AN ORBITAL-BASED GEOMETRIC DEEP LEARNING FRAMEWORK FOR PERIODIC MATERIALS

Minhyuk Kang[1,2*], **Beom Seok Kang**[1*], **Sergei Kliavinek**[1], **Maurice D. Hanisch**[1], **William A. Goddard III**[1], **Anima Anandkumar**[1]

[1]California Institute of Technology [2]Imperial College London
minhyuk.kang21@imperial.ac.uk
{bkang,skliavin,mhanisch,wag,anima}@caltech.edu
*Equal contribution

ABSTRACT

Orbital-based deep learning has achieved notable success in molecular systems by integrating quantum mechanical information into geometric deep learning. In this study, we introduce OrbNet-Crystal, which extends the OrbNet-Equi framework from molecular systems to periodic crystal graphs. Our approach derives orbital features from lower-level semi-empirical quantum mechanical calculations, converts reciprocal-space orbital features into real-space orbital features via inverse Fourier transforms, and embeds them into SE(3)-equivariant graph neural networks for periodic materials. We evaluate OrbNet-Crystal on the Computational 2D Materials Database for band gap prediction at multiple electronic-structure levels of theory. Using semi-empirical orbital features, OrbNet-Crystal achieves strong accuracy across all targets, both for direct prediction and delta-learning across multiple levels of theory. Furthermore, we show that transfer learning substantially improves performance for data-scarce, higher-level theory targets. Overall, this work establishes the first orbital-based deep learning paradigm for crystalline systems and demonstrates the potential of orbital features for transferable, multi-fidelity learning in solid-state materials discovery.

1 INTRODUCTION

Recent advances in machine learning for materials science have demonstrated that data-driven models can predict quantum chemical properties at reasonable accuracy with dramatically reduced computational cost (Ramprasad et al., 2017). Most notably, graph neural networks (GNNs) have shown their strengths in materials design by directly learning atomic structures (Xie & Grossman, 2018; Kaba & Ravanbakhsh, 2022). Despite this progress, predicting electronic properties such as band gaps remains challenging, as models relying solely on geometric information often struggle to fully capture the underlying electronic structure (Cheng et al., 2021).

Orbital-based deep learning methods, such as OrbNet and its variants, have shown notable success in predicting molecular quantum chemical properties (Qiao et al., 2020; Kang et al., 2025). By incorporating quantum mechanics information, orbital learning achieves strong generalization, high accuracy, and data efficiency, while also exhibiting robust extrapolation capabilities. Despite the promise shown in machine learning in chemistry, to the best of our knowledge, OrbNet-style orbital-informed learning has not yet been broadly explored for periodic crystal graphs, especially using semi-empirical tight-binding features embedded in equivariant GNNs.

Here, we introduce OrbNet-Crystal, which extends OrbNet-Equi to periodic boundary conditions (PBCs). We compute quantum mechanical matrices (QMMs) via a semi-empirical quantum mechanical method and convert reciprocal-space matrices to localized real-space blocks via inverse Fourier transforms (Grimme et al., 2017). We embed the real-space QMMs into SE(3)-equivariant periodic crystal graphs and process with an equivariant GNN (Qiao et al., 2022).

We evaluate OrbNet-Crystal on the Computational 2D Materials Database (C2DB) across varying fidelities, including PBE, HSE06, and G_0W_0. Across all targets, OrbNet-Crystal achieves state-of-

the-art accuracy for both direct prediction and delta-learning. Notably, for G_0W_0 band gaps, transfer learning across fidelities further improves accuracy, demonstrating the benefit of leveraging lower-level data for high-level predictions. These results indicate that orbital-informed representations improve data efficiency and offer a scalable alternative to DFT-derived electronic descriptors (Tawfik & Russo, 2022; Haotian Liu, 2025), making OrbNet-Crystal well suited for data-scarce, high-fidelity materials prediction.

Beyond extending orbital learning to periodic systems, this framework offers a principled setting for multi-fidelity learning of electronic structure. By learning corrections between different levels of theory directly from orbital features, the model enables transferable representations of physically meaningful electronic corrections, providing a pathway toward accurate, data-efficient prediction of high-fidelity properties in crystalline materials (Ramakrishnan et al., 2015).

