Deep Learning Accelerated Quantum Transport Simulations in Nanoelectronics: From Break Junctions to Field-Effect Transistors

Jijie Zou [2,3], Zhanghao Zhouyin,[2,4] Dongying Lin,[5] Linfeng Zhang,[2,6] Shimin Hou,[3,5,*] and Qiangqiang Gu[1,2,†]

[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]*Centre for Nanoscale Science and Technology, Academy for Advanced*
Interdisciplinary Studies, Peking University, Beijing 100871, China
[4]*Department of Physics, McGill University, Montreal, Quebec, Canada H3A2T8*
[5]*Key Laboratory for the Physics and Chemistry of Nanodevices,*
School of Electronics, Peking University, Beijing 100871, China
[6]*DP Technology, Beijing 100080, China*

Quantum transport calculations are essential for understanding and designing nanoelectronic devices, yet the trade-off between accuracy and computational efficiency has long limited their practical applications. We present a general framework that combines the deep learning tight-binding Hamiltonian (DeePTB) approach with the non-equilibrium Green's Function (NEGF) method, enabling efficient quantum transport calculations while maintaining first-principles accuracy. We demonstrate the capabilities of the DeePTB-NEGF framework through two representative applications: comprehensive simulation of break junction systems, where conductance histograms show good agreement with experimental measurements in both metallic contact and single-molecule junction cases; and simulation of carbon nanotube field effect transistors through self-consistent NEGF-Poisson calculations, capturing essential physics including the electrostatic potential and transfer characteristic curves under finite bias conditions. This framework bridges the gap between first-principles accuracy and computational efficiency, providing a powerful tool for high-throughput quantum transport simulations across different scales in nanoelectronics.

I. INTRODUCTION

Quantum transport simulations are essential for understanding and designing nanoelectronic devices. The non-equilibrium Green's function (NEGF) method[1–3], combined with density functional theory (DFT)[4, 5], has established itself as the standard approach for investigating quantum transport phenomena at the first-principles level[6, 7]. However, the computational burden of DFT-NEGF self-consistent field (SCF) iterations severely constrains its practical applications. This limitation becomes particularly critical in studying large-scale systems and dynamic processes, in which extensive sampling of configuration space is required. A prime example is the break junction experiments[8, 9], where quantum transport through nanoscale contacts evolves from metallic contacts to single-molecule junctions. Thousands of conductance measurements throughout this process are essential for statistical analysis, making first-principles calculations computationally prohibitive.

Machine learning (ML) approaches have emerged as promising alternatives to accelerate quantum transport calculations. There are several ML attempts including direct prediction of conductance from atomic structures using local descriptors[10], and mapping combined structural and electronic features to transmission spectra near the Fermi level[11]. While these approaches demonstrate

the potential of ML in transport calculations, they are often limited to specific systems or single transport property, lacking the generality required for broad applications in nanoelectronics. A systematic and comprehensive framework to accelerate DFT-NEGF simulations and capable of predicting various transport properties - from transmission spectra to potential profiles and current-voltage characteristics - remains elusive.

The key to addressing this challenge lies in bypassing the time-consuming SCF iterations while maintaining first-principles accuracy. Our recently developed deep learning represented tight-binding (TB) Hamiltonian method, DeePTB[12, 13], provides an efficient route to accurate electronic structure calculations. By combining deep neural networks with physics-based principles, DeePTB achieves remarkable accuracy in predicting both local environment-dependent Slater-Koster (SK) tight-binding Hamiltonians[12] and DFT Hamiltonians, overlap and density matrix in the linear combination of atomic orbitals (LCAO) basis[13]. Building on these capabilities, we present a general DeePTB-NEGF framework for efficient quantum transport calculations. This framework enables direct prediction of TB models or DFT Hamiltonians as well as the overlap matrices, eliminating the need for SCF iterations while maintaining first-principles accuracy.

We demonstrate the versatility and accuracy of our approach through two representative applications: (1) break junction systems, where we successfully simulate the dynamic process in metallic contacts and single-molecule junctions, showing good agreement

with both DFT-NEGF calculations and experimental measurements[14]; and (2) carbon nanotube field-effect transistors (CNT-FETs), where we implement Poisson calculations in DeePTB-NEGF framework to study transport characteristics under finite bias, validating our results against the established NanoTCAD ViDES solver[15, 16]. These applications demonstrate that our framework can efficiently handle quantum transport across different scales and materials systems, opening new possibilities for high-throughput quantum transport calculations in nanoelectronics.

II. METHODOLOGY AND WORKFLOW

The key to efficient quantum transport calculations lies in obtaining accurate electronic structure information while avoiding expensive SCF iterations. Our approach combines the NEGF formalism for transport calculations with the DeePTB method for efficient Hamiltonian prediction, as illustrated in Fig. 1. DeePTB offers two complementary strategies for constructing electronic Hamiltonians, ie. the environment-dependent SK tight-binding Hamiltonians[12] and DFT Hamiltonians in the LCAO basis with E(3) equivariant graph neural networks[13].

