Efficient Approximations of Complete Interatomic Potentials for Crystal Property Prediction

Yuchao Lin [1] **Keqiang Yan** [1] **Youzhi Luo** [1] **Yi Liu** [2] **Xiaoning Qian** [1 3] **Shuiwang Ji** [1]

Abstract

We study property prediction for crystal materials. A crystal structure consists of a minimal unit cell that is repeated infinitely in 3D space. How to accurately represent such repetitive structures in machine learning models remains unresolved. Current methods construct graphs by establishing edges only between nearby nodes, thereby failing to faithfully capture infinite repeating patterns and distant interatomic interactions. In this work, we propose several innovations to overcome these limitations. First, we propose to model physics-principled interatomic potentials directly instead of only using distances as in many existing methods. These potentials include the Coulomb potential, London dispersion potential, and Pauli repulsion potential. Second, we model the complete set of potentials among all atoms, instead of only between nearby atoms as in existing methods. This is enabled by our approximations of infinite potential summations with provable error bounds. We further develop efficient algorithms to compute the approximations. Finally, we propose to incorporate our computations of complete interatomic potentials into message passing neural networks for representation learning. We perform experiments on the JARVIS and Materials Project benchmarks for evaluation. Results show that the use of interatomic potentials and complete interatomic potentials leads to consistent performance improvements with reasonable computational costs. Our code is publicly available as part of the AIRS library (https://github.com/divelab/AIRS).

[1]Department of Computer Science & Engineering, Texas A&M University, College Station, TX, USA [2]Department of Computer Science, Florida State University, Tallahassee, FL, USA [3]Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX, USA. Correspondence to: Shuiwang Ji <sji@tamu.edu>.

Proceedings of the 40th International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023. Copyright 2023 by the author(s).

1. Introduction

The past decade has witnessed a surge of interests and rapid developments in machine learning for molecular analysis (Duvenaud et al., 2015). These initial studies mainly focus on the prediction and generation problems of small molecules. To enable computational analyses, molecules need to be featurized in an appropriate mathematical representation form. Recently, with the advances of graph neural networks (GNNs) (Gilmer et al., 2017; Battaglia et al., 2018; Gao et al., 2021; Liu et al., 2020), molecules are more commonly represented as graphs in which each node corresponds to an atom, and each edge corresponds to a chemical bond (Stokes et al., 2020; Wang et al., 2022b). A variety of molecular graph prediction (Stokes et al., 2020; Wang et al., 2022b) and generation (Shi et al., 2019; Jin et al., 2018; Luo et al., 2021) methods have been developed based on 2D molecular graph representations. A key limitation of the 2D graph representations is that the 3D geometries of molecules are not captured, but such information may be critical in many molecular property prediction problems (Hu et al., 2021). To enable the encoding of 3D molecular geometries in GNNs, a series of 3D GNN methods have been developed for prediction (Schütt et al., 2017; Gasteiger et al., 2019; Liu et al., 2022b; Wang et al., 2022a) and generation (Liu et al., 2022a; Luo & Ji, 2022; Hoogeboom et al., 2022) problems. In these 3D graph representations, each node is associated with the corresponding atom's coordinate in 3D space. Geometric information, such as distances between nodes and angles between edges, is used during message passing in GNNs. Recently, these methods have been extended to learn representations for proteins (Jing et al., 2020; Wang et al., 2023).

Inspired by the success of GNNs on small molecules, (Xie & Grossman, 2018) developed the crystal graph convolutional neural network (CGCNN) for crystal material property prediction. Different from small molecules and proteins, crystal materials are typically modeled by a minimal unit cell (similar to a small molecule) that is repeated in 3D space with certain directions and step sizes. In theory, the unit cell is repeated infinitely in 3D space, but any real-world material has finite size. However, given that our modeling is at the atomic level, modeling crystal materials as infinite repeti-

tions of unit cells is approximately accurate. Therefore, a key challenge in crystal material modeling is how to accurately capture the infinite-range interatomic interactions resulted from the repetitions of unit cells in 3D space. Current GNN-based crystal property prediction methods construct graphs by creating edges only between atoms within a pre-specified distance threshold (Xie & Grossman, 2018; Chen et al., 2019; Louis et al., 2020; Schmidt et al., 2021; Choudhary & DeCost, 2021). Thus, they fail to capture interactions between distant atoms explicitly.

In this work, we propose a new graph deep learning method, **PotNet**, with several innovations to significantly advance the field of crystal material modeling. First, we propose to model interatomic potentials directly as edge features in PotNet, instead of using distance as in prior methods. These potentials include the Coulomb potential (West, 1988), London dispersion potential (Wagner & Schreiner, 2015), and Pauli repulsion potential (Krane, 1991). Second, a distinguishing feature of PotNet is to model the **complete** set of potentials among all atoms, instead of only between nearby atoms as in prior methods. This is enabled by our approximations of infinite potential summations with provable error bounds. We further develop efficient algorithms to compute the approximations. Finally, we propose to incorporate our computations of interatomic potentials and complete interatomic potentials into message passing neural networks for representation learning. We performed comprehensive experiments on the JARVIS and Materials Project benchmarks to evaluate our methods. Results show that the use of interatomic potentials and complete interatomic potentials in our methods leads to consistent performance improvements with reasonable computational costs.

2. Background and Related Work

2.1. Crystal Representation and Property Prediction

A crystal structure can be represented as periodic repetitions of unit cells in the three-dimensional (3D) Euclidean space, where the unit cell contains the smallest repeatable structure of a given crystal. Specifically, let n be the number of atoms in the unit cell, a crystal can be represented as $\boldsymbol{M} = (\boldsymbol{A}, \boldsymbol{L})$. Here, $\boldsymbol{A} = \{\boldsymbol{a}_i\}_{i=1}^n = \{(\boldsymbol{x}_i, \boldsymbol{p}_i)\}_{i=1}^n$ describes one of the unit cell structures of $\boldsymbol{M}$, where $\boldsymbol{x}_i \in \mathbb{R}^b$ and $\boldsymbol{p}_i \in \mathbb{R}^3$ denote the b-dimensional feature vector and the 3D Cartesian coordinates of the i-th atom in the unit cell, respectively. $\boldsymbol{L} = [\boldsymbol{l}_1, \boldsymbol{l}_2, \boldsymbol{l}_3] \in \mathbb{R}^{3\times 3}$ is the lattice matrix describing how a unit cell repeats itself in the 3D space. In the complete crystal structure, every atom in a unit cell repeats itself periodically in the 3D space. Specifically, from an arbitrary integer vector $\boldsymbol{k} \in \mathbb{Z}^3$ and the unit cell structure $\boldsymbol{A}$, we can always obtain another repeated unit cell structure $\boldsymbol{A}^{\boldsymbol{k}} = \{\boldsymbol{a}_i^{\boldsymbol{k}}\}_{i=1}^n = \{(\boldsymbol{x}_i^{\boldsymbol{k}}, \boldsymbol{p}_i^{\boldsymbol{k}})\}_{i=1}^n$, where $\boldsymbol{x}_i^{\boldsymbol{k}} = \boldsymbol{x}_i$, $\boldsymbol{p}_i^{\boldsymbol{k}} = \boldsymbol{p}_i + \boldsymbol{L}\boldsymbol{k}$. Hence, the complete crystal structure $\widetilde{\boldsymbol{A}}$ of $\boldsymbol{M}$ with all unit cells can be described as

$$\widetilde{\boldsymbol{A}} = \bigcup_{\boldsymbol{k}\in\mathbb{Z}^3} \boldsymbol{A}^{\boldsymbol{k}}. \tag{1}$$

In this work, we study the problem of crystal property prediction. Our objective is to learn a property prediction model $f : \boldsymbol{M} \to y \in \mathbb{R}$ that can predict the property y of the given crystal structure $\boldsymbol{M}$. We will focus on predicting the total energy, or other energy-related properties of crystals.

2.2. Crystal Property Prediction with Interatomic Potentials

Most of the classical crystal energy prediction methods are based on interatomic potentials. According to existing studies in physics (West, 1988; Daw et al., 1993; Brown, 2016), the total energy of a crystal structure can be approximated by the summation of interatomic potentials in the crystal. Particularly, the three following categories of interatomic potentials are widely used in crystals, and they can be considered sufficient for accurate energy approximation.

- **Coulomb potential** is caused by the electrostatic interaction of two atoms with charges. Coulomb potentials are closely related to ionic bonding and metallic bonding in crystals (West, 1988). For any two atoms $\boldsymbol{a}$ and $\boldsymbol{b}$, let $z_{\boldsymbol{a}}$ and $z_{\boldsymbol{b}}$ denote the number of charges in the atom $\boldsymbol{a}$ and $\boldsymbol{b}$, and let $d(\boldsymbol{a}, \boldsymbol{b})$ be the Euclidean distance between the atom $\boldsymbol{a}$ and $\boldsymbol{b}$. The Coulomb potential V_{Coulomb} is defined as $V_{\text{Coulomb}}(\boldsymbol{a}, \boldsymbol{b}) = -\frac{z_{\boldsymbol{a}} z_{\boldsymbol{b}} e_0^2}{4\pi\epsilon_0 d(\boldsymbol{a},\boldsymbol{b})}$. Here e_0 is the elementary charge constant, and ϵ_0 is the permittivity constant of free space.
- **London dispersion potential** describes the Van der Waals interaction between atoms. It is often considered in energy estimation since its contribution is cumulative over the volume of crystals (Wagner & Schreiner, 2015) and can be sometimes very strong in bulk crystals, such as sulfur and phosphorus. The mathematical form of this potential is described as $V_{\text{London}}(\boldsymbol{a}, \boldsymbol{b}) = -\epsilon/d^6(\boldsymbol{a}, \boldsymbol{b})$, where ϵ is a hyperparameter.
- **Pauli repulsion potential** results from the Pauli exclusion principle that generally exists in all crystal structures. The Pauli exclusion principle forces any two atoms to be sufficiently far away from each other so that their electron orbits do not overlap. Such exclusion interactions lead to Pauli repulsion potential with the form of $V_{\text{Pauli}}(\boldsymbol{a}, \boldsymbol{b}) = e^{-\alpha d(\boldsymbol{a},\boldsymbol{b})}$, where α is a hyperparameter (Buckingham, 1938; Slater, 1928).

2.3. Crystal Property Prediction with Deep Learning

Physics-based methodologies have long been employed for crystal energy prediction, albeit with a certain degree of

specificity. Typically, these methods are highly specific to a particular type of crystal, implying that a single methodology can only deliver precise approximations for one distinct crystal type. Drawing inspiration from the field of physics, Coulomb matrices, as elucidated by (Rupp et al., 2012; Elton et al., 2018), assume a pivotal function in the prediction of crystal energy. Nevertheless, their application is constrained, primarily modeling a specific subset of materials, namely, ionic and metallic materials. Moreover, a significant limitation of these matrices is their lack of permutation invariance. Recently, thanks to the advances of deep learning, many studies have been done to develop a general crystal property predictor for a variety of different crystals with powerful deep neural network models. Some studies (Behler & Parrinello, 2007; Wang et al., 2021; Jha et al., 2018; 2019; Goodall & Lee, 2020) represent crystals as chemical formulas, and adopt sequence models to predict properties from these string representations. However, more recent studies consider crystals as 3D graphs and employ expressive 3D GNNs (Schütt et al., 2017; Klicpera et al., 2020b; Liu et al., 2022b), a family of deep neural networks specifically designed for 3D graph-structured data, to crystal representation learning. CGCNN (Xie & Grossman, 2018) is the first method that proposes to represent crystals with radius graphs and adopts a graph convolutional network to predict the property from the graph. Based on the pioneering exploration of CGCNN, subsequent studies (Schmidt et al., 2021; Louis et al., 2020; Chen et al., 2019; Choudhary & DeCost, 2021; Batzner et al., 2022; Omee et al., 2022) propose various 3D GNN architectures to achieve more effective crystal representation learning. Particularly, by enhancing the input features with periodic invariance and periodic patterns, Matformer (Yan et al., 2022) develops the currently most powerful 3D GNN architecture for crystals and achieves the best crystal property prediction performance.

3. Method

Although existing GNN-based methods have achieved impressive performance in crystal property prediction, they struggle in further boosting the performance due to the approximation of interatomic interactions using functional expansions based on distances and failing in capturing complete interatomic interactions. In this section, we present PotNet, a novel crystal representation model that can overcome these limitations of prior methods. Based on the physical modeling of crystal energy, PotNet explicitly uses interatomic potentials and complete interatomic potentials as input features to capture complete interatomic interactions. The interatomic potentials and complete interatomic potentials are incorporated into the message passing mechanism of graph neural networks and efficiently approximated by an efficient algorithm.

3.1. Approximating Crystal Energy with Complete Interatomic Potentials

According to the density functional theory (DFT) in physics, for any crystal $\boldsymbol{M} = (\boldsymbol{A}, \boldsymbol{L})$ with the complete structure $\widetilde{\boldsymbol{A}}$ defined in Eqn. (1), its total energy $E(\boldsymbol{M})$ can be approximated by the embedded atom method (Daw & Baskes, 1984; Daw et al., 1993; Baskes, 1987; Lee et al., 2016; Riffe et al., 2018) in the form of

$$E(\boldsymbol{M}) = \frac{1}{2} \sum_{\boldsymbol{a} \in \boldsymbol{A}} \sum_{\boldsymbol{b} \neq \boldsymbol{a}, \boldsymbol{b} \in \widetilde{\boldsymbol{A}}} V(\boldsymbol{a}, \boldsymbol{b}) + \sum_{\boldsymbol{a} \in \boldsymbol{A}} F(\rho_{\boldsymbol{a}}), \quad (2)$$

where $V(\boldsymbol{a}, \boldsymbol{b})$ denotes the interatomic potentials between the atoms $\boldsymbol{a}$ and $\boldsymbol{b}$, capturing the magnitude of interactions; $\rho_{\boldsymbol{a}}$ is the local electron density of the atom $\boldsymbol{a}$, determined by the coordinate and number of charges of the atom $\boldsymbol{a}$ according to the Hohenberg-Kohn theorem; $F(\cdot)$ is a parametrized function to embed the electron density $\rho_{\boldsymbol{a}}$. Actually, existing studies (Jalkanen & Müser, 2015) show that $\rho_{\boldsymbol{a}}$ can be considered as a function of $\sum_{\boldsymbol{b} \neq \boldsymbol{a}, \boldsymbol{b} \in \widetilde{\boldsymbol{A}}} V(\boldsymbol{a}, \boldsymbol{b})$ mathematically[1]. Hence, Eqn. (2) can be rewritten in the following form:

$$E(\boldsymbol{M}) = \sum_{\boldsymbol{a} \in \boldsymbol{A}} \left[\frac{1}{2} \sum_{\boldsymbol{b} \neq \boldsymbol{a}, \boldsymbol{b} \in \widetilde{\boldsymbol{A}}} V(\boldsymbol{a}, \boldsymbol{b}) + G\left(\sum_{\boldsymbol{b} \neq \boldsymbol{a}, \boldsymbol{b} \in \widetilde{\boldsymbol{A}}} V(\boldsymbol{a}, \boldsymbol{b}) \right) \right], \quad (3)$$

where $G(\cdot)$ is a parametrized function. Eqn. (3) can be considered as a way to compute the energy from the complete interatomic potential summation $\sum_{\boldsymbol{b} \neq \boldsymbol{a}, \boldsymbol{b} \in \widetilde{\boldsymbol{A}}} V(\boldsymbol{a}, \boldsymbol{b})$ of every atom $\boldsymbol{a}$ in the unit cell $\boldsymbol{A}$. However, in practice, the function G is computationally expensive if not infeasible. Hence, more and more studies have turned to the powerful learning capability of modern deep neural network models to approximate it effectively.

3.2. Limitations of Existing Deep Learning Methods

Currently, most of the existing graph deep learning methods for crystals (Xie & Grossman, 2018; Chen et al., 2019; Louis et al., 2020; Choudhary & DeCost, 2021) use radius graph representations and distance-based features as inputs to predict the crystal energy in Eqn. (3). Specifically, for a crystal $\boldsymbol{M} = (\boldsymbol{A}, \boldsymbol{L})$, the radius graph is constructed by adding edges between any atom $\boldsymbol{a}$ in the unit cell $\boldsymbol{A}$ and

[1]Under zeroth-order approximation, the electron density $\rho_{\boldsymbol{a}}$ is represented as the aggregate of functions, analogous in type to those used for potential energy calculations. Due to computational efficiency, the approximation form of $e^{-\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2}$ is intentionally excluded from this study. This series type can be computed using the Riemann Theta Function as described in Appendix A.

any other atom $\boldsymbol{b}$ in the complete crystal structure $\widetilde{\boldsymbol{A}}$ whose distances are smaller than a pre-specified distance threshold r. In addition, some functional expansions of distances, e.g., radial basis functions (RBF), are used to model interatomic interactions and form the input edge features to 3D GNN models. Hence, let $\boldsymbol{a} = (\boldsymbol{x_a}, \boldsymbol{p_a}), \boldsymbol{b} = (\boldsymbol{x_b}, \boldsymbol{p_b})$, the crystal energy prediction $\hat{E}(\boldsymbol{M})$ of these methods can be generally described as

$$\hat{E}(\boldsymbol{M}) = \sum_{\boldsymbol{a} \in \boldsymbol{A}} \sum_{\boldsymbol{b} \in \mathcal{N}_r(\boldsymbol{a})} H\left(\phi\left(||\boldsymbol{p_a} - \boldsymbol{p_b}||_2\right)\right), \quad (4)$$

where $\mathcal{N}_r(\boldsymbol{a}) = \{\boldsymbol{b} : \boldsymbol{b} \neq \boldsymbol{a}, \boldsymbol{b} \in \widetilde{\boldsymbol{A}}, ||\boldsymbol{p_a} - \boldsymbol{p_b}||_2 < r\}$, $\phi(\cdot)$ denotes the functional expansions, and $H(\cdot)$ is a non-linear function based on 3D GNN models.

However, we argue that predicting or approximating the energy with Eqn. (4) is a suboptimal solution. Actually, compared with Eqn. (3), which is physics-principled, there exist non-negligible approximation errors in Eqn. (4). First, Eqn. (4) captures the interatomic interactions based on interatomic distances, while the energy can be more accurately approximated by a function of interatomic potentials as in Eqn. (3). Though according to Sec. 2.2, interatomic potentials themselves are also functions of distances, we argue that directly using functional expansions of distances is not the best solution to crystal energy prediction. The commonly used functional expansions in existing methods, such as RBF $\phi(\cdot)$, have different mathematical forms from potentials defined in Sec. 2.2. Intuitively, this poses more challenges to 3D GNN models since they need to learn a mapping from $\phi(\cdot)$ to the energy $\boldsymbol{E}$, while the energy $\boldsymbol{E}$ is not a direct function of $\phi(\cdot)$. Hence, we argue that directly employing the physics-principled potential functions instead of $\phi(\cdot)$ as input features is more suitable for crystal energy prediction.