2 METHODS

Reciprocal-space to real-space In periodic systems, Bloch's theorem implies that electronic operators such as the Hamiltonian and overlap are naturally represented in reciprocal space as $\mathbf{O}(\mathbf{k})$, indexed by crystal momentum $\mathbf{k}$. While this formulation allows for solving the self-consistent field (SCF) method, the $\mathbf{k}$-space matrices must be resolved in real-space for geometric deep learning.

Within the linear combination of atomic orbitals (LCAO) framework, real-space interaction blocks between atomic orbitals can be recovered from a discrete $\mathbf{k}$-point mesh via an inverse Fourier transform (Slater & Koster, 1954),

$$\mathbf{O}_{AB}^{\mu\nu}(\mathbf{R}) = \frac{1}{N_k} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} \mathbf{O}_{AB}^{\mu\nu}(\mathbf{k}), \tag{1}$$

where A, B index atoms in the reference cell, μ, ν index atomic orbitals, and $\mathbf{R}$ labels lattice offsets. For localized atomic-orbital bases, these real-space blocks decay rapidly with $|\mathbf{R}|$, yielding a sparse set of short-range interactions well suited for message passing in periodic crystal graphs (Karandashev & von Lilienfeld, 2022). Full theoretical background is provided in Appendix A.

SEQM Calculation Configuration In our work, we apply this real-reciprocal space relationship to generate orbital features for crystalline materials. First, we perform a periodic SCF calculation using the GFN1-xTB method implemented in the DFTB+ software to obtain the QMMs for each material (Grimme et al., 2017; Hourahine et al., 2020; 2025). After SCF convergence, the $\mathbf{k}$-space Hamiltonian matrix is converted to a real-space ($\mathbf{r}$-space) Hamiltonian matrix via an inverse Fourier transform as in equation 1. This pipeline yields $\mathbf{r}$-space QMM blocks $\mathbf{O}_{AB}(\mathbf{R})$ for atom pairs and relevant lattice offsets, which are embedded into an E(3)-equivariant GNN forming the basis of the crystal graph described below.

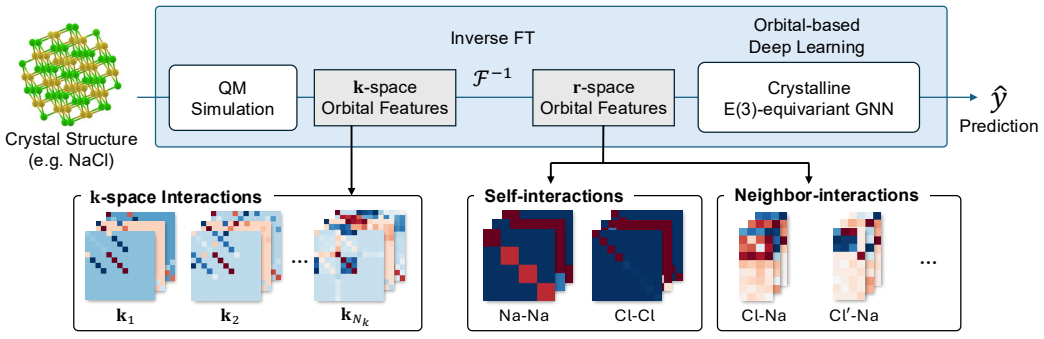

Figure 1: Overview of the Framework.

The OrbNet Framework The OrbNet framework is a hybrid of a semi-empirical method and geometric deep learning, enabling the incorporation of QM information into molecular representations. This physics-grounded approach yields strong generalization and accuracy, excelling in

challenging systems such as open-shell or charged molecules. It has demonstrated robust extrapolation and speedups of 1,000–10,000x over DFT while requiring roughly 10 times less data than other state-of-the-art models to achieve chemical accuracy (Qiao et al., 2020; Kang et al., 2025).

