PDDFormer: Pairwise Distance Distribution Graph Transformer for Crystal Material Property Prediction

Xiangxiang Shen[1†], **Zheng Wan**[2†], **Lingfeng Wen**[1], **Licheng Sun**[1], **Jian Yang**[3], **Xuan Tang**[4], **Shing-Ho J. Lin**[5], **Xiao He**[2], **Mingsong Chen**[1] and **Xian Wei**[1*]

[1]Software Engineering Institute, East China Normal University
[2]School of Chemistry and Molecular Engineering, East China Normal University
[3]School of Geospatial Information, Information Engineering University
[4]School of Communication and Electronic Engineering, East China Normal University
[5]School of Artificial Intelligence, University of Chinese Academy of Sciences
xian.wei@tum.de

Abstract

Crystal structures can be simplified as a periodic point set that repeats across three-dimensional space along an underlying lattice. Traditionally, crystal representation methods characterize the structure using descriptors such as lattice parameters, symmetry, and space groups. However, in reality, atoms in materials always vibrate above absolute zero, causing their positions to fluctuate continuously. This dynamic behavior disrupts the fundamental periodicity of the lattice, making crystal graphs based on static lattice parameters and conventional descriptors discontinuous under slight perturbations. Chemists proposed the pairwise distance distribution (PDD) method to address this problem. However, the completeness of PDD requires defining a large number of neighboring atoms, leading to high computational costs. Additionally, PDD does not account for atomic information, making it challenging to apply it directly to crystal material property prediction tasks. To tackle these challenges, we introduce the atom-Weighted Pairwise Distance Distribution (WPDD) and Unit cell Pairwise Distance Distribution (UPDD) and apply them to the construction of multi-edge crystal graphs. We demonstrate the continuity and general completeness of crystal graphs under slight atomic position perturbations. Moreover, by modeling PDD as global information and integrating it into matrix-based message passing, we significantly reduce computational costs. Comprehensive evaluation results show that WPDDFormer achieves state-of-the-art predictive accuracy across tasks on benchmark datasets such as the Materials Project and JARVIS-DFT.

1 Introduction

Crystals are solids with a regular geometric shape formed by atoms, ions, or molecules arranged periodically in space during the crystallization process. Their structure is typically described using repeating unit cells and lattice vectors. However, this description method brings a fundamental challenge: the same crystal structure can be represented by different unit cells and lattice vectors, as shown in Figure 1(a). Additionally, in real-world scenarios, the experimental coordinates of unit cells and atoms are inevitably affected by atomic vibrations and measurement noise. These subtle disturbances can lead to discontinuous changes in any simplified unit cell [Kurlin, 2024], resulting in numerous unit cells for a given crystal structure, as shown in Figure 1(b), thereby introducing ambiguity in the representation of crystal data [Widdowson and Kurlin, 2022]. Currently, many graph neural networks [Batzner *et al.*, 2022; Yan *et al.*, 2022; Yan *et al.*, 2024a; Yan *et al.*, 2024b] typically use unit cell parameters, simplified cell parameters, symmetry, and space groups to represent the periodic structure of crystals. However, these features are either non-invariant or discontinuous [Zwart *et al.*, 2008] invariants, leaving the issue of ambiguity in crystal data unresolved [Patterson, 1944; Widdowson *et al.*, 2022; Groom *et al.*, 2016; Bartók *et al.*, 2013; Wassermann *et al.*, 2010; Ahmad *et al.*, 2018].

The continuous and complete invariant—Pairwise Distance Distribution (PDD)—proposed by [Widdowson and Kurlin, 2022] addresses the ambiguity in crystal data representation by distinguishing all periodic structures in the world's largest real material collection, the Cambridge Structural Database. To achieve completeness, PDD requires a predetermined number of sufficient neighbors, which is computationally expensive and difficult to directly apply for predicting crystal properties [Balasingham *et al.*, 2022]. [Balasingham *et al.*, 2024] employed distance distribution graphs (DDGs) based on PDD to predict the properties of crystal materials, but they did not achieve satisfactory performance (only slightly better than CGCNN), and although this approach reduced computational costs, it compromised the completeness of PDD. In contrast, crystal graph representations based on multi-edge crystal graphs and unit cell parameters [Yan *et al.*, 2024a] achieve completeness, more accurately characterizing crystal structures, and achieving state-

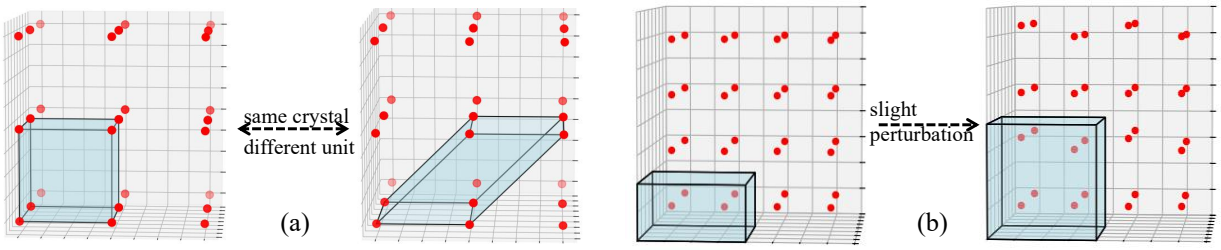

Figure 1: Illustrations of different unit cells and lattice representations of the same crystal structure. Figure (a) shows several possible choices among the infinitely many unit cells for the same crystal structure in the undisturbed case. Figure (b) illustrates that for almost any perturbation, the symmetry group and any reduced unit cell (with minimal volume) will undergo discontinuous changes.

of-the-art performance in crystal material property prediction tasks. However, using unit cell parameters leads to discontinuities in the crystal graphs.

Since PDD does not account for atomic types, it is challenging to use it directly for effective crystal property prediction. To better represent crystal structures, we first introduce WPDD and UPDD. Furthermore, we integrate WPDD and UPDD into the construction of multi-edge crystal graphs and propose the PDD Graph Transformer (including WPDDFormer and UPDDFormer) based on the transformer architecture. We model WPDD as global information and incorporate it into matrix-based message passing without introducing additional information to capture periodicity, significantly reducing computational costs (as shown in Table 3). Finally, we employ the Earth Mover's Distance (EMD) [Rubner *et al.*, 2000] to assess the continuity of crystal graphs, demonstrating that WPDD crystal graphs constructed using only Euclidean distances maintain continuity and general completeness under slight atomic position perturbations, providing a more accurate depiction of actual crystal structures. Ablation experiments show the crucial role of (W/U)PDD in constructing crystal graphs. Through comprehensive evaluations, our method achieves state-of-the-art predictive accuracy across various tasks in the Materials Project and JARVIS datasets. This advancement highlights the effectiveness of WPDDFormer in bridging the gap between traditional crystal descriptors and dynamic atomic behavior, leading to more accurate and reliable predictions in materials science.

