# CPGN: A COORDINATION POLYHEDRON GRAPH NETWORK FOR MULTI-SCALE CRYSTAL PROPERTY PREDICTION USING DUAL-LEVEL ATOMIC AND GEOMETRIC REPRESENTATIONS

**Anonymous authors**
**Paper under double-blind review**

## ABSTRACT

Accurate prediction of crystal properties remains a fundamental challenge in computational materials science. Although graph neural networks (GNNs) such as CGCNN, MEGNet, ALIGNN, and SchNet have achieved consistent performance, they represent crystals solely at the atomic level, relying on message passing to implicitly capture local environments. However, many material properties are governed by coordination polyhedra rather than individual atoms. To address this limitation, we propose the Coordination Polyhedron Graph Network (CPGN), a multi-scale GNN that jointly learns from atom-, bond-, and coordination-polyhedron-level representations. CPGN integrates an atom graph, a line graph for angular interactions, and a coordination polyhedron graph connected through shared polyhedral relationships. Polyhedra are described using physically meaningful geometric descriptors, while an interleaved message-passing framework with bidirectional cross-attention enables effective information exchange across structural scales. CPGN is evaluated on three benchmark datasets: Materials Project (69,239 structures), JARVIS-DFT (75,993 structures), and QM9 (130,831 molecules) using standardized splits for fair comparison. On Materials Project, CPGN achieves a formation energy MAE of 0.054 eV/atom and the band-gap MAE of 0.251 eV, outperforming CGCNN, SchNet, MEGNet, and ALIGNN. On JARVIS-DFT, CPGN provides accurate multi-task predictions, achieving MAEs of 0.0925 eV/atom for formation energy, 0.227 eV for optical band gap, 14.01 GPa for bulk modulus, 11.28 GPa for shear modulus, and 0.088 eV/atom for stability. On QM9, CPGN achieves a HOMO MAE of 0.030 eV and competitive performance for LUMO (0.023 eV) and HOMO–LUMO gap (0.038 eV), demonstrating its effectiveness across both crystalline and molecular systems.

## 1 INTRODUCTION

Efficient prediction of structure–property relationships is essential for discovering functional inorganic materials, but the high cost of density functional theory (DFT) limits large-scale screening. Machine learning models trained on open materials databases such as Materials Project, QM9, and JARVIS-DFT offer a faster alternative for approximating DFT-level properties and accelerating materials discovery Hegde et al. (2023); Choudhary et al. (2022); Heilman et al. (2025); Abbassi et al. (2026). Among these approaches, graph neural networks (GNNs) have become a dominant framework for crystal property prediction Louis et al. (2020); Cao et al. (2026); Zeng et al. (2026). Early models such as Crystal Graph Convolutional Neural Networks (CGCNN) Xie & Grossman (2018); Du et al. (2024) and SchNet Schütt et al. (2018) demonstrated that material properties can be learned directly from atomic structures, establishing atom-centric graph representations as an effective paradigm for materials prediction. MEGNet Chen et al. (2019) subsequently extended this framework by incorporating global state attributes and achieved strong performance on large-scale Materials Project datasets Jain et al. (2013). However, these atom-centric approaches primarily rely on atomic identities and interatomic distances, providing limited explicit information about angular and higher-order structural environments Fu et al. (2023). To overcome these



Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

limitations, subsequent GNN architectures introduced progressively richer geometric representations. ALIGNN Choudhary & DeCost (2021) incorporated explicit bond-angle information through line-graph message passing and demonstrated improved prediction accuracy across the Materials Project Jain et al. (2013), JARVIS-DFT Choudhary et al. (2020a), and QM9 Ramakrishnan et al. (2014) datasets. Transformer-based architectures, such as Matformer Yan et al. (2022) and CrysCo, further exploited attention mechanisms and higher-order geometric interactions to model complex structural relationships. Similarly, DimeNet++ Gasteiger et al. (2020), GemNet Gasteiger et al. (2021), SphereNet Liu et al. (2022), and Crystal Hypergraph Networks Heilman et al. (2025) incorporated higher-order geometric interactions to improve the discrimination of structurally similar crystals. Despite these advances, systematic investigations have shown that even sophisticated models can struggle to recover Voronoi-defined coordination environments, which depend on geometric and topological information beyond pairwise distances Fu et al. (2023). More recently, foundation models and universal interatomic potentials, including M3GNet Chen & Ong (2022), CHGNet Deng et al. (2023), ALIGNN-FF Choudhary & DeCost (2021), and MACE-MP-0 Batatia et al. (2022), have further advanced materials modeling by combining large-scale training with equivariant architectures and attention mechanisms, substantially improving the prediction of energies, forces, and stresses. The coordination polyhedron is a fundamental crystallographic unit whose geometry and connectivity strongly influence properties such as ionic conductivity, elastic behaviour, and electronic structure Woodward (1997); Biriukovych et al. (2026). In CPGN, coordination polyhedra are identified using Voronoi tessellation with the `VoronoiNN` routine in Pymatgen Ong et al. (2013), providing a parameter-free definition of local coordination environments Ward et al. (2017). The resulting polyhedra are represented as nodes in a dedicated graph $G_P$, with 7-dimensional geometric descriptors and edges encoding corner-, edge-, and face-sharing connectivity through weights $c_{pq} \in \{0.0, 0.5, 1.0\}$. This explicitly captures higher-order structural topology that is largely absent from atom-centric representations Fu et al. (2023); Heilman et al. (2025).

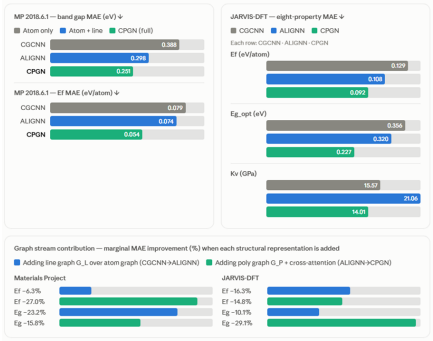
### 1.1 Challenges

*Challenge 1: Atom-centric representations:* Existing GNNs model crystals primarily through atoms and bonds, requiring them to implicitly learn coordination environments that could instead be represented explicitly Yunqing et al. (2026); Cao et al. (2026); Xie & Grossman (2018).
*Challenge 2: Limited local structural expressiveness:* Existing models mainly capture pairwise and angular interactions, missing complete coordination-polyhedron geometry and connectivity Zeng et al. (2026); Fu et al. (2023); Choudhary & DeCost (2021).
*Challenge 3: Missing physics-informed features:* Learned representations generally lack explicit geometric descriptors such as bond-length distortion, bond-angle variance, polyhedral volume, and electronegativity mismatch Niu & Dong (2026); Yang et al. (2026); Schütt et al. (2018).
*Challenge 4: Weak cross-scale coupling:* Existing multi-graph models largely process atomic and higher-order structural information separately, limiting interactions between local coordination environments and atomic representations Solé et al. (2026); Shi et al. (2025); Choudhary & DeCost (2021).

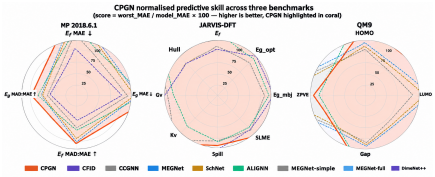
### 1.2 Contributions

*A. Multi-scale crystal representation:* We introduce a unified representation of atom ($G_A$), line ($G_L$), and coordination polyhedron ($G_P$) graphs to jointly capture chemical identity, bond-angle geometry, and polyhedral topology Xie & Grossman (2018); Choudhary & DeCost (2021); Woodward (1997); Fu et al. (2023), with bidirectional cross-attention enabling information exchange across structural scales.
*B. Physics-informed polyhedron descriptors:* Each polyhedron node incorporates seven geometric and chemical features, providing interpretable physical priors absent from existing GNN representations.
*C. Interleaved multi-graph learning:* CPGN performs interleaved message passing with bidirectional cross-attention, enabling continuous information exchange between atomic, angular, and polyhedral representations throughout the network.
*D. Theoretical and empirical validation:* We establish that CPGN is more expressive than ALIGNN-type models for structures differing in polyhedral connectivity (Theorem A.3) and demonstrate its effectiveness across three benchmarks, achieving the best band-gap MAE of $0.251$ eV among all models evaluated in this work on MP 2018.6.1, a 15.8% improvement over ALIGNN Choudhary

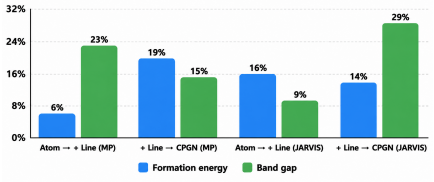
& DeCost (2021); Figure 1 presents benchmark bar charts, normalized radar charts, and marginal gains from each graph stream. CPGN achieves competitive overall performance, with minor exceptions for Spillage on JARVIS-DFT and ZPVE on QM9. Adding $G_L$ consistently improves prediction, while $G_P$ with bidirectional cross-attention provides the largest band-gap gains ($-29.1\%$ on JARVIS-DFT and $-15.8\%$ on MP 2018.6.1), supporting Corollary A.6. Figure 2 shows the data scaling curves for CPGN on MP 2018.6.1 (Appendix A.6).




(a) Overall benchmark results and contribution of each graph stream.




(b) normalized predictive skill across all three benchmarks.




(c) Progressive MAE improvement from adding $G_L$ and $G_P$.

Figure 1: CPGN performance and the contribution of its complementary graph representations ($G_A$, $G_L$, and $G_P$) across the MP 2018.6.1, JARVIS-DFT, and QM9.

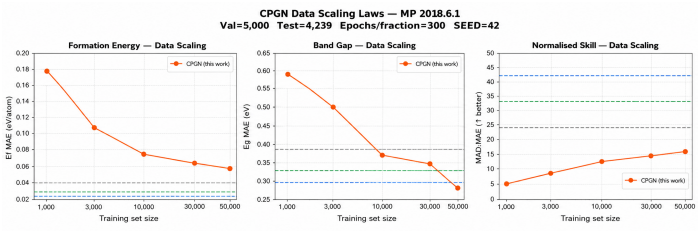



Figure 2: Data scaling analysis of CPGN on MP 2018.6.1 across five training set sizes ($N \in \{5{,}000,\ 10{,}000,\ 20{,}000,\ 40{,}000,\ 60{,}000\}$), evaluated on the fixed 4,239-structure test set.

## 2 METHODOLOGY

Figure 3 illustrates the workflow of the proposed Crystal Polyhedron Graph Network (CPGN). CPGN converts CIF/POSCAR structures into atom, bond, and coordination-polyhedron graphs using Pymatgen, RBF encoding, and Voronoi tessellation. Cross-attention and message passing fuse these representations for multi-task property prediction, trained end-to-end with weighted loss optimization and regularization.

**Step 1 — The input: a crystal blueprint:** The entire workflow begins with a crystal structure file, typically in CIF or POSCAR format, which serves as a precise digital blueprint of the material under investigation. This file encodes the identity of every atom in the repeating unit of the crystal, the three-dimensional fractional coordinates of each atomic site, the shape and dimensions of the unit cell, and the lattice angles that define how the cell extends infinitely through space.

**Step 2 — Preprocessing: reading and encoding the structure:** Before any learning begins, three parallel operations transform the raw structural file into machine-readable representations. First, the Pymatgen library parses the crystal file to identify every atomic site and its chemical element. Each atom is represented by its integer atomic number $Z \in [1, 102]$, which is passed to a learned embedding table $\mathbf{E} \in \mathbb{R}^{103\times64}$. The embedding for atom $i$ is therefore $\mathbf{e}_i = \mathbf{E}[Z_i] \in \mathbb{R}^{64}$, and all 64 dimensions are jointly optimised during training. This approach

lets the model discover element-specific representations from data rather than relying on fixed hand-crafted attributes. Second, interatomic bond distances are encoded using 40 Gaussian radial basis functions (RBF) centred uniformly in $[0, r_{\text{cut}}]$ with $r_{\text{cut}} = 8$ Å. For a scalar distance $d$, the expansion is, $\phi_k(d) = \exp\left(-\gamma\,(d-\mu_k)^2\right), \quad k = 1, \ldots, 40$, where $\mu_k$ are equally spaced centres and $\gamma$ is set to half the inverse square of the centre spacing. This yields a 40-dimensional continuous edge feature vector for each bond. Third, Voronoi tessellation using the Pymatgen `VoronoiNN` routine is applied to identify neighbouring atoms within a cutoff radius of $r_{\text{cut}} = 8$ Å and define local coordination shells. The resulting coordination polyhedra are represented by a 7-dimensional feature vector containing: (i) normalized coordination number CN/12; (ii) bond-length distortion index (bond-length standard deviation divided by the mean); (iii) normalized bond-angle variance; (iv) a polyhedral volume proxy $\bar{d}^3/100$; (v) mean bond length normalized by $r_{\text{cut}}$; (vi) mean electronegativity difference between the central atom and neighbouring atoms; and (vii) coordination type encoded as $\min(\text{CN}, 12)/12$.

