ARTICLE   OPEN

Check for updates

Determining dimensionalities and multiplicities of crystal nets

Hao Gao [1], Junjie Wang[1], Zhaopeng Guo[1] and Jian Sun [1]✉

Low-dimensional materials have attracted significant attention over the past decade. To discover new low-dimensional materials, high-throughput screening methods for structures with target dimensionality have been applied in different materials databases. For this purpose, the reliability of dimensionality identification is therefore highly important. In this work, we find that the existence of self-penetrating nets may lead to incorrect results by previous methods. Instead of this, we use the quotient graph to analyse the topologies of structures and compute their dimensionalities. Based on the quotient graph, we can calculate not only the dimensionality but also the multiplicity of self-penetrating structures. As a demonstration, we screened the Crystallography Open Database using the method and find hundreds of structures with different dimensionalities and high multiplicities up to 11. Some of the self-penetrating materials may have application values in gas storage, selective catalysis or photocatalysis because of their high gas sorption capacities and various electronic structures.

INTRODUCTION

The research of low-dimensional materials is an active field in recent years because of their novel properties and wide applications. For instance, some quantum phenomena, such as fractional quantum hall effects[1] and Luttinger liquids[2], can be realized in the low-dimensional systems. Due to their novel electronic properties induced by the geometric limit, low-dimensional materials have also been widely applied in various research areas including batteries, catalysis, electronics, and photonics[3,4], etc. One way to design new low-dimensional materials is the top-down approach, in which the low-dimensional materials are exfoliated from known bulk phases. Many recent investigations have focused on searching in large databases for compounds containing low-dimensional components and using high-throughput computational methods to discover new low-dimensional materials with appealing properties[5–16]. For example, previous work has identified thousands of layered structures from more than 100,000 compounds[13], and about 2000 of them are exfoliable.

Therefore, a correct and efficient identification method of structure dimensionality is highly desired for high-throughput mining of low-dimensional materials. Topology-scaling algorithm (TSA)[9] and rank determination algorithm (RDA)[13] are widely used to determine the dimensionality of a crystal structure. In TSA, an original cluster in the unit cell contains N_1 atoms, then the expanded cluster in an $n \times n \times n$ supercell has N_2 atoms. The component dimensionality is determined by the scaling factor N_2/N_1. The factor is expected to be n^d ($d = 0, 1, 2, 3$), where d is the dimensionality. Different from TSA, RDA computes the rank of the set composed by connected and equivalent atoms in the component to determine the dimensionality. Both methods do not require prior information like the stacking direction for the layers and can deal with complex situations where components with different dimensionalities coexist. However, they require building supercells to connect periodic images. If the supercell is not large enough, TSA and RDA will underestimate the dimensionalities of self-penetrating structures[15]. To solve the problem, Larsen et al.[15] have proposed a modified RDA method.

Actually, early works[17,18] based on quotient graph (QG)[19] have already proposed correct algorithms for dimensionality, which have been implemented in ToposPro[20] and Systre[21]. Interpenetration analysis is also available through this approach[18,22]. QG is a powerful method and has been applied in many different fields, such as crystal structure prediction[23–26], structure decomposition[27] and machine-learning models for materials property prediction[28–31], etc. Although the algorithms based on QG for dimensionality and self-penetrating multiplicity have been implemented before, however, the detailed discussions about these methods are still insufficient. This hinders the understanding and applying the QG method for computational material scientists.

Here, we revisited QG as a powerful tool to discuss dimensionalities and multiplicities of crystals including self-penetrating structures. We demonstrated and discussed a systematical approach based on the quotient graph to compute correct dimensionality. We compared the QG method with previous dimensionality determination algorithms. The QG method can deal with the multiplicities of self-penetrating structures correctly as well. Moreover, we mined structures with high multiplicities up to 11 in the Crystallography Open Database (COD)[32] to show the reliability of the method.

RESULTS

Definition of QG and dimensionality

For screening large materials databases, we used interatomic distances to identify the bonds. A bond between atom i and j exists if

$$d_{ij} < k(r_i^{cov} + r_j^{cov}), \qquad (1)$$

where d_{ij} is the interatomic distance, r_i^{cov} and r_j^{cov} are the atomic covalent radii, and k is the bond-length tolerance parameter.

For a crystal structure, the atoms and bonds can be viewed as nodes and edges, and they compose an infinite, undirected graph, called a net. Because of the translation symmetry in the crystal, a net can be described by a finite QG[17,19]. A QG is a labeled and directed graph containing N_{at} nodes, where N_{at} is the number of

[1]National Laboratory of Solid State Microstructures, School of Physics and Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, China. ✉email: jiansun@nju.edu.cn

atoms in the unit cell. To distinguish the translationally equivalent atoms, we use a notation $n_i(\mathbf{v})$ ($1 \leq i \leq N_{at}$) to represent the ith atom with a Cartesian position $(\mathbf{x}_i + \mathbf{v})\mathbf{h}$, where $\mathbf{x}_i$ is the fractional coordination of atom i, $\mathbf{h}$ is the cell matrix and $\mathbf{v}$ is an integer vector representing the coordinate of cell. If a bond exists between $n_i(\mathbf{v}')$ and $n_j(\mathbf{v}'')$, the corresponding QG has an edge $n_i \xrightarrow{\mathbf{v}=\mathbf{v}''-\mathbf{v}'} n_j$ labeled by $\mathbf{v}$. The edge denotes equivalent bonds between $n_i(\mathbf{v}_0)$ and $n_j(\mathbf{v}_0 + \mathbf{v})$ with an arbitrary integer vector $\mathbf{v}_0$. Obviously, the edge $n_i \xrightarrow{\mathbf{v}} n_j$ is equivalent to $n_j \xrightarrow{-\mathbf{v}} n_i$ with the opposite direction. Taking graphene as an example, we showed how to build a QG from a net in Fig. 1.

A connected component $\mathbf{X}$ in a crystal might contain multiple equivalent atoms $\{n_i(\mathbf{0}), n_i(\mathbf{v}_1), n_i(\mathbf{v}_2), n_i(\mathbf{v}_3), \ldots\}$ and its
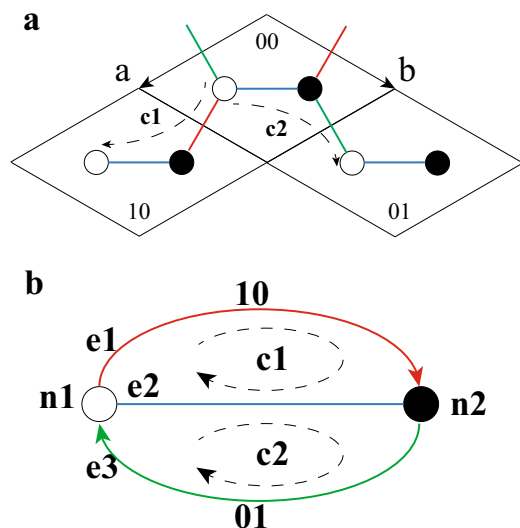
dimensionality is defined by the dimension of the subspace spanned by these connected and translationally equivalent atoms[13]. Let $\mathbf{V}$ be a matrix whose rows are $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots$, then the dimensionality equals to the rank of $\mathbf{V}$:

$$\dim(\mathbf{X}) = \dim(\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots\}) = \text{rank}(\mathbf{V}). \tag{2}$$

For an extensive (1D, 2D, 3D) component, the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots\}$ is infinite, so we cannot use Eq. (2) to compute the dimensionality directly. Before identifying the dimensionality, we should find the basis set. In this work, a systematical approach based on cycles of QG is used.

