






Machine learned Hückel theory: Interfacing physics and deep neural networks

Cite as: J. Chem. Phys. **154**, 244108 (2021); <https://doi.org/10.1063/5.0052857>

Submitted: 01 April 2021 . Accepted: 10 June 2021 . Published Online: 25 June 2021

 Tetiana Zubatiuk,  Benjamin Nebgen, Nicholas Lubbers, Justin S. Smith, Roman Zubatyuk,  Guoqing Zhou, Christopher Koh,  Kipton Barros,  Olexandr Isayev, Sergei Tretiak, et al.

COLLECTIONS

Paper published as part of the special topic on [Computational Materials Discovery](#)



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Perspective on integrating machine learning into computational chemistry and materials science](#)

The Journal of Chemical Physics **154**, 230903 (2021); <https://doi.org/10.1063/5.0047760>

[Chemical physics software](#)

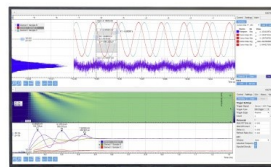
The Journal of Chemical Physics **155**, 010401 (2021); <https://doi.org/10.1063/5.0059886>

[Machine learning meets chemical physics](#)

The Journal of Chemical Physics **154**, 160401 (2021); <https://doi.org/10.1063/5.0051418>

Challenge us.

What are your needs for
periodic signal detection?



Zurich
Instruments

Machine learned Hückel theory: Interfacing physics and deep neural networks

Cite as: J. Chem. Phys. 154, 244108 (2021); doi: 10.1063/5.0052857

Submitted: 1 April 2021 • Accepted: 10 June 2021 •

Published Online: 25 June 2021



View Online



Export Citation



CrossMark

Tetiana Zubatiuk,¹ Benjamin Nebgen,^{2,4} Nicholas Lubbers,^{3,5} Justin S. Smith,^{2,3} Roman Zubatyuk,¹ Guoqing Zhou,⁶ Christopher Koh,² Kipton Barros,² Olexandr Isayev,^{1,a)} and Sergei Tretiak^{2,3,4}

AFFILIATIONS

¹Department of Chemistry, Mellon College of Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

²Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87544, USA

³Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

⁴Center for Integrated Nanotechnologies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

⁵Computer, Computational, and Statistical Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87544, USA

⁶Department of Physics and Astronomy, University of Southern California, Los Angeles, California 90089, USA

Note: This paper is part of the JCP Special Topic on Computational Materials Discovery.

^{a)}Author to whom correspondence should be addressed: olexandr@olexandrisayev.com

ABSTRACT

The Hückel Hamiltonian is an incredibly simple tight-binding model known for its ability to capture qualitative physics phenomena arising from electron interactions in molecules and materials. Part of its simplicity arises from using only two types of empirically fit physics-motivated parameters: the first describes the orbital energies on each atom and the second describes electronic interactions and bonding between atoms. By replacing these empirical parameters with machine-learned dynamic values, we vastly increase the accuracy of the extended Hückel model. The dynamic values are generated with a deep neural network, which is trained to reproduce orbital energies and densities derived from density functional theory. The resulting model retains interpretability, while the deep neural network parameterization is smooth and accurate and reproduces insightful features of the original empirical parameterization. Overall, this work shows the promise of utilizing machine learning to formulate simple, accurate, and dynamically parameterized physics models.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0052857>

I. INTRODUCTION

“The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws lead to equations too complicated to be soluble.”⁵⁹

The combination of quantum mechanics (QM) and machine learning (ML) is an emerging research field in the interdisciplinary areas of physics, mathematics, chemistry, and materials science. The rapidly growing pool of ML algorithms (see reviews^{1–3} and references therein) along with a robust deep learning software ecosystem (TensorFlow,⁴ PyTorch,⁵ etc.) suggests that we are approaching the next key milestone in the evolution of

computational chemistry. The development of neural-network (NN) ML interatomic potentials (MLIPs) and their successful application to high-dimensional systems have been demonstrated by many research groups.^{6–14} Among the diverse subjects in ML, robust and transferable prediction of the molecular electronic structure is now one of the most intensively studied directions, making ML methods more accessible for the solution of the Schrödinger equation (SE).¹⁵ Modern computational methods of quantum chemistry mainly deal with the molecular orbital (MO) theory first formulated by Hückel in 1930s.^{16–18} Hückel revealed an importance of a very basic locality in molecular and/or atomistic systems, with his insights now being implemented in almost every QM computational method. In particular, he suggested that the QM model built only on a subgroup of functions (in that particular case—wave functions of π -orbitals), which are responsible for the QM properties of the substance omitting the rest of the system components (as σ -electrons), has an

ability to predict those properties quite efficiently. From the mathematical point of view, such a suggestion is equivalent to the separation of variables in the one-electron Schrödinger equation (SE). It results in the restriction of the form of the wave function of *all* electrons to the antisymmetrized product of the functions of electron groups containing fewer electrons. For example, in semi-empirical (SQM) and pseudopotential methods, the wave functions of atomic core elements are not considered. In the *ab initio* methods, the aforementioned approximation is implemented implicitly due to the employment of a finite set of basis functions (the interactions of the rest of atomic orbital functions are omitted). From the viewpoint of practical applications of ML methods in QM, the idea of the locality of chemical properties in atomistic systems is of significant importance. The implementation of ML algorithms is mainly based on the fact that QM properties can be learned by splitting them into features of individual atoms with their local environment within some cutoff.^{7,19} Modern deep NN models automatically learn important features of the local atomic environment (AE) from local atomic vectors. Featurization of AE is general enough and does not require manual engineering of task-specific features to achieve a state-of-the-art performance in predictions of QM properties.^{1,2} Additionally, the training time scaling becomes polynomial against the number of datapoints (in contrast to the exponential time scaling of QM methods), making it practical to learn from datasets of millions of QM data.

By training MLIPs to approximate QM results, it is now possible to reproduce high-level QM properties at a rate of microseconds per atom and get very accurate results in specific applications. With that, the quantity, quality, and types of interactions included in the training dataset dictate the accuracy of MLIPs. As mentioned by Dral,³ MLIPs trained to reproduce properties of neutral molecules may not be appropriate for the description of charged molecules and radicals, and once trained for ground-state energies, they will fail predicting excitation energies. Although the advanced MLIPs are entirely data-driven, they do not necessarily capture the underlying physics and chemistry of atoms in molecules. As discussed by Manzhos,¹⁵ the highly accurate solutions and acceptability for QM applications will be likely achieved by combining generic properties

of ML with modifications specific to the problem of solving the SE or building functionals, including the task-specific training dataset and adapted architectures of NN.

Basically, integrating ML with QM can be achieved in multiple ways (Fig. 1). In series of works,^{20–22} the NN directly predicts wave functions based on QM data. A fundamentally different way is to use ML algorithms to improve the electronic Hamiltonian. In this light, Dral²³ introduced a hybrid ML-SQM automatic parameterization technique (APT). It is based on ML models of parameters as a function of molecular structure. The authors used APT for predicting molecule-specific corrections to the QM parameters that improved accuracy of semiempirical QM (SQM) calculations. Li *et al.*²⁴ used NN and spline-based ML models to tune matrix elements of the self-consistent-charge tight-binding (DFTB) Hamiltonian. With the hybrid ML-DFTB APT model, they succeeded to substantially reduce prediction errors on energy and dipole moments for small hydrocarbons. We suggest, however, that APT could be used to build a general-purpose and transferable ML-SQM model for electronic structure prediction across vast chemical space with the accuracy of *ab initio* methods. Hence, we here introduce a ML-based semi-empirical model based on the extended Hückel method (ML-EHM), where ML is used as an APT to determine the on-the-fly optimal matrix elements of the one-electron tight-binding Hamiltonian as a function of local AE.

In the present article, we start with overviewing a new generation of computational methods based on ML and QM interfaces. The EHM serves us as a starting point, and we further extend it considering the complications arising in the ML domain. In this proof-of-concept study, our first objective is to determine the ML-EHM accuracy and transferability of the ML Hamiltonian. We detail the ML-EHM architecture and training procedure in Sec. II. In Sec. III, we benchmark datasets of organic molecules containing four atomic species {H, C, N, and O}. In Sec. IV, we demonstrate how ML-EHM retains the conceptual strength of QM by performing covalent bond extension in the methane molecule. The second case study examines the behavior of the Hamiltonian eigenvalues upon internal rotation around single bonds in butadiene and aza-butadiene molecules.

