On the topology of the space of coordination geometries

John Çamkıran$^{1,a}$, Fabian Parsch$^{2,1}$, and Glenn D. Hibbard$^1$

$^1$ Department of Materials Science and Engineering, University of Toronto, Toronto, ON M5S 3E4, Canada
$^2$ Department of Mathematics, University of Toronto, Toronto, ON M5S 2E4, Canada

Received 11 August 2022 / Accepted 29 March 2023 / Published online 8 June 2023
© The Author(s), under exclusive licence to EDP Sciences, SIF and Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract. Coordination geometries describe the arrangement of the neighbours of a central particle. Such geometries can be thought to lie in an abstract topological space, a model of which could provide a mathematical basis for understanding physical transformations in crystals, liquids, and glasses. Through the generalisation of a recently proposed local order parameter, the present work conceives a metric model of the space of three-dimensional coordination geometries. This model appears to be consistent with elementary geometry and suggests a taxonomy of coordination geometries with five main classes. A quantifier of coordination-geometric typicality is derived from the metric. By the statement of a postulate on the topology of the space being modelled, the range of structures that are possible to resolve using the local order parameter is greatly increased.

1 Introduction

In both crystalline and noncrystalline materials, the geometry of the immediate surrounding of a particle often undergoes transformation between distinct geometries of coordination [1–4]. A scientific understanding of precisely which transformations are possible is precluded by the lack of a mathematical model of the abstract topological space that coordination geometries can be thought to lie in. Prior discussion on such a space is limited [5–7], presumably because there is little of interest to be said about it from the traditional perspective of symmetry.

In a recent article, the authors of this work discussed a local structural phenomenon called extracopularity, the tendency of particles in condensed phases to have far fewer different bond angles than combinatorially possible [8]. Extracopularity is an informational redundancy that accompanies local orientational order and is quantified by the extracopularity coefficient $E$, an order parameter in the sense of the work by Steinhardt and colleagues [9]. Being a strictly local quantity, the extracopularity coefficient provides limited insight on the relationships between coordination geometries, a point that is illustrated in Fig. 1.

To enable the study of these relationships, the present work generalises $E$ to $n$ particles. The generalised quantity is a statistical expansion of the original one and therefore enjoys a similar, information-theoretic interpretation. A metric for three-dimensional coordination geometries arises naturally from the generalisation, and the space endowed with this metric exhibits clustering around geometries that are similar in construction, thereby validating the latter as a model of the hypothetical abstract space of coordination geometries. The space also exhibits features that do not follow directly from similarities in construction, making it a possible new source of theoretical insight on the atomic-scale structure of crystals, liquids, and glasses [6,7,10].

The main computational challenge with a metric based on an $n$-particle generalisation of $E$ is evaluating the latter quantity, which requires the explicit discretisation of bond angles. In this work, we consider a method for discretising bond angles that is in a certain sense optimal. This method is based on the observation that bond angles are more likely to take certain values than others. The method can also be used to compute one-particle extracopularity coefficients with greater resolution than previously possible. It therefore has immediate practical implications on structural analysis in the molecular dynamics setting.

The remainder of this work is organised as follows: Section 2 reviews $E$ and discusses its generalisation to $n$ particles. Section 3 proposes a metric for three-dimensional coordination geometries based on this generalisation. Section 4 addresses the computational problem of bond angle discretisation. Section 5 studies the topology of the space of three-dimensional coordination geometries. Finally, Section 6 discusses a few implications of our results.
2 Generalisation

We begin this section by briefly reviewing the one-particle extracopularity coefficient. We then make precise the behaviour expected of its \( n \)-particle generalisation. Lastly, we consider one such generalisation, obtain its closed form, and check that it behaves in the expected way.

