

Review of machine learning tight-binding models: Route to accurate and scalable electronic simulations

Jijie Zou(邹暨捷), Zhanghao Zhouyin(周寅张皓), Shishir Kumar Pandey, and Qiangqiang Gu(顾强强)
Citation: Chin. Phys. B, 2026, 35 (**1**): 017101. DOI: 10.1088/1674-1056/ae15ef
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Electronic structure and disorder effect of $La_3Ni_2O_7$ superconductor

Yuxin Wang(王郁欣), Yi Zhang(张燚), and Kun Jiang(蒋坤)
Chin. Phys. B, 2025, 34 (**4**): 047105. DOI: 10.1088/1674-1056/adbacc

Classifying extended, localized and critical states in quasiperiodic lattices via unsupervised learning

Bohan Zheng(郑博涵), Siyu Zhu(朱思宇), Xingping Zhou(周兴平), and Tong Liu(刘通)
Chin. Phys. B, 2025, 34 (**1**): 017103. DOI: 10.1088/1674-1056/ad8cb9

Prediction of impurity spectrum function by deep learning algorithm

Ting Liu(刘婷), Rong-Sheng Han(韩榕生), and Liang Chen(陈亮)
Chin. Phys. B, 2024, 33 (**5**): 057102. DOI: 10.1088/1674-1056/ad3908

Quantitative structure-plasticity relationship in metallic glass: A machine learning study

Yicheng Wu(吴义成), Bin Xu(徐斌), Yitao Sun(孙奕韬), and Pengfei Guan(管鹏飞)
Chin. Phys. B, 2021, 30 (**5**): 057103. DOI: 10.1088/1674-1056/abdda5

The effects of combining alloying elements on the elastic properties of γ-Ni in Ni-based superalloy: High-throughput first-principles calculations

Baokun Lu(路宝坤), Chongyu Wang(王崇愚)
Chin. Phys. B, 2018, 27 (**7**): 077104. DOI: 10.1088/1674-1056/27/7/077104

TOPICAL REVIEW — AI + Physical Science

Review of machine learning tight-binding models: Route to accurate and scalable electronic simulations

Jijie Zou(邹暨捷)[2,3], Zhanghao Zhouyin(周寅张皓)[4], Shishir Kumar Pandey[5,6], and Qiangqiang Gu(顾强强)[1,2,7,8,†]

[1] *School of Artificial Intelligence and Data Science, University of Science and Technology of China, Hefei* 230026, *China*
[2] *AI for Science Institute, Beijing* 100080, *China*
[3] *Center for Nanoscale Science and Technology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing* 100871, *China*
[4] *Department of Physics, McGill University, Montreal, Quebec H3A2T8, Canada*
[5] *Department of General Sciences, Birla Institute of Technology and Science, Pilani, Dubai Campus, Dubai International Academic City, Dubai, United Arab Emirates*
[6] *Department of Physics, Birla Institute of Technology and Science, Pilani, Hyderabad Campus, Jawahar Nagar, Kapra Mandal, Medchal District, Telangana* 500078, *India*
[7] *Suzhou Institute for Advanced Research, University of Science and Technology of China, Suzhou* 215123, *China*
[8] *Suzhou Big Data & AI Research and Engineering Center, Suzhou* 215123, *China*

(Received 30 July 2025; revised manuscript received 10 October 2025; accepted manuscript online 22 October 2025)

The rapid advancement of machine learning based tight-binding Hamiltonian (MLTB) methods has opened new avenues for efficient and accurate electronic structure simulations, particularly in large-scale systems and long-time scenarios. This review begins with a concise overview of traditional tight-binding (TB) models, including both (semi-)empirical and first-principles approaches, establishing the foundation for understanding MLTB developments. We then present a systematic classification of existing MLTB methodologies, grouped into two major categories: direct prediction of TB Hamiltonian elements and inference of empirical parameters. A comparative analysis with other ML-based electronic structure models is also provided, highlighting the advancement of MLTB approaches. Finally, we explore the emerging MLTB application ecosystem, highlighting how the integration of MLTB models with a diverse suite of post-processing tools from linear-scaling solvers to quantum transport frameworks and molecular dynamics interfaces is essential for tackling complex scientific problems across different domains. The continued advancement of this integrated paradigm promises to accelerate materials discovery and open new frontiers in the predictive simulation of complex quantum phenomena.

Keywords: machine learning, tight-binding model, electronic simulations

PACS: 71.15.–m, 31.15.A–, 31.15.E, 31.15.es

DOI: 10.1088/1674-1056/ae15ef **CSTR:** 32038.14.CPB.ae15ef

1. Introduction

The accurate and efficient modeling of electronic structures lies at the heart of modern physical sciences, underpinning advancements across condensed matter physics, materials science, chemistry, nanotechnology, etc. While first-principles methods such as density functional theory (DFT)[1,2] provide a robust framework for describing the full complexity of electronic structures, its computational demands often make it impractical for larger or more intricate systems. This situation usually leads to unnecessary complexity when the fundamental physics governing a material is predominantly determined by its low-energy electronic states, which is often the case. This means that while DFT offers a comprehensive picture, its exhaustive approach can be overkill for phenomena primarily driven by these more accessible energy levels. In this context, tight-binding (TB) methods[3,4] offer a more practical alternative by representing the electronic Hamiltonian in a simplified form. By specifically focusing on the low-energy space and describing the electronic structure in a local basis with a manageable size and hoppings to only a few nearest neighbors, TB methods are significantly more computationally efficient than the DFT method. While the formal scaling for the eigenvalue diagonalization step is $\mathcal{O}(N_{\rm orb}^3)$ for both approaches, the vastly smaller basis size ($N_{\rm orb}$) in TB results in a dramatic, orders-of-magnitude speedup in practice. Hence TB provides an efficient and physically intuitive approach to approximate the quantum behavior of electrons in solids and molecules.

Generally, a parameterized TB Hamiltonian operator in the second quantization form can be written as

$$\hat{H} = \sum_i \varepsilon_i c_i^\dagger c_i + \sum_{i \neq j} t_{ij} c_i^\dagger c_j, \tag{1}$$

where i, j is the compact index indicating the site, orbital as well as spin degree of freedom, and $c_i^\dagger$ (c_i) is the creation (annihilation) operator. ε_i and t_{ij} are the on-site energy and hopping amplitudes, respectively. To accurately describe the

†Corresponding author. E-mail: guqq@ustc.edu.cn

© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

electronic structure of a given system, it is crucial to determine the appropriate and meaningful values for these parameters. For realistic systems involving a large number of atoms and complex chemical environments, determining these parameters is a challenging task. Traditional TB methods rely on (semi-)empirical or first-principles parameterization strategies, which often require extensive human efforts and are difficult to generalize across different chemical environments.

In recent years, machine learning-based TB (MLTB) methods have emerged as a powerful and versatile approach to overcome this parameterization challenge.[5–16] The central philosophy of MLTB is to build a computationally efficient surrogate model that inherits the accuracy of first-principles calculations while retaining the efficiency and physical intuition of the tight-binding framework. Compared to traditional TB parameterization schemes, MLTB eliminates the need for manual parameter fitting and allows for systematic improvement and transferability across diverse systems. This has spurred the development of a diverse array of methodologies. One major branch of research focuses on directly learning the TB Hamiltonian matrix elements,[7–9,11,14–16] often by training neural networks on first-principles band structures or pre-computed Wannier Hamiltonians. A parallel and equally powerful approach aims to learn the underlying physical parameters of the TB model itself,[5,6,10,12,13] such as the Slater–Koster integrals (SKIs). Recent advances in this direction have even gone beyond the traditional two-center approximation by incorporating environment-dependent corrections.

