Continuous Chiral Distances for 2-Dimensional Lattices

Matthew J Bright, Andrew I Cooper, Vitaliy A Kurlin
Computer Science department and Materials Innovation Factory, University of Liverpool, UK
E-mail: mattb242@gmail.com, vitaliy.kurlin@gmail.com

Abstract

Chirality was traditionally considered a binary property of periodic lattices and crystals. However, the classes of 2-dimensional lattices modulo rigid motion form a continuous space, which was recently parametrized by three geographic-style coordinates. The four non-oblique Bravais classes of 2-dimensional lattices form low-dimensional singular subspaces in the full continuous space.

Now the deviations of a lattice from its higher symmetry neighbours can be continuously quantified by real-valued distances satisfying metric axioms. This paper analyses these and newer G-chiral distances for millions of 2-dimensional lattices that are extracted from publicly available databases of 2-dimensional structures and real materials in the Cambridge Structural Database and others.
Keywords: 2-dimensional lattice, rigid motion, isometry, continuous metric, chiral distance.

1 Introduction: continuous metrics for 2D lattices

The term *chirality* has for most of its history been used as a binary property of molecular structures. Two objects with the same chemical structure have one of two possible chiralities if, considered as a set of atoms, one can be mapped to the other via a reflection, but not by any rigid motion (a composition of translations and rotations). However, the connection between the physical, chemical, and biological properties of materials and the extent to which their structures deviate from mirror symmetry[1] has seen a growing interest in the development of materials where this deviation can be continuously measured and controlled[2].

A quantification of chirality for finite molecules based on continuous atomic positions was proposed by Osipov et al[3,4] while Zabrodsky et al[5,6,7] developed a computationally easier approach by considering the molecule as a discrete atomic centres. The general idea is to

define, for a fixed crystallographic group G, a minimum distance from a given configuration of atoms to a configuration that is symmetric under the action of G. An overall asymmetry of a molecule was then defined as the minimum of these distances over all groups G.

The requirement for continuous distances between *periodic crystals* is motivated by the growing number of simulated and experimental structures being published. The Cambridge Structural Database (CSD) contains over 1.2 million known structures[8], while Crystal Structure Prediction software can over-predict several million hypothetical crystals[9]. Recent continuous Pointwise Distance Distributions[10] (stronger than the Pair Distribution Function) detected unexpected duplicates[11] in the CSD and predicted a new material by analogy[12].

Almost all past work on the chirality of periodic crystals[13] has studied discrete versions of chirality based on symmetry groups, which are discontinuous under perturbations.

This paper completes in the non-trivial case of 2-dimensional lattices[14] and uses geographical-style mapping[15] to define continuous chiral distances that are not based on discrete symmetry groups of the lattice. A point in the Euclidean plane $\mathbb{R}^2$ is identified with a vector $v = (x, y)$, whose length is $|v| = \sqrt{x^2 + y^2}$ and starting point is at the origin of $\mathbb{R}^2$.

Definition 1.1 (lattice, rigid motion and isometry)**.** *For any basis of vectors $v_1, v_2 \in \mathbb{R}^2$, the* lattice $\Lambda = \{c_1 v_1 + c_2 v_2 \in \mathbb{R}^2 \mid c_1, c_2 \in \mathbb{Z}\}$ *is the set of all integer combinations of the basis vectors. Shifting Λ by a vector $p \in \mathbb{R}^2$ produces the lattice $\Lambda + p$, whose origin is shifted to the point p. The composition of such a shift with a rotation is a* rigid motion*. If we compose a rigid motion with a mirror reflection, we get an* isometry*, which is a map $f : \mathbb{R}^2 \to \mathbb{R}^2$ maintaining distances: $|f(p) - f(q)| = |p - q|$ for any points $p, q \in \mathbb{R}^2$.* ■

Since crystal structures are determined in a rigid form, there is no practical sense in distinguishing lattices that differ by a rigid motion. However, structures related by a reflection may or may not differ in their properties, and so this paper explores the differences between the rigid motion and isometry cases. Lattices are therefore considered as subsets of $\mathbb{R}^2$, as in Definition 1.1, so that reflections are geometrically distinct from rotations. Definition 2.1 will introduce a sign (binary chirality or orientation) of a lattice. Then we can study lattices modulo isometry and still distinguish them modulo rigid motion by their sign. Problem 1.2 formalises the requirements for a metric on the space of lattices modulo either rigid motion

or isometry (one may simply exchange one term for the other in the problem statement).

Problem 1.2 (continuous metric for lattices)**.** *Design a metric d satisfying the axioms:*

(a) coincidence *: $d(\Lambda,\Lambda')=0$* if and only if *lattices $\Lambda,\Lambda'\subset\mathbb{R}^2$ are related by rigid motion;*

(b) symmetry *: $d(\Lambda,\Lambda')=d(\Lambda',\Lambda)$ for any lattices $\Lambda,\Lambda'\subset\mathbb{R}^2$;*

(c) triangle inequality *: $d(\Lambda,\Lambda')+d(\Lambda',\Lambda'')\geq d(\Lambda,\Lambda'')$ for any lattices $\Lambda,\Lambda',\Lambda''\subset\mathbb{R}^2$.*

If lattices Λ,Λ' are given by Selling[16] or Delone[17] reduced bases, $d(\Lambda,\Lambda')$ should be computable in a constant time. The new requirement is the continuity *under perturbations of the lattice, for example, the* Hölder continuity *: there are constants $C,\alpha>0$ such that if basis vectors of lattices Λ,Λ' differ from each other by a small perturbation δ, then $d(\Lambda,\Lambda')\leq C\delta^{\alpha}$.*

The Bilbao Crystallographic Server computes the strain tensor to quantify the similarity of periodic lattices. This tensor measures the degree of distortion required to transform one unit cell into another[18], but fails the triangle axiom in Problem 1.2(c). This axiom is absolutely necessary to justify outputs of clustering[19], see an easy counterexample in appendix B.

The first axiom in Problem 1.2(a) already allows us to detect non-isometric lattices by checking if $d(\Lambda,\Lambda')=0$. For hundreds of years, crystallography attempted this detection by reducing a lattice basis to a canonical (or conventional) form. For example, Selling[16] and Delone[17] proposed reductions to an obtuse superbase, as named by Conway and Sloane[20]. Recall that the *scalar product* of any vectors $u=(a,b),v=(x,y)\in\mathbb{R}^2$ is $u\cdot v=ax+by$.

Definition 1.3 (obtuse superbase)**.** *For any basis v_1,v_2 of a lattice $\Lambda\subset\mathbb{R}^2$, define the extra vector $v_0=-v_1-v_2$. The unordered triple v_0,v_1,v_2 is called an* obtuse superbase[20] *if all $v_i\cdot v_j\leq 0$ or, equivalently, all pariwise angles between v_0,v_1,v_2 are at least 90° (non-acute). The unordered triple v_0,v_1,v_2 is called an* obtuse superbase *if all $v_i\cdot v_j\leq 0$ or, equivalently, all pairwise angles between v_0,v_1,v_2 are at least 90°. A lattice is* oblique *if $v_i\cdot v_j<0$ and $|v_i|\neq|v_j|$ for all $i,j\in\{0,1,2\}$, and otherwise* non-oblique *(mirror-symmetric).* ■

Since a pseudo-symmetric classification[21] of lattices is typically done to within some numerical tolerance (that is, a lattice is assigned a particular symmetry group if its parameters deviate by some defined small quantity from one with that symmetry group), we clarify that

all inequalities in the above and any subsequent definitions are *precise*. We aim to map equivalence classes of lattices up to rigid motion (or isometry) to a continuous metric space - if we classify two lattices that differ by a small perturbation as equivalent then all lattices collapse into the same equivalence class, since any lattice can be deformed to any other via a series of arbitrarily small perturbations, represented by a continuous path in the space.

Any lattice $\Lambda \subset \mathbb{R}^2$ has a unique obtuse superbase up to isometry, but not up to rigid motion[14]. The non-oblique lattice bases $(2,0),(0,1)$ and $(2,0),(0,-1)$ generate the same rectangular lattice but are related by mirror reflection with respect to the x-axis, not related by rigid motion. In $\mathbb{R}^3$, an obtuse superbase is not unique even up to isometry[22]. So Selling/Delone reductions cannot directly provide a metric satisfying the coincidence axiom in Problem 1.2(a). Niggli's reduced cell[23] has a more complicated definition and is considered unique in $\mathbb{R}^3$ but has been known since 1965 to be discontinuous under basis perturbations, see page 80 in the comprehensive report of Lawton[24] and the easier 2D case in Fig. 1.

In $\mathbb{R}^3$, Andrews and Bernstein used the Niggli and Selling/Delone reduction in numerous attempts to define a metric between lattices by examples instead of proofs. The DC7 similarity[25] fails the first metric axiom - see Example 6.5 in the recent work by Kurlin[22]. The $4! = 24$ permutations of 4 superbase vectors v_0, v_1, v_2, v_3 of a 3D lattice are insufficient to relate non-isometric Delone cells. For instance, $v_1 = (1,0,0), v_2 = (0,1,0), v_3 = (0,0,1), v_0 = (-1,-1,-1)$ and v_1, v_2 together with $v_3' = (-1,0,1), v_0' = (0,-1,-1)$ are obtuse superbases of the same cubic lattice with vector lengths $(|v_1|,|v_2|,|v_3|,|v_0|) = (1,1,1,\sqrt{3})$, while $(|v_1|,|v_2|,|v_3'|,|v_0'|) = (1,1,\sqrt{2},\sqrt{2})$, so the superbases are not isometric. Problem 1.2 was solved by Theorems 4.2, Corollary 4.6, and Theorem 7.4 in the mathematical paper[14]. The 3D analogue will be solved in our forthcoming work extending the Voronoi classification[22].

For each of the crystallographic point groups $G = D_2, D_4, D_6$, which are the symmetry groups of non-oblique or mirror-symmetric lattices in Definition 1.3, section 6 in the past work[14] introduces the real-valued chiral distances RC[G] and PC[G] as shortest distances from a lattice $\Lambda \subset \mathbb{R}^2$ to the subspace of higher-symmetry lattices that have the group G.

Section 2 describes the mathematical methods. Section 3 analyses histograms of values for the G-chiral distances of millions of 2D lattices extracted from the Cambridge Structural Database, demonstrating the computational simplicity of our approach. Section 4 visualises

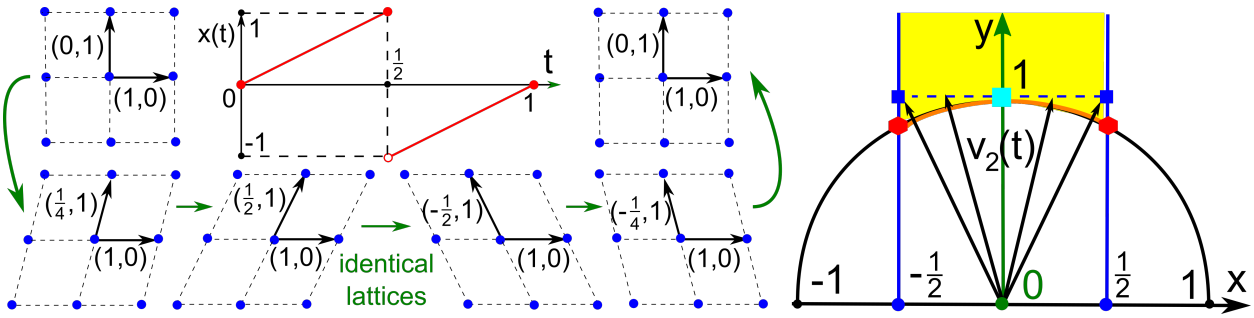

Figure 1: For the equivalence class of lattices considered up to rigid motion $\Lambda(t)$ generated by $v_1 = (1, 0)$, $v_2(t) = (t, 1)$ for $t \in [0, 1]$, a reduced basis with fixed v_1 (and v_2 that belongs to the yellow fundamental domain in the last image) such that $-\frac{1}{2} \leq v_1 \cdot v_2 \leq 0$ and $|v_2| \leq |v_1|$ discontinuously changes when the coordinate $x(t)$ of v_2 at $t = \frac{1}{2}$ drops down from $\frac{1}{2}$ to $-\frac{1}{2}$.

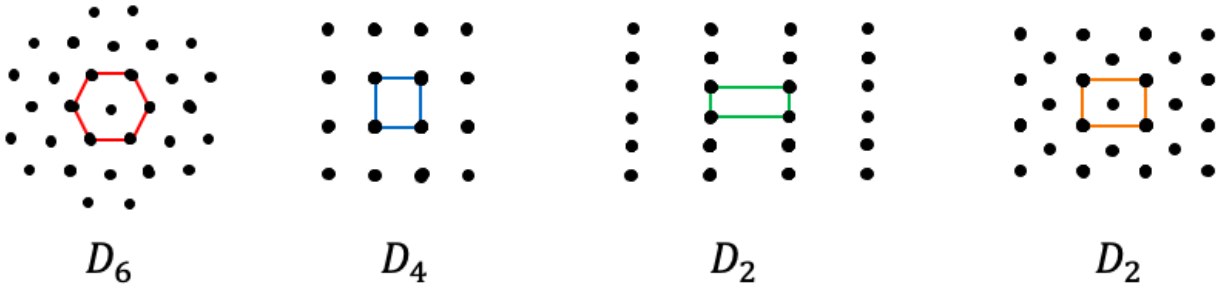

Figure 2: Any 2D lattice has one of four crystallographic point groups: oblique or generic (C_2), primitive or centred rectangular (D_2), square or tetragonal (D_4), hexagonal (D_6).

large databases of 2D materials on new continuous maps. Section 5 concludes with a discussion of future work. Two appendices to the paper contain a deeper analysis of lattices from the CSD, illustrating effects that are visible using continuous G-chiral distances which can be missed under a an approach based only on discrete symmetry groups, and the etra mathematical details and proofs for statements made in the main body of the paper.

2 Methods: lattice invariants and G-chiral distances

This section first recalls the concepts from sections 3-6 of the mathematical companion paper[14] and then introduces new invariants and G-chiral metrics in Definitions 2.5-2.7.