2.1 Review

Let a \textit{system} be a discrete nonempty set \( S \) of points in three-dimensional Euclidean space, and call each point \( p \in S \) a \textit{particle}. The \textit{neighbourhood} \( N(p) \) of a particle \( p \) is some subset of \( S \) not containing \( p \), and its elements \( q \in N(p) \) are called the \textit{neighbours} of \( p \). Determining precisely which subset of a system corresponds to the neighbourhood of a given particle is a task that continues to elude our ability. In the empirical context, it can be estimated using one of various algorithms \cite{11}. In the present work, prior knowledge of the neighbourhood is assumed. Given a particle \( p \) and any of its neighbours \( q \), call the vector \( q-p \) a \textit{bond}. Consider a particle \( p_i \) with at least two neighbours. Let \( B_i \) denote the set of unordered pairs of bonds associated with it, called its \textit{bond pairs}, and let \( \Theta_i \) denote the set of smaller angles made by these pairs, called its \textit{bond angles}. Define the \textit{information content} \( I(A) \) of a discrete set \( A \) by \( I(A) = \log_2 |A| \) \cite{12}. The \textit{extracopularity coefficient} \( E_i \) of the particle \( p_i \) is defined by the following difference:

\[
E_i = I(B_i) - I(\Theta_i). \tag{1}
\]

Information theoretically, this quantity can be interpreted as the difficulty that remains in the hypothetical search for a specific bond pair once the angle made by that pair is known. Letting \( k_i = |B_i| \) and \( m_i = |\Theta_i| \), one can write \( E_i \) more explicitly as follows:

\[
E_i = \log_2 \left( \frac{k_i^2 - k_i}{2m_i} \right). \tag{2}
\]

Below we generalise this quantity to \( n \) particles.

2.2 Desideratum

The simple fact that some quantity generalises the extracopularity coefficient to \( n \) particles does not necessarily make it useful. Hence, before devising such a quantity, we must carefully consider the behaviour needed to render it so.

Let \( E_{1...n} \) denote the extracopularity coefficient of a collection of particles. If each of these particles were to have the same coordination geometry, it would only be natural for the coefficient of the collection to be equal to the coefficient corresponding to that geometry. For instance, if these \( n \) particles are arranged, say, as an FCC crystal, then one would expect to have \( E_{1...n} = E_{\text{FCC}} \approx 4.04 \).

\[
E_{\text{FCC}} \approx 4.04. \tag{3}
\]

In addition to the above, it would not make much sense to suggest a particle collection to be more ordered as a whole than locally around any of its constituent particles. We capture both of the above behaviours in the following upper bound:

\[
E_{1...n} \leq \max\{E_1, \ldots, E_n\}, \tag{3}
\]

with equality if and only if \((k_i, \Theta_i) = (k_j, \Theta_j)\) for all \(i,j\).

A necessary condition for this upper bound to hold with equality is that \( E_i = E_j \) for all \(i,j\). This condition is not, however, sufficient, as the fraction in the argument of the logarithm in Eq. (2) is not in general unique to the pair \((k_i, \Theta_i)\).

2.3 Definition

Having formalised the desired behaviour, we are now ready to devise a generalisation able to satisfy it. Define the \textit{extracopularity coefficient} \( E_{1...n} \) of a collection of \( n \) particles by

\[
E_{1...n} = \langle I(B_i) \rangle - I(\Theta_{1...n}), \tag{4}
\]

where \( \langle \cdot \rangle \) denotes the arithmetic average and \( \Theta_{1...n} \) denotes the set comprising the bond angle of all \( n \) particles.

Much like the one-particle coefficient \( E_i \), the \( n \)-particle coefficient \( E_{1...n} \) can be expressed more explicitly in terms of the number of bonds and different bond angles. From elementary combinatorics, we have

\[
I(B_i) = \log_2 |B_i| \tag{5}
\]

\[
= \log_2 \left( \frac{k_i^2 - k_i}{2} \right). \tag{6}
\]

Taking the arithmetic average of this quantity gives us the first term in the definition of \( E_{1...n} \).

\[
\langle I(B_i) \rangle = \frac{1}{n} \sum_{i=1}^{n} \log_2 \left( \frac{k_i^2 - k_i}{2} \right) \tag{7}
\]

\[
= \frac{1}{n} \log_2 \left[ \prod_{i=1}^{n} \left( \frac{k_i^2 - k_i}{2} \right) \right] \tag{8}
\]