2 Preliminaries

2.1 The Structure of Crystals

By selecting an appropriate structural unit, the entire crystal structure can be viewed as the periodic repetition of this unit in space. This property, where atoms within a crystal repeat in three-dimensional space according to a specific pattern, is called periodicity, with the smallest repeatable structural unit being the unit cell. The unit cell can be defined as $\mathcal{U} = (\mathcal{X}, \mathcal{P})$, where $\mathcal{X}$ and $\mathcal{P}$ can be represented in matrix form. Typically, $\mathcal{X} = [x_1, x_2 \cdots x_{n-1}, x_n]^T \in \mathbb{R}^{n\times 1}$, where n represents the number of atoms and $x_i \in \mathbb{R}^1$ represents the atomic type of atom i in the unit cell. $\mathcal{P} = [p_1, p_2 \cdots p_{n-1}, p_n]^T \in \mathbb{R}^{n\times 3}$ is the atomic position matrix, where $p_i \in \mathbb{R}^3$ represents the Cartesian coordinates of the atom i in the unit cell in 3D space. The lattice vectors $\mathcal{L} = [l_1, l_2, l_3]^T \in \mathbb{R}^{3\times 3}$ can reflect the way the unit cell repeats in three directions to map the periodic crystal structure. Therefore, in 3D space, the infinite crystal structure $\mathcal{S}$ can be represented as $(\mathcal{U}, \mathcal{L})$.

2.2 Definitions

Definition 1. ***Pointwise Distance Distribution.*** *For the infinite crystal structure $\mathcal{S} = (\mathcal{U}, \mathcal{L})$ mentioned in Section 2.1, fix a neighbor count $k \geq 1$. For each point x_i in the unit cell $\mathcal{U}$, let $d_{i1} \leq \cdots \leq d_{ik}$ be the Euclidean distances from $\boldsymbol{p}_i$ to its k nearest neighbors in the infinite crystal structure. Consider an $n \times k$ matrix composed of n rows of distance vectors, where each point $x_i \in \mathcal{U}$ corresponds to one row. If the matrix contains $m \geq 1$ identical rows, they are merged into one row with a weight of $\frac{m}{n}$. The resulting matrix can be regarded as a weighted distribution of rows, which is called the Pairwise Distance Distribution $\mathcal{PDD}(S; k) \in \mathbb{R}^{n\times(k+1)}$.*

According to [Widdowson and Kurlin, 2022] and [Yan *et al.*, 2024a], we present Definitions 2-3. According to [Widdowson and Kurlin, 2022], we present Definitions 4-6.

Definition 2. ***Isometric Crystal Graphs.*** *An isometric transformation is a mapping that preserves Euclidean distances, denoted as $f(x) = Rx + b$. Any isometric transformation f can be decomposed into translation, rotation, and reflection. Specifically, suppose there exists a rotation matrix $R \in \mathbb{R}^{3\times 3}$, with a determinant of 1 ($|R| = 1$), and a translation vector $b \in \mathbb{R}^3$, then two crystal structures $\mathcal{S} = (\mathcal{U}, \mathcal{L})$ and $\mathcal{Q} = (\mathcal{U}', \mathcal{L}')$ are isometric, satisfying $\mathcal{U}' = R\mathcal{U} + b$, where $R\mathcal{U} + b$ denotes the application of the rotation R and translation b to each element in the infinite set $\mathcal{U}$.*

Suppose $\mathcal{S}$ and $\mathcal{Q}$ are isometric. In that case, their crystal graph representations satisfy $\mathcal{G}(\mathcal{S}) = \mathcal{G}(\mathcal{Q})$, which means that the graphical representation of the crystal structure produces no false positives; that is, there are no isometric pairs where $\mathcal{G}(\mathcal{S}) \neq \mathcal{G}(\mathcal{Q})$ but $\mathcal{S} \simeq \mathcal{Q}$. Conversely, if $\mathcal{G}(\mathcal{S}) = \mathcal{G}(\mathcal{Q})$, then $\mathcal{S}$ and $\mathcal{Q}$ are isometric, meaning f produces no false negatives, i.e., there are no non-isometric pairs where $\mathcal{G}(\mathcal{S}) = \mathcal{G}(\mathcal{Q})$ but $\mathcal{S} \not\simeq \mathcal{Q}$. That is, if the crystal graph representations of artificially constructed crystal structures are identical under isometric transformations, then they are geometrically equivalent.

Definition 3. ***Geometrically Complete Crystal Graphs.*** *If we construct crystal graphs $\mathcal{G}(\mathcal{S}) = \mathcal{G}(\mathcal{Q}) \Longrightarrow \mathcal{S} \simeq \mathcal{Q}$, where*

$\simeq$ denotes the isomorphism of two crystals as defined in Definition 2, then the crystal graph $\mathcal{G}$ is geometrically complete. This means that if two crystal graphs $\mathcal{G}(\mathcal{S})$ and $\mathcal{G}(\mathcal{Q})$ are identical, the infinite crystal structures represented by $\mathcal{G}(\mathcal{S})$ and $\mathcal{G}(\mathcal{Q})$ are also identical. If the constructed crystal graph $\mathcal{G}$ can distinguish any subtle structural differences between different crystal materials, it is said to be geometrically complete.

Definition 4. ***Metric.*** *The metric d between crystal graphs $\mathcal{G}$ satisfies all the axioms: 1) $d(\mathcal{G}(\mathcal{S}) = \mathcal{G}(\mathcal{Q})) = 0$ if and only if $\mathcal{G}(\mathcal{S}) = \mathcal{G}(\mathcal{Q})$; 2) Symmetry: $d(\mathcal{G}(\mathcal{S}), \mathcal{G}(\mathcal{Q})) = d(\mathcal{G}(\mathcal{Q}), \mathcal{G}(\mathcal{S}))$; 3) Triangle inequality: $d(\mathcal{G}(\mathcal{S}), \mathcal{G}(\mathcal{Q})) + d(\mathcal{G}(\mathcal{Q}), \mathcal{G}(\mathcal{K})) \geq d(\mathcal{G}(\mathcal{S}), \mathcal{G}(\mathcal{K}))$.*

Definition 5. ***Lipschitz continuity of crystal graphs.*** *If $\mathcal{Q}$ is obtained by moving each point in the periodic crystal $\mathcal{S} \subset \mathbb{R}^n$ by no more than ϵ, and the distance of the constructed crystal graph structures satisfies $d(\mathcal{G}(\mathcal{S}), \mathcal{G}(\mathcal{Q})) \leq C\epsilon$, where C is a constant, then the crystal graph is continuous, and $\mathcal{Q}, \mathcal{S} \subset \mathbb{R}^n$ can be any periodic crystal structures.*