For the SK tight-binding Hamiltonians, DeePTB trains local environment-dependent SK integrals $h_{ll'\zeta}^{\text{env}}$ as[12]:

$$h_{ll'\zeta}^{\text{env}} = h_{ll'\zeta}(r_{ij}) \times \left[1 + \Phi_{ll'\zeta}^{o_i,o_j}\left(r_{ij}, \mathcal{D}^{ij}\right)\right] \quad (1)$$

where $h_{ll'\zeta}(r_{ij})$ is the conventional SK integral, $\Phi_{ll'\zeta}^{o_i,o_j}$ introduces neural network-based environmental corrections through the descriptor $\mathcal{D}^{ij}$. Here, l, l' denote atomic orbitals, ζ the bond types, and r_{ij} the interatomic distance. By training on the DFT eigenvalues, the SK integrals as functions of the local environment can be predicted for unseen structures.

For the LCAO DFT Hamiltonians, DeePTB predicts multiple quantum operators (Hamiltonian, overlap and density matrices) through E(3) equivariant graph neural networks[13], preserving rotational symmetry through angular momentum coupling:

$$O_{l_1,l_2,m_1,m_2}^{i,j} = \sum_{l_3,m_3} C_{(l_1,m_1)(l_2,m_2)}^{(l_3,m_3)} o_{l_3,m_3}^{i,j} \quad (2)$$

where O represents the quantum operators (Hamiltonian, overlap and density matrices), $C_{(l_1,m_1)(l_2,m_2)}^{(l_3,m_3)}$ are the Clebsch-Gordan coefficients and $o_{l_3,m_3}^{i,j}$ are the output node or edge features from the equivariant graph neural networks. Both strategies provide efficient access to electronic Hamiltonians while maintaining first-principles accuracy, though based on different theoretical frameworks.

With the Hamiltonians efficiently predicted by DeePTB, quantum transport properties are calculated using the NEGF formalism. The NEGF approach provides a rigorous framework for quantum transport by
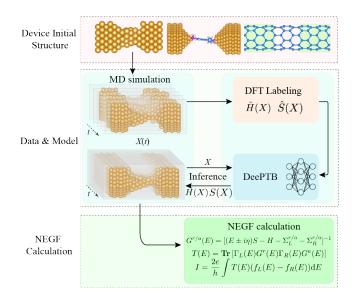

FIG. 1. Schematic illustration of the DeePTB-NEGF workflow. Atomic configurations $X(t)$ are generated from molecular dynamics simulations. Reference Hamiltonians $\hat{\boldsymbol{H}}(X)$ and overlap matrices $\hat{\boldsymbol{S}}(X)$ are obtained from DFT-NEGF calculations, while $\boldsymbol{H}(X)$ and $\boldsymbol{S}(X)$ represent predictions from the DeePTB model. $T(E)$ denotes the transmission coefficient at energy E and I the current through the two-terminal device.

incorporating open boundary conditions through self-energies $\boldsymbol{\Sigma}$ that describe the coupling between the device and semi-infinite electrodes. For a two-terminal system, the transmission spectrum $T(E)$ between the left and right electrodes is given by:

$$T(E) = \mathbf{Tr}\left[\Gamma_L(E)G^r(E)\Gamma_R(E)G^a(E)\right] \quad (3)$$

where $\Gamma_{L/R} = i(\boldsymbol{\Sigma}_{L/R}^r - \boldsymbol{\Sigma}_{L/R}^a)$ is the broadening function for the left (L) and right (R) electrodes and $G^{r/a}(E) = [(E \pm i\eta)S - H - \boldsymbol{\Sigma}_L^{r/a} - \boldsymbol{\Sigma}_R^{r/a}]^{-1}$ is the retarded/advanced Green's function with η being a positive infinitesimal and H/S being the Hamiltonian/overlap matrices. The current (I) and zero-bias conductance (G) at the zero-temperature limit are then given by[8]:

$$I = \frac{2e}{h} \int T(E)(f_L(E) - f_R(E))\mathrm{d}E \quad (4)$$

$$G = \frac{2e^2}{h} T(E_f) = T(E_f)\mathrm{G}_0 \quad (5)$$

where $f_{L/R}(E)$ are Fermi-Dirac distributions for the left and right electrodes, E_f is the Fermi energy, and $\mathrm{G}_0 = 2e^2/h$ is the quantum conductance, e is the elementary charge, and h is the Planck's constant.

Building upon the efficient Hamiltonian prediction by DeePTB, we further optimize the NEGF calculations through addressing its two most time-consuming operations: the calculation of electrode self-energies and matrix inversion for Green's functions. For the electrode self-energy calculations, we exploit the fact that in na-

noelectronics, particularly in molecular electronics, electrodes can often be considered periodic in the directions perpendicular to the current. By applying Bloch's theorem, we calculate the self-energy within the smallest electrode unit cell[17], significantly reducing both memory requirements and computational costs. For the Green's function calculations, we utilize the sparse nature of the Hamiltonian and overlap matrices in the local basis representation, which leads to block tridiagonal matrix configurations in quasi-one-dimensional systems[18]. We implement a greedy algorithm with recursive processing[19] to optimize the matrix structure, which substantially enhances the efficiency of the recursive Green's function calculations[20].

The complete workflow is illustrated in Fig. 1. Molecular dynamics (MD) simulations are used to generate atomic configurations $X(t)$ at finite temperatures along the trajectories at different time t, which are then divided into training and prediction sets. For the training set, configurations are labeled with converged Hamiltonians $\hat{\boldsymbol{H}}(X)$ and overlap matrices $\hat{\boldsymbol{S}}(X)$ from DFT-NEGF calculations using TranSIESTA[6, 17, 21]. The trained DeePTB model then efficiently predicts Hamiltonians $\boldsymbol{H}(X)$ and overlap matrices $\boldsymbol{S}(X)$ for new configurations, enabling high-throughput quantum transport calculations through the optimized NEGF framework.

