MS38-2-2 The pointwise distance distribution is stronger than the pair distribution function #MS38-2-2

V. Kurlin [1], D. Widdowson [1]
[1]University of Liverpool - Liverpool (United Kingdom)

Abstract

A periodic crystal is usually given in a Crystallographic Information File as a pair of a (primitive, conventional, or reduced) unit cell and a motif of atoms with fractional coordinates with respect to a cell basis. This traditional representation of crystals doesn't allow us to quickly and continuously quantify the similarity [1] between nearly identical crystals because a reduced cell can discontinuously increase under almost any tiny displacement of atoms.

The Pair Distribution Function (PDF) was a convenient tool [2] to distinguish periodic crystals up to rigid motion or isometry maintaining all interatomic distances. The exact PDF can be defined as the infinite distribution of pairwise distances between all atomic centres with multiplicities for repeated distances but avoiding all repetitions due to lattice translations. For convenience, this exact PDF can be written as an infinite sequence of increasing distances.

Figure 1 (left) shows the classical example of non-isometric sets of 4 points in the plane that have the same PDF of 6 distances: root(2), root(2), 2, root(10), root(10), 4. Figure 1 (right) shows a similar example of 1-dimensional periodic sequences S(r) and Q(r), which both have the unit cell [0,8] with 4 points in their motifs and depend on a parameter 0<r<1. These are examples of homometric sets that have the same PDF and the Patterson function [3].

The recently discovered Pointwise Distance Distribution (PDD) extends PDF to a stronger isometry invariant that is provably continuous under any small perturbations of points [4].

The PDD(S;k) of a periodic point set S is a matrix of rows consisting of increasing distances from every motif point of S to its k neighbours within the infinite set S, see Figure 2 (left).

PDD matrices are cpntinuously compared by the Earth Mover's Distance in Figure 2 (right).

All periodic sets in general position, including the sets in Figure 1, are provably distinguished by PDD for a large enough k with an explicit upper bound [4]. The PDF can be reconstructed from PDD, not vice versa. These theoretical results have been confirmed by 200B+ pairwise comparisons of all 660K+ periodic crystals in the Cambridge Structural Database.

This huge experiment took only a couple of days on a modest desktop [4] and detected five pairs of unexpected duplicate structures that are truly isometric to the last decimal place, but one atom is replaced by a different one, for example, Cd by Mn in HIFCAB vs JEPLIA.

References

[1] Sacchi, P., Matteo Lusi, M., Cruz-Cabeza, A.J., Nauhac, E.Joel Bernstein, J. Same or different – that is the question: identification of crystal forms from crystal structure data. CrystEngComm 22 (43), 7170-7185, 2020.

[2] Terban, Maxwell and Billinge, Simon. Structural analysis of molecular materials using the pair distribution function. Chemical Reviews 122 (2022), 1208-1272.

[3] Patterson, AL. Homometric structures, Nature 143 (1939), 939-940.

[4] Widdowson, D., Kurlin, V., Pointwise Distance Distributions of finite and periodic point sets. Arxiv.org:2108.04798. The latest version is at http://kurlin.org/projects/periodic-geometry-topology/PDD.pdf

Non-isometric sets with the same PDF

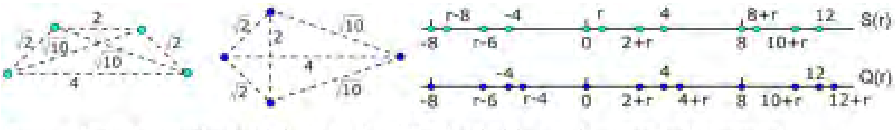

Figure 1: **Left**: the sets $K = \{(\pm 2, 0), (\pm 1, 1)\}$ and $T = \{(\pm 2, 0), (-1, \pm 1)\}$ can not be distinguished by their six pairwise distances $\sqrt{2}, \sqrt{2}, 2, \sqrt{10}, \sqrt{10}, 4$. **Right**: the 1D periodic sets $S(r) = \{0, r, 2+r, 4\} + 8\mathbb{Z}$ and $Q(r) = \{0, 2+r, 4, 4+r\} + 8\mathbb{Z}$ for $0 < r \leq 1$ have the same Patterson function and Pair Distribution Function. Both pairs are distinguished by the Pointwise Distance Distribution.

Pointwise distance distribution with a metric

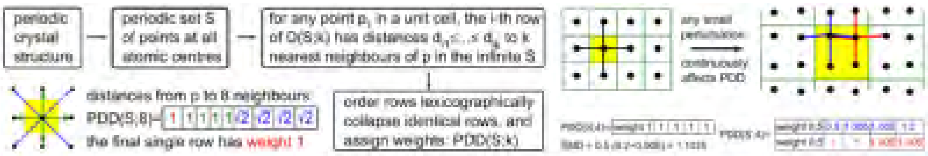

Figure 2: **Left**: pipeline of the Pointwise Distance Distribution (PDD). **Right**: example computation of the Earth Mover's Distance (EMD) between PDDs of a square lattice and its perturbation for $k = 4$ neighbors.