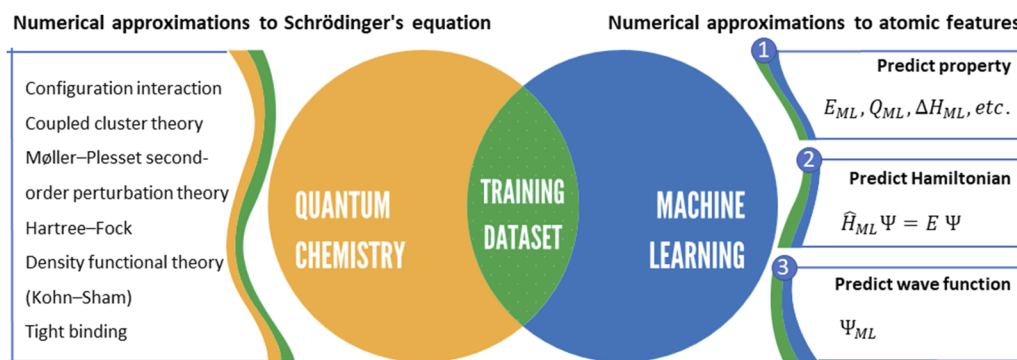


FIG. 1. A typical use of ML in quantum chemistry. The left block denotes QM calculations used to generate training datasets. Right block: (1) ML can be used to predict QM properties (energies, charges, enthalpies, etc.). (2) The Hamiltonian itself can be parameterized by ML. (3) ML can be used to predict the wave function.

II. ML-EHM FRAMEWORK

A. The extended Hückel method

Similar to the Hückel method,^{16–18} the EHM^{25,26} calculations employ a simplistic tight-binding form of the Hamiltonian, taken as a sum of single-electron interactions. The linear combination of atomic orbitals (LCAOs) $\Psi_i = \sum_j c_{ij} \phi_j$ identifies the molecular orbital (MO) without an anti-symmetrization operation. The basis set contains only the atomic-like orbitals for the valence shell. According to EHM, the SE was solved in a non-orthonormal basis of AO, thus requiring a generalized eigensolver algorithm to address the

following matrix diagonalization problem:

$$\mathbf{HC} = \varepsilon \mathbf{SC}, \quad (1)$$

where \mathbf{C} is the matrix of coefficients for the AO. Each column in matrix \mathbf{C} defines one MO in terms of AO basis functions. In the EHM, the overlap is not neglected and \mathbf{S} is the matrix of overlap integrals between the AO basis functions $|\chi_\mu\rangle$. ε is the diagonal matrix of orbital energies, E_{MO} . \mathbf{H} is a square matrix containing the one electron energy integrals. Specifically, the definition of the EHM Hamiltonian \mathbf{H} is given by

$$\mathbf{H} = \begin{bmatrix} \alpha_{11} & \frac{1}{2}K^\ddagger(\alpha_{22} + \alpha_{11})S_{21} & \vdots & \frac{1}{2}K^\ddagger(\alpha_{NN} + \alpha_{11})S_{N1} \\ \frac{1}{2}K^\ddagger(\alpha_{11} + \alpha_{22})S_{12} & \alpha_{22} & \vdots & \frac{1}{2}K^\ddagger(\alpha_{NN} + \alpha_{22})S_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}K^\ddagger(\alpha_{11} + \alpha_{NN})S_{1N} & \frac{1}{2}K^\ddagger(\alpha_{22} + \alpha_{NN})S_{2N} & \vdots & \alpha_{NN} \end{bmatrix}, \quad (2)$$

where N is the total number of AO basis functions. For EHM calculations, we need to set up all those matrices and then find the eigenvalues and eigenvectors [Eq. (1)]. The eigenvalues are the orbital energies, and the eigenvectors are the orbital coefficients that define the molecular orbital in terms of basis functions (i.e., the LCAO expansion).

In SQM, the diagonal elements of \mathbf{H} [Eq. (2)] $H_{\mu\mu} = \alpha_{\mu\mu}$ are parameterized as valence state ionization energies (IEs) with a minus sign to indicate binding. The off-diagonal Hamiltonian matrix elements are given by an approximation of the Wolfsberg–Helmholtz formula²⁷ that relates them to the diagonal elements and the overlap matrix element,

$$H_{\mu\nu} = \langle \chi_\mu | \hat{H} | \chi_\nu \rangle = \frac{1}{2}K^\ddagger(H_{\mu\mu} + H_{\nu\nu})S_{\mu\nu}, \quad (3)$$

where K^\ddagger is an empirically fit parameter, which scales the contribution of AO energy and overlap to the energy of MO.²⁷

1. Choice of $\alpha_{\mu\mu}$

The diagonal $H_{\mu\mu} = \alpha_{\mu\mu}$ elements in EHM [see Eq. (2)] are parameterized as negative of the valence state IEs for the appropriate isolated atom, as given by Koopmann's theorem.²⁸ In molecules, however, the atomic orbital energies are hybridization-dependent or, more generally, are environment-dependent. For example, IE for the isolated carbon atom differs from IE of the carbon atom in a saturated molecule because the associated configurations of those carbons are different: $1s^2 2s^2 2p^2$ vs $2sp^3$. Furthermore, the proper valence state for carbon in methane differs from that in ethylene, which, in turn, differs from that in acetylene, that is why various authors recommended molecule-specific sets of valence IEs.^{29,30}

2. Choice of K^\ddagger

The coefficient K^\ddagger [Eq. (3)] is an additional dimensionless empirical parameter that depends on the state of the molecule.^{27,31} The rationale for Eq. (3) is that the energy should be proportional

to the energy of the AO and should be greater when the overlap of the AOs is greater. The contribution of these effects to the energy is scaled by the K^\ddagger coefficient. For organic molecules, the EHM model was found to best agree with the experiment when $K^\ddagger = 1.75$,³² while for solids, a value of 2.3 is believed to be optimal.³¹ While K^\ddagger is usually being constant, the inclusion of overlaps into off-diagonal elements ($H_{\mu\nu}$) of the Hamiltonian \mathbf{H} [Eq. (2)] enforces that well-separated AOs do not interact.

B. ML parameterization of EHM

The idea behind the on-the-fly parameterization of EHM relies on the use of ML to locally learn the \mathbf{H} Hamiltonian matrix elements [Eq. (2)]. To this end, we have introduced the following procedure:

- (1) Find QM values of coefficients for the AO and MO energies for each individual molecule in the training dataset $\{C, E_{MO}\}$.
- (2) Train the \mathbf{H} [Eq. (2)] Hamiltonian with $\alpha_{\mu\mu}$ and K^\ddagger variables to reproduce QM values from the previous step. Solve the eigenvalue problem [Eq. (1)].
- (3) Use the ML \mathbf{H} Hamiltonian in the new eigenvalue equation [Eq. (1)] to predict $\{C, E_{MO}\}$ values for the target molecules.

As can be seen from above, interfacing ML with the EHM is accomplished by replacing the empirical matrix elements $\alpha_{\mu\mu}$ and coefficients K^\ddagger with their ML *learnable* counterparts. The choice of the learnable parameters is deliberate since the one-electron integral $\alpha_{\mu\mu}$ [Eq. (2)] is a function of nuclear charge and the type of AO; thus, it encodes the information about the *local chemical composition* given by the types of chemical elements. The coefficient K^\ddagger scales the contributions of the AO energies and AO overlaps to the MO energies. Thus, the K^\ddagger value is bond-dependent³³ and its role is to properly introduce the *i*th atom in the molecule in terms of its pairwise *local structure*. As the atomic chemical composition and the local atomic structure depend on higher-order features of atom's entire chemical environment (i.e., bond order,

conjugation, and hybridization), it is reasonable to approximate those unknown complex functions with NN. In the proposed ML-EHM scheme, NN models are aimed to reproduce the DFT reference orbital energies and respective MO densities for organic molecules containing four atom types: H, C, N, and O, as will be outlined in Subsection II C. Our NN model utilizes a DFT labeled training dataset to automatically learn representations of local AEs, which comes enough to predict Hamiltonian matrix elements.¹³

C. Neural network for Hamiltonian

The generic ML-EHM is based on a hierarchically interacting particle neural network (HIP-NN, Fig. 2).¹³ In general, HIP-NN implements the idea of atomic environment vectors (AEVs) and encodes the atomic configuration as a vector. Basically, AEV introduces the i th atom in terms of its (i) *local structure* R , described by a set of coordinates of all atoms within a cutoff, and (ii) *local chemical composition* Z , given by the types of chemical elements. HIP-NN encodes the atomic configurations using pairwise distances $R_{ij} = |R_i - R_j|$ with i th atom neighbors within a cutoff radius R_c into invariant fixed-length AEV, $\tilde{S}_v^Z = \{S_1, S_2, S_3, \dots, S_v\}$. Elements S_v probe specific regions of i th atom *radial* chemical environment and are assigned as *spatial sensitivity functions*.¹³ The key point of HIP-NN is that it encodes *local chemical composition* with learnable *feature vectors* (FVs), $\tilde{Z}_i^\ell = \{Z_1^0, Z_2^1, Z_3^2, \dots, Z_i^\ell\}$. The learning of \tilde{Z}_i^ℓ occurs through the *interaction layer* (green block) and *on-site layer* (red blocks), Fig. 2(b). On-site layers operate on the FVs of a single atom. The interaction layers additionally transmit information between atoms. Thus, the “message passing” in the interaction layers learns the atomic representations. “Message passing” not only captures the chemical environment of the i th atom but also accounts for the properties of the entire molecule. That makes HIP-NN a valuable tool capable of quantitative predicting of non-local molecular properties, such as, for instance, behavior of molecular orbitals.