Second, different from Eqn. (3), Eqn. (4) does not capture the complete set of interatomic interactions because the summation set $\mathcal{N}_r(\boldsymbol{a})$ of atoms $\boldsymbol{b}$ is constrained to be the atoms whose distances to the atom $\boldsymbol{a}$ are smaller than r. This can lead to an approximation error due to ignoring long-range interatomic potentials, i.e., interatomic potentials between distant atoms[2]. By the first principles in physics, interatomic potentials decay algebraically when pairwise interatomic distances become larger. Hence, for a finite structure like molecules, the potentials from atoms that are far away from a given atom are limited and can be ignored. However, long-range interatomic potentials cannot be ignored for crystals since they accumulate infinitely. As a result, long-range interatomic potentials can have a significant effect on a given atom in the infinite crystal structure. Let d be the distance between atoms $\boldsymbol{a}$ and $\boldsymbol{b}$ and p be a positive real number. Considering interatomic potentials $V \propto 1/d^p$, and assuming a 3D crystal structure containing an atom repeating itself with a Euclidean distance of 1, then the energy contribution by considering all its repetitions to it is simply the summation of all these interatomic potentials. To be concrete, the total potential summation $\widetilde{V}$ satisfies $\widetilde{V} \propto \sum_{\boldsymbol{k} \in \mathbb{Z}^3} 1/\|\boldsymbol{k}\|^p$. Considering potentials of the pairwise atoms within the distance threshold r, i.e., a sphere S_r, we have the smallest possible prediction error ΔV satisfying $\Delta V \propto \sum_{\boldsymbol{k} \in \mathbb{Z}^3/S_r} 1/\|\boldsymbol{k}\|^p$. Different from the geometry series, ΔV decays at an approximately algebraical rate rather than exponentially. This suggests that a large radius r is needed to accurately approximate $\widetilde{V}$. Taking London dispersion potential $p = 6$ as an example and it can be calculated that to approximate $\sum_{\boldsymbol{k} \in \mathbb{Z}^3} 1/\|\boldsymbol{k}\|^6 \approx 8.40$ with 0.01 absolute error, we need at least $r = 56$, while in common radius crystal graph construction (Xie & Grossman, 2018; Chen et al., 2019; Choudhary & DeCost, 2021; Louis et al., 2020; Schütt et al., 2017), the radius covers only a unit cell and its neighbors at average (see Appendix D.3), analogy to $r = 1$. In addition, a larger radius will consume much more time for crystal graph construction since it induces a cubic time complexity in the 3D space. We can observe from this example that the failure to capture complete interatomic potentials due to the use of radius is a key factor that prevents accurate energy prediction in existing GNN-based methods. In addition, we experimentally show that large cutoffs will produce better results for classic crystal-graph-based GNNs but consume more processing time in Appendix D.5. To remedy this problem, our efficient algorithm is presented in Sec. 3.4.

3.3. Message Passing with Complete Interatomic Potentials

It follows from the analysis in Sec. 3.2 that major limitations of existing deep learning methods for crystal representation learning lie in: (a) not making predictions from physics-principled interatomic potentials, and (b) not considering complete interatomic interactions. To overcome these limitations, we propose to explicitly use complete interatomic potential summations in GNN models. Since our proposed method is tightly related to potentials, we name it PotNet.

By reformulating Eqn. (3), our PotNet incorporates the crystal energy computation with complete interatomic potentials into the message passing scheme of GNN models. For any material structure $\boldsymbol{M} = (\boldsymbol{A}, \boldsymbol{L})$, we can rewrite the definition of its complete structure $\widetilde{\boldsymbol{A}}$ in Eqn. (1) as

$$\widetilde{\boldsymbol{A}} = \bigcup_{\boldsymbol{k} \in \mathbb{Z}^3} \boldsymbol{A}^{\boldsymbol{k}} = \bigcup_{\boldsymbol{k} \in \mathbb{Z}^3} \bigcup_{\boldsymbol{b} \in \boldsymbol{A}} \{\boldsymbol{b}^{\boldsymbol{k}}\} = \bigcup_{\boldsymbol{b} \in \boldsymbol{A}} \bigcup_{\boldsymbol{k} \in \mathbb{Z}^3} \{\boldsymbol{b}^{\boldsymbol{k}}\} = \bigcup_{\boldsymbol{b} \in \boldsymbol{A}} \boldsymbol{A_b}, \quad (5)$$

where $\boldsymbol{A_b} = \bigcup_{\boldsymbol{k} \in \mathbb{Z}^3} \{\boldsymbol{b}^{\boldsymbol{k}}\}$ denotes the set of atoms contain-

[2]Long-range potential often refers to Coulomb potential. Here, it is used to describe general interatomic potential outside radius r.

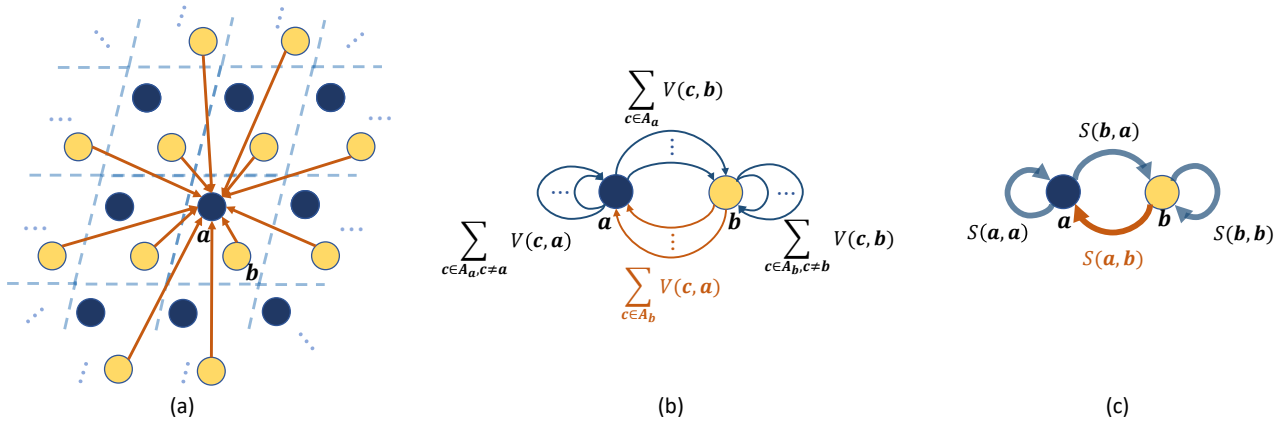

Figure 1. Schematic illustrations of how complete interatomic itneractions are captured in PotNet. Note that PotNet models 3D crystals while we have 2D illustration for simplicity. (a) An example crystal in which each unit cell contains two atoms a and b. In PotNet, the potentials between all pairs of atoms are captured. For simplicity, we only show the potentials from all b atoms to a a atom. (b) The complete set of potentials in (a) can be grouped into four categories, including $a \rightarrow b$, $b \rightarrow a$, $a \rightarrow a$, and $b \rightarrow b$. (c) We propose to compute an approximate summation for each category of potentials.

ing the atom $\boldsymbol{b}$ from the unit cell $\boldsymbol{A}$ and all its periodic repetitions in the complete crystal structure. With Eqn. (5), we can reformulate Eqn. (3) as

$$\begin{aligned} E(\boldsymbol{M}) &= \sum_{\boldsymbol{a}\in\boldsymbol{A}} \left[\frac{1}{2} \sum_{\boldsymbol{b}\in\boldsymbol{A}} \sum_{\boldsymbol{c}\neq\boldsymbol{a}, \boldsymbol{c}\in\boldsymbol{A}_{\boldsymbol{b}}} V(\boldsymbol{a},\boldsymbol{c}) \right. \\ &\qquad \left. + G\left(\sum_{\boldsymbol{b}\in\boldsymbol{A}} \sum_{\boldsymbol{c}\neq\boldsymbol{a}, \boldsymbol{c}\in\boldsymbol{A}_{\boldsymbol{b}}} V(\boldsymbol{a},\boldsymbol{c}) \right) \right] \\ &= \sum_{\boldsymbol{a}\in\boldsymbol{A}} \left[\frac{1}{2} \sum_{\boldsymbol{b}\in\boldsymbol{A}} S(\boldsymbol{a},\boldsymbol{b}) + G\left(\sum_{\boldsymbol{b}\in\boldsymbol{A}} S(\boldsymbol{a},\boldsymbol{b}) \right) \right], \end{aligned} \tag{6}$$

where the infinite potential summation $S(\boldsymbol{a},\boldsymbol{b}) = \sum_{\boldsymbol{c}\neq\boldsymbol{a}, \boldsymbol{c}\in\boldsymbol{A}_{\boldsymbol{b}}} V(\boldsymbol{a},\boldsymbol{c})$ denotes the sum of the interatomic potentials from the atom $\boldsymbol{b}$ together with its all periodic repetitions to the atom $\boldsymbol{a}$. Eqn. (6) can be integrated into the message passing scheme of GNN models. Specifically, we can create a graph $\boldsymbol{G}$ for $\boldsymbol{M} = (\boldsymbol{A}, \boldsymbol{L})$, where each atom in the unit cell $\boldsymbol{A}$ corresponds to a node in the graph. For any two nodes $\boldsymbol{u}, \boldsymbol{v}$ in the graph, there is an edge connecting them, and every node $\boldsymbol{u}$ in the graph is also connected to itself by a self-loop edge. If we consider the infinite potential summation $S(\boldsymbol{a},\boldsymbol{b})$ as the feature of the edge from node $\boldsymbol{b}$ to $\boldsymbol{a}$, we can use the message passing based nonlinear neural network model in GNN to fit the function $\frac{1}{2}\sum_{\boldsymbol{b}\in\boldsymbol{A}} S(\boldsymbol{a},\boldsymbol{b}) + G\left(\sum_{\boldsymbol{b}\in\boldsymbol{A}} S(\boldsymbol{a},\boldsymbol{b})\right)$.

Based on this design of directly using interatomic potentials as edge features, our PotNet employs a GNN model with multiple message passing layers on the graph $\boldsymbol{G}$ to predict the crystal energy of $\boldsymbol{M}$. The computational process of the ℓ-th message passing layer for the node $\boldsymbol{a}$ can be described as

$$\boldsymbol{h}_{\boldsymbol{a}}^{(\ell)} = g_{\varphi}\left(\boldsymbol{h}_{\boldsymbol{a}}^{(\ell-1)}, \sum_{\boldsymbol{b}\in\boldsymbol{A}} f_{\theta}\left(\boldsymbol{h}_{\boldsymbol{a}}^{(\ell-1)}, \boldsymbol{h}_{\boldsymbol{b}}^{(\ell-1)}, S(\boldsymbol{a},\boldsymbol{b}) \right) \right), \tag{7}$$

where $\boldsymbol{h}_{\boldsymbol{a}}^{(\ell)}$ denotes the embedding vector of node $\boldsymbol{a}$ generated from the ℓ-th message passing layer, $\boldsymbol{h}_{\boldsymbol{a}}^{(0)}$ is initialized to the atom feature vector of the atom $\boldsymbol{a}$, and $g_{\varphi}(\cdot)$, $f_{\theta}(\cdot)$ are both neural network models with trainable parameters φ and θ, respectively. Here, the model f_{θ} plays the role of capturing information from both atomic features and complete interatomic potentials. Detailed information about model architectures of f_{θ} and g_{φ} is provided in Appendix D.1. Note that our PotNet is actually a 3D GNN model even though 3D geometric information is not explicitly involved in Eqn. (7). This is because the edge feature $S(\boldsymbol{a},\boldsymbol{b})$ is related to potential functions, and by Sec. 2.2 we know that they are computed from interatomic distances. In other words, PotNet can be considered to encode 3D geometric information with potential functions, though our direct motivation of using potential functions comes from the physical modeling of crystal energy.

Intuitively, the message passing process in Eqn. (7) over the graph $\boldsymbol{G}$ can be considered as a general case of employing a radius graph where the distance threshold r goes to infinity, i.e., $r \rightarrow +\infty$. In this case, as shown in Fig. 1(a), for any atom in the crystal, all the other atoms in the complete crystal structure have been included to interact with it. If we follow the radius graph construction process in the previous methods (Xie & Grossman, 2018; Chen et al., 2019; Louis et al., 2020; Choudhary & DeCost, 2021), we obtain a graph $\widetilde{\boldsymbol{G}}$ in which there exist an infinite number of edges between every pair of nodes. However, PotNet simplifies this complicated graph $\widetilde{\boldsymbol{G}}$ to the graph $\boldsymbol{G}$ in which only one edge exists between every node pair. Specifically, PotNet directly

Table 1. Comparison between our method and other baselines in terms of test MAE on the Materials Project dataset. To make the comparison clear and fair, we follow Yan et al. (2022) and use the same dataset settings. The best results are shown in **bold** and the second best results are shown with underlines.

	Formation Energy	Band Gap	Bulk Moduli	Shear Moduli
Method	eV/atom	eV	log(GPa)	log(GPa)
CGCNN	0.031	0.292	0.047	0.077
SchNet	0.033	0.345	0.066	0.099
MEGNET	0.030	0.307	0.051	0.099
GATGNN	0.033	0.280	0.045	0.075
ALIGNN	0.0221	0.218	0.051	0.078
Matformer	0.0210	0.211	0.043	0.073
PotNet	**0.0188**	**0.204**	**0.040**	**0.065**

models interatomic interactions as potentials and for any two nodes in $\widetilde{\boldsymbol{G}}$, PotNet aggregates all edges between them to a single edge by the use of infinite potential summation $S(\boldsymbol{a}, \boldsymbol{b})$ (see Fig. 1(b)). In other words, PotNet provides an effective solution that enables GNN models to capture complete interatomic interactions through the use of infinite potential summations.

3.4. Efficient Computation of Infinite Potential Summation

Although the infinite potential summations have been effectively incorporated into the message passing based GNN models, the computation of these infinite potential summations is not trivial. Basically, there are two challenges to achieve accurate and efficient computation of the infinite potential summations. For accuracy, the computation algorithm requires provable error bounds. For efficiency, fast algorithm is needed to achieve scalable GNN training and fast crystal property prediction. To tackle these two challenges, we derive a fast approximation algorithm for infinite potential summations based on the Ewald summation method (Ewald, 1921). To be concrete, we unify the summations of three infinite potentials between the position of atom $\boldsymbol{a}$ and all repeated positions of atom $\boldsymbol{b}$ into an integral form, so that the Ewald summation method can be efficiently implemented in PotNet (Fig. 1(c)). The key idea of the Ewald summation is that a slowly converging summation in the real space is guaranteed to be converted into a quickly converging summation in the Fourier space (Woodward, 2014). Based on this, the Ewald summation method divides a summation into two parts. One part has a quicker converging rate in the real space than the original summation. The other slower-to-converge part is then transformed into the Fourier space and becomes quickly convergent. In our method, the Ewald summation method is used with the infinite summations by dividing the integral into two parts, including one part that converges quickly in the Fourier space and another part that converges quickly in the real space, to obtain a fast approximation with provable error bounds.

Following notations in Sec. 2 and Sec. 3.3, we denote the positions of atoms in the set $\boldsymbol{A_b}$ as $\boldsymbol{P_b} = \{\boldsymbol{p_b^k} \,|\, \boldsymbol{p_b^k} = \boldsymbol{p_b} + \boldsymbol{Lk}, \boldsymbol{k} \in \mathbb{Z}^3\}$. The Euclidean distances between the atom $\boldsymbol{a}$ and all atoms in $\boldsymbol{A_b}$ can be represented as $\{d \,|\, d = \|\boldsymbol{p_b} + \boldsymbol{Lk} - \boldsymbol{p_a}\|, \boldsymbol{k} \in \mathbb{Z}^3\}$. We then investigate the three types of potential mentioned in Sec. 2.2. Since charges and constants in the Coulomb potential function can be extracted outside the summation and modeled as part of atom features, we simplify the Coulomb potential function as $V_{\text{Coulomb}}(\boldsymbol{a}, \boldsymbol{b}) = -\epsilon'/d(\boldsymbol{a}, \boldsymbol{b})$, where ϵ' is a hyperparameter scaling the Coulomb potential. As a result, we can represent Coulomb potentials from all atoms in $\boldsymbol{A_b}$ to the atom $\boldsymbol{a}$ as $\{-\frac{\epsilon'}{d} \mid d = \|\boldsymbol{p_b} + \boldsymbol{Lk} - \boldsymbol{p_a}\|, d \neq 0, \boldsymbol{k} \in \mathbb{Z}^3\}$. Similarly, London dispersion potentials from all atoms in $\boldsymbol{A_b}$ to the atom $\boldsymbol{a}$ can be represented as $\{-\frac{\epsilon}{d^6} \mid d = \|\boldsymbol{p_b} + \boldsymbol{Lk} - \boldsymbol{p_a}\|, d \neq 0, \boldsymbol{k} \in \mathbb{Z}^3\}$. It is worth noting that Coulomb and London dispersion potentials can be represented in a unified view as $\{\frac{\text{constant}}{d^p} \mid d = \|\boldsymbol{v_{ab}} + \boldsymbol{Lk}\|, d \neq 0, \boldsymbol{v_{ab}} = \boldsymbol{p_b} - \boldsymbol{p_a}, \boldsymbol{k} \in \mathbb{Z}^3\}$ with a positive real number p. In addition, we represent Pauli potentials from all atoms in $\boldsymbol{A_b}$ to the atom $\boldsymbol{a}$ as $\{e^{-\alpha d} \mid d = \|\boldsymbol{v_{ab}} + \boldsymbol{Lk}\|, \boldsymbol{v_{ab}} = \boldsymbol{p_b} - \boldsymbol{p_a}, \boldsymbol{k} \in \mathbb{Z}^3\}$ with a hyperparameter α. We provide detailed proofs in Appendix C.1 that the summations of these three potentials can be unified in an integral form as

$$S(\boldsymbol{a}, \boldsymbol{b}) = D \int_0^\infty t^{C-1} (-\delta(\boldsymbol{v}, B) + \sum_{\boldsymbol{k} \in \mathbb{Z}^3} e^{-A\pi \|\boldsymbol{Lk} + \boldsymbol{v_{ab}}\|^2 t - \frac{B}{t}}) dt, \quad (8)$$

where A, B, C, D are constants derived from the corresponding specific potential forms and δ is the generalized delta function such that $\delta(\boldsymbol{v}, B) = 1$ if and only if $\boldsymbol{v} = \boldsymbol{0}$ and $B = 0$, otherwise $\delta(\boldsymbol{v}, B) = 0$. We then apply the Ewald summation method (Ewald, 1921) to Eqn. (8) and split it into two parts as

$$S(\boldsymbol{a}, \boldsymbol{b}) = S_{\text{Fourier}}(\boldsymbol{a}, \boldsymbol{b}) + S_{\text{direct}}(\boldsymbol{a}, \boldsymbol{b}), \quad (9)$$

Table 2. Comparison between our method and other baselines in terms of test MAE on JARVIS dataset. The best results are shown in **bold** and the second best results are shown with underlines.