The development of these predictive and transferable MLTB models has profound implications for computational simulation. By eliminating the need for manual, system-specific fitting, they are unlocking the ability to perform high-fidelity electronic structure simulations at unprecedented scales. As will be explored in this review, MLTB models are enabling simulations pushing towards the million-atom regime when combined with linear-scaling solvers,[13] allowing for high-fidelity device simulations when integrated into quantum transport frameworks,[17] and providing powerful tools to study the rich physics of electron–phonon interactions and nonadiabatic processes when coupled with molecular dynamics.[9,13,18] These advances are the central theme of this review: the emergence of a comprehensive MLTB application ecosystem that bridges the gap between accurate models and large-scale scientific simulations.

This review is organized as follows. We begin in Section 2 with a concise discussion of traditional TB approaches to establish the historical context. In Section 3, we present a systematic overview of modern MLTB methods, including their model architectures, descriptor design, and training strategies, and categorizes them by their prediction targets. We further situate them within the broader landscape of machine learning for electronic structure. We then shift focus in Section 4 to the emerging MLTB application ecosystem, analyzing how the integration of these models with efficient solvers, specialized transport frameworks, and dynamics interfaces enables the simulation of complex phenomena in large-scale systems. Finally, in Section 5, we conclude by summarizing the state of the art, highlighting key challenges, and outlining future directions for this rapidly advancing field.

2. Traditional TB approaches

Over the decades, various traditional TB approaches have been developed to balance computational efficiency with accuracy, ranging from (semi-)empirical parameterizations to first-principles-derived variants. In this section, we briefly review the key (semi-)empirical and first-principles TB approaches, highlighting their advantages and inherent limitations.

2.1. (Semi-)empirical TB methods

(Semi-)empirical TB methods offer a computationally efficient alternative to first-principles calculations by simplifying the electronic Hamiltonian into a parameterized form. The central idea is to replace complex quantum mechanical integrals with a smaller set of tractable parameters that are fitted against experimental data or higher-level theoretical results. Representative approaches, which we will discuss, include the density functional tight-binding (DFTB)[19–21] and the extended tight-binding (xTB) families of methods.[22–26]

A widely adopted parameterization strategy is the Slater–Koster (SK) formulation,[4] which employs the on-site energy and hopping terms with a minimal basis set under the two-center approximation.

The on-site energies are given by the l-dependent constants ε_l, where l is the angular quantum number of the corresponding orbital. For example, $l = 0, 1, 2$ correspond to s, p, and d orbitals, respectively. The hopping terms take the form

$$H_{i,j}^{lm,l'm'} = \sum_{\zeta} \left[\mathcal{U}_{\zeta}(\hat{\boldsymbol{r}}_{ij})\, h_{ll'\zeta} \right]_{mm'}, \tag{2}$$

where i, j are the site indices at positions $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$. $\mathcal{U}_{\zeta}(\hat{\boldsymbol{r}}_{ij})$ is the SK transformation matrix, which depends solely on the direction $\hat{\boldsymbol{r}}_{ij} = \boldsymbol{r}_{ij}/r_{ij}$ with $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$. Here, m refers to the magnetic quantum number. The parameter $h_{ll'\zeta}$ is the SKI corresponding to an ζ-type bond between two orbitals with angular momenta l and l'. Under the two-center approximation, SK formulation requires only $\min(l, l') + 1$ SKIs to describe all elements in $l - l'$ Hamiltonian matrix blocks with a shape of $(2l+1) \times (2l'+1)$. For example, as illustrated in Fig. 1(a), interactions involving s and p orbitals are described by only four types of SKIs ($h_{\mathrm{ss}\sigma}$, $h_{\mathrm{sp}\sigma}$, $h_{\mathrm{pp}\sigma}$, $h_{\mathrm{pp}\pi}$), which dramatically simplifies the parameter space. Consequently, only the $h_{\mathrm{sp}\sigma}$ term contributes to $H_{i,j}^{\mathrm{s,p}_z}$, as demonstrated by the SK transformation in Fig. 1(b). This formulation has been widely used since its

proposal in the 1950s.[28] Many SK-scheme-based parameterization strategies and parameter sets for various systems have emerged. A prominent example is the Navy Research Laboratory (NRL) TB model,[29,30] which fits not only the band structures but also the total energies. For more detailed discussion of the NRL method and its applications, see Ref. [28].

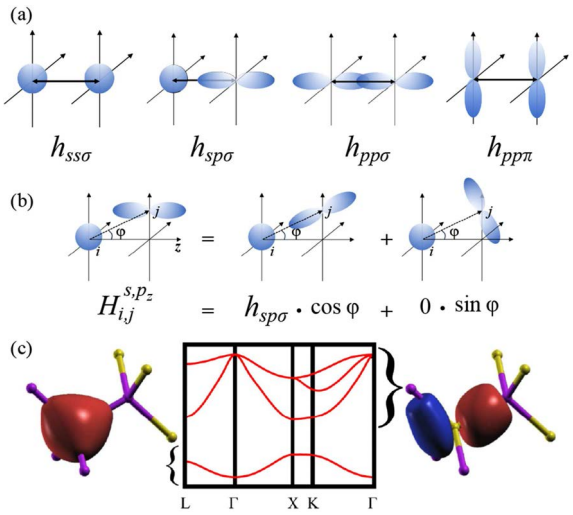


Fig. 1. Schematic illustration of the SK formalism and Wannierization. (a) The four types of SKIs involving s and p orbitals. (b) SK transformation for the coupling between s and p_z orbitals located at sites i and j. (c) Construction of Wannier functions from Bloch bands. Image adapted from Ref. [27].

Building upon the SK formulation, the density functional tight-binding (DFTB) method[19–21,31] was introduced as a systematic approximation to the DFT. DFTB models are formulated by expanding the DFT total energy functional E^{tot} around the reference density ρ_0 (superposition of neutral atomic densities) in a Taylor series

$$E^{\text{tot}}[\rho_0+\delta\rho] = E^0[\rho_0]+E^1[\rho_0,\delta\rho]+E^2[\rho_0,(\delta\rho)^2] + E^3[\rho_0,(\delta\rho)^3]+O((\delta\rho)^4), \quad (3)$$

where $\delta\rho$ is the density fluctuation with respect to ρ_0. Depending on the order of the Taylor expansion, different variants of DFTB have been proposed. DFTB1 (originally referred to as DFTB)[19,20] is a first-order, non-self-consistent model in which the Hamiltonian matrix elements are obtained from DFT calculations of isolated dimers as functions of interatomic distance. DFTB2[21] introduces a second-order correction $E^2[\rho_0,(\delta\rho)^2]$ by incorporating self-consistent charge effects. This allows for self-consistent treatment of $\delta\rho$ in systems. Furthermore, DFTB3[21,23–25] improves accuracy by incorporating the third-order term, which is important when local densities vary significantly. Since the 1990s, the DFTB method has established itself as a semi-empirical approach successfully improving the performance of the TB method across a broader range of materials systems.

Despite its theoretical elegance and improved accuracy, DFTB faces practical limitations due to two major drawbacks.[32] First, its pairwise parameterization scheme results in thousands of empirical parameters to be determined when covering even a portion of the periodic table. This makes the development of system-specific parameterizations time-consuming and requires expert knowledge, leading to fewer strategies succeeding in achieving broad elemental coverage.[33,34] Second, due to the use of a minimal basis set, not all desired properties can be described at a comparable accuracy. In the field of quantum chemistry, this problem has been, to some extent, circumvented by extended tight-binding (xTB) methods. A representative example is GFN-xTB,[22] the most widely used xTB method, originally designed to describe molecular properties in systems containing up to several thousand atoms. The acronym GFN reflects the methods' emphasis on accurately capturing geometries, vibrational frequencies, and non-covalent interactions.[32] Unlike DFTB, GFN-xTB primarily employs global and element-specific parameters, thereby minimizing the total number of required parameters. By fitting to reference data obtained from first-principles calculations, both GFN-xTB and its successor GFN2-xTB[26] have demonstrated robust accuracy across a broad range of chemical and biochemical applications. However, due to its underlying approximations, the accuracy of xTB methods tends to deteriorate for solids. In addition, GFN2-xTB calculations under periodic boundary conditions can become computationally demanding.