Definition 6. ***EMD.*** *Let $\mathcal{G}(\mathcal{S})$ and $\mathcal{G}(\mathcal{Q})$ be the crystal graph structures we construct for periodic crystals $\mathcal{S}$ and $\mathcal{Q} \in \mathbb{R}^n$. The flow from $\mathcal{G}(\mathcal{S})$ to $\mathcal{G}(\mathcal{Q})$ is represented by an $n(\mathcal{S}) \times n(\mathcal{Q})$ matrix, where the elements $f_{ij} \in [0, 1]$ indicate the partial flow from $\mathcal{R}_i(\mathcal{S})$ to $\mathcal{R}_j(\mathcal{Q})$. The Earth Mover's Distance (EMD) is defined as the minimum cost:*

$$\begin{gathered}\mathrm{EMD}(\mathcal{G}(\mathcal{S}), \mathcal{G}(\mathcal{Q})) = \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij} |R_i(\mathcal{S}) - R_j(\mathcal{Q})| \\ s.t. \sum_{i=0}^{n} f_{ij} \leq w_i(\mathcal{S}), \sum_{j=0}^{n} f_{ij} \leq w_j(\mathcal{Q}), \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij} = 1\end{gathered} \tag{1}$$

The first condition $\sum_{i=0}^{n} f_{ij} \leq w_i(\mathcal{S})$ means that not more than the weight $w_i(\mathcal{S})$ of the component $R_i(\mathcal{S})$ 'flows' into all components $R_j(\mathcal{Q})$ via 'flows' f_{ij} . Similarly, the second condition $\sum_{j=0}^{n} f_{ij} \leq w_j(\mathcal{Q})$ means that all 'flows' f_{ij} from $R_i(\mathcal{S})$ 'flow' Into $R_j(\mathcal{Q})$ up to the maximum weight $w_j(\mathcal{Q})$. The last condition $\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij} = 1$ forces to 'flow' all rows $R_i(\mathcal{S})$ to all rows $R_j(\mathcal{Q})$.

3 Related Work

Finite Crystal Graph Representation CGCNN [Xie and Grossman, 2018] represents crystal structures as finite multi-edge crystal graphs to model crystal structures and predict material properties. Building on the construction of multi-edge crystal graphs, MegNet [Chen *et al.*, 2019] introduced global state attributes into graph networks, while GATGNN [Louis *et al.*, 2020] utilized multiple graph attention layers (GAT) to learn the properties of local neighborhoods and employed global attention layers to weight global atomic features. ALIGNN [Choudhary and DeCost, 2021] and M3GNet [Chen and Ong, 2022] incorporated angular information into the message-passing process to generate richer and more discriminative representations. [Das *et al.*, 2022] proposed an interpretable deep property predictor called CrysXPP. CrysMMNet [Das *et al.*, 2023a] adopted a multimodal framework, integrating graph and text representations to produce joint multimodal representations of crystalline materials. [Das *et al.*, 2023b] proposed CrysGNN, a pretraining framework leveraging unlabeled data, while CrysDiff [Song *et al.*, 2024] introduced a diffusion model-based pretraining–fine-tuning framework. However, the methods above represent crystals as finite graph structures, which fail to effectively capture the periodicity of infinite crystals.

Periodic Representation of Crystals Recently, Matformer [Yan *et al.*, 2022] encoded periodic patterns by adding self-connecting edges to atoms based on lattice parameters. PotNet [Lin *et al.*, 2023] considered the infinite summation of interatomic interactions. Crystalformer [Taniai *et al.*, 2024] performed infinite summations of interatomic potentials through infinitely connected attention while also utilizing lattice parameters. ComFormer [Yan *et al.*, 2024a] constructed cell parameters by adding self-connecting edges to atoms and their copies in three different directions to encode periodic patterns, employing equivariant vector representations and invariant geometric descriptors of Euclidean distances and angles to represent the geometric information of crystals. GMTNet [Yan *et al.*, 2024b] aims to predict the tensor properties of crystalline materials while satisfying O(3) group equivariance and the symmetry of crystal space groups. However, the crystal structures they represent rely on non-invariant or discontinuous invariants, failing to resolve the issue of crystal data fuzziness.

Continuity and Complete Representations for Crystals Addressing the continuity and completeness of crystal representations is a critical issue. Recent advancements in AMD [Wang *et al.*, 2022] and PDD [Widdowson and Kurlin, 2022] have developed matrix forms that are both complete and continuous. However, in practical applications, using these matrix representations as inputs for predicting crystal properties without compromising continuity and completeness is challenging. The AMD and PDD representations are designed to distinguish stable crystal structures and do not consider atomic types; their completeness assumption only holds for stable structures. Additionally, to achieve completeness, a sufficiently large number of neighbors k must be predetermined for any test crystal. Directly modeling PDD as edge information is impractical and costly in real-world applications [Balasingham *et al.*, 2022].

4 PDDFormer

In this section, we propose two PDD variants, namely WPDD and UPDD, and then incorporate them into crystal graph construction. We finally present the PDDformer framework.

4.1 Atom-Weighted PDD (WPDD)

Since the PDD representation is designed for stable crystal structures and does not consider atomic types, it is not suitable for predicting crystal material properties. To account for the influence of atomic types, for a given crystal structure $\mathcal{S} = \mathcal{U} + \mathcal{L}$, where each atom $x_i \in \mathcal{U}$ is labeled with the atomic mass $t(x_i)$ corresponding to it, the final weight for each row is $\mathcal{W} = [w_1, \ldots, w_n]^T$, where $w_i = \frac{t(x_i)}{\sum_{u=1}^{n} t(x_u)}$. By concatenating $\mathcal{W}$ with $\mathcal{PDD} \in \mathbb{R}^{n \times k}$, an atomic-mass-weighted $\mathcal{WPDD}(S; k) \in \mathbb{R}^{n \times (k+1)}$ is formed, represented by the following equation:

$$\mathcal{WPDD} = (\mathrm{W}, \mathcal{PDD}) = \left(\bigcup_{i=1}^{n} \frac{t(x_i)}{\sum_{u=1}^{n} t(x_u)}, \bigcup_{j=1}^{k} d(p_i, p_j)\right) \tag{2}$$

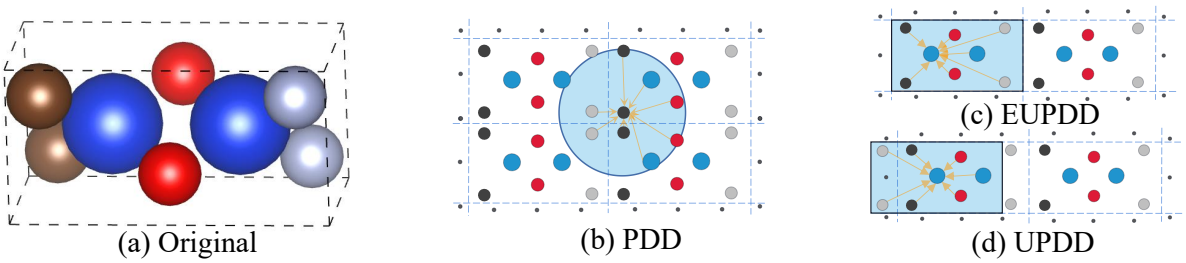

Figure 2: Schematic diagram of the selected neighbors in PDD. Figure (a) represents the 3D unit cell structure. The edges in Figure (b) show the neighbor selection for atom i in WPDD. By comparing Figures (c) and (d), we can see that we construct the unit cell centered around each atom and select neighbors, rather than being limited to the unit cell where the atoms are located.