	Formation Energy	Bandgap(OPT)	Total energy	Bandgap(MBJ)	Ehull
Method	eV/atom	eV	eV/atom	eV	eV
CFID	0.14	0.30	0.24	0.53	0.22
CGCNN	0.063	0.20	0.078	0.41	0.17
SchNet	0.045	0.19	0.047	0.43	0.14
MEGNET	0.047	0.145	0.058	0.34	0.084
GATGNN	0.047	0.17	0.056	0.51	0.12
ALIGNN	0.0331	0.142	0.037	0.31	0.076
Matformer	0.0325	0.137	0.035	0.30	0.064
PotNet	**0.0294**	**0.127**	**0.032**	**0.27**	**0.055**

where S_{direct} denotes the part that converges quickly in real space, S_{Fourier} denotes the other part that converges quickly in Fourier space, and the total summation converges as shown by Ewald (1921). We demonstrate in Appendix C.2 that S_{direct} and S_{Fourier} can be represented as sums of incomplete Bessel functions $K_\nu(x, y)$. Rigorous mathematical proofs are supplied in B.2, establishing the convergence of these summations of incomplete Bessel functions and their ability to be approximated with an error that remains within the bounds set by the Gaussian Lattice Sum. The practical application of the proposed summation algorithm is outlined in detail in Appendix C.5.

Notably, the summations of incomplete Bessel functions pertaining to London dispersion potentials and Pauli potentials can be approximated directly. In contrast, the summations related to Coulomb potentials are computed by leveraging established mathematical work (Terras, 1973; Kirsten, 1994), and employing analytic continuation to extend the domain of p, as elucidated in Appendix C.3 and C.4. Considering the crystal system's inherent tendency towards neutrality and equilibrium, we illustrate the application of Coulomb potential summation in Appendix E. It is of significance to note that PotNet stands as the pioneering methodology to apply the incomplete Bessel function for computing the Pauli potential summation, a feat unattainable by previous methods (Crandall, 1998; Lee & Cai, 2009; Nestler et al., 2015). Additionally, our method enables computation of other types of interatomic potential summations, including the Lennard-Jones potential, Morse potential, and screened Coulomb potential, as detailed in Appendix C.6.

4. Experimental Studies

4.1. Experimental Setup

We conduct experiments on two material benchmark datasets, including The Materials Project and JARVIS. Baseline methods include CFID (Choudhary et al., 2018), SchNet (Schütt et al., 2017), CGCNN (Xie & Grossman, 2018), MEGNET (Chen et al., 2019), GATGNN (Louis et al., 2020), ALIGNN (Choudhary & DeCost, 2021), and Matformer (Yan et al., 2022). All PotNet models are trained using the Adam (Kingma & Ba, 2014) optimizer with weight decay (Loshchilov & Hutter, 2017) and one cycle learning rate scheduler (Smith & Topin, 2019) with a learning rate of 0.001, training epoch of 500, and batch size of 64. We use Pytorch and Cython to implement our models. For all tasks on two benchmark datasets, we use one NVIDIA RTX A6000 48G GPU as well as Intel Xeon Gold 6258R CPU for computing. Other detailed configurations of PotNet are provided in Appendix D.1.

In the implementation, PotNet employs both local and infinite crystal graphs, allowing it to encapsulate global, infinite interactions without compromising the fidelity of local interactions. More specifically, for the local crystal graph, we adopt the radius crystal graph as introduced by CGCNN (Xie & Grossman, 2018), albeit with modifications; we substitute Euclidean distances with interatomic potentials to serve as edge features. Given that the impacts of both London dispersion and Pauli potentials are circumscribed within the radius crystal graph, and can be largely disregarded when focusing solely on radius regions, we restrict ourselves to using Coulomb potentials within the radius crystal graph. In contrast, the infinite crystal graph is formulated as detailed in Section 3.3, where we incorporate all forms of interatomic potentials - Coulomb, London dispersion, and Pauli potentials. This approach enables us to capture both local and long-range interactions simultaneously, increasing the model's capacity for crystal representation. In order to enhance our model by leveraging infinite potential features, we have also integrated two supplementary techniques, namely the incorporation of periodic table information and the implementation of transformer operations (Ying et al., 2021; Yan et al., 2022; Rampášek et al., 2022). It should be noted that these methods are not applied in the main body of this paper. They have been exclusively applied and discussed in the supplementary section, as elucidated in Appendix D.2.

Table 3. Training time per epoch, total training time, total testing time, and model complexity compared with ALIGNN and Matformer on JARVIS formation energy prediction.

Method	Time/Epoch	Total Training Time	Total Testing Time	Model Para.
ALIGNN	327 s	27.3 h	156 s	15.4MB
Matformer	64 s	8.9 h	59 s	11.0 MB
Ours	**42 s**	**5.8 h**	**31 s**	**6.7 MB**

4.2. Experimental Results

The Materials Project. We first evaluate PotNet on The Materials Project-2018.6.1, which is a widely used large-scale material benchmark with 69239 crystals. Following previous works (Xie & Grossman, 2018; Chen et al., 2019; Choudhary & DeCost, 2021; Louis et al., 2020; Yan et al., 2022), four crystal properties including formation energy, band gap, bulk moduli, and shear moduli, are used for evaluating our model. We notice that previous works (Xie & Grossman, 2018; Chen et al., 2019; Choudhary & DeCost, 2021; Louis et al., 2020; Schütt et al., 2017) compare with each other using different versions of splitting training, evaluation, and testing datasets with different random seeds. For instance, *the original CGCNN paper only uses 28046 training samples for formation energy prediction, resulting in the original result of 0.039.* To make the comparisons fair, we follow the settings of the previous state-of-the-art (SOTA) Matformer (Yan et al., 2022) for all tasks since they retrain all baselines using the same dataset settings and the same data splits from ALIGNN (Choudhary & DeCost, 2021). Note that most of the retrained results recorded in Yan et al. (2022) are better than their counterparts in their original papers. We present our results in Table 1, where PotNet consistently outperforms other SOTA methods on all four tasks. Furthermore, the impressive results of PotNet for Bulk Moduli and Shear Moduli tasks with only 4664 training samples demonstrate the robustness and adaptive ability of PotNet to the tasks with small training data.

JARVIS Dataset. We then evaluate PotNet on JARVIS-DFT-2021.8.18 3D dataset, a newly released benchmark dataset proposed by Choudhary et al. (2020) with 55722 crystals. We evaluate PotNet on five crystal property prediction tasks including formation energy, bandgap (OPT), bandgap (MBJ), total energy, and Ehull. We follow Matformer (Yan et al., 2022) and use the same training, validation, and test splits for all these tasks, and also use their retrained baseline results. As shown in Table 2, PotNet achieves the best performances on all five tasks consistently. The superior performances of PotNet show the effectiveness of interaction modeling using potentials and explicit modeling of infinite interactions for crystal structures.

Efficiency of PotNet. Beyond the superior modeling capacity for crystals, our PotNet is faster and more efficient than ALIGNN and Matformer. To demonstrate the efficiency of PotNet, we compare PotNet with ALIGNN and Matformer in terms of training time per epoch, total training time, inference time and model parameters for the task of JARVIS formation energy prediction. From Table 3, PotNet is four times faster in terms of total training time and inference time compared with ALIGNN, and also faster than Matformer by 34% and 47% in terms of training time per epoch and inference time, respectively.

Table 4. Prediction time cost of infinite potential summation algorithm on the JARVIS test dataset with 5572 crystals. The prediction time considers both preprocessing and model inference time.

Method	Total Prediction Time	Prediction Time/Per Crystal
ALIGNN	167s	30 ms
Matformer	67s	12ms
PotNet	91s	16ms

We also analyze the time cost of the infinite potential summation algorithm in Table 4. Since there lack baselines for comparison of our infinite potential summation, we compare the prediction time (involving both data preprocessing and model inference time) with the most recent methods ALIGNN and Matformer. To perform preprocessing, unlike previous methods, we need to compute the infinite potential summations besides constructing graphs. However, as shown in Table 4, the mean prediction time of PotNet for a single crystal is at the level of milliseconds, which is similar to ALIGNN and Matformer. Particularly, PotNet has a faster prediction speed than ALIGNN as the latter involves the computing of angles. PotNet is slightly slower than Matformer even though our PotNet needs to compute infinite summations, yet PotNet is much more effective. To provide more details, we further present numerical examples and their corresponding time cost of our infinite potential summation algorithm in Table 6 in Appendix C.5.

4.3. Ablation Studies

In this section, we demonstrate the importance of two core components of PotNet, including interaction modeling using interatomic potentials and infinite potential summations for crystal prediction. We conduct experiments on the JARVIS-DFT 3D formation energy task, and use test MAE for evaluation. In order to highlight the infinite potential summations, a comprehensive ablation study pertaining to various metrics, performed on the JARVIS-DFT 3D dataset, is detailed in Appendix D.4.

Interaction Modeling using Potentials. We demonstrate the importance of interaction modeling using potentials by directly replacing potentials with Euclidean distances used by previous works in our PotNet with exactly the same model architecture. Specifically, we denote PotNet with only local crystal graph as the base model. We use 'Base + Euclidean' to represent the base model with Euclidean distances as edge features and 'Base + Potential' to represent

Table 5. Ablation studies for the effects of adding Coulomb potentials and infinite summation.

Method	JARVIS Formation Energy eV/atom
Base + Euclidean	0.0363
Base + Potential	0.0301
Base + Potential + Infinite	**0.0294**

the base model using Coulomb potentials as edge features. As shown in Table 5, by replacing Euclidean distances with Coulomb potentials, PotNet without considering infinite potential summation already obtains a significant performance gain from 0.0363 to 0.0301, revealing the importance of interaction modeling using potentials in PotNet.

Infinite Potential Summations. The importance of infinite summation of potentials is demonstrated by comparing the previous base models with 'Base + Potential + Infinite', denoting the full PotNet model with infinite summation in infinite crystal graph. It can be seen from Table 5 that by using infinite crystal graphs introduced in Sec. 3.3, the global information of crystal structures is captured, resulting in a performance gain from 0.0301 to 0.0294 for formation energy prediction.

5. Limitation

Our research indicates an encouraging performance enhancement using interatomic potentials and complete interatomic potentials. However, we must acknowledge a limitation: the interatomic potential's inherent incapacity to account for interactions extending beyond two atomic participants. Even though the total potential of a material system can be estimated by the embedded atom method (Daw et al., 1993), which sidesteps the necessity for many-body interactions, explicit consideration of angular and many-body interactions may lead to more precise simulations. This is due to the potential influence of additional atoms on interatomic interactions, which could subsequently modify the total potential outcome. Future studies might gain from integrating many-body interactions into the model, such as modeling the three-body potential and extending this methodology to an infinite summation approach.

6. Conclusion

We study the problem of how to capture infinite-range interatomic potentials in crystal property prediction directly. As a radical departure from prior methods that only consider distances among nearby atoms, we develop a new GNN, PotNet, with a message passing scheme that considers interatomic potentials as well as efficient approximations to capture the complete set of potentials among all atoms. Experiments show that the use of complete potentials leads to consistent performance improvements. Altogether, our work provides a theoretically principled and practically effective framework for crystal modeling. In the future, we expect our approximations may be further improved to obtain a lower error bound. We also expect our algorithms for computing the summation could be further improved.

Acknowledgements

This work was supported in part by National Science Foundation grants IIS-1908220, CCF-1553281, IIS-1812641, DMR-2119103, and IIS-2212419, and National Institutes of Health grant U01AG070112.

References

Baskes, M. Application of the embedded-atom method to covalent materials: a semiempirical potential for silicon. *Physical review letters*, 59(23):2666, 1987.

Bateman, H. *Tables of integral transforms [volumes I & II]*, volume 1. McGraw-Hill Book Company, 1954.

Battaglia, P. W., Hamrick, J. B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., et al. Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*, 2018.

Batzner, S., Musaelian, A., Sun, L., Geiger, M., Mailoa, J. P., Kornbluth, M., Molinari, N., Smidt, T. E., and Kozinsky, B. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature communications*, 13(1):1–11, 2022.

Behler, J. and Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical review letters*, 98(14):146401, 2007.

Bétermin, L., Faulhuber, M., and Steinerberger, S. A variational principle for gaussian lattice sums. *arXiv preprint arXiv:2110.06008*, 2021.

Bolten, M., Fahrenberger, F., Halver, R., Heber, F., Hofmann, M., Kabadshow, I., Lenz, O., Pippig, M., and Sutmann, G. ScaFaCoS, C subroutine library. `http://www.scafacos.de/`. URL http://www.scafacos.de/.

Born, M. Über elektrostatische. gitterpotentiale. *Zeitschrift für Physik*, 7(1):124–140, 1921.

Borwein, D., Borwein, J. M., and Taylor, K. F. Convergence of lattice sums and madelung's constant. *Journal of mathematical physics*, 26(11):2999–3009, 1985.

Borwein, J. and Chan, O.-Y. Uniform bounds for the incomplete complementary gamma function. *Mathematical Inequalities and Applications*, 12:115–121, 2009.

Brown, I. D. *The chemical bond in inorganic chemistry: the bond valence model*, volume 27. Oxford University Press, 2016.

Buckingham, R. A. The classical equation of state of gaseous helium, neon and argon. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 168(933):264–283, 1938.

Chen, C., Ye, W., Zuo, Y., Zheng, C., and Ong, S. P. Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, 31(9): 3564–3572, 2019.

Choudhary, K. and DeCost, B. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7(1):1–8, 2021.

Choudhary, K., DeCost, B., and Tavazza, F. Machine learning with force-field-inspired descriptors for materials: Fast screening and mapping energy landscape. *Physical review materials*, 2(8):083801, 2018.

Choudhary, K., Garrity, K. F., Reid, A. C., DeCost, B., Biacchi, A. J., Hight Walker, A. R., Trautt, Z., Hattrick-Simpers, J., Kusne, A. G., Centrone, A., et al. The joint automated repository for various integrated simulations (jarvis) for data-driven materials design. *npj Computational Materials*, 6(1):1–13, 2020.

Crandall, R. E. Fast evaluation of epstein zeta functions. *manuscript*, 1998.

Crandall, R. E. and Buhler, J. P. Elementary function expansions for madelung constants. *Journal of Physics A: Mathematical and General*, 20(16):5497, 1987.

Daw, M. S. and Baskes, M. I. Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals. *Physical Review B*, 29(12):6443, 1984.

Daw, M. S., Foiles, S. M., and Baskes, M. I. The embedded-atom method: a review of theory and applications. *Materials Science Reports*, 9(7-8):251–310, 1993.

Deconinck, B., Heil, M., Bobenko, A., Van Hoeij, M., and Schmies, M. Computing riemann theta functions. *Mathematics of Computation*, 73(247):1417–1442, 2004.

Duvenaud, D., Maclaurin, D., Aguilera-Iparraguirre, J., Gómez-Bombarelli, R., Hirzel, T., Aspuru-Guzik, A., and Adams, R. P. Convolutional networks on graphs for learning molecular fingerprints. In *Proceedings of the 28th International Conference on Neural Information Processing Systems*, pp. 2224–2232, 2015.

Elton, D. C., Boukouvalas, Z., Butrico, M. S., Fuge, M. D., and Chung, P. W. Applying machine learning techniques to predict the properties of energetic materials. *Scientific reports*, 8(1):1–12, 2018.

Ewald, P. P. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der physik*, 369(3):253–287, 1921.

Galassi, M., Davies, J., Theiler, J., Gough, B., Jungman, G., Alken, P., Booth, M., Rossi, F., and Ulerich, R. *GNU scientific library*. Network Theory Limited Godalming, 2002.

Gao, H., Liu, Y., and Ji, S. Topology-aware graph pooling networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(12):4512–4518, 2021.

Gasteiger, J., Groß, J., and Günnemann, S. Directional message passing for molecular graphs. In *International Conference on Learning Representations*, 2019.

Gasteiger, J., Becker, F., and Günnemann, S. Gemnet: Universal directional graph neural networks for molecules. *Advances in Neural Information Processing Systems*, 34: 6790–6802, 2021.

Gaudreau, P., Slevinsky, R. M., and Safouhi, H. Computation of tail probabilities via extrapolation methods and connection with rational and padé approximants. *SIAM Journal on Scientific Computing*, 34(1):B65–B85, 2012.

Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. Neural message passing for quantum chemistry. In *International conference on machine learning*, pp. 1263–1272. PMLR, 2017.

Goodall, R. E. and Lee, A. A. Predicting materials properties without crystal structure: Deep representation learning from stoichiometry. *Nature communications*, 11(1): 1–9, 2020.

Harris, F. E. Incomplete bessel, generalized incomplete gamma, or leaky aquifer functions. *Journal of Computational and Applied Mathematics*, 215(1):260–269, 2008.

Hoogeboom, E., Satorras, V. G., Vignac, C., and Welling, M. Equivariant diffusion for molecule generation in 3D. In *International Conference on Machine Learning*, pp. 8867–8887. PMLR, 2022.

Hu, W., Fey, M., Ren, H., Nakata, M., Dong, Y., and Leskovec, J. OGB-LSC: A large-scale challenge for machine learning on graphs. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2021.

Jalkanen, J. and Müser, M. H. Systematic analysis and modification of embedded-atom potentials: case study of copper. *Modelling and simulation in materials science and engineering*, 23(7):074001, 2015.

Jha, D., Ward, L., Paul, A., Liao, W.-k., Choudhary, A., Wolverton, C., and Agrawal, A. Elemnet: Deep learning the chemistry of materials from only elemental composition. *Scientific reports*, 8(1):1–13, 2018.

Jha, D., Ward, L., Yang, Z., Wolverton, C., Foster, I., Liao, W.-k., Choudhary, A., and Agrawal, A. IRNet: A general purpose deep residual regression framework for materials discovery. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2385–2393, 2019.

Jin, W., Barzilay, R., and Jaakkola, T. Junction tree variational autoencoder for molecular graph generation. In *International conference on machine learning*, pp. 2323–2332. PMLR, 2018.

