SlaKoNet: A Unified Slater-Koster Tight-Binding Framework Using Neural Network Infrastructure for the Periodic Table

Kamal Choudhary*,†,‡

†*Department of Materials Science and Engineering, Whiting School of Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA*

‡*Department of Electrical and Computer Engineering, Whiting School of Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA*

E-mail: kchoudh2@jhu.edu

Abstract

Accurate and efficient prediction of electronic band structures is essential for designing materials with targeted properties. However, existing machine learning models often lack universality and struggle to predict detailed electronic structures, while traditional tight-binding models based on the Slater-Koster (SK) formalism suffer from (i) limited transferability, (ii) the need for manual parameterization, and (iii) training on low-fidelity electronic structure data. To address these challenges, I introduce SlaKoNet, a parameter optimization framework that learns SK-based Hamiltonian matrix elements across 65 elements of the periodic table using automatic differentiation. SlaKoNet is trained on density functional theory data from the JARVIS-DFT database using the Tran-Blaha modified Becke-Johnson (TBmBJ) functional, encompassing over 20000 materials. The framework achieves a mean absolute error (MAE) of 0.74 eV for bandgap predictions against experimental data, representing a reasonable improvement

over standard GGA functionals (MAE = 1.14 eV) while preserving the computational advantages and physical interpretability of tight-binding methods. SlaKoNet demonstrates promising scalability with up to 8.4× speedup on GPUs, enabling rapid electronic structure screening for materials discovery. SlaKoNet is publicly available at the website: `https://github.com/atomgptlab/slakonet`.

The rapid discovery and design of materials with tailored electronic properties represent one of the most pressing challenges in modern materials science.[1] Traditional experimental approaches to materials discovery are inherently slow and resource-intensive, often requiring years to decades to transition new materials from the conception to application.[2] This bottleneck has motivated the development of computational approaches that can accelerate materials discovery through high-throughput screening and inverse design strategies.[2–5]

Electronic structure calculations, particularly those based on density functional theory (DFT), have emerged as a widespread tool for predicting materials properties from first principles.[6,7] DFT provides quantum-mechanically accurate descriptions of electronic band structures, densities of states, and other fundamental properties that govern material behavior. However, despite significant advances in computational efficiency and the development of high-throughput frameworks such as JARVIS-DFT,[4] Materials Project[3] and Alexandria[8] etc., DFT calculations remain computationally expensive, preventing real-time exploration of vast chemical spaces.

To address these computational limitations, the materials science community has increasingly turned to machine learning (ML) approaches that learn complex structure-property relationships from existing DFT databases.[9–11] Graph-based neural networks have shown particular promise, with models such as SchNet,[12] Crystal Graph Convolutional Networks (CGCNN),[13] and the Atomistic Line Graph Neural Network (ALIGNN)[14] demonstrating impressive performance in predicting formation energies, bandgaps, and other scalar properties. More recently, sophisticated architectures such as ALIGNN-FF[15] and Mattersim[16] have extended ML capabilities to predicting forces and enabling molecular dynamics simu-

lations.[17]

Despite these advances, current universal ML models for materials face a fundamental limitation: they lack explicit incorporation of the underlying physics governing electronic structure. Most existing approaches treat materials as abstract graphs defined by bond distances and angles, learning statistical correlations without capturing the quantum mechanical principles that determine electronic behavior. This limitation becomes particularly pronounced when predicting detailed electronic properties like band structures and densities of states, where the spatial distribution and orbital character of electronic states are crucial.[18–21]

Physics-based approaches offer a complementary pathway that explicitly incorporates quantum mechanical principles into computational models. Among these, tight-binding (TB) methods represent a particularly attractive compromise between accuracy and computational efficiency.[22–27] TB models describe electronic structure using a minimal basis of atomic orbitals, with interactions parameterized through hopping integrals and on-site energies. Moreover, TB models integrate with other computational frameworks such as non-equilibrium Green's function (NEGF)[28–30] for current-voltage characteristic predictions and dynamical mean-field theory (DMFT)[31] for band structure predictions in strongly correlated systems. The Slater-Koster (SK)[22,23] parameterization provides a systematic framework for expressing these interactions in terms of on-site and hopping parameters.[32,33]

Traditional Slater-Koster tight-binding (SKTB) models have been successfully applied to diverse material systems, from semiconductors[34,35] to transition metal compounds[36,37] and two-dimensional materials.[38,39] However, these approaches suffer from significant limitations that have hindered their widespread adoption. First, parameter fitting typically requires manual optimization for each element or material class, making it difficult to develop transferable models across the periodic table.[40] Second, traditional parameterization schemes such as least-squares fitting often struggle to capture the complex chemical environments found in multicomponent systems, leading to poor transferability when moving from binary to

ternary compounds or when encountering new element combinations.[41]

Recent efforts have begun to address these limitations by combining TB approaches with machine learning. Large-scale TB databases and parameter sets across the periodic table are now being developed.[42,43] In parallel, machine learning techniques are being employed to accelerate TB models such as the TBMaLT project.[29,44,45] However, a deep-learning-based TB parameterization framework that offers a universal parameter set across the periodic table is still lacking. Additionally, existing TB parameterizations are typically fit to semi-local DFT data, which significantly underestimate electronic bandgaps. Recent efforts to develop more accurate and computationally feasible meta-generalized gradient approximation (GGA) functionals-such as GLLB-SC,[46] TBmBJ,[47] and R2SCAN[48] provide a promising middle ground between accuracy and throughput. A systematic parameterization of a universal TB model based on such meta-GGA datasets would be highly beneficial for the materials design community.

In this work, I introduce SlaKoNet, a unified framework that predicts Slater-Koster Hamiltonian matrix elements directly from the TBmBJ dataset using deep learning. SlaKoNet provides a universal TB parameter set inspired by the earlier three-body tight-binding (TB3Py) framework.[42] Similar to TB3Py, the parameter sets are made available for atomic numbers 1 to 65 (and their combinations), with plans to extend to the rest of the periodic table. By combining the physical principles of TB theory with the expressive strength of deep learning, SlaKoNet attempts to bridge the gap between physics-based and data-driven approaches to electronic structure prediction. The model is trained on diverse DFT data from the JARVIS-DFT database,[4,49,50] enabling it to capture complex structure-property relationships while maintaining the interpretability and computational efficiency of TB methods.

The source code and model parameters is made available on GitHub: `https://github.com/atomgptlab/slakonet`. The benchmarks used in this work will be made available in the JARVIS-Leaderboard platform especially the electronic structure component.[51] The model developed will also be available on interactive websites including, ChatGPT Material Ex-

plorer,[52] `https://atomgpt.org/` and `https://jarvis.nist.gov/`.[49]

The tight-binding method provides an efficient framework for describing electronic structure by expanding the crystal wavefunction $\psi_i(\mathbf{r})$ in terms of a linear combination of atomic orbitals (LCAO):

$$\psi_i(\mathbf{r}) = \sum_{\mu} c_{i\mu}\phi_{\mu}(\mathbf{r}), \tag{1}$$

where $c_{i\mu}$ are the expansion coefficients, and $\phi_{\mu}(\mathbf{r})$ represents the μ-th atomic orbital. This expansion leads to the generalized eigenvalue problem:

$$\sum_{\nu} H_{\mu\nu}c_{i\nu} = \varepsilon_i \sum_{\nu} S_{\mu\nu}c_{i\nu}, \tag{2}$$

where $H_{\mu\nu} = \langle\phi_{\mu}|\hat{H}|\phi_{\nu}\rangle$ and $S_{\mu\nu} = \langle\phi_{\mu}|\phi_{\nu}\rangle$ are the Hamiltonian and overlap matrices, respectively, in the basis of atomic orbitals such as s, p, d etc.