In a crystal net, if two equivalent atoms $n_{i_1}(\mathbf{u})$ and $n_{i_1}(\mathbf{u}')$ are connected, there must be a path between them. In the relative QG, the path connecting equivalent atoms is represented by a closed chain. Suppose a closed chain c is composed by a sequence of connected nodes:$(n_{i_1}, n_{i_2}, n_{i_3}, \ldots, n_{i_k}, n_{i_{k+1}}, \ldots, n_{i_M}, n_{i_1})$, in which the first and last nodes are same. Because of the equivalent relation: $n_i \xrightarrow{\mathbf{v}} n_j \equiv n_j \xrightarrow{-\mathbf{v}} n_i$, we can always adjust the direction of edges in c so that all the edges have the same direction. And the closed chain can be written as:
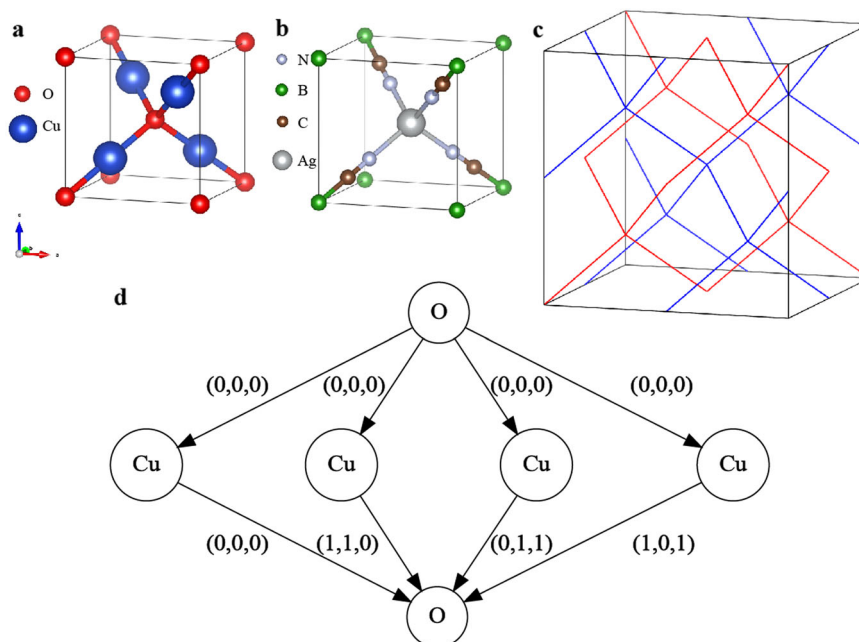
$$c = n_{i_1} \xrightarrow{\mathbf{v}_1} n_{i_2} \xrightarrow{\mathbf{v}_2} n_{i_3} \ldots n_{i_k} \xrightarrow{\mathbf{v}_k} n_{i_{k+1}} \ldots n_{i_M} \xrightarrow{\mathbf{v}_M} n_{i_1}. \tag{3}$$

The cycle sum[17] of the closed chain is defined as $\mathbf{s}(c) = \sum_k \mathbf{v}_k$, where k runs over all the edges in c. In the closed chain c, the kth edge connects atoms $n_{i_k}(\mathbf{u}_k)$ and $n_{i_{k+1}}(\mathbf{u}_{k+1})$, and the edge vector is $\mathbf{v}_k = \mathbf{u}_{k+1} - \mathbf{u}_k$. Then we obtain:

$$\mathbf{s}(c) = \sum_k \mathbf{u}_{k+1} - \mathbf{u}_k = (\mathbf{u}_2 - \mathbf{u}) + (\mathbf{u}_3 - \mathbf{u}_2) + \cdots + (\mathbf{u}_{k+1} - \mathbf{u}_k)$$
$$+ \cdots + (\mathbf{u}' - \mathbf{u}_M) = \mathbf{u}' - \mathbf{u}. \tag{4}$$

Therefore, the cycle sum $\mathbf{s}(c)$ equals to the cell offset between equivalent atoms $n_{i_1}(\mathbf{u})$ and $n_{i_1}(\mathbf{u}')$. The set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots\}$ in Eq. (2) is exactly the set of cycle sums of all the cycles in QG.

In a finite graph, all the closed chains compose a vector space called cycle space, and the generating subspace is composed by finite basic cycles[17]. Because the cycle sum function is linear: $\mathbf{s}(c_1 + c_2) = \mathbf{s}(c_1) + \mathbf{s}(c_2)$, all the cycle sums $\{\mathbf{s}(c)\}$ also compose a vector space. To compute the dimensionality of component $\mathbf{X}$, we


Fig. 1 Graphene net and its quotient graph. a Graphene net in 2D space; **b** The quotient graph of graphene net. n_1, n_2 are nodes. e_1, e_2, e_3 are edges. c_1, c_2 are the two basic cycles of graphene net.

Fig. 2 Self-penetrating nets and the QG. The crystal structures of Cu_2O (**a**) and $Ag(B(CN)_4)$ (**b**). **c** The net with multiplicity of 2 shown in a $2 \times 2 \times 2$ supercell. **d** The QG of Cu_2O. The disconnected networks are colored in red and blue, respectively.

only need to consider the cycles sums of finite basic cycles:

$$\dim(\mathbf{X}) = \dim(\{\mathbf{s}(c)|c \in \mathbf{F}(\mathbf{G})\}) = \text{rank}(\mathbf{S}), \qquad (5)$$

where $\mathbf{F}(\mathbf{G})$ is the set of basic cycles of the QG, $\mathbf{G}$ and $\mathbf{S}$ is a matrix whose rows are the cycle sums of the basic cycles.

We showed how to compute the dimensionality of graphene in Fig. 1 as an example. In the QG of graphene, as shown in Fig. 1, there are two nodes(n_1, n_2), three edges(e_1, e_2, e_3), and two basic
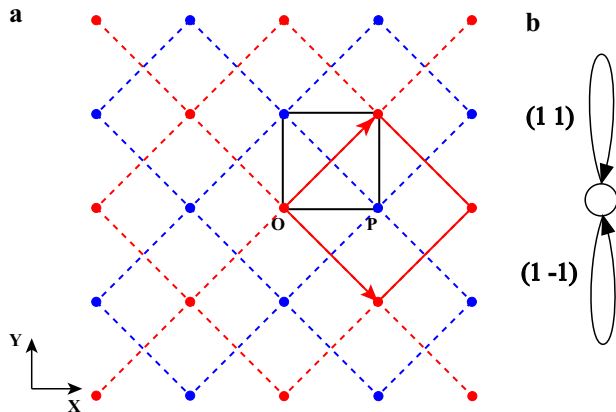
Fig. 3 A plane self-penetrating net with multiplicity of 2 and its QG. a A twofold self-penetrating net. The black box is the unit cell of the square lattice. The red and blue dash lines and points represent the disconnected subnets. The red solid box is the basic building block of the red subnet. **b** The relative quotient graph.