Herein, n represents the number of atoms in the unit cell, and p_i and p_j denote the spatial coordinates of an atom i and its neighbor j, respectively, and k is the number of nearest neighbors selected when constructing the PDD, sorted in ascending order of Euclidean distance as $d_{i1} \leq \cdots \leq d_{ik}$, as shown in Figure 2(b). WPDD is equivalent to the PDD of the crystal structure $\mathcal{S}$, except that the rows are not grouped as in the original version. This prevents the loss of atomic information when two primitive points have the same k-nearest neighbor distances but correspond to different atomic types. Therefore, $\mathcal{WPDD} \in \mathbb{R}^{n\times(k+1)}$, where n is the number of atoms in the constructed graph.

4.2 Unit-cell PDD (UPDD)

When ensuring the completeness of PDD, a large number of neighbors must be predefined, typically requiring information on hundreds of neighbors, and in extreme cases, the number must exceed the atom count in any unit cell within the dataset. The number of neighbors, k, is difficult to determine across different datasets, and for unit cells with fewer atoms, which constitute a larger proportion of the dataset, an excess of neighbor information may interfere with the speed of message aggregation, leading to greater resource consumption.

To address this issue, we introduce Unit-cell PDD (UPDD). We achieve this by reconstructing the unit cell around each atom and encoding the pairwise distances between the atom and other atoms within the reconstructed unit cell. This means that when constructing PDD, we focus more on the overall structure of the atoms within the reconstructed unit cell, thereby reducing interference from excessive neighbor information. UPDD is defined by the following formula:

$$\mathcal{UPDD} = \left\{ \bigcup_{i=1}^{n} \bigcup_{j=1}^{n} d(p_i, p_j)^{-1} \mid i, j \in \mathcal{Z}, p_i \neq p_j \right\} \quad (3)$$

Since the interaction energy between an atom and its neighboring atoms is usually inversely proportional to the distance, we take the reciprocal feature of the distance after removing zeros.

As shown in Figure 2(d), the selection is not based on Euclidean distances but on choosing atoms within the reconstructed unit cell for construction. The atoms in the unit cell determine the dimension of our UPDD and do not require consideration of the neighbor count, k, across different datasets, making it more generalizable. This UPDD covers unit cell structures with a larger number of atoms while ensuring that unit cell structures with fewer atoms are not disturbed by excessive neighbor information. It also reduces resource consumption. Due to this crystal-specific treatment, the UPDD dimensions of different crystal structures may not match, so dimension alignment is required before feeding them into the neural network.

4.3 Crystal Graph Construction

By introducing PDD, we constructed a general complete and continuous multi-edge crystal graph. In the graph, node features are x_i. An edge is established from node j to node i when the Euclidean distance $|e_{j'i}|^2$ between a duplicate of j and i satisfies $|e_{j'i}|^2 = |p_j + k_1' l_1 + k_2' l_2 + k_3' l_3 - p_i|^2 \leq r$, where $r \in \mathbb{R}$ is the cutoff radius. Next, we construct a PDD row for each atom. Since directly representing PDD as edge features is impractical, we retain its matrix form and incorporate it into the construction of the multi-edge crystal graph to reflect the global information of the crystal structure. Therefore, we represent the constructed crystal graph as $\mathcal{G} = (\mathcal{X}, \mathcal{XI}, \mathcal{E}, \mathcal{PDD})$. Therein, $x_i \in \mathcal{X}$ is the feature vector of the atom i, $e_{ij}^h \in \mathcal{E}$ is the feature vector of the h-th edge between nodes i and j, and we denote $\mathcal{XI}$ as the indices of the nodes i and j that form the edge. Sections 4.4 and 4.5 are our proofs of the continuity and general geometric completeness of PDD crystal graphs.

4.4 The Continuity of Proposed Crystal Graphs

The continuity of the constructed crystal graph $\mathcal{G}(\mathcal{S})$ under perturbations of the crystal structure $\mathcal{S}$ will be measured using the EMD [Rubner *et al.*, 2000], which applies to crystal graphs of any size. Definition 6 applies to any crystal graph $\mathcal{G}(\mathcal{S}) = ([w_1(S), R_1(S)], \ldots, [w_n(S), R_n(S)])$, where $[w_i(S), R_i(S)]$ represent the information extracted based on atom i in the unit cell. $R_i(S) = R_i(\mathcal{S}_\mathcal{X}, \mathcal{S}_\mathcal{XI}, \mathcal{S}_\mathcal{E}, \mathcal{S}_\mathcal{PDD})$ includes atomic information, neighbor information used in constructing the multi-edge crystal graph, and the PDD invariants of the crystal structure $\mathcal{S}$, with weights $w_i \in (0, 1]$ satisfying the normalization condition $\sum_{i=1}^{n} w_i(\mathcal{S}) = 1$.

Subsequently, we only consider the case where the weighted vector $[w_i, R_i]$ corresponds to the i-th row of $\mathcal{PDD}(S; k)$. Here, n denotes the number of rows in $\mathcal{PDD}(S; k)$. The size of each row $R_i(S)$ should be independent of $\mathcal{S}$ and depend solely on the number of neighbors k in $\mathcal{PDD}(S; k)$. For any vectors $R_i = (r_{i1}, \ldots, r_{ik})$ and