The Slater-Koster (SK) formalism provides a systematic approach to parameterize the hopping integrals between atomic orbitals based on their symmetries and the geometric arrangement of atoms. For a minimal sp^3 basis set, the SK parameters include on-site energies: ε_s, ε_p, and hopping integrals: $V_{ss\sigma}$, $V_{sp\sigma}$, $V_{pp\sigma}$, $V_{pp\pi}$, along with their corresponding overlap integrals, with σ and π as end-to-end and side-to-side overlaps respectively. For transition metals and more accurate descriptions, the basis is extended to include d orbitals, introducing additional SK parameters: d-orbital on-site energy ε_d, and hopping integrals $V_{sd\sigma}$, $V_{pd\sigma}$, $V_{pd\pi}$, $V_{dd\sigma}$, $V_{dd\pi}$, $V_{dd\delta}$, as well as corresponding overlap integrals: $S_{sd\sigma}$, $S_{pd\sigma}$, $S_{pd\pi}$, $S_{dd\sigma}$, $S_{dd\pi}$, $S_{dd\delta}$. Other orbitals can be included as well. These ten parameters for Hamiltonian and ten for overlap matrix elements are on a distance grid with cutoff upto 7 Å. I cast the distance-dependent parameters on a multivariate polynomial function. I use an interpolation scheme tailored for uniformly spaced grids, commonly used to interpolate distance-dependent quantities in tight-binding and DFTB models.

The training objective combines multiple targets to ensure accurate prediction of elec-

tronic properties across different material systems:

$$\mathcal{L} = \frac{1}{N_{\text{sys}}} \sum_{s=1}^{N_{\text{sys}}} w_s \left(\alpha \mathcal{L}\text{DOS}^{(s)} + \beta \mathcal{L}\text{gap}^{(s)} \right), \tag{3}$$

where N_{sys} is the number of training systems, α and β are set as 0.5, w_s is an optional system-specific weight, and the individual loss components are:

$$\mathcal{L}_{\text{DOS}}^{(s)} = \frac{1}{N_E} \sum_{E} \left(\rho_s^{\text{pred}}(E) - \rho_s^{\text{DFT}}(E) \right)^2, \tag{4}$$

$$\mathcal{L}_{\text{gap}}^{(s)} = \frac{1}{N_{\text{peak}}} \sum_{E \in \mathcal{E}_{\text{peak}}} \left(\rho_s^{\text{pred}}(E) - \rho_s^{\text{DFT}}(E) \right)^2 \tag{5}$$

More sophisticated loss functions, such as total energy, bulk modulus, phonons, magnons, spin-orbit coupling etc., can be added in the future as well. SlaKoNet is implemented in PyTorch[53] using a modular architecture that separates SKF parameter optimization from the automatic differentiation infrastructure. SlaKoNet employs direct parameterization of physically meaningful Slater-Koster tight-binding parameters. The model consists of trainable parameter dictionaries (implemented using PyTorch's `nn.ParameterDict`) that store distance-dependent Hamiltonian and overlap matrix elements for each element pair across the periodic table. Each element pair requires following trainable parameters: Hamiltonian and overlap, each represented as vectors corresponding to tabulated values on a uniform distance grid extending to 7 Å. The forward pass involves three key steps: (1) interpolating learned parameters using local polynomial fitting to construct Hamiltonian and overlap matrices for a given atomic configuration, (2) solving the generalized eigenvalue problem $H|\Psi\rangle = ES|\Psi\rangle$ using PyTorch's linear algebra routines, and (3) computing electronic properties from the resulting eigenvalues and eigenvectors. This approach preserves the physical interpretability of tight-binding methods while leveraging modern optimization techniques for parameter learning across diverse material systems. Model training employs the TBmBJ

dataset with AdamW optimizer with learning rate scheduling and gradient clipping for stability. Early stopping is used with a patience of $P_{\text{stop}} = 20$ epochs for a total of 200 epochs.

Note that the current implementation focuses on the electronic Hamiltonian and does not include optimized repulsive potentials. This represents a limitation of the current work, and potential terms could enhance performance, particularly for structural properties and total energies in the future. The tight-binding framework naturally provides access to derived properties such as density of states, optical transition matrix elements, and transport coefficients through the eigenvalues and eigenvectors. However, properties like atomic charges and dipole moments would require additional methodological development in the future.

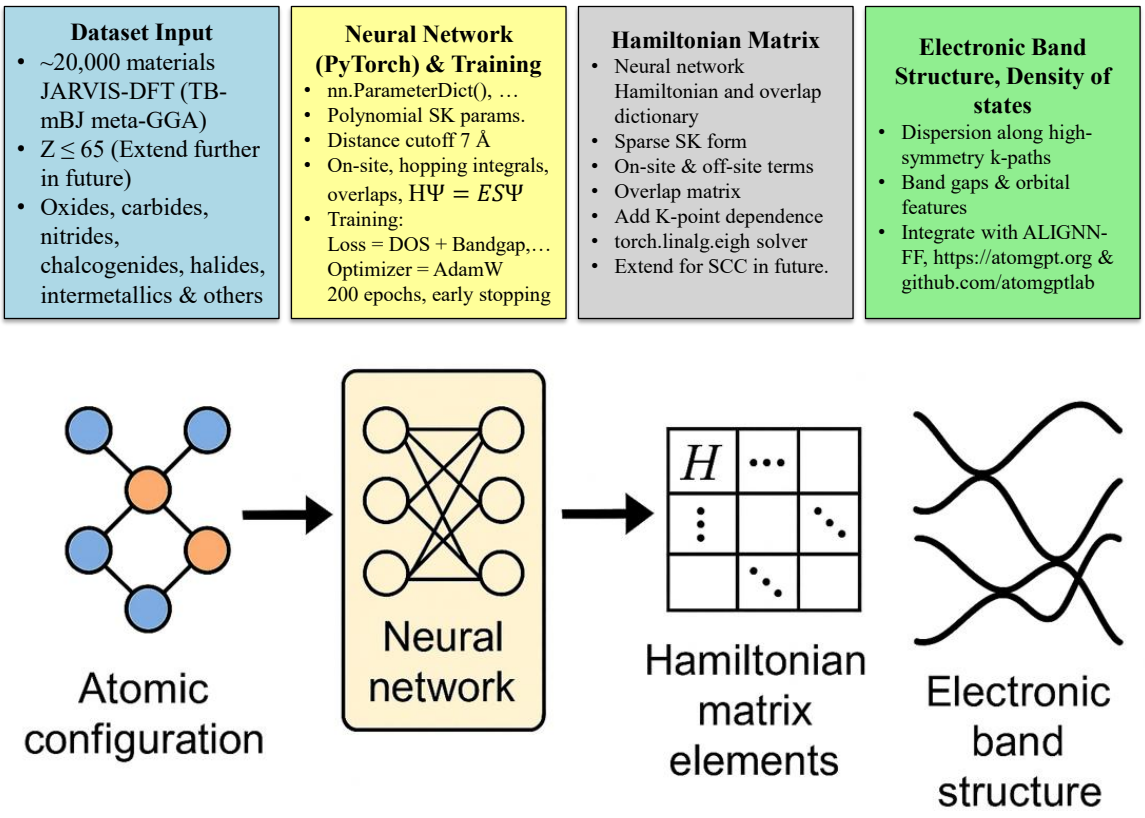

Figure 1: Schematic illustration of the SlaKoNet framework for predicting electronic band structures from atomic configurations.

The dataset used in this study was derived from the JARVIS-DFT database, with

TBmBJ-calculated bandgaps with nearly 20000 entries. After filtering out entries with Z less than 65 calculations, the final dataset spans a diverse range of chemistries and electronic structures. In terms of band-gap regimes, the composition is dominated by metallic systems (< 0.1 eV), which constitute approximately 55% of the dataset. Small-gap semiconductors (0.1 to 1 eV) represent about 6%, medium-gap semiconductors (1 to 3 eV) about 18%, large-gap semiconductors (3 to 6 eV) about 14%, and ultra-wide-gap insulators (> 6 eV) about 7%. Considering chemical complexity, the dataset consists of about 3% single-element systems, 27% binary compounds, 60% ternaries, 10% quaternaries, and only 1% quinary or higher-order systems. The Tran-Blaha modified Becke-Johnson (TBmBJ) exchange-correlation functional employed in this dataset has been extensively validated and shown to yield significantly improved predictions of electronic bandgaps and frequency-dependent dielectric functions compared to standard generalized gradient approximation (GGA) functionals, particularly for semiconductor and insulator materials where accurate bandgap prediction is crucial for technological applications. TBmBJ was shown to have accuracy close to hybrid functionals and computational speed in-between semi-local functionals and hybrid functionals.[47] The database is available on Figshare repository (`https://figshare.com/projects/JARVIS-DFT_TBmBJ/84020`) from which I extract materials with all materials with Z<65.