cycles(c_1, c_2):

$$
\begin{aligned}
e_1 &: n_1 \xrightarrow{(1,0)} n_2 \equiv n_2 \xrightarrow{(-1,0)} n_1 \\
e_2 &: n_1 \xrightarrow{(0,0)} n_2 \equiv n_2 \xrightarrow{(0,0)} n_1 \\
e_3 &: n_1 \xrightarrow{(0,1)} n_2 \equiv n_2 \xrightarrow{(0,-1)} n_1 \\
c_1 &= e_1 + e_2 = n_1 \xrightarrow{(1,0)} n_2 \xrightarrow{(0,0)} n_1 \\
c_2 &= e_2 + e_3 = n_1 \xrightarrow{(0,0)} n_2 \xrightarrow{(0,1)} n_1 \\
\mathbf{s}(c_1) &= (1,0) + (0,0) = (1,0) \\
\mathbf{s}(c_2) &= (0,0) + (0,1) = (0,1).
\end{aligned}
\qquad (6)
$$

We also showed the paths relative to the basic cycles in real space (Fig. 1a). The c_1 path connects equivalent atoms $n_1(0, 0)$ and $n_1(1, 0)$. The offset $(1, 0)$ equals to the cycle sum $\mathbf{s}(c_1)$. All the connected equivalent atoms can be represented by the basic cycles. For example, a cycle $c_3 = e_1 + e_3 = n_1 \xrightarrow{(1,0)} n_2 \xrightarrow{(0,1)} n_1$ is related to the path connecting atoms $n_1(0, 0)$ and $n_1(1, 1)$. It can be represented by a combination of the basic cycles: $c_3 = c_1 + c_2 = n_1 \xrightarrow{(1,0)} n_2 \xrightarrow{(0,0)} n_1 \xrightarrow{(0,0)} n_2 \xrightarrow{(0,1)} n_1$ and the cycle sum $\mathbf{s}(c_3) = \mathbf{s}(c_1) + \mathbf{s}(c_2) = (1, 1)$ is also a linear combination of basic cycle sums. The basic cycle sums matrix is $\mathbf{S} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and the dimensionality of graphene net equals to the rank of the cycle sum matrix: rank $(\mathbf{S}) = 2$.

Comparison with previous methods
RDA is based on the correct definition of component dimensionality in Eq. (2). However, the original version of RDA[13] only
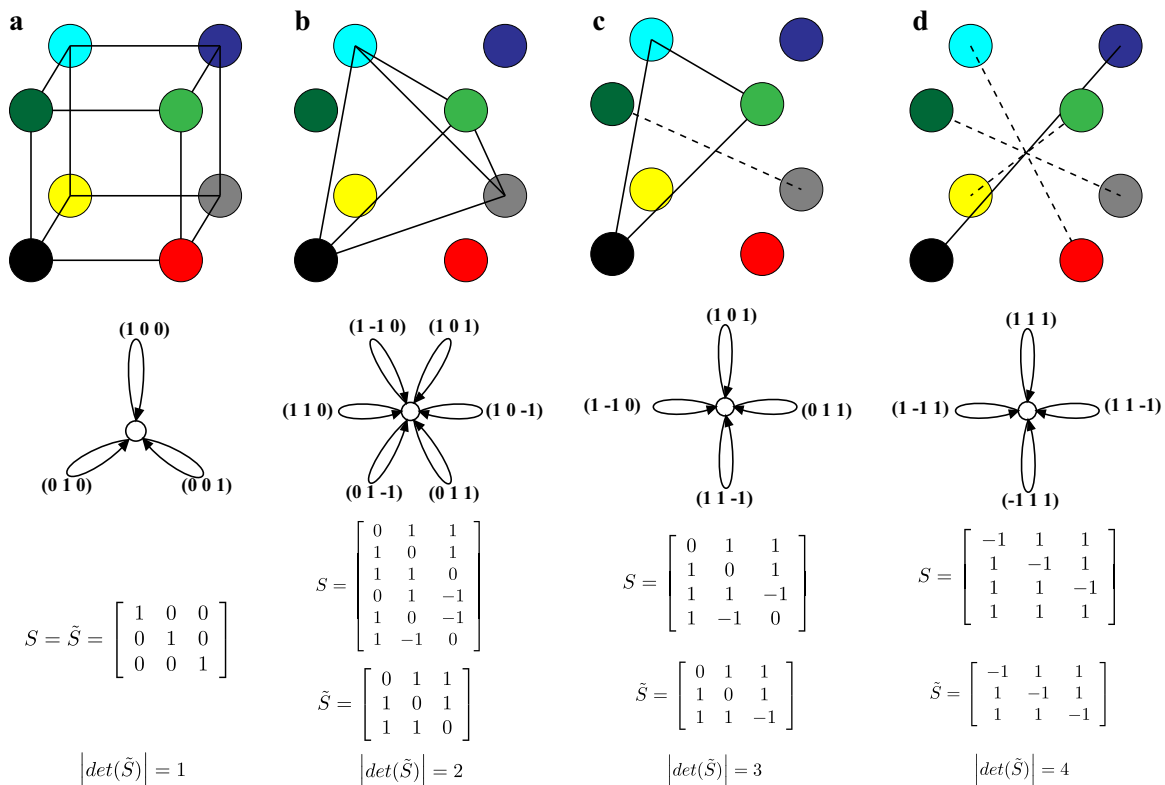
Fig. 4 Schematics, quotient graphs, cycle sums and their determinants of 3D nets. a Onefold net. **b** Twofold net. **c** threefold net. **d** Fourfold net. The black circle in schematics represents a cluster in the original cell. Other colored circles represent the seven images of the cluster for a $2 \times 2 \times 2$ supercell. The dash lines represent edges disconnected from solid lines.
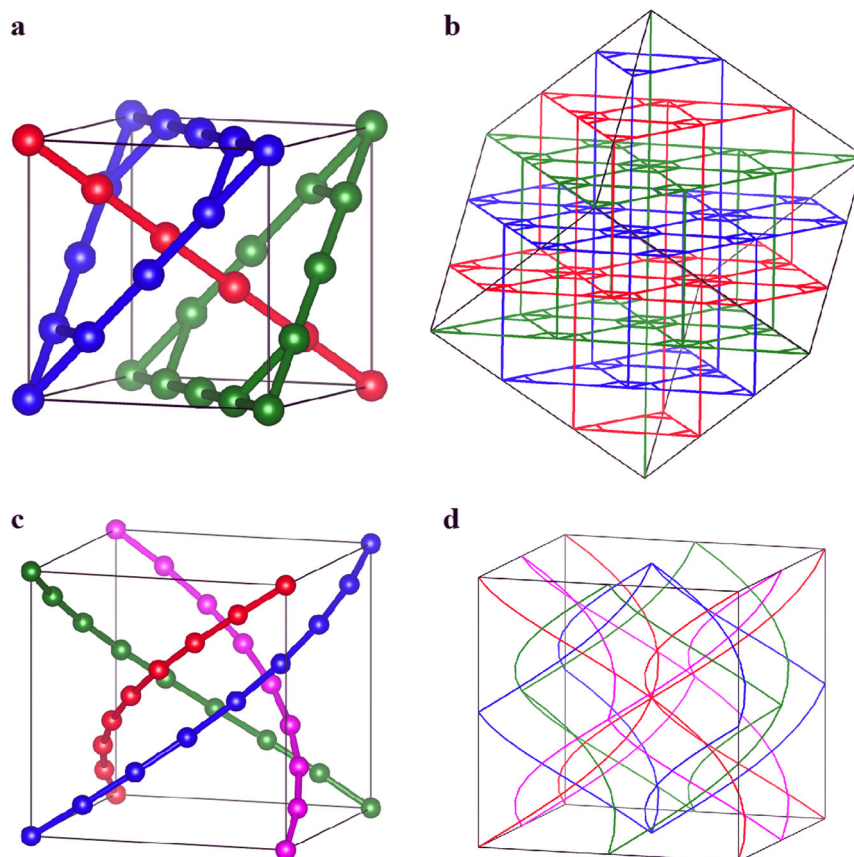
Fig. 5 The contrived examples of three and fourfold nets. Disconnected subnets are marked with different color. **a** A threefold net in the unit cell. **b** The threefold net in a $3 \times 3 \times 3$ supercell. **c** A fourfold net in the unit cell. **d** The fourfold net in a $2 \times 2 \times 2$ supercell.