To understand the distinction between isometry and rigid motion, we introduce the *sign*

of a 2D lattice, which will help to define continuous signed chiral distances later on.

Definition 2.1 (sign of a lattice)**.** *Any 2D lattice* Λ *has a (unique up to isometry) obtuse superbase* v_0, v_1, v_2*. If any of the vectors have equal lengths, then* Λ *is mirror-symmetric and has* $\operatorname{sign}(\Lambda) = 0$*. Otherwise* v_1, v_2, v_0 *are uniquely ordered by their length and* $\operatorname{sign}(\Lambda) \in \{\pm 1\}$ *is defined as the sign of the determinant of the* 2×2 *matrix whose columns are* v_1, v_2*.*

The lattice $\Lambda(\frac{1}{4})$ in the bottom left picture of Fig. 1 has the obtuse superbase $v_1 = (1,0), v_2 = (-\frac{1}{4}, -1)$, $v_0 = (-\frac{3}{4}, 1)$ and $\operatorname{sign} \det \begin{pmatrix} 1 & -\frac{1}{4} \\ 0 & -1 \end{pmatrix} < 0$. The mirror image $\Lambda(\frac{3}{4})$ has the obtuse superbase $v_1 = (1,0), v_2 = (-\frac{1}{4}, 1)$, $v_0 = (-\frac{3}{4}, -1)$ and $\operatorname{sign} \det \begin{pmatrix} 1 & -\frac{1}{4} \\ 0 & 1 \end{pmatrix} > 0$.

Definition 2.2 (root invariants $\mathrm{RI}(\Lambda)$ and $\mathrm{RI}^o(\Lambda)$)**.** *Any 2D lattice* Λ *has an obtuse superbase* v_0, v_1, v_2*. The* root invariant $\mathrm{RI}(\Lambda)$ *is the ordered triple of root products* $r_{ij} = \sqrt{-v_i \cdot v_j}$ *for distinct indices* $i, j \in \{0, 1, 2\}$*, where* $r_{ij} = r_{ji}$*. If any vectors have equal lengths, they can be swapped without affecting* $\mathrm{RI}(\Lambda)$*. The* oriented root invariant $\mathrm{RI}^o(\Lambda)$ *is obtained from* $\mathrm{RI}(\Lambda)$ *by attaching* $\operatorname{sign}(\Lambda)$ *as a subscript, which can be skipped for brevity if* $\operatorname{sign}(\Lambda) = 0$*.*

Definition 1.3 guarantees that all scalar products $v_i \cdot v_j \leq 0$ are non-positive, so that all root products $r_{ij} = \sqrt{-v_i \cdot v_j}$ are well-defined. In Appendix B we prove that the ordering of vector lengths $|v_1| \leq |v_2| \leq |v_0|$ guarantees the ordering $r_{12} \leq r_{01} \leq r_{02}$ of root products.

$\mathrm{RI}(\Lambda)$ is an isometry *invariant* in the sense that any isometric lattices $\Lambda \cong \Lambda'$ have $\mathrm{RI}(\Lambda) = \mathrm{RI}(\Lambda')$. The unit cell area $A(\Lambda)$ is an isometry invariant of Λ but is *not complete* in the sense that many non-isometric lattices $\Lambda \ncong \Lambda'$ have $A(\Lambda) = A(\Lambda')$. Theorem 4.2 in previous work[14] proves that $\mathrm{RI}(\Lambda)$ is a complete isometry invariant, hence uniquely determines Λ up to isometry. Similarly, $\mathrm{RI}^o(\Lambda)$ is a complete invariant of lattices under rigid motion.

The crucial advantage of the root invariant $\mathrm{RI}(\Lambda)$ is its continuity under perturbations of a lattice basis. The discontinuity of reduced bases is caused by the identification of vertical

boundaries $x = \pm\frac{1}{2}$ of the fundamental domain in the last picture of Fig. 1. The root invariant $\mathrm{RI}(\Lambda)$ takes values in the cone $\mathrm{TC} = \{0 \le r_{12} \le r_{01} \le r_{02}\}$. The Bravais classes of square and hexagonal lattices occupy the disjoint boundary lines $r_{02} = r_{01} > 0 = r_{12}$ and $r_{02} = r_{01} = r_{12} > 0$, respectively. The Bravais class of primitive rectangular lattices covers the open triangular boundary $r_{01} > 0 = r_{12}$. The Bravais class of centred rectangular lattices is represented by two disjoint boundaries $r_{02} = r_{01} \neq r_{12} > 0$. Oblique lattices occupy the interior of the cone TC. The boundary line $r_{01} = r_{12} = 0$ is excluded in Fig. 3 (left), but there are no identifications on other boundaries that can cause discontinuities[14].

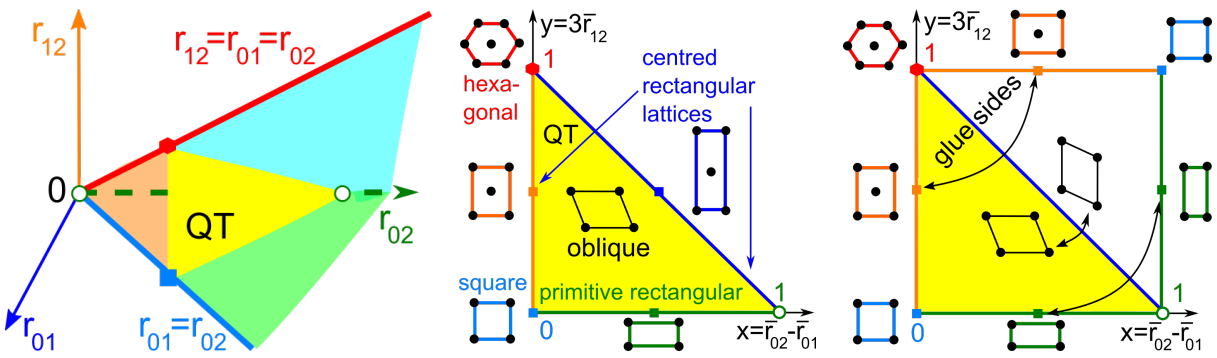

Figure 3: **Left**: all root invariants $\mathrm{RI}(\Lambda)$ live in the cone $\mathrm{TC} \subset \mathbb{R}^3$, which projects to the yellow triangle QT by uniform scaling of lattices $\Lambda \subset \mathbb{R}^2$. **Middle**: all projected invariants $\mathrm{PI}(\Lambda)$ live in the triangle QT parametrised by $x = \bar{r}_{02} - \bar{r}_{01} \in [0, 1)$ and $y = 3\bar{r}_{12} \in [0, 1]$. **Right**: mirror reflections $\Lambda^{\pm} \subset \mathbb{R}^2$ of any non-mirror-symmetric lattice can be represented by a pair of poits $(x, y) \leftrightarrow (1-y, 1-x)$ in the square QS symmetric in the diagonal $x+y = 1$.

The uniform scaling of a 2D lattice Λ by a factor $s > 0$ multiplies all root products r_{ij} by s, which motivates the normalisation of $\mathrm{RI}(\Lambda)$ to a simpler invariant. The most geometrically intuitive way to select a similarity class of lattices is to consider some hyperplane $r_{12} + r_{01} + r_{02} = \sigma$ for $\sigma \in \mathbb{R}$, intersecting the cone TC as shown in Fig. 3, allowing us to define both a numerical invariant of lattices up to similarity and a quantification of size as follows.

Definition 2.3 (projected invariants $\mathrm{PI}(\Lambda)$ and $\mathrm{PI}^o(\Lambda)$)**.** *Dividing all root products of* $\mathrm{RI}(\Lambda) = (r_{12}, r_{01}, r_{12})$ *by the size* $\sigma(\Lambda) = r_{12} + r_{01} + r_{02}$ *gives the ordered triple* $0 \le \bar{r}_{12} \le \bar{r}_{01} \le \bar{r}_{02}$ *such that* $\bar{r}_{12} + \bar{r}_{01} + \bar{r}_{02} = 1$. *Define the* projected invariant $\mathrm{PI}(\Lambda) = (x, y)$ *with* $x = \bar{r}_{02} - \bar{r}_{01}$ *and* $y = 3\bar{r}_{12}$ *takes values in the triangle* $\mathrm{QT} = \{0 \le x < 1, 0 \le y \le 1, x + y \le 1\}$. *The ori-*

ented projected invariant $\mathrm{PI}^o(\Lambda)$ *can be defined by attaching* $\mathrm{sign}(\Lambda)$ *as a subscript to* $\mathrm{PI}(\Lambda)$. *Equivalently, one can also map any point* $(x, y)_-$ *with a negative sign to* $(1-y, 1-x)$ *so that* $\mathrm{PI}^o(\Lambda)$ *takes values in the square* QS *whose sides are identified as in Fig. 3 (right).*

While the area of the unit cell of a lattice (that is, the absolute value of the determinant of its basis) is a more typical scaling quantity of Λ, the size $\sigma(\Lambda)$ has a number of advantages in this context. The size $\sigma(\Lambda)$ is easier to compute from the root invariant $\mathrm{RI}(\Lambda)$ as a sum of root products, while the area is the square root of a degree 4 polynomial in Lemma 4.1[14]. Also, $\sigma(\Lambda)$ retains the units of the original basis coordinates (typically Ångstroms) and is an upper bound on basis vector lengths in the obtuse superbase, see Lemma 7.4[14].

Fig. 3 (middle) shows all 2D lattices up to isometry and uniform scaling on the triangle QT whose interior is filled by oblique lattices. The boundary of QT represents the lower-dimensional subspaces of non-oblique 2D lattices that are achiral (mirror-symmetric): the horizontal edge of all primitive rectangular lattices (point group D_2), the vertical and diagonal (open) edges of all centred rectangular lattices (D_2), the right-angled vertex (0,0) of all square lattices (D_4) and the top vertex (0,1) of all hexagonal lattices (D_6). The excluded vertex $(1, 0)$, which maps from the intersection of the plane $r_{12} + r_{01} + r_{02} = 1$ with the excluded axis $r_{01} = 0$ in TC, represents the limit case when a minimal cell becomes infinitely long and thin. Any pair of lattices Λ_+, Λ_- of opposite signs that map to each other via a reflection are represented by a pair of points reflected in the diagonal $y = 1 - x$.

If mirror images of a lattice $\Lambda \subset \mathbb{R}^2$ are considered non-equivalent, the resulting space of 2D lattices modulo rigid motion and uniform scaling consists of two copies of the triangle QT that are glued along their boundaries. This space can be visualised as the square QS with identified edges in Fig. 3 (right) or a punctured sphere in Fig. 5 (middle).

The continuous spaces TC and QT allow many metrics to satisfy all axioms. The simplest Euclidean metric between root invariants RI satisfies the version of Problem 1.2 for isometry. The analytic formulae in Propositions 5.8-5.9[14] address the more complex case of oriented invariants. Any such metric easily quantifies a deviation from symmetry as a continuous distance from a lattice to a closest higher symmetry neighbour, hence measuring the minimum deformation required to move a given lattice into a more symmetric position. For

any 2D point group G, we call the resulting real value the *G-chiral distance*, with further qualification and distinct notation for different equivalences - rigid motion, isometry, and their composition with uniform scaling. Any G-chiral distance can be multiplied by the sign of a lattice, see Definition 2.1 to distinguish mirror images of 2D lattices.

Definition 2.4 (root and projected G-chiral distances $\mathrm{RC}_2[G](\Lambda)$ and $\mathrm{PC}_2[G](\Lambda)$)**.** *For a group $G \in \{D_2, D_4, D_6\}$ and a lattice $\Lambda \subset \mathbb{R}^2$, the* root G-chiral *distance* $\mathrm{RC}_2[G](\Lambda)$ *is the minimum Euclidean distance from* $\mathrm{RI}(\Lambda)$ *to the subspace of root invariants of all 2D lattices with the group G. The* projected G-chiral distance $\mathrm{PC}_2[G](\Lambda)$ *is similarly defined via the projected invariant* $\mathrm{PI}(\Lambda)$*, see Fig. 4. Any root or projected G-chiral distance can be multiplied by* $\mathrm{sign}(\Lambda)$ *to get the* signed distances*, which distinguish mirror images.*

The subscript 2 in $\mathrm{RC}_2, \mathrm{PC}_2$ refer to the Euclidean metric on the invariant spaces, see more general $\mathrm{RC}_q, \mathrm{PC}_q$ with a parameter $q \in [1, +\infty]$ in Definition 6.1[14]. The root chiral distances $\mathrm{RC}_2[G]$ have the same units as the original basis coordinates, usually Ångstroms. The projected chiral distances $\mathrm{PC}_2[G]$ are unitless for lattices modulo uniform scaling.

The D_4-chiral distance $\mathrm{RC}_2[D_4](\Lambda)$ is the Euclidean distance from $\mathrm{RI}(\Lambda)$ in the triangular cone TC to the boundary axis $\{r_{01} = r_{02} > 0 = r_{12}\}$. The hexagonal lattice Λ_6 with minimum inter-point distance 1 has an obtuse superbase $v_1 = (1, 0)$, $v_2 = (-\frac{1}{2}, \frac{\sqrt{3}}{2})$, $v_0 = (-\frac{1}{2}, -\frac{\sqrt{3}}{2})$, the root invariant $\mathrm{RI}(\Lambda_6) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ and D_4-chiral distance $\mathrm{RC}_2[D_4](\Lambda_6) = \frac{1}{\sqrt{2}}$ to a closest square lattice with $\mathrm{RI} = (0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ and basis $v_1 = (\frac{1}{\sqrt{2}}, 0)$, $v_2 = (0, \frac{1}{\sqrt{2}})$.

Mapping lattices to the unit square creates a simple and intuitive visualisation of lattice distributions. However, the resulting metrics are computationally awkward for lattices of opposite signs, since they still rely on exhaustive analysis of reflections across boundary subspaces. Identifying the sides of the square QS in Fig. 3 gives a punctured sphere, on which geographic coordinates lead to a new distance. The boundary of QT is a closed cycle of three edges and can be continuously mapped to the equatorial circle of the unit sphere S^2. There are many (indeed an infinite number of) ways to do this, so we map all mirror-symmetric lattices to the most obvious circle of symmetry, namely the equator of S^2.