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{Fig_1.png}
\caption{Comparing the bond angles of the face-centred cubic (FCC) and hexagonal close-packed (HCP) coordination geometries. The presence of many common bond angles suggests a strong similarity between the geometries. Such is not apparent from their one-particle coefficients, \( E_{\text{FCC}} \approx 4.04 \) and \( E_{\text{HCP}} \approx 3.46 \).}
\end{figure}
The second term in the definition can be expanded as follows:

\[ I(\Theta_1...n) = \log_2 |\Theta_1...n| \]

\[ = \log_2 \left| \bigcup_{i=1}^{n} \Theta_i \right|. \]

Combining Eqs. (4), (10), and (12) reveals a general formula for the extracopularity coefficient of \( n \) particles,

\[ E_{1...n} = \log_2 \left[ \frac{\sqrt[n]{\prod_{i=1}^{n} (k_i^2 - k_i)}}{2^{n-1}} \right]. \]

Observe that taking \( n = 1 \) recovers Eq. (2), demonstrating that what we have obtained is indeed a generalisation of the one-particle case. While the generalisation admits any whole number \( n \), the focus of the present work is on the special case of \( n = 2 \).

2.4 Bounds

Despite its clear desirability, the upper bound discussed in Sect. 2.2 is difficult to prove, as doing so would require a precise understanding of the relationship between bond angles and the number of bonds. There are nonetheless certain indications that it may be satisfied. Notably, the following looser bound is found to hold:

\[ E_{1...n} \leq \log_2 \left[ \frac{\max_i (k_i^2 - k_i)}{2 \min_i (m_i)} \right]. \]

with equality if and only if \( (k_i, \Theta_i) = (k_j, \Theta_j) \) for all \( i, j \).

Let us show that this is indeed the case. The argument of the logarithm in the formula for \( E_{1...n} \) is a positive fraction, which increases with its numerator and decreases with its denominator. Observe that the numerator of the argument is simply the geometric average of \( k_i^2 - k_i \). It is well known that a geometric average cannot exceed the largest of the numbers being averaged,

\[ \sqrt[n]{x_1...x_n} \leq \max\{x_1, ..., x_n\}, \]

with equality if and only if \( x_i = x_j \) for all \( i, j \). Next observe that the denominator of the argument is twice the cardinality of a union. Certainly, the cardinality of a union cannot be less than the cardinality of the sets under union,

\[ \bigcup_{i=1}^{n} A_i \geq \min\{|A_1|, ..., |A_n|\}, \]

with equality if and only if \( A_i \) are identical. Equation (14) is now immediate.

For completeness, let us also consider the corresponding lower bound. A geometric average cannot be smaller than the smallest of the numbers being averaged,

\[ \sqrt[x_1...x_n]{x_1...x_n} \geq \min\{x_1, ..., x_n\}. \]

And the cardinality of a union is no larger than the sum of the cardinalities of the sets under union,

\[ \bigcup_{i=1}^{n} A_i \leq \sum_{i=1}^{n} |A_i|, \]

Thus, \( E_{1...n} \) is bounded from below as follows:

\[ E_{1...n} \geq \log_2 \left[ \frac{\min_i (k_i^2 - k_i)}{2(m_1 + ... + m_n)} \right]. \]

3 Metric

A quantitative understanding of the relationships between coordination geometries is of both practical [13,14] and theoretical [7,15] interest. Prior efforts to capture these relationships have employed general dissimilarity functions [5,15,16]. However, such functions lack one or more of the properties expected of a quantifier of the degree of dissimilarity between two objects [17]. Functions that satisfy these properties are called metric dissimilarity functions, distance functions, or simply metrics. Given a set \( M \), a map \( d : M \times M \to \mathbb{R} \) is said to be a metric on \( M \) if it satisfies the following three properties:

\[ \begin{align*}
\text{I.} & \quad d(x, y) = 0 \text{ if and only if } x = y \\
\text{II.} & \quad d(x, y) = d(y, x) \\
\text{III.} & \quad d(x, z) \leq d(x, y) + d(y, z)
\end{align*} \]

Here we seek such a function on the set \( G \) of threedimensional coordination geometries. Given two geometries \( g, h \in G \), define the extracopularity distance \( d_E(g, h) \) between them by

\[ d_E(g, h) = \max\{E_g, E_h\} - E_{gh}, \]

where \( E_g \) is understood to be the extracopularity coefficient of a single particle with coordination geometry \( g \) and \( E_{gh} \) that of a pair of particles with geometries \( g \) and \( h \). Property II is immediate from the symmetry of the set maximum and the two-particle extracopularity
coefficient. While Properties I and III are difficult to prove analytically, they can be shown numerically for 22 of the most important geometries of coordination, which are outlined in Appendix A.