2.2. First-principles TB methods

Over the past decades, first-principles TB methods have been developed to systematically improve the accuracy and reliability of TB models. The central strategy involves projecting Bloch states obtained from first-principles calculations onto localized basis functions, such as Wannier functions[27] or quasi-atomic orbitals,[35,36] etc. Through projection, the TB parameters are directly obtained from the first-principles results, enabling an accurate reproduction of electronic band structures.

Among various first-principles TB schemes, one of the most widely adopted methodologies is the use of maximally localized Wannier functions (MLWFs).[27] MLWFs are obtained through a unitary transformation of Bloch states, typically selected from a low-energy window around the Fermi level, as illustrated in Fig. 1(c). This process is designed to minimize the spatial spread of the resulting orbitals, thereby ensuring the locality and orthogonality of Wannier functions. Moreover, an ecosystem of MLWF methods has flourished over the last two decades,[37] further advancing its development.

However, projection-based first-principles TB methods face several practical challenges. A key difficulty is the absence of a priori knowledge of the optimal minimal basis set, often requiring detailed analysis of the first-principles band

structure. In non-iterative projection approaches, such as direct projection[27] and quasi-atomic orbitals (QO),[35,36] the basis functions are predetermined, limiting the energy bands range that can be accurately reproduced. In contrast, iterative projection methods, including muffin-tin orbitals of arbitrary order (NMTO)[38] and MLWFs, require carefully choosing initial guesses for the basis functions and the energy window of the band to fit.[39] These steps require additional computational complexity and a certain level of expertise. Furthermore, as the first-principles TB Hamiltonian is optimized for specific reference structures, its transferability to unseen configurations is often limited.

In essence, although traditional tight-binding (TB) methods, from (semi-)empirical to first-principles, offer a significant reduction in computational complexity, they remain bound by the trade-off between efficiency and accuracy. While (semi-)empirical TB excels in speed and scalability, its reliability is limited. First-principles TB, on the other hand, achieves higher accuracy at the cost of increased complexity and user intervention. This fundamental compromise has become a critical bottleneck for modeling increasingly complex systems, motivating the recent exploration in machine learning approaches.

3. Machine learning TB methods

In recent years, ML methods have profoundly impacted various aspects of first-principles electronic structure simulations.[13,44,46–48] Within this broader context, MLTB approaches have emerged as a particularly active and promising direction. A growing number of studies have proposed diverse MLTB strategies, demonstrating great performance across different materials and device-level systems.[5,7–9,11,17,50]

To provide a clear overview of recent developments in MLTB methods, we categorize existing approaches based on their prediction targets and list them in Table 1. Specifically, we distinguish between models that directly predict the elements of the TB Hamiltonian matrix and those that infer empirical parameters used to construct the Hamiltonian within a specific parameterization scheme. Representative studies are introduced to highlight their methodological foundations and application domains. Lastly, we compare MLTB models with the capability to directly learn the Kohn–Sham (KS) Hamiltonian matrix elements from linear combination of atomic orbitals (LCAO)-based DFT packages, aiming to clarify their respective strengths and areas of applicability.

Table 1. Summarization of selected ML Hamiltonian approaches including MLTB and MLKS methods. The MLTB methods are categorized based on their prediction targets: TB Hamiltonian matrix elements or parameters in empirical TB formulas. All works are listed in chronological order of publication.

Category	Method/Authors	Year	Description
MLTB	TBHCNN[7]	2021	MLTB model trained on band structures, where neurons represent Hamiltonian elements.
	Schattauer *et al.*[8]	2022	A framework for defective systems by training a multilayer perceptron to map eigenvalues to TB elements.
	TBworks[9]	2022	Incorporating local structural information into training using eigenvalues from configurations along MD trajectories.
	DHNet[11,16]	2022	A divide-and-conquer approach to learn Wannier TB Hamiltonians from configurations along MD trajectories.
	WANDER[14]	2025	Employing layers in DeePMD model[40] to predict Wannier TB Hamiltonian elements.
	Fried *et al.*[15]	2025	Extension of Schattauers approach[8] by learning TB parameters from projected density of states instead of band structures.
	DFTB layer[5]	2018	Implements DFTB as differentiable tensor operations and learning TB parameters from structures to molecular properties.
	Nakhaee *et al.*[6]	2020	A linear model for the SKIs prediction in solids, which has been demonstrated in BiTeCl.
	DFTB-ML[10]	2022	A basis-function-optimization framework that learns the confinement radius in atomic orbital generation.
	Sun *et al.*[12]	2023	Optimization of SKIs via cubic spline interpolation using DOS as training labels for periodic systems.
	DeePTB[13]	2024	A deep-learning framework for onsite energies and SKIs with environment-dependent corrections beyond the two-center approximation.
MLKS	Hegde *et al.*[41]	2017	First attempt to use kernel ridge regression approach to fit DFT Hamiltonian of Cu system.
	SchNOrb[42]	2019	Extended SchNet,[43] the first equivariant neural network capable of representing high-order spherical tensors such as the Hamiltonian matrix.
	DeepH[44]	2022	A deep-learning approach for DFT Hamiltonian in crystalline materials with local coordinates based rotational equivariance.
	QHNet[45]	2023	A SE(3)-equivariant network for DFT Hamiltonian accelerated by significant reduction of the number of tensor products.
	DeepH-E3[46]	2023	An E(3)-equivariant deep-learning model for DFT Hamiltonian with the ability of handling SU(2) symmetry appears in the presence of spins.
	HamGNN[47]	2023	An E(3)-equivariant neural network for DFT Hamiltonian with explicit consideration for spin–orbital coupling.
	DeePTB-E3[48]	2025	An SE(3)-equivariant neural network for quantum operator matrices in DFT with the strictly localised equivariant message-passing.
	TraceGrad[49]	2025	A plug framework that enhances backbone models (e.g., DeepH-E3 and QHNet) by using a gradient-based mechanism to induce expressive SO(3)-equivariant features from learned SO(3)-invariant features.

3.1. TB Hamiltonian matrix element prediction

A straightforward way to integrate machine learning techniques into TB parameterization is to directly predict the on-site and hopping terms of the TB Hamiltonian in Eq. (1) using neural networks. These methods can be further categorized based on the type of training data used, either employing first-principles band structures as labels[7–9,15] or directly using first-principles TB Hamiltonian elements.[11,14,16]

Wang *et al.* proposed the TBHCNN workflow,[7] which learns the TB Hamiltonian elements from first-principles band structures and achieves higher accuracy than traditional empirical parameterizations, as illustrated in Figs. 2(a) and 2(b). For given systems, TBHCNN uses the neurons as on-site and hopping terms in TB Hamiltonian, taking the difference between predicted TB Hamiltonian's eigenvalues and the reference bands as a loss function to optimize these neurons. However, as mentioned in their paper, no information about the basis functions or atomic coordinates is used in TBHCNN. This results in a workflow that cannot ensure the symmetry characters and has no transferability to unseen structures.