considers equivalent atoms in a fixed $3 \times 3 \times 3$ supercell, which is not large enough for complicated structures. In practice, the size of the required supercell is unknown in advance. In the modified RDA proposed by Larsen et al.[15], a breadth-first-search (BFS) is used. The search starts from atoms in the original cell and then visits the nearest neighbors, second nearest neighbors, third nearest neighbors, …, successively. It terminates when the rank of the set of visited equivalent atoms does not increase. So the dimensionality can be determined in a finite number of steps for those components containing infinite atoms. As shown in the previous subsection, the offsets between equivalent atoms is the cycle sums of cycles in the QG. The rank stops increasing only when all the basic cycles have been considered. Therefore, both the modified RDA and the QG method aim to consider all the basic cell offsets(or basic cycle sums), but the latter is more elegant in theory and easier to understand. Moreover, the QG method provides an additional approach to calculating multiplicity as shown in the following subsection. Actually, the QG method described in the previous subsection is a type of RDA because it also starts from the definition Eq. (2).

Multiplicities of self-penetrating nets
Larsen et al.[15] have discussed the contrived self-penetrating helical networks and the improper connections between components which lead to incorrect dimensionality by TSA. Here we shall discuss self-penetrating nets following Thimm's approach[17] which provides more insights to this problem.

Cuprite with a space group of $Pn-3m$ is a typical example. As shown in Fig. 2a, the O atom with fractional coordinations of (0, 0, 0) in the original cell $n_{O_1}(0,0,0)$ is connected to the equivalent atoms $n_{O_1}(1,1,0)$, $n_{O_1}(0,1,1)$ and $n_{O_1}(1,0,1)$ through

copper and other oxygen atoms. This observation can be described by the cycle sum matrix from the QG of cuprite (Fig. 2d):

$$\mathbf{S} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \tag{7}$$

We found the O atom $n_{O_1}(1,0,0)$ is disconnected from $n_{O_1}(0,0,0)$ because (1, 0, 0) cannot be represented as a linear combination of three basic vectors in $\mathbf{S}$ if the coefficients are limited to integers. Actually, cuprite is composed by two disconnected subnets as shown in Fig. 2c, and they are equivalent because of the translational symmetry of crystal. Therefore, using TSA, we find one cluster containing 6 atoms ($N_1 = 6$) in the unit cell and in a $2 \times 2 \times 2$ supercell, the cluster expands to 24 atoms ($N_2 = 24$). The scaling factor N_2/N_1 is 4 which leads to an incorrect dimensionality of 2. Another example is $Ag(B(CN)_4)$ (Fig. 2b) found by Larsen et al.[15], which has the same topology as cuprite.

The net of cuprite contains two translationally equivalent but disconnected subnets, so its multiplicity is 2. For cuprite, the multiplicity $\hat{m}$ equals to absolute value of the determinant of $\mathbf{S}$ ($|\det(\mathbf{S})| = 2$)[17]. However, in general, the cycle sums matrix is not square. Instead, we should find the basic cycle sums $\tilde{\mathbf{S}}$. For a 3D net, $\tilde{\mathbf{S}}$ is a 3×3 matrix with minimum non-zero absolute value of determinant among all combinations of cycle sums. Then, the multiplicity equals to the absolute value of the determinant of $\tilde{\mathbf{S}}$:

$$\hat{m} = \left| \det(\tilde{\mathbf{S}}) \right|. \tag{8}$$

The definition of basic cycle sums for nets with arbitrary dimensionality is proposed by Thimm[17].

Thimm has proposed Eq. (8) however, he has not provided an explanation[17]. Here we demonstrate the relation between
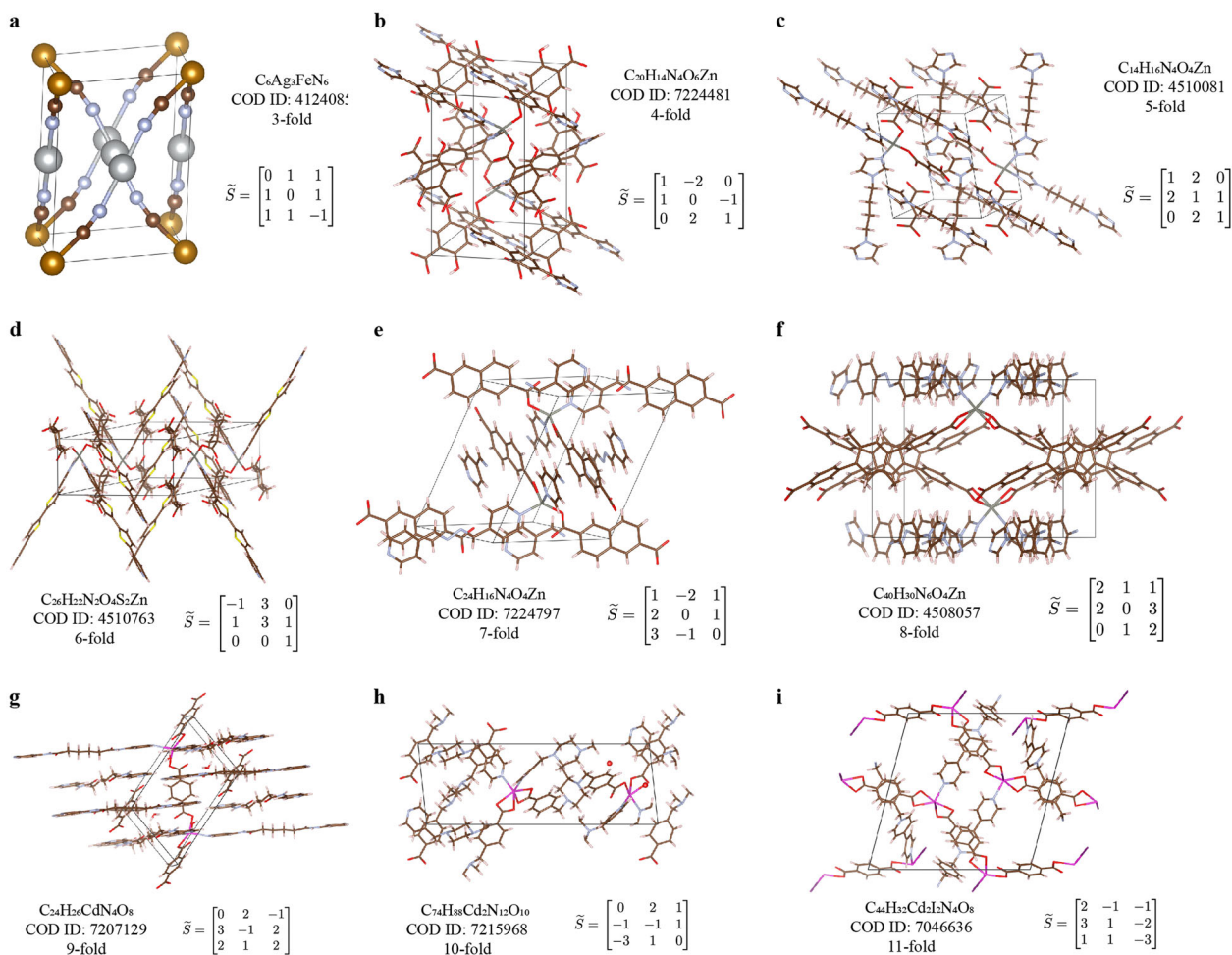
Fig. 6 Examples of self-penetrating 3D nets. The crystal structures, chemical formula, multiplicities, and basic cycle sums are presented for nets with multiplicities from 3 to 11.