4 Bond angle discretisation

As visible from Eq. (21), the extracopularity distance is ultimately the difference between a one-particle and a two-particle extracopularity coefficient. The key to computing both kinds of coefficients is in determining the number of different bond angles $|\Theta|$ for the particle(s) being studied. This is typically much lower than the naive count of bond angles due to equivalences between angles. For example, the three instances of the 180° degree bond angle in a simple cubic neighbourhood may be measured as 176°, 178.7°, and 179.2°. Thus, a naive count of these angles would be overcounting 180° by two. Visibly, each instance of an angle must be mapped to the class representative (180° in our example) before an informative count can take place. We call this procedure bond angle discretisation. In this section, we discuss a bond angle discretisation method that is optimal in a specific sense.

4.1 Approach

The most natural approach to bond angle discretisation is what one might call fixed discretisation, wherein the interval (0, 180°) is partitioned into subintervals called bins, each having fixed endpoints, called edges. Such an approach is justified by the observation that bond angles do not take all values in (0, 180°) with equal probability density. Observe that the repulsive forces that prevent (physical) particles from getting too close to each other also make certain angles more likely than others. The angle of 60° for instance is quite common, as it is observed whenever three particles all neighbour each other with equal radial distance. Thus, 60° can be seen as an inherent bond angle of particle packings. Given a comprehensive list of such bond angles, fixed discretisation can be performed optimally by placing bin edges at their midpoints. A technicality of this approach is discussed in Appendix B.

4.2 Method

The numerical results of the present work were obtained using a method of bond angle discretisation that takes the fixed approach. The method begins by listing the bond angles corresponding to commonly encountered geometries of coordination. A geometry is omitted from this list if any other geometry can be described as a capping (or augmentation) of it. For instance, we include BSA but not SA and CSA, as the former constitutes a capping of the latter. This is done to avoid biasing the bond angle list towards clusters of geometries that are essentially identical in construction. Then, inherent angles are taken to be the points in (0, 180°] that are locally maximal in bond angle density (0 is taken to be inherent by convention). These points are established using the algorithm DBSCAN [18], which takes the following two arguments: the minimum number of nearby points needed for a point to be considered an inherent angle candidate, abbreviated minPts, and the search radius \(\varepsilon\) that defines the term ‘nearby’.

For the first argument, we simply pick the most conservative value, minPts = 1, in order to avoid disqualifying any point from inherent angle candidacy. The choice of the second argument \(\varepsilon\) is more subtle. Visibly, larger values of \(\varepsilon\) will lead to coarser and hence more statistically robust discretisations. If chosen too large, however, substantially different angles will be equilibrated, rendering the result of the discretisation uninformative. To help establish the optimal compromise, we introduce a two-part postulate regarding the topology of the space of coordination geometries:

1. (a) FCC is closer to HCP than it is to BCC.
   (b) HCP is closer to FCC than it is to BCC.
2. (a) CSA and BSA are the two closest to SA.
   (b) CSP and BSP are the two closest to HDR.

The first part asserts the close resemblance between the constructions of the convex polyhedra corresponding to the FCC and HCP coordination geometries, namely the cuboctahedron and anticuboctahedron. The second part is justified by the fact that capping only constitutes a minimal change to any geometry. We thus arrive at our choice of \(\varepsilon = 2.85\), the largest value to two decimal places for which the above postulates are satisfied.