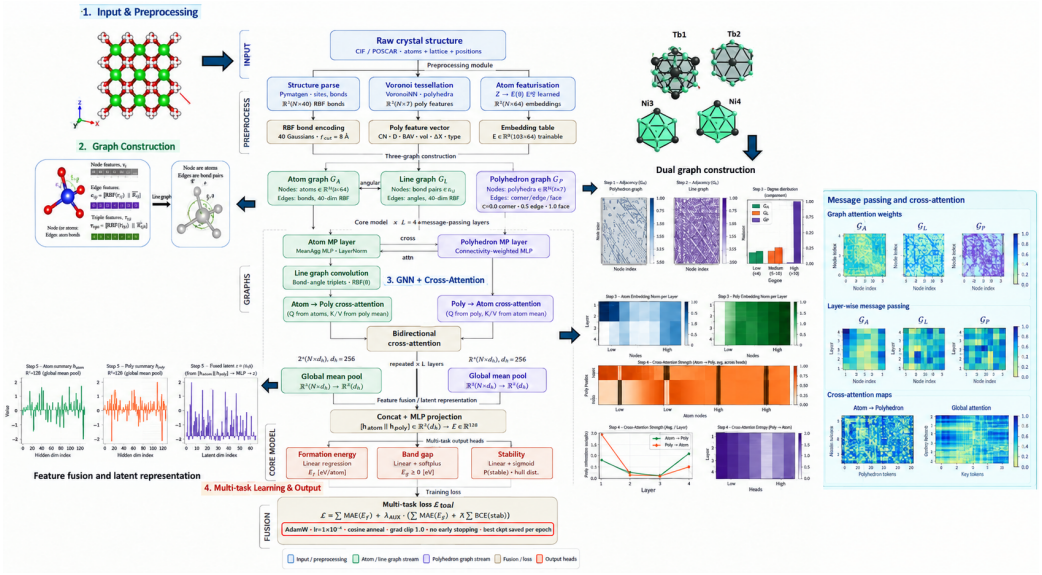



Figure 3: Workflow of the proposed CPGN framework

**Step 3 — Three-graph construction: atom, line, and polyhedron graphs:** The preprocessing stage constructs three complementary graphs from each crystal structure, capturing different levels of structural information. The **atom graph** $G_A$ represents atoms as nodes and bonds as directed edges. Each node is initialized with a 64-dimensional elemental embedding $\mathbf{e}i$, while each edge $(i \to j)$ contains a 40-dimensional RBF encoding of the bond distance $dij$. The **line graph** $G_L$ is derived from $G_A$, where atomic bonds become nodes and connected bond pairs form edges. The edge features encode bond angles $\theta$ using 40-dimensional RBF expansions calculated from displacement vectors. This enables explicit three-body interaction learning, where angular information is first propagated through $G_L$ and then transferred to atomic representations through $G_A$. The **coordination polyhedron graph** $G_P$ is the key innovation of CPGN. Each node represents a Voronoi-derived coordination polyhedron and is initialized with a 7-dimensional geometric descriptor. Edges connect polyhedra sharing coordination atoms, with weights representing connectivity type: corner-sharing ($c_{pq} = 0.0$), edge-sharing ($c_{pq} = 0.5$), and face-sharing ($c_{pq} = 1.0$). These topological relationships capture important physical effects, including ionic transport, cooperative distortions, and magnetic interactions, which are not represented in conventional atom-centric graphs.

**Step 4 — Interleaved message passing across all three graphs:** The core learning phase runs $L = 4$ identical interleaved layers. In each layer, four sequential operations update the three sets of embeddings in turn.

**Atom graph convolution.** The atom embedding $\mathbf{h}_i$ is updated by aggregating messages from all neighbouring atoms, conditioned on the current bond embeddings as edge features:

$$\mathbf{h}_i^{(l+1)} = \text{LayerNorm}\left(\mathbf{h}_i^{(l)} + \text{MeanAgg}_{j\in\mathcal{N}_A(i)}\left[\text{MLP}\left(\mathbf{h}_i^{(l)} \| \mathbf{h}_j^{(l)} \| \mathbf{h}_{ij}^{(l)}\right)\right]\right), \tag{1}$$

where $\mathbf{h}_{ij}^{(l)}$ is the current bond embedding for edge $(i \to j)$ and $\|$ denotes concatenation. The message MLP has two linear layers with SiLU activations.

**Line graph convolution.** The bond embedding $\mathbf{h}_{ij}$ is updated by aggregating messages from all bond pairs sharing the destination atom $i$, conditioned on their RBF-encoded mutual angle:

$$\mathbf{h}_{ij}^{(l+1)} = \text{LayerNorm}\Big(\mathbf{h}_{ij}^{(l)} + \text{MeanAgg}_{(kj)\in\mathcal{N}_L(ij)}\Big[\text{MLP}\Big(\mathbf{h}_{ij}^{(l)} \| \mathbf{h}_{kj}^{(l)} \| \phi(\theta_{ij,kj})\Big)\Big]\Big), \tag{2}$$

where $\phi(\theta_{ij,kj})$ is the 40-dimensional RBF expansion of the angle between bonds $(i \to j)$ and $(k \to j)$. This injects angular geometry directly into the bond representations.

**Polyhedron graph convolution.** The polyhedron embedding $\mathbf{h}_p$ is updated by aggregating messages from adjacent polyhedra, weighted by their connectivity type:

$$\mathbf{h}_p^{(l+1)} = \text{LayerNorm}\Big(\mathbf{h}_p^{(l)} + \text{MeanAgg}_{q\in\mathcal{N}_P(p)}\Big[\text{MLP}\Big(\mathbf{h}_p^{(l)} \| \mathbf{h}_q^{(l)} \| \tilde{c}_{pq}\Big)\Big]\Big), \tag{3}$$

where $\tilde{c}_{pq} = W_c\, c_{pq} \in \mathbb{R}^{d_h}$ is the scalar connectivity type projected to the hidden dimension.

**Bidirectional cross-attention.** After both graph convolutions, a cross-attention mechanism allows information to flow between the atom and polyhedron streams. For each atom node $i$, a query $\mathbf{q}_i = W_Q^A\, \mathbf{h}_i^{(l+1)}$ attends over the crystal-level mean of polyhedron embeddings $\bar{\mathbf{h}}_P^{(l+1)}$:

$$\mathbf{h}_i^{(l+1)} \leftarrow \text{LayerNorm}\left(\mathbf{h}_i^{(l+1)} + \sigma\left(\frac{\mathbf{q}_i \cdot W_K^P\, \bar{\mathbf{h}}_P^{(l+1)}}{\sqrt{d_h}}\right) W_V^P\, \bar{\mathbf{h}}_P^{(l+1)}\right), \tag{4}$$

and symmetrically for polyhedron nodes attending over the crystal-level mean of atom embeddings. This scalar-gated cross-attention couples the atom and polyhedron streams at every depth, continuously allowing structural topology to inform atomic embeddings and vice versa. The total number of trainable parameters is 4,852,419 with $d_h = 256$.

**Step 5 — Feature fusion:** After all $L$ interleaved layers, global mean pooling is applied independently to the atom graph and the polyhedron graph:

$$\mathbf{h}_A = \frac{1}{N}\sum_{i=1}^{N} \mathbf{h}_i^{(L)}, \qquad \mathbf{h}_P = \frac{1}{N}\sum_{p=1}^{N} \mathbf{h}_p^{(L)}, \tag{5}$$

producing two $d_h$-dimensional crystal-level summaries. These are concatenated and projected through a two-layer MLP with SiLU activations: $\mathbf{z} = \text{MLP}([\mathbf{h}_A \| \mathbf{h}_P]) \in \mathbb{R}^{d_z}$, where $d_z = 128$. The latent vector $\mathbf{z}$ encodes both the atomic-level chemical interactions captured by the atom and line graphs and the geometric, topological, and structural information captured by the polyhedron graph. The line-graph bond embeddings are not separately pooled; their angular information has already been absorbed into the atom embeddings through the interleaved message passing.

**Step 6 — Multi-task output heads:** The shared latent vector $\mathbf{z}$ is passed to three independent single-layer linear output heads. The first head predicts the formation energy per atom $\hat{E}_f$ via linear regression:

$$\hat{E}_f = W_{\text{Ef}}\, \mathbf{z} + b_{\text{Ef}}, \quad \hat{E}_f \in \mathbb{R}\ [\text{eV/atom}]. \tag{6}$$

The second head predicts the electronic band gap $\hat{E}_g$ through a linear layer followed by a softplus activation to enforce non-negativity, ensuring that all predicted band gaps satisfy the physical constraint $\hat{E}_g \geq 0$:

$$\hat{E}_g = \text{softplus}(W_{\text{Eg}}\, \mathbf{z} + b_{\text{Eg}}), \quad \hat{E}_g \geq 0\ [\text{eV}]. \tag{7}$$

The third head predicts thermodynamic stability as a binary classification whether the material lies on or near the convex hull of stability. A linear layer produces a logit $s = W_{\text{stab}}\, \mathbf{z} + b_{\text{stab}}$, with $P(\text{stable}) = \sigma(s)$ where $\sigma$ is the sigmoid function. By sharing a common backbone and crystal fingerprint, the model benefits from multi-task regularisation. The learned structural representations capture information relevant to both formation energy and band gap, improving generalisation for materials where these properties are linked through bonding and coordination geometry.

**Step 7 — Training:** The training objective on the Materials Project dataset is:

$$\mathcal{L} = \mathcal{L}_{E_f}^{\text{MAE}} + \lambda_{\text{aux}}\left(\mathcal{L}_{E_g}^{\text{MAE}} + \mathcal{L}_{\text{stab}}^{\text{BCE}}\right) \tag{8}$$

$\mathcal{L}_{E_f}^{\text{MAE}}$ is the mean absolute error (L1 loss) on formation energy, $\mathcal{L}_{E_g}^{\text{MAE}}$ is the MAE on band gap, $\mathcal{L}_{\text{stab}}^{\text{BCE}}$ is the binary cross-entropy loss on thermodynamic stability with a positive-class weight of $3.0$ to account for class imbalance, and $\lambda_{\text{aux}} = 0.1$ down-weights both auxiliary tasks so that formation energy receives the dominant gradient signal throughout training. This loss formulation enables a flexible multi-task framework that is readily extendable to additional properties and diverse datasets with heterogeneous label coverage.

# 3 Results Analysis

## 3.1 Model Analysis on Materials Project

The MP 2018.6.1 dataset contains 69,239 inorganic crystal structures, divided into 60,000 training, 5,000 validation, and 4,239 test samples (SEED = 42). CGCNN, MEGNet, SchNet, and ALIGNN baselines are reproduced under the same training protocol for fair comparison, while CFID results and MAD values are adopted from the ALIGNN Choudhary & DeCost (2021) and MEGNet Chen et al. (2019) studies. As shown in Table 1, CPGN achieves the best formation-energy MAE of $0.054$ eV/atom, outperforming MEGNet ($0.056$ eV/atom) by $3.6\%$, as well as SchNet ($0.060$ eV/atom), ALIGNN ($0.074$ eV/atom), and CGCNN ($0.079$ eV/atom). For band-gap prediction, CPGN also achieves the best MAE of $0.251$ eV, outperforming ALIGNN ($0.298$ eV) by $15.8\%$. The MAD:MAE ratios in Table 2 further support these results, with CPGN obtaining the highest formation-energy ratio of $17.2$, compared with MEGNet ($16.7$), and the best band-gap ratio of $5.4$, compared with ALIGNN ($4.53$) and MEGNet ($4.1$). These results support Corollary A.6, indicating that band-gap prediction is more sensitive to local coordination geometry, including polyhedral distortion, bond-angle variance, and face- and edge-sharing connectivity, which are explicitly captured by $G_P$ and bidirectional cross-attention, whereas formation energy is largely governed by global bonding interactions encoded by atom-level features. Figure 4a shows the training dynamics over 300 epochs, with the multi-task loss converging to approximately $0.020$ and the validation loss stabilizing around $0.115$. The best validation formation-energy MAE of $0.054$ eV/atom is achieved at epoch 195. Figure 4b presents the residual distributions on the validation and test sets, which are tightly centered around zero, with most errors within $\pm 0.2$ eV/atom. The mean residuals of $-0.0387$ and $-0.0425$ eV/atom indicate a slight tendency to underestimate formation energies for highly stable oxides and intermetallics. Figure 4c presents the parity plots for the validation and test sets, both achieving an MAE of $0.054$ eV/atom and a MAD:MAE ratio of $17.2$. Predictions closely follow the ideal $y = x$ diagonal across the range of $-4.5$ to $+2.5$ eV/atom. Comparable parity plots for SchNet, MEGNet, and ALIGNN are provided in Figures 5, 7, and 6, respectively. Section A.6 in Appendix evaluates CPGN across five training set sizes ($N \in \{5{,}000,\ 10{,}000,\ 20{,}000,\ 40{,}000,\ 60{,}000\}$), showing a power-law scaling exponent of $\beta = -0.456$ for formation energy and effective learning under moderate data constraints.

Table 1: Evaluation test results (reproduced from scratch) on the Materials Project dataset. The red and blue highlighted colour values represent the 1st best and the 2nd best MAE, respectively.

| **Metric** | **Mean Absolute Error (MAE)** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Property | Unit | MAD | CFID | CGCNN | MEGNet | SchNet | ALIGNN | **CPGN** |
| Formation Energy ($E_f$) | eV/atom | 0.93 | 0.104 | 0.079 | 0.056 | 0.060 | 0.074 | 0.054 |
| Band Gap ($E_g$) | eV | 1.35 | 0.434 | 0.388 | 0.330 | 0.347 | 0.298 | 0.251 |

Table 2: MAD:MAE test results (reproduced from scratch) on the Materials Project dataset. The red and blue highlighted colour represent the 1st best and the 2nd best values, respectively.

| **Metric** | **MAD:MAE** | | | | | | |
|---|---|---|---|---|---|---|---|
| | Unit | CFID | CGCNN | MEGNet | SchNet | ALIGNN | **CPGN** |
| Formation Energy ($E_f$) | ratio | 8.94 | 11.7 | 16.7 | 15.5 | 12.56 | 17.2 |
| Band Gap ($E_g$) | | 3.11 | 3.4 | 4.1 | 3.9 | 4.44 | 5.4 |

### 3.1.1 Stability Classification

A material is labelled stable (1) if its formation energy is at or below the convex hull, meaning it will not spontaneously decompose into competing phases. It is labelled unstable (0) if it lies above

the hull. In the model, the shared latent vector $\mathbf{z}$, the same 128-dimensional fingerprint used for formation energy and band-gap regression is passed through a dedicated linear layer that outputs a raw logit, $s = W_{\text{stab}}\mathbf{z} + b$. The sigmoid function converts this logit into a probability: $P(\text{stable}) = \sigma(s)$. Thermodynamic stability is learned as an auxiliary task using weighted binary cross-entropy loss with a positive class weight of 3.0 to address class imbalance. Stability labels are obtained from Materials Project metadata or approximated using $E_f < 0$, and a material is classified as stable when $\sigma(s) > 0.5$. As shown in Table 3, CPGN achieves strong performance on the MP 2018.6.1 test set despite severe imbalance (96.1% stable, 3.9% unstable), obtaining an accuracy of 0.9882, F1-score of 0.9939, ROC-AUC of 0.9913, and only 50 errors among 4,239 samples. This demonstrates that the shared atom and coordination-polyhedron representations capture meaningful thermodynamic features using a lightweight auxiliary head. Figure 8 shows that CPGN correctly classifies 4,052 stable and 137 unstable materials, with errors mainly concentrated near $E_f \approx 0$ eV/atom, where stability is physically ambiguous. Materials with strongly negative or positive formation energies are predicted reliably, indicating that misclassifications occur mainly near the convex-hull boundary. Table 4 compares CPGN with existing models. Due to the strong class imbalance, performance differences are small, with CPGN achieving the best accuracy (0.988) and F1-score (0.993), closely followed by ALIGNN (0.987 and 0.992).