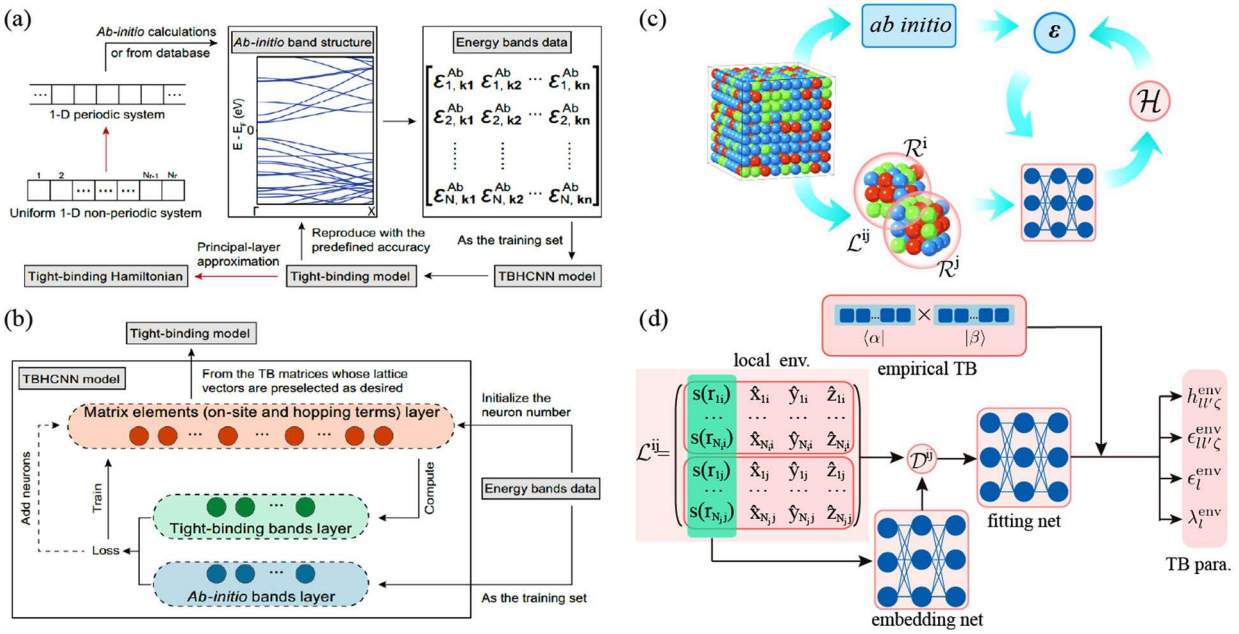


Fig. 2. Illustration of the typical workflows for two categories of MLTB methods. (a), (b) The workflow of the TBHCNN[7] that directly learns TB Hamiltonian matrix elements from band energies. (c), (d) the workflow of DeePTB approach[13] that learns environment-dependent SK TB parameters with local structural information as inputs.

In the approach of Schattauer*et al.*, the process begins with Wannierization of a pristine cell to obtain an initial TB Hamiltonian $H_{\mathrm{TB}}^{(0)}$. In the next step, a multilayer perceptron is then trained to map eigenvalues to TB parameters, enabling the extraction of TB elements for the target defective supercell based on its band structure. Although each network must be trained specifically for a given defect type, this approach offers a significantly faster alternative to Wannier-based parameterization especially for large systems. More recently, an extension to the work of Schattauer *et al.* was proposed.[15] In this ML framework for systems with point defects, the learning of TB parameters from the projected density of states (PDOS) was proposed. Since large supercells with defects often exhibit entangled bands and localized in-gap states, fitting the full band structure immediately becomes challenging. By using the PDOS instead, this issue can be naturally circumvented. However, these methods do not incorporate any structural information while learning the TB parameters. This results in non-transferability of the ML model in predicting the TB Hamiltonians for unseen structures. Up to some extent, the above-mentioned ML strategies can be viewed as an ML variant of Wannierization.

An earlier attempt towards incorporating local structural information as input in an MLTB framework, TBworks, was proposed in 2022 by Gu *et al.*[9] It learns TB Hamiltonians from first-principles band structures with local structural information as input. In TBworks, the Wannier Hamiltonians of a reference configuration (typically the equilibrium crystal structure) serve as the starting point for predicting Hamiltonians of unseen structures along MD trajectories. Due to the locality of Wannier functions, the hopping elements are non-negligible only within a finite range, making the local chemical environment around sites within a certain cutoff radius sufficient for electronic structure prediction. By this design, TBworks accurately predict electronic structures for configurations drawn from unseen temperatures and system sizes. The authors employed TBworks combined with ML force field molecular dynamics DeePMD[40] to calculate time-dependent spectral functions and optical conductivity in carbyne, which successfully captures the phenomena beyond the Born–Oppenheimer limit. TBworks achieves up to 5 orders of magnitude speedup.

An alternative pathway to obtain MLTB is to employ the first-principles TB Hamiltonian elements as training labels. Wang *et al.* proposed a divide-and-conquer strategy, termed the quasi-diabatic Hamiltonian network (DHNet).[11,16] The core idea is to partition the large structural configurations along the MD trajectories into smaller blocks and perform the Wannierization procedure for training labels on these small blocks after imposing periodic boundary conditions. This approach leverages the locality of electronic interactions and provides a scalable, accurate route for constructing first-principles TB Hamiltonians in large or disordered systems. Furthermore, the authors introduced an improved, well-defined descriptor within the DHNet framework that captures both distance and orientation information, preserving translational and rotational invariance.[16] DHNet achieved the accuracy at the Wannier TB level while gaining a significant speedup over the conventional DFT-Wannier workflow with 5 orders of magnitude.

More recently, in 2025, a "dual-functional" ML approach to predict Wannier TB Hamiltonian elements, WANDER,[14] was introduced, bridging the ML force fields and Wannier Hamiltonian predictions. In this study, the layers of the DeePMD model[40] are fixed and repurposed as input layers for predicting the elements of the Wannier Hamiltonian. Based on the classification of Hamiltonian elements, different internal layers of the DeePMD model are chosen for the on-site and intra/inter-layer hopping terms. A comparative analysis between the converged WANDER model and the conventional DFT-Wannier workflow reveals that WANDER achieves accuracy with a mean absolute error of 1.5 meV, while delivering an acceleration of 3 to 4 orders of magnitude. Given the preceding advancement of ML force fields approach, directly leveraging model architectures and training protocols from existing ML force fields offer a promising route to accelerate the progress of MLTB methods.

3.2. MLTB for empirical parameters prediction

Analogous to the development of empirical TB methods, it is natural to extend ML techniques to the empirical parameterization process, most notably within the SK framework. As discussed earlier, the SK formulation greatly simplifies the TB parameterization under the two-center approximation, reducing it to a small set of distance-dependent SKIs. Moreover, SK parameters are rotationally invariant (i.e., SO(3)-invariant), ensuring that the computed electronic properties remain independent of the choice of coordinate system. Several studies[6–8,15] have proposed SK-based variants using their own ML strategies.

Building on the SK framework, many MLTB works revolve around the DFTB framework to improve the parameterization process. Li *et al.*[5] proposed the DFTB layer, which revisits the DFTB procedure as differentiable tensor operations. The SKIs were represented using neurons and then fed to the DFTB layer to obtain the Hamiltonians, from which some of the target properties, such as energy, atomic charge and dipole moment, were calculated and optimized while comparing them with the corresponding labels. They employed two ways to represent SKIs using neurons: one using a spline model that relates SKIs to interatomic distances, and the second with a neural network that maps atomic geometry directly to SKIs. These designs enable fully differentiable end-to-end learning from atomic structures to target properties, thereby enabling data-driven refinement of DFTB parameters. This framework succeeded in providing a MLTB strategy for molecular systems, reducing the root-mean-square error in energy by 67% and in dipole moments by 59% compared to the initial DFTB parameters on 15700 hydrocarbons. In a similar study targeting periodic systems, the density of states (DOS) is used as the training label. The SKIs are expressed through cubic spline interpolation, and the corresponding coefficients are optimized accordingly.[12]