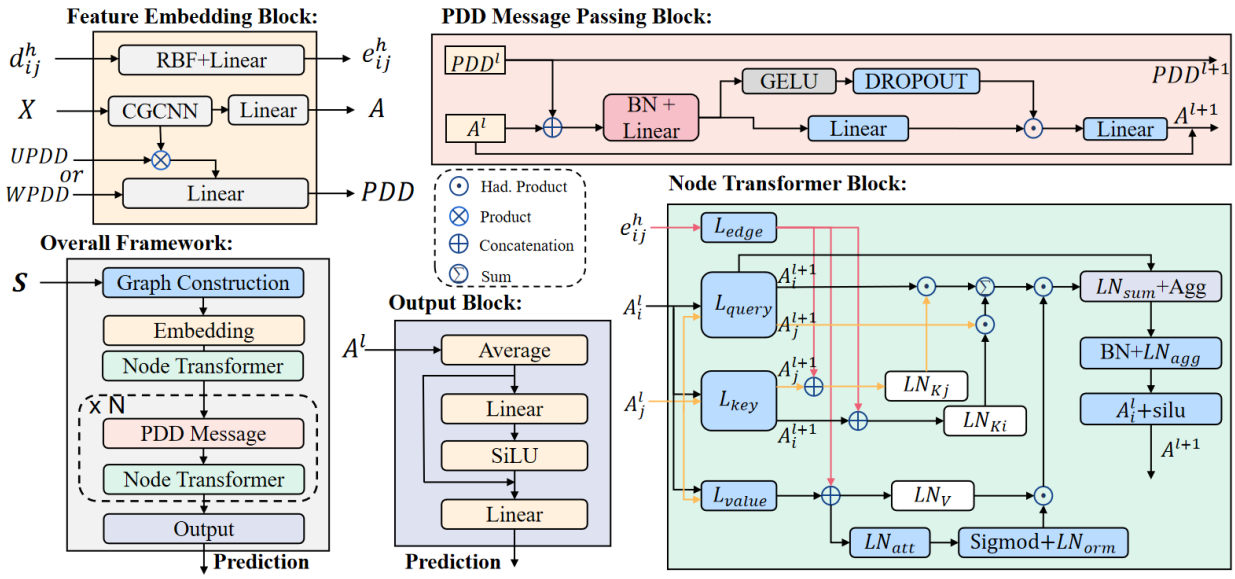

Figure 3: Architecture Overview. PDDFormer accepts an input crystal structure S. During the prediction process, it first undergoes a graph construction step to generate a continuous and general complete crystal graph structure, followed by an embedding block, then multiple blocks of node-wise Transformer and PDD Message Passing, and finally, an output block.

$R_j = (r_{j1}, \ldots, r_{jk})$ of length k, we use the L_∞- distance $|R_i - R_j|_\infty = \max_{l=0,\ldots k} |r_{il} - r_{jl}|_\infty$.

Proposition 1. *The WPDD and UPDD multi-edge crystal graph is continuous.*

Proof. For any neighbor count $k \geq 1$, if the periodic crystal $\mathcal{S}, \mathcal{Q} \in \mathbb{R}^n$ satisfy $d_B(\mathcal{S}, \mathcal{Q}) < r(\mathcal{S})$, then we have:

$$\begin{aligned} &\mathrm{EMD}\left(\mathcal{G}(\mathcal{S}), \mathcal{G}(\mathcal{Q})\right) = \\ &\mathrm{EMD}\left((\mathcal{S}_\mathcal{X}, \mathcal{S}_{\mathcal{XI}}, \mathcal{S}_\mathcal{E}, \mathcal{S}_{\mathcal{PDD}}), (\mathcal{Q}_\mathcal{X}, \mathcal{Q}_{\mathcal{XI}}, \mathcal{Q}_\mathcal{E}, \mathcal{Q}_{\mathcal{PDD}})\right) = \\ &\mathrm{EMD}\left((\mathcal{S}_\mathcal{X}, \mathcal{Q}_\mathcal{X})\right) + \mathrm{EMD}\left((\mathcal{S}_{\mathcal{XI}}, \mathcal{Q}_{\mathcal{XI}})\right) + \\ &\mathrm{EMD}\left((\mathcal{S}_\mathcal{E}, \mathcal{Q}_\mathcal{E})\right) + \mathrm{EMD}\left((\mathcal{S}_{\mathcal{PDD}}, \mathcal{Q}_{\mathcal{PDD}})\right). \end{aligned} \tag{4}$$

Since disturbances only change the positions of atoms and do not alter their types, therefore $\mathrm{EMD}\left((\mathcal{S}_\mathcal{X}, \mathcal{Q}_\mathcal{X})\right) = 0$ and $\mathrm{EMD}\left((\mathcal{S}_{\mathcal{XI}}, \mathcal{Q}_{\mathcal{XI}})\right) = 0$, refer to *Appendix* F for details. So, we obtain $\mathrm{EMD}\left(\mathcal{G}(\mathcal{S}), \mathcal{G}(\mathcal{Q})\right) = \mathrm{EMD}\left((\mathcal{S}_\mathcal{E}, \mathcal{Q}_\mathcal{E})\right) + \mathrm{EMD}\left((\mathcal{S}_{\mathcal{PDD}}, \mathcal{Q}_{\mathcal{PDD}})\right) \leq 2d_B(\mathcal{S}, \mathcal{Q})$. □

The bottleneck distance $d_B(\mathcal{S}, \mathcal{Q}) < r(\mathcal{S})$ is defined as: $d_B(\mathcal{S}, \mathcal{Q}) = \inf_{g:\mathcal{S}\to\mathcal{Q}} \sup_{p\in\mathcal{S}} |p - g(p)|$ and the envelope radius $r(\mathcal{S})$ is the minimum half-distance between any two points in $r(\mathcal{S})$. In other words, $r(\mathcal{S})$ is the maximum radius of non-overlapping open balls centered at all points in $\mathcal{S}$. This implies that any small perturbation in atomic positions under the d_B [Carstens *et al.*, 1999] will lead to minor changes in the distribution of distances between points in the EMD.

Since the EMD between the constructed crystal graphs only relates to Euclidean distance. Euclidean distance itself is continuous, Theorem 1 extends the following fact: for a unit cell structure with two atoms, when the number of neighbors $k = 1$, if we perturb at most two points by ϵ, the change in distance between the two points will be at most 2ϵ. Extending this to n atomic points with k neighbors, if we perturb at most n points by ϵ, the change in distance between n points will be at most $2nk\epsilon$. This aligns with Definition 5, hence, the constructed WPDD and UPDD multi-edge crystal graph is continuous.

4.5 General Geometric Completeness

Proposition 2. *The WPDD multi-edge crystal graph is generally geometrically complete.*

We discuss the limitations of general completeness in *Appendix* G. We prove this by categorical induction. Crystal structures can be classified into stable and unstable types and further divided into the following three categories: 1. Stable Crystal Structures(i.e., where no two crystals can have the same structure with only a difference in atomic types). 2. Unstable Crystal Structures with Differences in Atomic Coordinates. 3. Unstable Crystal Structures with Identical Structures but Differences in Atomic Types.

Proof. Since UPDD is constructed based on the size of the unit cell, when the number of atoms in the unit cell is relatively small, it could theoretically result in different crystal structures, where all atoms have the same Euclidean distances and atom types but inconsistent atomic positions, sharing the same crystal graph representation.