**Table 1.** Number of nets with different multiplicities and dimensionalities $\hat{m}$ in COD.

$\hat{m}$	2	3	4	5	6	7	8	9	10	11
N_{3D}	259	105	28	23	3	2	1	1	1	1
N_{2D}	92	33	1	1						
N_{1D}	11	3								

multiplicity and determinant using a plane self-penetrating net shown in Fig. 3. The basic self-penetrating cell of the red subnet (relative to the unit cell) is defined by the basic cycle sums $\tilde{\mathbf{S}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. The cell contains two points, O and P. Since P is in the interior of the cell, its coordinate $(1, 0)$ is not an integer linear combination of the basic self-penetrating vectors $(1, 1)$ and $(1, -1)$. So it is disconnected from the red subnet. For an arbitrary self-penetrating net, we can always find an $\tilde{\mathbf{S}}$ to build the basic self-penetrating cell of subnets. Because of the translational symmetry, for each subnet, there is only one point in the basic self-penetrating cell. Therefore, the multiplicity of the net equals to the number of points in the cell, which is the volume/area of the basic self-penetrating cell and the volume/area is $\left|\det(\tilde{\mathbf{S}})\right|$. Blatov et al. have proposed a general algorithm to compute the multiplicities of interpenetrating nets[18,33–35]. We found the basic self-penetrating cell is actually the primitive interpenetration cell

proposed by Blatov et al.[18] Actually, the self-penetration discussed here is a special class of interpenetration with only translations.

In Fig. 4, we have listed examples of 3D net with different multiplicities. The schematics, QGs and the cycle sums of a usual 3D net are shown in Fig. 4a. The original cluster is connected to images in $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ cell and the multiplicity is 1. In Fig. 4b, the QG contains six edges which are along face diagonals in the schematic. It describes nets with multiplicity of 2 like cuprite. Although the QG is different from that of cuprite, their basic cycle sums are similar. Nets with larger multiplicities of 3 and 4 are also possible, as shown in Fig. 4c–d. We can implement edges in the QGs using carbon atomic chains. Multiple disconnected and equivalent components in supercells can be identified, as shown in Fig. 5. For fourfold nets, helical atomic chains are used to avoid intersections between edges in the schematic (Fig. 4d).

Usually, the maximum multiplicity of inorganic 3D nets is four[17] and it is related to Hadamard's maximum determinant problem[36], which requires finding the largest determinant for any square matrix composed by elements from a set. For inorganic crystals, the elements in basic cycle sums $\tilde{\mathbf{S}}$ are usually limited in $\{-1, 0, 1\}$, so $\tilde{\mathbf{S}}$ is a $(-1, 0, 1)$-matrix[37]. For $n = 1, 2, 3, 4, 5, \ldots$, the largest possible determinant for an $n \times n$ $(-1, 0, 1)$-matrix[37] is 1, 2, 4, 16, 48, $\ldots$. The sequence is the same to maximum multiplicities for n-dimensional nets[17]. If the elements in $\tilde{\mathbf{S}}$ are allowed to be larger than 1 or smaller than -1, the maximum multiplicity becomes higher. Such structures are shown in Fig. 6 and discussed below.
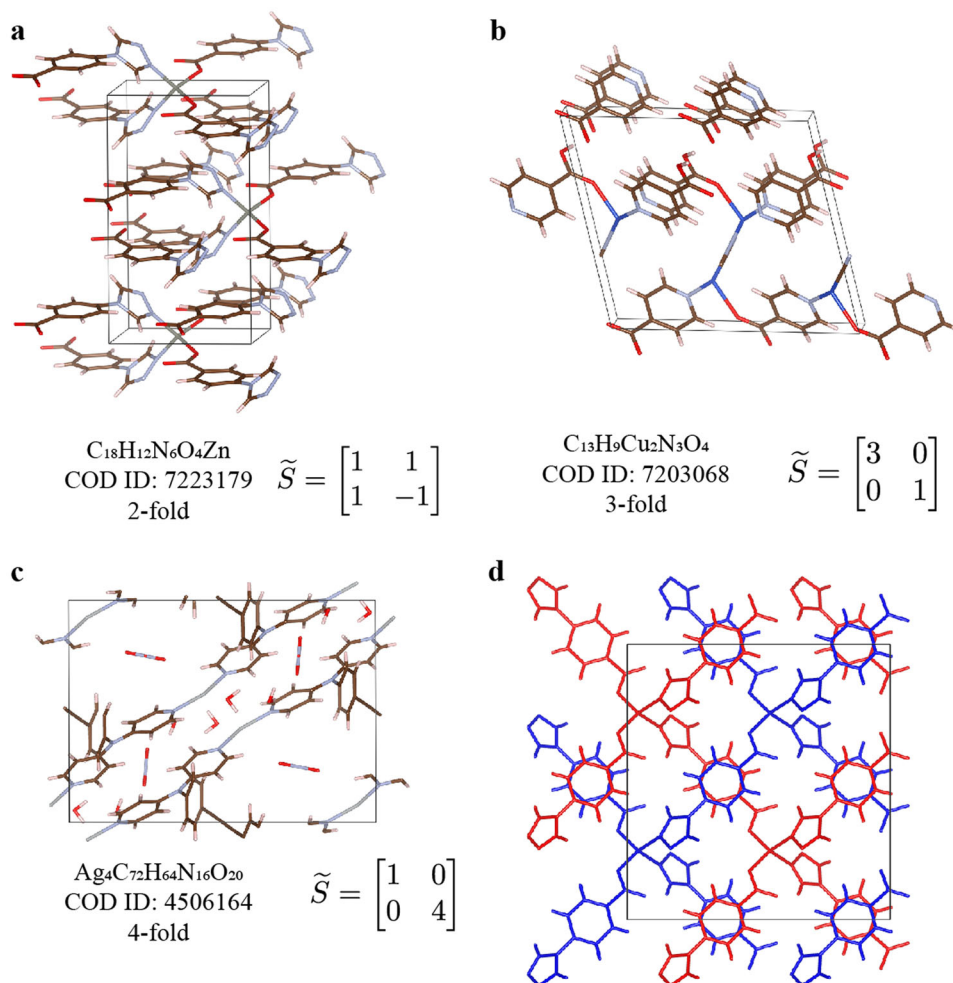
Fig. 7 Examples of self-penetrating 2D nets. a–c The crystal structures, chemical formula, multiplicities and basic cycle sums of 2D nets with multiplicities from 2 to 4. **d** The 2D twofold net in a $1 \times 2 \times 2$ supercell.

Based on the database built by Larsen et al.[38], we have found 3D, 2D and 1D nets with high multiplicities in COD using our method. The QGs of crystals highly depend on the value of k(Eq. (1)). For instance, if $k \to \infty$, all structures are identified as 3D and onefold nets. Larsen et al.[15] proposed a scoring parameter to determine the dimensionalities and k intervals and the results are provided in the database[38]. In this work, for each crystal, we used the low bound for the relative k interval to build the QG and determine the multiplicity. The screening results are shown in Table 1. Note that some nets are self-penetrating only when k is in a narrow interval.