III. RESULTS

A. Break Junction Systems

Break junction experiments, primarily implemented through mechanically controllable break junctions (MCBJ) and scanning tunneling microscopy break junctions (STM-BJ)[8, 9], provide a powerful platform for investigating quantum transport at the atomic and molecular scales. During a typical break junction process, the configuration evolves from the metallic contacts to the molecular junctions. In the metallic contacts, conductance exhibits characteristic quantization features reflecting the discrete nature of atomic contacts. As the junction breaks, individual molecules can bridge the nanoscale gap forming a stable molecular junction[22]. Statistical analysis of these transport features requires thousands of conductance measurements over repeated breaking events, making first-principles calculations computationally prohibitive. The DeePTB-NEGF framework developed here enables efficient simulation of the break junction process while maintaining the first-principles accuracy. Below we demonstrate its capabilities in the two cases of metallic contacts and molecular junctions, providing direct comparison with experimental measurements and atomic-level insights into the quantum transport mechanisms.

1. Quantum Transport in Metallic Contacts

Gold atomic contacts formed in break junction experiments exhibit distinct quantum transport features such as conductance quantization and ballistic transport[23]. During electrode separation, the conductance evolution reveals characteristic plateaus, reflecting the formation of atomic-scale contact configurations[24]. Understanding these quantum transport features requires statistical analysis over thousands of breaking events, making theoretical reproduction of conductance histograms at the first-principles level computationally challenging.

To simulate quantum transport through gold contacts, we construct a junction model consisting of a central scattering region connected to two semi-infinite electrodes, as shown in the top panel of Fig. 2(a). The scattering region contains 304 gold atoms arranged along the [100] direction, with two bulk-like sections (100 atoms each in $5 \times 5 \times 4$ supercells) serving as electrode extensions. To simulate the break junction formation, we performed molecular dynamics simulations using the DeePMD potential[25] at 150K in the NVT ensemble, with the simulation box being the scattering region under periodic boundary conditions. During elongation, the electrode sections maintain their bulk structure while moving in opposite directions, with stretching speeds ranging from 0.2 to 5.0 m/s. The structural evolution exhibits characteristic bipyramidal shapes consistent with experimental high-resolution transmission electron microscopy observations of gold contacts under stretching [26].

For constructing the DeePTB model for gold contacts, we randomly sampled 122 configurations from 4 independent trajectories at $v = 5$ m/s and calculated their electronic Hamiltonians and overlaps using DFT-NEGF with PBE functional and SZP basis in TranSIESTA. The model's cutoff radius for the local environment of gold atom r_{cut}^{Au} was set to 7.4Å to match the maximum radius of gold atomic orbitals in DFT calculations. Although the training configurations only include the scattering region, periodic boundary conditions ensure that atoms near the simulation box edges experience a bulk-like environment within the cutoff radius as shown in the lower panel of Fig. 2(a), crucial for accurate self-energy calculations in the NEGF framework. Details of data preparation and model training are provided in the supplementary materials (SM).

We then validate the model's predictive capability for electronic properties. Despite being trained only on scattering regions, the model successfully predicts Hamiltonian elements for larger systems shown in the top panel of Fig. 2(a) with additional electrode layers, exhibiting excellent agreement with DFT results (RMSE $= 4.3 \times 10^{-4}$ eV for Hamiltonian elements, see SM for details). More importantly, we examine the model's generalization ability across different elongation speeds. Although trained solely on configurations from trajectories with $v = 5.0$ m/s, the model demonstrates excellent transferability.
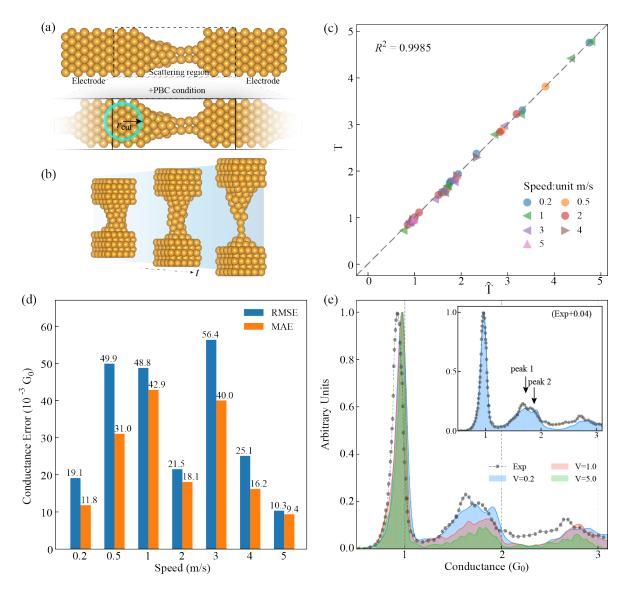
FIG. 2. DeePTB-NEGF simulation results for gold contacts. (a) Atomic configuration of gold contacts. Top panel: configuration used for transport calculation consisting of a central scattering region connected to two semi-infinite electrodes. Bottom panel: configuration used for DeePTB model training with periodic boundary conditions (PBC). The local environment inside the cutoff radius for the atoms near the simulation box edges is bulk-like. (b) Three representative snapshots in one breaking junction process. (c) Comparison for transmission at the Fermi level between the DeePTB-NEGF method (T) and DFT-NEGF calculations($\hat{T}$) in structures obtained with different elongation speeds. (d) RMSE and MAE for zero-bias conductance between DeePTB-NEGF and DFT-NEGF results. (e) Conductance histograms from DeePTB-NEGF simulations at different elongation speeds ($v = 0.2$, 1.0, and 5.0 m/s) compared with experimental data. Inset: Experimental $1G_0$ peak aligns with the $v = 0.2$ m/s case after shifting the experimental histogram by $+0.04G_0$.