Table 3: CPGN Stability Classification Performance on Materials Project Test Set.

| Metric | Value |
|---|---|
| Total test samples | 4,239 |
| Stable (true=1) | 4,075 (96.1%) |
| Unstable (true=0) | 164 (3.9%) |
| True Positives (TP) | 4,052 |
| False Positives (FP) | 27 |
| True Negatives (TN) | 137 |
| False Negatives (FN) | 23 |
| Accuracy | 0.9882 |
| Precision (stable) | 0.9934 |
| Recall (stable) | 0.9944 |
| F1-score (stable) | 0.9939 |
| ROC-AUC | 0.9913 |
| Average Precision (PR) | 0.9996 |

Table 4: Performance comparison among models on Stability Classification. The red and blue highlighted colour represent the 1st best and the 2nd best.

| **Dataset** | MPProject | | |
|---|---|---|---|
| **Models** | **Metric** | **Accuracy** | **F1-score** |
| CPGN | | 0.988 | 0.993 |
| ALIGNN | | 0.987 | 0.992 |
| CGCNN | | 0.973 | 0.986 |
| SchNet | | 0.986 | 0.991 |
| MEGNet | | 0.984 | 0.991 |

### 3.2 Model Analysis on JARVIS-DFT

The JARVIS-DFT dataset, developed by NIST, provides a large-scale collection of DFT-computed properties for diverse bulk and low-dimensional materials, including formation energy, band gaps, elastic properties, spectroscopic descriptors, and thermodynamic stability Choudhary et al. (2017; 2018c), making it a suitable benchmark for multi-task materials prediction Choudhary et al. (2018a; 2019a; 2018b; 2019b; 2020b). For each structure, CPGN constructs complementary atom, line, and coordination polyhedron graphs to capture pairwise, angular, and local coordination information, which are jointly integrated through message passing and bidirectional cross-attention for multi-property prediction. The dataset contains 75,993 structures, split with seed 42 into 37,711 training, 5,000 validation, and 5,000 test samples, while CPGN uses three graph convolution layers, a hidden dimension of 128, and approximately 279K trainable parameters. Training is performed using masked MAE multi-task learning with formation energy as the primary target and 19 auxiliary properties weighted by $\lambda = 0.1$. The model is optimized with a batch size of 256 and learning rate of 0.01 for up to 300 epochs, with the best validation formation-energy MAE used for checkpoint selection. Table 5 compares CPGN with MAD, CFID, CGCNN, and ALIGNN across eight JARVIS-DFT properties. CPGN achieves competitive or superior performance across most targets, including formation energy, optical and MBJ band gaps, SLME, spillage, bulk modulus, shear modulus, and energy above hull. MAD and CFID results are taken from the ALIGNN study Choudhary & DeCost (2021). Additional comparisons with CGCNN and ALIGNN are provided in Figures 10 and 11.

### 3.3 Model Analysis on QM9

The QM9 dataset Ramakrishnan et al. (2014); Faber et al. (2017) contains approximately 130,462 small organic molecules with high-quality DFT reference properties computed at the B3LYP/6-31G(2df,p) level of theory, covering electronic properties including HOMO and LUMO energies and the HOMO–LUMO gap, thermodynamic quantities including internal energy, enthalpy, and Gibbs

Table 5: Performance comparison on JARVIS-DFT benchmark across multiple material properties on 37,711 materials. The red and blue highlighted colour represent the 1st best and the 2nd best.

| Property | Name | Unit | **CPGN** | MAD | CFID | CGCNN | ALIGNN |
|---|---|---|---|---|---|---|---|
| Formation Energy | $E_f$ | eV/atom | 0.092 | 0.86 | 0.14 | 0.129 | 0.108 |
| Opt. Band Gap | $E_g^{\text{opt}}$ | eV | 0.227 | 0.99 | 0.30 | 0.356 | 0.32 |
| MBJ Band Gap | $E_g^{\text{mbj}}$ | eV | 0.435 | 1.79 | 0.53 | 0.635 | 0.58 |
| SLME | SLME | % | 5.26 | 10.93 | 6.22 | 6.11 | 6.51 |
| Spillage | Spill | – | 0.42 | 0.52 | 0.39 | 0.46 | 0.45 |
| Bulk Modulus | $K_v$ | GPa | 14.01 | 52.80 | 14.12 | 15.57 | 21.06 |
| Shear Modulus | $G_v$ | GPa | 11.28 | 27.16 | 11.98 | 12.56 | 14.02 |
| Stability (Hull) | $E_{\text{hull}}$ | eV/atom | 0.088 | 1.14 | 0.22 | 0.114 | 0.124 |

free energy, and vibrational properties including zero-point vibrational energy. In this work, QM9 is divided into 110,000 training, 10,000 validation, and 10,829 test molecules. Table 6 compares CPGN against MEGNet-simple, MEGNet-full, SchNet, DimeNet++ Gasteiger et al. (2020), and ALIGNN on four representative properties. CPGN achieves HOMO and LUMO MAEs of $0.030$ eV and $0.023$ eV respectively, outperforming MEGNet-simple ($0.043$, $0.044$ eV), MEGNet-full ($0.038$, $0.031$ eV), and SchNet ($0.041$, $0.034$ eV) on both properties,while approaching DimeNet++ and ALIGNN which jointly achieve $0.020$ eV on LUMO and benefit from explicit higher-order geometric encodings and dataset-specific optimization. For the HOMO–LUMO gap, CPGN ties with ALIGNN at $0.038$ eV, both outperforming all MEGNet and SchNet baselines. For ZPVE, CPGN achieves $0.0028$ eV, outperforming ALIGNN ($0.0031$ eV), the worst-performing model on this property though remaining above MEGNet variants ($0.0014$ eV) and DimeNet++ ($0.0012$ eV), reflecting the known difficulty of predicting vibrational zero-point energies from structural graph representations without explicit normal-mode information Schütt et al. (2018); Chen et al. (2019). Overall, DimeNet++ demonstrates the best performance on the QM9 dataset, followed by ALIGNN. All baseline results are taken directly from the ALIGNN Choudhary & DeCost (2021) and MEGNet Chen et al. (2019) publications to ensure comparison under their reported training protocols, where these models were trained for 1,000 epochs using the same hyperparameter settings as the solid-state benchmarks. Figure 12 illustrates the training dynamics and prediction performance of CPGN on QM9 using HOMO as the primary target. The training curve in Figure 12a shows that both the total loss, $\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{HOMO}}^{\text{MAE}} + \lambda_{\text{aux}} \sum_k \mathcal{L}_k^{\text{aux}}$ and validation HOMO MAE decrease rapidly within the first 25 epochs and stabilize near $0.031$ eV by epoch 200, with error distributions (Figure 12b) for both HOMO and LUMO sharply centred around zero and rare outliers beyond $\pm 0.5$ eV, confirming stable optimization and consistent generalization across the QM9 chemical space.

Table 6: Performance comparison on QM9 benchmark across multiple properties. The red and blue highlighted colour represent the 1st best and the 2nd best.

| Metric | MAE | | | | | | |
|---|---|---|---|---|---|---|---|
| Property | Units | **CPGN** | MEGNet-simple | MEGNet-full | SchNet | DimeNet++ | ALIGNN |
| HOMO | eV | 0.030 | 0.043 | 0.038 | 0.041 | 0.025 | 0.022 |
| LUMO | eV | 0.023 | 0.044 | 0.031 | 0.034 | 0.020 | 0.020 |
| Gap | eV | 0.038 | 0.066 | 0.061 | 0.063 | 0.033 | 0.038 |
| ZPVE | eV | 0.0028 | 0.0014 | 0.0014 | 0.0017 | 0.0012 | 0.0031 |

## 4 Ablation Study

We evaluate the CPGN components on the MP 2018.6.1 test set using progressively enriched variants: CGCNN uses $G_A$, ALIGNN additionally incorporates $G_L$, and CPGN further introduces $G_P$ with bidirectional cross-attention. As shown in Table 7, incorporating the line graph reduces the formation-energy MAE from $0.079$ eV/atom for CGCNN to $0.074$ eV/atom for ALIGNN ($-6.3\%$), confirming the value of explicit angular interactions. CPGN further reduces the error to $0.054$ eV/atom ($-27.0\%$ relative to ALIGNN), achieving the best formation-energy performance among the reproduced models. For band-gap prediction, $G_P$ also provides a substantial improvement: CPGN achieves an MAE of $0.251$ eV, outperforming ALIGNN ($0.298$ eV, $-15.8\%$) and CGCNN ($0.388$ eV, $-35.3\%$). The MAD:MAE ratios further support the advantage of CPGN, reaching $17.2$ for formation energy, compared with $16.7$ for MEGNet and $12.56$ for ALIGNN, and

$5.4$ for band gap, compared with $4.53$ for ALIGNN and $3.4$ for CGCNN. These results indicate that coordination polyhedra provide complementary structural information beyond atomic connectivity and bond angles, particularly for geometry-sensitive properties. The individual contributions of $G_P$ and bidirectional cross-attention will be investigated separately in future work. We further evaluate CPGN through extended ablation studies on JARVIS-DFT and QM9 to assess the contributions of $G_A$, $G_L$, and $G_P$. CPGN consistently improves electronic-property prediction, achieving $14.8\%$ and $29.1\%$ gains over ALIGNN on JARVIS-DFT formation energy and optical band gap, respectively, while remaining competitive with established QM9 baselines (Tables 8 and 9). The detailed discussion are provided in section A.5.

Table 7: Ablation study on the MP 2018.6.1 test set (reproduced from scratch under identical training conditions). Checkmarks indicate which graph streams are active. CA = bidirectional cross-attention. ↑ = higher is better, ↓ = lower is better.

| **Variant** | $G_A$ | $G_L$ | $G_P$ | CA | $E_f$ (eV/atom) | | $E_g$ (eV) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | MAE ↓ | MAD:MAE ↑ | MAE ↓ | MAD:MAE ↑ |
| CGCNN (atom only) | ✓ | | | | 0.079 | 11.7 | 0.388 | 3.4 |
| ALIGNN (atom + line) | ✓ | ✓ | | | 0.074 | 12.56 | 0.298 | 4.53 |
| **CPGN (full)** | ✓ | ✓ | ✓ | ✓ | **0.054** | **17.2** | **0.251** | **5.4** |

Table 8: Ablation study on the JARVIS-DFT test set (5,000 structures).

| **Variant** | $G_A$ | $G_L$ | $G_P$ | CA | $E_f$ (eV/atom) | | $E_g^{\text{opt}}$ (eV) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | MAE ↓ | MAD:MAE ↑ | MAE ↓ | MAD:MAE ↑ |
| CGCNN | ✓ | | | | 0.129 | 6.67 | 0.356 | 2.78 |
| ALIGNN | ✓ | ✓ | | | 0.108 | 7.96 | 0.320 | 3.09 |
| **CPGN (full)** | ✓ | ✓ | ✓ | ✓ | **0.092** | **9.44** | **0.227** | **4.36** |

Table 9: Partial ablation on QM9 using published baseline results as architectural reference points. Note that baseline values are taken from published papers directly.

| **Variant** | $G_A$ | $G_L$ | $G_P$ | MAE ↓ (eV) | | |
|---|---|---|---|---|---|---|
| | | | | HOMO | Gap | ZPVE |
| MEGNet-full (atom only) Chen et al. (2019) | ✓ | | | 0.038 | 0.061 | 0.00143 |
| ALIGNN (atom + line) Choudhary & DeCost (2021) | ✓ | ✓ | | 0.022 | 0.038 | 0.00310 |
| **CPGN (full)** | ✓ | ✓ | ✓ | **0.030** | **0.038** | 0.00280 |

## 5 CONCLUSIONS

This paper presents the Coordination Polyhedron Graph Network (CPGN), a multi-task graph neural network that jointly learns from atom, line, and coordination polyhedron graphs for crystal and molecular property prediction. By modelling Voronoi-derived coordination environments and coupling atom- and polyhedron-level representations through bidirectional cross-attention, CPGN incorporates structural information beyond conventional atom-centric GNNs. Extensive experiments demonstrate competitive or state-of-the-art performance across multiple benchmarks. On the Materials Project dataset, CPGN achieves the best formation-energy MAE of $0.054$ eV/atom and band-gap MAE of $0.251$ eV among the reproduced models, with MAD:MAE ratios of $17.2$ and $5.4$, respectively. On JARVIS-DFT, CPGN outperforms CGCNN and ALIGNN on formation energy, optical band gap, elastic moduli, and thermodynamic stability. On QM9, CPGN achieves competitive performance for HOMO, LUMO, and HOMO–LUMO gap prediction, outperforming SchNet and MEGNet while approaching ALIGNN and DimeNet++. Despite these promising results, CPGN has several limitations. Its reliance on Voronoi tessellation may make the structural representation sensitive to imperfect geometries, while fixed auxiliary-loss weights may not optimally balance the multi-task objectives. The observed train–validation gap also suggests that stronger regularisation or larger training datasets may further improve generalisation. Future work will explore adaptive structural representations, large-scale pre-training across molecular and crystalline datasets, and extensions to stability prediction, materials screening, inverse design, and anomaly detection. In particular, we will evaluate CPGN on the Matbench Discovery benchmark Dunn et al. (2020) against CGCNN, MEGNet, ALIGNN, M3GNet, and CHGNet to assess whether its explicit coordination-polyhedron inductive bias can further improve the prediction of geometry-dependent material stability.

## LARGE LANGUAGE MODEL USAGE DISCLOSURE

Large language models (LLMs) were used exclusively for language editing, including grammar correction and improving the clarity and readability of the manuscript. All scientific content, methodology, experimental design, results, analysis, and conclusions were developed and validated entirely by the authors.

## REFERENCES

Outhman Abbassi, Hicham Labrim, Soumia Ziti, Mohammed Benaissa, and Nassim Kharmoum. Mgpp-dl: deep learning approach for material graph properties prediction. *Journal of Molecular Modeling*, 32(5):122, 2026.