However, existing OrbNet implementations have been limited to finite systems (molecules), limiting their direct application to periodic materials. We now address this gap by extending the OrbNet-Equi framework to periodic systems by adapting it to operate on periodic crystal graphs (Qiao et al., 2022). To our knowledge, this is among the first works to adapt the OrbNet-Equi paradigm to periodic crystal graphs using localized real-space orbital blocks.

Orbital Features In this work, we utilize a set of QMMs, $\mathbf{T} = (\mathbf{H}^0, \mathbf{S}, \mathbf{P})$, where $\mathbf{H}^0$ is the initial Hamiltonian matrix, $\mathbf{S}$ is the overlap matrix, and $\mathbf{P}$ is the density matrix, following the OrbNet framework (Qiao et al., 2022). While QMMs can be defined in both $\mathbf{k}$-space and $\mathbf{r}$-space (Slater & Koster, 1954), in this study, we use the $\mathbf{r}$-space representation to construct localized and QM-informed messages.

Unlike finite molecular systems, periodic materials contain infinitely many atomic images due to lattice repetition, making neighborhood definition nontrivial (Xie & Grossman, 2018). To construct a finite yet physically meaningful crystal graph, we select neighbors based on orbital interaction strength rather than a purely geometric distance cutoff, prioritizing electronically relevant bonded and near-bonded couplings. For each atom, candidate neighbors across lattice translations are ranked according to the magnitude of their associated $\mathbf{r}$-space quantum mechanical matrix blocks (from $\mathbf{H}^0$, $\mathbf{S}$, and $\mathbf{P}$), and only the top-k strongest interactions are retained for message passing.

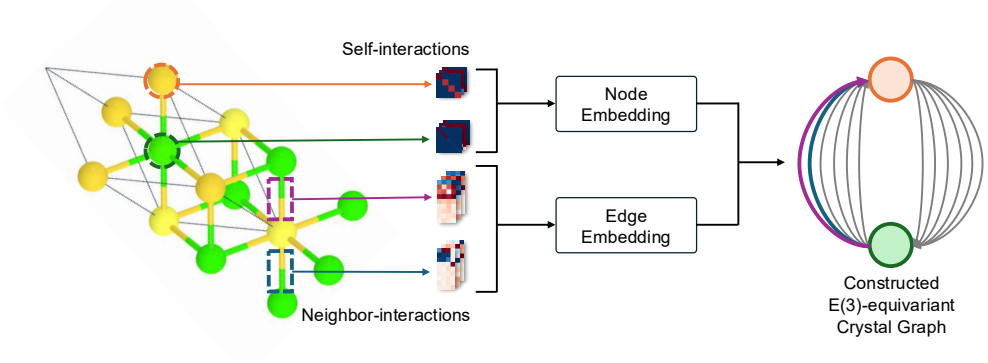

Figure 2: Crystal Graph Embedding of Orbital Features.

Embedding Orbital Features We embed the orbital features following the OrbNet-Equi architecture (Qiao et al., 2022). The self-interaction block $\mathbf{O}_{AA}$ is used to initialize node features, while QMMs describing interactions between different atoms are used to construct messages and attention weights during message passing. In crystal graphs, a node may connect to another node through multiple edges corresponding to different lattice offsets (Xie & Grossman, 2018). Within our framework, these distinct interactions are represented explicitly by separate QMM blocks, as illustrated in Figure 2. Both node information (self-interactions) and edge information (neighbor interactions) are embedded using SE(3)-equivariant operations (Kaba & Ravanbakhsh, 2022), following the approach of Qiao et al. (2022). Depending on the crystal structure and the neighboring strategy, atoms may connect to neighbors in adjacent unit cells and, in some cases, to their own periodic images.