However, these models remain constrained by the inherent two-center approximation of the SK formalism. To overcome this limitation, recent studies have successfully incorporated information beyond the two-center approximation into the SK parameterization through different ML strategies. Fan *et al.* proposed a basis-function-optimization strategy, DFTB-ML.[10] Specifically, they focused on learning the confinement radius r_0 in the confinement potential $(r/r_0)^n$, which is added to the atomic Kohn–Sham equation when generating the basis functions. Three optimization schemes were introduced and compared: (1) spline model fitting, which directly fits the two-center integrals, (2) global optimization, where all atoms of a given element share the same optimized basis (consistent with standard DFTB parameterization), and (3) local optimization, where basis functions may differ across atoms to obtain the environment-dependent parameters. Given the freedom and sensitivity to the local environment, the local optimization approach excels in providing an accurate and flexible MLTB model. This entire workflow has been implemented in an open-source package, tight binding machine learning toolkit (TBMaLT).[50]

Gu *et al.* proposed a deep-learning approach named DeePTB[13] to optimize the on-site energies as well as SKIs, with incorporation of environment-dependent corrections beyond the two-center approximation, as illustrated in Figs. 2(c) and 2(d). In this method, the local environment-dependent SKIs, $h^{\mathrm{env}}_{ll'\zeta}$, are expressed as functions of symmetry-preserving local environment descriptors,

$$h^{\mathrm{env}}_{ll'\zeta} = h_{ll'\zeta}(r_{ij}) \times \left[1 + \Phi^{o_i,o_j}_{ll'\zeta}\left(r_{ij}, \mathcal{D}^{ij}\right)\right], \tag{4}$$

where $h_{ll'\zeta}$ is the empirical SKIs, r_{ij} is the bond length and $\mathcal{D}^{ij}$ denotes the local environment descriptor preserving the translational, permutational, and rotational symmetries. The map-

ping function $\Phi_{ll'\zeta}^{o_i,o_j}$ introduces environment-dependent corrections based on both bond length r_{ij} and descriptor $\mathcal{D}^{ij}$, where o_i, o_j are the chemical species of atoms i and j. By incorporating local environment dependence into SKIs, DeePTB accurately captures realistic underlying physics. Trained on structural datasets labeled with corresponding first-principles eigenvalues, DeePTB can accurately predict TB Hamiltonians for unseen structures. This provides flexibility to DeePTB of being independent of the choice of bases or exchange–correlation functionals. This method has been successfully validated on large-scale systems (up to 10^6 atoms) at finite temperatures, demonstrating excellent agreement with both experimental data and first-principles simulations.

3.3. MLTB vs. other ML-based electronic structure approaches

To fully appreciate the unique role of MLTB, it is helpful to place it within the broader landscape of ML for electronic structure simulation. These modern approaches can be broadly distinguished by their prediction targets: some aim to directly predict a final, macroscopic physical property, while others, use a physically inspired framework, learn a more fundamental, intermediate representation of the quantum system, from which a wide range of properties can be derived.

The first category of methods builds a direct mapping from atomic structure to some specific physical properties of interest, such as the density of states (DOS),[51,52] band gaps[53] and conductivity.[54,55] Despite the efficiency and powerful representative capacity from a proper chemical and material descriptor, these models are often highly specialized for certain downstream tasks. This also infers a huge demand of data consumption. Besides, their generality is often constrained, and predicting a new physical property typically requires the design and training of a new, specialized model. Recently, several works have tried to overcome these challenges by pretraining an atomic structure model with a designed supervised or unsupervised task, and fine-tuning the model on a downstream task dataset.[56] In comparison, MLTB is considered more versatile by learning the representation of the system's Hamiltonian. The Hamiltonian serves as a unified starting point, allowing for the subsequent calculation of a wide range of electronic properties without retraining the model. Alternatively, quantities other than the TB Hamiltonian can also act similarly as a fundamental physical quantity. Recent studies that contribute to their modeling include charge density,[57–59] density matrix,[48] wave function,[57] KS Hamiltonian[41,44–49] and self-energy.[60,61] These quantities are fundamentally linked to the underlying electronic Hamiltonian within their respective computational framework. Among this family of approaches, learning Hamiltonians has emerged as a particularly prominent and successful strategy, leading to two major but distinct branches of researches, i.e., directly learning TB and KS Hamiltonians.

The most direct comparison to MLTB lies in machine learning methods that learn the Kohn–Sham Hamiltonian (MLKS) from LCAO basis DFT calculations, such as Hegde's work,[41] SchNOrb,[42] DeepH,[44,46] QHNet,[45] HamGNN,[47] DeePTB-E3[48] and TraceGrad,[49] as summarized in Table 1. While both approaches predict a Hamiltonian matrix, their underlying nature and objectives differ significantly. MLTB learns a simplified, effective Hamiltonian represented in a minimal basis set in low-energy space, fundamentally performing a model reduction to gain computational efficiency. Furthermore, it is relatively straightforward to incorporate additional interaction terms to the TB Hamiltonian, such as the spin–orbit coupling effect,[62] magnetic fields[63] or Hubbard interactions[64,65] due to the flexibility of TB formalism. Additionally, MLTB methods based on orthogonal basis enable more efficient and flexible computational algorithms,[66] which are critical for large-scale simulations. In contrast, MLKS methods aim to accurately reproduce the full KS Hamiltonian in the original, much larger LCAO basis set from the parent DFT calculations. This distinction has profound consequences for post-processing: calculations using an MLTB Hamiltonian are computationally inexpensive due to its small size, whereas those using an MLKS Hamiltonian still involve "DFT-level" computational costs.

Ultimately, these two Hamiltonian-based strategies represent complementary philosophies. MLTB prioritizes the creation of a computationally efficient, scalable, and physically flexible effective model, making it exceptionally well-suited for large-scale simulations and the exploration of model physics through terms like the Hubbard U. MLKS methods, on the other hand, prioritize high-fidelity reproduction of a specific first-principles result, making them a powerful choice when accuracy against a reference DFT calculation is paramount. Together, these parallel lines of development offer a versatile suite of tools, tailored to different scientific goals and reflecting the growing sophistication of ML in electronic structure modeling.

4. The MLTB application ecosystem

The success of MLTB methods in rapidly and accurately generating electronic Hamiltonians marks a significant paradigm shift in computational physics. However, this achievement reframes the central challenge: the computational bottleneck, once focused on constructing the Hamiltonian matrix H, now migrates to the subsequent post-processing stage for physical observables. The crucial question becomes how to efficiently extract meaningful physical properties from H, especially for large-scale and complex systems that MLTB

is designed to handle. The answer lies not in a single, superior method, but in a diverse methodological and software ecosystem where the MLTB model acts as a powerful engine within an integrated workflow. The following exploration of this ecosystem reveals how distinct computational challenges have given rise to a suite of complementary and sophisticated solutions as shown in Fig. 3.

The most fundamental of these challenges is that of **scale**, a requirement common to frontier topics like surface science, heterostructures, moiré materials, and systems with sparse defects. The choice of solvers depends critically on both the system size and the specific physical property of interest. For small-to-medium-sized systems, direct diagonalization algorithms are commonly used, such as the QR method,[67–69] the Jacobi method,[70] and divide-and-conquer strategies.[71,72] Despite its $\mathcal{O}(N^3)$ complexity, these methods remain invaluable as they provide the complete set of eigenvalues and eigenvectors, which is indispensable when the full electronic spectrum is required. For larger systems where only spectral features like the band gap or states near the Fermi level are of interest, iterative solvers like the Lanczos algorithm[73,74] are highly effective. For sparse Hamiltonians, they require only $\mathcal{O}(N)$ memory and $\mathcal{O}(d \cdot N)$ computations per iteration (where d denotes the average row sparsity), efficiently targeting extremal eigenvalues. For the massive systems that MLTB can describe, linear-scaling methods are essential for calculating electronic properties. Techniques like the kernel polynomial method (KPM)[75–79] and the tight-binding propagation method (TBPM)[80–82] compute spectral quantities such as the density of states (DOS), optical conductivity, etc. with $\mathcal{O}(N)$ effort, bypassing the diagonalization process. A compelling demonstration is the simulation of a million-atom GaP supercell (with $\sim 10^7$ orbitals and a cell length of ~ 27.2 nm), where the DeePTB model was paired with the TBPLaS package[66] to compute finite-temperature properties at an unprecedented scale.

