (1964).
16. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138, doi:10.1103/PhysRev.140.A1133 (1965).
17. Borysov, S. S., Geilhufe, R. M. & Balatsky, A. V. Organic materials database: An open-access online database for data mining. *PLoS one* **12**, e0171501 (2017).
18. Young, S. M. & Kane, C. L. Dirac semimetals in two dimensions. *Phys. Rev. Lett.* **115**, 126803, doi:10.1103/PhysRevLett.115.126803 (2015).
19. Schoop, L. M. *et al.* Dirac cone protected by non-symmorphic symmetry and three-dimensional dirac line node in ZrSiS. *Nat. communications* **7** (2016).
20. Yang, B.-J., Bojesen, T. A., Morimoto, T. & Furusaki, A. Topological semimetals protected by off-centered symmetries in nonsymmorphic crystals. *Phys. Rev. B* **95**, 075135 (2017).
21. Zak, J. Method to obtain the character tables of nonsymmorphic space groups. *J. Math. Phys.* **1**, 165–171 (1960).
22. Bradlyn, B. *et al.* Beyond dirac and weyl fermions: Unconventional quasiparticles in conventional crystals. *Sci.* **353**, aaf5037 (2016).
23. Herring, C. Accidental degeneracy in the energy bands of crystals. *Phys. Rev.* **52**, 365 (1937).
24. Wieder, B. J. & Kane, C. Spin-orbit semimetals in the layer groups. *Phys. Rev. B* **94**, 155108 (2016).
25. Pinkerton, D. M., Banwell, M. G. & Willis, A. C. Chemoenzymatic access to versatile epoxyquinol synthons. *Org. Lett.* **11**, 4290–4293, doi:10.1021/ol9016657 (2009).
26. Bae, J.-Y., Lee, H.-J., Youn, S.-H., Kwon, S.-H. & Cho, C.-W. Organocatalytic asymmetric synthesis of chiral pyrrolizines by cascade conjugate addition-aldol reactions. *Org. Letters* **12**, 4352–4355 (2010).
27. Dong, L. *et al.* Asymmetric nitroallylation of arylboronic acids with nitroallyl acetates catalyzed by chiral rhodium complexes and its application in a concise total synthesis of optically pure (+)- γ -lycorane. *Org. Letters* **7**, 4285–4288 (2005).
28. Butin, A. V. *et al.* Furan ring opening-pyrrole ring closure: a new synthetic route to aryl (heteroaryl)-annulated pyrrolo [1, 2-a] [1, 4] diazepines. *Org. & biomolecular chemistry* **8**, 3316–3327 (2010).
29. Šterk, D., Stephan, M. & Mohar, B. Highly enantioselective transfer hydrogenation of fluoroalkyl ketones. *Org. Letters* **8**, 5935–5938 (2006).
30. Eloi, A. *et al.* Cationic planar chiral (η^6 -arene) Mn (CO)₃⁺ complexes: resolution, nmr study in chiral-oriented solvents, and applications to the enantioselective synthesis of 4-substituted cyclohexenones and (η^6 -phosphinoarene) Mn (CO)₃⁺ complexes. *Organometallics* **29**, 3876–3886 (2010).
31. Pfeiffer, M. *et al.* Doped organic semiconductors: Physics and application in light emitting diodes. *Org. Electron.* **4**, 89–103, doi:10.1016/j.orgel.2003.08.004, High Efficiency Light Emitters (2003).
32. Zhou, X. *et al.* Very-low-operating-voltage organic light-emitting diodes using a p-doped amorphous hole injection layer. *Appl. Phys. Lett.* **78**, 410–412, doi:10.1063/1.1343849 (2001).
33. Blochwitz, J., Pfeiffer, M., Fritz, T. & Leo, K. Low voltage organic light emitting diodes featuring doped phthalocyanine as hole transport material. *Appl. Phys. Lett.* **73**, 729–731, doi:10.1063/1.121982 (1998).
34. Hoppe, H. & Sariciftci, N. S. Organic solar cells: An overview. *J. Mater. Res.* **19**, 1924–1945 (2004).
35. Drechsel, J. *et al.* Efficient organic solar cells based on a double p-i-n architecture using doped wide-gap transport layers. *Appl. Phys. Lett.* **86**, 244102 (2005).
36. Kim, G.-H., Shao, L., Zhang, K. & Pipe, K. P. Engineered doping of organic semiconductors for enhanced thermoelectric efficiency. *Nat. Mater.* **12**, 719–723 (2013).