Based on the learned representation of each atom, we trained HIP-NN to predict the contribution of this atom to the matrix

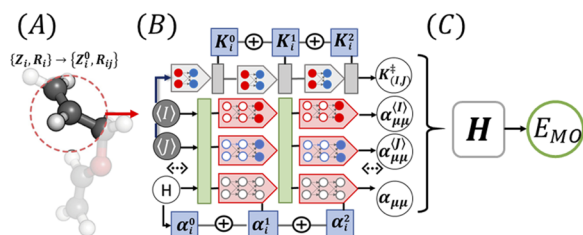


FIG. 2. (a) The molecular geometry is encoded in terms of feature vectors (FVs) to characterize the chemical environment of atoms $\{i, j\}$ and pairwise distances within the radial cutoff $\{Z_i^0, R_{ij}\}$. (b) HIP-NN builds sets of FVs through the interactions and on-site layers (green and red blocks, respectively). The FVs of an atomic pair are built through the pair regression layer (upper gray blocks). The total AO energies of each atom $\alpha_{\mu\mu}^{(i,j)}$ and pairwise $K_{(i,j)}^\ddagger$ parameter include their contributions α_i^n and K_i^n at all sites i and hierarchical levels n . (c) The EHM Hamiltonian, H , is built with the parameters $\{K^\ddagger, \alpha_{\mu\mu}\}$ of each atom and atomic pair. These parameters provide the eigenvalue problem to be solved next to find MO energies, E_{MO} .

elements of the Hamiltonian, $\{K^\ddagger, \alpha_{\mu\mu}\}$. Once the FVs are known, HIP-NN uses linear regression [Fig. 2(b), blue boxes] on the i th atom to model the local hierarchical output,

$$\alpha_i^n = \sum_{k=1}^N \omega_k^n Z_{i,k}^{\ell_n} + b^n, \quad (4)$$

where ω_k^n and b^n are learnable parameters. $Z_{i,k}^{\ell_n}$ from the layer is indexed by $\ell = 0, \dots, N_{\text{layer}}$. Finally, these hierarchical contributions [Eq. (4)] are combined together to recover the target diagonal matrix element of the Hamiltonian

$$H_{\mu\mu} = \alpha_{\mu\mu} = \sum_{n=0}^{N_{\text{interaction}}} \alpha_i^n. \quad (5)$$

The hierarchical decomposition is non-unique but should be designed such that α_i^n rapidly vanishes with increasing order n . To encourage this, we used a hierarchical regularization scheme developed earlier.¹³ The learned elements $\{K^\ddagger, \alpha_{\mu\mu}\}$ are then used for solving an eigenvalue problem [Eq. (1)], i.e., finding E_{MO} and coefficients, C_v , of AO according to a secular equation [Fig. 2(c)],

$$\sum_{v=1}^{N_{AO}} (H_{\mu\nu} - E_{MO} S_{\mu\nu}) C_v = 0, \quad \mu = 1, \dots, N_{AO}. \quad (6)$$

D. Extension of HIP-NN for K^\ddagger

The original HIP-NN potential has been constructed to predict atomic energies,¹³ charges,³⁴ and dipoles³⁵ in terms of local atomic charges or energy approximations. To predict K^\ddagger , a quantity that depends on the interaction of two atoms in the off-diagonal elements [Eq. (3)], we have introduced a new type of layer into the HIP-NN framework [Fig. 2(b)]. The latter enables HIP-NN to learn several matrix elements $\{K^\ddagger, \alpha_{\mu\mu}\}$ of the Hamiltonian within the same framework. This *pair regression layer* in Fig. 2(b) produces a set of features for each pair of atoms in the molecule. At every hierarchical level ℓ , there is a learnable matrix as a function of distance $u_{ab}^\ell(r)$, where the a and b indices contract over the pair of atomic features,

$$u_{ab}^\ell(r) = \sum_v U_{\tau,ab}^\ell s_\tau^\ell(r). \quad (7)$$

The sensitivity functions $s_\tau^\ell(r)$ are parameterized as in the *interaction layers* (the technical details are given in Ref. 13). $U_{v,ab}^\ell$ is a three-dimensional tensor of learnable parameters. The pair-valued parameter $K_{(I,J)}^\ddagger(\ell)$ is computed as a bilinear function,

$$K_{(I,J)}^\ddagger(\ell) = \sum_a \sum_b z_{I,a}^\ell u_{ab}^\ell(r_{IJ}) z_{J,b}^\ell + (I \leftrightarrow J). \quad (8)$$

Here, a and b indices feature vectors on neighboring atoms I and J , respectively. The second term in Eq. (8) carries out the symmetrization of the $K_{(I,J)}^\ddagger$ matrix. The prediction for $K_{(I,J)}^\ddagger$ is constructed by summing over the contributions from each hierarchical block ℓ , similarly to Eq. (5),

$$K_{(I,J)}^\ddagger = \sum_\ell K_{(I,J)}^{\ddagger\ell}. \quad (9)$$

The predicted $K_{(IJ)}^\ddagger$ matrix exists over atom pairs, whereas $K_{\mu\nu}^\ddagger$ needed for EHM exists over orbital pairs. To resolve this dimension mismatch, the predicted $K_{(IJ)}^\ddagger$ is expanded over the orbital indices such that all orbital pairs that share atom centers also share the same $K_{\mu\nu}^\ddagger$ parameter. Thus, from a high-level perspective, a modified HIP-NN framework in Fig. 2(b) allows for learning the entire matrix, including diagonal (on-site values) and off diagonal (interaction) elements. Recently, this model was successfully applied to learn chemical bond orders³⁶ relevant to off-diagonal matrix elements of the single-electron density matrix.

E. Loss function for training

We train the ML-EHM by varying the HIP-NN parameters [Eqs. (4) and (8)] to minimize a loss, which quantifies the error of a model in performing these tasks. Since our goal with the ML-EHM is to learn QM-based molecular physics, we train it to reproduce properties of DFT MO energies. We start by including an error term corresponding to the MO energy spectrum. However, our reference DFT data and the EHM model are defined over different basis sets. To determine a matching between the two sets of MOs, we examine only the MOs surrounding the occupation threshold. We take X occupied MOs and Y virtual MOs and order them by energy to compute the square difference between the ML-EHM and DFT orbital energies, E_{MO} ,

$$\Delta E_{MO} = \sum_{\chi=HOMO-X+1}^{LUMO+Y-1} (E_{\chi}^{DFT} - E_{\chi}^{ML-EHM})^2, \quad (10)$$

where index χ labels the MOs. However, such a loss term does not fully constrain the ML-EHM model. The Hamiltonian may reproduce the eigen-spectrum of DFT without reproducing the distribution of the electronic density over MOs. This is, particularly, important for preserving the orbital identities when MO energies cross depending on changes in molecular geometry. We address the orbital density distribution by defining an *orbital occupation fraction* for orbital χ on atom I ,

$$\mathcal{F}_{\chi,I} = \frac{\sum_{\mu \in \{I\}} C_{\chi,\mu}^2}{\sum_{\mu} C_{\chi,\mu}^2}. \quad (11)$$

The notation $\mu \in \{I\}$ denotes that the sum in the numerator runs over all atomic orbitals associated with atom I ; the sum in the denominator runs over all atomic orbitals in the molecule. This definition provides positive-definite atom-local charge densities for each MO, which can be compared between basis sets of different sizes. The sum-squared error of $\mathcal{F}_{\chi,I}$ is the *density error*

$$\Delta \wp_{MO} = \sum_{\chi}^{N_{orb}} \sum_I^{N_{atom}} (\mathcal{F}_{\chi,I}^{DFT} - \mathcal{F}_{\chi,I}^{ML-EHM})^2. \quad (12)$$

The full loss function \mathcal{L} is a combination of ΔE_{MO} , $\Delta \wp_{MO}$, and an L_2 regularization penalty using importance factors λ ,

$$\mathcal{L} = \lambda_{\Delta E_{MO}} \times \Delta E_{MO} + \lambda_{\Delta \wp_{MO}} \times \Delta \wp_{MO} + \lambda_{L_2} \times \mathcal{L}_{L_2}. \quad (13)$$

The \mathcal{L}_{L_2} regularization penalizes large weight values and promotes smoothness of the network. λ values were empirically chosen to be

$\lambda_{\Delta E_{MO}} = 1.0$ and $\lambda_{\Delta \wp_{MO}} = 10^{-3}$, $\lambda_{L_2} = 10^{-6}$. Interestingly, the inclusion of the $\Delta \wp_{MO}$ term during training results in better minimization of the ΔE_{MO} term when compared to a model trained without $\Delta \wp_{MO}$. We take this as evidence that the inclusion of $\Delta \wp_{MO}$ successfully promotes the construction of physically meaningful Hamiltonians that correctly track the character of MOs.