![Fig. 2 Heatmap of extracopularity distances for a selection of commonly encountered geometries. Values are rounded to one decimal place. Rows and columns are ordered as per Ref. [19]](image)
5 The topology

We studied the topology implied by $d_E$ numerically for 22 of the most commonly encountered geometries of coordination (see Appendix A). The 231 extracopularity distances between these geometries are summarised in Fig. 2 and given in full by Fig. 5. These distances were found to be consistent with the properties of a metric, stated in Eq. (20), and the extracopularity coefficients they were computed from were found to obey the desired upper bound, given in Eq. (3). We employed two common techniques to make sense of the distance data, namely hierarchical clustering and multidimensional scaling.

5.1 Hierarchical clustering

Hierarchical clustering is the construction of a binary tree through the iterative pairing of a set of items given the distances between them [20]. Each iteration pairs the two items that are closest to each other, which are thereafter treated as a single item. The distance to/from this new item is calculated as the average distance to/from its constituents. The iteration stops when only one item remains. Fig. 3 illustrates the result of hierarchical clustering for commonly encountered coordination geometries based on the extracopularity distance.

The tree was found to exhibit clustering around collections of geometries with similar construction, namely those pentagonal prismatic, square antiprismatic, trigonal prismatic, bipyramidal, cubic-cored, and pentagonal antiprismatic. Among the less expected pairings were that of HCP with the pentagonal prisms, FCC with the pentagonal antiprisms, and SDS with the trigonal prisms. The technique we discuss next offers some perspective on these pairings.

5.2 Multidimensional scaling

Multidimensional scaling is the practice of embedding a set of items into an abstract Cartesian space given their pairwise distances [21]. We used nonclassical multidimensional scaling with metric stress, which is the algorithm appropriate for non-Euclidean metrics. We established eight as a sufficient number of dimensions for the scaling, as additional dimensions were not found to lead to further reductions in stress. Fig. 4 depicts the graph implied by the Delaunay triangulation [22] of the two-dimensional approximation of the full, eight-dimensional scaling.

5.2.1 Taxonomy

The graph appeared to suggest five main classes of coordination geometries: (1) spheroidals, characterised by their high sphericity; (2) ellipsoidals, characterised by their high moment of inertia; (3) cuboidals, characterised by their cubic core; (4) bipyramidals, characterised by their bipyramidal form; and (5) the tetrahe-
Table 1 For each class, averages of sphericity Ψ, moment of inertia per neighbour I/k, and typicality τ

| Class             | ⟨Ψ⟩  | ⟨I/k⟩ | ⟨τ⟩  |
|-------------------|------|-------|------|
| 1. Spheroidal      | 0.896| 1.09  | 0.607|
| 2. Ellipsoidal     | 0.836| 1.32  | 0.519|
| 3. Bipyramidal     | 0.838| 1.00  | 0.612|
| 4. Cuboidal        | 0.849| 1.12  | 0.597|
| 5. Tetrahedral     | 0.671| 1.00  | 0.390|

Upon close inspection, we found the positions of individual geometries within classes to exhibit consistency with the overall taxonomy. For example, the proximity of TET to the cuboids makes sense given that its four bonds correspond to four of the eight edges of a cube. Likewise, the proximity of PBP and HBP to the spheroids is justified by the fact that they are the bipyramidal geometries of highest sphericity, as is the proximity of BTP and SDS to the spheroids, given that they are the trigonal prisms of highest sphericity.

5.2.2 Typicality

An important observation concerning one-particle E is that it implies ICO to be the maximally ordered coordination geometry among those commonly encountered and FCC to be that among those that fill three-dimensional Euclidean space [8]. An interesting question along this line of inquiry is how ‘typical’ a given coordination geometry is relative to others. One can quantify the typicality τ(g) of a geometry g through the base-2 exponential of the additive inverse of its extra-copularity distance from the centroid ⟨h⟩ of a set of reference geometries,

$$\tau(g) = 2^{-d_{E}(g,⟨h⟩)}.$$  \hspace{2cm} (22)$$

We computed typicalities numerically using the eight-dimensional coordinates obtained through multidimensional scaling, taking ⟨h⟩ to be the componentwise average over all 22 coordination geometries under study. We found TET to be the geometry of lowest typicality, which is not surprising given its anomalously low coordination number and bond angle count. Meanwhile, HBP was found to be the geometry of highest typicality, which also makes sense given that it is quite usual in both regards. Table 1 states class-average typicalities, while Table 2 provides typicalities for individual geometries.