By analyzing 55 randomly selected configurations spanning speeds from 0.2 m/s to 5.0 m/s, we compare zero-bias conductances predicted by DeePTB-NEGF with DFT-NEGF results. The model achieves remarkable accuracy with $R^2 = 0.9985$ in the parity plot and maintains low prediction errors (RMSE and MAE $< 60 \times 10^{-3} G_0$) across all elongation speeds, as shown in Fig. 2(c) and (d).

Having validated the model's accuracy, we then applied the model to calculate the gold contact conduc-

tance and demonstrate its efficiency in high-throughput calculations. Using the DeePTB model trained only on 122 configurations, we successfully generated conductance histograms by analyzing 9224 configurations sampled from 105 elongation processes at speeds of 0.2, 1.0, and 5.0 m/s. The calculations were accelerated using the Bloch theorem for self-energy evaluation and exploiting the tridiagonal block structure of Hamiltonian matrices. The histograms, normalized to their respective $1G_0$ peaks, are presented in Fig.2(e). The prominent peak at

$1G_0$ corresponds to monatomic chains, a well-established signature in metallic break junctions.

Different elongation speeds lead to distinct variations in conductance peak shapes, particularly in the range from $1G_0$ to $2G_0$. This variation arises from the suppression of collective atomic relaxations at higher elongation speeds, which prolongs the stability of monatomic chains before breaking[10, 27]. Consequently, the relative intensity of conductance features beyond $1G_0$ is systematically reduced in the normalized histograms. Note that the simulation speeds (0.2-5.0 m/s) are necessarily higher than experimental rates (10 pm/s to 100 nm/s)[8], as the experimental breaking processes would require prohibitively large MD simulation times[26]. Nevertheless, even at these elevated speeds, our simulated conductance histogram at $v = 0.2$ m/s shows good agreement with experimental measurements, despite small shifts in peak positions. These shifts can be attributed to the fixed perfect lattice structure imposed on the electrode boundary atoms in the simulation, which neglects the experimental structural perturbation and minor defects. By shifting the experimental histogram by $+0.04G_0$, we achieve excellent alignment of the $1G_0$ peaks as shown in the inset of Fig. 2(e). The aligned histograms reveal remarkable agreement between calculation and experiment [26], particularly in reproducing two characteristic conductance peaks between $1G_0$ and $2G_0$. These results demonstrate that our DeePTB-NEGF framework can efficiently and accurately predict the quantum transport properties of gold contacts, capturing the statistical features of the conductance histograms consistent with the experimentally observed characteristics.

2. Quantum Transport through Single-molecule Junctions

Single-molecule junctions present additional challenges for quantum transport calculations compared to metallic contacts due to their complex chemical environments and dynamic metal-molecule interfaces. Their transmission spectra are particularly sensitive to molecular conformations and contact geometries, requiring extensive sampling to capture the dynamic features during junction formation and breaking in the break junction experiments. While DFT-NEGF calculations can accurately predict transmission spectra for static configurations, their computational cost has restricted most studies to selected relaxed structures, missing crucial dynamic aspects of the transport process. We demonstrate that our DeePTB-NEGF framework efficiently addresses this challenge by enabling high-throughput transmission spectra calculations throughout the entire junction evolution.

The molecule investigated here is a classical π-conjugated system with sulfur methyl as anchoring groups, which contains rich chemical environments and flexible configurations due to three types of carbon bonds and the dihedral angle between the two benzene planes (top left inset of Fig. 3(c)). The scattering region consists of the molecule connected to gold electrodes with a bi-pyramidal shape, where each electrode principal layer contains 144 atoms ($6 \times 6 \times 4$ supercell). To simulate the breaking junction process, we performed MD simulations at 300K using the ReaxFF reactive force field[28], which can efficiently describe chemical bonding through its empirical bond-order formalism. During stretching, the gold electrodes move in opposite directions at a speed of $v = \pm 2$ m/s, with three representative snapshots at t_1, t_2, t_3 in equal time intervals shown in Fig. 3(a). To construct the DeePTB model for molecular junctions, we randomly sampled 268 configurations from 14 breaking junction trajectories as the training set and calculated their electronic Hamiltonians and overlaps using DFT-NEGF with PBE functional and SZP basis in TranSIESTA. Similarly, the model's cutoff radius was set as the largest radius of atomic orbitals for each element in DFT-NEGF calculations, ensuring accurate description of electronic interactions. More details of data preparation and model training are provided in the SM.

To understand how the training set size affects model performance, we trained a series of models using varying numbers of configurations (30, 60, 90, 123, 188, 208, 268) randomly sampled from the 14 breaking junction trajectories. The prediction accuracy was evaluated using the Spearman correlation coefficient between DeePTB-NEGF and DFT-NEGF transmission spectra, which measures how well two variables can be described by a monotonic function[29]. The coefficients are evaluated in the energy range of -0.5 to 0.5 eV with a step of 0.01 eV, utilizing a validation set of 60 configurations from different trajectories to test out-of-distribution generalization. As shown in Fig. 3(c), both the mean value and standard deviation of the correlation coefficient improve with increasing training set size, reaching a high correlation of 0.912 with 268 configurations. The gradually reduced improvement rate suggests that our model has achieved sufficient accuracy in capturing the structural complexity present in the breaking junction process.