Jing, B., Eismann, S., Suriana, P., Townshend, R. J. L., and Dror, R. Learning from protein structure with geometric vector perceptrons. In *International Conference on Learning Representations*, 2020.

Jones, D. Incomplete bessel functions. i. *Proceedings of the Edinburgh Mathematical Society*, 50(1):173–183, 2007.

Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Kirichenko, E. and Stephanovich, V. The influence of coulomb interaction screening on the excitons in disordered two-dimensional insulators. *Scientific Reports*, 11 (1):1–14, 2021.

Kirsten, K. Generalized multidimensional epstein zeta functions. *Journal of Mathematical Physics*, 35(1):459–470, 1994.

Klicpera, J., Giri, S., Margraf, J. T., and Günnemann, S. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. *arXiv preprint arXiv:2011.14115*, 2020a.

Klicpera, J., Groß, J., and Günnemann, S. Directional message passing for molecular graphs. *arXiv preprint arXiv:2003.03123*, 2020b.

Krane, K. S. *Introductory nuclear physics*. John Wiley & Sons, 1991.

Kung, J. P. and Yang, C.-C. Complex analysis. *Encyclopedia of Physical Science and Technology*, 2003.

Lee, E., Lee, K.-R., Baskes, M., and Lee, B.-J. A modified embedded-atom method interatomic potential for ionic systems: 2 nnmeam+ qeq. *Physical Review B*, 93(14): 144110, 2016.

Lee, H. and Cai, W. Ewald summation for coulomb interactions in a periodic supercell. *Lecture Notes, Stanford University*, 3(1):1–12, 2009.

Lennard-Jones, J. and Dent, B. M. Cohesion at a crystal surface. *Transactions of the Faraday Society*, 24:92–108, 1928.

Liu, M., Luo, Y., Uchino, K., Maruhashi, K., and Ji, S. Generating 3D molecules for target protein binding. In *Proceedings of The 39th International Conference on Machine Learning*, pp. 13912–13924, 2022a.

Liu, Y., Yuan, H., Cai, L., and Ji, S. Deep learning of high-order interactions for protein interface prediction. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 679–687, 2020.

Liu, Y., Wang, L., Liu, M., Lin, Y., Zhang, X., Oztekin, B., and Ji, S. Spherical message passing for 3d molecular graphs. In *International Conference on Learning Representations*, 2022b.

Loshchilov, I. and Hutter, F. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

Louis, S.-Y., Zhao, Y., Nasiri, A., Wang, X., Song, Y., Liu, F., and Hu, J. Graph convolutional neural networks with global attention for improved materials property prediction. *Physical Chemistry Chemical Physics*, 22(32): 18141–18148, 2020.

Luo, Y. and Ji, S. An autoregressive flow model for 3D molecular geometry generation from scratch. In *International Conference on Learning Representations*, 2022.

Luo, Y., Yan, K., and Ji, S. GraphDF: A discrete flow model for molecular graph generation. In *Proceedings of The 38th International Conference on Machine Learning*, pp. 7192–7203, 2021.

Morse, P. M. Diatomic molecules according to the wave mechanics. ii. vibrational levels. *Physical review*, 34(1): 57, 1929.

Nestler, F., Pippig, M., and Potts, D. Fast ewald summation based on nfft with mixed periodicity. *Journal of Computational Physics*, 285:280–315, 2015.

Omee, S. S., Louis, S.-Y., Fu, N., Wei, L., Dey, S., Dong, R., Li, Q., and Hu, J. Scalable deeper graph neural networks for high-performance materials property

prediction. *Patterns*, 3(5):100491, 2022. ISSN 2666-3899. doi: https://doi.org/10.1016/j.patter.2022.100491. URL https://www.sciencedirect.com/science/article/pii/S2666389922000769.

Rampášek, L., Galkin, M., Dwivedi, V. P., Luu, A. T., Wolf, G., and Beaini, D. Recipe for a general, powerful, scalable graph transformer. *Advances in Neural Information Processing Systems*, 35:14501–14515, 2022.

Riffe, D., Christensen, J. D., and Wilson, R. Vibrational dynamics within the embedded-atom-method formalism and the relationship to born–von-kármán force constants. *Journal of Physics: Condensed Matter*, 30(45):455702, 2018.

Rupp, M., Tkatchenko, A., Müller, K.-R., and Von Lilienfeld, O. A. Fast and accurate modeling of molecular atomization energies with machine learning. *Physical review letters*, 108(5):058301, 2012.

Schmidt, J., Pettersson, L., Verdozzi, C., Botti, S., and Marques, M. A. L. Crystal graph attention networks for the prediction of stable materials. *Science Advances*, 7 (49):eabi7948, 2021. doi: 10.1126/sciadv.abi7948.

Schütt, K., Kindermans, P.-J., Sauceda Felix, H. E., Chmiela, S., Tkatchenko, A., and Müller, K.-R. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in neural information processing systems*, 30, 2017.

Schütt, K., Unke, O., and Gastegger, M. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International Conference on Machine Learning*, pp. 9377–9388. PMLR, 2021.

Selberg, A. and Chowla, S. On epstein's zeta-function. 1967.

Shi, C., Xu, M., Zhu, Z., Zhang, W., Zhang, M., and Tang, J. GraphAF: a flow-based autoregressive model for molecular graph generation. In *International Conference on Learning Representations*, 2019.

Slater, J. C. The normal state of helium. *Physical Review*, 32(3):349, 1928.

Slevinsky, R. M. and Safouhi, H. A recursive algorithm for an efficient and accurate computation of incomplete bessel functions. *arXiv preprint arXiv:2204.11197*, 2022.

Smith, L. N. and Topin, N. Super-convergence: Very fast training of neural networks using large learning rates. In *Artificial intelligence and machine learning for multi-domain operations applications*, volume 11006, pp. 369–386. SPIE, 2019.

Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., MacNair, C. R., French, S., Carfrae, L. A., Bloom-Ackermann, Z., et al. A deep learning approach to antibiotic discovery. *Cell*, 180(4): 688–702, 2020.

Terras, A. A. Bessel series expansions of the epstein zeta function and the functional equation. *Transactions of the American Mathematical Society*, 183:477–486, 1973.

Wagner, J. P. and Schreiner, P. R. London dispersion in molecular chemistry—reconsidering steric effects. *Angewandte Chemie International Edition*, 54(42):12274–12296, 2015.

Wang, A. Y.-T., Kauwe, S. K., Murdock, R. J., and Sparks, T. D. Compositionally restricted attention-based network for materials property predictions. *Npj Computational Materials*, 7(1):1–10, 2021.

Wang, L., Liu, Y., Lin, Y., Liu, H., and Ji, S. ComENet: Towards complete and efficient message passing for 3D molecular graphs. In *The 36th Annual Conference on Neural Information Processing Systems*, 2022a.

Wang, L., Liu, H., Liu, Y., Kurtin, J., and Ji, S. Learning hierarchical protein representations via complete 3d graph networks. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=9X-hgLDLYkQ.

Wang, Z., Liu, M., Luo, Y., Xu, Z., Xie, Y., Wang, L., Cai, L., Qi, Q., Yuan, Z., Yang, T., and Ji, S. Advanced graph and sequence neural networks for molecular property prediction and drug discovery. *Bioinformatics*, 38(9): 2579–2586, 2022b.

West, A. R. *Basic Solid State Chemistry*. Wiley, 1988.

Woodward, P. M. *Probability and information theory, with applications to radar: international series of monographs on electronics and instrumentation*, volume 3. Elsevier, 2014.

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

Yan, K., Liu, Y., Lin, Y., and Ji, S. Periodic graph transformers for crystal material property prediction. In *The 36th Annual Conference on Neural Information Processing Systems*, 2022.

Ying, C., Cai, T., Luo, S., Zheng, S., Ke, G., He, D., Shen, Y., and Liu, T.-Y. Do transformers really perform badly for graph representation? In *NeurIPS 2021*, December 2021.

A. Gaussian Lattice Sum

Let $\boldsymbol{L} \in \mathbb{R}^{d\times d}$ be the full rank lattice matrix and $\boldsymbol{v} \in \mathbb{R}^d$. By extending the definition of *Gaussian Lattice Sum* (Bétermin et al., 2021), we consider a summation of d-dimensional Gaussian functions on a given shifted lattice,

$$E(\boldsymbol{L}, \boldsymbol{v}, \alpha) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-\pi\alpha\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2}, \quad \alpha > 0. \tag{10}$$

One of the characteristics of the Gaussian Lattice Sum is the term $e^{-\pi\alpha\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2}$ rapidly decays as $\boldsymbol{k}$ becomes large, leading to a fast converging rate of $E(\boldsymbol{L}, \boldsymbol{v}, \alpha)$. By simple derivation, Gaussian Lattice Sum can be written w.r.t *Riemann Theta Function*,

$$E(\boldsymbol{L}, \boldsymbol{v}, \alpha) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-\pi\alpha\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2} = \theta\left(i\alpha\boldsymbol{L}^T\boldsymbol{v}|i\alpha\boldsymbol{L}^T\boldsymbol{L}\right) e^{-\pi\alpha\boldsymbol{v}^T\boldsymbol{v}}, \tag{11}$$

where $\theta(\boldsymbol{z}|\boldsymbol{\Omega}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{2\pi i(\frac{1}{2}\boldsymbol{k}\cdot\boldsymbol{\Omega}\cdot\boldsymbol{k}+\boldsymbol{k}\boldsymbol{z})}$ denotes *Riemann Theta Function* (Deconinck et al., 2004), and it is easy to verify that $\Im(\boldsymbol{\Omega}) = \alpha\boldsymbol{L}^T\boldsymbol{L}$ is positive definite that guarantees convergence of Gaussian Lattice Sum. Deconinck et al. (2004) further shows several upper bounds. In particular, given $R > 0$,

$$\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|\geq R} \|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^p e^{-\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2} \leq \frac{d+p}{2}\left(\frac{2}{\rho}\right)^d \Gamma\left(\frac{d}{2}, \left(R-\frac{\rho}{2}\right)^2\right), \tag{12}$$

where $\Gamma(z, x) = \int_x^\infty t^{z-1}e^{-t}dt$ is the incomplete Gamma function and $\rho = \min\{\|\boldsymbol{L}\boldsymbol{k}\| \mid \boldsymbol{k} \in \mathbb{Z}^d, \boldsymbol{k} \neq \boldsymbol{0}\}$. Let $p = 0$ and a lattice matrix be $\sqrt{\pi\alpha}L$. We obtain

$$\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \sqrt{\pi\alpha}\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|\geq R} e^{-\pi\alpha\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2} \leq \frac{d}{2}\left(\frac{2}{\rho}\right)^d \Gamma\left(\frac{d}{2}, \left(R-\frac{\rho}{2}\right)^2\right), \tag{13}$$

where $\rho = \min\{\sqrt{\pi\alpha}\|\boldsymbol{L}\boldsymbol{k}\| \mid \boldsymbol{k} \in \mathbb{Z}^d, \boldsymbol{k} \neq \boldsymbol{0}\}$. In addition, given $R = 0$, we can obtain the upper bound of the Gaussian Lattice Sum by incomplete Gamma function,

$$\sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-\pi\alpha\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2} \leq \frac{d}{2}\left(\frac{2}{\rho}\right)^d \Gamma\left(\frac{d}{2}, \left(\frac{\rho}{2}\right)^2\right). \tag{14}$$

B. Incomplete Bessel Function

Let $x, y, \nu \in \mathbb{R}$, $x > 0$, $y \geq 0$, $\nu > 0$. The *incomplete Bessel function* (Harris, 2008) has the below integral form

$$K_\nu(x, y) = \int_1^\infty t^{-\nu-1}e^{-xt-y/t}dt, \tag{15}$$

In existing studies (Harris, 2008; Slevinsky & Safouhi, 2022; Jones, 2007), $K_\nu(x, y)$ is analytically continued to $\nu \in \mathbb{R}$. In this work, we follow these studies and also use the analytic continuation for our calculations, i.e., we also consider $\nu \in \mathbb{R}$.

B.1. Fast Approximation of the Incomplete Bessel Function

Computing the incomplete Bessel function is challenging as there does not exist direct closed-form solution. In this work, we investigate a fast approximation of the incomplete Bessel function. Concretely, we adopt the algorithm in work by Nestler et al. (2015); Slevinsky & Safouhi (2022), where the incomplete Bessel function $K_\nu(x, y)$ is approximated by the $G_n^{(1)}$ transformation with the linear time complexity of $O(n)$, and n here is the number of iterations. To begin with, the approximation $G_n^{(m)}$ to $\int_0^\infty f(t)dt$ is given as a solution of

$$\frac{d^l}{dx^l}\left\{G_n^{(m)} - \int_0^x f(t)dt - \sum_{k=0}^{m-1} x^{\sigma_k} f^{(k)}(x) \sum_{i=0}^{n-1} \frac{\beta_{i,k}}{x^i}\right\} = 0, \quad l = 0, 1, \cdots, mn, \tag{16}$$

where it is assumed that $\frac{d^l}{dx^l}G_n^{(m)} \equiv 0, \forall l > 0$, $\sigma_k = \min(s_k, k+1)$, and s_k is the largest of $s \in \mathbb{Z}$ such that $\lim_{x\to\infty} x^s f^{(k)}(x) = 0$ for $k = 0, 1, \cdots, m-1$. And we aim to obtain the unknown coefficients $\beta_{i,k}$ to approximate $\int_x^\infty f(t)dt$ from these $mn+1$ linear equations. By considering $l = 0$ in the Eqn. (16), we obtain

$$G_n^{(1)} - \int_0^x f(t)dt = x^{\sigma_0} f(x) \sum_{i=0}^{n-1} \frac{\beta_{i,k}}{x^i}. \tag{17}$$

To eliminate the summation, Slevinsky & Safouhi (2022) applied the $(x^2 \frac{d}{dx})$ operator n times such that

$$(x^2 \frac{d}{dx})^n \left[\frac{G_n^{(1)} - \int_0^x f(t)dt}{x^{\sigma_0} f(x)} \right] = 0. \tag{18}$$

By doing this,

$$G_n^{(1)} = \frac{(x^2 \frac{d}{dx})^n (\frac{\int_0^x f(t)dt}{x^{\sigma_0} f(x)})}{(x^2 \frac{d}{dx})^n (\frac{1}{x^{\sigma_0} f(x)})} = \frac{\mathcal{N}_n(x)}{\mathcal{D}_n(x)}, \tag{19}$$

where it sets

$$\mathcal{N}_n(x) = (x^2 \frac{d}{dx})\mathcal{N}_{n-1}(x) \quad \text{and} \quad \mathcal{D}_n(x) = (x^2 \frac{d}{dx})\mathcal{G}_{n-1}(x), \tag{20}$$

and

$$\mathcal{N}_0(x) = \frac{\int_0^x f(t)dt}{x^{\sigma_0} f(x)} \quad \text{and} \quad \mathcal{D}_0(x) = \frac{1}{x^{\sigma_0} f(x)}. \tag{21}$$

This leads to a recursive algorithm for $G_n^{(1)}$ transformation. To compute the incomplete Bessel function $K_\nu(x, y)$, Slevinsky & Safouhi (2022) investigated the following property

$$K_\nu(x, y) + x^\nu \int_0^x t^{-\nu-1} e^{-t-\frac{xy}{t}} dt = x^\nu \int_0^\infty t^{-\nu-1} e^{-t-\frac{xy}{t}} dt, \tag{22}$$

where the term $\int_0^\infty t^{-\nu-1} e^{-t-\frac{xy}{t}} dt$ can be approximated by $G_n^{(1)}$ transformation. Therefore, to approximate $K_\nu(x, y)$, we have corresponding approximation

$$\begin{aligned} \tilde{G}_n^{(1)} &= x^\nu (G_n^{(1)} - \int_0^x t^{-\nu-1} e^{-t-\frac{xy}{t}} dt) \\ &= x^\nu \frac{\mathcal{N}_n(x, y, \nu) - \mathcal{D}_n(x, y, \nu) \int_0^x t^{-\nu-1} e^{-t-\frac{xy}{t}} dt}{\mathcal{D}_n(x, y, \nu)} \\ &= x^\nu \frac{\tilde{\mathcal{N}}_n(x, y, \nu)}{\mathcal{D}_n(x, y, \nu)}. \end{aligned} \tag{23}$$

As a result, the approximation $\tilde{G}_n^{(1)}$ to $K_\nu(x, y)$ is obtained by recursively solving $\tilde{\mathcal{N}}_n(x, y, \nu)$ and $\mathcal{D}_n(x, y, \nu)$ (Slevinsky & Safouhi, 2022; Gaudreau et al., 2012). The detailed expressions of $\tilde{\mathcal{N}}_n(x)$ and $\mathcal{D}_n(x)$ are given in Slevinsky & Safouhi (2022) Proposition 2.2. In addition, Nestler et al. (2015) further optimizes the approximation of the incomplete Bessel function when $\nu = 0$ and x, y are both small. In this case, the remaining part of the Taylor expansion of $K_0(x, y)$ is small and we can approximate $K_0(x, y)$ by the first m terms of the Taylor series such that

$$K_0(x, y) \approx \sum_{n=0}^{m} \frac{(-1)^n}{n!} x^n y^n \Gamma(-n, x). \tag{24}$$

The detailed error bound by this expansion is shown in Nestler et al. (2015).

B.2. Convergence of Incomplete Bessel Function Summation

Let $\nu, \alpha, \beta, \gamma \in \mathbb{R}$ and be constants. Given a d-dimensional lattice, we define the *incomplete Bessel function summation* on the lattice as

$$\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma>0} K_\nu(\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma, \beta), \tag{25}$$

where $\alpha > 0$, $\beta \geq 0$, $\gamma \geq 0$, $\boldsymbol{v} \in \mathbb{R}^d$ and $\boldsymbol{L} \in \mathbb{R}^{d\times d}$ denoting the full rank lattice matrix. We aim to prove this summation of incomplete Bessel functions converges and can be approximated with an error bounded by the Gaussian Lattice Sum as introduced in Appendix A. And this convergence property can be easily extended to the summation with $e^{2\pi i\zeta}$ coefficients since $|e^{2\pi i\zeta}K_\nu(x, y)| \leq |K_\nu(x, y)|$.