Table 2 Commonly encountered geometries of coordination, sorted by increasing E

| Description            | Coordination geometry | Polyhedral classification | Parameters |
|------------------------|-----------------------|---------------------------|------------|
| Abbrev.                |                       |                           | k  m  E  τ  Ψ  I/k |
| TBP                    | Trigonal bipyramidal  | Deltahedral, bipyramidal  | 5 3 1.737 0.4704 0.7563 1.000 |
| SDS                    | Snub disphenoidal     | Deltahedral               | 6 2 2.222 0.4866 0.8413 1.444 |
| PBP                    | Pentagonal bipyramidal| Deltahedral, bipyramidal  | 7 4 2.392 0.6054 0.8696 1.000 |
| CTP                    | Capped trigonal prismatic | Prismatic               | 7 4 2.392 0.5080 0.8025 1.333 |
| BTP                    | Bicapped trigonal prismatic | Prismatic       | 8 5 2.485 0.5616 0.8630 1.191 |
| TET                    | Regular tetrahedral   | Platonic, deltahedral     | 4 1 2.585 0.3899 0.6711 1.000 |
| HBP                    | Hexagonal bipyramidal | Deltahedral               | 8 4 2.807 0.7815 0.8787 1.000 |
| CSA                    | Capped square antiprismatic | Antiprismatic        | 9 5 2.248 0.6700 0.7878 1.238 |
| CSP                    | Capped square prismatic | Prismatic               | 9 5 2.248 0.6286 0.8272 1.279 |
| TTP                    | Tricapped trigonal prismatic | Prismatic, deltahedral | 9 5 2.248 0.6198 0.9062 1.000 |
| SC                     | Regular octahedral    | Platonic, deltahedral, bipyramidal | 6 2 2.907 0.5922 0.8456 1.000 |
| BSA                    | Bicapped square antiprismatic | Deltahedral, antiprismatic | 10 6 2.907 0.7276 0.8853 1.176 |
| BSP                    | Bicapped square prismatic | Prismatic               | 10 5 3.170 0.6168 0.8579 1.067 |
| CPP                    | Capped pentagonal prismatic | Prismatic              | 11 6 3.196 0.6635 0.8695 1.200 |
| SA                     | Square antiprismatic  | Antiprismatic             | 8 3 2.222 0.6187 0.8595 1.000 |
| HDR                    | Regular hexahedral    | Platonic, prismatic       | 8 3 3.222 0.6325 0.8600 1.000 |
| BPP                    | Bicapped pentagonal prismatic | Prismatic             | 12 7 3.237 0.6913 0.9095 1.000 |
| HCP                    | Anticuboctahedral     | Bicupolar                 | 12 6 3.459 0.6889 0.9050 1.000 |
| BCC                    | Rhombic dodecahedral  | Catalan                    | 14 6 3.923 0.5086 0.9047 1.143 |
| FCC                    | Cuboctahedral         | Bicupolar                 | 12 4 4.044 0.5519 0.9050 1.000 |
| CPA                    | Capped pentagonal antiprismatic | Antiprismatic        | 11 3 4.196 0.4949 0.8967 1.200 |
| ICO                    | Regular icosahedral   | Platonic, deltahedral, antiprismatic | 12 3 4.459 0.4467 0.9393 1.000 |

a—Trigonal dodecahedral, b—Square bipyramidal, C—Square bipyramidal, Square prismatic, cubic, d—Triangular orthobicupolar e—Triangular gyrobicupolar, f—Bicapped pentagonal antiprismatic.
6 Discussion