For each material in this dataset, I systematically extract key structural and electronic properties including complete atomic structures with crystallographic symmetries, electronic bandgaps, and electronic densities of states projected onto atomic orbitals. The dataset encompasses chemical diversity, spanning oxides (including binary, ternary, and complex perovskites), carbides (from simple binary compounds to complex MAX phases), nitrides (ranging from binary III-V semiconductors to complex oxynitrides), chalcogenides (including layered materials and thermoelectric compounds), halides, intermetallics, and various other material classes representing virtually every region of the periodic table and most known crystal structure types.

This chemical breadth ensures that the resulting SlaKoNet model captures the wide spectrum of bonding environments encountered in solid-state materials, from highly ionic systems to purely covalent networks, with the extensive coverage enabling robust interpolation and extrapolation capabilities across the materials space. The systematic inclusion of materials with varying degrees of structural complexity, from simple binary compounds to complex quaternary phases, provides the model with training examples that span the range of coordination environments and chemical environments likely to be encountered in materials discovery applications.

Building upon our previous theoretical developments,[42] I leverage a universal tight-binding model framework originally developed for the periodic table encompassing elements and their combinations with atomic numbers up to 65, thereby covering the vast majority of technologically relevant elements. The original parameterization of this model was meticulously constructed using the PBEsol exchange-correlation functional and employed conventional least-squares optimization techniques to fit TB parameters against high-quality DFT reference calculations. In the current work, I adopt a sophisticated hybrid approach that preserves the validated two-body interaction framework from our previous model while enhancing its accuracy and transferability through automatic differentiation refinement. Specifically, I reuse the carefully optimized two-body parameters that capture the fundamental orbital overlap integrals and hopping mechanisms, then systematically refine these parameters using advanced neural network architectures to better reproduce the more accurate TBmBJ-calculated bandgaps and densities of states. This refinement process focuses on both on-site terms (representing atomic orbital energies and crystal field effects) and hopping or off-site terms (describing electron delocalization and band formation), ensuring improvement across all aspects of the electronic structure description.

An important aspect of our approach lies in the sophisticated treatment of distance-dependent Slater-Koster parameters, which must accurately capture the exponential decay of orbital overlaps with interatomic separation while maintaining smooth, differentiable

functional forms suitable for gradient-based optimization. I employ a uniform distance grid extending to 7 $\mathring{A}$, where parameters are stored as discrete values and interpolated using local polynomial fitting with 8 neighboring points. Beyond the 7 $\mathring{A}$ cutoff, polynomial tail functions ensure smooth decay to zero. Such schemes are commonly used in DFTB+ frameworks[44]. This grid-based representation provides several key advantages: it ensures smooth derivatives necessary for stable gradient computation, allows direct storage of tabulated values that can capture complex distance dependencies without imposing restrictive functional forms, and maintains computational efficiency during both training and inference phases through efficient interpolation algorithms adapted from the DFTB+ implementation.

Once the parameter optimization is complete, the resulting SlaKoNet model can generate complete Hamiltonian and overlap matrices for any element or compound within its training domain at arbitrary k-points throughout the Brillouin zone. This universality represents an advancement over traditional tight-binding approaches, which typically require system-specific parameterization and manual adjustment for each new material class.

Solving the resulting generalized eigenvalue problem $H|\Psi\rangle = ES|\Psi\rangle$ yields the complete electronic band structure, including eigenvalues (energy levels) and eigenvectors (electronic wavefunctions) at each k-point. This solution provides access to all derived electronic properties, including density of states, Fermi surfaces, optical transition matrix elements, and transport coefficients, making SlaKoNet a comprehensive tool for electronic structure analysis.

Figure 1 illustrates the SlaKoNet framework for predicting electronic band structures from atomic configurations. The workflow starts with an atomic structure, where atoms are represented by colored spheres indicating different elements. The framework is trained on a comprehensive database of approximately 20000 materials from the JARVIS-DFT TB-mBJ dataset, encompassing diverse material classes including oxides, carbides, nitrides, chalcogenides, halides, and intermetallics with elements up to atomic number of 65. This structural input is processed through a bond-distance based neural network framework that learns to

predict Slater-Koster tight-binding parameters. The network outputs are used to construct the Hamiltonian matrix elements, which are then diagonalized to obtain the electronic band structure. The framework integrates PyTorch with tight-binding theory to enable rapid electronic structure calculations across diverse material systems.

This structural representation serves as input to a neural network architecture (center panel, highlighted in yellow), which has been trained to learn the complex mapping from atomic positions, chemical identities, and local coordination environments directly to the corresponding Slater-Koster Hamiltonian matrix elements in the form of PyTorch's nn.ModuleDict() and nn.ParameterDict() methods. The model learns representations that naturally encode both the radial dependencies of orbital overlaps (arising from the spatial extent and decay of atomic orbitals) and the angular dependencies that reflect the directional nature of orbital interactions (such as the characteristic angular patterns of s, p, d, and f orbitals).

The system output consists of the complete set of Hamiltonian matrix elements (center-right panel), which are schematically visualized as a sparse matrix structure where dots represent non-zero elements corresponding to orbital overlaps and hopping terms between atomic sites within the chosen cutoff radius. The sparsity pattern reflects the localized nature of atomic orbitals and the finite range of interatomic interactions, leading to computational efficiency during eigenvalue solution. The matrix elements encode essential physics of electron behavior, including on-site energies (diagonal elements) that reflect atomic orbital energies and crystal field effects, and off-site hopping terms (off-diagonal elements) that describe electron delocalization and band formation. These constructed Hamiltonian matrix elements are then used to solve the generalized eigenvalue problem using PyTorch's linear algebra routines, yielding the complete electronic band structure (rightmost panel). The band structure is depicted as energy dispersion curves plotted along high-symmetry paths in the Brillouin zone, showing the characteristic crossing patterns, band gaps, and dispersion relationships that determine all electronic and optical properties of the material.

This unified SlaKoNet framework, trained on extensive DFT data spanning the periodic table and incorporating diverse material classes, enables rapid and accurate prediction of electronic properties for a large range of materials while maintaining the computational efficiency essential for high-throughput screening applications. The approach fundamentally eliminates the need for system-specific parameter fitting or manual parameter tuning that has historically limited the applicability of tight-binding methods, thereby bridging the critical gap between computationally expensive first-principles methods and efficient empirical approaches.

The resulting methodology combines the accuracy of modern density functional theory with the computational speed of tight-binding models, enabling electronic structure calculations that are orders of magnitude faster than DFT while maintaining comparable accuracy for most properties of interest. This combination makes SlaKoNet particularly valuable for materials discovery applications where thousands or millions of candidate materials must be rapidly screened to identify promising compounds for experimental investigation.