To further validate the statistical assessment of model accuracy indicated by the Spearman correlation coefficient of 0.912, we examine the performance of our model trained with 268 configurations by comparing transmission spectra for three representative configurations captured at times t_1, t_2, and t_3 during the breaking junction process shown in Fig.3(a). As illustrated in Fig.3(b), the DeePTB-NEGF predicted transmission spectra closely match the DFT-NEGF results across different junction configurations, accurately reproducing both the positions and heights of transmission peaks. Analysis of these transmission spectra further reveals distinct behavior of frontier molecular orbital contributions: the HOMO-dominated transmission peak maintains its energetic position relative to the Fermi level, while the LUMO-dominated peak shifts toward the Fermi level during junction elongation. This characteristic evolution of transmission peaks indicates that the conductance is LUMO-dominated, which is consistent with previous
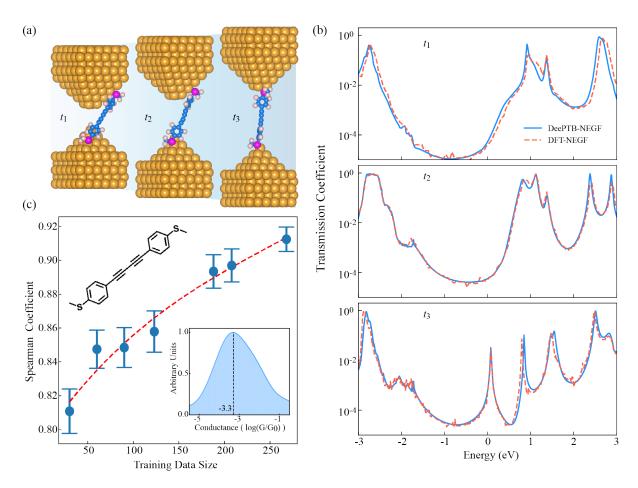
FIG. 3. DeePTB-NEGF simulation results for molecular junctions. (a) Three sequential snapshots of the molecular junction breaking process at equal time intervals (t_1, t_2, and t_3). (b) Transmission spectra corresponding to the three snapshots in (a). (c) Spearman correlation coefficient of transmission spectra calculated for validation set configurations across varying training data sizes, within the energy range of -0.5 to 0.5 eV. Blue dots represent mean values with error bars showing standard deviations. Top left inset: Molecular structure in the junction. Bottom right inset: Conductance histogram computed from 590 configurations across 11 stretching processes using DeePTB-NEGF, showing the highest peak at $10^{-3.3}G_0$.

experimental and theoretical studies[14]. Furthermore, to enable statistical comparison with experimental measurements, we calculated the zero-bias conductance for 590 configurations sampled from 11 breaking junction trajectories. The resulting conductance histogram (bottom right inset of Fig. 3(c)) shows a pronounced peak at $10^{-3.3}G_0$, in good agreement with the experimental value of $10^{-3.6}G_0$[14].

In summary, we have demonstrated the applicability and accuracy of the DeePTB-NEGF approach in simulating quantum transport properties in the two cases of metallic atomic contacts and single-molecule junctions. Our method successfully captures the key features of the conductance evolution during the junction-breaking process and reproduces the main characteristics of the experimental conductance histograms. The good agreement between the DeePTB-NEGF and DFT-NEGF results further validates its accuracy in describing the electronic structure of the junctions. These results establish DeePTB-NEGF as a powerful and efficient method for

studying quantum transport in break junction systems, bridging the gap between first-principles accuracy and statistical analysis.