Since the crystal structures of the first and second categories differ, PDD alone can effectively distinguish them. As WPDD includes PDD, different WPDD representations can be constructed for the two crystals, thereby achieving differentiation. For the third category, we incorporate WPDD as global information into the construction of a multi-edge crystal graph and encode atomic information, such that the WPDD crystal graph $\mathcal{G}$ is represented as $\mathcal{G} = (\mathcal{X}, \mathcal{XI}, \mathcal{E}) +$ WPDD, where $\mathcal{X}$ represents atomic information embedded through CGCNN. For any atom in the unit cell, a WPDD

Method	Formation Energy	Bandgap(OPT)	Total Energy	Ehull	Bandgap(MBJ)	Bulk Moduli(Kv)	Shear Moduli(Gv)
	eV/atom	eV	eV/atom	eV	eV	GPa	GPa
PotNet(2023)	0.0294	0.127	0.032	0.055	0.27	10.06	8.883
CrysMMNet(2023)	0.028	0.128	0.034	–	0.278	9.625	<u>8.471</u>
CrysDiff (2024)	0.029	0.131	0.034	0.062	0.287	9.875	9.193
Crystalformer(2024)	0.0306	0.128	0.032	0.046	0.274	–	–
eComFormer(2024)	0.0284	0.124	0.032	0.044	0.28	10.79	9.826
iComFormer(2024)	0.0272	<u>0.122</u>	0.0288	0.047	0.26	9.617	9.098
UPDDFormer	<u>0.0267</u>	<u>0.122</u>	<u>0.0287</u>	<u>0.0406</u>	<u>0.260</u>	<u>9.456</u>	8.738
WPDDFormer	**0.0257**	**0.119**	**0.0276**	**0.0355**	**0.249**	**9.224**	**8.441**

Table 1: Comparison between UPDDFormer, WPDDFormer, and other baselines in terms of test MAE on the JARVIS dataset. The best results are shown in **bold** and the second-best results are shown with <u>underlines</u>.

Method	Formation Energy	Band Gap	Bulk	Shear
	eV/atom	eV	log(GPa)	log(GPa)
PotNet	0.0188	0.204	0.040	0.065
CrysMMNet	0.0200	0.197	0.038	<u>0.062</u>
Crystalformer	0.0186	0.198	0.0377	0.0689
eComFormer	0.01816	0.202	0.0417	0.0729
iComFormer	0.01826	<u>0.193</u>	0.0380	0.0637
UPDDFormer	<u>0.01696</u>	<u>0.189</u>	<u>0.0370</u>	0.0670
WPDDFormer	**0.01604**	**0.187**	**0.0323**	**0.0605**

Table 2: Comparison of test MAE between UPDDFormer, WPDDFormer, and other baselines on the Materials Project dataset.

row vector needs to be constructed along with the corresponding atomic information embedding. This ensures that for any two crystal structures with identical crystal structures but differing atomic types at corresponding coordinates, the $(\mathcal{X}, \mathcal{XI}, \mathcal{E})$ in their WPDD crystal graphs will differ. On the contrary, if two crystals have the same WPDD crystal graph representation, they share the same WPDD and multigraph representations. It indicates that their crystal structures and the atomic information at corresponding coordinates are identical, thus confirming that they are the same crystal. This contradicts our premise. Hence, the proof is complete. Therefore, the proposed identical crystal graph can represent only the same infinite crystal structure. Then, based on Definition 3, we complete the proof of Proposition 2.

Finally, we conclude that the UPDD crystal graph can only guarantee continuity, while the WPDD crystal graph can ensure both continuity and general completeness. □

4.6 Network Architecture

Based on the graph in Section 4.3, we propose the information propagation scheme of PDDFormer. Figure 3 illustrates the overall framework architecture of PDDFormer.

Feature Embedding Block First, we introduce the construction of the graph embedding Block. We use atomic encoding from CGCNN for embedding. For the edge information e_{ij}^h, we employ radial basis functions to encode the distance between two adjacent nodes in the graph, represented by Equation 5, where γ and μ are hyperparameters. For UPDD, due to the varying feature dimensions of UPDD for different crystals, we perform matrix multiplication on UPDD to align the structural information of different crystals, obtaining information for the PDD message passing layer. Thus, we obtain the graph embedding as:

$$\begin{gathered} \mathrm{A} = \mathrm{CGCNN}(\mathcal{X}), \quad e_{ij}^h = \exp\left(-\gamma\left(\frac{\|p_i - p_j\|^2}{\mu}\right)\right), \\ \mathcal{PDD} = \mathcal{UPDD} \otimes \mathrm{A} \quad or \quad \mathcal{WPDD} \end{gathered} \tag{5}$$

Node Transformer Block Building upon the constructed graph, we aggregate the node information. Let a_i^l be the input feature vector of node i at layer l in PDDFormer. The information propagation of layer l is formulated as follows:

$$\begin{gathered} k_i^l = \left(LK\left(a_i^l\right) \oplus LE\left(e_{ij}^h\right)\right), k_j^l = \left(LK\left(a_j^l\right) \oplus LE\left(e_{ij}^h\right)\right), \\ v = \left(LV\left(a_i^l\right) \oplus LV\left(a_j^l\right) \oplus LE\left(e_{ij}^h\right)\right), \; q_i^l = LQ\left(a_i^l\right), \\ q_j^l = LQ\left(a_j^l\right), v_{ij}^l = v \odot \mathrm{Sigmoid}\left(LN_{orm}\left(LN_{att}\left(v\right)\right)\right), \\ att^l = \frac{q_i^l \odot LN_{Kj}\left(k_j^l\right) + q_j^l \odot LN_{Ki}\left(k_i^l\right)}{\sqrt{d_{q_i^l}}}, \\ m_{ij}^h = q_i^l + \mathrm{Sigmoid}\left(\mathrm{BN}\left(att^l\right)\right) \odot LN_V\left(v_{ij}^l\right) \end{gathered} \tag{6}$$

where LQ, LK, LV, LE are the linear transformations for query, key, value, and edge features. LN_K, LN_V are the nonlinear transformations for key and value, including two linear layers and an activation layer in between. LN_{att} represents the linear transformation for updating messages, and $LNorm$ denotes the layer normalization [Ba, 2016] operation. BN denotes the batch normalization layer [Ioffe, 2015], and $d_{q_i^l}$ is the dimension of q_i^l.

Then, we obtain the message M_i^l by aggregating the information from the neighborhood of node i over multiple edges, and A_i^{l+1} is realized as follows :

$$\begin{gathered} M_i^l = \mathrm{BN}\left(\sum_{j \in A_i} \sum_h LN_{sum}\left(m_{ij}^h\right)\right), \\ A_i^{l+1} = \mathrm{SiLU}\left(a_i^l + LN_{agg}\left(M_i^l\right)\right) \end{gathered} \tag{7}$$

where LN_{sum} is the linear transformation used for updating the edge messages.