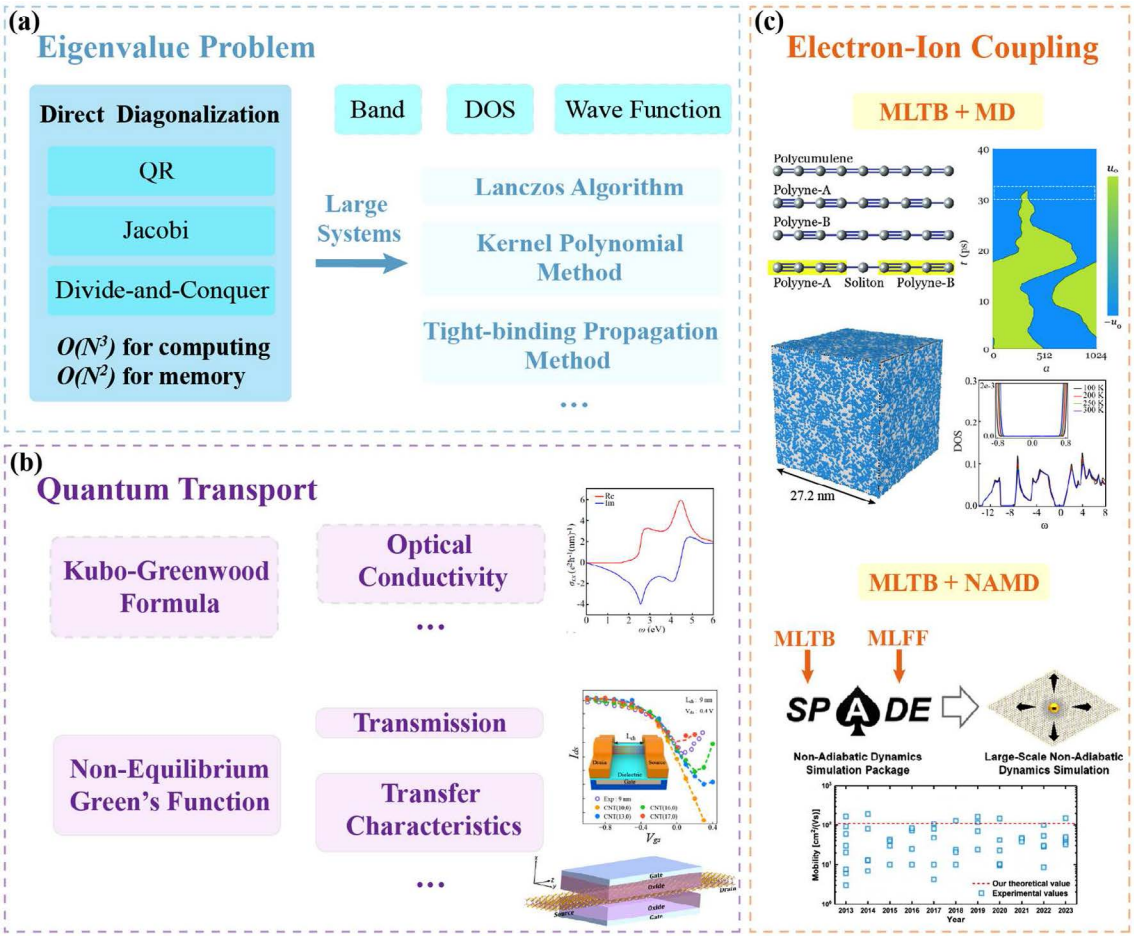


Fig. 3. The MLTB application ecosystem. (a) Methods for solving eigenvalue problems, including the QR, Jacobi, divide-and-conquer, Lanczos, kernel polynomial methods and tight-binding propagation methods, etc. (b) Methods for quantum transport simulations, including Kubo–Greenwood formula and non-equilibrium Green's function methods. The inset figures are from Res. [7,13,17]. (c) Methods for electron–ion coupling simulations including integration of MLTB methods with MD and NAMD approaches. The inset figures are from Refs. [9,13,16].

While the solvers above are powerful for static systems even at large scale, a distinct set of methodologies is required to address quantum transport.[17,66,83–90] The ecosystem for these simulations is primarily built upon two complementary cornerstones: the Kubo–Greenwood formula[91] and the non-equilibrium Green's function (NEGF) formalism.[92,93] The

former, grounded in the linear response theory, is ideal for calculating intrinsic transport coefficients such as conductivity in materials near equilibrium conditions. By combining the Kubo formula with the random state technique from TBPM,[66,80] Gu *et al.* employed the DeePTB-predicted TB Hamiltonian to simulate the optical conductivity of bulk GaP systems containing up to 10^6 atoms.[13] In contrast, the NEGF formalism is indispensable for device physics, as it is designed to model systems far from equilibrium, explicitly handling open boundary conditions with source and drain electrodes under finite bias. A more profound complexity arises in the field-effect transistor (FET) controlled by gate voltage. Here, the external field and the device's internal charge density mutually influence each other and hence the Hamiltonian, which requires a self-consistent treatment that typically couples the NEGF formalism with a Poisson solver.[93] The evolution of this research area reflects the growing sophistication, from early works coupling MLTB models with transport simulators for predicting transistor characteristics in sub-10-nm 2D materials based FETs,[7,94,95] to more recent frameworks like DeePTB-NEGF[17] investigating the device structures comparable to those in experiments. Using DeePTB-predicted Hamiltonians,[13,48] DeePTB-NEGF achieve a 2–3 order-of-magnitude speedup over conventional DFT-NEGF methods.[17] This enables high-fidelity simulations of complex systems under operational conditions, such as electron transport through over 10^4 atomic configurations in a break junction MD trajectory, and in experimentally-scaled carbon nanotube transistors with 41 nm channel lengths.[96,97]

The ecosystem further extends to capture the explicit interplay between electronic and nuclear degrees of freedom, which requires robust interfaces for structural samplings with computational approaches such as molecular dynamics (MD) or Monte Carlo (MC) simulations. This synergy is crucial as it recovers the electronic structure information that is inherently discarded by ML force fields and is powerful for studying the thermal vibrations and electron–phonon coupling influences on material properties. For instance, the TBworks[9] was combined with MD to investigate the soliton–antisoliton annihilation process in one-dimensional charge-density wave material carbyne, capturing the significant renormalization of low-energy modes due to the electron–lattice coupling beyond the Born–Oppenheimer approximation. More recently, the combination of DeePMD[40] and DeePTB[13] was used to calculate phonon-assisted optical absorption spectra in Si and GaAs[18] with the accuracy of HSE and SCAN XC functionals across temperature 100–400 K. The method accurately captures phonon-induced bandgap renormalization and indirect/direct absorption processes which are in excellent agreement with experimental findings over five orders of magnitude.[18] Beyond these adiabatic effects, MLTB can also be extended to interface with nonadiabatic molecular dynamics (NAMD) simulations. As demonstrated by Wang *et al.*,[11,16] the DHNet approach combined with the DeePMD supports the large-scale NAMD simulations based on the surface hopping method integrated in the `SPADE`[98] package.

From a broader perspective, the development of these distinct computational solutions paints a picture of a rapidly maturing field. However, a critical perspective reveals that significant challenges persist within each component of this ecosystem. For instance, while linear-scaling methods grant access to unprecedented system sizes, they are tailored for some specific spectral functions and cannot readily provide other physical quantities that have not been implemented. In device simulations, even with a computationally inexpensive Hamiltonian, the NEGF-SCF cycle itself, particularly involvement of the Green's function and Poisson equations, remains a bottleneck that limits the ultimate scale of transport calculations. Similarly, the reliance on approximate methods like surface hopping for nonadiabatic processes means that phenomena involving strong quantum decoherence may not yet be fully captured. In addition, underpinning all these applications is the persistent need for robust training data, making the development of active learning workflows tailored for Hamiltonian models a crucial ongoing challenge. Therefore, the future trajectory of this research field will be defined not only by the creation of more powerful MLTB models but also by the concerted effort to overcome these limitations, refining the entire ecosystem to achieve truly predictive power across all scales and phenomena.

