

Significance of the Chemical Environment of an Element in Nonadiabatic Molecular Dynamics: Feature Selection and Dimensionality Reduction with Machine Learning

Wei Bin How, Bipeng Wang, Weibin Chu, Alexandre Tkatchenko, and Oleg V. Prezhdo*



Cite This: *J. Phys. Chem. Lett.* 2021, 12, 12026–12032



Read Online

ACCESS |



Metrics & More

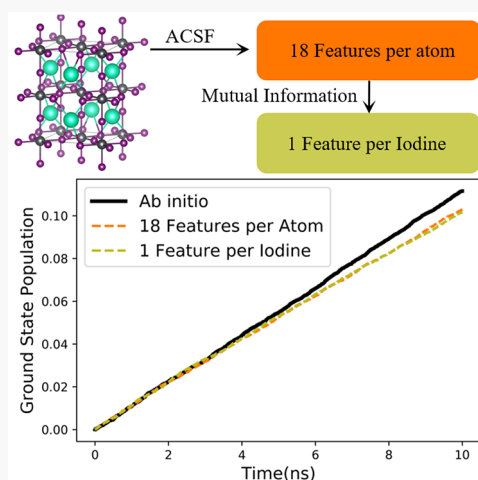


Article Recommendations



Supporting Information

ABSTRACT: Using supervised and unsupervised machine learning (ML) on features generated from nonadiabatic (NA) molecular dynamics (MD) trajectories under the classical path approximation, we demonstrate that mutual information with the NA Hamiltonian can be used for feature selection and model simplification. Focusing on CsPbI₃, a popular metal halide perovskite, we observe that the chemical environment of a single element is sufficient for predicting the NA Hamiltonian. The conclusion applies even to Cs, although Cs does not contribute to the relevant wave functions. Interatomic distances between Cs and I or Pb and the octahedral tilt angle are the most important features. We reduce a typical 360-parameter ML force-field model to just a 12-parameter NA Hamiltonian model, while maintaining a high NA-MD simulation quality. Because NA-MD is a valuable tool for studying excited state processes, overcoming its high computational cost through simple ML models will streamline NA-MD simulations and expand the ranges of accessible system size and simulation time.



Machine learning (ML) has been attracting more interest and traction in the field of physical and quantum chemistry since its boom over the past decade.^{1–4} Its ability to deduce nontrivial trends from data has contributed to its popularity in a wide range of applications from design of materials for catalysis or solar power to the fundamental advancement of our chemical knowledge.^{5–15} ML techniques can be broadly classified into two categories, supervised and unsupervised. Supervised learning techniques typically involve the prediction of a target variable based on a set of predictors, while unsupervised learning is mainly used to deduce patterns and trends that may be hidden in the data set.

Recently, it has been demonstrated that supervised ML can be used to accelerate nonadiabatic (NA) molecular dynamics (MD) simulations^{16–20} and that an increase in efficiency of ≤ 2 orders of magnitude can be realized by interpolation of NA Hamiltonians along NA-MD trajectories generated under the classical path approximation.^{21–24} NA-MD modeling can provide key insights into excited state dynamics, whereby the Born–Oppenheimer approximation breaks down as the electronic and nuclear degrees of freedom cannot be adiabatically separated.^{25,26} NA-MD simulations can directly yield macroscopic observables, such as quantum yield, without any prior knowledge of the mechanism, making it a particularly powerful tool for large systems with strong coupling, in which the choice of a reaction coordinate might be unfeasible or

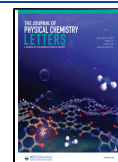
difficult.^{26–33} However, NA-MD simulations typically require intensive *ab initio* calculations of geometry-dependent energies and forces for the different states and the NA coupling (NAC) between them. This drawback has been partially alleviated through the use of ML to predict bandgap and NAC from a small fraction of the data, significantly reducing the computational load for *ab initio* calculations.^{17,19,20,34–37}

Unsupervised ML has been applied on trajectories of NA-MD simulations to elucidate important structural factors in determining the physical properties of materials.^{38–40} One technique is the use of mutual information (MI), which is a measure of the degree of the relationship between data sets composed of large numbers of important features. MI can detect unanticipated relationships between data sets, and its results are easy to interpret. MI is supported by an information-theoretic background and is insensitive to the size of the data set.^{41,42} For example, unsupervised MI uncovered that NAC is determined by geometric features of