PDD Message Passing Block $\mathcal{A}^l$ and PDD^l represent the atomic features and 3D periodic pattern encoding at layer l, respectively. Its message-passing mechanism is as follows:

$$\begin{gathered} \mathcal{PDD}^{l+1} = \mathcal{PDD}^l + \mathcal{A}^{l+1}, \; A_1, A_2 = LN_{\mathcal{PDD}}\left(\mathrm{BN}\left(\mathcal{PDD}^{l+1}\right)\right), \\ A^{l+1} = A^l + LN_{A2}\left(LN_{A1}\left(A_1\right) \odot \mathrm{Drop}\left(\mathrm{GELU}\left(A_2\right)\right)\right) \end{gathered} \tag{8}$$

Finally, we use average pooling to aggregate the features of all nodes in the graph, followed by a nonlinear layer, and then

Models	Time/epoch	Total time	Inference time	GPU memory usage	Complexity	Model Parameter
eComformer	97s	13.4h	361.5s	17GB	$O(nk)$	12.4M
iComformer	103s	14.3h	365.6s	12GB	$O(nk)$	5.0M
WPDDFormer	66s	7.4h	240.6s	6GB	$O(nk)$	4.52M

Table 3: Efficiency comparison with ComFormer on the Jarvis Ehull task. We show the training time per epoch, inference time for the whole test set, total training time, time complexity, GPU memory consumption, and total number of parameters. The experiments were conducted using a 3090 RTX 24GB GPU. The data in the table is averaged over three experiments.

Method	Num. Block	Ehull	Band Gap
NO PDD Block	3,0	0.0430	0.194
NO PDD	3,2	0.0426	0.193
UPDDFormer	3,2	0.0406	0.189
WPDDFormer	3,2	**0.0355**	**0.187**

Table 4: Num. Block represents the number of Node transformer blocks and PDD message passing blocks.

a linear layer to obtain the scalar output of the graph as described above. A detailed description of the PDDFormer architecture can be found in $Appendix$ B.

5 Experiments

We conducted experiments on two material benchmark datasets, namely the Materials Project [Chen *et al.*, 2019] and Jarvis [Choudhary *et al.*, 2020] datasets. Detailed descriptions of the datasets can be found in $Appendix$ A. More information about the experimental settings of PDDFormer can be found in $Appendix$ C. Baseline methods include PotNet [Lin *et al.*, 2023], CrysMMNet [Das *et al.*, 2023a], CrysDiff [Song *et al.*, 2024], Crystalformer [Taniai *et al.*, 2024], and ComFormer [Yan *et al.*, 2024a], The complete baselines can be found in $Appendix$ E.1. For all baselines on the datasets, we report the results provided in the cited papers.

5.1 Experimental Results

JARVIS The quantitative results for JARVIS [Choudhary *et al.*, 2020] are shown in Table 1. WPDDformer achieves the best performance across all tasks. UPDD achieved the second-best results. Notably, WPDDFormer outperforms eComFormer by 19% respectively in the Ehull task.

The Materials Project (MP) The experimental results on MP [Chen *et al.*, 2019] are shown in Table 2. WPDDFormer significantly outperforms previous works across all tasks, with a 11.7% improvement over the second-best model in the critical formation energy task and a 14.3% improvement in the bulk moduli task. Additionally, the excellent prediction accuracy of WPDDFormer in the bulk modulus and shear modulus tasks, using only $4,664$ training samples, demonstrates the expressiveness and robustness of WPDD multi-edge crystal graphs under limited training samples. Overall, our methods are compared with 14 existing methods across the two datasets. Our WPDDFormer consistently outperforms all methods in all tasks. Additionally, WPDDFormer shows a significant improvement in prediction accuracy compared to UPPDFormer. This improvement is not only because the WPDD graph structure is complete and continuous, while UPPD can only ensure continuity, but also because UPPD requires dimensional alignment as mentioned in 4.6, which results in some loss of the expression of global information about the unit cell.

Efficiency We compare model efficiency with Comformer under the best configurations, as reported in Table 3. They have a time complexity of $O(nk)$, where n is the number of atoms in the unit cell and k is the average number of neighbors. Compared to ComFormer, WPDDFormer has fewer parameters and achieves significantly superior experimental results with lower computational cost and faster training and inference speeds, demonstrating the remarkable superiority of our method.

5.2 Ablation Studies

In this section, we demonstrate the impact of introducing (W/U)PDD on the representation learning of crystal materials through ablation studies. Specifically, we conducted experiments on the MP and JARVIS datasets, using testing mean absolute error (MAE) as the quantitative evaluation metric, comparing the results for **Band Gap** and **Ehull** tasks, as shown in Table 4. By comparing (W/U)PDDFormer models without PDD message passing blocks to models that retain the PDD message passing blocks but lack (W/U)PDD information, we validated the importance of (W/U)PDD. The results show that compared to models without the PDD message passing blocks, WPDDForemer achieved improvements of 17.4% and 3.6% in the Bulk Moduli and Ehull tasks, respectively. Compared to models that retain only the PDD message passing blocks but lack (W/U)PDD information, we achieved improvements of 16.7% and 3.1% in these two tasks, respectively. More data can be found in $Appendix$ E.3.

6 Conclusion

In this study, we integrated WPDD and UPDD into the representation of crystal structures, achieving a general complete and continuous construction of crystal graphs. This resolves the ambiguity in crystal graph representations for predicting the properties of crystalline materials and bridges the gap between traditional crystal descriptors and dynamic atomic behavior. Experimental results demonstrate the significant advantage of our PDDFormer in various property prediction tasks. Achieving absolute completeness under perturbations is a problem that will be further explored in the future.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (No.42130112, No.42371479), General Program of Shanghai Natural Science Foundation(Grant No.24ZR1419800, No.23ZR1419300), Science and the Technology Commission of Shanghai Municipality (Grant No.22DZ2229004), Beijing Natural Science Foundation (No.QY23187), and Shanghai Frontiers Science Center of Molecule Intelligent Syntheses.

Contribution Statement

M.C., X.T., X.H., J.Y., and X.W. supervised the research project throughout its duration. Z.W. and X.S. conceived the experimental design and jointly developed the model framework. M.C., X.T., X.H., J.Y., S.L., and X.W. provided revisions to the manuscript. The experimental results were collaboratively analyzed by L.S., L.W., Z.W., and X.S. Data interpretation and manuscript writing were carried out by Z.W. and X.S. X.S. and Z.W. are co-first authors, and X.W. is the corresponding author. All authors participated in the discussion of the results and contributed to the final manuscript preparation.

References

[Ahmad *et al.*, 2018] Zeeshan Ahmad, Tian Xie, Chinmay Maheshwari, Jeffrey C Grossman, and Venkatasubramanian Viswanathan. Machine learning enabled computational screening of inorganic solid electrolytes for suppression of dendrite formation in lithium metal anodes. *ACS central science*, 4(8):996–1006, 2018.

[Ba, 2016] Jimmy Lei Ba. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016.

[Balasingham *et al.*, 2022] Jonathan Balasingham, Viktor Zamaraev, and Vitaliy Kurlin. Compact graph representation of molecular crystals using point-wise distance distributions. *arXiv preprint arXiv:2212.11246*, 2022.