5. Discussion and summary

The rapid advancement of MLTB methods has provided powerful tools for electronic structure simulations in large-scale and long-time simulations. This review has traced the evolution from traditional TB approaches, with their inherent trade-offs between accuracy and efficiency, to the rise of MLTB as a solution that combines the strengths of both. We systematically categorized the main branches of MLTB development and emphasized these methods within the broader landscape of AI-driven simulation. We contrast them with other ML-based electronic structure methods to highlight MLTB's distinct advantages in terms of physical versatility and computational efficiency. Ultimately, this exploration revealed that the true impact of MLTB is realized not by the models in isolation, but through their integration into a rich and diverse ecosystem in applications to material's property calculations.

Looking forward, despite the significant progress, the continued evolution of the field hinges on addressing several key frontiers. The first is the development of the model architecture itself. A primary goal is the construction of a universal and transferable MLTB model, inspired by the success

of equivariant neural networks, that offers excellent data efficiency and broad applicability across the periodic table. The second frontier lies in the generation of training data. To ensure the reliability required for scientific discovery, it is crucial to move beyond manual data curation and develop automated, intelligent pipelines. Designing robust active learning workflows tailored specifically for Hamiltonian models, rather than for force fields, remains a critical and open challenge. The third frontier is the continued refinement of the application ecosystem. Future work will focus on overcoming the current limitations of post-processing tools, such as developing highly-optimized solvers for MLTB to capture more complex quantum effects for large-scale simulations. As these challenges are met, MLTB is emerging as a powerful and versatile paradigm in computational science. By offering a compelling balance of accuracy, efficiency, and physical interpretability, it is breaking longstanding computational barriers. The ongoing advances across all these fronts are not only accelerating materials discovery and device design but are also opening new avenues for the fundamental understanding of complex quantum systems.

Acknowledgments

This work is supported by the Advanced Materials-National Science and Technology Major Project (Grant No. 2025ZD0618401) and the National Natural Science Foundation of China (Grant No. 12504285), and the Natural Science Foundation of Jiangsu Province (Grant No. BK20250472). SKP thankfully acknowledge NFSG grant from BITS-Pilani, Dubai campus.

References

[1] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[2] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[3] Harrison W 2012 *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond* Dover Books on Physics (Dover Publications)
[4] Slater J C and Koster G F 1954 *Phys. Rev.* **94** 1498
[5] Li H, Collins C, Tanha M, Gordon G J and Yaron D J 2018 *Journal of Chemical Theory and Computation* **14** 5764
[6] Nakhaee M, Ketabi S A and Peeters F M 2020 *J. Appl. Phys.* **128** 215107
[7] Wang Z, Ye S, Wang H, He J, Huang Q and Chang S 2021 *npj Computational Materials* **7** 11
[8] Schattauer C, Todorović M, Ghosh K, Rinke P and Libisch F 2022 *npj Computational Materials* **8** 116
[9] Gu Q, Zhang L and Feng J 2022 *Science Bulletin* **67** 29
[10] Fan G, McSloy A, Aradi B, Yam C Y and Frauenheim T 2022 *The Journal of Physical Chemistry Letters* **13** 10132
[11] Wang Z, Dong J, Qiu J and Wang L 2022 *ACS Applied Materials & Interfaces* **14** 22929
[12] Sun W, Fan G, van der Heide T, McSloy A, Frauenheim T and Aradi B 2023 *Journal of Chemical Theory and Computation* **19** 3877
[13] Gu Q, Zhouyin Z, Pandey S K, Zhang P, Zhang L and E W 2024 *Nat. Commun.* **15** 6772
[14] Qi Y, Gong W and Yan Q 2025 *npj Computational Materials* **11** 177
[15] Fried H P, Barragan-Yani D, Libisch F and Wirtz L 2025 *npj Computational Materials* **11** 176
[16] Cao B, Dong J, Wang Z and Wang L 2025 *The Journal of Physical Chemistry Letters* **16** 4907
[17] Zou J, Zhouyin Z, Lin D, Huang Y, Zhang L, Hou S and Gu Q 2025 *npj Comput. Mater.* **11** 375
[18] Gu Q, Pandey S K and Zhouyin Z 2025 Preprint 2502.00798
[19] Seifert G, Porezag D and Frauenheim T 1996 *International Journal of Quantum Chemistry* **58** 185
[20] Porezag D, Frauenheim T, Köhler T, Seifert G and Kaschner R 1995 *Phys. Rev. B* **51** 12947
[21] Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S and Seifert G 1998 *Phys. Rev. B* **58** 7260
[22] Grimme S, Bannwarth C and Shushkov P 2017 *Journal of Chemical Theory and Computation* **13** 1989
[23] Yang Y, Yu H, York D, Cui Q and Elstner M 2007 *The Journal of Physical Chemistry A* **111** 10861
[24] Gaus M, Cui Q and Elstner M 2011 *Journal of Chemical Theory and Computation* **7** 931
[25] Gaus M, Goez A and Elstner M 2013 *Journal of Chemical Theory and Computation* **9** 338
[26] Bannwarth C, Ehlert S and Grimme S 2019 *Journal of Chemical Theory and Computation* **15** 1652
[27] Marzari N, Mostofi A A, Yates J R, Souza I and Vanderbilt D 2012 *Rev. Mod. Phys.* **84** 1419
[28] Papaconstantopoulos D and Mehl M 2003 *J. Phys.: Condens. Matter* **15** R413
[29] Cohen R E, Mehl M J and Papaconstantopoulos D A 1994 *Phys. Rev. B* **50** 14694
[30] Mehl M J and Papaconstantopoulos D A 1996 *Phys. Rev. B* **54** 4519
[31] Hourahine B, Aradi B, Blum V, Bonafe F, Buccheri A, Camacho C, Cevallos C, Deshaye M, Dumitrică T, Dominguez A *et al.* 2020 *J. Chem. Phys.* **152** 124101
[32] Bannwarth C, Caldeweyher E, Ehlert S, Hansen A, Pracht P, Seibert J, Spicher S and Grimme S 2021 *Wiley Interdisciplinary Reviews: Computational Molecular Science* **11** e1493
[33] Wahiduzzaman M, Oliveira A F, Philipsen P, Zhechkov L, van Lenthe E, Witek H A and Heine T 2013 *Journal of Chemical Theory and Computation* **9** 4006
[34] Cui M, Reuter K and Margraf J T 2024 *Journal of Chemical Theory and Computation* **20** 5276
[35] Lu W C, Wang C Z, Chan T L, Ruedenberg K and Ho K M 2004 *Phys. Rev. B* **70** 041101
[36] Qian X, Li J, Qi L, Wang C Z, Chan T L, Yao Y X, Ho K M and Yip S 2008 *Phys. Rev. B* **78** 245112
[37] Marrazzo A, Beck S, Margine E R, Marzari N, Mostofi A A, Qiao J, Souza I, Tsirkin S S, Yates J R and Pizzi G 2024 *Rev. Mod. Phys.* **96** 045008
[38] Andersen O and Saha-Dasgupta T 2000 *Phys. Rev. B* **62** R16219
[39] Agapito L A, Ferretti A, Calzolari A, Curtarolo S and Buongiorno Nardelli M 2013 *Phys. Rev. B* **88** 165127
[40] Zhang L, Han J, Wang H, Car R and E W 2018 *Phys. Rev. Lett.* **120** 143001
[41] Hegde G and Bowen R C 2017 *Scientific Reports* **7** 42669
[42] Schütt K T, Gastegger M, Tkatchenko A, Müller K R and Maurer R J 2019 *Nat. Commun.* **10** 5024
[43] Schütt K T, Arbabzadah F, Chmiela S, Müller K R and Tkatchenko A 2017 *Nat. Commun.* **8** 13890
[44] Li H, Wang Z, Zou N, Ye M, Xu R, Gong X, Duan W and Xu Y 2022 *Nature Computational Science* **2** 367
[45] Yu H, Xu Z, Qian X, Qian X and Ji S 2023 *Proceedings of the 40th International Conference on Machine Learning (Proceedings of Machine Learning Research* vol. 202) ed Krause A, Brunskill E, Cho K, Engelhardt B, Sabato S and Scarlett J (PMLR) pp. 40412–40424
[46] Gong X, Li H, Zou N, Xu R, Duan W and Xu Y 2023 *Nat. Commun.* **14** 2848
[47] Zhong Y, Yu H, Su M, Gong X and Xiang H 2023 *npj Computational Materials* **9** 1
[48] Zhouyin Z H, Gan Z X, Pandey S K, Zhang L F and Gu Q Q 2025 *Learning local equivariant representations for quantum operators* in *The Thirteenth International Conference on Learning Representations*
[49] Yin S, Pan X Y, Wang F Y and He L X 2025 *Tracegrad: a framework learning expressive SO(3)-equivariant non-linear representations for electronic-structure hamiltonian prediction* in *Proceedings of the 42nd International Conference on Machine Learning (ICML)*