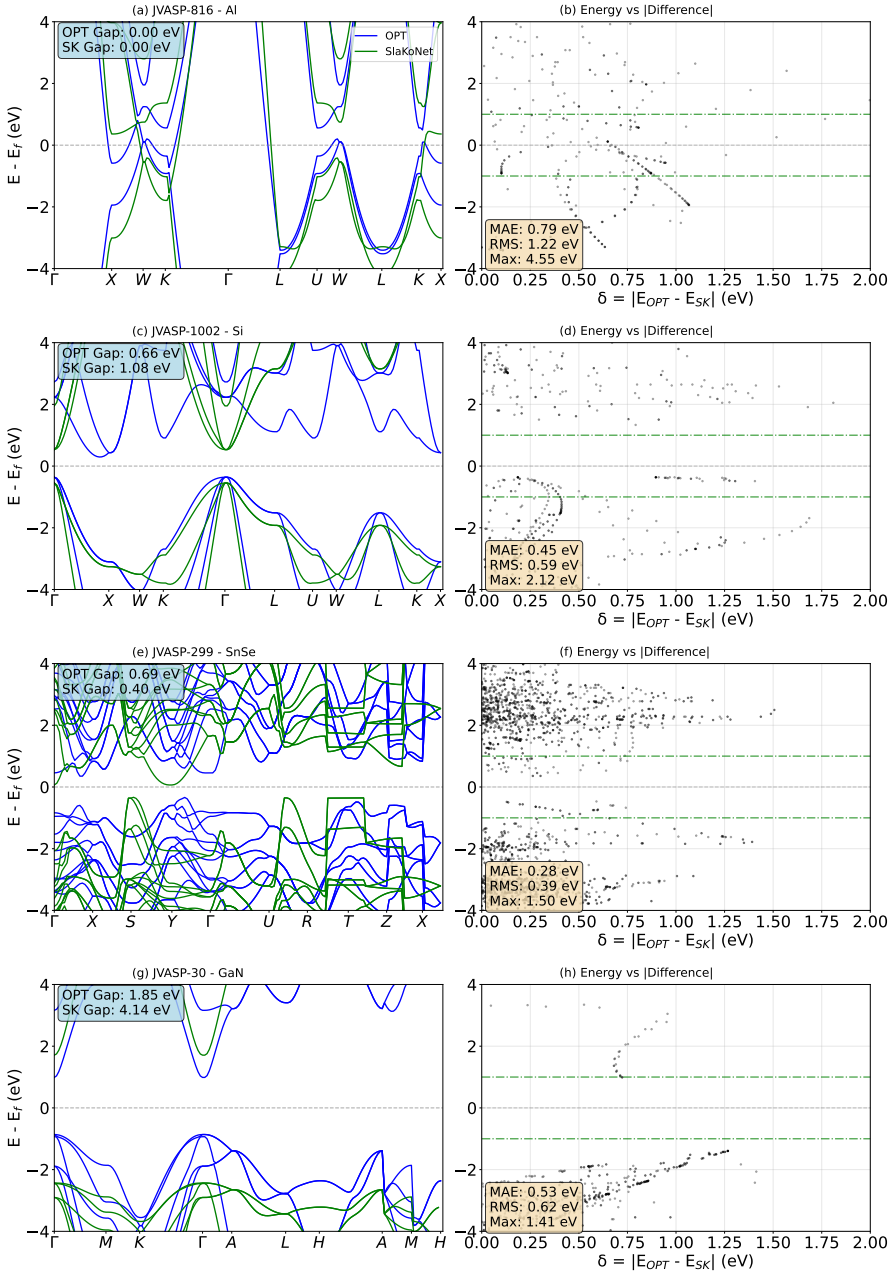

Figure 2: Comparison of electronic band structures predicted by SlaKoNet (green lines) and reference DFT calculations using the OptB88vdW (OPT) functional (blue lines) for representative materials along with band difference analysis (using mean absolute error (MAE), root mean square (RMS), maximum difference(Max): (a) Al (JVASP-816), (b) band difference for Al, (c) Si (JVASP-1002), (d) band difference for Si, (e) SnSe (JVASP-299), (f) band difference for SnSe, and (g) GaN (JVASP-30), and (h) band difference for GaN. The JVASP identifiers correspond to entries in the JARVIS-DFT database.

Now, I present example comparisons of electronic band structures predicted by SlaKoNet against reference DFT calculations in Fig. 2. These comparisons showcase SlaKoNet pre-

dictions (green lines) versus reference DFT calculations using the OptB88vdW (OPT) functional (blue lines) for four representative materials. The maximum band structure deviation is defined as:

$$\Delta E_{\max} = \max_{k,i \in \mathcal{S}_k^{\mathrm{SK}}} \left\{ \min_{j \in \mathcal{S}_k^{\mathrm{OPT}}} \left| \tilde{E}_{k,i}^{\mathrm{SK}} - \tilde{E}_{k,j}^{\mathrm{OPT}} \right| \right\} \tag{6}$$

where $\tilde{E}_{k,i} = E_{k,i} - E_F$ are the Fermi-level-adjusted eigenvalues, and $\mathcal{S}_k = \{i : |\tilde{E}_{k,i}| < \epsilon\}$ represents the set of bands within energy tolerance ϵ around the Fermi level.

Figure 2(a) presents the band structure of face-centered cubic aluminum (Al), a prototypical simple metal that serves as an ideal benchmark for testing the model's ability to capture nearly-free-electron behavior characteristic of metallic systems. Overall agreement between SlaKoNet and DFT is immediately apparent, with the model reproducing the essential features of Al's electronic structure: multiple band crossings throughout the Brillouin zone, the characteristic parabolic band dispersions near the Fermi level (EF = 0, indicated by the horizontal dashed line). The band difference plot in Fig. 2(b) shows less difference in valence band region than in the conduction bands.

Figure 2(c) showcases silicon (Si), the archetypal covalent semiconductor whose band structure has been extensively studied both theoretically and experimentally. This comparison illustrates the model's capability to predict the electronic structure of materials dominated by directional sp^3 hybridization and strong covalent bonding. SlaKoNet reproduces the overall band topology along the cubic Brillouin zone path. The complex valence band structure, featuring the characteristic splitting and dispersion patterns arising from sp^3 hybridization, shows good agreement with DFT calculations. However, discrepancies are observed in the conduction band minimum region, where SlaKoNet exhibits deviations from the DFT reference. This limitation reflects a known challenge in tight-binding approaches for silicon, where the accurate description of excited states requires careful parameterization of higher-order interactions and excited state orbital contributions that can be challenging to capture with standard Slater-Koster formalism. In the band difference analysis, higher differences in the conduction band as well than in the valence bands are observed. While there

is not a comprehensive benchmarking on all the DFTB parameters, DFTB with standard matsci/pbc parameter set predicts indirect gap nearly 0.45 to 0.9 eV which are underestimated), specialized si-band parameter set improves the description to nearly 1.0 to 1.1 eV.[54] However, systematic comparison is challenging because most DFTB parameter sets are system-specific and not publicly available for the broad range of materials in our evaluation set.

Figure 2(e) presents tin selenide (SnSe), a layered chalcogenide. This material is of significant technological interest due to its high thermoelectric performance, which arises from the complex interplay of its electronic and phononic structures. The model accurately reproduces the high density of states near the Fermi level, multiple band crossings, and the intricate dispersion patterns along the orthorhombic Brillouin zone path. However, the exact band dispersion slightly deviates from DFT. This can be attributed to the fact that the loss function is based on the density of states instead of the detailed bandstructure. In the future, I plan to include the entire bandstructure and band differences as shown in panels Fig. 2 b,d,f,h.

Figure 2(g) illustrates results for wurtzite gallium nitride (GaN), a wide-bandgap semiconductor of high importance for modern optoelectronic and power electronic applications. The material's technological significance in LED technology, laser diodes, and high-power electronics makes accurate band structure prediction crucial for device design and optimization. SlaKoNet achieves reasonable agreement in reproducing GaN's electronic structure, accurately predicting the direct bandgap at the Γ point that is fundamental to the material's optoelectronic properties. The model successfully captures the characteristic dispersion of both valence and conduction bands along the hexagonal Brillouin zone path, including the complex valence band structure arising from the interplay of Ga and N's s and p states. The errors in the band difference analysis are low compared to above cases.