F. ML-EHM training dataset

To train the HIP-NN model to predict the ML-EHM parameters, we use a small fraction of the ANI-1x dataset.³⁷ The ANI-1x dataset consists of small organic molecules at non-equilibrium conformations. The ANI-1x dataset was generated in an active learning process with the ANI potential.⁷ The final dataset contains $\sim 5 \times 10^6$ DFT (wB97x/6-31G*) calculations, each computed with a neutral charge state and a singlet spin state. All QM calculations in this work were performed with Gaussian 09.³⁸ Molecules are obtained from various sources, such as GDB-11^{39,40} and ChEMBL,⁴¹ and processed with the RDKit software package.⁴² The details about the dataset can be found in our previous publication.³⁷

The HIP-NN model training was performed with all molecules smaller than five heavy atoms. This was found to aid HIP-NN in producing realistic parameters for the ML-EHM. In addition to these molecules, a random sample of the remaining molecules with 18 atoms or fewer was added to the training set. The 18-atom cutoff was used for training due to poor scaling of the training code to larger system sizes: the overlap matrices are zero-padded to the largest possible Hamiltonian size. Three distinct models were trained, each to a different number of occupied and virtual orbitals. The final size of the training datasets and target orbitals is given in Sec. S1 of the [supplementary material](#).

III. RESULTS AND DISCUSSION

A. Test for dataset accuracy

Figure 3 compares the accuracy of three ML-EHM models with different orbital configurations from Table I. The comparison is performed by predicting on the remainder of ANI-1x (described in Sec. II F) with 20 atoms or fewer. The size limit was imposed so that we would have a measure of how well the models perform on molecules similar to those in the training set. Obviously, the (4×0) model has the best predictive performance on the remaining ANI-1x data. It can predict the energies of the four highest occupied orbitals in the training set with a MAD of 0.092 eV. In contrast, the (4×0) model is completely unable to predict the bandgap of organic molecules, which is anticipated since it was never trained to virtual orbitals. (4×2) and (2×2) models trained to virtual orbitals show a lower accuracy when predicting orbital energies. This is likely due to the ML-EHM form lacking explicit Coulomb and exchange terms, making the model insufficiently complex to describe both the occupied and virtual MOs from DFT (we recall that the bandgap has a very steep dependence on the fraction of the orbital exchange in the hybrid DFT models^{43,44}). Despite this, the (4×2) model can predict bandgaps with an accuracy of about 0.2 eV.

B. ML-EHM validation on COMP-6 dataset

Figure 4 shows our ML-EHM model's performance on a portion of the COMP6 benchmark³⁷ to validate the ML-EHM ability

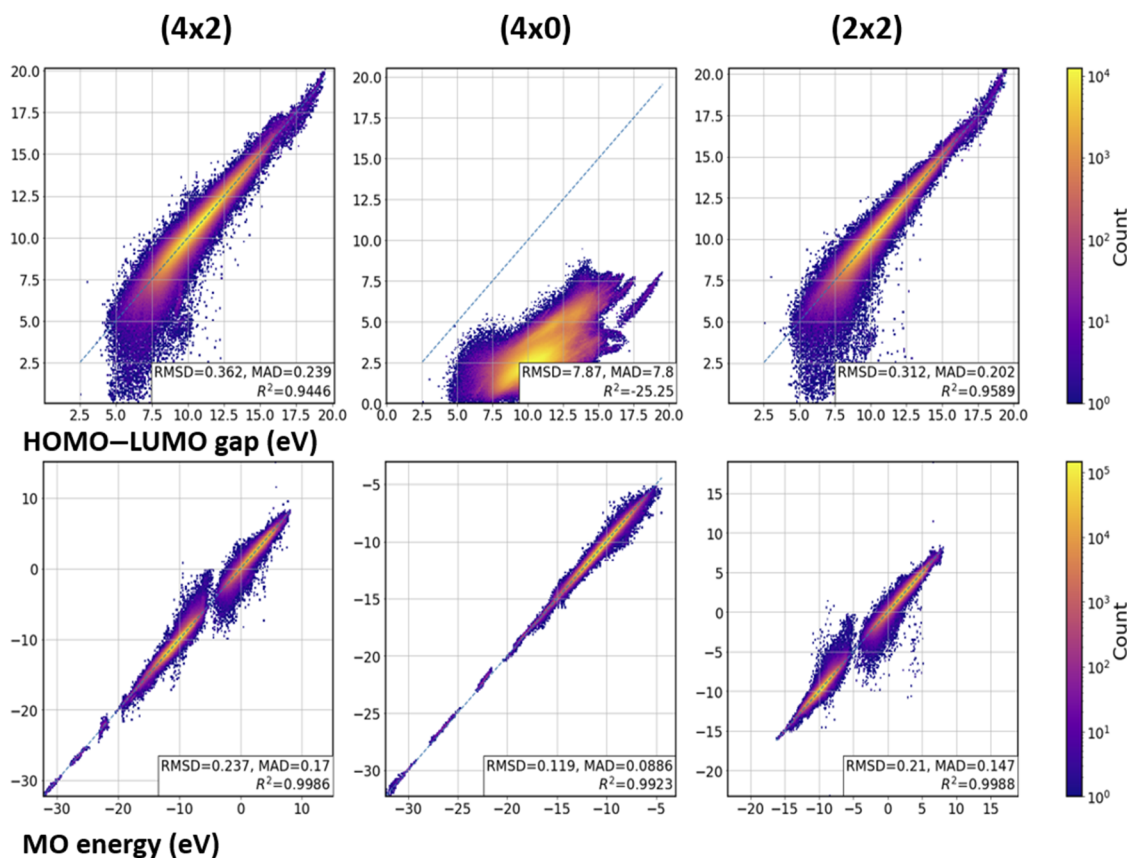


FIG. 3. Predictions of ML-EHM on the remaining ANI-1x datapoints that were not included in the training sets. This includes some large molecules that were intentionally left out of training. While the (4 × 0) model makes an excellent occupied orbital prediction, it is unable to predict the bandgaps. Other models trained to virtual orbitals give worse predictions on orbital energies but are able to make quantitative predictions on the bandgap.

to extend its predictions to molecular systems larger than those in the training dataset. It consists of three separate benchmarks. The GDB benchmarks (i.e., GDB7-8 and GDB10-13) are randomly subsampled molecules from the GDB-13 database. Each GDB subset contains a few thousand of different molecules. The s66 × 8 benchmark⁴⁵ is a well-known benchmark for testing the intermolecular interaction for dimer molecules. All test sets are described in more detail in the previous work.³⁷ Naturally, all comparisons of ML-EHM were done to the same reference DFT level (wb97x/6-31G*).

Generally, the performance of ML-EHM is quite good. The MO energies' predictions on the subsets that are most similar to ANI-1x,

GDB 7–9 and GDB 10–13, by (4 × 0) have mean absolute deviations of 0.089 and 0.13 eV, respectively [Figs. 4(a)–4(f)]. The performance degrades by roughly a factor of 2 when applying the model to the larger dimer systems found in s66 × 8 [Figs. 4(g)–4(i)]. The molecules generating the highest errors in the s66 × 8 dataset contain ring structures (see Sec. S2 of the [supplementary material](#)).

Figure 5 shows histograms of the {K[‡], α_{μμ}} learnable parameters, generated by HIP-NN when applied to the COMP-6 dataset. These are the α_{μμ} values for s- and p-orbitals and K[‡] values fed into the ML-EHM models. The histograms make it clear that the (4 × 0) model has the clearest distinction between atom types, with each atom type in Fig. 5(b) having a distinct distribution. The incorporation of virtual orbitals into the training procedure makes the distributions of parameter predictions indistinct, with all atom types having s-orbital energies over roughly the same range [Fig. 5(c)]. One interesting point is an agreement between the ML-EHM and the original EHM in terms of prediction of the hydrogen s-orbital energy. In EHM, this parameter was taken from the ionization energy of a hydrogen atom, which is 13.6 eV. By training the HIP-NN on molecular orbital energies, the (4 × 0) network produces

TABLE I. ML-EHM models and training datasets explored in this paper.