Proof. An upper bound of the incomplete Bessel function can be derived as

$$\big|K_\nu(x, y)\big| = \Big|\int_1^\infty t^{-\nu-1}e^{-xt-y/t}dt\Big| \leq \Big|\int_1^\infty t^{-\nu-1}e^{-xt}dt\Big| = \big|x^{-\nu}\Gamma(-\nu, x)\big|, \tag{26}$$

where Γ is the incomplete Gamma function described in Appendix A. Based on this, we obtain

$$\big|K_\nu(\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma, \beta)\big| \leq \Big|\frac{\Gamma(-\nu, \pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma)}{(\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma)^\nu}\Big|. \tag{27}$$

Given $x > 0$, $|\frac{\Gamma(z,x)}{x^z}|$ has an upper bound (Borwein & Chan, 2009) such that

$$\begin{aligned}\Big|\frac{\Gamma(z, x)}{x^z}\Big| &= \Big|e^{-x}\int_0^\infty e^{-xs}(1+s)^{z-1}ds\Big| \\ &\leq \begin{cases} \frac{e^{-x}}{x-z+1}, & z > 1 \\ \frac{e^{-x}}{x}, & z \leq 1 \end{cases}.\end{aligned} \tag{28}$$

Considering a prefixed value $R \in \mathbb{R}$, $R^2 > -\nu$, $R^2 > 1$ and $R^2 > \gamma$, we have

$$\begin{aligned}\epsilon(R) &= \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \pi\alpha|\boldsymbol{Lk}+\boldsymbol{v}|^2+\gamma\geq R^2} K_\nu(\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma, \beta) \\ &\leq \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma\geq R^2} |K_\nu(\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma, \beta)| \\ &\leq \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma\geq R^2} e^{-\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2-\gamma} \\ &\leq e^{-\gamma}E(\boldsymbol{L}, \boldsymbol{v}, \alpha) \\ &\leq E(\boldsymbol{L}, \boldsymbol{v}, \alpha),\end{aligned} \tag{29}$$

where $E(\boldsymbol{L}, \boldsymbol{v}, \alpha)$ is the Gaussian Lattice Sum described in Appendix A. Therefore, the incomplete Bessel function summation can be divided into two parts

$$\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma>0} K_\nu(\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma, \beta) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d, 0<\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma\leq R^2} K_\nu(\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma, \beta) + \epsilon(R), \tag{30}$$

where the first part is a finite part inside an ellipsoid with a size of $\sqrt{R^2/\pi\alpha - \gamma/\pi\alpha}$, and the second part $\epsilon(R)$ is bounded by the Gaussian Lattice Sum $E(\boldsymbol{L}, \boldsymbol{v}, \alpha)$ and is convergent. Therefore, the incomplete Bessel function summation is convergent. Consequently, to approximate the incomplete Bessel function summation, we can choose to evaluate the summation inside an ellipsoid with the size of $\sqrt{R^2/\pi\alpha - \gamma/\pi\alpha}$ for a prefixed $R \in \mathbb{R}$, such that $R^2 > -\nu$, $R^2 > 1$ and $R^2 > \gamma$. Then the error $\epsilon(R)$ is bounded by Gaussian Lattice Sum $E(\boldsymbol{L}, \boldsymbol{v}, \alpha)$. We can further bound the error by the inequality (13) such that

$$\epsilon(R) \leq \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \sqrt{\pi\alpha}\|\boldsymbol{Lk}+\boldsymbol{v}\|\geq\sqrt{R^2-\gamma}} e^{-\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2} \leq \frac{d}{2}(\frac{2}{\rho})^d\Gamma(\frac{d}{2}, (\sqrt{R^2-\gamma} - \frac{\rho}{2})^2), \tag{31}$$

where $\rho = \min\{\sqrt{\pi\alpha}\|\boldsymbol{Lk}\| \mid \boldsymbol{k} \in \mathbb{Z}^d, \boldsymbol{k} \neq \boldsymbol{0}\}$. This completes the proof. □

C. Fast Algorithm of Potential Summation

C.1. Integral Transformation

In the main sections of our paper, it is much clear to use node positions to deliver our ideas. However, in the below sections, we are more interested in the vectors between nodes rather than the positions of those nodes. Given a full rank lattice matrix $\boldsymbol{L} \in \mathbb{R}^{d\times d}$ and a vector $\boldsymbol{v} \in \mathbb{R}^d$ between two atoms inside the unit cell, we denote $G(\boldsymbol{L}, \boldsymbol{v})$ as the potential summation and $U(\boldsymbol{L}, \boldsymbol{v})$ as the potential function. Here, if a vector $\boldsymbol{v}'$ is not inside the unit cell, we can apply a simple transformation (Crandall, 1998) to $\boldsymbol{v}'$ by converting it into the fractional coordinate, reducing mod 1, and then converting it back, i.e.,

$$\boldsymbol{v} = \boldsymbol{L}(\boldsymbol{L}^{-1}\boldsymbol{v}' \mod 1) \tag{32}$$

For a potential summation $S(\boldsymbol{a}, \boldsymbol{b})$ of a crystal with a lattice matrix $\boldsymbol{L}$, we have $S(\boldsymbol{a}, \boldsymbol{b}) = G(\boldsymbol{L}, \boldsymbol{v}_{\boldsymbol{ab}})$. Based on these notations, we aim to prove that the summation of the three introduced potentials can be transformed into an integral form as

$$G(\boldsymbol{L}, \boldsymbol{v}) = D\int_0^\infty t^{C-1}(-\delta(\boldsymbol{v}, B) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-A\pi\|\boldsymbol{Lk}+\boldsymbol{v}\|^2 t - \frac{B}{t}})dt, \tag{33}$$

where A, B, C, D are constants derived from the corresponding specific potential forms and δ is the generalized delta function such that $\delta(\boldsymbol{v}, B) = 1$ if and only if $\boldsymbol{v} = \boldsymbol{0}$ and $B = 0$, otherwise $\delta(\boldsymbol{v}, B) = 0$.

Proof. 1). For the potentials in the form of $U(\boldsymbol{L}, \boldsymbol{v}) = 1/\|\boldsymbol{Lk} + \boldsymbol{v}\|^{2p}$ and $\|\boldsymbol{Lk} + \boldsymbol{v}\| \neq 0$, we apply the Mellin transform such that

$$\begin{aligned} M\{U\}(\boldsymbol{L}, \boldsymbol{v}) &= \int_0^\infty t^{p-1} e^{-t\|\boldsymbol{Ln}+\boldsymbol{v}\|^2} dt \\ &= \frac{\Gamma(p)}{\|\boldsymbol{Ln} + \boldsymbol{v}\|^{2p}}. \end{aligned} \tag{34}$$

Consequently, we obtain

$$U(\boldsymbol{L}, \boldsymbol{v}) = \frac{1}{\|\boldsymbol{Lk} + \boldsymbol{v}\|^{2p}} = \frac{1}{\Gamma(p)}\int_0^\infty t^{p-1} e^{-t\|\boldsymbol{Ln}+\boldsymbol{v}\|^2} dt. \tag{35}$$

Next, by deriving the summation,

$$G(\boldsymbol{L}, \boldsymbol{v}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{Lk}+\boldsymbol{v}\|\neq 0} \frac{1}{\|\boldsymbol{Lk} + \boldsymbol{v}\|^{2p}} = \frac{1}{\Gamma(p)}\int_0^\infty t^{p-1}(-\delta(\boldsymbol{v}) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-\|\boldsymbol{Lk}+\boldsymbol{v}\|^2 t})dt. \tag{36}$$

Apparently, we can obtain $A = 1/\pi$, $B = 0$, $C = p$ and $D = 1/\Gamma(p)$ for the summation of $U(\boldsymbol{L}, \boldsymbol{v}) = 1/\|\boldsymbol{Lk} + \boldsymbol{v}\|^{2p}$. The same result is also given in Crandall (1998).

2). For the potentials in the form of $U(\boldsymbol{L}, \boldsymbol{v}) = e^{-\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|}$, we consider the inverse Laplace transform on $e^{-\alpha\sqrt{s}}$ as shown by Bateman (1954), such that

$$\mathcal{L}^{-1}\{e^{-\alpha\sqrt{s}}\} = \frac{a}{2\sqrt{\pi}} t^{-\frac{3}{2}} e^{-\frac{\alpha^2}{4t}}. \tag{37}$$

Therefore, by applying the Laplace transform in Eqn. (37) we derive the integral form of $e^{-\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|}$:

$$\begin{aligned} U(\boldsymbol{L}, \boldsymbol{v}) &= e^{-\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|} \\ &= \frac{a}{2\sqrt{\pi}}\int_0^\infty t^{-\frac{3}{2}} e^{-t\|\boldsymbol{Lk}+\boldsymbol{v}\|^2 - \frac{\alpha^2}{4t}} dt \\ &= \frac{\alpha}{2\pi}\int_0^\infty t^{-\frac{3}{2}} e^{-\pi t\|\boldsymbol{Lk}+\boldsymbol{v}\|^2 - \frac{\alpha^2}{4\pi t}} dt \, (t \leftarrow \pi t). \end{aligned} \tag{38}$$

Next, by deriving the summation,

$$G(\boldsymbol{L}, \boldsymbol{v}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|} = \frac{\alpha}{2\pi}\int_0^\infty t^{-\frac{3}{2}} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-\pi t\|\boldsymbol{Lk}+\boldsymbol{v}\|^2 - \frac{\alpha^2}{4\pi t}} dt. \tag{39}$$

Apparently, we can obtain $A = 1$, $B = \frac{\alpha^2}{4\pi}$, $C = -\frac{1}{2}$ and $D = \frac{\alpha}{2\pi}$ for the summation of $U(\boldsymbol{L}, \boldsymbol{v}) = e^{-\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|}$.

This completes the proof. □

C.2. Calculating Integral Summation

As shown in Sec. 3.4, $G(\boldsymbol{L}, \boldsymbol{v})$ can be written as the summations in the direct space and the Fourier space:

$$\begin{aligned} G(\boldsymbol{L}, \boldsymbol{v}) &= D \int_0^{\infty} t^{C-1}(-\delta(\boldsymbol{v}, B) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}})dt \\ &= D \int_0^{1} t^{C-1}(-\delta(\boldsymbol{v}, B) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}})dt + D \int_1^{\infty} t^{C-1}(-\delta(\boldsymbol{v}, B) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}})dt \\ &= G_{\text{Fourier}}(\boldsymbol{L}, \boldsymbol{v}) + G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v}), \end{aligned} \tag{40}$$

where $G_{\text{Fourier}}(\boldsymbol{L}, \boldsymbol{v}) = D \int_0^1 t^{C-1}(-\delta(\boldsymbol{v}, B) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}})dt$ denotes the summation in Fourier space, and $G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v}) = D \int_1^{\infty} t^{C-1}(-\delta(\boldsymbol{v}, B) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}})dt$ denotes the summation in direct space. Below, we prove that both $G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v})$ and $G_{\text{Fourier}}(\boldsymbol{L}, \boldsymbol{v})$ can be deduced into the incomplete Bessel function summation.

Proof. For $G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v})$, by deriving in direct space,

$$\begin{aligned} G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v}) &= D \int_1^{\infty} t^{C-1}(-\delta(\boldsymbol{v}, B) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}})dt \\ &= D \int_1^{\infty} t^{C-1}(-\delta(\boldsymbol{v}, B) + e^{-A\pi\|\boldsymbol{v}\|^2 t - \frac{B}{t}})dt + D \int_1^{\infty} t^{C-1} \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}} dt \\ &= D \int_1^{\infty} t^{C-1}(-\delta(\boldsymbol{v}, B) + e^{-A\pi\|\boldsymbol{v}\|^2 t - \frac{B}{t}})dt + D \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} \int_1^{\infty} t^{C-1} e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}} dt \\ &= D \int_1^{\infty} t^{C-1}(-\delta(\boldsymbol{v}, B) + e^{-A\pi\|\boldsymbol{v}\|^2 t - \frac{B}{t}})dt + D \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} K_{-C}(A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2, B). \end{aligned} \tag{41}$$

If $\boldsymbol{v} = \boldsymbol{0}$ and $B = 0$,

$$G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v}) = D \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} K_{-C}(A\pi\|\boldsymbol{L}\boldsymbol{k}\|^2, 0); \tag{42}$$

And if $\boldsymbol{v} = \boldsymbol{0}$ and $B \neq 0$,

$$\begin{aligned} G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v}) &= D \int_1^{\infty} t^{C-1} e^{-B/t} dt + D \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} K_{-C}(A\pi\|\boldsymbol{L}\boldsymbol{k}\|^2, B) \\ &= DB^C(\Gamma(-C) - \Gamma(-C, B)) + D \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} K_{-C}(A\pi\|\boldsymbol{L}\boldsymbol{k}\|^2, B); \end{aligned} \tag{43}$$

Otherwise,

$$G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v}) = D \sum_{\boldsymbol{k}\in\mathbb{Z}^d} K_{-C}(A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2, B). \tag{44}$$

Overall,

$$G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v}) = D\delta(\boldsymbol{v})B^C(\Gamma(-C) - \Gamma(-C, B)) + D \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|\neq 0} K_{-C}(A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2, B). \tag{45}$$

Here we obtain $G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v})$ as the incomplete Bessel function summation. Inspired by the Ewald summation (Ewald, 1921; Crandall, 1998), we consider $G_{\text{Fourier}}(\boldsymbol{L}, \boldsymbol{v})$ on the reciprocal lattice using the Poisson summation (Crandall, 1998):

$$\sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-2\pi i \boldsymbol{w}\cdot\boldsymbol{L}\boldsymbol{k} - \pi t\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2} = \frac{t^{-\frac{d}{2}} e^{2\pi i \boldsymbol{w}\cdot\boldsymbol{v}}}{\det \boldsymbol{L}} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{2\pi i \boldsymbol{L}'\boldsymbol{k}\cdot\boldsymbol{v} - \frac{\pi}{t}\left\|\boldsymbol{L}'\boldsymbol{k}+\boldsymbol{w}\right\|^2}, \tag{46}$$

where $\boldsymbol{w} \in \mathbb{R}^d$ and $\boldsymbol{w} = \boldsymbol{0}$ in our case, and $\boldsymbol{L}' = \boldsymbol{L}(\boldsymbol{L}^T\boldsymbol{L})^{-1}$ is the lattice matrix of the reciprocal lattice. Therefore, we derive

$$\begin{aligned}
G_{\text{Fourier}}(\boldsymbol{L},\boldsymbol{v}) &= D\int_0^1 t^{C-1}(-\delta(\boldsymbol{v},B) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}})dt \\
&= -D\delta(\boldsymbol{v},B)\int_0^1 t^{C-1}dt + D\int_0^1 t^{C-1}\sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}}dt \\
&= -D\delta(\boldsymbol{v},B)\int_0^1 t^{C-1}dt + D\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\int_0^1 t^{C-1}e^{-A\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{B}{t}}dt \\
&= -D\delta(\boldsymbol{v},B)\int_0^1 t^{C-1}dt + \frac{D}{A^C}\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\int_0^A t^{C-1}e^{-\pi\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{AB}{t}}dt(t \leftarrow \frac{t}{A}) \\
&= -D\delta(\boldsymbol{v},B)\int_0^1 t^{C-1}dt + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^C}\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\int_0^A t^{C-\frac{d}{2}-1}e^{2\pi i\boldsymbol{L}'\boldsymbol{n}\cdot\boldsymbol{v} - \frac{\pi}{t}\|\boldsymbol{L}'\boldsymbol{k}\|^2 - \frac{AB}{t}}dt(\text{Eqn. (46)}) \\
&= -D\delta(\boldsymbol{v},B)\int_0^1 t^{C-1}dt + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\int_0^1 t^{C-\frac{d}{2}-1}e^{2\pi i\boldsymbol{L}'\boldsymbol{k}\cdot\boldsymbol{v} - \frac{\pi}{At}\|\boldsymbol{L}'\boldsymbol{k}\|^2 - \frac{B}{t}}dt(t \leftarrow At) \\
&= -D\delta(\boldsymbol{v},B)\int_0^1 t^{C-1}dt + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\sum_{\boldsymbol{k}\in\mathbb{Z}^d}\int_1^\infty t^{\frac{d}{2}-C-1}e^{2\pi i\boldsymbol{L}'\boldsymbol{n}\cdot\boldsymbol{v} - \frac{\pi t}{A}\|\boldsymbol{L}'\boldsymbol{k}\|^2 - Bt}dt(t \leftarrow \frac{1}{t}) \\
&= -D\delta(\boldsymbol{v},B)\int_0^1 t^{C-1}dt + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\int_1^\infty t^{\frac{d}{2}-C-1}\sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{2\pi i\boldsymbol{L}'\boldsymbol{n}\cdot\boldsymbol{v} - \frac{\pi t}{A}\|\boldsymbol{L}'\boldsymbol{k}\|^2 - Bt}dt.
\end{aligned} \tag{47}$$

If $B = 0$, we obtain

$$\begin{aligned}
&G_{\text{Fourier}}(\boldsymbol{L},\boldsymbol{v}) \\
&= -D\delta(\boldsymbol{v})\int_0^1 t^{C-1}dt + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\int_1^\infty t^{\frac{d}{2}-C-1}(1 + \sum_{\boldsymbol{k}\in\mathbb{Z}^d,\boldsymbol{k}\neq\boldsymbol{0}} e^{2\pi i\boldsymbol{L}'\boldsymbol{n}\cdot\boldsymbol{v} - \frac{\pi t}{A}\|\boldsymbol{L}'\boldsymbol{k}\|^2})dt \\
&= -D\delta(\boldsymbol{v})\int_0^1 t^{C-1}dt + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\int_1^\infty t^{\frac{d}{2}-C-1}dt + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\sum_{\boldsymbol{k}\in\mathbb{Z}^d,\boldsymbol{k}\neq\boldsymbol{0}}\int_1^\infty t^{\frac{d}{2}-C-1}e^{2\pi i\boldsymbol{L}'\boldsymbol{n}\cdot\boldsymbol{v} - \frac{\pi t}{A}\|\boldsymbol{L}'\boldsymbol{k}\|^2}dt \\
&= -D\delta(\boldsymbol{v})\int_0^1 t^{C-1}dt + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\int_1^\infty t^{\frac{d}{2}-C-1}dt + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\sum_{\boldsymbol{k}\in\mathbb{Z}^d,\boldsymbol{k}\neq\boldsymbol{0}} e^{2\pi i\boldsymbol{L}'\boldsymbol{k}\cdot\boldsymbol{v}}K_{C-\frac{d}{2}}(\frac{\pi\|\boldsymbol{L}'\boldsymbol{k}\|^2}{A}, 0).
\end{aligned} \tag{48}$$

By applying analytic continuation to domain of C as discussed in Appendix C.3, we obtain

$$G_{\text{Fourier}}(\boldsymbol{L},\boldsymbol{v}) = -\frac{D\delta(\boldsymbol{v})}{C} + \frac{1}{\det\boldsymbol{L}}\frac{D}{(C-\frac{d}{2})A^{\frac{d}{2}}} + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\sum_{\boldsymbol{k}\in\mathbb{Z}^d,\boldsymbol{k}\neq\boldsymbol{0}} e^{2\pi i\boldsymbol{L}'\boldsymbol{k}\cdot\boldsymbol{v}}K_{C-\frac{d}{2}}(\frac{\pi\|\boldsymbol{L}'\boldsymbol{k}\|^2}{A}, 0) \tag{49}$$

with poles $C = 0$ and $C = \frac{d}{2}$; Otherwise, if $B \neq 0$,

$$G_{\text{Fourier}}(\boldsymbol{L},\boldsymbol{v}) = \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{2\pi i\boldsymbol{L}'\boldsymbol{k}\cdot\boldsymbol{v}}K_{C-\frac{d}{2}}(\frac{\pi\|\boldsymbol{L}'\boldsymbol{k}\|^2}{A} + B, 0). \tag{50}$$

Together, we obtain

$$G_{\text{Fourier}}(\boldsymbol{L},\boldsymbol{v}) = -\frac{D\delta(\boldsymbol{v},B)}{C} + \frac{\delta(B)}{\det\boldsymbol{L}}\frac{D}{(C-\frac{d}{2})A^{\frac{d}{2}}} + \frac{1}{\det\boldsymbol{L}}\frac{D}{A^{\frac{d}{2}}}\sum_{\boldsymbol{k}\in\mathbb{Z}^d,\pi\|\boldsymbol{L}'\boldsymbol{k}\|^2/A+B>0} e^{2\pi i\boldsymbol{L}'\boldsymbol{k}\cdot\boldsymbol{v}}K_{C-\frac{d}{2}}(\frac{\pi\|\boldsymbol{L}'\boldsymbol{k}\|^2}{A} + B, 0). \tag{51}$$

Apparently, $G_{\text{Fourier}}(\boldsymbol{L},\boldsymbol{v})$ is deduced into the incomplete Bessel function summation, and this completes the proof. □

Therefore, both $G_{\text{Fourier}}(\boldsymbol{L}, \boldsymbol{v})$ and $G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v})$ can be expressed by the incomplete Bessel function summation. In addition, as shown in Appendix B.2, the incomplete Bessel function summation $\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma>0} K_\nu(\pi\alpha\|\boldsymbol{Lk}+\boldsymbol{v}\|^2+\gamma, \beta)$ converges and can be approximated. Therefore, $G(\boldsymbol{L}, \boldsymbol{v})$ converges and also can be approximated.