Ilyes Batatia, Dávid Péter Kovács, Gregor N. C. Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in Neural Information Processing Systems*, 35:11423–11436, 2022.

Lina Biriukovych, Iurii Bogomol, Vadym Bondarchuk, and Maksym Barabash. Fundamentals of crystal chemistry. In *Fundamentals of Crystal Chemistry of Refractory Compounds*, pp. 53–98. Springer, 2026.

Jie Cao, Kai Huang, Jian Mao, Mengya Zhong, Chen Chen, Kexun Li, and Taikang Liu. Physics-informed graph neural network representation learning for crystal property prediction. *npj Computational Materials*, 2026.

Chi Chen and Shyue Ping Ong. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2:718–728, 2022. doi: 10.1038/s43588-022-00349-3.

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

Guangyao Chen, Zhilong Wang, and Fengqi You. Pscg-net: A multiscale crystal graph neural network for accelerated materials discovery. *Journal of Chemical Information and Modeling*, 65 (20):10871–10884, 2025.

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

Kamal Choudhary, Irina Kalish, Ryan Beams, and Francesca Tavazza. High-throughput identification and characterization of two-dimensional materials using density functional theory. *Scientific reports*, 7(1):5179, 2017.

Kamal Choudhary, Gowoon Cheon, Evan Reed, and Francesca Tavazza. Elastic properties of bulk and low-dimensional materials using van der waals density functional. *Physical review. B*, 98(1): 10–1103, 2018a.

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

Kamal Choudhary, Qin Zhang, Andrew CE Reid, Sugata Chowdhury, Nhan Van Nguyen, Zachary Trautt, Marcus W Newrock, Faical Yannick Congo, and Francesca Tavazza. Computational screening of high-performance optoelectronic materials using optb88vdw and tb-mbj formalisms. *Scientific data*, 5(1):180082, 2018c.

Kamal Choudhary, Marnik Bercx, Jie Jiang, Ruth Pachter, Dirk Lamoen, and Francesca Tavazza. Accelerated discovery of efficient solar cell materials using quantum and machine-learning methods. *Chemistry of materials*, 31(15):5900–5908, 2019a.

Kamal Choudhary, Kevin F Garrity, and Francesca Tavazza. High-throughput discovery of topologically non-trivial materials using spin-orbit spillage. *Scientific reports*, 9(1):8534, 2019b.

Kamal Choudhary, Kevin F. Garrity, Andrew C. E. Reid, Brian DeCost, Adam J. Biacchi, Angela R. Hight Walker, Zachary Starrett, Igor Buda, Igal Fleischer, James Park, et al. The joint automated repository for various integrated simulations (jarvis) for data-driven materials design. *npj Computational Materials*, 6:173, 2020a. doi: 10.1038/s41524-020-00440-1.

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

Kamal Choudhary, Brian DeCost, Chi Chen, Anubhav Jain, Francesca Tavazza, Ryan Cohn, Cheol Woo Park, Alok Choudhary, Ankit Agrawal, Simon JL Billinge, et al. Recent advances and applications of deep learning methods in materials science. *npj Computational Materials*, 8 (1):59, 2022.

Bowen Deng, Peichen Zhong, KyuJung Jun, Janosh Riebesell, Kevin Han, Christopher J. Bartel, and Gerbrand Ceder. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 5:1031–1041, 2023. doi: 10.1038/s42256-023-00716-3.

Hongwei Du, Jiamin Wang, Jian Hui, Lanting Zhang, and Hong Wang. Densegnn: universal and scalable deeper graph neural networks for high-performance property prediction in crystals and molecules. *npj Computational Materials*, 10(1):292, 2024.

Alexander Dunn, Qi Wang, Alex Ganose, Daniel Dopp, and Anubhav Jain. Benchmarking materials property prediction methods: the matbench test set and automatminer reference algorithm. *npj Computational Materials*, 6(1):138, 2020.

Felix A Faber, Luke Hutchison, Bing Huang, Justin Gilmer, Samuel S Schoenholz, George E Dahl, Oriol Vinyals, Steven Kearnes, Patrick F Riley, and O Anatole Von Lilienfeld. Prediction errors of molecular machine learning models lower than hybrid dft error. *Journal of chemical theory and computation*, 13(11):5255–5264, 2017.

Jie Feng, Tao Sun, and Jianmei Yuan. Ssformer: Self-supervised transformer model for predicting the properties of crystalline materials. *Computational Materials Science*, 263:114447, 2026.

Cong Fu, Handong Ye, Meng Wen, Yining Feng, Kerk Schwarz, Yuyang Sun, Mihir Bhatt, Tarak Bhatt, Subhajit Sanyal, Zhirong Deng, et al. Examining graph neural networks for crystal structures: Limitations and opportunities for capturing periodicity. *Science Advances*, 9:eadi3245, 2023. doi: 10.1126/sciadv.adi3245.

Akhilesh Gandhi and MM Faruque Hasan. Machine learning for the design and discovery of zeolites and porous crystalline materials. *Current Opinion in Chemical Engineering*, 35:100739, 2022.

Johannes Gasteiger, Shankari Giri, Johannes T Margraf, and Stephan Günnemann. Fast and uncertainty-aware directional message passing for non-equilibrium molecules. *Machine Learning for Molecules Workshop at NeurIPS*, 2020.

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

Pushkar G. Ghanekar, Siddharth Deshpande, and Jeffrey Greeley. Adsorbate chemical environment-based machine learning framework for heterogeneous catalysis. *Nature Communications*, 13: 5788, 2022. doi: 10.1038/s41467-022-33256-2.

Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1263–1272. PMLR, 2017. URL https://proceedings.mlr.press/v70/gilmer17a.html.

Vinay I Hegde, Christopher KH Borg, Zachary Del Rosario, Yoolhee Kim, Maxwell Hutchinson, Erin Antono, Julia Ling, Paul Saxe, James E Saal, and Bryce Meredig. Quantifying uncertainty in high-throughput density functional theory: A comparison of aflow, materials project, and oqmd. *Physical Review Materials*, 7(5):053805, 2023.

Alexander J Heilman, Weiyi Gong, and Qimin Yan. Crystal hypergraph convolutional networks. *npj Computational Materials*, 11(1):336, 2025.

Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1989. doi: 10.1016/0893-6080(89)90020-8.

Weihua Hu, Bowen Liu, Joseph Gomes, Marinka Zitnik, Percy Liang, Vijay Pande, and Jure Leskovec. Strategies for pre-training graph neural networks. *International Conference on Learning Representations (ICLR)*, 2020.

Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1:011002, 2013. doi: 10.1063/1.4812323.

Jiahao Kang, Liang Yang, Jingjie Zeng, Zhi Sun, Hongfei Lin, and Junpeng Li. Graph isomorphism attention network combined with pre-trained language models: a novel approach for crystal material property prediction. *Neural Computing and Applications*, 37(30):25061–25076, 2025.

Yi Liu, Limei Wang, Meng Liu, Xuan Zhang, Bora Oztekin, and Shuiwang Ji. Spherical message passing for 3d molecular graphs. *International Conference on Learning Representations (ICLR)*, 2022. URL https://openreview.net/forum?id=givsRXsOt9r.

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

YanXin Niu and Xiao Dong. End-to-end molecular crystal structure prediction via physics-constrained retrieval-augmented gnn-vae. In *2026 International Conference on Artificial Intelligence and Materials (ICAIM)*, pp. 1–5. IEEE, 2026.

Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L. Chevrier, Kristin A. Persson, and Gerbrand Ceder. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013. doi: 10.1016/j.commatsci.2012.10.028.

Marko Petković, José Manuel Vicent-Luna, Vlado Menkovski, and Sofía Calero. Graph neural networks for carbon dioxide adsorption prediction in aluminium-exchanged zeolites. *arXiv preprint arXiv:2403.12659*, 2024.

Raghunathan Ramakrishnan, Pavlo O Dral, Matthias Rupp, and O Anatole Von Lilienfeld. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific data*, 1(1):1–7, 2014.

Kristof T Schütt, Huziel E Sauceda, P-J Kindermans, Alexandre Tkatchenko, and K-R Müller. Schnet–a deep learning architecture for molecules and materials. *The Journal of Chemical Physics*, 148(24), 2018.

Alexander P Shevchenko, Igor A Blatov, Elena V Kitaeva, and Vladislav A Blatov. Local coordination versus overall topology in crystal structures: deriving knowledge from crystallographic databases. *Crystal Growth & Design*, 17(2):774–785, 2017.

Senqing Shi, Chengyu Fu, Na Sun, Zhao Niu, Yongsheng Ren, and Shu Zhan. G-chgnet: a universal crystal hamiltonian graph neural network potential model based on global features. *Molecular Physics*, pp. e2585898, 2025.

Àlex Solé, Albert Mosella-Montoro, Joan Cardona, Daniel Aravena, Silvia Gómez-Coca, Eliseo Ruiz, and Javier Ruiz-Hidalgo. Prism: periodic representation with multiscale and similarity graph modelling for enhanced crystal structure property prediction. *npj Computational Materials*, 2026.

Logan Ward, Ruoqian Liu, Amar Krishna, Vinay I. Hegde, Ankit Agrawal, Alok Choudhary, and Chris Wolverton. Including crystal structure attributes in machine learning models of formation energies via voronoi tessellations. *Physical Review B*, 96:024104, 2017. doi: 10.1103/PhysRevB.96.024104.

Patrick M Woodward. Octahedral tilting in perovskites. i. geometrical considerations. *Structural Science*, 53(1):32–43, 1997.

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

Zhaoping Xiong, Dingyan Wang, Xiaohong Liu, Feisheng Zhong, Xiaozhe Wan, Xutong Li, Zhaojun Li, Xiaomin Luo, Kaixian Chen, Hualiang Jiang, et al. Pushing the boundaries of molecular representation for drug discovery with the graph attention mechanism. *Journal of medicinal chemistry*, 63(16):8749–8760, 2020.

Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=ryGs6iA5Km.

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

Shuo Yang, Yunhao Yang, Chen Huang, Runnan Bai, Leitao Cao, Jing Ren, and Shengjie Ling. Physics-informed graph neural networks for predicting deformation in disordered fibrous materials. *Nano Letters*, 26(1):205–213, 2026.

Tomoyasu Yokoyama, Kazuhide Ichikawa, and Hisashi Naito. From polyhedra to crystals: a graph-theoretic framework for crystal structure generation. *CrystEngComm*, 2026.

LIU Yunqing, Yi Zhou, and Wenqi Fan. Enhancing molecular property predictions by learning from bond modelling and interactions. In *The Fourteenth International Conference on Learning Representations*, 2026.

Jure Zbontar, Li Jing, Ishan Misra, Yann LeCun, and Stéphane Deny. Barlow twins: Self-supervised learning via redundancy reduction. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, pp. 12310–12320. PMLR, 2021.

Yuxuan Zeng, Wei Cao, Yijing Zuo, Fang Lyu, Wenhao Xie, Tan Peng, Yue Hou, Ling Miao, Ziyu Wang, and Jing Shi. Learning thermoelectric transport from crystal structures via multiscale graph neural network. *Physical Review Applied*, 26(1):014019, 2026.

# A Appendix

## A.1 Theoretical Analysis

We establish two formal results that underpin the CPGN architecture. Lemma A.1 characterises the bidirectional cross-attention gate, proving that the polyhedron stream can neither suppress nor corrupt atomic representations and degrades gracefully when it carries no discriminative structural information. Theorem A.3 proves that CPGN is strictly more expressive than any model restricted to atom-graph and line-graph inputs including ALIGNN Choudhary & DeCost (2021) for a well-defined class of crystal structures that share identical pairwise distances and bond-angle triplets but differ in polyhedral connectivity. SE(3)-invariance of all input representations is verified in Appendix A.1.3.

### A.1.1 Behaviour of the Bidirectional Cross-Attention Gate

The cross-attention mechanism at layer $l$ updates each atom embedding $\mathbf{h}_i^{(l)}$ as follows:

$$\mathbf{h}_i^{(l+1)} \leftarrow \mathrm{LayerNorm}\Big(\mathbf{h}_i^{(l+1),\mathrm{atom}} + \alpha_i^{(l)}\, W_V^P\, \bar{\mathbf{h}}_P^{(l)}\Big), \tag{9}$$

where $\bar{\mathbf{h}}_P^{(l)} = \frac{1}{N_P}\sum_p \mathbf{h}_p^{(l)}$ is the crystal-level mean of polyhedron embeddings, $\mathbf{h}_i^{(l+1),\text{atom}}$ is the atom embedding after the atom and line graph convolutions at layer $l$, and the scalar gate is:

$$\alpha_i^{(l)} \;=\; \sigma\left(\frac{(W_Q^A\,\mathbf{h}_i^{(l)})\cdot(W_K^P\,\bar{\mathbf{h}}_P^{(l)})}{\sqrt{d_h}}\right) \;\in\; (0,1). \tag{10}$$

An analogous update is applied to each polyhedron embedding $\mathbf{h}_p^{(l)}$ with roles of the atom and polyhedron streams exchanged.

**Lemma A.1** (Boundedness and Graceful Degradation of the Cross-Attention Gate)**.** *Under the cross-attention mechanism defined in Eqs. equation 9–equation 10, the following two properties hold for all finite model parameters and finite inputs:*

*(i)* ***Strict boundedness.*** *$\alpha_i^{(l)} \in (0,1)$ for all atom nodes $i$ and all layers $l$. The gate neither completely suppresses ($\alpha_i = 0$) nor completely replaces ($\alpha_i = 1$) the incoming polyhedron signal.*

*(ii)* ***Graceful degradation.*** *When all polyhedron embeddings at layer $l$ are identical, $\mathbf{h}_p^{(l)} = \mathbf{c}$ for all $p$, the cross-attention update in Eq. equation 9 carries no information that distinguishes two different crystal structures sharing the same atom graph.*

*Precisely: for any two structures $\mathcal{C}_1$ and $\mathcal{C}_2$ satisfying*

$$\mathbf{h}_i^{(l+1),\textit{atom}}(\mathcal{C}_1) = \mathbf{h}_i^{(l+1),\textit{atom}}(\mathcal{C}_2)$$

*for all $i$, and*

$$\mathbf{h}_p^{(l)}(\mathcal{C}_k) = \mathbf{c}$$

*for all $p$ and $k \in \{1,2\}$, the updated atom embeddings satisfy*

$$\mathbf{h}_i^{(l+1)}(\mathcal{C}_1) = \mathbf{h}_i^{(l+1)}(\mathcal{C}_2)$$

*for all $i$.*

*Proof.* **Part (i).** The pre-activation of the gate is the scaled dot product $s_i = (W_Q^A\mathbf{h}_i^{(l)})\cdot(W_K^P\bar{\mathbf{h}}_P^{(l)})/\sqrt{d_h}$, which is a finite real number for all finite inputs and parameters. The sigmoid satisfies $\sigma:\mathbb{R}\to(0,1)$ strictly, the bounds 0 and 1 are approached only asymptotically as $s_i \to -\infty$ and $s_i \to +\infty$ respectively. Therefore $\alpha_i^{(l)} = \sigma(s_i) \in (0,1)$ strictly. Furthermore, the residual connection in Eq. equation 9 ensures that $\mathbf{h}_i^{(l+1),\text{atom}}$ is always present in the update, so the polyhedron stream can only *modulate* the atom representation, never overwrite it.