Model	Training orbitals	Molecules in training set
(4 × 2)	HOMO–3 to LUMO+1	172 588
(4 × 0)	HOMO–3 to HOMO	210 933
(2 × 2)	HOMO–1 to LUMO+1	210 652

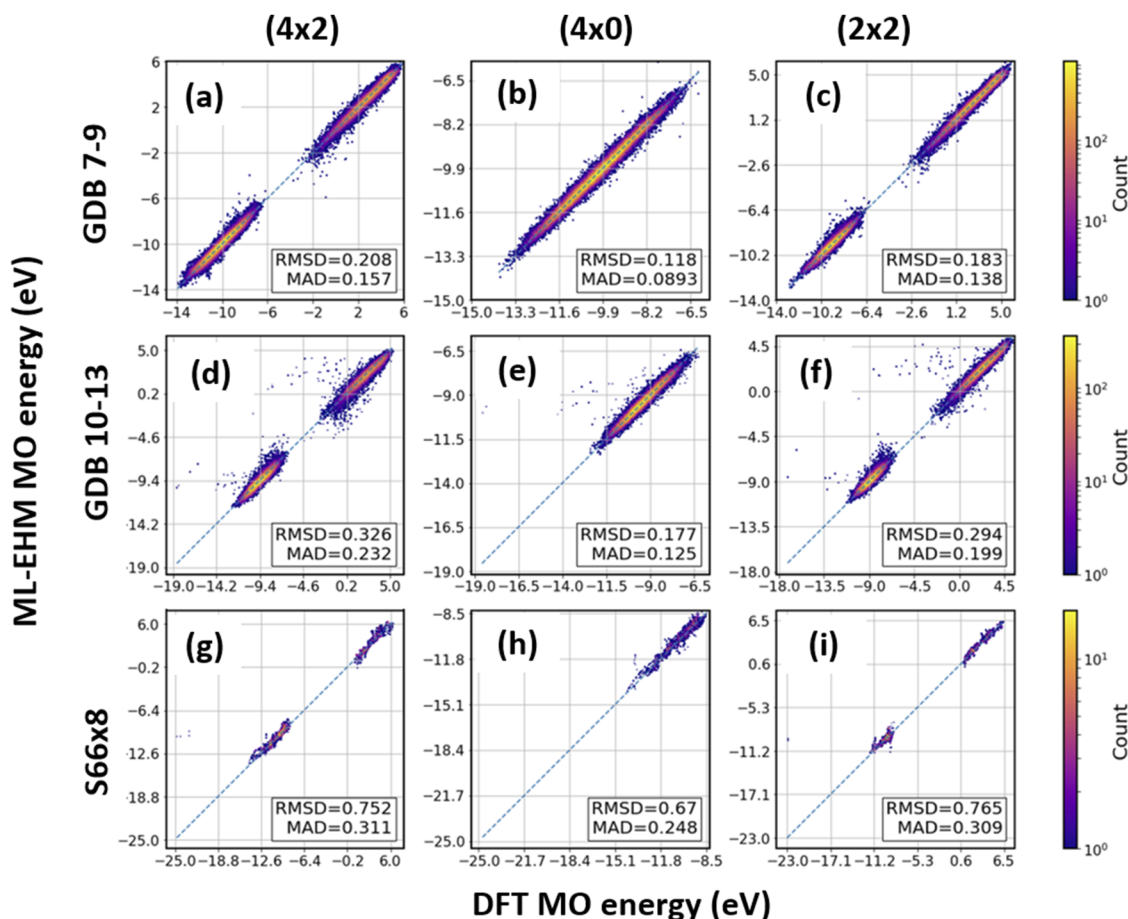


FIG. 4. Prediction of the EHM-ML models on the COMP-6 test set. The portion tested consists of three separate benchmarks: GDB 7–9 [(a)–(c)], GDB 10–13 [(d)–(f)], and $s66 \times 8$ [(g)–(i)]. Generally, the models perform well on all panels [(a)–(i)].

an average hydrogen s-orbital energy of 12.5 eV, close to the EHM value. Moreover, the 12.5 eV value is consistent with more sophisticated semiempirical methods, such as MNDO⁴⁶ (11.9 eV) and OMx models⁴⁷ (~12.5 eV). The proton in molecules typically comes out in semiempirical methods as a weaker electron acceptor than in vacuum, rationalizing the difference with 13.6 eV of a hydrogen atom. Another observation is that the sequence of orbital energies ($O < N < C < H$) is in line with the expected series from a simple electrostatic argument, showing that the oxygen S-orbital is lower in energy than the S-orbital of hydrogen. Additionally, the HIP-NN reproduces a bifurcated distribution of the K^\ddagger off-diagonal matrix elements with an average value of 1.54, which matches the EHM parameterization of 1.75. While there are two maxima in the distribution of predicted K^\ddagger values, these maxima do not necessarily correlate with atom types in the pair to which the K^\ddagger value is assigned. However, it probably points to the sp^2 and sp^3 C atom hybridization in the molecules of the training dataset. The appearance of these two maxima requires further investigation.

IV. REPRESENTATIVE MOLECULAR CASE STUDIES

A. Molecular orbital view of C–H stretching in methane

The orbital shapes and energies of valence electrons are crucial to obtaining a wide variety of quantum mechanical properties. Examples of these applications are Fukui's frontier electron orbital theory of reactivity⁴⁸ and the Woodward–Hoffmann rules,⁴⁹ where information on orbital distributions and phase is required (not just total electron density). Effective Hamiltonian models that correctly describe the behavior of valence electrons are necessary for such analyses. In this section, we examine ML-EHM accuracy in describing the orbital structure of a C–H stretch of a methane molecule.

The (4×0) model was employed to find the orbital energies and wave functions of the four highest energy molecular orbitals of methane when one of hydrogen is stretched out from carbon