C.3. Analytic Continuation of Potential Summations

To represent the series that is not well-defined in its original domain, including the Coulomb potential summation $\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{Lk}+\boldsymbol{v}\|\neq 0} \frac{1}{\|\boldsymbol{Lk}+\boldsymbol{v}\|}$, we need to investigate the analytic continuation of the potential summation. Analytic continuation is a technique to extend the domain P of a given analytic function $f(x)$. If there exists a domain Q containing P, and a function $\hat{f}(x)$ that is analytic on Q, and $\hat{f}(x) = f(x)$ holding for all x in P, consequently, $\hat{f}(x)$ is an analytic continuation of $f(x)$ to Q. As shown by Kung & Yang (2003), the analytic continuation is unique and satisfies the permanence of functional relationships, i.e., the equations holding for $f(x)$ will also hold for $\hat{f}(x)$.

In our case, we can expand the domain of C in $G(\boldsymbol{L}, \boldsymbol{v})$ to $C \in \mathbb{R}/\{0, \frac{d}{2}\}$ such that $G(\boldsymbol{L}, \boldsymbol{v})$ is well-defined for any $C \in \mathbb{R}/\{0, \frac{d}{2}\}$. This is enabled by analytic continuation in Eqn. (48). To be concrete, the original domain of C is $(\frac{d}{2}, +\infty)$ in Eqn. (48) and we can extend the domain of C for $\int_0^1 t^{C-1}dt$, $\int_1^\infty t^{\frac{d}{2}-C-1}dt$, and $K_{C-\frac{d}{2}}(\frac{\pi\|\boldsymbol{L}'\boldsymbol{k}\|^2}{A}, 0)$, respectively. By applying the analytically continued incomplete Bessel function we can extend the domain of C to $\mathbb{R}$. On the other hand, the analytic continuation on $\int_0^1 t^{C-1}dt$ and $\int_1^\infty t^{\frac{d}{2}-C-1}dt$ will result in two poles $C = 0$ and $C = \frac{d}{2}$ on $\mathbb{R}$. Therefore, the final analytically continued domain of C is $\mathbb{R}/\{0, \frac{d}{2}\}$. For the potential summation $\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{Lk}+\boldsymbol{v}\|\neq 0} 1/\|\boldsymbol{Lk}+\boldsymbol{v}\|^{2p}$, we have $A = 1/\pi, B = 0, C = p, D = 1/\Gamma(p)$ as shown in Appendix C.1. By analytic continuation, we are able to compute the summation for any $p \in \mathbb{R}/\{\frac{d}{2}, 0, -1, -2, -3, \cdots\}$. Since $p = 1$ and $d = 3$ for our crystal dataset, we are able to compute the Coulomb potential summation. In next section, a more general analytic continuation result $p \in \mathbb{C}/\{0, \frac{d}{2}\}$ is given.

C.4. Generalized Epstein Zeta Function and Analytic Continuation

In fact, the potential summation $\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{Lk}+\boldsymbol{v}\|\neq 0} 1/\|\boldsymbol{Lk}+\boldsymbol{v}\|^{2p}$ is a special case of *generalized Epstein zeta function* (Crandall & Buhler, 1987; Terras, 1973; Crandall, 1998), and is a generalization of the Riemann zeta function. Let $s \in \mathbb{C}/\{0, d\}$, $\boldsymbol{L} \in \mathbb{R}^{d\times d}$ and $\boldsymbol{v}, \boldsymbol{u} \in \mathbb{R}^d$ inside the unit cell and reciprocal unit cell respectively. The generalized Epstein zeta function (Crandall, 1998) has the below summation form

$$Z_{\boldsymbol{L}}(s; \boldsymbol{u}, \boldsymbol{v}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{Lk}-\boldsymbol{v}\|\neq 0} \frac{e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{Lk}}}{\|\boldsymbol{Lk}-\boldsymbol{v}\|^s}. \tag{52}$$

Apparently, the potential summation $\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{Lk}+\boldsymbol{v}\|\neq 0} 1/\|\boldsymbol{Lk}+\boldsymbol{v}\|^{2p}$ can be expressed in terms of generalized Epstein zeta function as $Z_{\boldsymbol{L}}(2p; \boldsymbol{0}, -\boldsymbol{v})$. In addition, $Z_{\boldsymbol{L}}(s; \boldsymbol{u}, \boldsymbol{v})$ has an analytic continuation to the entire complex plane, except for simple poles at $s = 0$ and $s = d$ (Crandall & Buhler, 1987), which is corresponding to our result in Appendix C.3. Moreover, $Z_{\boldsymbol{L}}(s; \boldsymbol{u}, \boldsymbol{v})$ can be written in the form of an integral summation (Crandall, 1998) similar to Eqn. (35) such that

$$Z_{\boldsymbol{L}}(s; \boldsymbol{u}, \boldsymbol{v}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{Lk}-\boldsymbol{v}\|\neq 0} \frac{e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{Lk}}}{\|\boldsymbol{Lk}-\boldsymbol{v}\|^s} = \frac{\pi^{s/2}}{\Gamma(s/2)} \int_0^\infty t^{s/2-1}(-\delta(\boldsymbol{v}) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{Lk} - \pi t\|\boldsymbol{Lk}-\boldsymbol{v}\|^2})dt. \tag{53}$$

Based on this, we can also split the integral and apply Poisson summation in Eqn. (46) to obtain two summations of incomplete Bessel functions to evaluate this series. To be concrete, similar to Eqn. (41), the summation of generalized Epstein zeta function on direct space is

$$\begin{aligned}
G_{\text{direct}}(s; \boldsymbol{L}, \boldsymbol{u}, \boldsymbol{v}) &= \frac{\pi^{s/2}}{\Gamma(s/2)} \int_1^\infty t^{s/2-1}(-\delta(\boldsymbol{v}) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{Lk} - \pi t\|\boldsymbol{Lk}-\boldsymbol{v}\|^2})dt \\
&= \frac{\pi^{s/2}}{\Gamma(s/2)} \int_1^\infty t^{s/2-1}(-\delta(\boldsymbol{v}) + e^{-\pi t\|\boldsymbol{v}\|^2})dt + \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} \frac{\pi^{s/2}}{\Gamma(s/2)} \int_1^\infty t^{s/2-1} e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{Lk} - \pi t\|\boldsymbol{Lk}-\boldsymbol{v}\|^2} dt \\
&= \frac{\pi^{s/2}}{\Gamma(s/2)} \int_1^\infty t^{s/2-1}(-\delta(\boldsymbol{v}) + e^{-\pi t\|\boldsymbol{v}\|^2})dt + \frac{\pi^{s/2}}{\Gamma(s/2)} \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{Lk}} K_{-s/2}(\pi\|\boldsymbol{Lk}-\boldsymbol{v}\|^2, 0).
\end{aligned} \tag{54}$$

By further exploring two cases of $\delta(\boldsymbol{v})$, we obtain

$$G_{\text{direct}}(s;\boldsymbol{L},\boldsymbol{u},\boldsymbol{v}) = \frac{\pi^{s/2}}{\Gamma(s/2)} \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{L}\boldsymbol{k}-\boldsymbol{v}\|\neq 0} e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{k}} K_{-s/2}(\pi\|\boldsymbol{L}\boldsymbol{k}-\boldsymbol{v}\|^2, 0). \tag{55}$$

The summation of generalized Epstein zeta function on Fourier space is

$$\begin{aligned}
G_{\text{Fourier}}(s;\boldsymbol{L},\boldsymbol{u},\boldsymbol{v}) &= \frac{\pi^{s/2}}{\Gamma(s/2)} \int_0^1 t^{s/2-1}(-\delta(\boldsymbol{v}) + \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{k} - \pi t\|\boldsymbol{L}\boldsymbol{k}-\boldsymbol{v}\|^2})dt \\
&= -\frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)} \int_0^1 t^{s/2-1}dt + \frac{\pi^{s/2}}{\Gamma(s/2)} \int_0^1 t^{s/2-1} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{k} - \pi t\|\boldsymbol{L}\boldsymbol{k}-\boldsymbol{v}\|^2} dt \\
&= -\frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)} \int_0^1 t^{s/2-1}dt + \\
&\quad \frac{\pi^{s/2} e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{v}}}{\Gamma(s/2)\det\boldsymbol{L}} \int_0^1 t^{s/2-d/2-1} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-2\pi i \boldsymbol{v}\cdot\boldsymbol{L}'\boldsymbol{k} - \frac{\pi}{t}\|\boldsymbol{L}'\boldsymbol{k}-\boldsymbol{u}\|^2} dt (Eqn.\ (46)) \\
&= -\frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)} \int_0^1 t^{s/2-1}dt + \\
&\quad \frac{\pi^{s/2} e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{v}}}{\Gamma(s/2)\det\boldsymbol{L}} \int_1^\infty t^{d/2-s/2-1} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-2\pi i \boldsymbol{v}\cdot\boldsymbol{L}'\boldsymbol{k} - \pi t\|\boldsymbol{L}'\boldsymbol{k}-\boldsymbol{u}\|^2} dt (t \leftarrow \frac{1}{t}).
\end{aligned} \tag{56}$$

if $\boldsymbol{u} = \boldsymbol{0}$,

$$\begin{aligned}
G_{\text{Fourier}}(s;\boldsymbol{L},\boldsymbol{u},\boldsymbol{v}) &= -\frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)} \int_0^1 t^{s/2-1}dt + \\
&\quad \frac{\pi^{s/2}}{\Gamma(s/2)\det\boldsymbol{L}} \int_1^\infty t^{d/2-s/2-1} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-2\pi i \boldsymbol{v}\cdot\boldsymbol{L}'\boldsymbol{k} - \pi t\|\boldsymbol{L}'\boldsymbol{k}\|^2} dt \\
&= -\frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)} \int_0^1 t^{s/2-1}dt + \\
&\quad \frac{\pi^{s/2}}{\Gamma(s/2)\det\boldsymbol{L}} \int_1^\infty t^{d/2-s/2-1}(1 + \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} e^{-2\pi i \boldsymbol{v}\cdot\boldsymbol{L}'\boldsymbol{k} - \pi t\|\boldsymbol{L}'\boldsymbol{k}\|^2})dt \\
&= -\frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)} \int_0^1 t^{s/2-1}dt + \frac{\pi^{s/2}}{\Gamma(s/2)\det\boldsymbol{L}} \int_1^\infty t^{d/2-s/2-1}dt + \\
&\quad \frac{\pi^{s/2}}{\Gamma(s/2)\det\boldsymbol{L}} \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} \int_1^\infty t^{d/2-s/2-1} e^{-2\pi i \boldsymbol{v}\cdot\boldsymbol{L}'\boldsymbol{k} - \pi t\|\boldsymbol{L}'\boldsymbol{k}\|^2} dt \\
&= -\frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)} \int_0^1 t^{s/2-1}dt + \frac{\pi^{s/2}}{\Gamma(s/2)\det\boldsymbol{L}} \int_1^\infty t^{d/2-s/2-1}dt + \\
&\quad \frac{\pi^{s/2}}{\Gamma(s/2)\det\boldsymbol{L}} \sum_{\boldsymbol{k}\in\mathbb{Z}^d, \boldsymbol{k}\neq\boldsymbol{0}} e^{-2\pi i \boldsymbol{v}\cdot\boldsymbol{L}'\boldsymbol{k}} K_{s/2-d/2}(\pi\|\boldsymbol{L}'\boldsymbol{k}\|, 0);
\end{aligned} \tag{57}$$

Otherwise,

$$\begin{aligned}
G_{\text{Fourier}}(s;\boldsymbol{L},\boldsymbol{u},\boldsymbol{v}) &= -\frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)} \int_0^1 t^{s/2-1}dt + \\
&\quad \frac{\pi^{s/2} e^{2\pi i \boldsymbol{u}\cdot\boldsymbol{v}}}{\Gamma(s/2)\det\boldsymbol{L}} \sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-2\pi i \boldsymbol{v}\cdot\boldsymbol{L}'\boldsymbol{k}} K_{s/2-d/2}(\pi\|\boldsymbol{L}'\boldsymbol{k}-\boldsymbol{u}\|, 0).
\end{aligned} \tag{58}$$

Then

$$G_{\text{Fourier}}(s;\boldsymbol{L},\boldsymbol{u},\boldsymbol{v}) = -\frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)}\int_0^1 t^{s/2-1}dt + \frac{\delta(\boldsymbol{u})\pi^{s/2}}{\Gamma(s/2)\det\boldsymbol{L}}\int_1^\infty t^{d/2-s/2-1}dt + \frac{\pi^{s/2}e^{2\pi i\boldsymbol{u}\cdot\boldsymbol{v}}}{\Gamma(s/2)\det\boldsymbol{L}}\sum_{\boldsymbol{k}\in\mathbb{Z}^d,\boldsymbol{k}\neq\boldsymbol{0}} e^{-2\pi i\boldsymbol{v}\cdot\boldsymbol{L}'\boldsymbol{k}}K_{s/2-d/2}(\pi\|\boldsymbol{L}'\boldsymbol{k}-\boldsymbol{u}\|,0); \tag{59}$$

Therefore,

$$\begin{aligned} Z_{\boldsymbol{L}}(s;\boldsymbol{u},\boldsymbol{v}) &= G_{\text{direct}}(s;\boldsymbol{L},\boldsymbol{u},\boldsymbol{v}) + G_{\text{Fourier}}(s;\boldsymbol{L},\boldsymbol{u},\boldsymbol{v}) \\ &= \frac{\pi^{s/2}}{\Gamma(s/2)}\sum_{\boldsymbol{k}\in\mathbb{Z}^d,\|\boldsymbol{L}\boldsymbol{k}-\boldsymbol{v}\|\neq 0} e^{2\pi i\boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{k}}K_{-s/2}(\pi\|\boldsymbol{L}\boldsymbol{k}-\boldsymbol{v}\|^2,0) - \frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)}\int_0^1 t^{s/2-1}dt + \\ &\quad \frac{\delta(\boldsymbol{u})\pi^{s/2}}{\Gamma(s/2)\det\boldsymbol{L}}\int_1^\infty t^{d/2-s/2-1}dt + \frac{\pi^{s/2}e^{2\pi i\boldsymbol{u}\cdot\boldsymbol{v}}}{\Gamma(s/2)\det\boldsymbol{L}}\sum_{\boldsymbol{k}\in\mathbb{Z}^d,\boldsymbol{k}\neq\boldsymbol{0}} e^{-2\pi i\boldsymbol{v}\cdot\boldsymbol{L}'\boldsymbol{k}}K_{s/2-d/2}(\pi\|\boldsymbol{L}'\boldsymbol{k}-\boldsymbol{u}\|,0). \end{aligned} \tag{60}$$

Similar to Eqn. (48), we apply analytic continuation and obtain

$$\begin{aligned} Z_{\boldsymbol{L}}(s;\boldsymbol{u},\boldsymbol{v}) &= G_{\text{direct}}(s;\boldsymbol{L},\boldsymbol{u},\boldsymbol{v}) + G_{\text{Fourier}}(s;\boldsymbol{L},\boldsymbol{u},\boldsymbol{v}) \\ &= \frac{\pi^{s/2}}{\Gamma(s/2)}\sum_{\boldsymbol{k}\in\mathbb{Z}^d,\|\boldsymbol{L}\boldsymbol{k}-\boldsymbol{v}\|\neq 0} e^{2\pi i\boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{k}}K_{-s/2}(\pi\|\boldsymbol{L}\boldsymbol{k}-\boldsymbol{v}\|^2,0) - \frac{\delta(\boldsymbol{v})\pi^{s/2}}{\Gamma(s/2)s/2} + \\ &\quad \frac{\delta(\boldsymbol{u})\pi^{s/2}}{\Gamma(s/2)\det\boldsymbol{L}(d/2-s/2)} + \frac{\pi^{s/2}e^{2\pi i\boldsymbol{u}\cdot\boldsymbol{v}}}{\Gamma(s/2)\det\boldsymbol{L}}\sum_{\boldsymbol{k}\in\mathbb{Z}^d,\boldsymbol{k}\neq\boldsymbol{0}} e^{-2\pi i\boldsymbol{v}\cdot\boldsymbol{L}'\boldsymbol{k}}K_{s/2-d/2}(\pi\|\boldsymbol{L}'\boldsymbol{k}-\boldsymbol{u}\|,0). \end{aligned} \tag{61}$$

Here, Crandall (1998) gives the same result Eqn. (61) with incomplete Gamma function, while we provide detailed derivation on $Z_{\boldsymbol{L}}(s;\boldsymbol{u},\boldsymbol{v})$. For more details of the generalized Epstein zeta function, we refer readers to Crandall & Buhler (1987); Terras (1973); Crandall (1998); Kirsten (1994); Selberg & Chowla (1967). The generalized Epstein zeta function is used for the computation of Madelung constants (Crandall, 1998) as described in Appendix E. This is because one can view the term $e^{2\pi i\boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{k}} = \cos(2\pi\boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{k}) + i\sin(2\pi\boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{k})$ as the charge distribution in the crystal system. If the phase $2\boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{k}\in\mathbb{Z}$ for any $\boldsymbol{k}$, the generalized Epstein zeta function will become an alternating series. This is useful for the ionic crystal systems where the unit cell is generally neutral and the Coulomb potentials cancel each other. A famous example is the Madelung constant of $NaCl$, which is derived by the summation of Coulomb potentials among Na and Cl ions. By using the generalized Epstein zeta function, we have the Madelung constant of $NaCl$ as $M = Z_{1_3}(1;(\frac{1}{2})_3,0_3)$ (Crandall, 1998) and here a_n denotes a diagonal matrix with all main diagonal values as the scalar a.