B. Double-gated Carbon Nanotube Field-effect Transistor

Having demonstrated the capability of DeePTB-NEGF in simulating quantum transport in break junction systems, we now extend our framework to semiconductor devices where the interplay between quantum transport and electrostatic effects becomes crucial. Here we focus on carbon nanotube field-effect transistors (CNT-FETs), which are promising candidates for sub-5 nm technology nodes due to their excellent carrier transport properties[30, 31]. The quasi-one-dimensional structure and long mean-free path of CNT enable near-ballistic transport at nanoscale dimensions, making them ideal channel materials for high-performance electronic
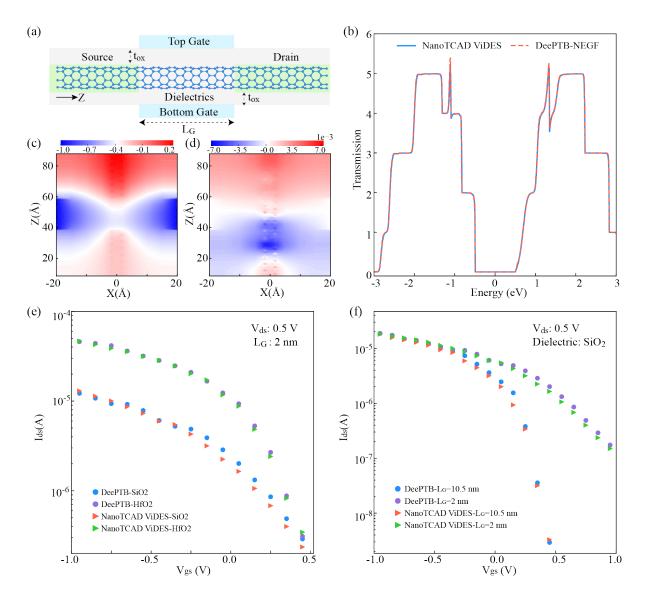
FIG. 4. DeePTB-NEGF-Poisson SCF simulation results for double-gated carbon nanotube field-effect transistor. (a) Schematic of the transistor geometry, where L_G denotes gate length and t_{ox} represents gate dielectric thickness. (b) Transmission spectrum at gate voltage $V_{gs} = 1$ V and drain-source bias $V_{ds} = 0.5$ V. (c) Electrostatic potential energy distribution in the xz plane. (d) Difference in electrostatic potential energy between DeePTB-NEGF and NanoTCAD ViDES calculations in the xz plane, using identical device geometry and boundary conditions ($L_G = 2$ nm, $V_{gs} = 1$ V). (e) Transfer characteristics comparing DeePTB-NEGF and NanoTCAD ViDES results for SiO$_2$ and HfO$_2$ dielectrics with $L_G = 2$ nm and $V_{ds} = 0.5$ V. (f) Transfer characteristics comparing DeePTB-NEGF and NanoTCAD ViDES results for different gate lengths ($L_G = 2$ nm and 10.5 nm) with SiO$_2$ dielectric and $V_{ds} = 0.5$ V.

devices[32]. The realistic simulation of CNT-FETs requires addressing the electrostatic effects on quantum transport under finite gate voltage and drain-source bias, which can be achieved through the NEGF-Poisson approach[33, 34].

While DFT-NEGF methods have been successful in quantum transport calculations, they become computationally prohibitive for experimentally relevant device dimensions. The TB method offers a more efficient alternative by employing smaller and sparser Hamiltonian matrices. Here we implement the DeePTB model which en-

hances traditional TB approaches by incorporating local environment-dependent corrections to the SK parameters (Eq.1). This approach goes beyond the conventional two-center approximation while maintaining the computational efficiency of TB methods[12]. Combined with the NEGF-Poisson formalism, our method enables efficient self-consistent quantum transport calculations for realistic devices under finite bias conditions.

The self-consistent calculation begins with the equilibrium TB Hamiltonian H_0 and iteratively solves for the non-equilibrium potential profile using the Poisson equa-

tion [15, 35]:

$$\nabla[\epsilon(\vec{r})\nabla\phi(\vec{r})] = -e[\rho_{\text{free}}(\vec{r}) + \rho_{\text{fix}}(\vec{r})] \tag{6}$$

Here, ϵ is the dielectric constant, ϕ is the electrostatic potential, ρ_{free} is the free charge concentration, and ρ_{fix} represents the fixed charge from doping. Specifically, ρ_{fix} corresponds to the average density of donated or accepted electrons per atomic site in the doped region. The NEGF-Poisson self-consistency is achieved through the Gummel iteration scheme [36], with the Poisson equation solved using the Newton-Raphson method on a discrete real-space grid. Within our TB-NEGF framework, the spatial charge distribution is approximated by Mulliken charges localized at atomic sites.

To demonstrate the capability of our implementation, we study quantum transport in a double-gated CNT-FET, as illustrated in Fig. 4(a). The device consists of a carbon nanotube channel with p-doped source and drain regions ($\rho_{\text{fix}} = 5 \times 10^{-3}$ per atomic site), sandwiched between top and bottom gates. The gates encompass the entire channel length L_{G} and are separated from the CNT by dielectric layers of thickness $t_{\text{ox}} = 1.7$ nm. The electronic structure is described using the DeePTB model implemented with the conventional Slater-Koster parameters ($h_{ll'\zeta}(r_{ij})$ term in Eq. 1) without environmental corrections, to enable direct comparison with established TB-NEGF codes. A simplified π-bond model is adopted with one p_z orbital per carbon atom and hopping energy $t = 1.09$ eV. The electrostatic boundary conditions are set as follows: Dirichlet conditions at the gate regions (fixed to gate voltage), von Neumann conditions (zero electric field) at the exposed dielectric surfaces, and flat potential profiles with von Neumann conditions at the edges of the semi-infinite source and drain regions. The dielectric environment is controlled through the spatial distribution of the dielectric constant $\epsilon(\vec{r})$.

We first validate our NEGF-Poisson implementation by comparison with the established TB-NEGF code NanoTCAD ViDES[15, 16]. Under identical simulation conditions ($L_{\text{G}} = 2$ nm, $V_{\text{gs}} = 1$ V, and $V_{\text{ds}} = 0.5$ V), both the transmission spectra (Fig.4(b)) and electrostatic potential energy distribution (Fig.4(c)) show excellent agreement. The difference in potential energy distribution (Fig.4(d)) shows a maximum deviation of less than 7×10^{-3} eV, confirming the accuracy of our self-consistent implementation in the DeePTB-NEGF framework.