The most geometrically intuitive way to complete a map $\mathrm{QS} \to S^2$ is to send the incentre P (centre of the inscribed circle) of QT to the north pole of S^2. So the further away a lattice

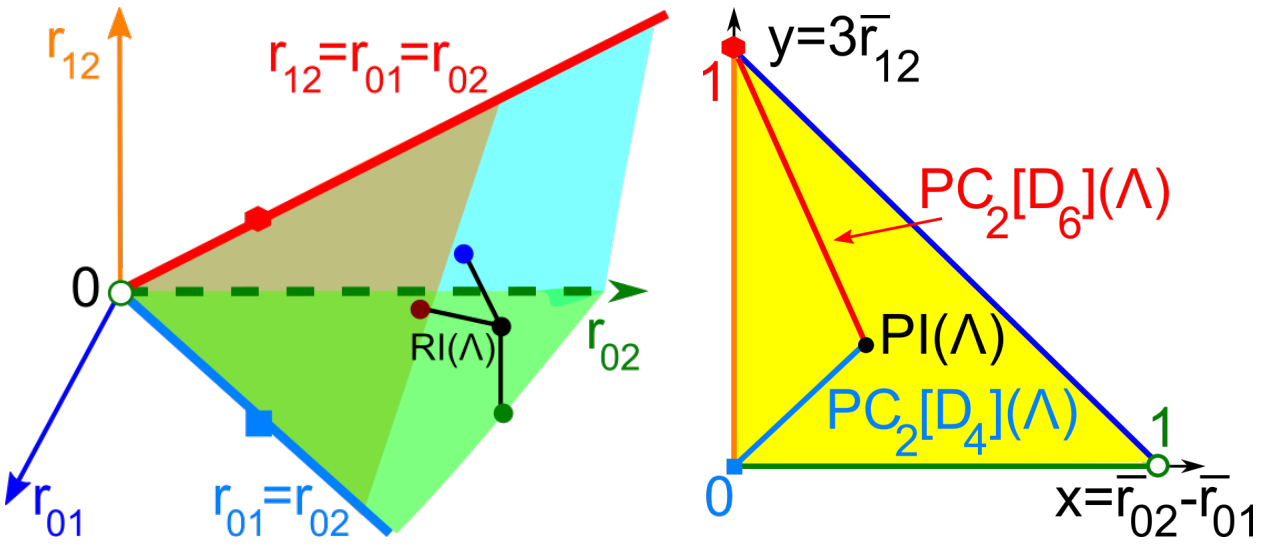

Figure 4: Visualisation of Definition 2.4. **Left**: $\mathrm{RC}_2[D_2](\Lambda)$ and $\mathrm{PC}_2[D_2](\Lambda)$ equal the minimum of three Euclidean distances from the points $\mathrm{RI}(\Lambda) \in \mathrm{TC}$ and $\mathrm{PI}(\Lambda) \in \mathrm{QT}$ to the boundary of the cone TC (along perpendiculars to the three boundary planes) and triangle QT, respectively. **Right**: $\mathrm{PC}_2[D_4](\Lambda), \mathrm{PC}_2[D_6](\Lambda)$ equal the Euclidean distances from $\mathrm{PI}(\Lambda)$ to the vertices $(0,0), (0,1) \in \mathrm{QT}$ representing all square and hexagonal lattices, respectively.

is from any mirror-symmetric lattice on the boundary of QT, the higher its latitude on S^2. We map QT to the northern hemisphere of QT linearly along radial rays emanating from P. The spherical map $\mathrm{SM} : \mathrm{QS} \to S^2$ symmetrically maps the second triangle of QS to the southern hemisphere. Since the vertex $(1,0)$ is excluded from QS, the *Greenwich meridian* is chosen as the image (under SM) of the green line through the incentre P and $(1,0)$, hence intersecting the vertical edge of QT at the *Greenwich point* $\mathrm{GP} = (0, \sqrt{2}-1)$.

Definition 2.5 (spherical projected invariant SPI)**.** *For any lattice* $\Lambda \subset \mathbb{R}^2$*, the* latitude $\varphi(\Lambda) \in [-90°, 90°]$ *is the angle measured from the equatorial plane* EP *(upwards in the positive direction) to the vector from the centre of* S^2 *to* $\mathrm{SM}(\mathrm{PI}(\Lambda))$*, see Figure 5. For a fixed orientation of* EP*, the* longitude $\mu(\Lambda) \in [-180°, 180°]$ *is the anti-clockwise angle from the Greenwich meridian to the vector from the centre of* S^2 *to the orthogonal projection of* $\mathrm{SM}(\mathrm{PI}(\Lambda))$ *to* EP*. The* spherical projected invariant *is* $\mathrm{SPI}(\Lambda) = (\mu(\Lambda), \varphi(\Lambda)) = \mathrm{SM}(\mathrm{PI}(\Lambda))$.

Proposition 5.2 of our crystallographic paper[15] gives analytic formulae for $\mathrm{SPI}(\Lambda)$. The advantage of the spherical map over QS is that all lattices of positive and negative signs live

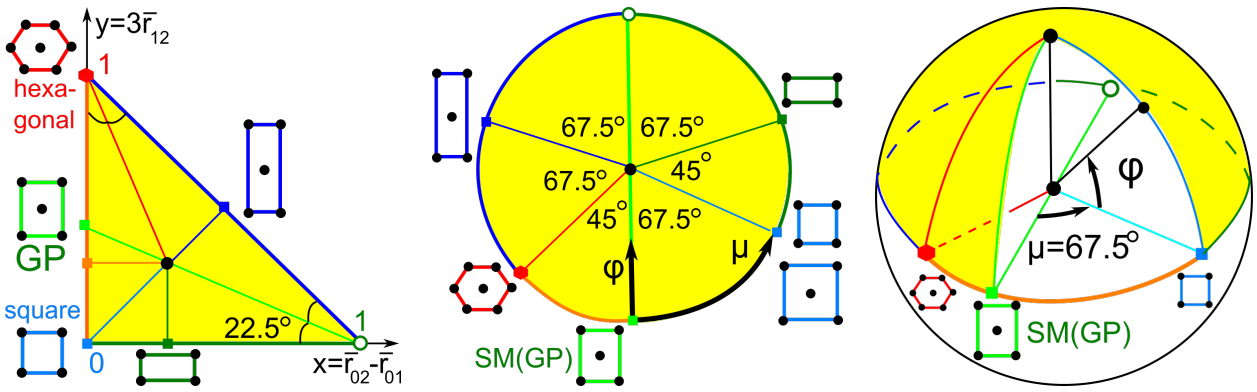

Figure 5: The incentre and boundary edges of the triangle QT on the left are mapped to the centre and boundary circle of the disk in the middle and further mapped to the north pole and equator of the unit sphere, whose southern hemisphere is obtained by a similar map of the second triangle, which represents all lattices with negative signs, in Fig. 3 (right).

in a common sphere without any need for the additional gluing operation of Definition 2.3. Hence chiral distances can be easily defined using known metrics on the sphere S^2. The *haversine function* $\mathrm{hav}(\theta) = \sin^2 \frac{\theta}{2}$ defines the angular distance along a great circle[26] on S^2 as $h((\mu_1, \varphi_1), (\mu_2, \varphi_2)) = 2 \arcsin \sqrt{\mathrm{hav}(\varphi_1 - \varphi_2) + \mathrm{hav}(\mu_1 - \mu_2) \cos \varphi_1 \cos \varphi_2} \in [0°, 180°]$. Tables of haversine distances have been used[26,27] since the 19th century for points on S^2. Since the sphere S^2 in Definition 2.5 has radius 1, the haversine distance is in degrees.

Definition 2.6 (spherical projected chiral distances)**.** *For any lattice* $\Lambda \subset \mathbb{R}^2$ *and a group* $G \in \{D_2, D_4, D_6\}$*, the* spherical projected G-chiral distance $\mathrm{SPC}[G](\Lambda)$ *is the minimum haversine distance from* $\mathrm{SPI}(\Lambda) \in S^2$ *to the subspace of lattices whose point group is* G*.*

Finally, we introduce invariants modulo rigid motion without uniform scaling by taking into account the size $\sigma(\Lambda) = r_{12} + r_{01} + r_{02}$. The resulting chiral distances will be measured in the units of the original basis coordinates (such as Ångstroms) instead of degrees.

Definition 2.7 (spherical root chiral distances)**.** *Let* Λ *be a lattice with root invariant* $\mathrm{RI}(\Lambda) = (r_{12}, r_{01}, r_{02})$*. The* spherical root invariant *is* $\mathrm{SRI}(\Lambda) = (\mu, \varphi, \sigma)$*, where* $\mathrm{SPI}(\Lambda) = (\mu, \varphi)$ *was introduced in Definition 2.5. Map the spherical coordinates* (μ, φ, σ) *to the standard Euclidean coordinates* $(x, y, z) = (\sigma \cos\varphi \cos\mu, \sigma \cos\varphi \sin\mu, \sigma \sin\varphi)$*. For any lattices* $\Lambda_1, \Lambda_2 \subset \mathbb{R}^2$*, the* spherical root metric $\mathrm{SRM}(\Lambda_1, \Lambda_2)$ *is the Euclidean distance between the*

points $(x, y, z) \in \mathbb{R}^3$ *obtained from* $\mathrm{SRI}(\Lambda_1), \mathrm{SRI}(\Lambda_2)$ *as above. For a group* $G \in \{D_2, D_4, D_6\}$ *and any lattice* $\Lambda \subset \mathbb{R}^2$*, the* spherical root G-chiral distance $\mathrm{SRC}[G](\Lambda)$ *is the minimum distance* $\mathrm{SRM}(\Lambda, \Lambda')$ *for a lattice* Λ' *that has* $\sigma(\Lambda') = \sigma(\Lambda)$ *and the crystallographic group* G*.*

Proposition 2.8 (proved in appendix B)**.** *For a lattice* Λ *with* $\mathrm{SRI}(\Lambda) = (\mu, \varphi, \sigma)$*, the distances from Definitions 2.5 and 2.7 can be computed as follows:* $\mathrm{SPC}[D_2](\Lambda) = |\varphi|$,
$\mathrm{SPC}[D_4](\Lambda) = 2 \arcsin \sqrt{\mathrm{hav}(\varphi - 67.5^\circ) + \mathrm{hav}(\mu) \cos\varphi \cos 67.5^\circ}$*, where* $\mathrm{hav}(\theta) = \sin^2 \frac{\theta}{2}$,
$\mathrm{SPC}[D_6](\Lambda) = 2 \arcsin \sqrt{\mathrm{hav}(\varphi + 45^\circ) + \mathrm{hav}(\mu) \cos\varphi \cos 45^\circ}$*, and* $\mathrm{SRC}[D_2](\Lambda) = 2\sigma |\sin \frac{\varphi}{2}|$,
$\mathrm{SRC}[D_4](\Lambda) = \sigma \sqrt{(\cos\varphi \cos\mu - \cos 67.5^\circ)^2 + (\cos\varphi \sin\mu - \sin 67.5^\circ)^2 + \sin^2\varphi}$,
$\mathrm{SRC}[D_6](\Lambda) = \sigma \sqrt{(\cos\varphi \cos\mu - \cos 67.5^\circ)^2 + (\cos\varphi \sin\mu + \sin 45^\circ)^2 + \sin^2\varphi}$.

All the above defined distances arise from the mapping illustrated in Figure 5, which is chosen for its geometrically intuitive nature. An infinite number of other mappings exist, and could be considered as a composition of our chosen mapping with some arbitrary homeomorphism acting on the punctured sphere. More suitable metrics may arise on such surfaces, or alternatively, the distorting effect of the chosen homeomorphism could be computed.

3 Results: G-chiral distances for 2D lattices of crystals

This section illustrates the ease with which the chiral distances can be computed for very large datasets by visualising histograms of G-chiral distances from section 2 for millions of 2D lattices extracted from all 870+ thousand crystals in the Cambridge Structural Database[8] (CSD), which have full lattice data. Any such crystal is deposited with a lattice basis of vectors v_1, v_2, v_3, which are ordered by length. We took three 2D lattices generated by three pairs of basis vectors $\{v_1, v_2\}, \{v_2, v_3\}, \{v_1, v_3\}$. For consistency, we used the same indices as in the CSD, though a re-ordering of vectors can change the sign of the extracted 2D lattices, see Definition 2.1. This choice of basis ordering explains slight dissymmetries in Fig. 6-17. All plots were produced on a standard laptop in a few minutes by the code at `https://github.com/MattB-242/Lattice_Invariance`.

Since any non-oblique (mirror-symmetric) 2D lattice has all D_2-chiral distance equal to zero, we first focus on oblique (generic) lattices. Their total number is 1177678, which is

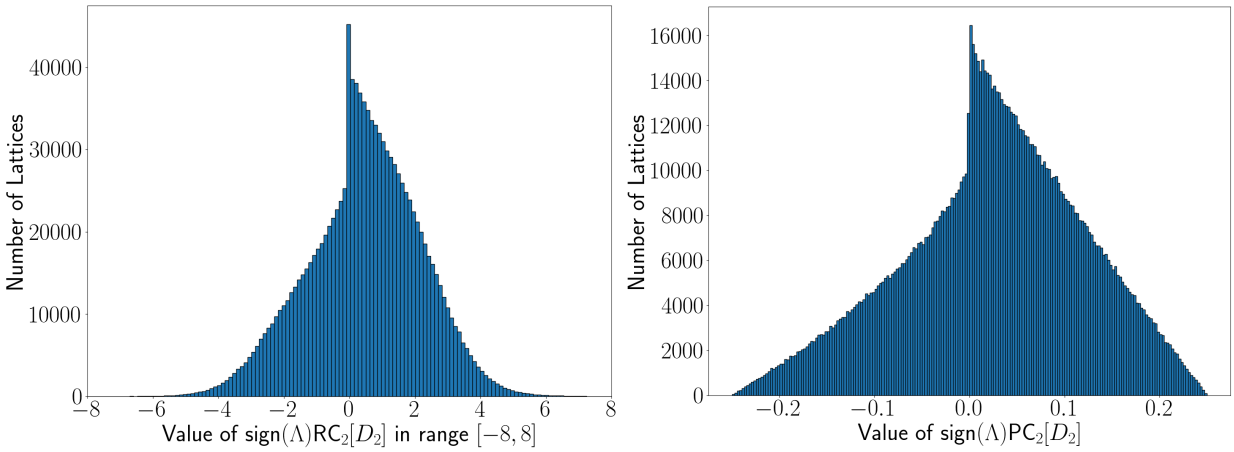

Figure 6: Signed D_2-chiral distances of all oblique 2D lattices found in the CSD, see Definition 2.4 **Left**: $\text{sign}(\Lambda)\text{RC}_2[D_2](\Lambda)$ in Ångstroms. **Right**: $\text{sign}(\Lambda)\text{PC}_2[D_2](\Lambda)$ is unitless.

about 45% of all 2D lattices found in the CSD. Fig. 6-8 show the histograms of the signed distances $\text{sign}(\Lambda)\text{RC}_2[G](\Lambda)$ and $\text{sign}(\Lambda)\text{PC}_2[G](\Lambda)$ for the point groups $G \in \{D_2, D_4, D_6\}$.