37. Aroyo, M. I., Kirov, A., Capillas, C., Perez-Mato, J. & Wondratschek, H. Bilbao crystallographic server. II. representations of crystallographic point groups and space groups. *Acta Crystallogr. Sect. A: Foundations Crystallogr.* **62**, 115–128 (2006).
38. Wieder, B. J., Kim, Y., Rappe, A. & Kane, C. Double Dirac semimetals in three dimensions. *Phys. Rev. Lett.* **116**, 186402 (2016).
39. Nielsen, H. B. & Ninomiya, M. Absence of neutrinos on a lattice: (I). proof by homotopy theory. *Nucl. Phys. B* **185**, 20–40 (1981).
40. Nielsen, H. B. & Ninomiya, M. Absence of neutrinos on a lattice: (II). intuitive topological proof. *Nucl. Phys. B* **193**, 173–194 (1981).
41. Bouhon, A. & Black-Schaffer, A. Global band topology of simple and double Dirac-point semimetals. *Phys. Rev. B* **95**, 241101 (2017).
42. Triola, C., Zhu, J.-X., Migliori, A. & Balatsky, A. V. Many-body instabilities and mass generation in slow Dirac materials. *Phys. Rev. B* **92**, 045401 (2015).
43. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953 (1994).
44. Vanderbilt, D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **41**, 7892 (1990).
45. Kresse, G. & Hafner, J. Norm-conserving and ultrasoft pseudopotentials for first-row and transition elements. *J. Physics: Condens. Matter* **6**, 8245 (1994).
46. Kresse, G. & Hafner, J. *Ab initio* molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561, doi:10.1103/PhysRevB.47.558 (1993).
47. Kresse, G. & Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
48. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758 (1999).
49. Giannozzi, P. *et al.* QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Physics: Condens. Matter* **21**, 395502 (2009).
50. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. review letters* **77**, 3865 (1996).
51. Merkys, A. *et al.* Cod::CIF::Parser: an error-correcting CIF parser for the perl language. *J. applied crystallography* **49** (2016).
52. Gražulis, S., Merkys, A., Vaitkus, A. & Okulič-Kazarinas, M. Computing stoichiometric molecular composition from crystal structures. *J. applied crystallography* **48**, 85–91 (2015).
53. Gražulis, S. *et al.* Crystallography open database (COD): an open-access collection of crystal structures and platform for world-wide collaboration. *Nucleic acids research* **40**, D420–D427 (2012).
54. Ong, S. P. *et al.* Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319, doi:10.1016/j.commatsci.2012.10.028 (2013).
55. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **13**, 5188 (1976).

Acknowledgements

The work at Los Alamos is supported by the US Department of Energy, BES E3B7. Furthermore, the work was supported by the Swedish Research Council Grant No. 638-2013-9243, the Knut and Alice Wallenberg Foundation, and the European Research Council under the European Union's Seventh Framework Program (FP/2207-2013)/ERC Grant Agreement No. DM-321031, Dr. Max Rössler, the Walter Haefner Foundation and the ETH Zürich Foundation. The authors acknowledge computational resources from the Max Planck Institute of Microstructure Physics in Halle (Germany) and the Swedish National Infrastructure for Computing (SNIC) at the National Supercomputer Centre at Linköping University.

Author Contributions

A.V.B. designed the study. R.M.G. performed the DFT calculations as well as the group theoretical analysis. S.S.B. developed and applied the data mining algorithms. A.B. worked on the group theoretical analysis and topological aspects. All authors worked on the text and reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-07374-7](https://doi.org/10.1038/s41598-017-07374-7)

Competing Interests: The authors declare that they have no competing interests.

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



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) 2017