Having validated our implementation, we investigate the device transfer characteristics under different operating conditions, focusing on the effects of dielectric materials and gate lengths. Fig. 4(e) compares the transfer characteristics with two different gate dielectric materials, SiO$_2$ and HfO$_2$, at $V_{\text{ds}} = 0.5$ V and $L_{\text{G}} = 2$ nm. For direct comparison, the gate work functions are adjusted to achieve the same off-current $I_{\text{OFF}} = 2.5 \times 10^{-7}$ A at $V_{\text{gs}} = 0.45$ V. When switching from SiO$_2$ to HfO$_2$(dielectric constant changing from 3.9 to 20 in gate dielectric area), the subthreshold swing (SS), which characterizes the inverse slope of the trans-

fer characteristics curve on a semi-logarithmic scale, decreases from 696.56 mV dec^{-1} to 213.72 mV dec^{-1}. This improvement aligns with the fact that high-κ gate dielectrics reduce the effective screening length of the channel $\lambda_{\text{ch}} \propto \sqrt{(\epsilon_{\text{ch}}/\epsilon_{\text{ox}})t_{\text{ox}}t_{\text{ch}}}$ where ϵ_{ch} and ϵ_{ox} are the dielectric constants for channel and gate dielectric, and t_{ch} and t_{ox} represent their respective thicknesses[37]. The higher ϵ_{ox} of HfO$_2$ leads to shorter effective screening length, indicating stronger gate electrostatic control over the channel. The effect of gate length is investigated using SiO$_2$ as the dielectric material (Fig. 4(f)). Increasing the gate length from 2 nm to 10.5 nm improves the SS from 449.47 mV dec^{-1} to 92.12 mV dec^{-1}. This significant improvement occurs because longer gates provide better electrostatic control with larger gate capacitance.

The excellent agreement between our implementation and NanoTCAD ViDES validates our NEGF-Poisson self-consistent framework. The physically reasonable response to different dielectric materials and gate lengths further demonstrates its capability to capture essential device physics. These results establish DeePTB-NEGF as a versatile approach for quantum transport simulations in realistic semiconductor devices, combining the efficiency of TB methods with the accuracy needed for device design and optimization under finite bias conditions.

IV. SUMMARY AND DISCUSSION

We have developed a deep learning-accelerated framework, DeePTB-NEGF, for efficient quantum transport calculations in nanoelectronic devices. By integrating deep learning Hamiltonian prediction with the NEGF method, our approach enables both direct quantum transport calculations and self-consistent solutions incorporating electrostatic effects. Most importantly, this framework achieves comprehensive simulation of the dynamic break junction process, extending to realistic semiconductor device simulations through NEGF-Poisson self-consistency. This versatility, combined with the method's efficiency and accuracy, bridges the gap between atomic-scale quantum transport simulations to statistical analysis of experimental measurements and practical device applications.

The broad applicability and accuracy of this framework have been rigorously validated through two representative applications. In break junction systems, we demonstrate the capability of our method to comprehensively simulate the junction evolution in both the metallic contacts and single-molecule junction systems. Our predicted conductance histograms show good agreement with experimental measurements in both cases, reproducing characteristic features such as the quantized conductance peaks between 1.0G_0 and 2.0G_0 for metallic contacts and the pronounced peak at $10^{-3.3}\text{G}_0$ for molecular junctions. Building upon this success in break junction systems, the framework's extension to semiconduc-

tor devices is validated through carbon nanotube FET simulations, where our NEGF-Poisson implementation accurately captures device electrostatics and transport characteristics under various conditions, as confirmed by comparison with established quantum transport codes.

Based on these achievements, the demonstrated performance of DeePTB-NEGF establishes a new paradigm for high-throughput quantum transport calculations in nanoelectronics. The framework's dual capability in simulating break junction processes and transistors offers opportunities to advance both molecular electronics research and practical device engineering. Further development of this methodology could incorporate more sophisticated quantum effects and extend to larger-scale systems, potentially enabling quantum transport simulations at industrially relevant scales. The successful integration of deep learning efficiency with quantum mechanical accuracy presents a significant step toward bridging fundamental physics and technological applications in nanoelectronics.

CODE AVAILABILITY

The DeePTB-NEGF code is openly available in the github deepmodeling community at https://github.com/deepmodeling/DeePTB. For data preparation, we also developed and released an open-source tool dftio.

ACKNOWLEDGMENTS

[1] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics: Green's Function Methods in Equilibrium and Non-Equilibrium Problems* (New York: W.A. Benjamin, 1962).

[2] L. V. Keldysh *et al.*, Diagram technique for nonequilibrium processes, Sov. Phys. JETP **20**, 1018 (1965).

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

[4] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, Phys. Rev. **136**, B864 (1964).

[5] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. **140**, A1133 (1965).

[6] M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, Density-functional method for nonequilibrium electron transport, Phys. Rev. B **65**, 165401 (2002).

[7] E. Louis, J. A. Vergés, J. J. Palacios, A. J. Pérez-Jiménez, and E. SanFabián, Implementing the Keldysh formalism into *ab initio* methods for the calculation of quantum transport: Application to metallic nanocontacts, Phys. Rev. B **67**, 155321 (2003).

[8] E. Scheer and J. C. Cuevas, *Molecular electronics: an introduction to theory and experiment*, Vol. 15 (World Scientific, 2017).

[9] P. Gehring, J. M. Thijssen, and H. S. van der Zant, Single-molecule quantum-transport phenomena in break junctions, Nature Reviews Physics **1**, 381 (2019).

[10] M. Bürkle, U. Perera, F. Gimbert, H. Nakamura, M. Kawata, and Y. Asai, Deep-learning approach to first-principles transport simulations, Physical Review Letters **126**, 177701 (2021).