The 3D nets with different multiplicities are shown in Fig. 6. The threefold structure, $Ag_3[Fe(CN)_6]$ (Fig. 6a), is similar to the contrived model shown in Fig. 5a since they have the same basic cycle sums. $Ag_3[Fe(CN)_6]$ and the isomorphic compound $Ag_3[Co(CN)_6]$ have been reported to be colossal thermal expansion materials[39,40].
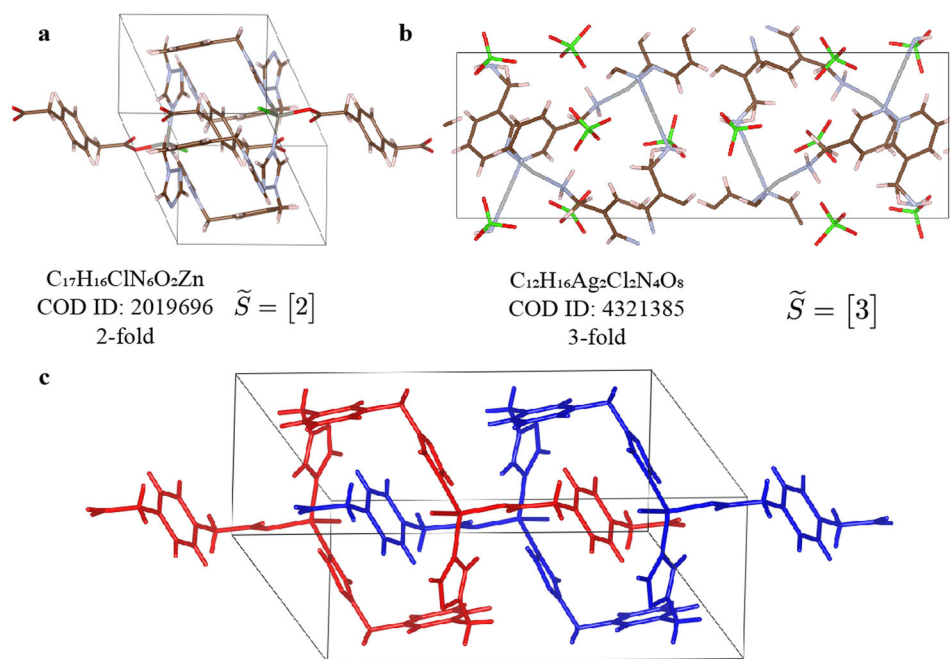
For 3D nets in the database, as shown in Table 1, the maximum multiplicity is 11, contrary to the conclusion that the maximum multiplicity of 3D nets is 4. We found that the crystals with multiplicities higher than 3 are all coordination complexes. The long chains allow connections between equivalent atoms in remote cells. Thus, the elements in the basic cycle sums are not limited in $\{-1, 0, 1\}$ and the compounds have high multiplicities.

We have also found low-dimensional self-penetrating structures shown in Table 1. In a 2D space, we cannot implement a twofold net since the edges will always intersect (Fig. 3). But in 3D crystals,

atomic chains can curve to form self-penetrating nets. For example, the 2D twofold complex (Fig. 7a) is similar to the plane self-penetrating net shown in Fig. 3. There are two 2-fold monolayers in the unit cell of the complex and they are stacked along the a axis. In Fig. 7d, we displayed one of the monolayers in a $1 \times 2 \times 2$ supercells and marked the two components using different colors. Two-dimensional nets with multiplicity of 3 and 4 also exist and the examples are shown in Fig. 7. We have also found a 5-fold 2D net but the compound (COD ID: 7216004) is self-penetrating only when k is in a very narrow interval [1.056, 1.085]. So it is not regarded as a penetrating polymer in the original reference[41]. For 1D structures, self-penetrating nets are very rare and we can only find two and threefold structures in the database, as shown in Fig. 8. The 1D twofold complex extends along the a axis and two translationally equivalent components are found to be entangled (Fig. 8c).

DISCUSSION
Mixed-dimensional materials contain multiple components which have different dimensionalities. We have done a statistical analysis on single and mixed dimensionalities in the self-penetrating structure set, as shown in Table 2. Mixed-dimensional structures are rare in the whole database[15]. However, the proportion of self-penetrating structures in the mix-dimensional set is relatively higher. For instance, 4.9% of 0D + 3D structures are self-penetrating while only 1.4% of "pure" 3D structures are

$C_{17}H_{16}ClN_6O_2Zn$
COD ID: 2019696 $\widetilde{S} = [2]$
2-fold

$C_{12}H_{16}Ag_2Cl_2N_4O_8$
COD ID: 4321385 $\widetilde{S} = [3]$
3-fold

Fig. 8 Example of self-penetrating 1D nets. a, b The crystal structures, chemical formula, multiplicities and basic cycle sums of 1D nets with multiplicities of 2 and 3. **c** The 1D twofold net in a $2 \times 1 \times 1$ supercell.

Table 2. Number of self-penetrating structures with different crystal dimensionality types.

Dimensions	0	1	2	3
0	0			
1	8 (0.3%)	6 (0.2%)		
2	45 (1.8%)	0	82 (2.3%)	
3	112 (4.9%)	0	0	311 (1.4%)

The percentages in the brackets represent the proportions of self-penetrating structures to all the structures with each dimensionality type. Items in the diagonal are the numbers of self-penetrating structures with a single dimension while items in the off-diagonal represent numbers of self-penetrating structures containing components of two different dimensionalities. Besides these self-penetrating structures with one or two different dimensional components, a complicated twofold structure containing 0D, 1D, and 3D components(COD ID: 4311765) can also be identified by the quotient graph method.

self-penetrating. This phenomenon is related to gas sorption capacities of penetrating structures. Compared with non-penetrating nets, self-penetrating nets have less pore size and pore volume. The small pore size and volume in a self-penetrating structure usually leads to stronger interactions between the 3D framework and 0D gas molecules, which improves gas sorption capacity[42]. The 0D + 3D structures usually contain 0D gas molecules and 3D frameworks, so the proportion of self-penetrating nets in the structure set with this dimensionality type is larger.

As shown in Fig. 9, we presented the distribution of crystal systems for the whole self-penetrating set and 3D twofold structures, respectively. More than 70% of self-penetrating structures are triclinic or monoclinic because most of them are organic polymers which have low symmetries. For 3D twofold structures, the distribution is roughly similar but the proportion of

cubic crystals is much larger. Actually, there are 19 cubic crystals in the self-penetrating set and all of them are 3D and twofold. These structures are isomorphic to Cu_2O and $Ag(B(CN)_4)$ shown in Fig. 2.

We selected 17 3D twofold compounds to evaluate their band gaps. There are two selection criteria. One is that the chosen structures should be distinct self-penetrating. It means they are self-penetrating in a relatively large k interval. The other is that the structure contains <30 atoms in the unit cell. After relaxation by density functional theory (DFT), 16 of the chosen structures maintain the original nets. Then we calculated the electronic structures of them using DFT and the results are shown in Table 3. Among the chosen structures, five compounds are metals or semimetals, and others have band gaps in a wide range from 0.5 to 6.3 eV. These self-penetrating materials might have application values in gas storage and selective catalysis because of their high gas sorption capacities[42]. Some of them with suitable band gaps also have potential applications in photocatalysis[11]. In addition, we found both Pb_2O and Ag_2O are topological semimetals, according to Materiae Database[43,44]. Pb_2O is a high symmetry point semimetal with band degeneracy at the R point while Ag_2O is a semimetal with band degeneracy on the X-Γ path. Their IDs in Materiae database are MAT00026156(Pb_2O) and MAT00026168 (Ag_2O), respectively.