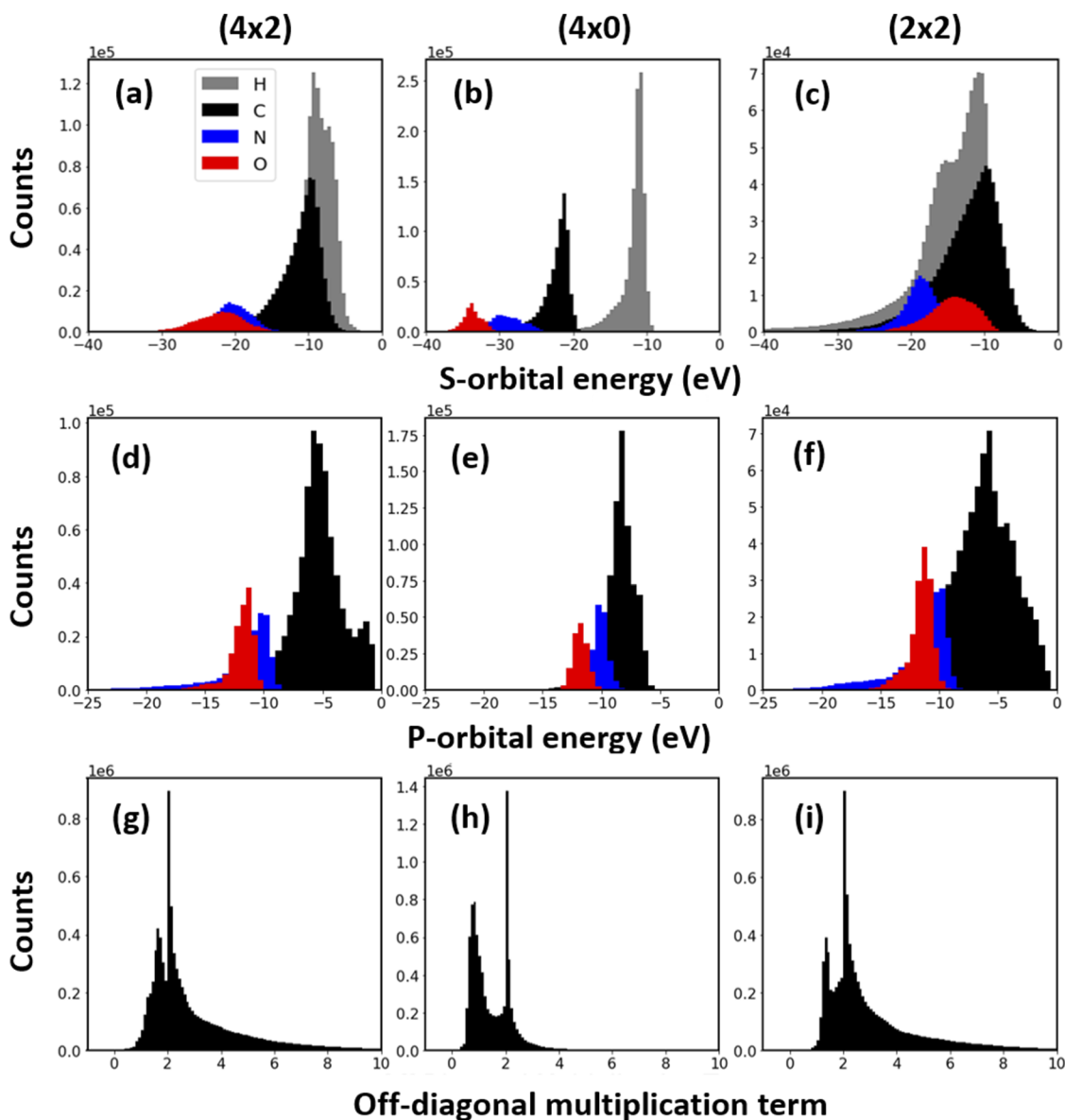


FIG. 5. Histograms of ML-EHM learnable parameters over the COMP-6 test dataset. The hydrogen s-orbital energy [(a)–(c)] is in remarkable agreement with the EHM assignment of -13.6 eV. Additionally, although the off diagonal factor is bifurcated in the ML-EHM [(g)–(i)], its average value is in agreement with the EHM empirical K^{\ddagger} value of 1.75.

[Fig. 6(a)]. In testing, we found that only (4×0) was able to capture the orbital physics detailed below. This is likely due to the EHM Hamiltonian's simplicity: it is unable to make simultaneous quantitative predictions on occupied and virtual orbitals. Figure 6(a) displays the ML-EHM predicted $\alpha_{\mu\mu}$ on-site energies along with the empirical parameterization used by Pople and Segal.³⁰ In the original approach, $\alpha_{\mu\mu}$ are the constant values of the valence state IEs for the appropriate isolated atoms, whereas ML-EHM produces a smooth function for $\alpha_{\mu\mu}$ depending on the bond length. As shown in Fig. 6(a), the dissociating hydrogen (H^a) 1S orbital energy increases

as the hydrogen distances from the carbon. As a result of increasing site energy, one would expect the hydrogen to decouple from the lowest energy molecular orbitals, reflecting a dissociation of bonding between the hydrogen and the CH_3 fragment. Simultaneously, the carbon 2S orbital energy decreases by a similar amount, while the carbon 2P orbital energies remain approximately constant. Near the equilibrium bond distance (~ 1.09 Å for our reference DFT), all on-site energies are not variable. This provides a strong evidence that the empirical parameterization of EHM is valid at the equilibrium geometries.

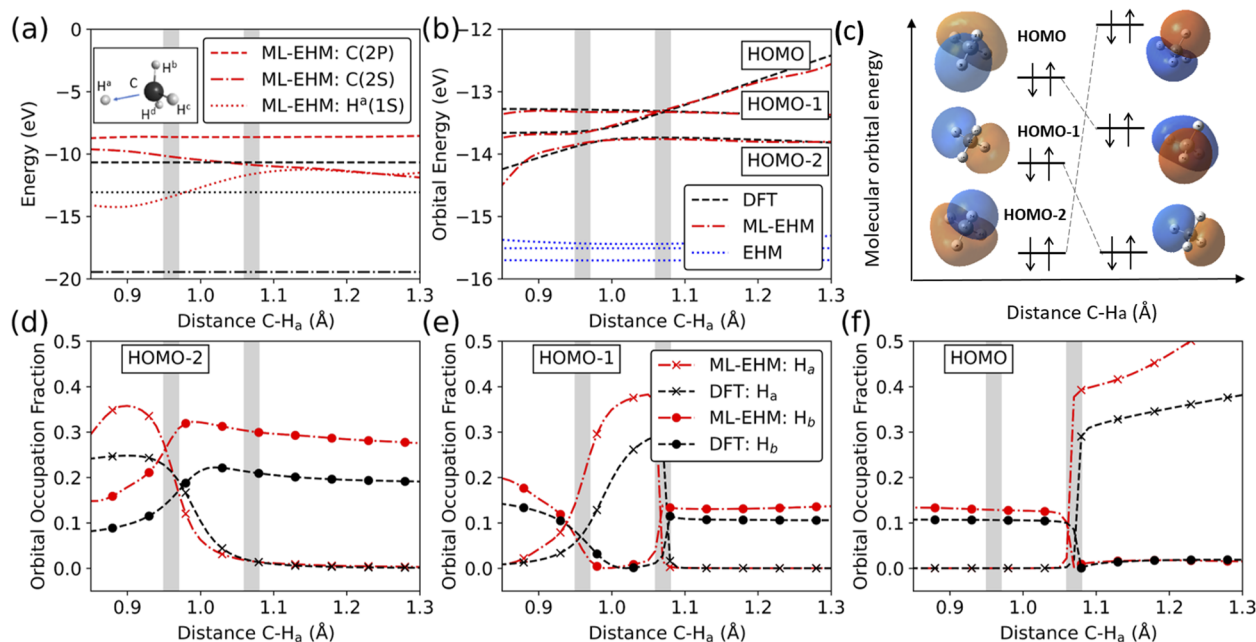


FIG. 6. ML-EHM $\alpha_{\mu\mu}$ energies and corresponding Pople and Segal $\alpha_{\mu\mu}$ values shown in black (a) and MO energies (b) as a function of hydrogen extraction coordinate. The partial MO diagram (c) illustrates how the frontier MO (FMO) changes in the process of C-H^a bond stretching. The two gray areas on the diagrams indicate the FMO crossings between HOMO-2 and HOMO-1 and HOMO and HOMO-1. [(d)-(f)] The FMO occupation fractions for H_a and H_b.

In Fig. 6, we show that the dynamic ML-EHM parameters correctly reproduce the electronic structure changes witnessed in the source DFT throughout the hydrogen abstraction coordinate. This is shown through MO energies and charge occupation via the *orbital occupation fractions*, $\mathcal{F}_{\chi,I}$ [Eq. (11)]. Figure 6(b) indicates that in contrast to EHM, both DFT and ML-EHM produce two orbital crossings throughout the methane dissociation coordinate: one at ≈ 0.95 Å between HOMO-2 and HOMO-1 and one at ≈ 1.1 Å between HOMO-1 and HOMO. By plotting the DFT orbitals [Fig. 6(c)], it becomes clear that HOMO-2 transforms to the HOMO. Meanwhile, the HOMO and HOMO-1 are shifted downward but maintain their order as the C-H_a bond stretches.

The correspondence between orbital energy and orbital character as determined by $\mathcal{F}_{\chi,I}$ [Eq. (11)] is of particular interest. As HOMO-2 and HOMO-1 approach in energy near a C-H_a separation of ≈ 0.95 Å, the character of these two orbitals switches as indicated in Figs. 6(d) and 6(e). HOMO-1 H_a $\mathcal{F}_{\chi,I}$ starts near zero and transitions to 0.4 e⁻. The opposite is observed for HOMO-2 H_a $\mathcal{F}_{\chi,I}$, indicating that around 0.95 Å, these two orbitals switch character. At 1.1 Å, another closer crossing between the HOMO and HOMO-1 appears. Correspondingly, a more abrupt exchange of orbital character, as indicated by H_a $\mathcal{F}_{\chi,I}$, is shown in Figs. 6(e) and 6(f) near 1.1 Å. Both the gradual and abrupt crossings at ≈ 0.95 and at ≈ 1.1 Å, respectively, are captured by the ML-EHM without discontinuous changes in EHM parameters, as shown in Figs. 6(a) and 6(b). This example illustrates how the NN and EHM parts of ML-EHM can harmonize to capture the orbital physics concomitant to large geometry changes. The EHM can capture

rapid changes in orbital physics using diagonalization, and the NN parameters $\{K^{\ddagger}, \alpha_{\mu\mu}\}$ provide smooth modulation of the underlying Hamiltonian.