[11] R. Topolnicki, R. Kucharczyk, and W. Kaminski, Combining multiscale MD simulations and machine learning methods to study electronic transport in molecular junctions at finite temperatures, The Journal of Physical Chemistry C **125**, 19961 (2021).

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

[13] Z. Zhouyin, Z. Gan, S. K. Pandey, L. Zhang, and Q. Gu, Learning local equivariant representations for quantum operators (2024), arXiv:2407.06053 [cond-mat.mtrl-sci].

[14] M.-M. Guo, Y. Jiang, J.-Y. Wang, Z.-N. Chen, S. Hou, and Q.-C. Zhang, Effectively enhancing the conductance of asymmetric molecular wires by aligning the energy level and symmetrizing the coupling, Langmuir **40**, 3759 (2024).

[15] G. Fiori, G. Iannaccone, and G. Klimeck, A three-dimensional simulation study of the performance of carbon nanotube field-effect transistors with doped reservoirs and realistic geometry, IEEE Transactions on electron devices **53**, 1782 (2006).

[16] D. Marian, E. G. Marin, M. Perucchini, G. Iannaccone, and G. Fiori, Multi-scale simulations of two dimensional material based devices: the nanotcad vides suite, Journal of Computational Electronics **22**, 1327 (2023).

[17] N. Papior, N. Lorente, T. Frederiksen, A. García, and M. Brandbyge, Improvements on non-equilibrium and transport Green function techniques: The next-generation transiesta, Computer Physics Communications **212**, 8 (2017).

[18] A. Pecchia, G. Penazzi, L. Salvucci, and A. D. Carlo, Non-equilibrium green's functions in density functional tight binding: Method and applications, New Journal of Physics **10**, 065022 (2008).

[19] M. Klymenko, J. Vaitkus, J. Smith, and J. Cole, Nanonet: An extendable python framework for semi-empirical tight-binding models, Computer Physics Communications **259**, 107676 (2021).

[20] M. Anantram, M. S. Lundstrom, and D. E. Nikonov, Modeling of nanoscale devices, Proceedings of the IEEE **96**, 1511 (2008).

[21] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, The SIESTA

method for *ab initio* order-n materials simulation, Journal of Physics: Condensed Matter **14**, 2745 (2002).

[22] B. Xu and N. J. Tao, Measurement of single-molecule resistance by repeated formation of molecular junctions, science **301**, 1221 (2003).

[23] Y. Yu, F. Cui, J. Sun, and P. Yang, Atomic structure of ultrathin gold nanowires, Nano letters **16**, 3078 (2016).

[24] L. Rego, A. Rocha, V. Rodrigues, and D. Ugarte, Role of structural evolution in the quantum conductance behavior of gold nanowires during stretching, Physical Review B **67**, 045412 (2003).

[25] C. M. Andolina, M. Bon, D. Passerone, and W. A. Saidi, Robust, multi-length-scale, machine learning potential for Ag–Au bimetallic alloys from clusters to bulk materials, The Journal of Physical Chemistry C **125**, 17438 (2021).

[26] M. J. Lagos, F. Sato, P. A. S. Autreto, D. S. Galvão, V. Rodrigues, and D. Ugarte, Temperature effects on the atomic arrangement and conductance of atomic-size gold nanowires generated by mechanical stretching, Nanotechnology **21**, 485702 (2010).

[27] T. N. Todorov and A. P. Sutton, Force and conductance jumps in atomic-scale metallic contacts, Phys. Rev. B **54**, R14234 (1996).

[28] T. P. Senftle, S. Hong, M. M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. M. Aktulga, *et al.*, The ReaxFF reactive force-field: development, applications and future directions, npj Computational Materials **2**, 1 (2016).

[29] C. Spearman, The proof and measurement of association between two things, The American journal of psychology **100**, 441 (1987).

[30] S. Iijima, Helical microtubules of graphitic carbon, Nature **354**, 56 (1991).

[31] G. Pitner, Z. Zhang, Q. Lin, S.-K. Su, C. Gilardi, C. Kuo, H. Kashyap, T. Weiss, Z. Yu, T.-A. Chao, *et al.*, Sub-0.5 nm interfacial dielectric enables superior electrostatics: 65 mv/dec top-gated carbon nanotube fets at 15 nm gate length, in *2020 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2020) pp. 3–5.

[32] L.-M. Peng, Z. Zhang, and S. Wang, Carbon nanotube electronics: recent advances, Materials today **17**, 433 (2014).

[33] A. Sanchez-Soares, T. Kelly, G. Fagas, J. C. Greer, and E. Chen, Top-gated carbon nanotube fets from quantum simulations: Comparison with experiments, in *2021 International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA)* (IEEE, 2021) pp. 1–2.

[34] J. Guo, A. Javey, H. Dai, and M. Lundstrom, Performance analysis and design optimization of near ballistic carbon nanotube field-effect transistors, in *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004.* (IEEE, 2004) pp. 703–706.

[35] P.-H. Ahn and S.-M. Hong, A fully coupled scheme for a self-consistent Poisson-NEGF solver, IEEE Transactions on Electron Devices **70**, 239 (2022).

[36] H. K. Gummel, A self-consistent iterative scheme for one-dimensional steady state transistor calculations, IEEE Transactions on electron devices **11**, 455 (1964).

[37] F. Pikus and K. Likharev, Nanoscale field-effect transistors: An ultimate size analysis, Applied Physics Letters **71**, 3661 (1997).