Dataset and learning objectives We train OrbNet-Crystal on the Computational 2D Materials Database (C2DB) (Haastrup et al., 2018; Gjerding et al., 2021). Starting from 3,363 candidate semiconducting materials with available HSE06 label, periodic GFN1-xTB calculations and orbital feature generation succeeded for 3,191 materials (94.9%), which we retain for all experiments. Each selected material also has a corresponding PBE label. A smaller subset of 328 materials additionally has nonzero G_0W_0 band gaps. For the PBE and HSE06 targets, we use an 8:1:1 train/validation/test split. For the data-scarce G_0W_0 target, we employ five-fold cross-validation. To prevent compositional leakage, materials are grouped by stoichiometry patterns (e.g., A, AB_2, ABC), constructed by mapping chemical formulas to element-agnostic templates, and groups are

split consistently across folds (Gjerding et al., 2021). Models are trained to predict either absolute band gaps (E_g^{PBE}, E_g^{HSE06}, $E_g^{G_0W_0}$) or band gap corrections between levels of theory, such as $\Delta E_g^{\mathrm{HSE06-PBE}} = E_g^{\mathrm{HSE06}} - E_g^{\mathrm{PBE}}$, which is the delta learning target for HSE06 with the PBE baseline Ramakrishnan et al. (2015). All models are trained using supervised regression with mean absolute error (MAE) loss. For G_0W_0, we further employ transfer learning by initializing from a model trained on PBE→HSE06 corrections and fine-tuning on HSE06→G_0W_0 corrections.

3 RESULTS

Table 1: Performance of OrbNet-Crystal on band gaps for C2DB using GFN1-xTB orbital features ($\mathbf{H}^0, \mathbf{S}, \mathbf{P}$).

Target	Strategy	N	MAE (eV)
PBE	Direct prediction	3191	0.26
HSE06	Direct prediction	3191	0.31
HSE06	Δ(HSE06−PBE)	3191	0.12
G_0W_0	Direct prediction (5-fold CV)	328	0.46 ± 0.07
G_0W_0	Δ(G_0W_0−PBE) (5-fold CV)	328	0.23 ± 0.04
G_0W_0	Δ(G_0W_0−HSE06) (5-fold CV)	328	0.20 ± 0.04
G_0W_0	TL: Δ(HSE06−PBE) → Δ(G_0W_0 − HSE06) (5-fold CV)	328	0.17 ± 0.04

We evaluate the performance of OrbNet-Crystal on the C2DB (Haastrup et al., 2018; Gjerding et al., 2021) across three target levels of theory (PBE, HSE06, and G_0W_0). Table 1 summarizes results using post-SCF GFN1-xTB orbital features, which consistently yield the strongest performance. Pre-SCF variants that use only the initial Hamiltonian and overlap matrices ($\mathbf{H}^0, \mathbf{S}$) already achieve competitive accuracy (Appendix B), indicating that approximate orbital information provides a useful physical prior. Incorporating the converged density matrix further improves performance, suggesting that the GFN1-xTB-level electronic structure can provide meaningful physical information that the model can exploit.

Comparison to Prior Work On semiconducting materials in C2DB, prior PBE band gap models typically relied on elemental or compositional descriptors and report MAEs in the range of 0.33–0.35 eV (Wang et al., 2025; Dau et al., 2023), whereas OrbNet-Crystal achieves an MAE of 0.26 eV using GFN1-xTB orbital features. For higher-fidelity targets, most existing approaches incorporate PBE-level DFT information: reported HSE06 errors span 0.23–0.37 eV for predicting Δ(HSE06−PBE), often requiring PBE electron densities or DOS-level descriptors (Tawfik & Russo, 2022; Haotian Liu, 2025). In contrast, OrbNet-Crystal predicts HSE06 band gaps directly from semi-empirical orbital features with an MAE of 0.31 eV, and further reduces the error to 0.12 eV via delta-learning using only a scalar PBE baseline.

In the data-scarce G_0W_0 regime, where prior methods typically depend on extended DFT features such as full band structures or wavefunctions to achieve MAEs of 0.15–0.28 eV (Knøsgaard & Thygesen, 2022; Tawfik & Russo, 2022; Wang et al., 2025; Javed & Ali, 2025), our orbital-based transfer-learning approach attains a mean MAE of 0.17 ± 0.04 eV, with the best test set achieving 0.12 eV, demonstrating transferable correction learning without expensive DFT inputs.