[50] McSloy A, Fan G, Sun W, Hölzer C, Friede M, Ehlert S, Schütte N E, Grimme S, Frauenheim T and Aradi B 2023 *The Journal of Chemical Physics* **158** 034801

[51] Fung V, Ganesh P and Sumpter B G 2022 *Chemistry of Materials* **34** 4848

[52] Kong S, Ricci F, Guevarra D, Neaton J B, Gomes C P and Gregoire J M 2022 *Nat. Commun.* **13** 949

[53] Zhuo Y, Mansouri Tehrani A and Brgoch J 2018 *The Journal of Physical Chemistry Letters* **9** 1668

[54] Bürkle M, Perera U, Gimbert F, Nakamura H, Kawata M and Asai Y 2021 *Phys. Rev. Lett.* **126** 177701

[55] Lin D, Zou J, Dong Y, Wang Y, Wang Y, Sanvito S and Hou S 2025 *Phys. Chem. Chem. Phys.* **27** 7728

[56] Zhou G, Gao Z, Ding Q, Zheng H, Xu H, Wei Z, Zhang L and Ke G 2023 *Uni-mol: A universal 3d molecular representation learning framework* in *The Eleventh International Conference on Learning Representations*

[57] Unke O, Bogojeski M, Gastegger M, Geiger M, Smidt T and Müller K R 2021 *Advances in Neural Information Processing Systems* Vol. 34 ed. Ranzato M, Beygelzimer A, Dauphin Y, Liang P and Vaughan J W (Curran Associates, Inc.) pp. 14434–14447

[58] Lv T, Zhong Z, Liang Y, Li F, Huang J and Zheng R 2023 *Phys. Rev. B* **108** 235159

[59] Song F and Feng J 2024 Preprint 2406.15873

[60] Dong X, Gull E and Wang L 2024 *Phys. Rev. B* **109** 075112

[61] Venturella C, Li J, Hillenbrand C, Leyva Peralta X, Liu J and Zhu T 2025 *Nature Computational Science* **5** 502

[62] Gu Q, Pandey S K and Tiwari R 2023 *Computational Materials Science* **221** 112090

[63] Vonsovsky S V and Katsnelson M I 1989 *Quantum solid-state physics* (Berlin: Springer-Verlag)

[64] Hubbard J and Flowers B H 1963 *Proc. Roy. Soc. Lond. Ser. A Math. Phys. Sci.* **276** 238

[65] Qin M, Schöfer T, Andergassen S, Corboz P and Gull E 2022 *Annual Review of Condensed Matter Physics* **13** 275

[66] Li Y, Zhan Z, Kuang X, Li Y and Yuan S 2023 *Computer Physics Communications* **285** 108632

[67] Francis J G 1961 *The Computer Journal* **4** 265

[68] Francis J G F 1962 *The Computer Journal* **4** 332

[69] Kublanovskaya V 1962 *USSR Computational Mathematics and Mathematical Physics* **1** 637

[70] Rutishauser H 1966 *Numerische Mathematik* **9** 1

[71] Cuppen J J M 1980 *Numerische Mathematik* **36** 177

[72] Demmel J W 1997 *Applied numerical linear algebra* (SIAM)

[73] Lanczos C 1950 *Journal of Research of the National Bureau of Standards* **45** 255

[74] Cullum J K and Willoughby R A 2002 *Lanczos algorithms for large symmetric eigenvalue computations: Vol. I: Theory* (SIAM)

[75] Röder H and Silver R N 1995 *Densities of States and Thermodynamics of Mega-Dimensional Sparse Matrices* (Boston: Springer) pp. 301–308

[76] Wang L W 1994 *Phys. Rev. B* **49** 10154

[77] Wang L W and Zunger A 1994 *Phys. Rev. Lett.* **73** 1039

[78] Silver R, Roeder H, Voter A and Kress J 1996 *Journal of Computational Physics* **124** 115

[79] Weiße A, Wellein G, Alvermann A and Fehske H 2006 *Rev. Mod. Phys.* **78** 275

[80] Yuan S, De Raedt H and Katsnelson M I 2010 *Phys. Rev. B* **82** 115448

[81] Hams A and De Raedt H 2000 *Phys. Rev. E* **62** 4365

[82] Yuan S, Roldán R and Katsnelson M I 2011 *Phys. Rev. B* **84** 035439

[83] Fiori G, Iannaccone G and Klimeck G 2006 *IEEE Transactions on Electron Devices* **53** 1782

[84] Marian D, Marin E G, Perucchini M, Iannaccone G and Fiori G 2023 *Journal of Computational Electronics* **22** 1327

[85] Luisier M and Klimeck G 2008 *2008 8th IEEE Conference on Nanotechnology* pp. 354–357

[86] Pecchia A, Penazzi G, Salvucci L and Carlo A D 2008 *New J. Phys.* **10** 065022

[87] Steiger S, Povolotskyi M, Park H H, Kubis T and Klimeck G 2011 *IEEE Transactions on Nanotechnology* **10** 1464

[88] Fan Z, Uppstu A, Siro T and Harju A 2014 *Computer Physics Communications* **185** 28

[89] Groth C W, Wimmer M, Akhmerov A R and Waintal X 2014 *New J. Phys.* **16** 063065

[90] João S M, Anđelković M, Covaci L, Rappoport T G, Lopes J M V P and Ferreira A 2020 *Royal Society Open Science* **7** 191809

[91] Kubo R 1957 *J. Phys. Soc. Jpn.* **12** 570

[92] Keldysh L V *et al.* 1965 *Sov. Phys. JETP* **20** 1018

[93] Datta S 1997 *Electronic Transport in Mesoscopic Systems* (Cambridge University Press)

[94] Yang N and Guo J 2024 *Solid-State Electronics* **213** 108859

[95] Spinazze M and Yoon Y 2025 *IEEE Transactions on Electron Devices* **72** 3287

[96] Franklin A D and Chen Z 2010 *Nat. Nanotechnol.* **5** 858

[97] Franklin A D, Luisier M, Han S J, Tulevski G, Breslin C M, Gignac L, Lundstrom M S and Haensch W 2012 *Nano Lett.* **12** 758

[98] Dong J, Qiu J, Bai X, Wang Z, Xiao B and Wang L 2025 *Journal of Chemical Theory and Computation* **21** 3300