In essence, this work advances a redundancy-based perspective on coordination geometries as an alternative to traditional symmetry-based thinking. Symmetry is certainly a very useful notion in many respects [23]. However, there are at least two ways in which it is unsuited to the study of the relationships between coordination geometries. Firstly, it is degenerate to the extent of not being able to distinguish between basic crystal structures (e.g. FCC, BCC, and SC). Secondly, small changes to a given geometry, such as the removal of a neighbour, can lead to large differences in its symmetry (cf. ICO and CPA). In assigning positive distances to all 231 pairs of coordination geometries herein studied (Fig. 5), our redundancy-based perspective does not appear to evidence any degeneracies. And in suggesting a geometrically meaningful taxonomy (Fig. 4), it does not appear to avoid degeneracies at the expense of informativeness. The central result of this work is our metric model of the space of three-dimensional coordination geometries. It is conceivable that the metric captures, if only partially, the energetic or entropic cost that would be associated with a transformation from one geometry to another [3]. Certainly, the well-known pathways of Bain [24] and Burgers [25] are consistent with this hypothesis, as the geometries they involve lie in close proximity to one another in our model. However, further investigation is needed before any conclusive statements can be made.

In addition to the theoretical implications discussed above, our results also have practical implications on the problem of local structural indication [13,14]. Through the statement of a postulate on the topology of the space of coordination geometries, we arrive at a method of discretising bond angles that is optimal in a certain sense. Bond angle discretisation makes it possible to compute one-particle  $E$ with greater resolution than by the workaround discussed in Ref. [8]. This, in turn, greatly increases the range of structures that it can indicate in the computational setting.

Acknowledgements. The authors would like to thank Alán Aspuru-Guzik and Chandra Veer Singh for their helpful discussions.

Fig. 5 Heatmap of extracopularity distances for all 22 commonly encountered geometries. Values are rounded to one decimal place. Rows and columns are ordered as per Ref. [19]
Author contributions

JC conceived the project, obtained the results, and wrote the manuscript. FP and GDH contributed to the results and the writing.

Data availability statement The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Appendix A: Commonly encountered geometries

Table 2 lists 22 of the most commonly encountered geometries of coordination. These include 11 zero-lone-pair molecular geometries predicted by valence shell electron pair repulsion theory [26] and eight solutions of the Thomson problem [27], the latter of which also happen to be minimum-energy sphere packings for \( k \leq 12 \) [28]. Mathematically, these correspond to the first four Platonic solids, all eight strictly convex deltahedra, 12 capped (anti)prisms, two circumscribable bicupolae, and the rhombic dodecahedron (a Catalan solid). Figure 5 depicts their extracopularity distances.

Appendix B: A technicality of fixed discretisation

One of the inevitabilities of the fixed approach to bond angle discretisation is the confusion of angles that are close yet unequal even in the ideal form of a coordination geometry. Among the geometries herein studied, this issue is observed to afflict those of type (X)PP, (X)SA, (X)TP, and SDS. In the analysis described in Sect. 5, we corrected their bond angle counts as follows. Let \( f_a(\theta) \) denote the number of close yet unequal angles that are mapped to the class representative \( \theta \) for a coordination geometry \( g \). Then, our correction to \( |\Theta_{gh}| \) is given by

\[
|\Theta_{gh}|^* = \sum_{\theta \in \Theta_{gh}} \max\{f_a(\theta), f_h(\theta)\}. \tag{B1}
\]

Appendix C: Miscellaneous parameters

1. Sphericity

The sphericity \( \Psi \) of a given coordination geometry is defined by the ratio of the surface area of a sphere with the same volume \( V \) as the convex hull of that geometry to the surface area \( A \) of the boundary of its convex hull [29],

\[
\Psi = \frac{\pi^{1/3}(6V)^{2/3}}{A}. \tag{C1}
\]

2. Moment of inertia per neighbour

The moment of inertia per neighbour \( I/k \) of a particle with a given coordination geometry can be computed as follows:

1. Calculate the centroid \( c \) of the particle’s neighbourhood as the average position of its neighbours, \( c = \langle q \rangle \).
2. Determine the Euclidean distance \( \ell(q) \) of each neighbour \( q \) from the centroid, \( \ell(q) = ||q - c|| \).
3. Take the sum of the square of these distances to get the moment of inertia, \( I = \sum q \ell(q)^2 \).
4. Divide \( I \) by \( k \) to get the per-neighbour value.