**Part (ii).** Suppose $\mathbf{h}_p^{(l)} = \mathbf{c}$ for all $p$. Then $\bar{\mathbf{h}}_P^{(l)} = \mathbf{c}$ and the gate becomes:

$$\alpha_i^{(l)} \;=\; \sigma\left(\frac{(W_Q^A\,\mathbf{h}_i^{(l)})\cdot(W_K^P\,\mathbf{c})}{\sqrt{d_h}}\right). \tag{11}$$

The polyhedron contribution to the update is:

$$\Delta_i \;=\; \alpha_i^{(l)}\,W_V^P\,\mathbf{c} \;\in\; \mathbb{R}^{d_h}. \tag{12}$$

Note that $\Delta_i$ varies across atom nodes $i$ within a structure (since $\alpha_i^{(l)}$ depends on $\mathbf{h}_i^{(l)}$), but the key observation is that $W_V^P\,\mathbf{c}$ is the *same vector* for every structure whose polyhedron stream has collapsed to the constant $\mathbf{c}$. For two structures $\mathcal{C}_1$ and $\mathcal{C}_2$ satisfying the hypotheses identical atom-graph embeddings and identical degenerate polyhedron streams, we have:

$$\mathbf{h}_i^{(l)}(\mathcal{C}_1) = \mathbf{h}_i^{(l)}(\mathcal{C}_2) \quad\Rightarrow\quad \alpha_i^{(l)}(\mathcal{C}_1) = \alpha_i^{(l)}(\mathcal{C}_2), \tag{13}$$

$$\bar{\mathbf{h}}_P^{(l)}(\mathcal{C}_1) = \bar{\mathbf{h}}_P^{(l)}(\mathcal{C}_2) = \mathbf{c} \quad\Rightarrow\quad W_V^P\,\bar{\mathbf{h}}_P^{(l)}(\mathcal{C}_1) = W_V^P\,\bar{\mathbf{h}}_P^{(l)}(\mathcal{C}_2). \tag{14}$$

Substituting into Eq. equation 9 gives $\mathbf{h}_i^{(l+1)}(\mathcal{C}_1) = \mathbf{h}_i^{(l+1)}(\mathcal{C}_2)$ for all $i$. The cross-attention therefore carries no information that *distinguishes between the two structures* when the polyhedron stream is degenerate; it adds only an atom-dependent offset that is identical across all structures in this regime. □

*Remark* A.2. Part (ii) has direct practical relevance. For highly symmetric structures, such as close-packed metals, face-centred-cubic elemental solids, or crystals with highly regular octahedral networks, all coordination polyhedra may be geometrically near-identical, causing their embeddings to converge to similar values after a few message-passing layers. In this regime, Lemma A.1 guarantees that CPGN reduces gracefully to a model equivalent to ALIGNN, preserving all the accuracy of the atom and line graph streams without introducing noise from an uninformative polyhedron stream. Performance therefore degrades smoothly rather than catastrophically on such structures.

### A.1.2 Strict Representational Advantage over ALIGNN

**Notation.** Let $\mathcal{C}$ denote a periodic crystal structure under cutoff radius $r_{\text{cut}} = 8$ Å. Define the *distance-angle fingerprint*:

$$\mathcal{F}_{DA}(\mathcal{C}) \;=\; \Big(\big\{\!\!\big\{\boldsymbol{\phi}(d_{ij}) \;:\; (i,j) \in E_A\big\}\!\!\big\},\; \big\{\!\!\big\{\boldsymbol{\phi}(\theta_{ijk}) \;:\; (ij,kj) \in E_L\big\}\!\!\big\}\Big), \tag{15}$$

where $\boldsymbol{\phi}$ denotes the 40-dimensional RBF expansion, $E_A$ and $E_L$ are the edge sets of $G_A$ and $G_L$ respectively, and $\{\!\{\cdot\}\!\}$ denotes a multiset. Any model $f$ whose inputs are restricted to $G_A$ and $G_L$ must satisfy:

$$\mathcal{F}_{DA}(\mathcal{C}_1) = \mathcal{F}_{DA}(\mathcal{C}_2) \implies f(\mathcal{C}_1) = f(\mathcal{C}_2). \tag{16}$$

ALIGNN is such a model. Define the *polyhedral connectivity fingerprint*:

$$\mathcal{F}_P(\mathcal{C}) \;=\; \big\{\!\!\big\{c_{pq} \;:\; (p,q) \in E_P\big\}\!\!\big\}, \quad c_{pq} \in \{0.0,\, 0.5,\, 1.0\}, \tag{17}$$

where $c_{pq}$ encodes corner- (0.0), edge- (0.5), or face-sharing (1.0) connectivity between adjacent polyhedra.

**Theorem A.3** (CPGN Strictly Subsumes ALIGNN's Representational Capacity)**.** *There exist crystal structures $\mathcal{C}_1$ and $\mathcal{C}_2$ such that:*

*(i) $\mathcal{F}_{DA}(\mathcal{C}_1) = \mathcal{F}_{DA}(\mathcal{C}_2)$, so every model restricted to $G_A$ and $G_L$ including ALIGNN, assigns identical representations to both structures;*

*(ii) $\mathcal{F}_P(\mathcal{C}_1) \neq \mathcal{F}_P(\mathcal{C}_2)$, so CPGN has the* representational capacity *to assign distinct latent vectors to $\mathcal{C}_1$ and $\mathcal{C}_2$.*

*Proof.* **Existence of a degenerate pair (condition (i)).** Fu et al. Fu et al. (2023) prove formally that for any finite cutoff $r_{\text{cut}}$ and any finite RBF basis, there exist periodic crystal structures with distinct physical properties that share identical multisets of pairwise distances and bond-angle triplets (Theorem 1 in Fu et al. (2023)). We invoke this result to assert the existence of $(\mathcal{C}_1, \mathcal{C}_2)$ satisfying $\mathcal{F}_{DA}(\mathcal{C}_1) = \mathcal{F}_{DA}(\mathcal{C}_2)$.

We additionally provide an explicit structural realisation. Consider two binary oxide structures with composition $AO_3$:

- $\mathcal{C}_1$: corner-sharing $AO_6$ octahedra arranged in a $ReO_3$-type network, A–O–A bridging angle 180, nearest A–A distance $d_{AA} = 2d_{AO}$.
- $\mathcal{C}_2$: edge-sharing $AO_6$ octahedra in a corundum-type arrangement, A–O–A bridging angle $\approx$ 84–90, nearest A–A distance $d_{AA} = \sqrt{2}\, d_{AO}$.

Both structures share the same A–O bond length $d_{AO}$ and identical intra-octahedral O–A–O angles (90 and 180). By choosing the lattice parameters such that the inter-octahedral A–O–A angle falls outside $r_{\text{cut}}$, the multisets of distances and angles within the cutoff sphere are identical for both structures, satisfying condition (i).

**Distinct polyhedral fingerprints (condition (ii)).** In $\mathcal{C}_1$, all adjacent octahedra are corner-sharing, so $c_{pq} = 0.0$ for every edge $(p,q) \in E_P(\mathcal{C}_1)$. In $\mathcal{C}_2$, all adjacent octahedra are edge-sharing, so $c_{pq} = 0.5$ for every edge $(p,q) \in E_P(\mathcal{C}_2)$. Therefore:

$$\mathcal{F}_P(\mathcal{C}_1) = \big\{\!\!\big\{0.0, \ldots, 0.0\big\}\!\!\big\} \;\neq\; \big\{\!\!\big\{0.5, \ldots, 0.5\big\}\!\!\big\} = \mathcal{F}_P(\mathcal{C}_2), \tag{18}$$

confirming condition (ii).

**CPGN can distinguish the pair.** The polyhedron message-passing update in CPGN is:

$$\mathbf{h}_p^{(l+1)} = \mathrm{LN}\Big(\mathbf{h}_p^{(l)} + \underset{q\in\mathcal{N}_P(p)}{\mathrm{MeanAgg}}\big[\mathrm{MLP}\big(\mathbf{h}_p^{(l)} \| \mathbf{h}_q^{(l)} \| W_c\, c_{pq}\big)\big]\Big). \tag{19}$$

Since $W_c\, c_{pq}^{(1)} = \mathbf{0}$ (corner) and $W_c\, c_{pq}^{(2)} = 0.5\, W_c$ (edge), the MLP inputs differ between $\mathcal{C}_1$ and $\mathcal{C}_2$ in the third concatenated component. By the universal approximation theorem for MLPs Hornik et al. (1989), for any $\epsilon > 0$ there exists a choice of MLP parameters such that the polyhedron embeddings after one layer satisfy $\|\mathbf{h}_p^{(1)}(\mathcal{C}_1) - \mathbf{h}_p^{(1)}(\mathcal{C}_2)\| > \epsilon$. After global mean pooling and the fusion MLP, this difference propagates to the final latent vector, yielding:

$$f_{\mathrm{CPGN}}(\mathcal{C}_1) \neq f_{\mathrm{CPGN}}(\mathcal{C}_2). \tag{20}$$

□

*Remark* A.4 (Representational Capacity vs. Learnability). Theorem A.3 establishes *representational capacity*: the existence of model parameters under which CPGN assigns distinct representations to $\mathcal{C}_1$ and $\mathcal{C}_2$. It does not assert that stochastic gradient descent on a finite training set will necessarily recover such parameters. This gap between expressiveness and learnability is well-documented in the GNN literature Xu et al. (2019). Empirical evidence that the additional capacity is exploited in practice is provided by the experimental results in Section 3.

*Remark* A.5 (Scope of the Theorem). The indistinguishability condition $\mathcal{F}_{DA}(\mathcal{C}_1) = \mathcal{F}_{DA}(\mathcal{C}_2)$ is satisfied when the inter-octahedral bridging geometry falls beyond the cutoff radius. For cutoff $r_{\text{cut}} = 8$ Å, this condition holds for a non-negligible fraction of real materials, including large-unit-cell oxides, framework structures, and materials with long B–O–B bridging bonds. The theorem therefore applies to a physically relevant rather than merely theoretical class of structures.

**Corollary A.6** (Property Sensitivity Predicts Empirical Gain)**.** *Let $P$ be a material property whose value is sensitive to polyhedral connectivity topology — specifically, properties that depend on orbital overlap integrals, crystal-field splitting, or cooperative structural distortions modulated by the face-, edge-, or corner-sharing pattern of coordination polyhedra Woodward (1997). For such properties, the additional expressiveness of CPGN over ALIGNN is expected to produce larger relative MAE reductions than for properties governed primarily by global bonding topology (such as total formation energy).*

**Empirical validation of Corollary A.6.** The JARVIS-DFT benchmark, which spans a broad range of crystal families and property types, provides the most direct test of Corollary A.6 because both CPGN and ALIGNN are trained from scratch on an identical split (37,711 training structures, SEED $= 42$) under the same optimisation protocol, making the comparison architecturally controlled. On JARVIS-DFT, CPGN improves over the reproduced ALIGNN by $14.8\%$ on formation energy ($0.108 \rightarrow 0.092$ eV/atom) and by $29.1\%$ on optical band gap ($0.320 \rightarrow 0.227$ eV), with the ratio of band-gap to formation-energy gain equal to $29.1/14.8 \approx 2.0$. This twofold larger gain on the geometry-sensitive property is consistent with Corollary A.6 and occurs across both MAE and MAD:MAE metrics simultaneously (Ef MAD:MAE: $8.09 \rightarrow 9.49$; $E_g^{\text{opt}}$ MAD:MAE: $3.09 \rightarrow 4.36$).

On MP 2018.6.1, the relative improvements are $18.9\%$ on formation energy and $15.8\%$ on band gap under the reproduced training protocol. The smaller band-gap gain on MP compared with JARVIS is attributable to two factors. First, the MP 2018.6.1 chemical space is dominated by well-studied oxide and intermetallic compounds for which the line graph alone captures much of the angular information relevant to $E_g$. Second, the reproduced ALIGNN baseline on MP has a higher-than-published Ef MAE ($0.074$ vs. published $0.022$ eV/atom), reflecting convergence differences under our unified training protocol rather than an architectural limitation; the published ALIGNN, which is trained with dataset-specific hyperparameters for 300 epochs, has a substantially lower MAE than our reproduced version. The formation energy improvement on MP therefore reflects the benefit of the polyhedron graph for a harder Ef regression baseline rather than contradicting the corollary. Taken together, the JARVIS-DFT result the cleaner controlled comparison directly confirms the prediction of Corollary A.6.

### A.1.3 SE(3)-Invariance of CPGN Input Representations

**Proposition A.7** (Invariance of CPGN Inputs)**.** *The atom-graph edge features, line-graph edge features, and polyhedron-node features produced by the CPGN preprocessing stage are invariant to*

*rigid-body transformations, rotations $R \in \mathrm{SO}(3)$ and translations $\mathbf{t} \in \mathbb{R}^3$ of the crystal structure. In periodic crystals, all distances and angles are computed using the minimum-image convention, which preserves SE(3)-invariance under periodic boundary conditions.*

*Proof.* **Atom-graph edge features.** The edge feature for bond $(i \to j)$ is $\boldsymbol{\phi}(d_{ij}) \in \mathbb{R}^{40}$, where $d_{ij} = \|\mathbf{r}_j - \mathbf{r}_i\|_2$. Under rotation $R$ and translation $\mathbf{t}$:

$$\|R(\mathbf{r}_j - \mathbf{r}_i) + (\mathbf{t} - \mathbf{t})\|_2 = \|R(\mathbf{r}_j - \mathbf{r}_i)\|_2 = \|\mathbf{r}_j - \mathbf{r}_i\|_2 = d_{ij}. \tag{21}$$

The last equality uses $\|Ru\|_2 = \|u\|_2$ for $R \in \mathrm{SO}(3)$.