Fig. 6 quantifies the continuous tendency towards non-oblique lattices, which have $\text{RC}_2[D_2] = 0 = \text{PC}_2[D_2]$. The mathematical paper[14] proves the continuity of all G-chiral distances under perturbations of a lattice basis (Proposition 7.10) and the upper bound $\text{PC}_2[D_2](\Lambda) \leq \dfrac{1}{2+\sqrt{2}} \approx 0.29$ (Proposition 6.6c), which is respected by the distances in Fig. 6 (right). Also notable is a preponderance of lattices with positive sign. This is an artefact of our choice of lattice extraction - given cells with lengths a, b, c and angles α, β, γ, we generate lattices using the parameters $(a, b, \gamma), (b, c\alpha)$ and (a, c, β). The use of this convention interacts with consistent ordering of lattice length entries in the CSD to give rise to the imbalance.

Though there is no theoretical upper bound on $\text{RC}_2[D_2]$, it is notable that the vast majority of lattices (99.9%) occupy a fairly narrow range of signed distances between ± 8.

In Fig. 6-17, the histograms of all root G-chiral distances have the bin size $0.1\mathring{A}$, while the histograms of all projected G-chiral distances have the bin size of 0.01 (unitless). There are a number of structural points to note in Figures 7 and 8.

Figure 9 (centre) previously showed that if we plot the number of $2D$ lattices extracted from the CSD in the QT, we observe a strong preference for higher symmetry structures (on the boundary) and a concentration of lattices towards the point $(0, 1)$ representing the

hexagonal lattice, while the number decreases towards the point $(1, 0)$ representing infinitely long, thin cells[15]. Figure 9 (left) illustrates a geometric reason for this by showing how chiral distances change for the centred rectangular lattice Λ_θ with basis vectors of length 1 as its angle θ varies in $[90°, 120°]$. Close to 90°, $\mathrm{PC}_2[D_4](\Lambda_\theta)$ increases rapidly with a small change in angle - the rate of increase slows as we move closer to $\theta = 120°$. The plot thus concentrates lattices closer to the point $(0, 1)$ representing the hexagonal lattice, explaining the apparent rarity of lattices with low D_4-chiral distances compared to the plot for $\mathrm{PC}_2[D_6]$.

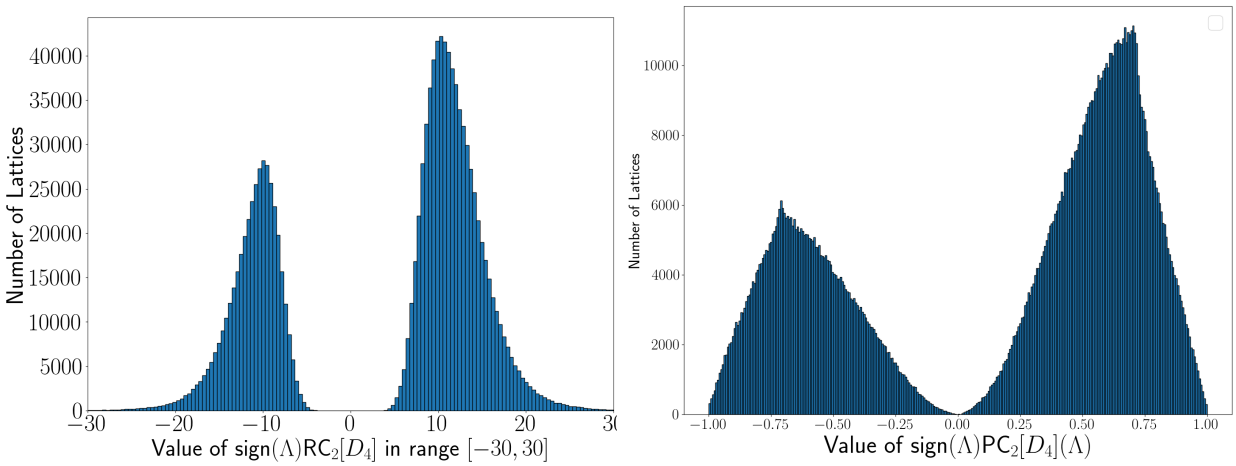

Figure 7: Signed D_4-chiral distances of 1,177,678 oblique 2D lattices in the CSD, see Definition 2.4 **Left**: $\mathrm{sign}(\Lambda)\mathrm{RC}_2[D_4](\Lambda)$ in Ångstroms. **Right**: $\mathrm{sign}(\Lambda)\mathrm{PC}_2[D_4](\Lambda)$ is unitless.

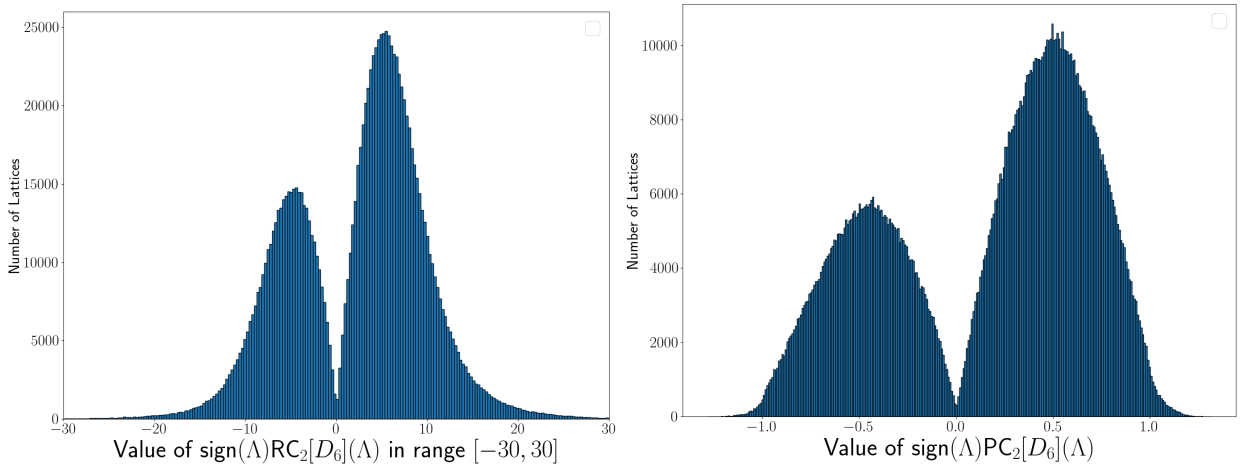

Figure 8: Signed D_6-chiral distances of 1,177,678 oblique 2D lattices in the CSD, see Definition 2.4. **Left**: $\mathrm{sign}(\Lambda)\mathrm{RC}_2[D_6](\Lambda)$ in Ångstroms. **Right**: $\mathrm{sign}(\Lambda)\mathrm{PC}_2[D_6](\Lambda)$ is unitless.

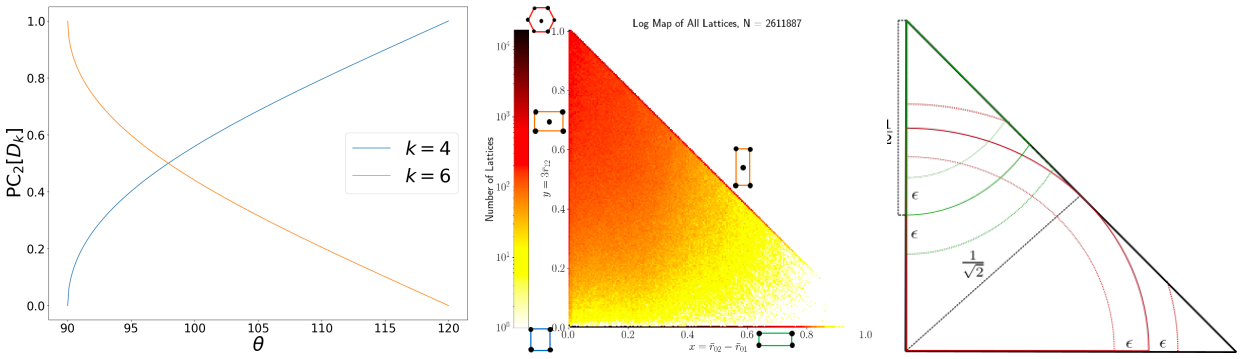

Figure 9: **Left**: values of $\mathrm{PC}_2[D_4]$ and $\mathrm{PC}_2[D_6]$ for a lattice with parameters $a = 1, b = 1$ and angle $\theta \in [90°, 120°]$ **Centre**: heat map of all 2D lattices in the quotient triangle, extracted from crystals in the CSD. **Right**: intersection of QT with red and green circles centred at the origin $(0, 0)$ and the vertex $(0, 1)$, respectively, of the radii $r_i = \frac{1}{2}$ and $r_i \pm \epsilon$ for $i = 1, 2$.

To explain the peaks of $\mathrm{PC}_2[D_4](\Lambda) \approx \frac{1}{\sqrt{2}}$ and $\mathrm{PC}_2[D_6] \approx \frac{1}{2}$, we observe the intersection of circles at the origin and the points $(0, 0)$ with the QT - see Figure 9 (right). We would expect the frequency of a particular chiral distance value to be a function of both the length of the intersection of the circle whose radius corresponds to that value and the density, in terms of CSD-derived 2D lattices, of the region through which the circle passes. The intersection of the circle at the origin of radius $\frac{1}{\sqrt{2}}$ with the QT gives a curve of maximal length. Thus a larger number of lattices in the plot in the centre will intersect with this curve.

The maximal length circle of radius 1 centred at the point $(1, 0)$ would pass mostly through a low-density area - as its radius decreases the density of lattices increases but the length of the intersection decreases - an optimum is evidently reached at radius $\frac{1}{2}$.

Both peaks are strengthened due to the presence of a large number of experimental determinations of Oxalic Acid (a standard test molecule) deposited in the CSD. The lattice Λ_{OX} derived from the single pair of non-orthogonal vectors in this primitive monoclinic lattice has parameters $a = 6.1143$Å, $b = 12.0109$Å, $\gamma = 106.1°$. Computation of chiral distances does indeed give the expected values $\mathrm{PC}_2[D_4](\Lambda_{OX}) \approx \frac{1}{\sqrt{2}}$ and $\mathrm{PC}_2[D_6](\Lambda_{OX}) \approx \frac{1}{2}$.

4 Results: continuous maps of monolayer structures

A more chemically practical application of chiral distances is for 2D structures. This has been an area of great interest in chemistry since such materials are predicted to have many useful physical properties[28]. While the lattice parameters for 2D monolayers are very often close to those of high symmetry lattices (square or hexagonal), there is a growing interest in stable 2D structures with more generic lattice geometries [29]. Our chiral distances allow a more formal definition of this problem - we wish to find materials whose lattices have a high chiral distance, and whose projected invariants will thus occupy the interior of the QT.

2DMatPedia [30] is one of the largest open-source databases of such materials available - currently containing $6,351$ crystal structures that have the potential to form monolayers. Of these, two were retrieved from existing literature, $2,940$ were found through a *layer detection* approach (referred to as the 'top-down' process), in which separable 2-dimensional features were detected from geometric data [31]. The remaining structures were then generated from this list by the *substitution* of atoms in the same group - the 'bottom-up' discovery process. Potential physical properties of the 2D structures are simulated using DFT calculations.

A natural first question is whether the database can tell us anything about the feasibility of synthesising 2D structures with oblique lattices. The database calculates two key properties which determine this potential - the *decomposition energy*, which is the energy required to split a structure into its most stable components, and the *exfoliation energy*, which is the average energy per atom required to separate the modelled layer from its parent material. The former should be high, and the latter low, to guarantee stability - typically $0.2eV$ is considered an acceptable upper bound for exfoliation energy in 2D structures[32].

Figure 10 shows the positions of the oriented projected invariants $\mathrm{PI}^o(\Lambda)$, see QS in Fig. 3, for 2D lattices of all structures in 2DMatPedia. It is immediately clear that strongly oblique 2D lattices are relatively rare in this dataset. The majority of 2D lattices are non-oblique, with very few occupying the interior of the square: 66% of lattices in the database have sign 0, and 75% have extremely small D_2-chiral distances $\mathrm{RC}_2[D_2]$ less than $3 \times 10^{-8}\mathring{A}$.