Penetration in materials is usually related to mechanical properties. For example, interpenetrating polymer network (IPN) is a type of elastomer(rubber) which is composed by two or more network polymers[45,46]. IPNs based on two polymer materials can improve mechanical properties like tensile and tear strength[45]. Thus, IPNs have many applications and some commercial materials are IPNs. In recent years, researchers have also proposed new applications of IPNs such as high-performance electroelasto-mer artificial muscles[47]. The self-penetrating polymers screened in this work might have good performances in mechanical properties and wide potential applications. Because the interactions in atomic scale is much more complex, penetration in inorganic atomic networks is highly different from that in polymer materials. For example, we calculated Vickers hardnesses of inorganic 3D twofold structures which contain B, N, or C, but the highest
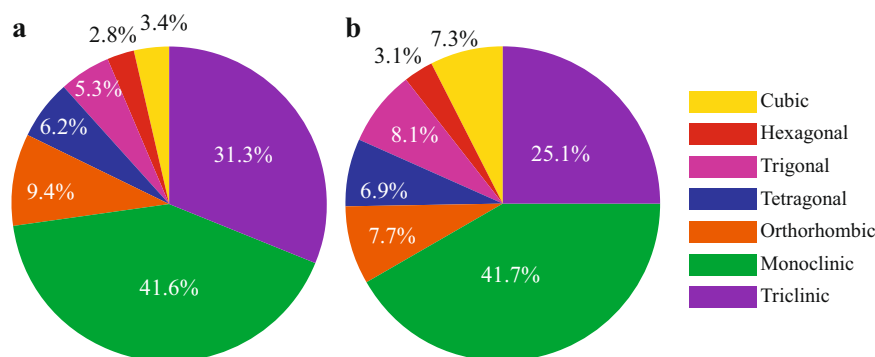
Fig. 9 The distributions of crystal systems. a All self-penetrating structures. **b** 3D twofold structures.

Table 3. Chemical formulas, COD IDs, space groups and band gaps for the chosen 3D twofold self-penetrating structures.

Formula	COD ID	Space group	Band gap (eV)
Cu_2O	9007497	$Pn-3m$	0.5
Pb_2O	1010382	$Pn-3m$	0
Ag_2O	4318188	$Pn-3m$	0
Au_2S	1510461	$Pn-3m$	2.0
CaC_2	9014958	$Pn-3m$	1.8
SiC_2N_4	1526364	$Pn-3m$	4.2
Fe_4N	9004225	$P-43m$	0
$Zn(CN)_2$	4124708	$P-43m$	5.9
$LiB(CN)_4$	4308584	$P-43m$	6.3
$CuB(CN)_4$	4308583	$P-43m$	3.2
$ZnPd(CN)_4$	7211957	$P4_2/mcm$	4.2
$ZnPt(CN)_4$	7211958	$P4_2/mcm$	4.0
$ZnNi(CN)_4$	7211956	$P4_2/mcm$	3.1
$AgAu(CN)_4$	4345897	$P2/c$	2.8
LiN_4C_6	4121263	$P2/c$	0
Ag_3BO_3	1509905	$R32$	0

The materials with zero band gaps are metals or semimetals.

hardnesses is only 13 GPa. So there are no simple relations between penetration in inorganic materials and their mechanical properties.

In summary, we discuss different dimensionality identification algorithms, such as TSA and RDA. And we find self-penetration in crystal nets will affect the reliability of previous methods. In this work, we use a method based on QG theory to determine the multiplicities of self-penetrating nets by the absolute value of the determinant of QG's basic cycle sums. Our approach allows for screening structures with target dimensionality and multiplicity in large databases. As a showcase, we have identified 1D, 2D, and 3D self-penetrating crystals in the COD database. The self-penetrating structures have various properties, which may lead to potential applications in different aspects, such as gas storage, selective catalysis or photocatalysis, etc.

METHODS
Graph theory methods
We have implemented the algorithms of QG based on Python packages NetworkX[48], NumPy[49], and ASE[50]. We also used ToposPro[20] to confirm the results.

Density functional theory details
Structural optimizations and electronic structure calculations are performed by VASP[51] using PBE functional[52]. Elastic properties are computed by the energy-strain method implemented in VASPKIT[53] interfaced with VASP. We used Tian's model[54] to evaluate Vickers hardnesses of materials.

DATA AVAILABILITY
The code and data in this article are available from the corresponding author upon request.

REFERENCES
1. Laughlin, R. B. Anomalous quantum Hall effect: an incompressible quantum fluid with fractionally charged excitations. *Phys. Rev. Lett.* **50**, 1395–1398 (1983).
2. Luttinger, J. M. An exactly soluble model of a many-fermion system. *J. Math. Phys.* **4**, 1154–1162 (1963).
3. Ferrari, A. C. et al. Science and technology roadmap for graphene, related two-dimensional crystals, and hybrid systems. *Nanoscale* **7**, 4598–4810 (2015).
4. Bhimanapati, G. R. et al. Recent advances in two-dimensional materials beyond graphene. *ACS Nano* **9**, 11509–11539 (2015).
5. Björkman, T., Gulans, A., Krasheninnikov, A. V. & Nieminen, R. M. Van der Waals bonding in layered compounds from advanced density-functional first-principles calculations. *Phys. Rev. Lett.* **108**, 235502 (2012).
6. Lebègue, S., Björkman, T., Klintenberg, M., Nieminen, R. M. & Eriksson, O. Two-dimensional materials from data filtering and Ab Initio calculations. *Phys. Rev. X* **3**, 031002 (2013).
7. Mitina, T. G. & Blatov, V. A. Topology of 2-periodic coordination networks: toward expert systems in crystal design. *Cryst. Growth Des.* **13**, 1655–1664 (2013).
8. Gorai, P., Toberer, E. S. & Stevanović, V. Computational identification of promising thermoelectric materials among known quasi-2D binary compounds. *J. Mater. Chem. A* **4**, 11110–11116 (2016).
9. Ashton, M., Paul, J., Sinnott, S. B. & Hennig, R. G. Topology-scaling identification of layered solids and stable exfoliated 2D materials. *Phys. Rev. Lett.* **118**, 106101 (2017).
10. Cheon, G. et al. Data mining for new two- and one-dimensional weakly bonded solids and lattice-commensurate heterostructures. *Nano Lett.* **17**, 1915–1923 (2017).
11. Zhang, X. et al. Computational screening of 2D materials and rational design of heterojunctions for water splitting photocatalysts. *Small Methods* **2**, 1700359 (2018).
12. Zhang, X. et al. An effective method to screen sodium-based layered materials for sodium ion batteries. *npj Comput. Mater.* **4**, 13 (2018).
13. Mounet, N. et al. Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds. *Nat. Nanotechnol.* **13**, 246–252 (2018).
14. Haastrup, S. et al. The Computational 2D Materials Database: high-throughput modeling and discovery of atomically thin crystals. *2D Mater.* **5**, 042002 (2018).
15. Larsen, P. M., Pandey, M., Strange, M. & Jacobsen, K. W. Definition of a scoring parameter to identify low-dimensional materials components. *Phys. Rev. Materials* **3**, 034003 (2019).