4 DISCUSSION

Effectiveness of Orbital Features The strong performance of OrbNet-Crystal indicates that GFN1-xTB electronic structure provides physically meaningful information that the model can learn from. Explicit multi-edge representations are essential in periodic systems, as the same atom index pair can couple through distinct lattice offsets. Collapsing these interactions would discard physically meaningful directional information (Chen et al., 2019). The success in delta-learning and transfer learning across multiple levels of theory further suggests that orbital-based representations

enable the model to efficiently learn the systematic differences between electronic structure methods, which is particularly valuable for data-scarce domains like G_0W_0 data.

However, several previous studies have also demonstrated the limitations of GFN1-xTB in solid systems (Cui et al., 2024; Buccheri et al., 2025). Hence, the performance may further improve if the semi-empirical method parameters are extensively optimized for solid systems (Hourahine et al., 2025), or if orbital features from a higher level of theory, such as the cheap GGA functional (PBE), are used (Tawfik & Russo, 2022). These remain as promising future works.

Conclusion OrbNet-Crystal extends orbital-based geometric deep learning paradigm to periodic materials by embedding $\mathbf{r}$-space orbital blocks into SE(3)-equivariant multi-edge crystal graphs (Kaba & Ravanbakhsh, 2022). With the GFN1-xTB orbital features, it achieves accurate band-gap learning across multiple levels of theory and demonstrates the effectiveness of delta-learning strategies for accessing more accurate electronic structure prediction.

REFERENCES

Alexander Buccheri, Rui Li, J Emiliano Deustua, S Mohamad Moosavi, Peter J Bygrave, and Frederick R Manby. Periodic gfn1-xtb tight binding: A generalized ewald partitioning scheme for the klopman–ohno function. *Journal of Chemical Theory and Computation*, 21(4):1615–1625, 2025.

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.

Jiucheng Cheng, Chunkai Zhang, and Lifeng Dong. A geometric-information-enhanced crystal graph network for predicting properties of materials. *Communications Materials*, 2(1):92, 2021.

Mengnan Cui, Karsten Reuter, and Johannes T Margraf. Obtaining robust density functional tight-binding parameters for solids across the periodic table. *Journal of Chemical Theory and Computation*, 20(12):5276–5290, 2024.

Minh Tuan Dau, Mohamed Al Khalfioui, Adrien Michon, Antoine Reserbat-Plantey, Stéphane Vézian, and Philippe Boucaud. Descriptor engineering in machine learning regression of electronic structure properties for 2d materials. *Scientific Reports*, 13:5426, 2023. doi: 10.1038/s41598-023-31928-7.

Morten N. Gjerding, Alireza Taghizadeh, Anders Rasmussen, Shabbir Ali, Federico Bertoldo, Thorsten Deilmann, Uriel P. Holguin, N. R. Knøsgaard, Matthias Kruse, Anders H. Larsen, Simone Manti, Thomas G. Pedersen, Troels Skovhus, Mikkel K. Svendsen, Jens Jørgen Mortensen, Thomas Olsen, and Kristian S. Thygesen. Recent progress of the computational 2d materials database (c2db). *2D Materials*, 8(4):044002, 2021.

Stefan Grimme, Christoph Bannwarth, and Philip Shushkov. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-block elements (z= 1–86). *Journal of chemical theory and computation*, 13(5):1989–2009, 2017.

Sten Haastrup, Mikkel Strange, Mohnish Pandey, Thorsten Deilmann, Per S. Schmidt, Nicki F. Hinsche, Morten N. Gjerding, Daniele Torelli, Peter M. Larsen, Anders C. Riis-Jensen, Jakob Gath, Karsten W. Jacobsen, Jens Jørgen Mortensen, Thomas Olsen, and Kristian S. Thygesen. The computational 2d materials database: High-throughput modeling and discovery of atomically thin crystals. *2D Materials*, 5(4):042002, 2018.

Yunning Huang Yijia Luo Yongping Zheng Haotian Liu, Mingjun Weng. Self supervised ensemble learning models for 2d semiconductors bandgap prediction. *Computational Materials Science*, 255, 2025.

B Hourahine, M Berdakin, JA Bich, FP Bonafé, C Camacho, Q Cui, MY Deshaye, G Díaz Mirón, S Ehlert, M Elstner, et al. Recent developments in dftb+, a software package for efficient atomistic quantum mechanical simulations. *The Journal of Physical Chemistry A*, 2025.