[Balasingham *et al.*, 2024] Jonathan Balasingham, Viktor Zamaraev, and Vitaliy Kurlin. Material property prediction using graphs based on generically complete isometry invariants. *Integrating Materials and Manufacturing Innovation*, pages 1–14, 2024.

[Bartók *et al.*, 2013] Albert P Bartók, Risi Kondor, and Gábor Csányi. On representing chemical environments. *Physical Review B—Condensed Matter and Materials Physics*, 87(18):184115, 2013.

[Batzner *et al.*, 2022] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature communications*, 13(1):2453, 2022.

[Carstens *et al.*, 1999] Hans-Georg Carstens, Walter A Deuber, Wolfgang Thumser, and Elke Koppenrade. Geometrical bijections in discrete lattices. *Combinatorics, Probability and Computing*, 8(1-2):109–129, 1999.

[Chen and Ong, 2022] Chi Chen and Shyue Ping Ong. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, 2022.

[Chen *et al.*, 2019] Chi Chen, Weike Ye, Yunxing Zuo, Chen Zheng, and Shyue Ping Ong. Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, 31(9):3564–3572, 2019.

[Choudhary and DeCost, 2021] Kamal Choudhary and Brian DeCost. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7(1):185, 2021.

[Choudhary *et al.*, 2020] Kamal Choudhary, Kevin F Garrity, Andrew CE Reid, Brian DeCost, Adam J Biacchi, Angela R Hight Walker, Zachary Trautt, Jason Hattrick-Simpers, A Gilad Kusne, Andrea Centrone, et al. The joint automated repository for various integrated simulations (jarvis) for data-driven materials design. *npj computational materials*, 6(1):173, 2020.

[Das *et al.*, 2022] Kishalay Das, Bidisha Samanta, Pawan Goyal, Seung-Cheol Lee, Satadeep Bhattacharjee, and Niloy Ganguly. Crysxpp: An explainable property predictor for crystalline materials. *npj Computational Materials*, 8(1):43, 2022.

[Das *et al.*, 2023a] Kishalay Das, Pawan Goyal, Seung-Cheol Lee, Satadeep Bhattacharjee, and Niloy Ganguly. Crysmmnet: multimodal representation for crystal property prediction. In *Uncertainty in Artificial Intelligence*, pages 507–517. PMLR, 2023.

[Das *et al.*, 2023b] Kishalay Das, Bidisha Samanta, Pawan Goyal, Seung-Cheol Lee, Satadeep Bhattacharjee, and Niloy Ganguly. Crysgnn: Distilling pre-trained knowledge to enhance property prediction for crystalline materials. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 7323–7331, 2023.

[Groom *et al.*, 2016] Colin R Groom, Ian J Bruno, Matthew P Lightfoot, and Suzanna C Ward. The cambridge structural database. *Structural Science*, 72(2):171–179, 2016.

[Ioffe, 2015] Sergey Ioffe. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

[Kurlin, 2024] Vitaliy Kurlin. Mathematics of 2-dimensional lattices. *Foundations of Computational Mathematics*, 24(3):805–863, 2024.

[Lin *et al.*, 2023] Yuchao Lin, Keqiang Yan, Youzhi Luo, Yi Liu, Xiaoning Qian, and Shuiwang Ji. Efficient approximations of complete interatomic potentials for crystal property prediction. In *International Conference on Machine Learning*, pages 21260–21287. PMLR, 2023.

[Louis *et al.*, 2020] Steph-Yves Louis, Yong Zhao, Alireza Nasiri, Xiran Wang, Yuqi Song, Fei Liu, and Jianjun Hu. Graph convolutional neural networks with global attention for improved materials property prediction. *Physical Chemistry Chemical Physics*, 22(32):18141–18148, 2020.

[Patterson, 1944] A Lindo Patterson. Ambiguities in the x-ray analysis of crystal structures. *Physical Review*, 65(5-6):195, 1944.

[Rubner *et al.*, 2000] Yossi Rubner, Carlo Tomasi, and Leonidas J Guibas. The earth mover's distance as a metric for image retrieval. *International journal of computer vision*, 40:99–121, 2000.

[Song *et al.*, 2024] Zixing Song, Ziqiao Meng, and Irwin King. A diffusion-based pre-training framework for crystal property prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 8993–9001, 2024.

[Taniai *et al.*, 2024] Tatsunori Taniai, Ryo Igarashi, Yuta Suzuki, Naoya Chiba, Kotaro Saito, Yoshitaka Ushiku, and Kanta Ono. Crystalformer: Infinitely connected attention for periodic structure encoding. In *The Twelfth International Conference on Learning Representations*, 2024.

[Wang *et al.*, 2022] Rui Wang, Robin Walters, and Rose Yu. Approximately equivariant networks for imperfectly symmetric dynamics. In *International Conference on Machine Learning*, pages 23078–23091. PMLR, 2022.

[Wassermann *et al.*, 2010] Anne Mai Wassermann, Mathias Wawer, and Juŕgen Bajorath. Activity landscape representations for structure- activity relationship analysis. *Journal of medicinal chemistry*, 53(23):8209–8223, 2010.

[Widdowson and Kurlin, 2022] Daniel Widdowson and Vitaliy Kurlin. Resolving the data ambiguity for periodic crystals. *Advances in Neural Information Processing Systems*, 35:24625–24638, 2022.

[Widdowson *et al.*, 2022] Daniel Widdowson, Marco M Mosca, Angeles Pulido, Andrew I Cooper, and Vitaliy Kurlin. Average minimum distances of periodic point sets–foundational invariants for mapping periodic crystals. *MATCH Commun. Math. Comput. Chem*, 87(3):529–559, 2022.

[Xie and Grossman, 2018] Tian Xie and Jeffrey C Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical review letters*, 120(14):145301, 2018.

[Yan *et al.*, 2022] Keqiang Yan, Yi Liu, Yuchao Lin, and Shuiwang Ji. Periodic graph transformers for crystal material property prediction. *Advances in Neural Information Processing Systems*, 35:15066–15080, 2022.

[Yan *et al.*, 2024a] Keqiang Yan, Cong Fu, Xiaofeng Qian, Xiaoning Qian, and Shuiwang Ji. Complete and efficient graph transformers for crystal material property prediction. In *The Twelfth International Conference on Learning Representations*, 2024.

[Yan *et al.*, 2024b] Keqiang Yan, Alexandra Saxton, Xiaofeng Qian, Xiaoning Qian, and Shuiwang Ji. A space group symmetry informed network for o(3) equivariant crystal tensor prediction. In *Forty-first International Conference on Machine Learning*, 2024.

[Zwart *et al.*, 2008] Peter H Zwart, Ralf W Grosse-Kunstleve, Andrey A Lebedev, Garib N Murshudov, and Paul D Adams. Surprises and pitfalls arising from (pseudo) symmetry. *Acta Crystallographica Section D: Biological Crystallography*, 64(1):99–107, 2008.