Now, to extend validation against experimental measurements and demonstrate the transferability of SlaKoNet across different levels of electronic structure theory, I compare

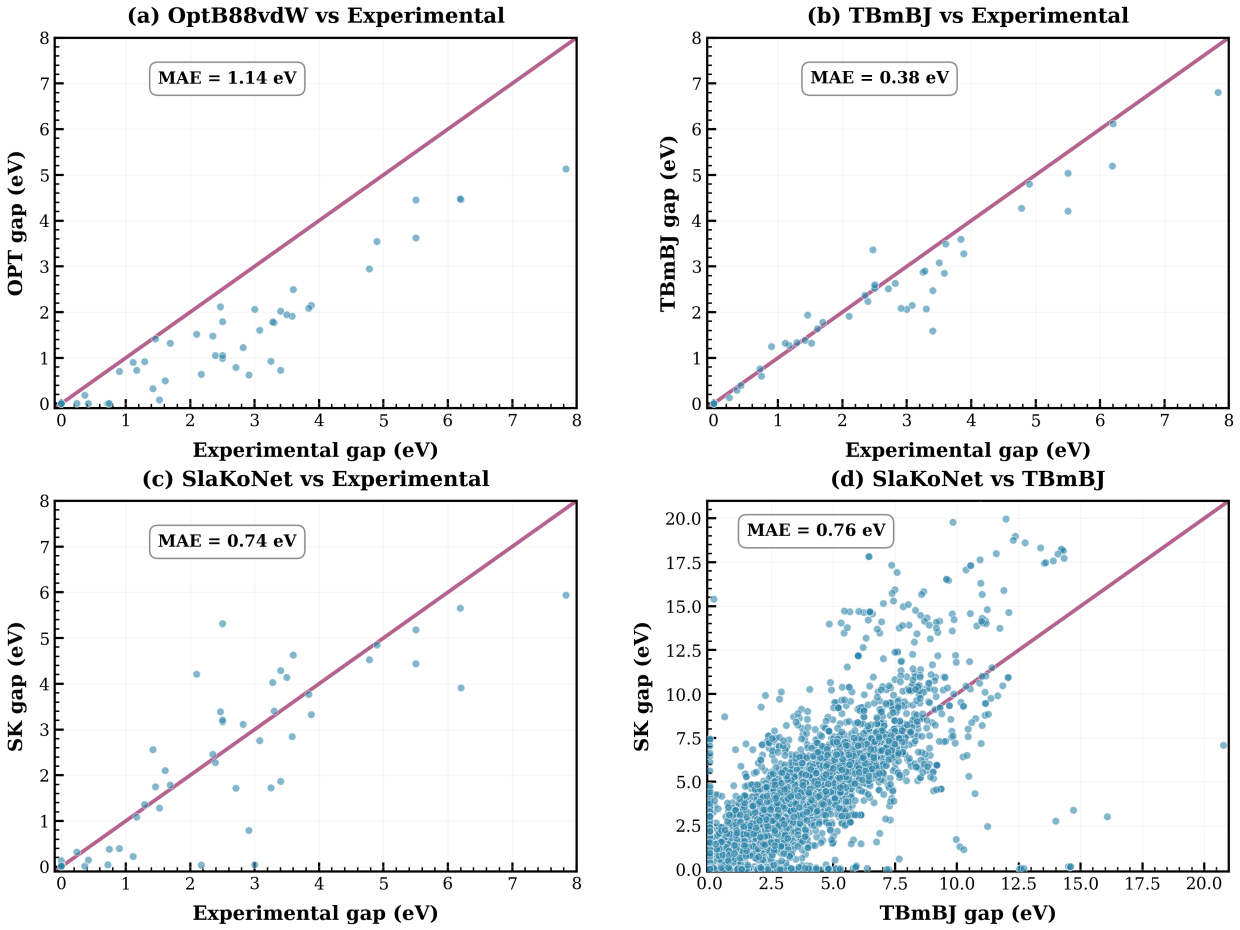

Figure 3: Comparison of bandgap predictions against experimental measurements. (a) OptB88vdW functional vs experimental data (MAE = 1.14 eV), (b) TBmBJ vs experimental data (MAE = 0.38 eV), (c) SlaKoNet predictions vs experimental data for 52 materials (MAE = 0.74 eV), (d) SlaKoNet predictions vs TBmBJ reference data from the training set (MAE = 0.76 eV). Diagonal lines indicate perfect agreement between predicted and target values.

the bandgaps obtained from DFT using two exchange-correlation functionals-OptB88vdW (OPT) and Tran-Blaha modified Becke-Johnson (TBmBJ)-with those predicted by SlaKoNet models for a diverse set of 52 semiconductor and insulator compounds, as shown in Fig. 3. This analysis spans materials with experimental bandgaps ranging from narrow-gap semiconductors (0.1 eV) to wide-gap insulators (>6 eV), providing a rigorous test of SlaKoNet's predictive capabilities across the wide spectrum of electronic behavior. The scatter plots compare predicted bandgaps from SlaKoNet models trained on different DFT functionals against experimental values, with the black diagonal line indicating perfect agreement be-

tween theory and experiment.

Figure 3(a) presents results from the OptB88vdW (OPT) functional, a generalized gradient approximation (GGA) that includes van der Waals corrections. The systematic underestimation of bandgaps is clearly evident, with most data points falling below the parity line and yielding a mean absolute error (MAE) of 1.14 eV for 52 materials in the JARVIS-Leaderboard electronic structure (ES) benchmark.[51] This substantial deviation is consistent with the well-documented bandgap problem of standard and semi-local GGA functionals, which stems from the spurious self-interaction error in the exchange-correlation functional.

Figure 3(b) displays results using the Tran-Blaha modified Becke-Johnson (TBmBJ) potential, which incorporates a momentum-dependent exchange potential designed to better approximate the exact exchange potential and improve bandgap predictions. The improvement in agreement with experimental values is clearly observed, with the data points clustering much more closely around the parity line and achieving a reduced MAE of 0.38 eV. This represents an upper limit for the accuracy that SlaKoNet should achieve. In Fig. 3(c), SlaKoNet model predictions are compared against experimental data, yielding an MAE of 0.74 eV. This intermediate performance demonstrates that SlaKoNet captures much of the improved bandgap behavior from TBmBJ training data, though with somewhat reduced accuracy compared to the reference functional and there is room for further improvement.

To better characterize the limitations of the model, a systematic error analysis against TBmBJ bandgaps was performed. Across the full dataset, the model achieves a global MAE of 0.76 eV and RMSE of 1.59 eV as shown in Fig. 3(d). Breaking down by true bandgap classes, I find that the largest errors occur for ultra-wide-gap materials ($>$ 6 eV, MAE = 2.29 eV), followed by the 3 to 6 eV range (MAE = 1.32 eV) and the 1 to 3 eV range (MAE = 1.13 eV). The model performs well for metallic systems ($\sim$0 eV, MAE = 0.17 eV), where distinguishing metals from small-gap semiconductors is critical. In terms of chemical complexity, single-element systems and compounds with $\geq$4 species tend to have higher errors (MAE $\approx$ 1.1 eV), suggesting that both data scarcity and compositional complexity

contribute to poor performance. We also identified ∼5% of cases as severe failures (absolute error ≥ 2.97 eV). Representative examples include alkali halides (Na, K, Rb) and molecular salts ($NaSO_2$, $GeCl_4$, AsF_5), where the true TBmBJ gaps are extremely wide (> 12 eV) but the model predicts nearly metallic behavior. Conversely, false positives, e.g. $CrCdF_6$ are also observed, where the true TBmBJ gap is only 0.16 eV but the model predicts ∼15 eV. These failures highlight systematic challenges with (i) ultra-wide-gap insulators and (ii) compounds outside the dominant training distribution.