Ben Hourahine, Bálint Aradi, Volker Blum, Frank Bonafe, Alex Buccheri, Cristopher Camacho, Caterina Cevallos, MY Deshaye, T Dumitrică, A Dominguez, et al. Dftb+, a software package for efficient approximate density functional theory based atomistic simulations. *The Journal of chemical physics*, 152(12), 2020.

Ahsan Javed and Sajid Ali. Machine learning-driven insights into excitonic effects in 2d materials. *Next Research*, 2, 2025. doi: 10.1016/j.nexres.2025.100840.

Oumar Kaba and Siamak Ravanbakhsh. Equivariant networks for crystal structures. *Advances in Neural Information Processing Systems*, 35:4150–4164, 2022.

Beom Seok Kang, Vignesh C Bhethanabotla, Amin Tavakoli, Maurice D Hanisch, William A Goddard III, and Anima Anandkumar. Orbitall: A unified quantum mechanical representation deep learning framework for all molecular systems. *arXiv preprint arXiv:2507.03853*, 2025.

Konstantin Karandashev and O Anatole von Lilienfeld. An orbital-based representation for accurate quantum machine learning. *The Journal of Chemical Physics*, 156(11), 2022.

Nick R. Knøsgaard and Kristian S. Thygesen. Representing individual electronic states for machine learning GW band structures of 2d materials. *Nature Communications*, 13:468, 2022. doi: 10.1038/s41467-022-28122-0.

Hendrik J. Monkhorst and James D. Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13:5188–5192, Jun 1976. doi: 10.1103/PhysRevB.13.5188. URL https://link.aps.org/doi/10.1103/PhysRevB.13.5188.

Zhuoran Qiao, Matthew Welborn, Animashree Anandkumar, Frederick R Manby, and Thomas F Miller. Orbnet: Deep learning for quantum chemistry using symmetry-adapted atomic-orbital features. *The Journal of chemical physics*, 153(12), 2020.

Zhuoran Qiao, Anders S Christensen, Matthew Welborn, Frederick R Manby, Anima Anandkumar, and Thomas F Miller III. Informing geometric deep learning with electronic interactions to accelerate quantum chemistry. *Proceedings of the National Academy of Sciences*, 119(31): e2205221119, 2022.

Raghunathan Ramakrishnan, Pavlo O Dral, Matthias Rupp, and O Anatole Von Lilienfeld. Big data meets quantum chemistry approximations: the δ-machine learning approach. *Journal of chemical theory and computation*, 11(5):2087–2096, 2015.

Rampi Ramprasad, Rohit Batra, Ghanshyam Pilania, Arun Mannodi-Kanakkithodi, and Chiho Kim. Machine learning in materials informatics: recent applications and prospects. *npj Computational Materials*, 3:54, 2017. doi: 10.1038/s41524-017-0056-5.

John C Slater and George F Koster. Simplified lcao method for the periodic potential problem. *Physical review*, 94(6):1498, 1954.

Sherif A. Tawfik and Salvy P. Russo. Naturally-meaningful and efficient descriptors: machine learning of material properties based on robust one-shot ab initio descriptors. *Journal of Cheminformatics*, 14(1):78, 2022. doi: 10.1186/s13321-022-00658-9.

Jingfeng Wang, Zihe Li, Mengke Li, Wenyan Jiao, Yufeng Luo, Huijun Liu, and Ying Fang. Accurate prediction of band gap of two-dimensional monolayer materials via transfer learning. *Materials Today Physics*, 56:101774, 2025. ISSN 2542-5293. doi: 10.1016/j.mtphys.2025.101774.

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.

A Additional Theoretical Details

This appendix summarizes the reciprocal-to-real-space relations used to construct localized quantum mechanical matrix (QMM) blocks under periodic boundary conditions. These details complement, but do not repeat, the discussion in the main text.