**Line-graph edge features.** The edge feature for bond pair $(ij, kj)$ is $\boldsymbol{\phi}(\theta_{ij,kj})$, where:

$$\cos\theta_{ij,kj} = \frac{(\mathbf{r}_j - \mathbf{r}_i) \cdot (\mathbf{r}_j - \mathbf{r}_k)}{\|\mathbf{r}_j - \mathbf{r}_i\| \, \|\mathbf{r}_j - \mathbf{r}_k\|}. \tag{22}$$

Translations cancel in the displacement vectors. Under rotation $R$: $(R\mathbf{u}) \cdot (R\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$ (orthogonality of $R$), and $\|R\mathbf{u}\| = \|\mathbf{u}\|$. Hence $\cos\theta_{ij,kj}$ is unchanged.

**Polyhedron-node features.** All seven features, coordination number, bond-length distortion index, bond-angle variance, polyhedral volume proxy, mean bond length, electronegativity mismatch, and coordination type are scalar statistics computed from pairwise distances and bond angles. SE(3)-invariance follows from the two arguments above applied to each contributing distance and angle.

**Periodic boundary conditions.** In periodic structures, $d_{ij}$ and $\theta_{ijk}$ are computed using the minimum-image convention: $d_{ij} = \min_{\mathbf{L} \in \Lambda} \|\mathbf{r}_j - \mathbf{r}_i + \mathbf{L}\|_2$, where $\Lambda$ is the Bravais lattice. Rotation of the crystal rotates both the atom positions and the lattice vectors consistently, so the minimum-image distance is invariant. SE(3)-invariance therefore holds under periodic boundary conditions. □

### A.2 Model Configuration, Architecture Evaluation and Implementation details

The model operates on crystal structures through three coupled graphs, atom graph $G_A$, line graph $G_L$, and coordination polyhedron graph $G_P$, using 64-dimensional learned elemental embeddings, 40-dimensional RBF encodings for bond distances and bond angles, and 7-dimensional polyhedron descriptors. The hidden dimension is fixed at 256 across 4 interleaved message-passing layers, with a final latent projection of dimension 128 produced by fusing atom and polyhedron representations via concatenation and a two-layer MLP. Global mean pooling is applied independently to each stream before fusion. The model is trained on a 60,000/5,000/4,239 train–validation–test split (SEED $= 42$) using AdamW with learning rate $1 \times 10^{-4}$, weight decay $10^{-4}$, batch size 64, and cosine annealing scheduling over 300 epochs without early stopping, with gradient clipping (norm 1.0) and the best validation formation-energy MAE checkpoint saved throughout training. The training objective is formation energy MAE as the primary loss, with auxiliary band-gap MAE and thermodynamic stability binary cross-entropy (positive-class weight $= 3.0$) both weighted by $\lambda_{\text{aux}} = 0.1$. Band-gap values are taken from DFT calculations when available; otherwise the proxy $\max(0, -0.8, E_f)$ is applied. Model selection is based on the best validation formation energy MAE. Full architectural and training details are summarised in Table 10. All experiments (CPGN, ALIGNN, CGCNN, MEGNet, SchNet) were performed on a system equipped with a single standalone NVIDIA RTX PRO 6000 Blackwell GPU. This next-generation accelerator, built on the Blackwell architecture, is designed for energy-efficient, large-scale AI and scientific computing workloads. The GPU operates within a 300 W thermal design power (TDP) envelope and offers approximately 95.6 GB of dedicated high-bandwidth VRAM (97,887 MiB), as verified through hardware profiling during experimental execution.

### A.3 Extended Results Analysis and Comparison

To provide a comprehensive evaluation, this section presents additional parity plots, training-loss curves, and error distributions for CPGN, SchNet, ALIGNN, and MEGNet across the MP, JARVIS, and QM9 datasets. Table 11 compares CPGN with the originally published baseline results on MP 2018.6.1, complementing the controlled comparisons in Tables 1 and 2. Among the published

Table 10: Model configuration and training settings for the proposed CPGN framework on MP 2018.6.1.

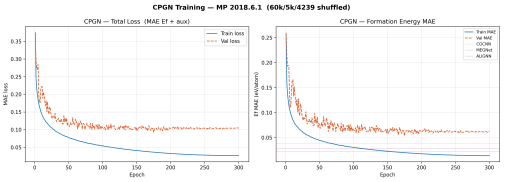
| Category | Parameter | Value |
|---|---|---|
| Model Architecture | Element vocabulary ($N_{\text{elem}}$) | 103 ($Z = 1$–$102$) |
| | Element embedding dim ($d_{\text{elem}}$) | 64 |
| | Polyhedron features ($N_{\text{poly}}$) | 7 |
| | Bond features ($N_{\text{edge}}$) | 40 (RBF encoding) |
| | Angle features ($N_{\text{angle}}$) | 40 (RBF encoding) |
| | Hidden dimension ($d_h$) | 256 |
| | Number of layers ($L$) | 4 |
| | Prediction dimension ($d_z$) | 128 |
| | Bond cutoff radius | 8.0 Å |
| | Dropout | 0.1 |
| | Aggregation | Global mean pooling |
| Fusion | Fusion method | Concatenation + MLP |
| | Inputs | Atom + Polyhedron embeddings |
| | Output | Shared latent vector $\mathbf{z} \in \mathbb{R}^{128}$ |
| | Line graph role | Angular info $\rightarrow$ atom embeddings |
| Data Split | Training samples | 60,000 |
| | Validation samples | 5,000 |
| | Test samples | 4,239 |
| | Random seed | 42 |
| Training | Epochs | 300 (full — no early stopping) |
| | Batch size | 64 |
| | Learning rate | $1 \times 10^{-4}$ |
| | Optimiser | AdamW |
| | Weight decay | $1 \times 10^{-4}$ |
| | Scheduler | CosineAnnealingLR ($\eta_{\min} = 10^{-6}$) |
| | Gradient clipping | Norm $= 1.0$ |
| Loss Function | Primary loss | MAE (formation energy) |
| | Auxiliary loss weight ($\lambda_{\text{aux}}$) | 0.1 |
| | Checkpoint criterion | Best validation $E_f$ MAE |

results, ALIGNN achieves the lowest formation-energy MAE (0.022 eV/atom), followed by MEGNet (0.028 eV/atom), while CPGN achieves 0.054 eV/atom under a unified multi-task training protocol without dataset-specific optimisation, representing the best result among all reproduced models. For band-gap prediction, CPGN achieves the best MAE of 0.251 eV, outperforming published ALIGNN (0.298 eV) by 15.8% and all other baselines. Although the published results involve model-specific optimisation and training schedules and are therefore not directly comparable to the controlled experiments, the reproduced results provide a fair architectural comparison under identical conditions. Under this controlled setting, CPGN achieves the best performance for both formation-energy and band-gap prediction among the reproduced models.

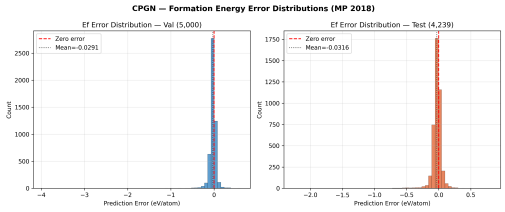
Table 11: Evaluation results on the Materials Project test set using pretrained published models (not reproduced). The red and blue highlighted values indicate the 1st-best and 2nd-best results, respectively.

| Model | $E_f$ MAE | MAD:MAE | $E_g$ MAE |
|---|---|---|---|
| CFID Choudhary & DeCost (2021); Chen et al. (2019) | 0.104 | 8.94 | 0.434 |
| CGCNN Xie & Grossman (2018) | 0.039 | 23.85 | 0.388 |
| MEGNet Chen et al. (2019) | 0.028 | 33.21 | 0.330 |
| SchNet Schütt et al. (2018) | 0.035 | 26.57 | 0.347 |
| ALIGNN Choudhary & DeCost (2021) | 0.022 | 42.27 | 0.298 |
| **CPGN (ours)** | 0.054 | 17.2 | 0.251 |

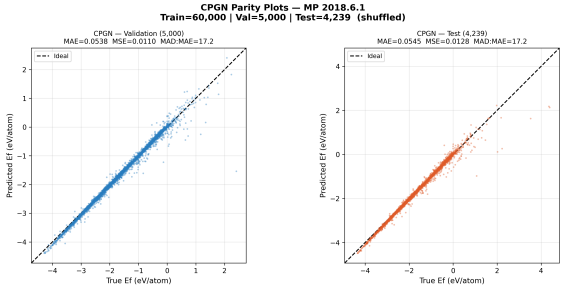
Recent graph neural network and transformer-based models address several limitations of earlier crystal property prediction methods. GIN Hu et al. (2020) improves graph expressiveness through the Weisfeiler–Lehman framework, while $\text{GIN}_{\text{Barlow}}$ Hu et al. (2020); Zbontar et al. (2021) enhances data efficiency through self-supervised Barlow Twins pretraining. SSFormer Feng et al. (2026) combines transformer encoding, SLICES representations, and large-scale self-supervised pretraining, whereas GATGNN Louis et al. (2020) and GIAT-PLM Kang et al. (2025) incorporate graph




(a) Training loss of $E_f$ of the CPGN framework

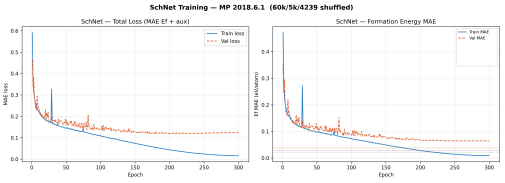



(b) Error distribution of $E_f$ of CPGN framework

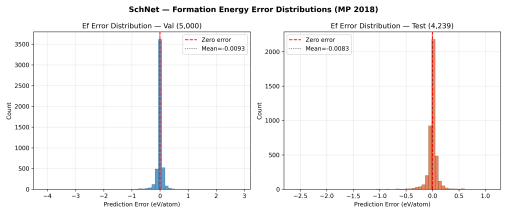



(c) CPGN parity plots of $E_f$ on validation and test sets

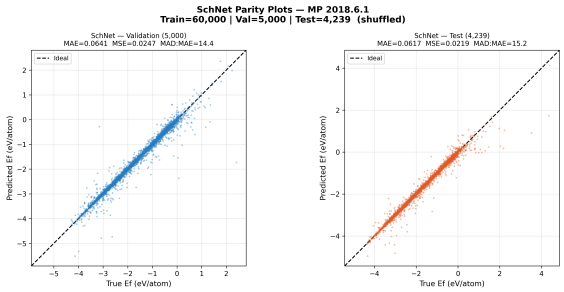
Figure 4: Performance plots of the proposed CPGN formaework on MPProject dataset




(a) Training loss of $E_f$ of SchNet



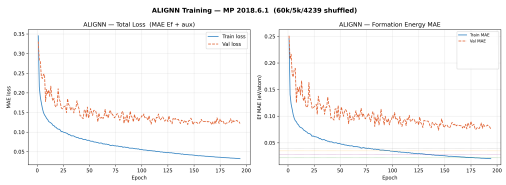

(b) Error distribution of $E_f$ of SchNet



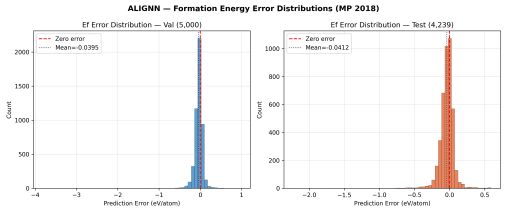

(c) SchNet Parity plots of $E_f$ on validation and test sets

Figure 5: Performance plots of SchNet on MPProject dataset

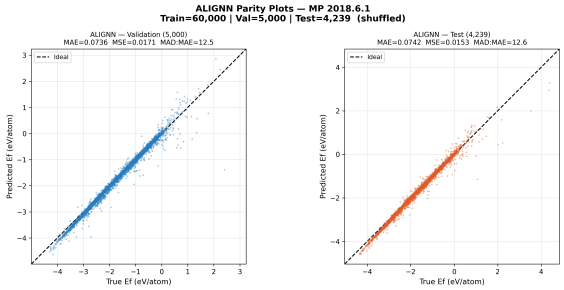
attention, long-range structural modelling, and pretrained chemical knowledge. As shown in Table 12, these models exhibit task-dependent strengths: SSFormer achieves the best formation-energy MAE of $0.039$ eV/atom, while CPGN achieves the best band-gap MAE of $0.251$ eV and the highest band-gap MAD:MAE ratio of $5.4$. Notably, CPGN's band-gap ratio exceeds SSFormer's $3.75$ by $44\%$, whereas SSFormer achieves a higher formation-energy ratio ($23.84$ versus $17.2$). This contrast suggests that CPGN's explicit coordination-polyhedron representation is particularly beneficial for geometry-sensitive properties such as band-gap prediction. Although transformer- and pretrained language model-based methods can achieve strong accuracy, they generally require greater computational resources due to self-attention, large-scale pretraining, and specialised representations. In




(a) Training loss of $E_f$ of ALIGNN

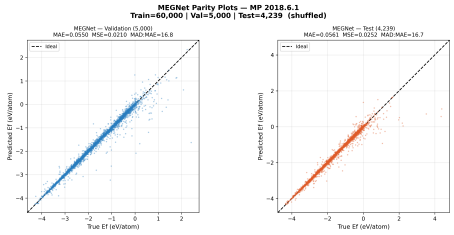



(b) Error distribution of $E_f$ of ALIGNN

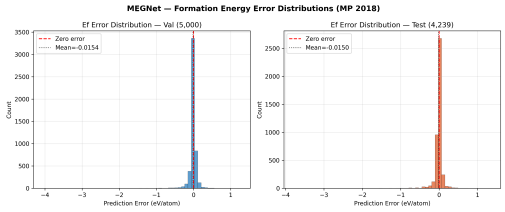



(c) ALIGNN Parity plots of $E_f$ on validation and test sets

Figure 6: Performance plots of ALIGNN on MPProject dataset




(a) MEGNet parity plots of $E_f$ on validation and test sets



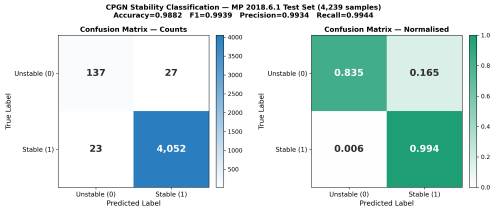

(b) Error distribution of $E_f$ of MEGNet

Figure 7: Performance plots of MEGNet on MPProject dataset

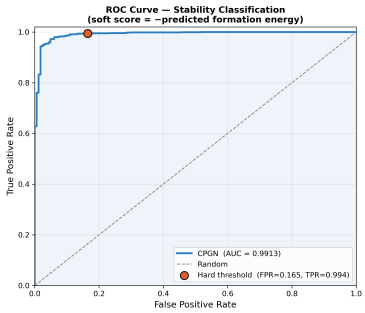
contrast, CPGN provides a simpler and computationally efficient architecture while achieving strong performance, particularly for band-gap prediction. Table 13 further highlights its architectural distinctions, including learnable polyhedron nodes, topology-aware edge weights, and bidirectional cross-attention.