In Figures 11 and 12, we see the relationship between overall G-chiral distance (using the $\mathrm{RC}_2[D_2]$ metric) and both indicators of the feasibility of potential $2D$ materials. The

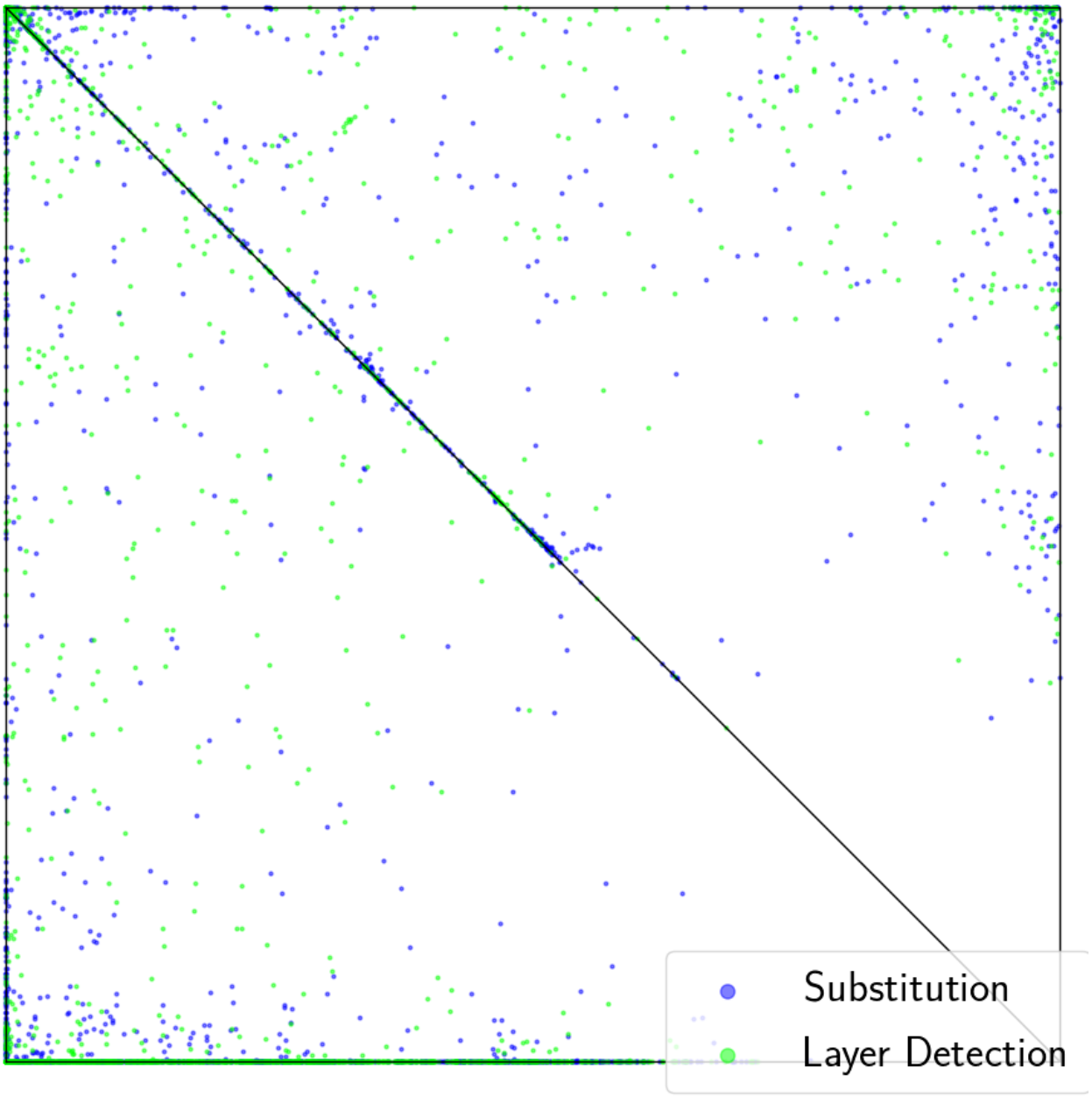

Figure 10: Invariants $\mathrm{PI}^o(\Lambda)$ in the square QS for 2D lattices of 6,351 monolayer structures[30] isolated from 3D crystals by layer detection or generated by atomic substitution.

most obvious thing to note is that the highest decomposition energies and lowest exfoliation energies occur with materials whose layers are non-oblique. It is also notable that the range of chiral distances overall is within $[-3, 3]$ - very narrow compared to the full CSD analysis.

Strongly oblique lattices are not structurally preferred or strongly stable in candidate 2D materials. It is also interesting to note that on the whole atomic substitution gives rise to

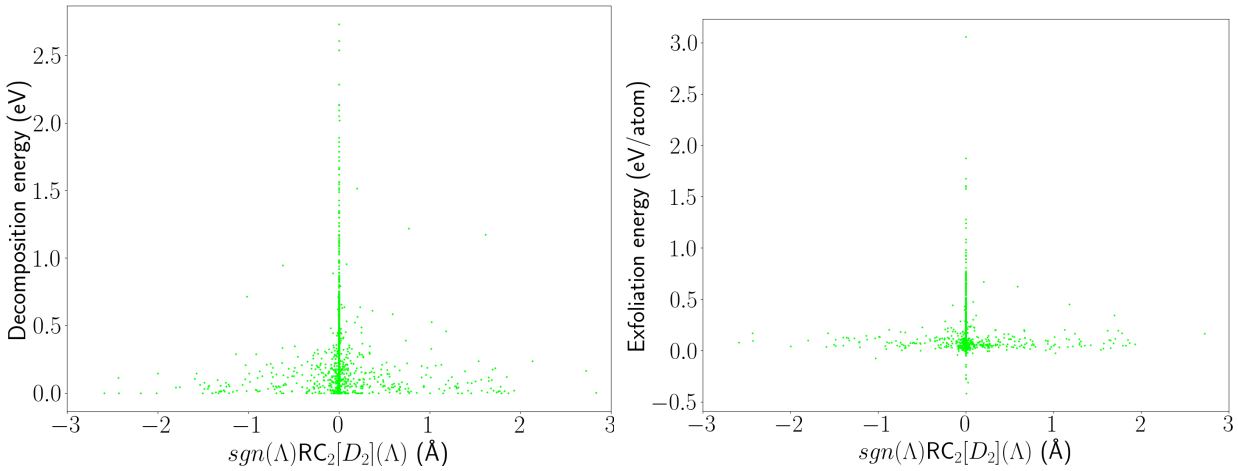

Figure 11: Scatter plot of physical properties of materials discovered by layer detection in 2DMatPedia vs sign(Λ)RC[D_2](Λ). **Left**: Decomposition energy in meV. **Right**: Exfoliation energy in meV per atom.

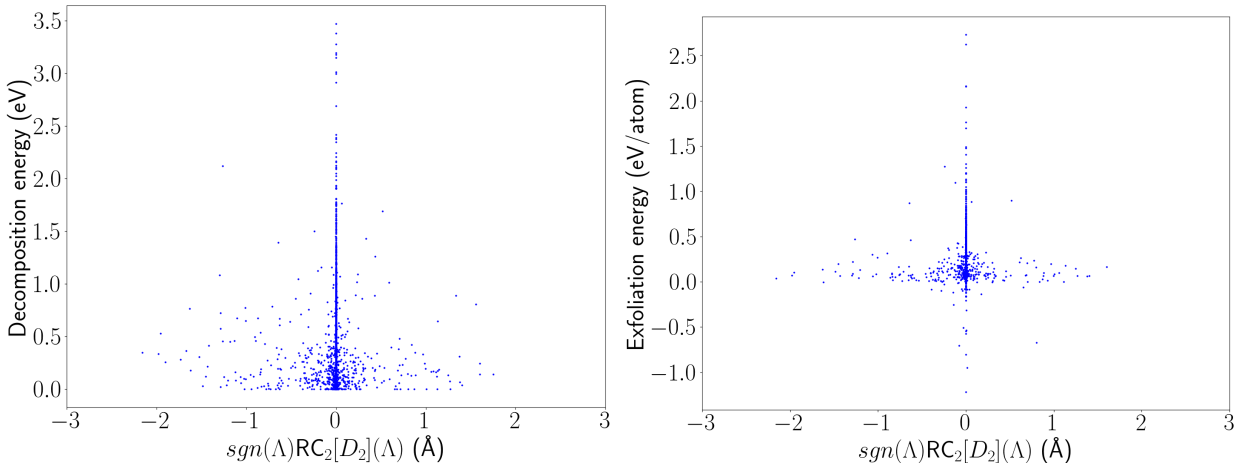

Figure 12: Scatter plot of physical properties of materials discovered by atomic substitution in 2DMatPedia vs sign(Λ)RC[D_2](Λ). **Left**: Decomposition energy in meV. **Right**: Exfoliation energy in meV per atom.

structures with lower chiral distances - lying in the range $[-2.1, 2.0]$ while structures found by identifying layers geometrically lie in the wider range of $[-2.5, 2.8]$.

In Figure 13 we try to isolate high G-chiral distance 2D structures that may feasibly be synthesised by plotting both exfoliation and decomposition energy on the same 2D plot. The right hand plot shows only those molecules with a D_2-chiral distance above 1.5, with three

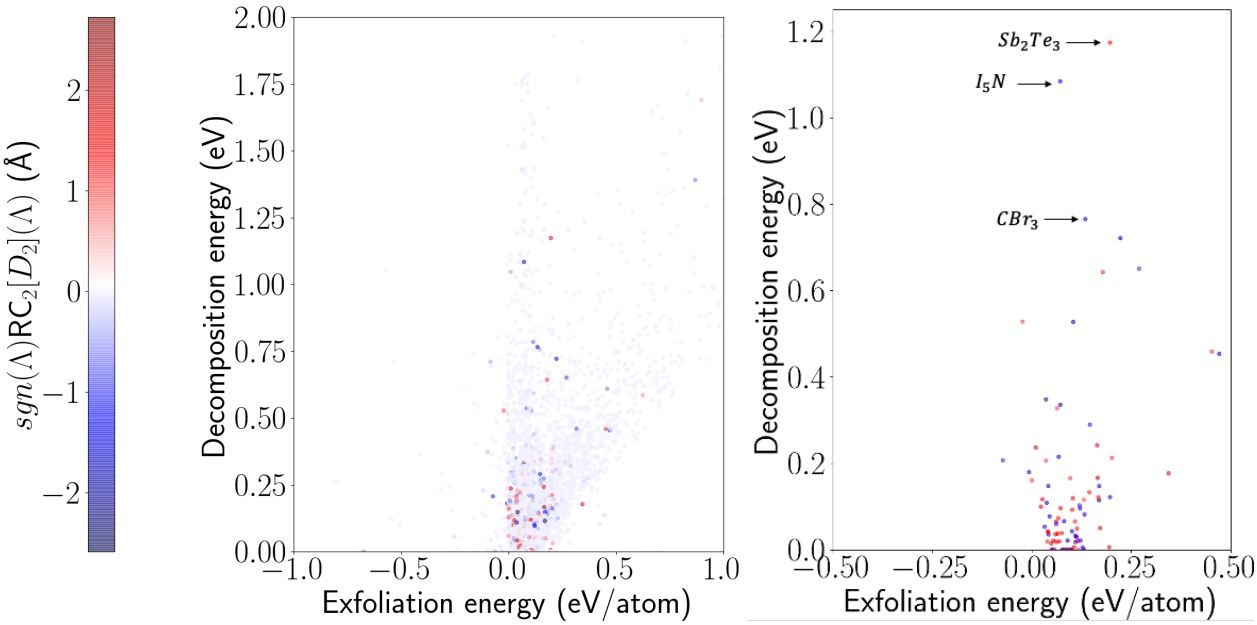

Figure 13: $2D$ scatter plot of exfoliation vs.decomposition energy, with G-chiral distance indicated by colour **Left**: All structures in 2DMatPedia. **Right**: Structures where $\mathrm{RC}_2[D_2] \geq 1$

structures labelled that have low exfoliation and relatively high decomposition energy.

Of these three monolayer structures, only one, antimony telluride (Sb_2Te_3) has been reported in the literature. In its monolayer form, the structure of Sb_2Te_3 has been found to exhibit reversible state changes which switch its electrical resistance from low to high values[33]. There are no reports in the literature on any anisotropic properties it may possess due to its highly oblique 2D lattice. The other two structures I_5N and CBr_3, which have been generated by atomic substitution, were not reported in the literature.

While 2DMatPedia is the largest open database, the 2D materials database[34] is smaller, containing only 1728 structures, some of which overlap with 2DMatPedia. It is of interest since of these structures, 183 were additionally 'relaxed': further DFT computations simulate the likely final 2D structures once isolated from the parent crystal. Our main analysis here investigates the effect of such relaxation on resulting G-chiral distances.

Fig. 14 and 15 show the invariants PI(Λ) of 1726 original 2D lattices and 183 'relaxed' structures in the square QS with coordinates $(x, y) \in [0, 1]$, see Definition 2.3. The elemental structures (monolayers of a single element) are highlighted in red, the MX2 structures[35],

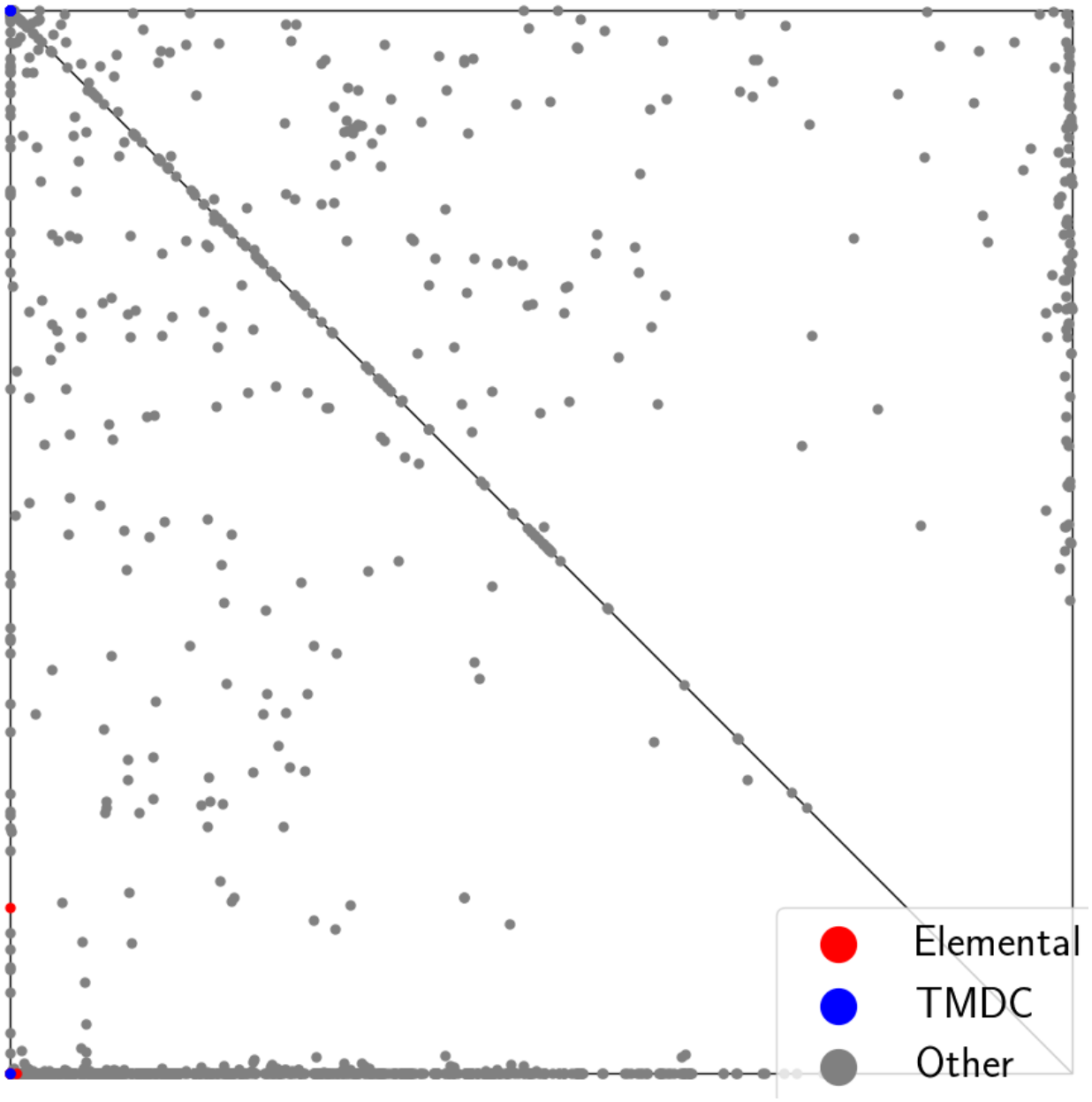

Figure 14: 2D lattices of 1726 monolayer structures[34] isolated from 3D crystals and shown by the invariants PI(Λ) in the square QS, see Definition 2.3 and labels of structures in Fig. 15.