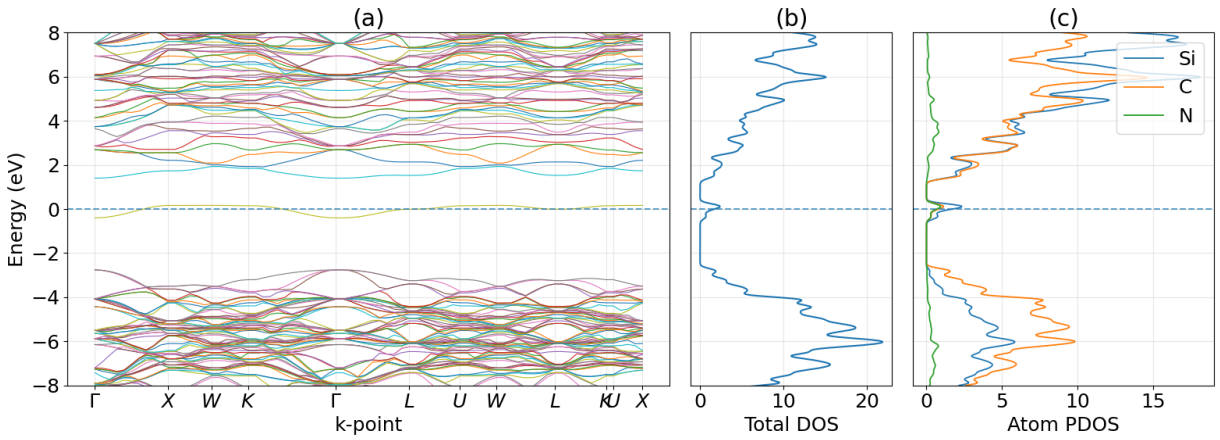

Figure 4: Electronic structure prediction of N-doped SiC (3×3×3 supercell with one C atom replaced by N) using SlaKoNet. a) band structure along high-symmetry k-path, b) total density of states, c) atom-projected density of states.

The results also illuminate the importance of training data quality in machine learning approaches to electronic structure. Most of the previous tight-binding parametrizations are based on semi-local DFT which systematically underestimates bandgaps. The difference in performance between the OPT-trained model (MAE = 1.14 eV) and the MBJ-trained model (MAE = 0.38 eV) demonstrates that SlaKoNet's predictive accuracy is fundamentally limited by the accuracy of the underlying DFT calculations used for training. This observation has important implications for the development of models in computational materials science: while such models can provide considerable computational speedups, they cannot transcend the inherent limitations of their training data. Therefore, efforts to improve machine learning predictions must be coupled with advances in the underlying electronic structure methods.

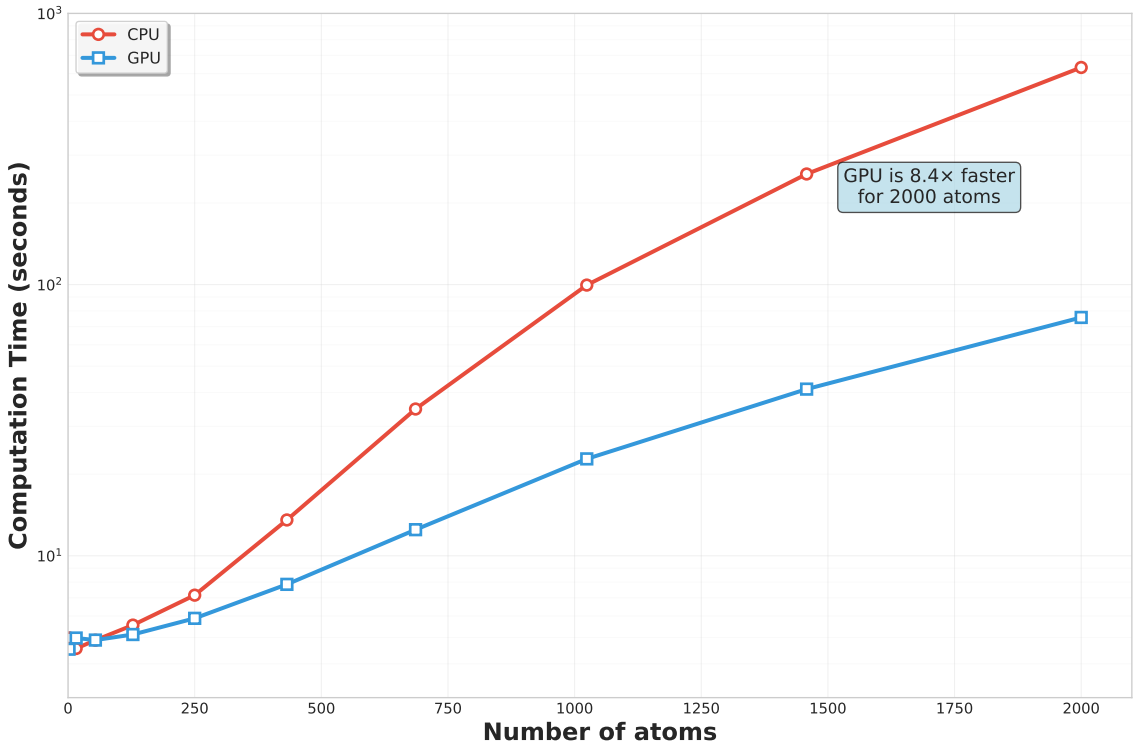

Figure 5: Computational scaling of SlaKoNet with supercell size of silicon. Wall-clock time for SlaKoNet electronic structure calculations as a function of silicon supercell size, ranging from 2×2×2 (16 atoms) to 10×10×10 (2000 atoms) are shown for (a) GPU, (b) CPU.

In addition to studying bulk materials, SlaKoNet can be useful for studying defects[55] and interfaces.[56] In Fig. 4, the electronic structure analysis of an example system: nitrogen-doped SiC ($F\bar{4}3m$) is shown. A SiC structure (JVASP-8158) was used with a 3×3×3 supercell where one of the C atoms was randomly replaced with N and relaxed with ALIGNN-FF.[15] Panel (a) shows the band structure along high-symmetry k-points, revealing that N-doping introduces additional electronic states within the band gap region around the Fermi level (0 eV), effectively reducing the semiconductor's band gap compared to pristine SiC. The total density of states (DOS) in panel (b) confirms this band gap narrowing, with enhanced electronic density near the Fermi level. Panel (c) provides the atom-projected DOS (PDOS), decomposing the electronic contributions from silicon (blue), carbon (orange), and nitrogen (green) atoms. The PDOS analysis reveals that nitrogen atoms contribute to states near the Fermi level, indicating that the N dopant introduces donor levels that

modify the electronic properties of the SiC crystal. This electronic structure modification is crucial for understanding how nitrogen doping enhances the n-type conductivity of SiC, making it suitable for semiconductor device applications. This behavior is similar to that observed in previous studies,[57,58] where the doped system was observed to exhibit potential superconducting/metallic behavior.

Wall-clock computation time for SlaKoNet electronic structure calculations as a function of silicon supercell dimensions, ranging from 2×2×2 (16 atoms) to 10×10×10 (2000 atoms) is shown in Fig. 5 to demonstrate the method's computational scalability and efficiency across different hardware architectures: (a) GPU-accelerated calculations and (b) CPU-only calculations. This systematic scaling analysis provides crucial insights into the practical computational limits and optimal deployment strategies for SlaKoNet across the range of system sizes commonly encountered in materials discovery applications.

The scaling analysis reveals fundamentally distinct computational behaviors between hardware architectures, reflecting the different strengths and limitations of each platform for machine learning-based electronic structure calculations. GPU calculations (panel a) exhibit remarkably modest time increases up to moderate supercell sizes (6×6×6, corresponding to 432 atoms), with computation times remaining impressively low-well below 10 seconds for systems up to 8×8×8 (1024 atoms). This performance in the small-to-moderate system size regime demonstrates the effectiveness of GPU acceleration for the tensor operations and matrix algebra that dominate SlaKoNet computations.

However, a pronounced inflection point occurs at larger system sizes, with a sharp increase in computational cost observed for the largest 10×10×10 supercell (2000 atoms), reaching approximately 70 seconds. This performance transition likely indicates the onset of memory bandwidth limitations, where the GPU's high-bandwidth memory becomes saturated, or algorithmic bottlenecks where certain computational operations do not scale favorably with the massively parallel GPU architecture. The transition suggests that for GPU implementations, there exists a practical system size threshold beyond which the computational

advantages begin to diminish rapidly.