Bloch basis and periodic operators. Under periodic boundary conditions, electronic states are indexed by crystal momentum $\mathbf{k}$ in the first Brillouin zone and satisfy Bloch's theorem,

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}), \tag{2}$$

where $u_{n\mathbf{k}}(\mathbf{r})$ is lattice-periodic. In a linear-combination-of-atomic-orbitals (LCAO) representation, atomic orbitals are Bloch-summed over lattice translations $\mathbf{R}$,

$$\chi^{\mu}_{\mathbf{k},A}(\mathbf{r}) = \frac{1}{\sqrt{N_{\text{cell}}}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \phi^{\mu}_{A}(\mathbf{r} - \mathbf{R}), \tag{3}$$

leading to $\mathbf{k}$-resolved operator matrices $\mathbf{O}(\mathbf{k})$ (e.g., Hamiltonian, overlap).

Real-space operator blocks. For localized atomic-orbital bases, $\mathbf{k}$-space matrices admit an equivalent representation in terms of real-space blocks indexed by lattice offsets,

$$\mathbf{O}^{\mu\nu}_{AB}(\mathbf{R}) = \frac{1}{N_k} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} \mathbf{O}^{\mu\nu}_{AB}(\mathbf{k}), \tag{4}$$

where A, B label atoms in the reference cell. These real-space blocks decay rapidly with increasing $|\mathbf{R}|$, yielding a sparse and localized description of orbital couplings that is well suited for message passing in periodic graphs.

B Extended Results

This section provides additional quantitative results and ablations that support the conclusions in the main text.

B.1 Pre-SCF vs. Post-SCF Orbital Features

We compare models trained using:

- **Pre-SCF features:** $(\mathbf{H}^0, \mathbf{S})$
- **Post-SCF features:** $(\mathbf{H}^0, \mathbf{S}, \mathbf{P})$

The comparison isolates the impact of including the self-consistent density matrix $\mathbf{P}$.

Table 2: Comparison of pre-SCF and post-SCF orbital features on C2DB. Pre-SCF uses $(\mathbf{H}^0, \mathbf{S})$; post-SCF uses $(\mathbf{H}^0, \mathbf{S}, \mathbf{P})$.

Target	Strategy	N	Pre-SCF MAE (eV)	Post-SCF MAE (eV)
PBE	Direct prediction	3191	0.32	0.26
HSE06	Direct prediction	3191	0.40	0.31
HSE06	Δ(HSE06−PBE)	3191	0.14	0.12
G_0W_0	Direct (5-fold CV)	328	0.62 ± 0.09	0.46 ± 0.07
G_0W_0	$\Delta(G_0W_0-\text{PBE})$ (5-fold CV)	328	0.27 ± 0.04	0.23 ± 0.04
G_0W_0	$\Delta(G_0W_0-\text{HSE06})$ (5-fold CV)	328	0.23 ± 0.03	0.20 ± 0.04
G_0W_0	Transfer Learning: $\Delta(\text{HSE06}-\text{PBE}) \rightarrow \Delta(G_0W_0-\text{HSE06})$	328	0.21 ± 0.05	0.17 ± 0.04

Overall, pre-SCF models remain competitive across all targets, indicating that even approximate orbital information provides a strong physical prior. Including the density matrix consistently improves accuracy, particularly for correction learning tasks.

B.2 Representative Prior Results on C2DB

For contextual reference, Table 3 summarizes representative errors reported in the literature. These values are shown for comparison only, as dataset filtering, splits, and targets may differ across studies.

Table 3: Representative reported errors on C2DB band-gap prediction from prior work.

Work	Target	MAE (eV)	Key inputs
Wang et al. (2025)	PBE	0.33	Elemental descriptors
Dau et al. (2023)	PBE	0.35	Elemental descriptors
Tawfik & Russo (2022)	Δ(HSE06$-$PBE)	0.23	PBE electronic descriptors
Haotian Liu (2025)	HSE06	0.26	PBE electronic descriptors
Knøsgaard & Thygesen (2022)	G_0W_0	0.15	PBE wavefunctions
Wang et al. (2025)	G_0W_0	0.28	Transfer from PBE
Javed & Ali (2025)	G_0W_0	0.20	PBE gap + geometry