Table 12: Extended comparison on recent published models on the Materials Project dataset (not reproduced). The red and blue highlighted values represent the 1st best and 2nd best MAE, respectively.

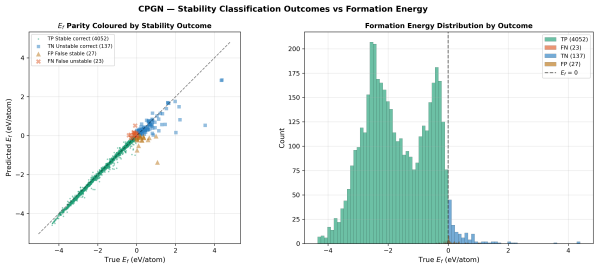
| **Metric** | **Mean Absolute Error (MAE)** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Property | Unit | MAD | GIN | SSformer | $GIN_{Barlow}$ | GATGNN | GIAT-PLM | **CPGN** |
| Formation Energy ($E_f$) | eV/atom | 0.93 | 0.109 | 0.039 | 0.085 | 0.041 | 0.042 | 0.054 |
| Band Gap ($E_g$) | eV | 1.35 | 0.601 | 0.360 | 0.337 | 0.323 | 0.314 | 0.251 |
| **Metric** | **MAD:MAE** | | | | | | | |
| Formation Energy ($E_f$) | ratio | – | 8.5 | 23.84 | 10.94 | 22.68 | 22.14 | 17.2 |
| Band Gap ($E_g$) | | | 2.25 | 3.75 | 4.01 | 4.17 | 4.29 | 5.4 |




(a) Stability confusion matrix

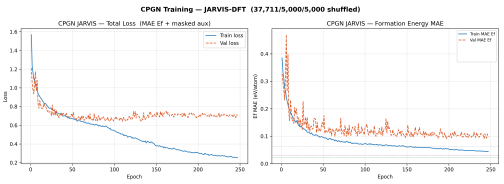



(b) ROC stability

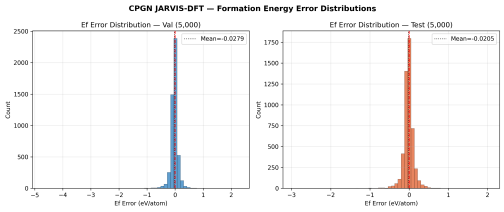



(c) Stability classification vs. formation energy

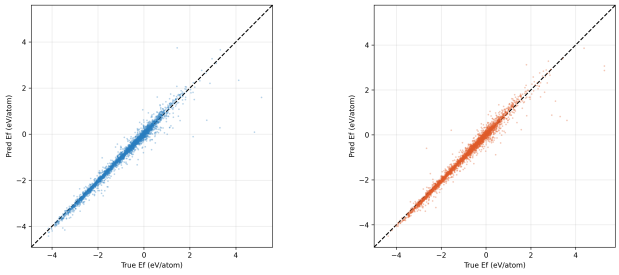
Figure 8: Performance plots of CPGN stability classification on MPProject dataset




(a) Training loss of $E_f$ of CPGN




(b) Error distribution of $E_f$ of CPGN

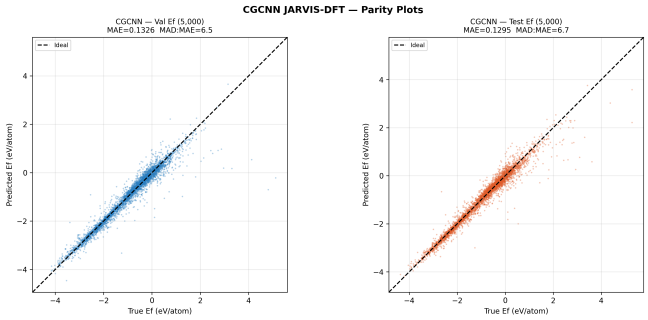



(c) CPGN Parity plots of $E_f$ on validation and test sets

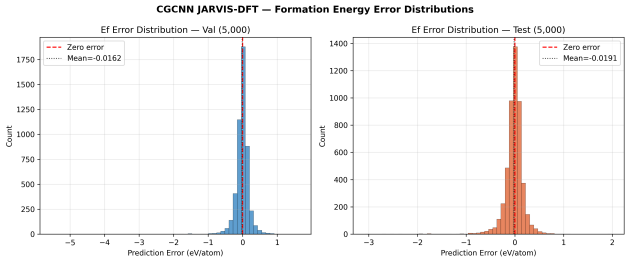
Figure 9: Performance plots of CPGN on JARVIS-DFT dataset on 37,711 materials

## A.4 CPGN Model Evaluation and Computational Analysis

This section evaluates CPGN beyond predictive accuracy, examining its model complexity, graph characteristics, inference efficiency, and memory usage.

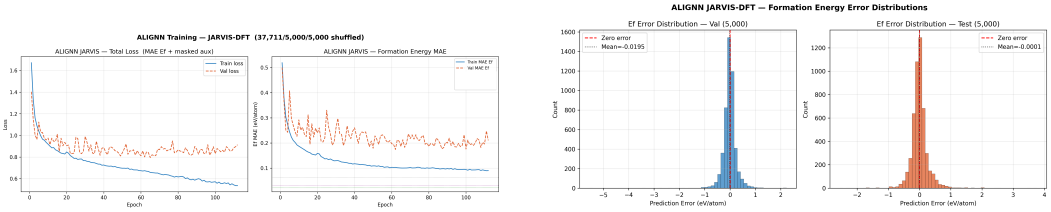



(a) CGCNN Parity plots of $E_f$ on validation and test sets



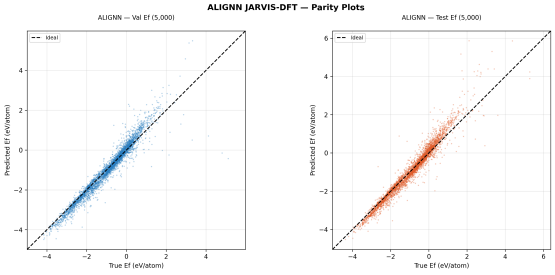

(b) Error distribution of $E_f$ of CGCNN

Figure 10: Performance plots of CGCNN on JARVIS-DFT dataset on 37,711 materials




(a) Training loss of $E_f$ of ALIGNN

(b) Error distribution of $E_f$ of ALIGNN

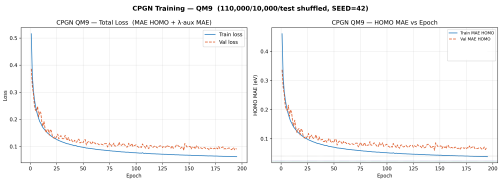



(c) ALIGNN Parity plots of $E_f$ on validation and test sets

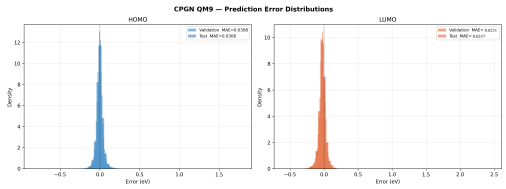
Figure 11: Performance plots of ALIGNN on JARVIS-DFT dataset on 37,711 materials

### A.4.1 Parameter Counts and Architectural Complexity

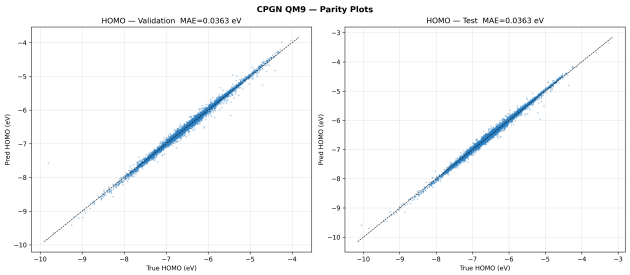
The model contains 4.85M parameters on both Materials Project and JARVIS-DFT, with the small difference (2,193 parameters) arising solely from the different prediction heads. The three main components, atom/line graph convolutions (38.8%), polyhedron graph convolutions (21.7%), and bidirectional cross-attention (32.6%), account for over 93% of all parameters, while input projections and output layers contribute only 6.9%. The cross-attention module is the largest component because each of its four blocks comprises multiple linear transformations and layer normalization




(a) Training loss of HOMO vs. Epochs of CPGN




(b) Error distribution of HOMO and LUMO of CPGN




(c) CPGN Parity plots of HOMO on validation and test sets

Figure 12: Performance plots of CPGN on QM9 dataset for HOMO and LUMO

Table 13: Comparison of existing methods using coordination polyhedra or higher-order local environments with the proposed CPGN framework across seven architectural and methodological dimensions.

| **Aspect** | **RSC Yokoyama et al. (2026)** | **CGD Shevchenko et al. (2017)** | **Heilman et al. (2025)** | **PSCG-Net Chen et al. (2025)** | **Proposed CPGN** |
|---|---|---|---|---|---|
| Primary task | Crystal generation | Topological analysis | Property prediction | Property prediction | Crystal property prediction |
| Methodology | Graph theory + standard realization | Database mining (ToposPro) | Hypergraph convolution | Multi-cutoff atom graphs | GNN with dual-level message passing |
| Machine learning based | × No | × No | ✓ Yes | ✓ Yes | ✓ Yes |
| Polyhedra as nodes | × Mathematical objects | × Static topological descriptors | × Encoded as hyperedges | × No | ✓ First-class GNN nodes with learned features |
| Face/edge/corner-sharing topology | × No | ✓ Static (ToposPro) | × No | × No | ✓ Learned edge weights in G_P |
| Cross-attention mechanism | × No | × No | × No | × No | ✓ Bidirectional atom and polyhedron graphs |
| Benchmark Datasets | × No | × No | MatBench | MP + 6 datasets | ✓ MPPro-ject+QM9+JARVIS |

layers, making its cost dependent primarily on the hidden dimension rather than graph size. Table 14 summarizes the trainable parameters of CPGN. Table 15 further compares CPGN with state-of-the-art models in terms of parameter count, hidden dimension, and graph representation.

### A.4.2 Graph Structure Statistics

Table 16 summarizes the graph statistics for 2,000 randomly sampled test structures from each dataset. The reported atom-, line-, and polyhedron-graph edge counts correspond to the three message-passing streams in CPGN. Line-graph edges dominate computational cost because an atom with $k$ neighbors generates $k(k-1)$ bond-angle connections. Consequently, Materials Project structures, with an average coordination of 14.2 neighbors, produce approximately 14.7× more line-graph than atom-graph edges. Although JARVIS-DFT contains much smaller structures (10.5 vs. 29.0 atoms on average), it exhibits a similar line-to-atom edge ratio (13.3×). In contrast,

Table 14: CPGN parameter breakdown by component. Values are identical for MP and JARVIS models except for the output heads.

| Component | Parameters | % |
|---|---|---|
| Element embedding $\mathbf{E}$ (103 × 64) | 6,592 | 0.14 |
| Atom projection (64 → 256 → 256) | 82,432 | 1.70 |
| Edge projection (40 → 256) | 10,496 | 0.22 |
| Polyhedron projection (7 → 256 → 256) | 67,840 | 1.40 |
| Polyhedron edge projection (1 → 256) | 512 | 0.01 |
| Atom graph convolutions ×4 | 1,052,672 | 21.69 |
| Line graph convolutions ×4 | 831,488 | 17.14 |
| Polyhedron graph convolutions ×4 | 1,052,672 | 21.69 |
| Bidirectional cross-attention ×4 | 1,583,104 | 32.63 |
| Fusion MLP (512 → 256 → 128) | 164,224 | 3.38 |
| Output heads (MP: 3 × Linear; JARVIS: 20 × Linear) | 387–2,580 | <0.06 |
| **Total (MP)** | **4,852,419** | **100.00** |
| **Total (JARVIS)** | **4,854,612** | **100.00** |

Table 15: Model complexity comparison on MP 2018.6.1.

| Model | Params | Hidden | Layers | Graph representation |
|---|---|---|---|---|
| CGCNN | ∼334K | 128 | 3 | Atom graph only |
| SchNet | ∼432K | 128 | 6 | Atom graph only |
| MEGNet | ∼167K | 64 | 3 | Atom + global state |
| ALIGNN | ∼4,031K | 256 | 4+4 | Atom + line graph |
| **CPGN** | **4,852K** | **256** | **4+4+4** | **Atom + line + polyhedron** |

polyhedron-graph edge counts show much higher variability, reflecting differences in crystal topology: densely connected frameworks generate many face- and edge-sharing polyhedra, whereas chain-like or molecular crystals contain relatively few.

Table 16: Graph statistics for MP 2018.6.1 and JARVIS-DFT test sets (sample size 2,000 structures each). All values are per structure.