where M is a metal and X is a halogen, are in green, and the transition metal dichalcogenide monolayers[36] (TMDC) are in blue. For illustration some of these have been labelled, although since several structures occupy higher symmetry lattice points at the vertics we have not labelled all of them. In Fig. 15, the green dot at the top left vertex $(0, 1)$ of the square QS indicates the hexagonal crystal of Nb3Br8, whose monolayer form was recently discovered

to have the long-sought-after property of acting as a superconducting diode[37].

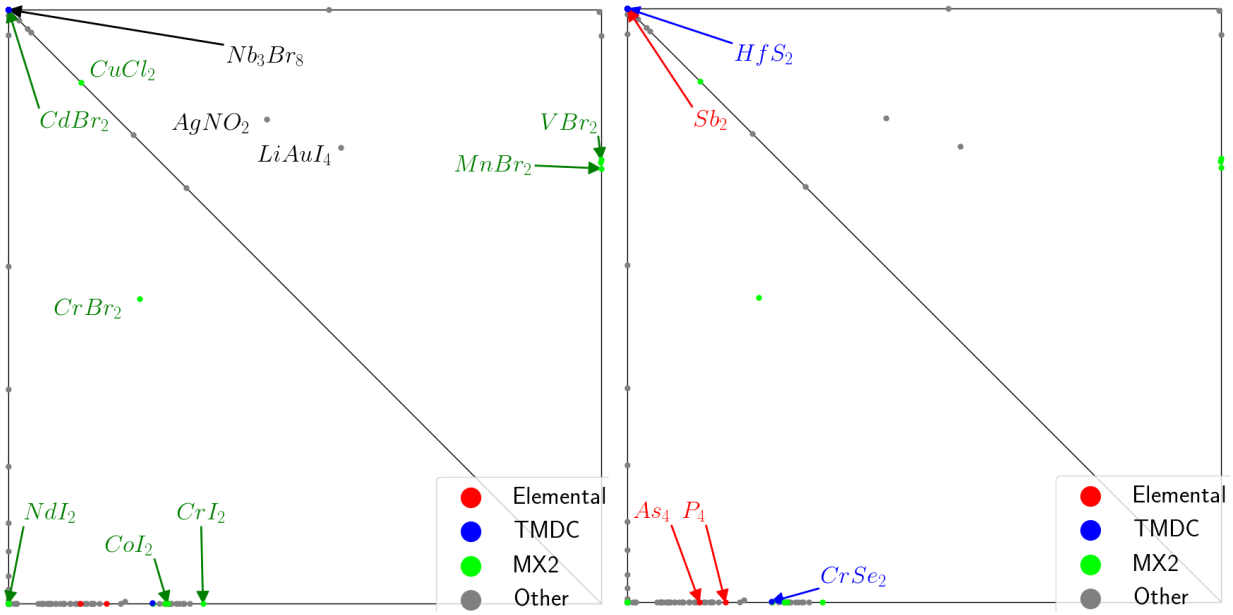

Figure 15: 183 'relaxed' by DFT 2D lattices from 1726 monolayer structures[34] in Fig. 14. **Left**: structures with oblique lattices and molecules of type MX2 are labelled in green. **Right**: elemental crystals and molecules of type TMDC are labelled in red and blue.

The key observation in this analysis is that while candidate 2D structures in their parent crystal can be both oblique and non-oblique, nearly all such structures simulated in isolation become non-oblique. This suggests that candidate materials with even larger D_2-chiral distances such as those selected above from 2DMatpedia may, when isolated, revert to a mirror-symmetric state. We have labelled the three chiral molecules that retain non-zero D_2-chiral distances with the formula of their parent structure. AgNO2 has in fact been shown to form chiral monolayers[37], but we are not aware of any publications specifically concerning the other two structures.

5 Conclusions and discussion of novelty and future work

This paper visualises new G-chiral distances for millions of 2D lattices extracted from real crystals in the Cambridge Structural Database, and for large databases of 2D monolayer structures. The key novelty is the continuity of these distances, which was proved under

any perturbations of lattice bases[14]. The resulting histograms and maps with meaningful coordinates in Ångstroms (or unitless under uniform scaling) reveal for the first time how 2D lattices in real crystals and monolayer structures populate the underlying continuous space of all possible 2D lattices. Though a strong preference for higher symmetry was expected, the maps in Fig. 14 and 15 also include many oblique 2D lattices. These continuous maps offer the opportunity to seek new materials based on the *novelty* of their structure - that is, to discover materials with potentially new properties by targeting a search specifically in regions where existing structures are sparse.

Even the high-level definitions and statements in section 2 demonstrate that the 2D case is far from trivial and should be fully understood before attacking the 3D case. Problem 1.2 makes sense for any dimension $n > 1$, see recent progress[22,38] for $n = 3$. Illustrating the impact of the distinction between isometry and rigid motion on the resulting geometric analysis is particularly important since there can be important functional differences between 3D structures which are mirror images of each other.

Previous work[15] concluded with a vision for a new branch of *continuous crystallography* that studies all periodic structures within a common continuous space. For general periodic crystals, this Crystal Isometry Space is parametrised by complete isometry invariants that are descriptors without false negatives and without false positives, hence playing the role of a DNA-style code or a materials genome to unambiguously identify any periodic structure.

We now have a hierarchy of continuous isometry invariants from simple and ultra-fast[10,11,39] to slower but provably complete invariants[40,41,42]. *Inverse design* aims to make complete invariants invertible, so that the space of all materials can be explored by trying new invariant values, which all give rise to 3D periodic structures - having found areas with low density (and thus high novelty) the proposed structure can be reconstructed directly from its invariant. The ultimate goal is then to describe a much smaller subspace of invariants whose crystals can be physically synthesised. By showing that chiral distances take a relatively narrow range of values for the space of real crystals and that certain values are more likely for structures with different chemical properties, the analysis of 2D lattices discussed in this paper represents a step towards the inverse design of materials.

Acknowledgements

We are very grateful to Marjorie Senechal and Nikolai Dobilin for fruitful discussions on this work, and to Feng Yuan Ping for their help in navigating the 2DMatPedia database. This research was funded in part by the following grants: EPSRC New Horizons "Inverse design of periodic crystals" (EP/X018474/1) and the Royal Academy of Engineering fellowship (IF2122/186) of the last author.

References

1. W. Ma, L. Xu, A. F. de Moura, X. Wu, H. Kuang, C. Xu, and N. A. Kotov, Chemical Reviews **117**, 8041 (2017).

2. P. Kumar, T. Vo, M. Cha, A. Visheratina, J.-Y. Kim, W. Xu, J. Schwartz, A. Simon, D. Katz, V. Nicu, et al., Nature **615**, 418 (2023).

3. M. A. Osipov, B. T. Pickup, and D. A. Dunmur, Molecular Physics **84**, 1193 (1995).

4. G. Millar, N. W. *, and K. M. *, Molecular Physics **103**, 2769 (2005).

5. H. Zabrodsky, S. Peleg, and D. Avnir, J. Am. Chem. Soc **114**, 7843 (1992).

6. H. Zabrodsky and D. Avnir, J. Am. Chem. Soc **117**, 462 (1995).

7. I. Tuvi-Arad, G. Alon, and D. Avnir, *Cosym*, `http://csm.ouproj.org.il`, accessed: 2022-05-26.

8. S. C. Ward and G. Sadiq, CrystEngComm **22**, 7143 (2020).

9. A. Pulido, L. Chen, T. Kaczorowski, D. Holden, M. Little, S. Chong, B. Slater, D. McMahon, B. Bonillo, C. Stackhouse, et al., Nature **543**, 657 (2017).

10. D. Widdowson, M. Mosca, A. Pulido, V. Kurlin, and A. Cooper, MATCH Communications in Mathematical and in Computer Chemistry **87**, 529 (2022).

11. D. Widdowson and V. Kurlin, Advances in Neural Information Processing Systems (Proceedings of NeurIPS 2022) **35** (2022).

12. Q. Zhu, D. W. Jay Johal, Z. Pang, B. Li, C. Kane, V. Kurlin, G. Day, M. Little, and A. Cooper, Journal of the American Chemical Society **144**, 9893–9901 (2022).

13. H. D. Flack, Helvetica Chimica Acta **86**, 905 (2003).

14. V. A. Kurlin, Foundations of Computational Mathematics (2022).

15. M. Bright, A. I. Cooper, and V. Kurlin, Acta Crystallographica Section A **79** (2023).

16. E. Selling, Journal für die reine und angewandte Mathematik **77**, 143 (1874).

17. B.N.Delone, N. Padurov, and A. Aleksandrov, *Mathematical Foundations of Structural Analysis of Crystals* (State Technical-Theoretical Press, USSR, 1934).

18. G. Flor, D. Orobengoa, E. Tasci, J. Perez-Mato, and M. Aroyo, J. Applied Crystallography **49**, 653 (2016).

19. S. Rass, S. König, S. Ahmad, and M. Goman, *Metricizing the euclidean space towards desired distance relations in point clouds* (2022), URL `https://arxiv.org/abs/2211.03674`.

20. J. H. Conway and N. J. Sloane, Proceedings of the Royal Society A **436**, 55 (1992).

21. P. H. Zwart, R. W. Grosse-Kunstleve, A. A. Lebedev, G. N. Murshudov, and P. D. Adams, Acta Cryst D **64**, 99 (2008).

22. V. Kurlin, arxiv:2201.10543 (2022), URL `http://kurlin.org/projects/lattice-geometry/lattices3Dmaths.pdf`.

23. P. Niggli, *Krystallographische und strukturtheoretische Grundbegriffe*, vol. 1 (Akademische verlagsgesellschaft mbh, 1928).

24. S. L. Lawton and R. A. Jacobson, Tech. Rep., Ames Lab., Iowa State Univ. of Science and Tech., US (1965).

25. H. J. Bernstein and L. C. Andrews, in *Acta Cryst A* (2021), vol. 77, p. C809.

26. K. Gade, Journal of Navigation **63**, 395 (2010).

27. J. Inman, *Navigation and Nautical Astronomy for the Use of British Seamen.* (C and J Rivington, 1835).

28. P. MirÃ³, M. Audiffred, and T. Heine, Chem. Soc. Rev. **43**, 6537 (2014).

29. H. Tian, J. Rice, R. Fei, V. Tran, X. Yan, L. Yang, and H. Wang, Nano Today **11**, 763 (2016), ISSN 1748-0132, URL `https://www.sciencedirect.com/science/article/pii/S1748013216302419`.

30. J. Zhou, L. Shen, M. Costa, and others., Scientific Data **86**, 669 (2019).

31. M. Ashton, J. Paul, S. B. Sinnott, and R. G. Hennig, Phys. Rev. Lett. **118**, 106101 (2017).

32. T. Barnowsky, A. V. Krasheninnikov, and R. Friedrich, Advanced Electronic Materials (2022).

33. R. B. Jacobs-Gedrim, M. T. Murphy, F. Yang, N. Jain, M. Shanmugam, E. S. Song, Y. Kandel, P. Hesamaddin, H. Y. Yu, M. Anantram, et al., Applied Physics Letters **112**, 133101 (2018).

34. N. Mounet, M. Gibertini, P. Schwaller, D. Campi, A. Merkys, A. Marrazzo, T. Sohier, I. E. Castelli, A. Cepellotti, G. Pizzi, et al., Nature nanotechnology **13**, 246 (2018).

35. V. Kulish and Y. Huang, J. Mater. Chem A **5**, 8734 (2017).

36. A. Eftekhari, J. Mater. Chem A **5**, 18299 (2017).

37. H. Wu, Y. Wang, Y. Xu, P. Sivakumar, C. Pasco, U. Filippozzi, S. Parkin, Y.-J. Zeng, M. T., and A. M.N., Nature **604**, 653 (2022).

38. M. Bright, A. I. Cooper, and V. Kurlin, arxiv:2109.11538 (early draft) (2021), URL `http://kurlin.org/projects/periodic-geometry-topology/lattices3Dmap.pdf`.

39. J. Ropers, M. M. Mosca, O. D. Anosova, V. A. Kurlin, and A. I. Cooper, in *International Conference on Data Analytics and Management in Data Intensive Domains* (2022), pp. 178–192.