In contrast, CPU calculations (panel b) exhibit more pronounced and earlier-onset scaling challenges, with computation times remaining reasonable (under 20 seconds) only for smaller systems up to approximately 7×7×7 supercells (686 atoms). Beyond this more restrictive threshold, CPU calculations display noticeable time increases, reaching over 300 seconds for the largest 10×10×10 system-representing more than a 4× increase in computation time compared to the GPU implementation. This steeper scaling curve indicates fundamentally less favorable scaling characteristics for CPU architectures when applied to SlaKoNet calculations.

The approximately 8× performance advantage of GPU over CPU for larger systems provides evidence for the effectiveness of GPU acceleration in machine learning-based electronic structure methods like SlaKoNet. This performance gap becomes particularly pronounced at larger system sizes, where the GPU's massively parallel architecture can efficiently handle the extensive tensor operations required for Hamiltonian matrix construction and eigenvalue problems inherent to tight-binding calculations.

The steeper scaling observed in CPU calculations may reflect several architectural limitations: less efficient parallelization of the underlying tensor operations due to limited core counts compared to GPU thread counts, suboptimal memory access patterns that fail to fully utilize CPU cache hierarchies, and the inherently sequential nature of certain algorithmic components that cannot be effectively parallelized on CPU architectures. In contrast, the GPU's specialized tensor processing units and high-bandwidth memory enable more efficient execution of the matrix operations that dominate SlaKoNet computations.

These evaluations were performed on NVIDIA H200 and a laptop for comparing the timings. All calculations employed identical model parameters, and numerical precision across all supercell sizes to ensure meaningful comparison. The benchmarks included full electronic structure calculations with gamma-point computation. The results demonstrate SlaKoNet's practical applicability for moderately large systems while clearly identifying computational

limits that provide valuable guidance for future algorithmic optimizations and hardware deployment strategies. The performance characteristics strongly suggest that GPU acceleration is not merely beneficial but essentially required for routine application to systems containing hundreds to thousands of atoms, making SlaKoNet a viable and competitive tool for materials discovery workflows requiring high-throughput electronic structure screening. The identified scaling transitions also provide important guidance for computational resource allocation in high-throughput screening campaigns. For systems below 1000 atoms, GPU implementations provide noticeable performance with sub-minute computation times, enabling rapid screening of large material databases. For larger systems approaching 2000 atoms, the computational cost becomes more significant but remains manageable for targeted investigations of specific materials of interest. These findings have important implications for the practical deployment of machine learning-based electronic structure methods in materials discovery pipelines. The clear performance advantages of GPU acceleration support the trend toward GPU-centric computational infrastructure in materials science, while the identified scaling limits provide realistic expectations for the system sizes that can be routinely studied with current hardware generations. Future algorithmic optimizations might focus on addressing the memory bandwidth limitations that appear to dominate performance for very large systems, potentially through improved memory access patterns, algorithmic reformulations that reduce memory requirements, or hybrid CPU-GPU and mixed-precision implementations that leverage the strengths of both architectures. Additionally, the scaling analysis provides a baseline for evaluating future hardware generations and assessing the potential benefits of emerging accelerator architectures for electronic structure calculations. The performance characteristics position SlaKoNet as an important tool that effectively bridges the gap between computationally expensive first-principles methods and the rapid screening requirements of modern materials discovery, enabling electronic structure calculations at scales and throughputs previously achievable only with much more approximate methods.

Acknowledgements

K.C. thanks computational resources from the Johns Hopkins University and Dr. Kevin F. Garrity at the National Institute of Standards and Technology for helpful discussion.

References

(1) Sze, S. M.; Li, Y.; Ng, K. K. *Physics of semiconductor devices*; John wiley & sons, 2021.

(2) Curtarolo, S.; Hart, G. L.; Nardelli, M. B.; Mingo, N.; Sanvito, S.; Levy, O. The high-throughput highway to computational materials design. *Nature materials* **2013**, *12*, 191–201.

(3) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; others Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL materials* **2013**, *1*.

(4) Choudhary, K.; Garrity, K. F.; Reid, A. C.; DeCost, B.; Biacchi, A. J.; Hight Walker, A. R.; Trautt, Z.; Hattrick-Simpers, J.; Kusne, A. G.; Centrone, A.; others The joint automated repository for various integrated simulations (JARVIS) for data-driven materials design. *npj computational materials* **2020**, *6*, 173.

(5) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (OQMD). *Jom* **2013**, *65*, 1501–1509.

(6) Sholl, D. S.; Steckel, J. A. *Density functional theory: a practical introduction*; John Wiley & Sons, 2022.

(7) Kittel, C.; McEuen, P. *Introduction to solid state physics*; John Wiley & Sons, 2018.

(8) Schmidt, J.; Cerqueira, T. F.; Romero, A. H.; Loew, A.; Jäger, F.; Wang, H.-C.; Botti, S.; Marques, M. A. Improving machine-learning models in materials science through large datasets. *Materials Today Physics* **2024**, *48*, 101560.

(9) Choudhary, K.; DeCost, B.; Chen, C.; Jain, A.; Tavazza, F.; Cohn, R.; Park, C. W.; Choudhary, A.; Agrawal, A.; Billinge, S. J.; others Recent advances and applications of deep learning methods in materials science. *npj Computational Materials* **2022**, *8*, 1–26.

(10) Schmidt, J.; Marques, M. R.; Botti, S.; Marques, M. A. Recent advances and applications of machine learning in solid-state materials science. *npj computational materials* **2019**, *5*, 83.

(11) Vasudevan, R. K.; Choudhary, K.; Mehta, A.; Smith, R.; Kusne, G.; Tavazza, F.; Vlcek, L.; Ziatdinov, M.; Kalinin, S. V.; Hattrick-Simpers, J. Materials science in the artificial intelligence age: high-throughput library generation, machine learning, and a pathway from correlations to the underpinning physics. *MRS communications* **2019**, *9*, 821–838.

(12) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. Schnet–a deep learning architecture for molecules and materials. *The Journal of chemical physics* **2018**, *148*.

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

(14) Choudhary, K.; DeCost, B. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials* **2021**, *7*, 185.

(15) Choudhary, K.; DeCost, B.; Major, L.; Butler, K.; Thiyagalingam, J.; Tavazza, F.

Unified graph neural network force-field for the periodic table: solid state applications. *Digital Discovery* **2023**, *2*, 346–355.

(16) Yang, H.; Hu, C.; Zhou, Y.; Liu, X.; Shi, Y.; Li, J.; Li, G.; Chen, Z.; Chen, S.; Zeni, C.; others Mattersim: A deep learning atomistic model across elements, temperatures and pressures. *arXiv preprint arXiv:2405.04967* **2024**,

(17) Wines, D.; Choudhary, K. Chips-ff: Evaluating universal machine learning force fields for material properties. *ACS Materials Letters* **2025**, *7*, 2105–2114.

(18) Li, H.; Wang, Z.; Zou, N.; Ye, M.; Xu, R.; Gong, X.; Duan, W.; Xu, Y. Deep-learning density functional theory Hamiltonian for efficient ab initio electronic-structure calculation. *Nature Computational Science* **2022**, *2*, 367–377.

(19) Qi, Y.; Gong, W.; Yan, Q. Bridging deep learning force fields and electronic structures with a physics-informed approach. *npj Computational Materials* **2025**, *11*, 177.

(20) Schütt, K. T.; Gastegger, M.; Tkatchenko, A.; Müller, K.-R.; Maurer, R. J. Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions. *Nature communications* **2019**, *10*, 5024.

(21) Schätzle, Z.; Szabó, P. B.; Mezera, M.; Hermann, J.; Noé, F. DeepQMC: An open-source software suite for variational optimization of deep-learning molecular wave functions. *The Journal of Chemical Physics* **2023**, *159*.

(22) Koskinen, P.; Mäkinen, V. Density-functional tight-binding for beginners. *Computational Materials Science* **2009**, *47*, 237–253.

(23) Slater, J. C.; Koster, G. F. Simplified LCAO method for the periodic potential problem. *Physical review* **1954**, *94*, 1498.

(24) Cohen, R. E.; Mehl, M. J.; Papaconstantopoulos, D. A. Tight-binding total-energy method for transition and noble metals. *Physical Review B* **1994**, *50*, 14694.