B. Internal rotation in conjugated systems

The key success of the EHM model was its predictions for the behavior of conjugated π -systems in electrocyclic reactions. This became the theoretical background for the Woodward-Hoffman rules and the subject of the 1981 Nobel Prize in Chemistry. While limitations in the training dataset preclude our model's application to reactive systems, it can be applied to bond rotations in conjugated π -systems. To illustrate this, we use ML-EHM to study the central bond rotation in 1,3-butadiene and 2-aza-1,3-butadiene. While these two systems have the same number of electrons, aza-butadiene has one fewer core σ -orbital. Thus, the LUMO in butadiene becomes the HOMO in aza-butadiene, with a similar shift happening between all valence MOs. This can be clearly seen when examining the π -system in Fig. 7. We have already checked that a change in ML-EHM energy of FMO reflects certain changes in calculated orbital occupation fractions of the corresponding AOs. We now consider the internal rotation in conjugated systems to check that such an ML-EHM FMO energy change agrees directly with the structural changes in the molecular system.

1. Orbital correlation between *s-trans/s-cis*-1,3-butadiene

We consider the *s-cis*-1,3-butadiene and *cis*-2-aza-1,3-butadiene conformers with a torsion angle Θ rotation from 180° to 0° . The ground electronic states of *s-trans*- and *s-cis*-1,3-butadiene

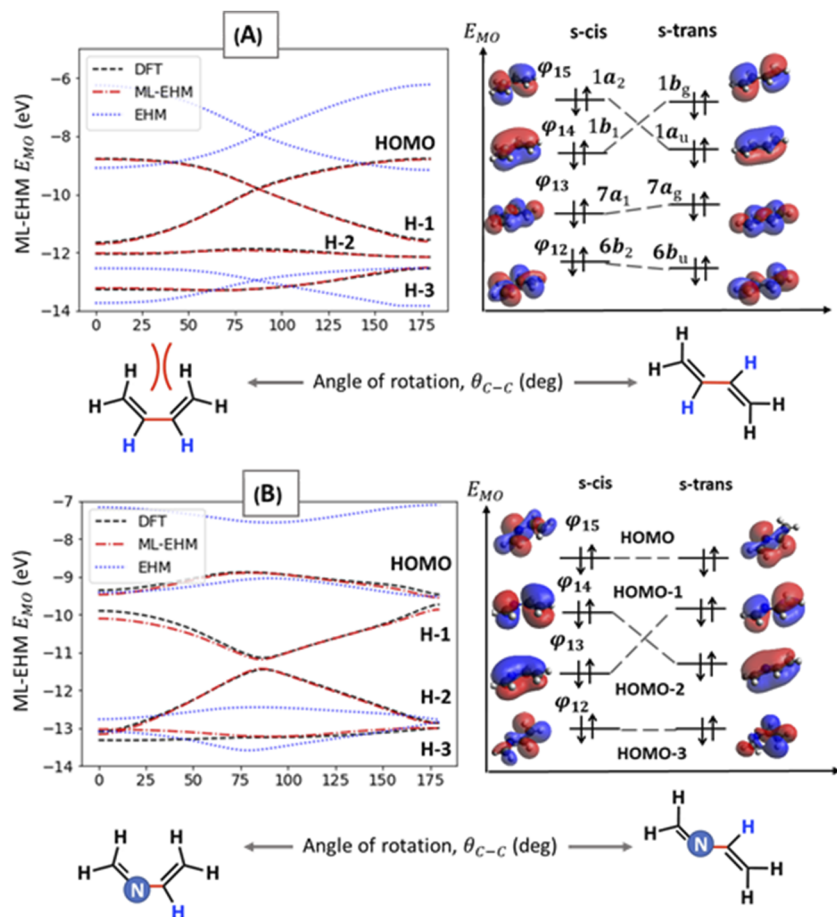


FIG. 7. ML-EHM FMO energies (left) and partial FMO diagram (right), which illustrates the orbital changes in the process of internal rotation around the central single bond for 1,3-butadiene (a) and 2-aza-1,3-butadiene (b).

and 2-aza-1,3-butadiene conformers have closed shells with 15 doubly occupied orbitals. There are four core orbitals and four inner valence orbitals; the remaining seven orbitals are outer valence orbitals. Here, we apply our (4×0) ML-EHM to predict the behavior of the four highest lying occupied orbitals in both butadiene and aza-butadiene as they undergo a torsional bond rotation. In *cis*-butadiene, the four highest orbitals have the following symmetry:

$$X^1 A_1 : (6b_2)(7a_1)(1b_1)(1a_2).$$

In *trans*-butadiene, the symmetries of the highest orbitals are as follows:

$$X^1 A_g : (6b_u)(7a_g)(1a_u)(1b_g).$$

Torsional rotation causes the butadiene to lose a mirror plane and gain an inversion center, causing the point group to change from C_{2h} to C_{2v} . Additionally, when the bond rotates, orbitals that were symmetric with respect to the mirror plane transform into orbitals that are symmetric under inversion. The same holds true for orbitals that are anti-symmetric under reflection and inversion. This simple symmetry argument predicts a crossing between the HOMO and HOMO-1 in butadiene as the molecule undergoes a torsional rotation. As seen in Fig. 7, both DFT and ML-EHM find that the HOMO and HOMO-1 orbitals cross when butadiene undergoes a torsional

rotation, and the ML-EHM model remains in a quantitative agreement with the energy levels of DFT. For an analysis of the $\alpha_{\mu\mu}$ orbital parameters constructed by HIP-NN, which change very little over the course of the rotation, see Sec. S3 of the [supplementary material](#).

A very similar event happens in aza-butadiene: here, HOMO-1 and HOMO-2 are anticipated to cross. Generally, aza-butadiene is a much more complicated case and does not lend itself to a simple molecular orbital description as was done with butadiene.⁵⁰ Nonetheless, the ML-EHM remains in a quantitative agreement with the underlying DFT throughout the torsional rotation in contrast to EHM results.

V. CONCLUSIONS

Machine learning is rapidly proving to be a powerful tool for computational chemists. In the past, accurate QM calculations were prohibitively expensive, preventing their application to many interesting problems. By training an ML model to replicate these results, it is now possible to compute high-level QM properties at a rate of microseconds per atom. However, the results produced from pure ML models are still difficult to interpret. By interfacing ML with a physics-based effective Hamiltonian model (EHM in this work) for electrons in molecules, we can retain the accuracy of ML, the interpretability of *ab initio* calculations, and the speed of a reduced

dimensionality description of quantum mechanics. Notably, the EHM used here for a molecular case is a representative of a general class of tight-binding Hamiltonians broadly exploited for both solids and finite systems. A modified HIP-NN model naturally trains to on-site energies and hopping integrals in the context of any tight-binding model, such as EHM. This training is dynamically parameterizing the EHM Hamiltonian: its parameters are varying smoothly as the molecule changes its conformation and passes through regions of orbital crossings.

Importantly, mapping the original quantum mechanical problem to a simple model Hamiltonian provides a viable approach to address the spatial locality challenge for conventional ML schemes targeting individual molecular properties. The latter frequently have an empirically determined spatial cutoff radius of 3–6 Å^{7,13,51–53} and may further utilize physics-based models (e.g., electrostatics) to extend beyond this range.^{54,55} In contrast, such interactions are naturally introduced through physically transparent terms of the reduced Hamiltonian form.

The ML-EHM provides many surprising and promising results. First, the ML model recovered many of the original parameters used in the EHM. Examples of this include the hydrogen s-orbital energy, as well as the empirical K^{\ddagger} parameter used to modify the off-diagonal elements of the EHM Hamiltonian. Additionally, the ML-EHM recovered many of the orbital signatures seen in both the h-vibration on methane and the bond rotation on butadiene and aza-butadiene. The ML-EHM model's largest shortcoming is the lack of explicit Coulomb and exchange terms in the underlying Hamiltonian. Neglecting these terms significantly simplifies training as electron density does not need to be iterated to achieve self-consistency (i.e., mean field convergence being the case for Hartree-Fock or DFT approaches). However, their absence also implies that the ML-EHM's total energy is simply a sum of orbital energies. This is not true for the reference DFT method, which prevented multi-objective training to both total molecular energy and valence orbital energy.

Future work should focus on interfacing ML models with more complex effective Hamiltonian representations of QM and targeting more QM properties in multi-objective training. These semiempirical effective Hamiltonian models (such as AM1,⁵⁶ PM series,⁵⁷ OMx⁴⁷, and others⁵⁸) will simultaneously allow for training to the total molecular energy, valence orbital energy, and molecular orbital density. These more complicated models will also facilitate accurate training to valence orbitals, which is not fully accomplished here without partly destroying the physics encoded in the EHM model. This all points to an auspicious future for ML enhanced physics models in quantum chemistry.

AUTHORS' CONTRIBUTIONS

T.Z. and B.N. contributed equally to this work.

SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for a detailed description of the training datasets and atomic site parameters for the butadiene molecule discussed in the main text.