40. O. Anosova and V. Kurlin, arXiv:2103.02749 (2021).

41. O. Anosova and V. Kurlin, in *Proceedings of Discrete Geometry and Mathematical Morphology* (2021), pp. 229–241.

42. O. Anosova and V. Kurlin, arXiv:2205.15298 (2022), URL `http://kurlin.org/projects/periodic-geometry/near-duplicate-periodic-patterns.pdf`.

43. A. Mattei, R. S. Hong, H. Dietrich, D. Firaha, J. Helfferich, Y. M. Liu, K. Sasikumar, N. S. Abraham, R. Miglani Bhardwaj, M. A. Neumann, et al., Journal of Chemical Theory and Computation **18**, 583 (2022).

44. W. Kruskal and W. Wallis, Journal of the American Statistical Association **47**, 583 (1952).

A Further analysis of chiral distances in the CSD

To further illustrate the behaviour of our proposed chiral distance, we present below some additional analysis for 2D lattices extracted from the CSD in the manner described in section 3.

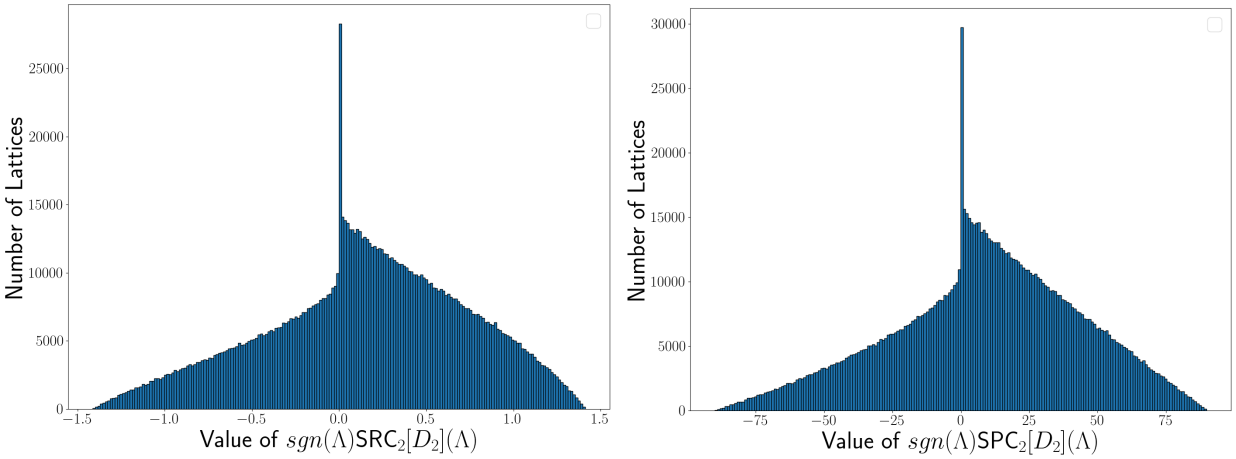

Figure 16: Histogram of signed spherical D_2-chiral distances of 2D lattices from all crystals in the CSD. **Left**: $\mathrm{SRC}[D_2](\Lambda)$ is in Ångstroms $\mathring{A}$. **Right**: $\mathrm{SPC}[D_2](\Lambda)$ is in degrees.

Figure 16 shows the spherical D_2-chiral distance on the complete dataset, exhibiting similar behaviour to the planar version. Any non-oblique lattice is either rectangular (primitive or centred) or has even higher symmetry (square or hexagonal).

Fig. 17 shows the histograms of root distances from rectangular 2D lattices in the CSD to their closest square and hexagonal lattices. The high bar to about 6000 lattices in Fig. 17 (right) indicates that the CSD has many centred rectangular lattices close to hexagonal ones and many fewer rectangular lattices close to square ones.

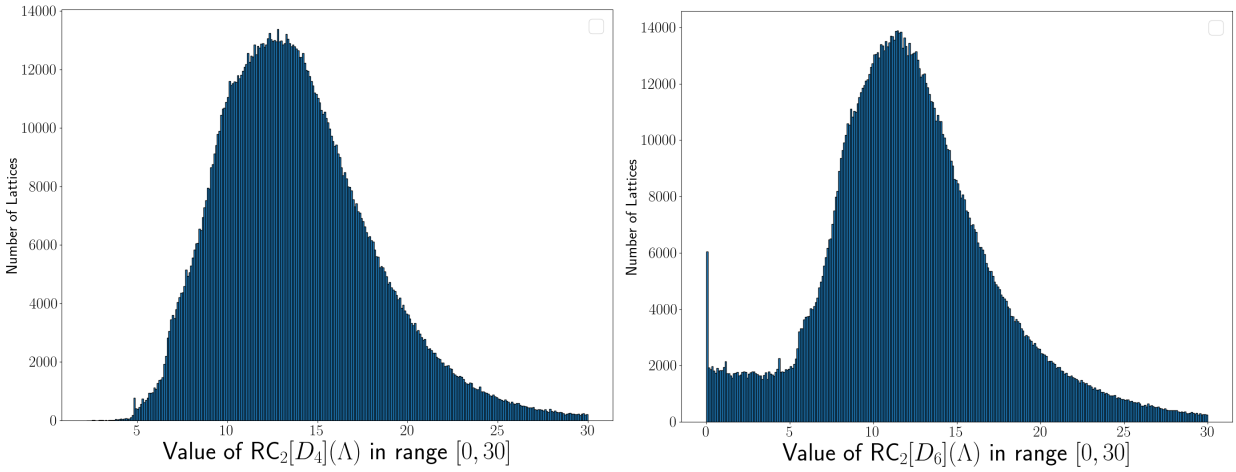

Figure 17: Distances (in Ångstroms) of rectangular (primitive and centred) 2D lattices. **Left**: $\mathrm{RC}_2[D_4](\Lambda)$ to a closest square lattice. **Right**: $\mathrm{RC}_2[D_6](\Lambda)$ to a closest hexagonal lattice.

We can investigate the relationship between continuous chiral distances and chemical characteristics such as molecular weight by isolating the root and projected invariants of 2D lattices derived from the $50,000$ crystals whose constituent molecules had the highest molecular weight, and the $50,000$ of lowest molecular weight.

A preliminary analysis of lattice parameters shows that among the non-oblique lattices (D_2-chiral distance 0) in this sample those with the lowest molecular weight form a higher proportion (57%) of the achiral molecules. By this discrete analysis, we might say that crystals of lower-weight molecules tend to form a more symmetric lattice.

Continuous analysis reveals a more nuanced picture. Fig. 18 compares the histograms of chiral distances for oblique 2D lattices extracted from crystals in the CSD whose main

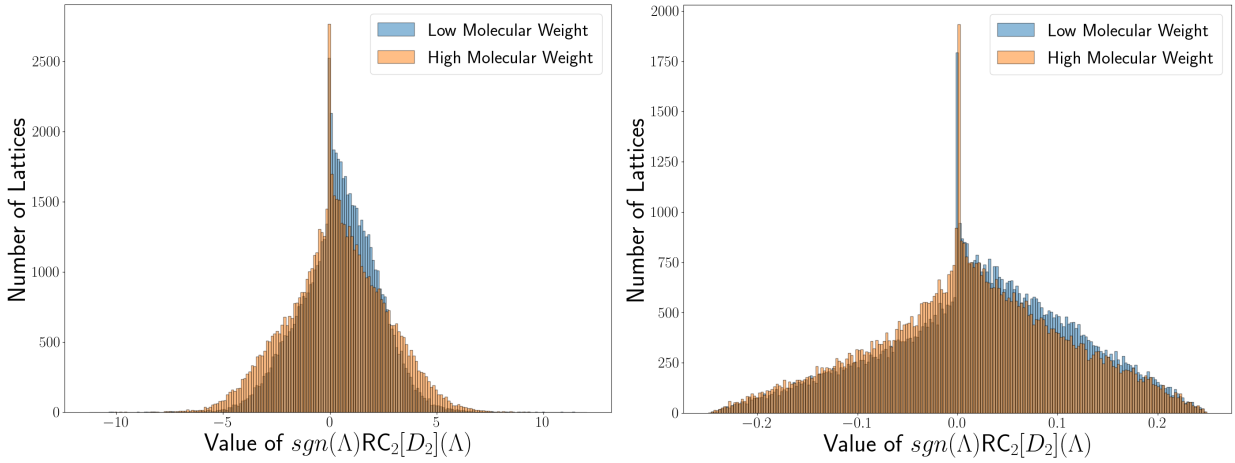

Figure 18: Histograms of $\mathrm{RC}_2[D_2]$ (**left**) and $\mathrm{PC}_2[D_2]$ (**right**) for oblique lattices in crystals whose molecular weight is among the the $50,000$ highest and lowest in the CSD.

molecules have extreme (low or high) weight. The overall distribution of chiral distances for higher molecular weight molecules is slightly wider, indicating that higher molecular weight crystals tend to form lattices with more extreme chiral distances. There is a stronger preference towards negative sign for high-weight molecules which is more clearly visible when comparing values of $\mathrm{RC}_2[D_2]$.

Organometallic crystals have unit cells containing an organic molecule non-covalently bonded to one or more metal ions. Their importance as new materials is illustrated by the fact that nearly half ($\sim 52\%$) of the deposited structures in the CSD which form oblique lattices are organometallic. They are often a target for materials design by crystal structure prediction experiments, and so there is a drive for methods that can help to categorise the outputs of such experiments[43].

In Fig. 19, it appears that organometallic lattices have a stronger tendency to form crystals whose derived 2D lattices have negative signs, and from the histogram of $\mathrm{RC}_2[D_2]$ values we observe a wider distribution of distances overall. Since we cannot assume a normal distribution, but our distributions are sampled from a variable known to be continuous, we apply the non-parametric tests such as Kruskal-Wallace[44] to determine if these distributions are significantly different - and indeed we find a p-value much less than 0.001 in this case.

These differences are relatively small, but they illustrate the sort of analysis that can be

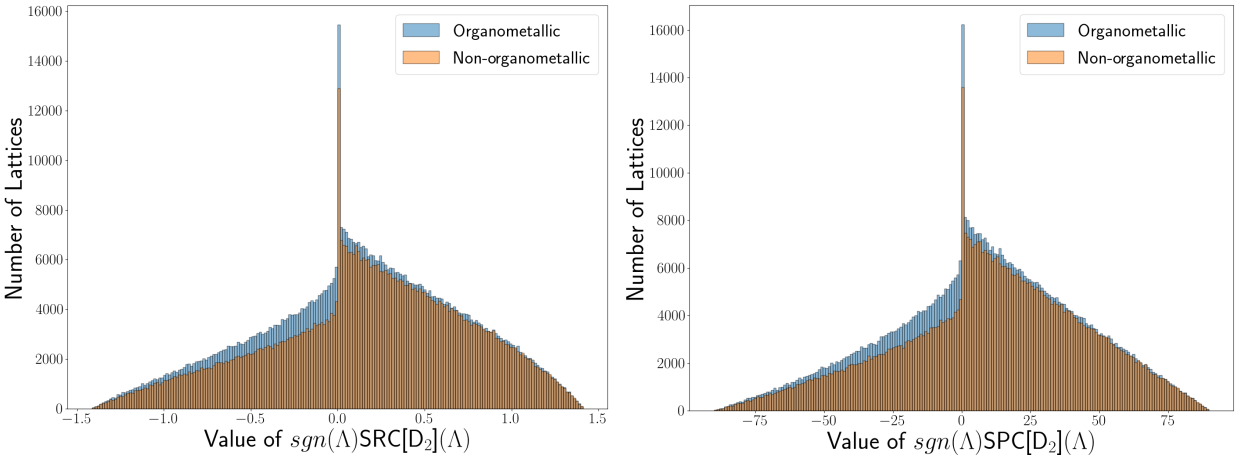

Figure 19: Histograms of spherical D_2-chiral distances of 2D lattices in the CSD, separated into organometallic and non-organometallic structures. **Left**: SRC[D_2]. **Right**: SPC[D_2].

done quickly to investigate possible differences between very large crystal datasets

B Extra Mathematical Details and Proofs

The strain tensor fails the triangle inequality. Section 1 mentioned the strain tensor, which measures the difference between crystal structures in the same space group and can be computed at the Bilbao Crystallographic Server. The simple example below shows that the lattice distortion obtained from this tensor is not a true metric since it fails the triangle inequality. Let $d_S(X, Y)$ denote the lattice distortion between lattices X, Y. Let

$$A = [1, 1, 1, 90°, 90°, 90°], B = [2, 2, 2, 90°, 90°, 90°], C = [3, 3, 3, 90°, 90°, 90°]$$

be the length and angle parameters of three cubic lattices. The online calculation at `https://cryst.ehu.es/cryst/strain.html` show that

$$d(A, B) = 0.86603, d(B, C) = 0.36084, d(A, C) = 2.30940$$

which violates the triangle axiom since $d(A, B) + d(B, C) = 1.22687 < d(A, C) = 2.30940$.

Ordering of Root Products. Definition 1.3 stated that for an obtuse superbase the vector length ordering induces an ordering $r_{12} \leq r_{01} \leq r_{02}$ of the root products. This fact is evident from the detailed mathematical development of the invariant[14], but since it is central to the

quantification of G-chiral distances we provide here a short proof based on the relationship between root products and lattice vector norms. First, we compute

$$r_{02}^2 = -v_0 \cdot v_2 = -v_0 \cdot v_2 = (v_1 + v_2) \cdot v_2 = |v_2|^2 - r_{12}^2,$$

and similarly $r_{01}^2 = -v_0 \cdot v_1 = (v_1 + v_2) \cdot v_1 = |v_1|^2 - r_{12}^2$. Taking the difference, $r_{02}^2 - r_{01}^2 = |v_2|^2 - |v_1|^2 > 0$ and thus $r_{01} \leq r_{02}$. We can also express

$$r_{01}^2 = -v_0 \cdot v_1 = v_0 \cdot (v_0 + v_2) = |v_0|^2 - r_{02}^2,$$

$$r_{12} = -v_1 \cdot v_2 = (v_0 + v_2) \cdot v_2 = |v_2|^2 - r_{02}^2.$$

Again, taking the difference $r_{01} - r_{12} = |v_0|^2 - |v_2|^2 > 0$ and thus $r_{12} \leq r_{01}$.