One can also calculate the Madelung constant by considering extracting terms in $Z_{\boldsymbol{L}}(s;\boldsymbol{u},\boldsymbol{v})$ with the same coefficient $e^{2\pi i\boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{k}}$ as individual potential summations and computing those individual summations (by analytic continuation). This is due to the fact that $Z_{\boldsymbol{L}}(s;\boldsymbol{u},\boldsymbol{v})$ and these individual summations are all calculated by incomplete Bessel functions and share the same analytically continued domain. In other words, $Z_{\boldsymbol{L}}(s;\boldsymbol{u},\boldsymbol{v})$ can be calculated by a linear combination of individual potential summations (by analytic continuation). It is useful for the case where $\boldsymbol{u}$ is initially unknown but can be learned. Here we give a simple numerical example to calculate eta function $\sum_{n=1}^\infty \frac{(-1)^{n+1}}{\sqrt{n}}$ by analytic continuation such that

$$\sum_{n=1}^\infty \frac{(-1)^{n+1}}{\sqrt{n}} = \sum_{n=0}^\infty \frac{(-1)^n}{\sqrt{n+1}} = \frac{1}{\sqrt{2}}\sum_{n=0}^\infty\left(\frac{1}{\sqrt{n+\frac{1}{2}}} - \frac{1}{\sqrt{n+1}}\right) = \frac{1}{\sqrt{2}}\left(\zeta(\frac{1}{2},\frac{1}{2}) - \zeta(\frac{1}{2},1)\right), \tag{62}$$

where $\zeta(s,a)$ is Hurwitz zeta function and $\zeta(s,a) = \sum_{n=0}^\infty \frac{1}{(n+a)^s}$ when $s>1, a\neq 0,-1,-2,...$, and its analytic continuation elsewhere. That is, we can use the analytically continued zeta function $\zeta(s,a)$ to precisely evaluate a conditionally convergent series (or alternating zeta function). Overall, the above shows that we can use analytically continued potential summations to approximate the total contribution of potentials where individual potential summations cancel each other. We further show the Madelung constant calculation of $NaCl$ by analytic continuation in Appendix E.

Table 6. Numerical examples of our algorithm. We approximate results by using all grid elements and approximating the error upper bound. The first column denotes the ground truth targets we aim to approximate. By choosing different grid length $2 * r$ in our algorithm, we can obtain different evaluation results and different estimated error upper bound of the approximation as described in the second, the third, and the fourth column. The fifth column gives the truth error between our approximation and ground truth. And the sixth column denotes time cost of our approximation. Here, $\zeta(x) = \sum_{n=1}^{\infty} 1/n^x$. Implementation details can be found in Appendix C.5.

Ground Truth	Evaluation	r	Estimated Error	True Error	Time
$2 * \zeta(2) = 3.28986813$	3.28068288	1	9e-1	9e-3	2 ms
$2 * \zeta(2) = 3.28986813$	3.28984070	2	6e-2	3e-5	3 ms
$2 * \zeta(2) = 3.28986813$	3.28986812	3	7e-4	1e-8	3 ms
$2 * \zeta(2) = 3.28986813$	3.28986813	4	1e-6	$<$ 1e-8	3 ms
$2 * \zeta(3) = 2.40411381$	2.40411381	4	1e-6	$<$ 1e-8	3 ms
$2 * \zeta(4) = 2.16464647$	2.16464647	4	1e-6	$<$ 1e-8	3 ms
$\sum_{\boldsymbol{n}\in\mathbb{Z}^2,\boldsymbol{n}\neq\boldsymbol{0}} \frac{1}{\lvert\boldsymbol{n}\rvert^4} = 6.02681204$	5.99068949	1	3	4e-2	3 ms
$\sum_{\boldsymbol{n}\in\mathbb{Z}^2,\boldsymbol{n}\neq\boldsymbol{0}} \frac{1}{\lvert\boldsymbol{n}\rvert^4} = 6.02681204$	6.02670959	2	4e-1	1e-4	3 ms
$\sum_{\boldsymbol{n}\in\mathbb{Z}^2,\boldsymbol{n}\neq\boldsymbol{0}} \frac{1}{\lvert\boldsymbol{n}\rvert^4} = 6.02681204$	6.02681199	3	7e-3	5e-8	3 ms
$\sum_{\boldsymbol{n}\in\mathbb{Z}^2,\boldsymbol{n}\neq\boldsymbol{0}} \frac{1}{\lvert\boldsymbol{n}\rvert^4} = 6.02681204$	6.02681204	4	2e-5	$<$ 1e-8	3 ms
$\sum_{n\in\mathbb{Z}} e^{-\lvert n\rvert} = 2.16395341$	2.16395326	1	4e-1	2e-7	2 ms
$\sum_{n\in\mathbb{Z}} e^{-\lvert n\rvert} = 2.16395341$	2.16395341	2	3e-4	$<$ 1e-8	4 ms
$\sum_{\boldsymbol{n}\in\mathbb{Z}^3} e^{-\lvert\boldsymbol{n}\rvert} = 25.39268269$	25.39268214	1	2.5	5e-7	3 ms
$\sum_{\boldsymbol{n}\in\mathbb{Z}^3} e^{-\lvert\boldsymbol{n}\rvert} = 25.39268269$	25.39268269	2	1e-2	$<$ 1e-8	3 ms

C.5. Implementation and Numerical Examples of Approximation

In this section, we explicate the methodology applied for the implementation of our algorithm. A predefined d-dimensional discrete grid, centered at the origin and having a length defined by $2r$, is applied for point selection of summation approximation. Typically, $r > R$ is delineated in the algorithm. Owing to the complexities associated with computing the inverse of the incomplete Gamma function, our primary objective is to establish the value of R, after which we calculate its corresponding error bound. To elaborate, we commence by determining the value of R within a crystalline structure, following which we ascertain the associated error bound, as prescribed by Eqn. (31). Within the d-dimensional discrete grid, a set of points inside an ellipsoid predicated on R is selected for the evaluation of the summation of incomplete Bessel functions. It is essential to underscore that determining R for each crystal is a nontrivial procedure. In cases where R is not predetermined, we employ all grid elements to approximate the result and estimate the error upper bound by approximating R. As $\boldsymbol{v}$ is inside the lattice, the approximation of R is achieved by considering r_{min}, which is the length of the minor axis of an ellipsoid. And the ellipsoid is constructed via the expression $\boldsymbol{L}\boldsymbol{x} + \boldsymbol{v}$ or $\boldsymbol{L}'\boldsymbol{x}$, with $\boldsymbol{x}$ denoting points derived from a sphere centered at the origin with a radius r. Formally, given a lattice matrix $\boldsymbol{L} \in \mathbb{R}^{d\times d}$, a vector $\boldsymbol{v} \in \mathbb{R}^d$ inside a unit cell, and the constants A, B, C, D derived from specific potential functions as described in Appendix C.1, we aim to evaluate these two parts $G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v})$ and $G_{\text{Fourier}}(\boldsymbol{L}, \boldsymbol{v})$.

To evaluate $G_{\text{direct}}(\boldsymbol{L}, \boldsymbol{v})$ according to Eqn. (45), we derive the following steps.

Step 1: Determine the grid residing in the d-dimensional integer space $\mathbb{Z}^d$ with length denoted as $2r$ and a value R such that the conditions $r > R$ and $R^2 \geq C$, $R^2 \geq 1$ are satisfied. Subsequently, select points denoted by $\boldsymbol{k}$ from the grid, adhering to the inequality $A\pi\|\boldsymbol{L}\boldsymbol{k} + \boldsymbol{v}\|^2 \leq R^2$. Once the points are selected, calculate the function represented by Eqn. (45). In the scenario where R remains undefined, it is recommended to use large r and select all available points on the grid for computation of Eqn. (45). R is then approximated by $\sqrt{A\pi}(r_{min} - \|\boldsymbol{v}\|)$.

Step 2: Compute the error bound denoted as ϵ by the following formula: $\epsilon = \frac{d}{2}(\frac{2}{\rho})^d\Gamma(\frac{d}{2}, (R - \frac{\rho}{2})^2)$, where ρ satisfies the condition $\rho = \min\{\sqrt{A\pi}\|\boldsymbol{L}\boldsymbol{k}\| \mid \boldsymbol{k} \neq \boldsymbol{0}\}$.

To evaluate $G_{\text{Fourier}}(\boldsymbol{L}, \boldsymbol{v})$ according to Eqn. (51), we derive the following steps.

Step 1: Determine the grid residing in the d-dimensional integer space $\mathbb{Z}^d$ with length denoted as $2r$ and a value R such that the conditions $r > R$ and $R^2 \geq \frac{d}{2} - C$, $R^2 \geq B$, $R^2 \geq 1$ are satisfied. Subsequently, select points denoted by $\boldsymbol{k}$ from the

grid, adhering to the inequality $\frac{\pi}{A}\|\boldsymbol{L}'\boldsymbol{k}\|^2 + B \leq R^2$. Once the points are selected, calculate the function represented by Eqn. (51). In the scenario where R remains undefined, it is recommended to use large r and select all available points on the grid for computation of Eqn. (51). R is then approximated by $\sqrt{\frac{\pi}{A}r_{min}^2 + B}$.

Step 2: Compute the error bound denoted as ϵ by the following formula: $\epsilon = \frac{d}{2}(\frac{2}{\rho})^d\Gamma(\frac{d}{2}, (\sqrt{R^2 - B} - \frac{\rho}{2})^2)$, where ρ satisfies the condition $\rho = \min\{\sqrt{\frac{\pi}{A}}\|\boldsymbol{L}'\boldsymbol{k}\| \mid \boldsymbol{k} \in \mathbb{Z}^d, \boldsymbol{k} \neq \boldsymbol{0}\}$.

Our implementation is based on Cython, GNU Scientific Library (Galassi et al., 2002) and ScaFaCoS (Bolten et al.), in which the native incomplete Gamma function and incomplete Bessel function are used. We show the evaluation examples in Table 6 with the corresponding error bound and evaluation time. The running time is at the scale of milliseconds.

C.6. Potential Summation Extensions

To highlight the versatility of our potential summation method, we incorporate additional potentials that can be computed using our algorithm. These include the Lennard-Jones potential, Morse potential, and screened Coulomb potential. These potentials typically find application in the analysis of specific categories of materials, thereby demonstrating the broad applicability of our method.

Lennard-Jones Potential (Lennard-Jones & Dent, 1928) is an intermolecular pair potential that is usually used for gas or organic materials. Let ϵ and σ be hyperparameters. The commonly used expression for the Lennard-Jones potential is

$$U_{\text{LJ}}(\boldsymbol{L}, \boldsymbol{v}) = 4\epsilon\left[\left(\frac{\sigma}{\|\boldsymbol{L}\boldsymbol{k} + \boldsymbol{v}\|}\right)^{12} - \left(\frac{\sigma}{\|\boldsymbol{L}\boldsymbol{k} + \boldsymbol{v}\|}\right)^{6}\right]. \tag{63}$$

And the summation of $U_{\text{LJ}}(\boldsymbol{L}, \boldsymbol{v})$ can be converted to two potential summations of $1/\|\boldsymbol{L}\boldsymbol{k} + \boldsymbol{v}\|^{2p}$ with $p = 3$ and $p = 6$ respectively, such that

$$G_{\text{LJ}}(\boldsymbol{L}, \boldsymbol{v}) = 4\epsilon\left[\sigma^{12}\left(\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|\neq 0}\frac{1}{\|\boldsymbol{L}\boldsymbol{k} + \boldsymbol{v}\|^{12}}\right) - \sigma^{6}\left(\sum_{\boldsymbol{k}\in\mathbb{Z}^d, \|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|\neq 0}\frac{1}{\|\boldsymbol{L}\boldsymbol{k} + \boldsymbol{v}\|^{6}}\right)\right], \tag{64}$$

where we already show the calculation of the potential summation of $1/\|\boldsymbol{L}\boldsymbol{k} + \boldsymbol{v}\|^{2p}$ in Appendix C.1.

Morse Potential (Morse, 1929) is an interatomic potential of diatomic molecules and is used for simple molecular materials. Let D_e and r_e be hyperparameters. The Morse potential has a mathematical form of

$$U_{\text{Morse}}(\boldsymbol{L}, \boldsymbol{v}) = D_e\left(e^{-2a(\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\| - r_e)} - 2e^{-a(\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\| - r_e)}\right). \tag{65}$$

Similarly, the summation of $U_{\text{Morse}}(\boldsymbol{L}, \boldsymbol{v})$ can be converted to two potential summations of $e^{-\alpha\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|}$ with $\alpha = a$ and $\alpha = 2a$ respectively, such that

$$G_{\text{Morse}}(\boldsymbol{L}, \boldsymbol{v}) = D_e\left(e^{2ar_e}\sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-2a\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|} - 2e^{ar_e}\sum_{\boldsymbol{k}\in\mathbb{Z}^d} e^{-a\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|}\right), \tag{66}$$

where we already show the calculation of the potential summation of $e^{-\alpha\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|}$ in Appendix C.1.

Screened Coulomb Potential represents the Coulomb interactions with damping of electric fields. It is an important potential reflecting the behaviors of charge-carrying fluids or particles in semiconductors (Kirichenko & Stephanovich, 2021). Let e_0 be elementary charge constant and α be a scaling hyperparameter. The screened Coulomb potential has an analytic form of $V(\boldsymbol{a}, \boldsymbol{b}) = \frac{z_{\boldsymbol{a}} z_{\boldsymbol{b}} e_0^2}{d(\boldsymbol{a}, \boldsymbol{b})} e^{-\alpha d(\boldsymbol{a}, \boldsymbol{b})}$, where $d(\boldsymbol{a}, \boldsymbol{b})$ is the distance between atom $\boldsymbol{a}$ and $\boldsymbol{b}$, and $z_{\boldsymbol{a}}, z_{\boldsymbol{b}}$ are charges of atom $\boldsymbol{a}$ and $\boldsymbol{b}$. Similar to our Coulomb potential case, since $z_{\boldsymbol{a}}, z_{\boldsymbol{b}}, e_0$ are constants and can be extracted outside the summation, we derive a simplified screened Coulomb potential

$$U_{\text{screened}}(\boldsymbol{L}, \boldsymbol{v}) = \frac{e^{-\alpha\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|}}{\|\boldsymbol{L}\boldsymbol{k} + \boldsymbol{v}\|}. \tag{67}$$

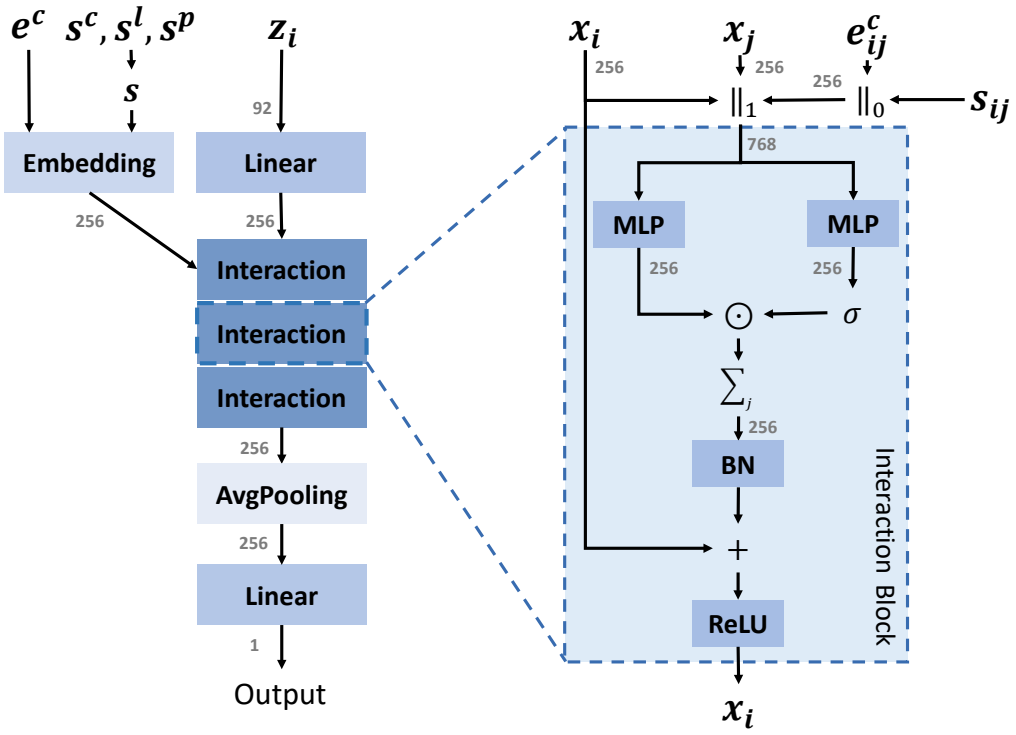

Figure 2. The developed network architecture for PotNet. The notations used in this figure are defined as follows: σ denotes the sigmoid function; $\|_a$ represents the concatenation operation along the dimension a; $\odot$ signifies the Hadamard product, or element-wise multiplication of two matrices or tensors of the same dimensions; And $\sum_j$ represents the aggregation operation over the index j.