ACKNOWLEDGMENTS

This work was done, in part, at the Center for Nonlinear Studies (CNLS) and the Center for Integrated Nanotechnologies (CINT) at Los Alamos National Laboratory (LANL). Funding for this work was provided by the LANL Laboratory Directed Research and Development (LDRD) program. We also acknowledge the LANL Institutional Computing (IC) program and ACL data team for providing computational resources. J.S.S. thanks the Advanced Simulation and Computing (ASC) program for the Nicholas C. Metropolis Postdoctoral Fellowship. O.I. thanks CNLS for their support and hospitality. O.I. acknowledges support from the DOD-ONR (Grant No. N00014-16-1-2311) and the NSF, Grant No. CHE-1802789. O.I. and T.Z. acknowledge support from the NSF EPSCoR RII, Grant No. OIA-1632899.

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its [supplementary material](#) and in GitHub at: <https://github.com/aiqm/EHM-ML>.

REFERENCES

- ¹T. Zubatiuk and O. Isayev, *Acc. Chem. Res.* **54**, 1575 (2021).
- ²T. W. Ko, J. A. Finkler, S. Goedecker, and J. Behler, *Acc. Chem. Res.* **54**, 808 (2021).
- ³P. O. Dral, *J. Phys. Chem. Lett.* **11**, 2336 (2020).
- ⁴M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, [arXiv:1603.04467](https://arxiv.org/abs/1603.04467) (2016).
- ⁵A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, NIPS work, 2017.
- ⁶J. S. Smith, B. T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev, and A. E. Roitberg, *Nat. Commun.* **10**, 2903 (2019).
- ⁷J. S. Smith, O. Isayev, and A. E. Roitberg, *Chem. Sci.* **8**, 3192 (2017).
- ⁸J. Klicpera, J. Groß, and S. Günnemann, [arXiv:12003.03123](https://arxiv.org/abs/12003.03123) (2020).
- ⁹S. Lorenz, A. Groß, and M. Scheffler, *Chem. Phys. Lett.* **395**, 210 (2004).
- ¹⁰J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007).
- ¹¹A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Phys. Rev. Lett.* **104**, 136403 (2010).
- ¹²M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, *Phys. Rev. Lett.* **108**, 058301 (2012).
- ¹³N. Lubbers, J. S. Smith, and K. Barros, *J. Chem. Phys.* **148**, 241715 (2018).
- ¹⁴O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, "Machine learning force field," *Chem. Rev.* (published online, 2021).
- ¹⁵S. Manzhos, *Mach. Learn.: Sci. Technol.* **1**, 013002 (2020).
- ¹⁶E. Hückel, *Z. Phys.* **83**, 632 (1933).
- ¹⁷E. Hückel, *Z. Phys.* **72**, 310 (1931).
- ¹⁸E. Hückel, *Z. Phys.* **76**, 628 (1932).
- ¹⁹J. Behler, *Int. J. Quantum Chem.* **115**, 1032 (2015).
- ²⁰K. T. Schütt, M. Gastegger, A. Tkatchenko, K.-R. Müller, and R. J. Maurer, *Nat. Commun.* **10**, 5024 (2019).
- ²¹G. Carleo and M. Troyer, *Science* **355**, 602 (2017).
- ²²M. Sugawara, *Comput. Phys. Commun.* **140**, 366 (2001).
- ²³P. O. Dral, O. A. von Lilienfeld, and W. Thiel, *J. Chem. Theory Comput.* **11**, 2120 (2015).

- ²⁴H. Li, C. Collins, M. Tanha, G. J. Gordon, and D. J. Yaron, *J. Chem. Theory Comput.* **14**, 5764 (2018).
- ²⁵R. Hoffmann, *J. Chem. Phys.* **39**, 1397 (1963).
- ²⁶R. Hoffmann, *Rev. Mod. Phys.* **60**, 601 (1988).
- ²⁷M. Wolfsberg and L. Helmholz, *J. Chem. Phys.* **20**, 837 (1952).
- ²⁸T. Koopmans, *Physica I*, 104 (1934).
- ²⁹J. Hinze and H. H. Jaffe, *J. Am. Chem. Soc.* **84**, 540 (1962).
- ³⁰J. A. Pople and G. A. Segal, *J. Chem. Phys.* **43**, S136 (1965).
- ³¹J. Cerdá and F. Soria, *Phys. Rev. B* **61**, 7965 (2000).
- ³²J. H. Ammeter, H. B. Bürgi, J. C. Thibeault, and R. Hoffmann, *J. Am. Chem. Soc.* **100**, 3686 (1978).
- ³³L. Rincón, A. Hasmy, C. A. Gonzalez, and R. Almeida, *J. Chem. Phys.* **129**, 044107 (2008).
- ³⁴B. Nebgen, N. Lubbers, J. S. Smith, A. E. Sifain, A. Likhov, O. Isayev, A. E. Roitberg, K. Barros, and S. Tretiak, *J. Chem. Theory Comput.* **14**, 4687 (2018).
- ³⁵A. E. Sifain, N. Lubbers, B. T. Nebgen, J. S. Smith, A. Y. Likhov, O. Isayev, A. E. Roitberg, K. Barros, and S. Tretiak, *J. Phys. Chem. Lett.* **9**, 4495 (2018).
- ³⁶S. Magedov, C. Koh, W. Malone, N. Lubbers, and N. Benjamin, *J. Appl. Phys.* **129**, 064701 (2021).
- ³⁷J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev, and A. E. Roitberg, *J. Chem. Phys.* **148**, 241733 (2018).
- ³⁸M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, *GAUSSIAN 09*, Revision E.01, 2009.
- ³⁹T. Fink and J.-L. Reymond, *J. Chem. Inf. Model.* **47**, 342 (2007).
- ⁴⁰T. Fink, H. Bruggesser, and J.-L. Reymond, *Angew. Chem., Int. Ed.* **44**, 1504 (2005).
- ⁴¹A. P. Bento, A. Gaulton, A. Hersey, L. J. Bellis, J. Chambers, M. Davies, F. A. Krüger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos, and J. P. Overington, *Nucleic Acids Res.* **42**, D1083 (2014).
- ⁴²G. Landrum, www.RdKit.Org (n.d.).
- ⁴³H. Sun and J. Autschbach, *J. Chem. Theory Comput.* **10**, 1035 (2014).
- ⁴⁴C. Sutton, J. S. Sears, V. Coropceanu, and J.-L. Brédas, *J. Phys. Chem. Lett.* **4**, 919 (2013).
- ⁴⁵J. Řezáč, K. E. Riley, and P. Hobza, *J. Chem. Theory Comput.* **7**, 3466 (2011).
- ⁴⁶M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.* **99**, 4899 (1977).
- ⁴⁷P. O. Dral, X. Wu, L. Spörkel, A. Koslowski, W. Weber, R. Steiger, M. Scholten, and W. Thiel, *J. Chem. Theory Comput.* **12**, 1082 (2016).
- ⁴⁸K. Fukui, in *Quantum Chemistry*, edited by P. O. Löwdin, O. Goscinski, and J.-L. Calais (Wiley, 1978), p. 277, a scientific melting pot: a symposium sponsored by the Quantum Chemistry Group to mark the 500th anniversary of the University of Uppsala, held 31 August through 4 September 1977.
- ⁴⁹R. B. Woodward and R. Hoffmann, *The Conservation of Orbital Symmetry* (Verlag Chemie and Academic Press, Weinheim, NY, 1971).
- ⁵⁰Z. Neiman, *J. Chem. Soc., Perkin Trans. 2* **2**(12), 1746 (1972).
- ⁵¹J. Behler, *J. Chem. Phys.* **134**, 074106 (2011).
- ⁵²R. Zubatyuk, J. S. Smith, J. Leszczynski, and O. Isayev, *Sci. Adv.* **5**, eaav6490 (2019).
- ⁵³K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, *J. Chem. Phys.* **148**, 241722 (2018).
- ⁵⁴O. T. Unke and M. Meuwly, *J. Chem. Theory Comput.* **15**, 3678 (2019).
- ⁵⁵T. Morawietz, V. Sharma, and J. Behler, *J. Chem. Phys.* **136**, 064103 (2012).
- ⁵⁶M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.* **107**, 3902 (1985).
- ⁵⁷J. J. P. Stewart, *Encyclopedia of Computational Chemistry* (Wiley, 1998).
- ⁵⁸T. Husch, A. C. Vaucher, and M. Reiher, *Int. J. Quantum Chem.* **118**, e25799 (2018).
- ⁵⁹P. A. M. Dirac, *Proc. R. Soc. London, Ser. A* **123**, 714 (1929).