C Calculation Settings

Orbital features were generated using periodic GFN1-xTB calculations (Grimme et al., 2017) performed with the DFTB+ software package (Hourahine et al., 2025). All calculations employed the xTB Hamiltonian with self-consistent charge (SCC) enabled and an SCF convergence tolerance of 10^{-6}. Crystal geometries were provided in VASP `POSCAR` format. Brillouin-zone sampling used a Γ-centered Monkhorst–Pack mesh (Monkhorst & Pack, 1976) with dimensions $k_x \times k_y \times k_z$, where $k_z = 1$ reflects the two-dimensional nature of the C2DB structures. Unless otherwise stated, a $5 \times 5 \times 1$ mesh was used.

D Training Configuration

All models were trained with PyTorch using the same periodic OrbNet-Crystal architecture and GFN1-xTB orbital features ($\mathbf{H}^0, \mathbf{S}, \mathbf{P}$) unless otherwise noted. Unless specified, we used seed 42, batch size 32 for PBE/HSE tasks and batch size 8 for G_0W_0 tasks.

Model and graph construction Node/edge orbital features were $\{\mathbf{H}^0, \mathbf{S}, \mathbf{P}\}$ for both one-body and two-body channels, with edge padding size 32. Periodic graphs used top-k neighbor selection with $k = 32$ and edge cutoffs of 12.0 (in bohr units) for H^0, S, P. Labels were normalized during training. Model parameters include `decoding_mode=scalarattentive` and `reduce_op=mean`.

PBE and HSE06 direct prediction For PBE and HSE06 direct prediction, we used AdamW with learning rate 5×10^{-4}, weight decay 5×10^{-4}, $\beta = (0.9, 0.95)$, and $\epsilon = 10^{-8}$. We clipped the gradient norm at 0.8 and used EMA with decay 0.999. The learning rate schedule was linear warmup + cosine annealing for 400 epochs (warmup 20 epochs, start factor 0.05, $\eta_{\min} = 10^{-6}$). The loss was MAE. Early stopping used patience 60 with ΔMAE threshold 0.5 meV (evaluated every epoch).

HSE06 delta-learning (Δ(HSE06$-$PBE)) For HSE06 delta-learning, we set base label E_g^{PBE} and target label E_g^{HSE06}. We used AdamW with learning rate 3×10^{-4} and weight decay 10^{-4}, $\beta = (0.9, 0.95)$, and $\epsilon = 10^{-8}$. We clipped gradient norm at 1.0 and used EMA with decay 0.9995. The learning rate schedule was linear warmup + cosine annealing for 400 epochs (warmup 20 epochs, start factor 0.05, $\eta_{\min} = 10^{-6}$). The loss was SmoothL1. Early stopping used patience 60 with ΔMAE threshold 0.5 meV.

G_0W_0 direct prediction (5-fold CV) For G_0W_0 direct prediction, we trained each fold with AdamW (learning rate 10^{-3}, weight decay 10^{-3}, $\beta = (0.9, 0.95)$, $\epsilon = 10^{-8}$), gradient clipping

at 0.8, and EMA decay 0.999. We used ReduceLROnPlateau (`mode=min`, factor 0.5, patience 12, threshold 0.001, minimum LR 10^{-6}), stepping on validation MAE. The loss was SmoothL1. Training ran up to 800 epochs with early stopping patience 80 and ΔMAE threshold 5 meV.

G_0W_0 transfer learning (5-fold CV) For transfer learning, we initialized the G_0W_0 model from a pretrained checkpoint (HSE-related model as specified in the config) and trained with base label E_g^{HSE06} and target label $E_g^{G_0W_0}$. We used AdamW (learning rate 5×10^{-4}, weight decay 5×10^{-3}, $\beta = (0.9, 0.95)$, $\epsilon = 10^{-8}$), gradient clipping at 0.6, EMA decay 0.999, and ReduceLROnPlateau (factor 0.5, patience 8, threshold 0.001, min LR 10^{-6}) stepped on validation MAE. The loss was SmoothL1. Training ran up to 500 epochs with early stopping patience 60 and ΔMAE threshold 1 meV.