16. Zhou, J. et al. 2DMatPedia, an open computational database of two-dimensional materials from top-down and bottom-up approaches. *Sci. Data* **6**, 86 (2019).

17. Thimm, G. *A Graph Theoretical Approach to the Analysis, Comparison, and Enumeration of Crystal Structures*. Ph.D. thesis (2008).

18. Blatov, V. A., Carlucci, L., Ciani, G. & Proserpio, D. M. Interpenetrating metal–organic and inorganic 3D networks: a computer-aided systematic investigation. Part I. Analysis of the Cambridge structural database. *CrystEngComm* **6**, 377–395 (2004).

19. Chung, S. J., Hahn, T. & Klee, W. E. Nomenclature and generation of three-periodic nets: the vector method. *Acta Cryst. A* **40**, 42–50 (1984).

20. Blatov, V. A., Shevchenko, A. P. & Proserpio, D. M. Applied topological analysis of crystal structures with the program package ToposPro. *Cryst. Growth Des.* **14**, 3576–3586 (2014).

21. The Gavrog Project. http://gavrog.org/.

22. Eon, J.-G. Topological features in crystal structures: a quotient graph assisted analysis of underlying nets and their embeddings. *Acta Cryst. A* **72**, 268–293 (2016).

23. Winkler, B., Pickard, C. J., Milman, V. & Thimm, G. Systematic prediction of crystal structures. *Chem. Phys. Lett.* **337**, 36–42 (2001).

24. Strong, R. T., Pickard, C. J., Milman, V., Thimm, G. & Winkler, B. Systematic prediction of crystal structures: an application to sp^3-hybridized carbon polymorphs. *Phys. Rev. B* **70**, 045101 (2004).

25. He, C. et al. Complex low energy tetrahedral polymorphs of group IV elements from first principles. *Phys. Rev. Lett.* **121**, 175701 (2018).

26. Shi, X., He, C., Pickard, C. J., Tang, C. & Zhong, J. Stochastic generation of complex crystal structures combining group and graph theory with application to carbon. *Phys. Rev. B* **97**, 014104 (2018).

27. Ahnert, S. E., Grant, W. P. & Pickard, C. J. Revealing and exploiting hierarchical material structure through complex atomic networks. *npj Comput. Mater.* **3**, 35 (2017).

28. Isayev, O. et al. Universal fragment descriptors for predicting properties of inorganic crystals. *Nat. Commun.* **8**, 15679 (2017).

29. Xie, T. & Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **120**, 145301 (2018).

30. Chen, C., Ye, W., Zuo, Y., Zheng, C. & Ong, S. P. Graph networks as a universal machine learning framework for molecules and crystals. *Chem. Mater.* **31**, 3564–3572 (2019).

31. Jørgensen, P. B., Garijo del Río, E., Schmidt, M. N. & Jacobsen, K. W. Materials property prediction using symmetry-labeled graphs as atomic position independent descriptors. *Phys. Rev. B* **100**, 104114 (2019).

32. Gražulis, S. et al. Crystallography Open Database (COD): an open-access collection of crystal structures and platform for world-wide collaboration. *Nucleic Acids Res.* **40**, D420–D427 (2012).

33. Baburin, I. A., Blatov, V. A., Carlucci, L., Ciani, G. & Proserpio, D. M. Interpenetrating metal-organic and inorganic 3D networks: a computer-aided systematic investigation. Part II [1]. Analysis of the Inorganic Crystal Structure Database (ICSD). *J. Solid State Chem.* **178**, 2452–2474 (2005).

34. Carlucci, L., Ciani, G., Proserpio, D. M., Mitina, T. G. & Blatov, V. A. Entangled two-dimensional coordination networks: a general survey. *Chem. Rev.* **114**, 7557–7580 (2014).

35. V. Alexandrov, E., A. Blatov, V. & M. Proserpio, D. How 2-periodic coordination networks are interweaved: entanglement isomerism and polymorphism. *CrystEngComm* **19**, 1993–2006 (2017).

36. Hadamard's Maximum Determinant Problem. http://mathworld.wolfram.com/HadamardsMaximumDeterminantProblem.html.

37. (-1,0,1)-Matrix. http://mathworld.wolfram.com/-101-Matrix.html.

38. Definition of a scoring parameter to identify low-dimensional materials components — COMPUTATIONAL MATERIALS REPOSITORY. https://cmr.fysik.dtu.dk/lowdim/lowdim.html#lowdim.

39. Goodwin, A. L. et al. Colossal positive and negative thermal expansion in the framework material Ag3[Co(CN)6]. *Science* **319**, 794–797 (2008).

40. Goodwin, A. L. et al. Argentophilicity-dependent colossal thermal expansion in extended Prussian Blue analogues. *J. Am. Chem. Soc.* **130**, 9660–9661 (2008).

41. Ge, Y. et al. Assembly of a series of zinc coordination polymers based on 1,4-bis[2-(4-pyridyl)ethenyl]-2,3,5,6-tetramethylbenzene and 1,3-benzenedicarboxylate derivatives. *CrystEngComm* **16**, 6621–6629 (2014).

42. Jiang, H.-L., Makal, T. A. & Zhou, H.-C. Interpenetration control in metal–organic frameworks for functional applications. *Coordination Chem. Rev.* **257**, 2232–2249 (2013).

43. Zhang, T. et al. Catalogue of topological electronic materials. *Nature* **566**, 475–479 (2019).

44. Materiae - Material Sciences Database | IOP and CNIC, CAS | Materiae. http://materiae.iphy.ac.cn/.

45. Sperling, L. H. In *Interpenetrating Polymer Networks*, vol. 239 (eds Klempner, D., Sperling, L. H., Utracki, L. A.), 3–38 (American Chemical Society, 1994).

46. Roland, C. M. In *Encyclopedia of Polymeric Nanomaterials* (eds obayashi, S. & Müllen, K.) 1–9 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2021).

47. Ha, S. M., Yuan, W., Pei, Q., Pelrine, R. & Stanford, S. Interpenetrating polymer networks for high-performance electroelastomer artificial muscles. *Adv. Mater.* **18**, 887–891 (2006).

48. Hagberg, A. A., Schult, D. A. & Swart, P. J. Exploring Network Structure, Dynamics, and Function using NetworkX. In *Proc. 7th Python in Science Conference (SciPy 2008)* (eds Varoquaux, G., Vaught, T. & Millman, J.) 11–15 (Pasadena, CA USA).

49. van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy array: a structure for efficient numerical computation. *Compu. Sci. Eng.* **13**, 22–30 (2011).

50. Larsen, A. H. et al. The atomic simulation environment—a Python library for working with atoms. *J. Phys.: Condens. Matter* **29**, 273002 (2017).

51. Kresse, G. & Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).

52. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

53. Wang, V., Xu, N., Liu, J.-C., Tang, G. & Geng, W. T. VASPKIT: a pre- and post-processing program for VASP code. Preprint at http://arxiv.org/abs/1908.08269 (2019).

54. Tian, Y., Xu, B. & Zhao, Z. Microscopic theory of hardness and design of novel superhard crystals. *Int. J. Refractory Metals Hard Mater.* **33**, 93–106 (2012).

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

J.S. and H.G. designed the study and methodology. H.G. wrote the code and performed the calculations. J.S. and H.G. made the analysis and wrote the manuscript. All authors discussed the results and commented on the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to J.S.

Reprints and permission information is available at http://www.nature.com/reprints

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