(25) Aradi, B.; Hourahine, B.; Frauenheim, T. DFTB+, a sparse matrix-based implementation of the DFTB method. *The Journal of Physical Chemistry A* **2007**, *111*, 5678–5684.

(26) Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D. Maximally localized Wannier functions: Theory and applications. *Reviews of Modern Physics* **2012**, *84*, 1419–1475.

(27) Chadi, D.; Cohen, M. L. Tight-binding calculations of the valence bands op diamond and zincblende crystals. **1974**,

(28) Pecchia, A.; Penazzi, G.; Salvucci, L.; Di Carlo, A. Non-equilibrium Green's functions in density functional tight binding: method and applications. *New Journal of Physics* **2008**, *10*, 065022.

(29) Gu, Q.; Zhouyin, Z.; Pandey, S. K.; Zhang, P.; Zhang, L.; E, W. Deep learning tight-binding approach for large-scale electronic simulations at finite temperatures with ab initio accuracy. *Nature Communications* **2024**, *15*, 6772.

(30) Zhao, Y.; Wan, Z.; Hetmanuik, U.; Anantram, M. Negative differential resistance in graphene boron nitride heterostructure controlled by twist and phonon-scattering. *IEEE Electron Device Letters* **2016**, *37*, 1242–1245.

(31) Singh, V.; Herath, U.; Wah, B.; Liao, X.; Romero, A. H.; Park, H. DMFTwDFT: An open-source code combining Dynamical Mean Field Theory with various density functional theory packages. *Computer Physics Communications* **2021**, *261*, 107778.

(32) Papaconstantopoulos, D. A.; others *Handbook of the Band Structure of Elemental Solids: From Z*; Springer, 1986.

(33) Mehl, M. J.; Papaconstantopoulos, D. A. Applications of a tight-binding total-energy method for transition and noble metals: Elastic constants, vacancies, and surfaces of monatomic metals. *Physical Review B* **1996**, *54*, 4519.

(34) Vogl, P.; Hjalmarson, H. P.; Dow, J. D. A semi-empirical tight-binding theory of the electronic structure of semiconductors. *Journal of physics and chemistry of solids* **1983**, *44*, 365–378.

(35) Jancu, J.-M.; Scholz, R.; Beltram, F.; Bassani, F. Empirical spds* tight-binding calculation for cubic semiconductors: General method and material parameters. *Physical Review B* **1998**, *57*, 6493.

(36) Andersen, O. K.; Jepsen, O. Explicit, first-principles tight-binding theory. *Physical Review Letters* **1984**, *53*, 2571.

(37) Wills, J. M.; Cooper, B. R. Synthesis of band and model Hamiltonian theory for hybridizing cerium systems. *Physical Review B* **1987**, *36*, 3809.

(38) Liu, G.-B.; Shan, W.-Y.; Yao, Y.; Yao, W.; Xiao, D. Three-band tight-binding model for monolayers of group-VIB transition metal dichalcogenides. *Physical Review B—Condensed Matter and Materials Physics* **2013**, *88*, 085433.

(39) Ridolfi, E.; Le, D.; Rahman, T.; Mucciolo, E.; Lewenkopf, C. A tight-binding model for MoS2 monolayers. *Journal of Physics: Condensed Matter* **2015**, *27*, 365501.

(40) Goringe, C.; Bowler, D.; Hernández, E. Tight-binding modelling of materials. *Reports on Progress in Physics* **1997**, *60*, 1447.

(41) Sutton, A. P.; Finnis, M. W.; Pettifor, D. G.; Ohta, Y. The tight-binding bond model. *Journal of Physics C: Solid State Physics* **1988**, *21*, 35.

(42) Garrity, K. F.; Choudhary, K. Fast and accurate prediction of material properties with three-body tight-binding model for the periodic table. *Physical review materials* **2023**, *7*, 044603.

(43) Cui, M.; Reuter, K.; Margraf, J. T. Obtaining Robust Density Functional Tight-Binding

Parameters for Solids across the Periodic Table. *Journal of Chemical Theory and Computation* **2024**, *20*, 5276–5290.

(44) McSloy, A.; Fan, G.; Sun, W.; Hölzer, C.; Friede, M.; Ehlert, S.; Schütte, N.-E.; Grimme, S.; Frauenheim, T.; Aradi, B. TBMaLT, a flexible toolkit for combining tight-binding and machine learning. *The Journal of Chemical Physics* **2023**, *158*.

(45) Burrill, D. J.; Liu, C.; Taylor, M. G.; Cawkwell, M. J.; Perez, D.; Batista, E. R.; Lubbers, N.; Yang, P. MLTB: Enhancing Transferability and Extensibility of Density Functional Tight-Binding Theory with Many-body Interaction Corrections. *Journal of Chemical Theory and Computation* **2025**, *21*, 1089–1097.

(46) Castelli, I. E.; Olsen, T.; Datta, S.; Landis, D. D.; Dahl, S.; Thygesen, K. S.; Jacobsen, K. W. Computational screening of perovskite metal oxides for optimal solar light capture. *Energy & Environmental Science* **2012**, *5*, 5814–5819.

(47) Choudhary, K.; Zhang, Q.; Reid, A. C.; Chowdhury, S.; Van Nguyen, N.; Trautt, Z.; Newrock, M. W.; Congo, F. Y.; Tavazza, F. Computational screening of high-performance optoelectronic materials using OptB88vdW and TB-mBJ formalisms. *Scientific data* **2018**, *5*, 1–12.

(48) Kingsbury, R.; Gupta, A. S.; Bartel, C. J.; Munro, J. M.; Dwaraknath, S.; Horton, M.; Persson, K. A. Performance comparison of r 2 SCAN and SCAN metaGGA density functionals for solid materials via an automated, high-throughput computational workflow. *Physical Review Materials* **2022**, *6*, 013801.

(49) Choudhary, K. The JARVIS infrastructure is all you need for materials design. *Computational Materials Science* **2025**, *259*, 114063.

(50) Wines, D.; Gurunathan, R.; Garrity, K. F.; DeCost, B.; Biacchi, A. J.; Tavazza, F.; Choudhary, K. Recent progress in the JARVIS infrastructure for next-generation data-driven materials design. *Applied Physics Reviews* **2023**, *10*.

(51) Choudhary, K.; Wines, D.; Li, K.; Garrity, K. F.; Gupta, V.; Romero, A. H.; Krogel, J. T.; Saritas, K.; Fuhr, A.; Ganesh, P.; others JARVIS-Leaderboard: a large scale benchmark of materials design methods. *npj Computational Materials* **2024**, *10*, 93.

(52) Choudhary, K. ChatGPT Material Explorer: Design and Implementation of a Custom GPT Assistant for Materials Science Applications. *Integrating Materials and Manufacturing Innovation* **2025**, 1–8.

(53) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; others Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems* **2019**, *32*.

(54) Markov, S.; Aradi, B.; Yam, C.-Y.; Xie, H.; Frauenheim, T.; Chen, G. Atomic level modeling of extremely thin silicon-on-insulator MOSFETs including the silicon dioxide: Electronic structure. *IEEE Transactions on Electron Devices* **2015**, *62*, 696–704.

(55) Choudhary, K.; Sumpter, B. G. Can a deep-learning model make fast predictions of vacancy formation in diverse materials? *AIP Advances* **2023**, *13*.

(56) Choudhary, K.; Garrity, K. F. InterMat: accelerating band offset prediction in semiconductor interfaces with DFT and deep learning. *Digital Discovery* **2024**, *3*, 1365–1377.

(57) Liu, N.; Wang, W.; Guo, L.; Peng, T.; Chen, X. Superconductivity in nitrogen-doped 3C–SiC from first-principles calculations. *Modern Physics Letters B* **2017**, *31*, 1750116.

(58) Rajabzadeh, Z.; Hammer, R.; Romaner, L. On the impact of nitrogen and aluminum doping on silicon carbide polytype stability: Insights from first principle calculations. *Journal of the American Ceramic Society* **2025**, *108*, e20284.