Consider the inverse Laplace transform on $e^{-\alpha\sqrt{s}}/\sqrt{s}$ by using Bromwich contour with branch points, we obtain

$$\mathcal{L}^{-1}\{e^{-\alpha\sqrt{s}}/\sqrt{s}\} = \frac{1}{\sqrt{\pi t}} e^{-\frac{\alpha^2}{4t}}. \tag{68}$$

Therefore, we can apply the Laplace transform in Eqn. (67) such that

$$U_{\text{screened}}(\boldsymbol{L}, \boldsymbol{v}) = \frac{e^{-\alpha\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|}}{\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|} = \frac{1}{\sqrt{\pi}} \int_0^\infty t^{-\frac{1}{2}} e^{-\|\boldsymbol{L}\boldsymbol{k}+\boldsymbol{v}\|^2 t - \frac{\alpha^2}{4t}} dt. \tag{69}$$

Then we obtain $A = 1/\pi$, $B = \alpha^2$, $C = \frac{1}{2}$ and $D = 1/\sqrt{\pi}$ in Eqn. (33) to fit screened Coulomb potential into our potential summation method.

D. Experimental Details

D.1. PotNet Implementation

The employed network architecture is shown in Fig. 2. Since our major contribution is to consider interatomic potentials and their complete form, we simply design our network architecture following the commonly used settings. Specifically, existing methods for 3D graphs (Xie & Grossman, 2018; Schütt et al., 2017; Klicpera et al., 2020b;a; Gasteiger et al., 2021; Schütt et al., 2021; Wang et al., 2022a; Liu et al., 2022b; Yan et al., 2022; Wang et al., 2023) share a similar architecture, which usually contains an input block, an interaction block, and an output block. Without loss of generality, we take the updating process for node i as an example to illustrate the network.

- *The Inputs* contain atomic features and potentials. $\mathbf{z}_i$ is the 92-dimensional atomic feature for any atom i following CGCNN (Xie & Grossman, 2018). Below the term d denotes interatomic distances, which form an integral part of our

potential features. As delineated in Sec. 4.1, our computational model employs both local and global graphs. In the context of the local graph, only the Coulomb potential is considered, as per the details provided in the aforementioned Sec. 4.1. On the other hand, the global graph encapsulates the summations of infinite potentials as its edge features. We consider all three categories of infinite potential summations in our model, namely Coulomb potentials, London dispersion potentials, and Pauli repulsion potentials. These are discussed in detail in Sec. 2.2. The hyperparameter ϵ' is employed to simplify the mathematical form of Coulomb potentials, expressed as $V_{\text{Coulomb}}(\boldsymbol{a}, \boldsymbol{b}) = -\frac{z_{\boldsymbol{a}} z_{\boldsymbol{b}} e_0^2}{4\pi\epsilon_0 d(\boldsymbol{a},\boldsymbol{b})}$, to $V_{\text{Coulomb}}(\boldsymbol{a}, \boldsymbol{b}) = -\epsilon'/d(\boldsymbol{a}, \boldsymbol{b})$. This reduction is permissible since e_0, π, ϵ_0 are all known constants, and $z_{\boldsymbol{a}}, z_{\boldsymbol{b}}$ can be derived from atomic features. In the context of the local crystal graph, the Coulomb potentials are denoted as $\mathbf{e}^c = -\epsilon''/d$. For the infinite crystal graph, the summations of Coulomb potentials, London dispersion potentials, and Pauli potentials are denoted as $\mathbf{s}^c = -\sum_d \epsilon'/d$, $\mathbf{s}^l = -\sum_d \epsilon/d^6$, and $\mathbf{s}^p = \sum_d e^{-\alpha d}$, respectively. We aggregate all three infinite features to derive the infinite features for the network, expressed as $s = s^c + s^l + ks^p$. In our experimentation, we fixed the values of the various parameters as follows: $\epsilon'' = 0.75$, $\epsilon' = 0.801$, $\epsilon = 0.074$, $\alpha = 3.0$, and $k = 0.145$. The selection of these specific constants was the result of manual grid searching.

- *The Input Block* of our model consists of two primary components: a ***Linear*** layer and an ***Embedding*** layer. For every node i in our model, we use the ***Linear*** layer to generate a 256-dimensional vector. This vector serves as the input node features for the first interaction layer. Simultaneously, we apply an ***Embedding*** layer for each edge in the model. This layer functions to map the Coulomb potentials and the summations of infinite potentials onto 256-dimensional embeddings. The Coulomb potentials undergo transformation using 256 Radial Basis Function (RBF) kernels, with the centers spanning a range from -4.0 to 4.0. Likewise, we transform the summations of infinite potentials using 64 RBF kernels, with centers also ranging from -4.0 to 4.0. Subsequently, we perform an up-projection on these transformed summations through a Multilayer Perceptron (MLP), resulting in a 256-dimensional output.

- *The Interaction Block* of our model comprises several ***Interaction*** layers. Each of these layers dynamically updates the feature vector of a given node i, taking into account the features of its neighboring nodes as well as the potential embeddings of the connecting edges. More specifically, for any neighboring node j of node i, the corresponding potential embeddings, denoted as $\mathbf{e}^c_{ij}$ and $\mathbf{s}_{ij}$, are generated by the ***Embedding*** layer. These embeddings are initially concatenated along the edge dimension. Subsequently, the embeddings are concatenated with the node features $\mathbf{x}_i$ and $\mathbf{x}_j$ along the feature dimension. The principal interaction pattern of a layer parallels the pattern adopted in the CGCNN (Xie & Grossman, 2018) as shown in the right side of Fig. 2.

- *The Readout Block* of our model incorporates an ***AvgPooling*** layer and a subsequent ***Linear*** layer. Initially, the ***AvgPooling*** layer is used to compile and aggregate features from all nodes within a graph. This aggregated feature set is then processed through the ***Linear*** layer. The function of this layer is to map the hidden dimension, represented as a 256-dimensional vector, to a final scalar output.

D.2. PotNet Improvement

In order to enhance the efficacy of the model, we incorporated two strategies: the inclusion of periodic table information and the application of a transformer structure for infinite potential features.

In our model, the charge information is inferred through atomic numbers, which could potentially introduce inaccuracies. To address this, we explicitly encode ten classes that are distinguished based on the categories of elements in the periodic table: alkali metals, alkaline earth metals, transition metals, post-transition metals, metalloids, reactive nonmetals, noble gases, lanthanides, and actinides. This encoding facilitates the model's learning of atomic properties. We have denoted this improved network model as PotNet-C.

Table 7. PotNet improving techniques by periodic table information and transformer structure on JARVIS formation energy and training time.

Model	MAE	Time/Epoch
	eV/atom	s
PotNet	0.0294	42
PotNet-C	0.0293	42
PotNet-T	0.0290	50

Furthermore, we incorporated a transformer operation (Ying et al., 2021; Yan et al., 2022; Rampášek et al., 2022) specifically applied to the infinite potential summation. This operation was particularly apt as the infinite potential summation is based on a fully-connected graph, an ideal fit for the transformer model. The transformer node output was then added to the local graph node output (Rampášek et al., 2022). This enhanced network structure, combined with the prior periodic table information, is referred to as PotNet-T.

Table 9. Ablation on our method with/without infinite potential summations in terms of test MAE on JARVIS dataset. The best results are shown in **bold**.

	Formation Energy	Bandgap(OPT)	Total energy	Bandgap(MBJ)	Ehull
Method	eV/atom	eV	eV/atom	eV	eV
PotNet w/o Infinite	0.0301	0.134	0.033	0.294	0.072
PotNet	**0.0294**	**0.127**	**0.032**	**0.272**	**0.055**

As illustrated in Table. 7, PotNet-C displays performance comparable to that of the original PotNet, while PotNet-T exhibits superior results to both PotNet and PotNet-C. However, it should be noted that the computational efficiency of PotNet-T is lower, with only a modest increase in performance.

D.3. Dataset Information

In Table 8, we detail the fundamental attributes of two distinct datasets: the Materials Project (MP) and the JARVIS 3D dataset. The considered characteristics comprise the size of each dataset, the average number of atoms per cell, the mean cell lengths, and the minimum cell length. As reflected in Table 8, the shortest cell length ranges approximately from 1$\AA$ to 2$\AA$. Many methods predicting crystal properties (Xie & Grossman, 2018; Schütt et al., 2017; Chen et al., 2019; Choudhary & DeCost, 2021) typically opt for a cutoff of either 4$\AA$ or 8$\AA$. This range, on average, only encompasses one-hop neighbors of a unit cell, and, at its most extensive, eight-hop neighbors.

Table 8. Dataset information on JARVIS and the Materials Project.

Information	JARVIS	MP
Size	55722	69239
Mean Atom Numbers	10.1	29.9
Mean Cell Length	5.95$\AA$	8.02$\AA$
Minimum Cell Length	0.99$\AA$	1.78$\AA$

D.4. Full Ablation on JARVIS Dataset

In addition to the results shown in the main body, we conduct a thorough ablation study of the infinite potential summations, employing the JARVIS dataset for this analysis. As highlighted in Table. 9, the incorporation of infinite summation features contributes to consistent performance augmentation across all metrics evaluated in this study. Remarkably, a substantial enhancement is observed in the predictive accuracy associated with the BandGap (MBJ) and Ehull properties.

D.5. Cutoff Experiments

In this section, we delve deeper into the investigation of two GNN approaches with differing cutoff values, with the aim of directly assessing the importance of complete interatomic potentials. The selection of these two methods is informed by their wide usage and the fact that they do not require additional graph construction techniques apart from radius crystal graph construction.

The first approach involves the application of the conventional GNN methodology, SchNet (Schütt et al., 2017). The second approach, GATGNN (Louis et al., 2020), incorporates the use of global attention. Additionally, we have incorporated our model without the infinite potential summation and local Coulomb potential into the study as the third approach. It should be noted that as the cutoff value increases, the feasibility of conducting training experiments diminishes due to the increase in time complexity. To facilitate training, we set the maximum number of an atom's neighboring atoms to 16.

Table 10. Experiments with varied cutoffs on JARVIS formation energy and dataset preprocessing time.

Cutoff	SchNet	GATGNN	PotNet	Time
Å	eV/atom	eV/atom	eV/atom	s
4	0.052	0.048	0.036	84
8	0.045	0.047	0.034	108
12	0.045	0.047	0.033	112
16	0.045	0.046	0.033	178
20	0.044	0.046	0.033	259
30	0.042	0.045	0.031	688
50	Unstable	0.043	0.030	3239

We conduct training and testing on these three methodologies using the JARVIS 3D dataset, maintaining the same dataset settings as in Matformer (Yan et al., 2022). We present the results pertaining to the formation energy in Table 10. As indicated in Table 10, it can be inferred that an increase in the cutoff value correlates with an enhancement in the performance of the three methodologies. However, at a cutoff value of 50$\AA$, the training of SchNet encounters a gradient explosion, rendering its final results unavailable. This issue may potentially stem from the modeling capacity of SchNet.

Further, we demonstrate the preprocessing time required for the entire JARVIS dataset, which comprises 55,722 crystals, in relation to different cutoff values in the fourth column of Table 10. It becomes evident that as the cutoff value increases, the preprocessing time escalates to an unmanageable extent.

E. Linear Energy Modeling using Infinite Potential Summation

In this section, we provide examples of calculations where the total energies of these materials can be directly approximated by linear combinations of infinite potential summations. These are special cases of Eqn. (6) with a linear embedded function G. Specifically, we evaluate the total energy per atom of $NaCl$ and two other materials (MgO, LiF) whose crystal structures are similar to $NaCl$. Since they are pure ionic crystals and Coulomb interactions dominate the system, we first consider their total electrostatic energy, i.e., the Coulomb potential summations.

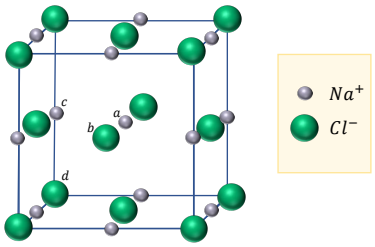

Figure 3. Crystal structure of $NaCl$.

Inspired by analytic continuation as discussed in Appendix C.4, we can approximate the total energy per atom of $NaCl$ by analytically continued infinite potential summations. The crystal structure of $NaCl$ is shown in Fig. 3. Due to the symmetry of the $NaCl$ cell, we only involve atoms a, b, c, d in our calculation. Here, atom a represents the body center Na^+, atom b represents the face center Cl^-, atom c represents the edge center Na^+, and atom d represents the corner Cl^-. Given N_A as Avogadro constant, r_0 as the minimum distance between Na and Cl, z_{Na}, z_{Cl} as charges of Na^+ and Cl^-, e_0 as the elementary charge constant, and ϵ_0 as the permittivity constant of free space, the total energy of Na^+ is approximated by the total Coulomb interactions with atom a such that

$$
\begin{aligned}
E_{Na} &= -N_A\Bigg[1\cdot\frac{z_{Na}z_{Na}e_0^2}{4\pi\epsilon_0}\sum_{\boldsymbol{u}\in\boldsymbol{A_a},\boldsymbol{u}\neq\boldsymbol{a}}\frac{1}{d(\boldsymbol{a},\boldsymbol{u})}+\frac{1}{2}6\cdot\frac{z_{Na}z_{Cl}e_0^2}{4\pi\epsilon_0}\sum_{\boldsymbol{u}\in\boldsymbol{A_b}}\frac{1}{d(\boldsymbol{a},\boldsymbol{u})} \\
&\quad+\frac{1}{4}12\cdot\frac{z_{Na}z_{Na}e_0^2}{4\pi\epsilon_0}\sum_{\boldsymbol{u}\in\boldsymbol{A_c}}\frac{1}{d(\boldsymbol{a},\boldsymbol{u})}+\frac{1}{8}8\cdot\frac{z_{Na}z_{Cl}e_0^2}{4\pi\epsilon_0}\sum_{\boldsymbol{u}\in\boldsymbol{A_d}}\frac{1}{d(\boldsymbol{a},\boldsymbol{u})}\Bigg] \\
&= -N_A\Bigg[\frac{z_{Na}z_{Na}e_0^2}{4\pi\epsilon_0 r_0}\sum_{\boldsymbol{u}\in\boldsymbol{A_a},\boldsymbol{u}\neq\boldsymbol{a}}\frac{1}{\tilde{d}(\boldsymbol{a},\boldsymbol{u})}+3\cdot\frac{z_{Na}z_{Cl}e_0^2}{4\pi\epsilon_0 r_0}\sum_{\boldsymbol{u}\in\boldsymbol{A_b}}\frac{1}{\tilde{d}(\boldsymbol{a},\boldsymbol{u})} \\
&\quad+3\cdot\frac{z_{Na}z_{Na}e_0^2}{4\pi\epsilon_0 r_0}\sum_{\boldsymbol{u}\in\boldsymbol{A_c}}\frac{1}{\tilde{d}(\boldsymbol{a},\boldsymbol{u})}+\frac{z_{Na}z_{Cl}e_0^2}{4\pi\epsilon_0 r_0}\sum_{\boldsymbol{u}\in\boldsymbol{A_d}}\frac{1}{\tilde{d}(\boldsymbol{a},\boldsymbol{u})}\Bigg] \\
&= -\frac{N_A|z_{Na}||z_{Cl}|e_0^2}{4\pi\epsilon_0 r_0}\Big[\tilde{\mathcal{S}}(\boldsymbol{a},\boldsymbol{a})-3\cdot\tilde{\mathcal{S}}(\boldsymbol{a},\boldsymbol{b})+3\cdot\tilde{\mathcal{S}}(\boldsymbol{a},\boldsymbol{c})-\tilde{\mathcal{S}}(\boldsymbol{a},\boldsymbol{d})\Big] \\
&\approx -\frac{N_A|z_{Na}||z_{Cl}|e_0^2}{4\pi\epsilon_0 r_0}\Big[-1.41864874-3\cdot(-0.04796615)+3\cdot(-0.29126077) \\
&\quad-(-0.40096799)\Big] \\
&\approx -\frac{N_A|z_{Na}||z_{Cl}|e_0^2}{4\pi\epsilon_0 r_0}\cdot(-1.7475646) \\
&\approx 8.81eV,
\end{aligned}
\tag{70}
$$

where $d(\boldsymbol{a},\boldsymbol{u})$ is the distance between atom $\boldsymbol{a}$ and $\boldsymbol{u}$, $\tilde{d}=d/r_0$ is the normalized distance, $\tilde{\mathcal{S}}$ is the infinite potential summation with $\tilde{d}$, approximated by our algorithm in Sec. 3.4, $\boldsymbol{A_a}$ denotes the set of atoms containing atom $\boldsymbol{a}$ and all its repetitions, and the coefficients $1, \frac{1}{2}6, \frac{1}{4}12, \frac{1}{8}8$ denote the fraction of atoms in a unit cell. We finally obtain a constant -1.7475646 and the total electrostatic energy approximation $8.81eV$ from our infinite potential summations. In fact, this constant -1.7475646 is exactly the famous Madelung constant M (Borwein et al., 1985) of $NaCl$. To obtain a more accurate total energy result by considering an additional repulsion term, we can derive the calculation result in Eqn. (70) to the famous Born-Landé equation (Born, 1921)

$$
E = -\frac{N_A|z^+||z^-|e_0^2 M}{4\pi\epsilon_0 r_0}(1-\frac{1}{n}), \tag{71}
$$

where z^+, z^- are the charges of cation and anion, M is the Madelung constant computed from Coulomb potential summations, and n is the Born exponent measuring the effect of repulsion. Choosing $n=9$, we can obtain an approximation for the total energy of $NaCl$ of 7.84 eV. Similarly, we also apply Eqn. (71) to MgO and LiF to approximate the total energy per atom of these crystals.

We show these approximations in Table 11 and it can be noticed that these approximations already give rough results compared to the ground truth total energy. This implies that our features can serve as a good starting point for machine learning models to learn the ground truth energy. Apparently, previous methods cannot achieve this due to the lack of such informative features. It is worth noting that the Madelung constant is typically unknown because those coefficients for the infinite potential summations depend on the charge distribution in the system, which we do not know at the beginning. Also, we already mention that these crystals are special cases of Eqn. (6) with a linear embedded function G, while G is typically a nonlinear function (Daw & Baskes, 1984). Therefore, the network serves the purpose of learning those coefficients to learn the Madelung constant and providing nonlinearity.

Table 11. Total energy per atom approximation of $NaCl$, MgO and LiF.

Formula	r_0	n	Ground Truth	Eqn. (71)
$NaCl$	282 pm	9	8.15 eV	7.84 eV
MgO	210 pm	6	39.33 eV	39.45 eV
LiF	201 pm	7	10.67 eV	10.60 eV