Proof of Proposition 2.8: For a plane crystallographic group $G \in \{D_2, D_4, D_6\}$, Proposition 2.8 expresses the root spherical chiral G-distance $\mathrm{SRC}[G](\Lambda)$ and the projected spherical chiral G-distance $\mathrm{SPC}[G](\Lambda)$ of any 2D lattice $\Lambda \subset \mathbb{R}^2$ in terms of its spherical root invariant $\mathrm{SRI}(\Lambda) = (\mu, \varphi, \sigma)$.

Case of $\mathrm{SPC}[D_2]$. Any lattice that has the crystallographic group D_2 is mirror-symmetric (primitive or centred rectangular). Hence its spherical projected invariant $\mathrm{SPI} = (\mu, 0)$ lies on the equator of S^2 for some longitude $\mu \in (-180°, 180]$. By Definition 2.6, for any other lattice Λ with $\mathrm{SPI} = (\mu, \varphi)$, the spherical projected D_2-chiral distance $\mathrm{SPC}[D_2](\Lambda)$ equals the haversine distance from (μ, φ) to the equator (minimal along the meridional arc where μ is constant): $\mathrm{SPC}[D_2](\Lambda) = h((\mu, \varphi), (\mu, 0)) = 2 \arcsin \sqrt{\mathrm{hav}(\varphi)} = 2 \arcsin \sqrt{\sin^2 \frac{\varphi}{2}} = |\varphi|$.

Cases of $\mathrm{SPC}[D_4]$ and $\mathrm{SPC}[D_6]$. By Definition 2.5 all square and hexagonal lattices are represented by the spherical projected invariants $(\mu, \varphi) = (67.5°, 0)$ and $(\mu, \varphi) = (-45°, 0)$, respectively, see Fig. 5. By Definition 2.6 the spherical projected chiral distances are $\mathrm{SPC}[D_4] = h((\mu, \varphi), (67.5°, 0°))$ and $\mathrm{SPC}[D_6] = h((\mu, \varphi), (-45°, 0°))$, where the haversine distance h below gives the required formulae in terms of $\mathrm{SPI}(\Lambda) = (\mu, \varphi)$ in Proposition 2.8:

$$h((\mu_1, \varphi_1), (\mu_2, \varphi_2)) = 2 \arcsin \sqrt{\mathrm{hav}(\varphi_1 - \varphi_2) + \mathrm{hav}(\mu_1 - \mu_2) \cos \varphi_1 \cos \varphi_2}, \ \mathrm{hav}(\theta) = \sin^2 \frac{\theta}{2}.$$

For any 2D lattice Λ with a spherical root invariant $\mathrm{SRI}(\Lambda) = (\mu, \varphi, \sigma)$, Definition 2.7 gives the Euclidean coordinates $(x, y, z) = \sigma(\cos \varphi \cos \mu, \cos \varphi \sin \mu, \sin \varphi) \in \mathbb{R}^3$ used below.

Case of SRC$[D_2]$. In the spherical coordinates (μ, φ, σ), the subspace of all mirror-symmetric lattices, which have the crystallographic group D_2, is the equatorial plane $\varphi = 0$ or the horizontal plane $z = 0$ in the Euclidean coordinates $(x, y, z) = \sigma(\cos\varphi\cos\mu, \cos\varphi\sin\mu, \sin\varphi)$. By Definition 2.7 for any 2D lattice Λ, the root spherical distance SRC$[D_2](\Lambda)$ to a closest mirror-symmetric lattice of the same size σ has the minimum Euclidean distance from the fixed point (x, y, z) to the equatorial circle $\sigma(\cos\mu', \sin\mu', 0)$ whose radius σ is fixed, but the longitude μ' is variable. Ignoring the fixed third coordinates and factor σ^2, we minimise the squared distance $s(\mu') = (\cos\varphi\cos\mu - \cos\mu')^2 + (\cos\varphi\sin\mu - \sin\mu')^2$ by differentiation:

$$\begin{aligned}\frac{ds}{d\mu'} &= 2\sin\mu'(\cos\varphi\cos\mu - \cos\mu') - 2\cos\mu'(cos\varphi\sin\mu - \sin\mu'),\\ &= 2\cos\varphi(\cos\mu\sin\mu' - \sin\mu\cos\mu') = 2\cos\varphi\sin(\mu' - \mu)\\ \frac{d^2s}{d\mu'^2} &= 2\cos\varphi\cos(\mu' - \mu).\end{aligned}$$

If $\dfrac{ds}{d\mu'} = 0$ for $\mu \in [-180°, 180°]$ then either $\cos\varphi = 0$ or $\sin(\mu' - \mu) = 0$, which implies that $\mu' - \mu = 180°n$ for $n \in \{-1, 0, 1\}$. If $\cos\phi = 0$, then $\phi = \pm 90°$ and $s = \cos^2\mu' + \sin^2\mu' = 1$ independent of μ, while $\sin\varphi = 1$ giving SRC$_2[D_2](\Lambda) = \sqrt{2}\sigma$.

If $\mu' - \mu = 180n$ and $n = \pm 1$ then $\frac{d^2s}{d\mu'^2} = \cos(\pm 180) = -1$ and s has a maximum. In the remaining case $\frac{d^2s}{d\mu'^2} = \cos(\pm 0) = 1$ and, substituting μ' for μ we find that s achieves a minimum at $s = \cos^2\mu(\cos\varphi - 1)^2 + \sin^2\mu(\cos\varphi - 1)^2 = (1 - \cos\varphi)^2$, so SRC$[D_2](\Lambda) = \sigma\sqrt{(1 - \cos\varphi)^2 + \sin^2\varphi} = \sigma\sqrt{2 - 2\cos\varphi} = \sigma\sqrt{4\sin^2\frac{\varphi}{2}} = 2\sigma|\sin\frac{\varphi}{2}|$.

Case of SRC$[D_4]$. In the spherical coordinates (μ, φ, σ), the subspace of square lattices is defined by $\varphi = 0$, $\mu = 67.5°$, see Fig. 5. The spherical root D_4-chiral distance SRC$[D_4](\Lambda)$ equals the Euclidean distance from $(x, y, z) = \sigma(\cos\varphi\cos\mu, \cos\varphi\sin\mu, \sin\varphi)$ to the point SRI$(\Lambda_4) = \sigma(\cos 67.5°, \sin 67.5°, 0)$, which represents the only square lattice Λ_4 of the same size σ, so SRC$[D_4](\Lambda) = \sigma\sqrt{(\cos\varphi\cos\mu - \cos 67.5°)^2 + (\cos\varphi\sin\mu - \sin 67.5°)^2 + \sin^2\varphi}$.

Case of SRC$[D_6]$. In the spherical coordinates, the subspace of hexagonal lattices is defined by $\varphi = 0$, $\mu = -45°$, see Fig. 5. The spherical root D_6-chiral distance SRC$[D_6](\Lambda)$ equals the Euclidean distance from $(x, y, z) = \sigma(\cos\varphi\cos\mu, \cos\varphi\sin\mu, \sin\varphi)$ to the point SRI$(\Lambda_6) = \sigma(\cos 45°, -\sin 45°, 0)$, which represents the only hexagonal lattice Λ_6 of the same size σ, so SRC$[D_6](\Lambda) = \sigma\sqrt{(\cos\varphi\cos\mu - \cos 45°)^2 + (\cos\varphi\sin\mu + \sin 45°)^2 + \sin^2\varphi}$. □

List of Figure Captions

Figure 1. For the lattices $\Lambda(t)$ generated by $v_1 = (1, 0)$, $(t, 1)$ for $t \in [0, 1]$, a reduced basis with fixed v_1 (and v_2 that belongs to the yellow fundamental domain in the last image) discontinuously changes when the coordinate $x(t)$ of v_2 at $t = \frac{1}{2}$ drops down from $\frac{1}{2}$ to $-\frac{1}{2}$.

Figure 2: Any 2D lattice has one of four crystallographic point groups: oblique or generic (C_2), primitive or centred rectangular (D_2), square or tetragonal (D_4), hexagonal (D_6).

Figure 3: **Left**: all root invariants $\mathrm{RI}(\Lambda)$ live in the cone $\mathrm{TC} \subset \mathbb{R}^3$, which projects to the yellow triangle QT by uniform scaling of lattices $\Lambda \subset \mathbb{R}^2$. **Middle**: all projected invariants $\mathrm{PI}(\Lambda)$ live in the triangle QT parametrised by $x = \bar{r}_{02} - \bar{r}_{01} \in [0, 1)$ and $y = 3\bar{r}_{12} \in [0, 1]$. **Right**: mirror reflections $\Lambda^{\pm} \subset \mathbb{R}^2$ of any non-mirror-symmetric lattice can be represented by a pair $(x, y) \leftrightarrow (1 - y, 1 - x)$ in the square QS symmetric in the diagonal $x + y = 1$

Figure 4: Visualisation of Definition 2.4. **Left**: $\mathrm{RC}_2[D_2](\Lambda)$ and $\mathrm{PC}_2[D_2](\Lambda)$ equal the minimum of three Euclidean distances from the points $\mathrm{RI}(\Lambda) \in \mathrm{TC}$ and $\mathrm{PI}(\Lambda) \in \mathrm{QT}$ to the boundary of the cone TC (along perpendiculars to the three boundary planes) and triangle QT, respectively. **Right**: $\mathrm{PC}_2[D_4](\Lambda), \mathrm{PC}_2[D_6](\Lambda)$ equal the Euclidean distances from $\mathrm{PI}(\Lambda)$ to the vertices $(0, 0), (0, 1) \in \mathrm{QT}$ representing all square and hexagonal lattices, respectively.

Figure 5: The incentre and boundary edges of the triangle QT on the left are mapped to the centre and boundary circle of the disk in the middle and further mapped to the north pole and equator of the unit sphere, whose southern hemisphere is obtained by a similar map of the second triangle, which represents all lattices with negative signs, in Fig. 3 (right).

Figure 6: Signed D_2-chiral distances of all oblique 2D lattices found in the CSD, see Definition 2.4 **Left**: $\mathrm{sign}(\Lambda)\mathrm{RC}_2[D_2](\Lambda)$ in Ångstroms. **Right**: $\mathrm{sign}(\Lambda)\mathrm{PC}_2[D_2](\Lambda)$ is unitless.

Figure 7: Signed D_4-chiral distances of 1,177,678 oblique 2D lattices in the CSD, see Definition 2.4 **Left**: $\mathrm{sign}(\Lambda)\mathrm{RC}_2[D_4](\Lambda)$ in Ångstroms. **Right**: $\mathrm{sign}(\Lambda)\mathrm{PC}_2[D_4](\Lambda)$ is unitless.

Figure 8: Signed D_6-chiral distances of 1,177,678 oblique 2D lattices in the CSD, see Definition 2.4. **Left**: $\mathrm{sign}(\Lambda)\mathrm{RC}_2[D_6](\Lambda)$ in Ångstroms. **Right**: $\mathrm{sign}(\Lambda)\mathrm{PC}_2[D_6](\Lambda)$ is unitless.

Figure 9: Behaviour of chiral distances and QT plots in relation to lattice parameters. **Left**: Values of $\mathrm{PC}_2[D_4]$ and $\mathrm{PC}_2[D_6]$ for a lattice with parameters $a = 1, b = 1$ and angle $\theta \in [90°, 120°]$. **Centre**: Map of all CSD lattices in the Quotient Triangle (reproduced from

our previous paper[15]). **Right**: Intersection of QT with circles centred at the origin (red) and the point $(0, 1)$ (green), of radii $r_1 = \frac{1}{2}, r_2 = \frac{1}{2}$ respectively, and $r_i \pm \epsilon$.

Figure 10: Two dimensional lattices of 6,351 monolayer structures[30] isolated from 3D crystals by layer detection or generated by atomic substitution, invariants PI(Λ) in the square QS, see Definition 2.3.

Figure 11: Scatter plot of physical properties of materials discovered by layer detection in 2DMatPedia vs $sgn(\Lambda)\mathrm{RC}[D_2](\Lambda)$. **Left**: Decomposition energy in meV. **Right**: Exfoliation energy in meV per atom.

Figure 12: Scatter plot of physical properties of materials discovered by atomic substitution in 2DMatPedia vs $sgn(\Lambda)\mathrm{RC}[D_2](\Lambda)$. **Left**: Decomposition energy in meV. **Right**: Exfoliation energy in meV per atom.

Figure 13: $2D$ scatter plot of exfoliation vs.decomposition energy, with chiralities indicated by colour **Left**: All structures in 2DMatPedia. **Right**: Structures where $\mathrm{RC}_2[D_2] \geq 1$

Figure 14: 2D lattices of 1726 monolayer structures[34] isolated from 3D crystals and shown by the invariants PI(Λ) in the square QS, see Definition 2.3 and labels of structures in Fig. 15.

Figure 15: 183 'relaxed' by DFT 2D lattices from 1726 monolayer structures[34] in Fig. 14. **Left**: structures with oblique lattices and molecules of type MX2 are labelled in green. **Right**: elemental crystals and molecules of type TMDC are labelled in red and blue.

Figure 16: Histogram of signed spherical D_2-chiral distances of 2D lattices of all crystals found in the CSD. **Left**: $\mathrm{SRC}[D_2](\Lambda)$ is in Ångstroms $\mathring{A}$. **Right**: $\mathrm{SPC}[D_2](\Lambda)$ is in degrees.

Figure 17: Distances (in Ångstroms) of rectangular (primitive and centred) 2D lattices. **Left**: $\mathrm{RC}_2[D_4](\Lambda)$ to a closest square lattice. **Right**: $\mathrm{RC}_2[D_6](\Lambda)$ to a closest hexagonal lattice.

Figure 18: Histograms of $\mathrm{RC}_2[D_2]$ (**left)** and $\mathrm{PC}_2[D_2]$ (**right**) for oblique lattices in crystals whose molecular weight is among the $50,000$ highest and lowest in the CSD.

Figure 19: Histograms of spherical D_2-chiral distances of 2D lattices in the CSD, separated into organometallic and non-organometallic structures. **Left**: $\mathrm{SRC}[D_2]$. **Right**: $\mathrm{SPC}[D_2]$.