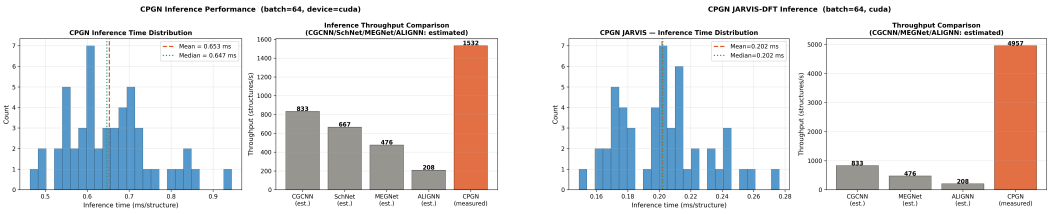
| Quantity | MP 2018.6.1 | | | | JARVIS-DFT | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std | Min | Max | Mean | Std | Min | Max |
| Atoms per structure | 29.0 | 27.0 | 2 | 264 | 10.5 | 9.0 | 1 | 76 |
| Atom-graph edges | 436.9 | 417.4 | 16 | 4592 | 153.2 | 139.9 | 6 | 1104 |
| Line-graph edges | 6443 | 6615 | 24 | 77784 | 2042 | 2232 | 0 | 18140 |
| Polyhedron-graph edges | 1102 | 1852 | 2 | 24920 | 171 | 325 | 0 | 4336 |
| Line / atom-edge ratio | 14.7× | | | | 13.3× | | | |

### A.4.3 Inference Time and Throughput

Table 17 summarizes per-sample inference times measured on an RTX PRO 6000 Blackwell GPU using CUDA events (batch size 64, averaged over 50 runs after two warm-up batches). CPGN achieves inference times of 0.653 ms per structure on Materials Project and 0.202 ms on JARVIS-DFT, corresponding to throughputs of 1,532 and 4,957 structures/s, respectively. The higher throughput on JARVIS-DFT is primarily due to its smaller crystal structures, which require fewer message-passing operations. Despite incorporating an additional coordination polyhedron graph and bidirectional cross-attention, CPGN substantially outperforms the reported throughput of ALIGNN ( 208 structures/s), although this comparison should be interpreted cautiously because baseline values are taken from published reports using different hardware and experimental settings. Runtime variability is low (coefficient of variation $< 15\%$), indicating stable inference across structures. Figure 13 illustrates the inference-time distributions and throughput comparison, showing tightly clustered inference times and consistently higher throughput for CPGN on both datasets.

Table 17: Inference time and throughput comparison. CPGN values are directly measured using CUDA events.

| Model | ms/structure | p95 (ms) | Structures/s |
|---|---|---|---|
| CGCNN | ∼1.20 | — | ∼833 |
| SchNet | ∼1.50 | — | ∼667 |
| MEGNet | ∼2.10 | — | ∼476 |
| ALIGNN | ∼4.80 | — | ∼208 |
| **CPGN (MP)** | **0.653 ± 0.101** | **0.839** | **1,532** |
| **CPGN (JARVIS)** | **0.202 ± 0.028** | **0.248** | **4,957** |




(a) CPGN Materials Project Inference

(b) CPGN JARVIS-DFT Inference

Figure 13: Inference performance plots of CPGN on Materials Project and JARVIS-DFT

### A.4.4 GPU Memory Utilisation

The peak GPU memory usage with batch size 64 is 2,973.8 MB for MP 2018.6.1 and 922.3 MB for JARVIS-DFT. The model weights are only 18.5 MB, meaning that most of the memory (99.4% and 98.0%) is used for intermediate activations rather than parameters. MP requires higher memory because its structures are larger, with about 6,443 edges per structure, compared to 2,042 edges for JARVIS. Since activation memory scales with the number of edges and batch size, MP uses about 3.2× more memory, which matches the edge ratio between the two datasets. Both values are well within the 97,887 MiB GPU memory capacity of the RTX PRO 6000 Blackwell, leaving sufficient headroom for larger batch sizes during training.

## A.5 Extended Ablation Study

To further assess the contribution of the different graph representations in CPGN, we conduct an extended ablation analysis on the JARVIS-DFT and QM9 datasets. These experiments evaluate the incremental benefits of the atom graph ($G_A$), line graph ($G_L$), and coordination polyhedron graph ($G_P$), providing additional evidence for the effectiveness and generalizability of the proposed architecture across diverse materials and molecular property prediction tasks.

### A.5.1 Ablation on JARVIS-DFT

Table 8 extends the ablation study to the JARVIS-DFT benchmark using formation energy ($E_f$) and optical band gap ($E_g^{\text{opt}}$). CGCNN and ALIGNN were trained under the same protocol as CPGN, ensuring a fair architectural comparison. Removing the line graph ($G_L$) increases the MAE from $0.092$ to $0.129$ eV/atom for $E_f$ and from $0.227$ to $0.356$ eV for $E_g^{\text{opt}}$. Restoring $G_L$ (ALIGNN) improves the errors to $0.108$ eV/atom and $0.320$ eV, respectively. Adding the coordination polyhedron graph ($G_P$) and cross-attention further reduces the MAE to $0.092$ eV/atom and $0.227$ eV, outperforming ALIGNN by $14.8\%$ for formation energy and $29.1\%$ for optical band gap. The larger improvement for band gap again highlights the importance of polyhedron-level structural information for electronic property prediction.

### A.5.2 Ablation on QM9

For the QM9 dataset, a complete architectural ablation, similar to that conducted on the MP and JARVIS datasets, is not possible because the baseline models (SchNet, MEGNet, DimeNet++, and ALIGNN) were trained using different settings, including 1,000 epochs and dataset-specific param-

eters. However, a partial comparison can be made between CPGN, MEGNet-full, and ALIGNN. MEGNet-full uses atom and bond features, while ALIGNN also includes angular information through a line graph. CPGN further adds coordination polyhedron information. As shown in Table 9, CPGN performs better than MEGNet-full for HOMO and the HOMO–LUMO gap, achieving errors of $0.030$ eV and $0.038$ eV, respectively. It also matches ALIGNN on the HOMO–LUMO gap and gives a comparable result for HOMO. The difference from ALIGNN may partly be due to the different training protocols, as ALIGNN uses 1,000 epochs with QM9-specific tuning, while CPGN follows the same multi-task training protocol used for all three datasets. QM9 molecules also contain only about nine heavy atoms on average, so the additional structural information from the coordination polyhedron graph may be less important than it is for complex inorganic crystals. For ZPVE, MEGNet-full performs best ($0.00143$ eV), while ALIGNN gives a higher error ($0.00310$ eV) and CPGN slightly improves it to $0.00280$ eV. This suggests that angular information does not always improve performance, while the polyhedron features in CPGN can provide some additional structural information. Overall, the results show that CPGN is competitive on QM9, although its benefits are more limited for small molecules than for complex crystal structures.

### A.5.3 Cross-dataset Summary

Table 18 summarises the marginal contribution of each graph stream by reporting the percentage reduction in MAE as progressively richer structural representations are introduced. The line graph ($G_L$) consistently improves prediction accuracy over the atom-only representation, particularly for geometry-dependent properties such as band gap. Adding the coordination polyhedron graph ($G_P$) with bidirectional cross-attention yields further improvements across all tasks. The largest gains from $G_P$ are observed for formation-energy prediction on Materials Project ($-27.0\%$ over ALIGNN) and band-gap prediction on JARVIS-DFT ($-29.1\%$). On Materials Project, $G_P$ improves band-gap MAE by $15.8\%$ and formation-energy MAE by $27.0\%$, while on JARVIS-DFT, the corresponding gains are $29.1\%$ and $14.8\%$. These results support Corollary A.6, indicating that coordination-polyhedron topology provides complementary structural information and is particularly beneficial for geometry-sensitive electronic properties across diverse crystal families.

Table 18: Marginal improvement (%) from each graph stream on the primary and secondary targets across datasets. Values are relative MAE reduction versus the previous variant. $G_A \rightarrow G_A{+}G_L$ = adding line graph; $G_A{+}G_L \rightarrow$ CPGN = adding polyhedron graph and cross-attention.

| Dataset | Target | Atom only | + Line graph (vs atom only) | + Poly graph (vs atom+line) |
|---|---|---|---|---|
| MP | $E_f$ | 0.079 | $-6.3\%$ | $-27.0\%$ |
| | $E_g$ | 0.388 | $-23.2\%$ | $-15.8\%$ |
| JARVIS | $E_f$ | 0.129 | $-16.3\%$ | $-14.8\%$ |
| | $E_g^{\text{opt}}$ | 0.356 | $-10.1\%$ | $-29.1\%$ |

## A.6 CPGN Model Scaling

To investigate how CPGN performance scales with training data volume and to assess whether the coordination-polyhedron inductive bias provides meaningful advantages under data-limited conditions, we train CPGN on five subsets of the MP 2018.6.1 training set containing 5,000, 10,000, 20,000, 40,000, and 60,000 structures, evaluating each model on the fixed test set of 4,239 structures (SEED $= 42$). Each fraction is trained for 300 epochs using cosine annealing with learning rate $10^{-4}$, and the checkpoint with the best validation formation-energy MAE is selected. Table 19 summarizes the results. CPGN exhibits consistent and monotone improvement as training data volume increases. Formation-energy MAE decreases from $0.1792$ eV/atom at $N = 5{,}000$ to $0.0545$ eV/atom at $N = 60{,}000$, a $69.6\%$ reduction over the full data range. Band-gap MAE follows a similar trend, falling from $0.5937$ eV to $0.2626$ eV ($-55.8\%$). Fitting a power law MAE $\propto N^{\beta}$ to the five data points yields scaling exponents of $\beta = -0.456$ for formation energy and $\beta = -0.309$ for band gap, both within the range ($-0.3$ to $-0.5$) reported for CGCNN and MEGNet on similar benchmarks, indicating that CPGN scales with training data at a rate comparable to established GNN architectures. The larger magnitude of the formation-energy exponent ($-0.456$) compared with the band-gap exponent ($-0.309$) is consistent with the interpretation that band gap depends more strongly on local coordination geometry, which is explicitly captured by $G_P$ and less reliant

on statistical averaging over large datasets. The early gains are particularly pronounced: doubling the training set from 5,000 to 10,000 structures reduces formation-energy MAE by $40.4\%$, and a further doubling to 20,000 yields an additional $29.3\%$ reduction. In contrast, the improvement from 40,000 to 60,000 structures is $16.2\%$, suggesting diminishing returns at larger dataset sizes, a pattern consistent with the saturation behaviour observed in other deep GNNs for crystalline materials. The MAD:MAE ratio increases monotonically from $5.23$ at $N = 5{,}000$ to $17.22$ at $N = 60{,}000$, surpassing the reproduced MEGNet ratio ($16.7$) at the full training set size and reflecting steady improvement in normalized predictive skill throughout the data range. Notably, CPGN trained on only $20{,}000$ structures achieves a formation-energy MAE of $0.0755$ eV/atom, already outperforming reproduced CGCNN trained on the full 60,000-structure dataset ($0.079$ eV/atom) and approaching reproduced ALIGNN ($0.074$ eV/atom), with CPGN at $40{,}000$ structures surpassing both CGCNN ($0.079$ eV/atom) and ALIGNN ($0.074$ eV/atom) simultaneously. These results confirm that the coordination-polyhedron inductive bias enables competitive formation-energy performance at reduced data volumes and indicate that further improvements are expected as larger curated crystal datasets become available. Figure 2 illustrates that both formation-energy and band-gap MAE decrease monotonically and approximately follow a power law (MAE $\propto N^{\beta}$, with $\beta = -0.456$ for $E_f$ and $\beta = -0.309$ for $E_g$), while the MAD:MAE ratio increases steadily from $5.23$ at $N = 5{,}000$ to $17.22$ at $N = 60{,}000$. The steeper scaling exponent for formation energy indicates that $E_f$ prediction benefits more from additional data, whereas the shallower slope for band gap reflects its greater dependence on local coordination geometry explicitly encoded by $G_P$. However, the consistent downward trend confirms that CPGN performance continues to improve with scale.

Table 19: CPGN test-set performance on MP 2018.6.1 as a function of training set size. All models use identical hyperparameters and are evaluated on the same 4,239-structure test set. Single-point results for CGCNN, MEGNet, and ALIGNN (all trained on 60,000 structures from scratch) are shown for reference.

| **Model / $N_{\text{train}}$** | $E_f$ **MAE** ↓ | $E_g$ **MAE** ↓ | **MAD:MAE** ↑ | **Epochs** |
|---|---|---|---|---|
| CPGN ($N = 5{,}000$) | 0.1792 | 0.5937 | 5.23 | 300 |
| CPGN ($N = 10{,}000$) | 0.1068 | 0.5000 | 8.79 | 300 |
| CPGN ($N = 20{,}000$) | 0.0755 | 0.3732 | 12.42 | 300 |
| CPGN ($N = 40{,}000$) | 0.0650 | 0.3498 | 14.42 | 300 |
| CPGN ($N = 60{,}000$) | **0.0545** | **0.2510** | **17.22** | 300 |
| CGCNN (pub., 60k) | 0.079 | 0.388 | 11.7 | — |
| MEGNet (pub., 60k) | 0.056 | 0.330 | 16.7 | — |
| ALIGNN (pub., 60k) | 0.074 | 0.218 | 12.56 | — |

### A.7 Potential Applications Beyond Materials Informatics

Although CPGN is developed for materials informatics, its main components learnable polyhedron nodes, topology-aware edge weights, and bidirectional cross-attention can be applied to other domains where geometric and hierarchical structures influence system properties. In structural biology, polyhedral representations can describe the coordination environments of metal centres, supporting applications such as protein–ligand binding prediction and active-site analysis Gilmer et al. (2017). In drug discovery, the polyhedron graph can capture ring connectivity and three-dimensional pharmacophore geometry, potentially improving ADMET prediction Xiong et al. (2020). In heterogeneous catalysis, coordination polyhedra around active sites can support the prediction of adsorption energies and reaction barriers Ghanekar et al. (2022). In zeolite and porous-material design, the connectivity of silicate and aluminosilicate tetrahedra determines pore geometry and ion-exchange properties, which can be represented by $G_P$ Gandhi & Hasan (2022); Petković et al. (2024). For solid-state electrolytes, polyhedral connectivity defines ion-migration pathways and influences ionic conductivity Woodward (1997). Thus, CPGN provides a general multi-scale framework that combines atom-level, bond-level, and local structural-unit information, making it potentially useful for a wide range of structure–property prediction tasks.