References

1. D. Moroni, P.R. ten Wolde, P.G. Bolhuis, Phys. Rev. Lett. 94, 235703 (2005). https://doi.org/10.1103/PhysRevLett.94.235703
2. A.R. Natarajan, A. Van der Ven, Phys. Rev. Lett. 121, 255701 (2018). https://doi.org/10.1103/PhysRevLett.121.255701
3. P.N. Plessow, Phys. Chem. Chem. Phys. 22, 12939 (2020). https://doi.org/10.1039/d0cp01651a
4. Y.-C. Hu, H. Tanaka, Nat. Commun. 13, 4519 (2022). https://doi.org/10.1038/s41467-022-32241-z
5. H. Terrones, A.L. Mackay, J. Math. Chem. 15, 157 (1994). https://doi.org/10.1007/bf01277557
6. S.K. Kolli, A.R. Natarajan, J.C. Thomas, T.M. Pollock, A. Van der Ven, Phys. Rev. Mater. 4, 113604 (2020). https://doi.org/10.1103/physrevmaterials.4.113604
7. J.C. Thomas, A.R. Natarajan, A. Van der Ven, Npj Comput. Mater. 7, 164 (2021). https://doi.org/10.1038/s41524-021-00627-0
8. J. Çakırman, F. Parsch, G.D. Hibbard, J. Chem. Phys. 156, 091101 (2022). https://doi.org/10.1063/5.0079985
9. P.J. Steinhardt, D.R. Nelson, M. Ronchetti, Phys. Rev. B 28, 784 (1983). https://doi.org/10.1103/physrevb.28.784
10. L. Pauling, J. Am. Chem. Soc. 51, 1010 (1929). https://doi.org/10.1021/ja01379a006
11. R. Staub, S.N. Steinmann, J. Chem. Phys. 152, 024124 (2020). https://doi.org/10.1063/5.0079985
12. R.V.L. Hartley, Bell Syst. Tech. J. 7, 535 (1928). https://doi.org/10.1002/j.1538-7305.1928.tb01236.x
13. A. Stukowski, Modell. Simul. Mater. Sci. Eng. 20, 045021 (2012). https://doi.org/10.1088/0965-0393/20/4/045021
14. H. Tanaka, H. Tong, R. Shi, J. Russo, Nat. Rev. Phys. 1, 333 (2019). https://doi.org/10.1038/s42254-019-0053-3
15. L. Yang, S. Dacek, G. Ceder, Phys. Rev. B 90, 054102 (2014). https://doi.org/10.1103/physrevb.90.054102
16. R. Hundt, J.C. Schön, M. Jansen, J. Appl. Crystallogr. 39, 6 (2006). https://doi.org/10.1107/S0021889805032450
17. M. Searcóid, Metric spaces (Springer, 2006), p.2
18. M. Ester, H.-P. Kriegel, J. Sander, X. Xu, in Proc. of 2nd International Conference on Knowledge Discovery and Data Mining, (1996) pp.226–231
19. Z. Bar-Joseph, D.K. Gifford, T.S. Jaakkola, Bioinformatics 17, S22 (2001). https://doi.org/10.1093/bioinformatics/17.suppl_1.s22
20. P.H.A. Sneath, R.R. Sokal, Numerical taxonomy (Freeman, 1973)
21. T.F. Cox, M.A.A. Cox, Multidimensional scaling (Chapman & Hall, 1994)
22. S.L. Devadoss, J. O’Rourke, Discrete and computational geometry (Princeton University Press, 2011), pp.98–117
23. C.J. Bradley, A.P. Cracknell, *The mathematical theory of symmetry in solids: representation theory for point groups and space groups* (Oxford University Press, Oxford, 2010)
24. E.C. Bain, N. Dunkirk, Trans. AIME. **70**, 24 (1924)
25. W. Burgers, Physica **1**, 561–586 (1934)
26. R.J. Gillespie, Coord. Chem. Rev. **252**, 1315 (2008). https://doi.org/10.1016/j.ccr.2007.07.007
27. J.J. Thomson, Lond. Edinb. **7**, 237 (1904). https://doi.org/10.1080/14786440409463107
28. M. Atiyah, P. Sutcliffe, Milan J. Math. **71**, 33 (2003). https://doi.org/10.1007/s00032-003-0014-1
29. H. Wadell, J. Geol. **43**, 250 (1935). https://doi.org/10.1086/624298

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.