Received: October 22, 2021

Accepted: December 9, 2021

Published: December 13, 2021



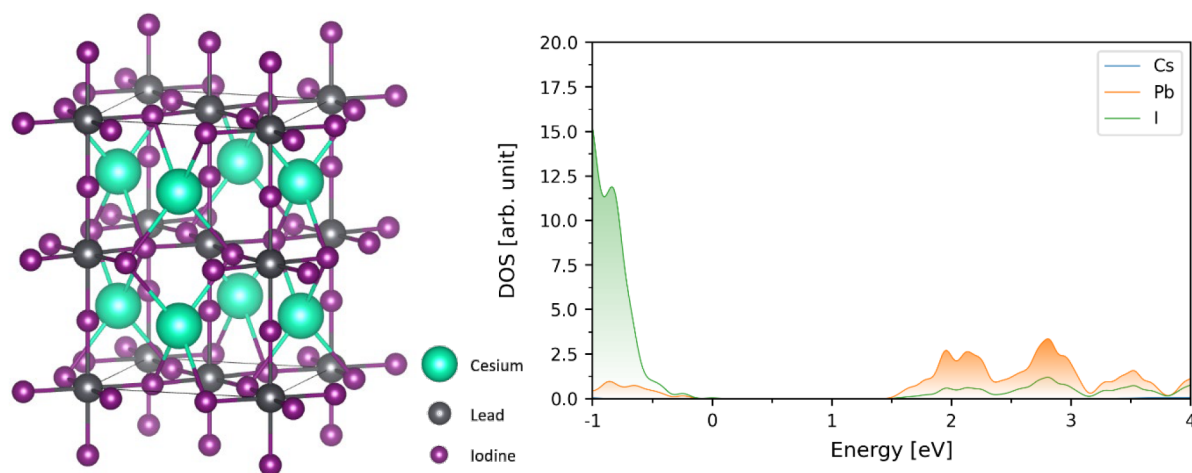


Figure 1. Geometric structure of pristine CsPbI₃ and projected density of states (DOS) for the pristine CsPbI₃ system. The zero energy is set to the VBM.

metal halide perovskites (MHPs) and is less sensitive to atomic velocities, even though it depends explicitly on the latter.³⁸

In this Letter, we used both supervised and unsupervised ML for feature selection and prediction of excitation energies and NACs of the CsPbI₃ MHP for NA-MD simulation. MHPs are a highly popular material for solar energy harvesting and optoelectronics due to their relatively low cost, high optical absorption, and long carrier diffusion.^{43–48} The ability to perform low-cost NA-MD on MHPs can pave the way toward better MHP design, integrating theory and simulation to predict the physical properties of MHPs, instead of the current trial-and-error approach.⁶ We use ML on a data set from NA-MD trajectories of the CsPbI₃ perovskite to establish the most important geometric features that determine the NA Hamiltonian. Then, we evaluate the performance of these features in predicting the bandgap and NAC using a kernel ridge regressor (KRR) model. We demonstrate that information about the chemical environment of any of the three elements present in CsPbI₃ can be used to predict the NA Hamiltonian and that even a single most important feature of a given element is already sufficient for this purpose. These rather surprising results allow us to reduce the dimensionality of the original 360-feature ML model borrowed from the methods used to develop ML force fields to just a 12-feature ML model capable of generating high-quality NA-MD simulation results. The choice of the features is corroborated by the theoretical knowledge of the MHP electronic structure. The reported analysis and significant reduction of model dimensionality help streamline application of ML to accelerate NA-MD simulations used for larger systems and longer processes.

The trajectory data set containing the geometric structures was obtained via ab initio calculations using the Vienna ab initio simulation package (VASP) and the Perdew–Burke–Ernzerhof exchange–correlation functional.^{49–52} The structure shown in Figure 1 was optimized, heated, and equilibrated at room temperature. Next, 7 ps trajectories were generated with a time step of 1 fs in the microcanonical ensemble. Then, 4% of the data, representing 280 data points equally spaced along the trajectory, was selected for training, while the remaining data are used for testing. The NA-MD simulation details are provided in the Supporting Information.

To extract features from the trajectory data set, a symmetry function is used to obtain information regarding the physical and chemical environment of the atoms in the lattice. Here, the symmetry function is adapted from the work of Smith, Isayev, and Roitberg,⁵³ implementing a modified version of Behler and Parrinello's symmetry functions to include radial and angular information.⁵⁴ The following symmetry function is used:

$$G_i^{\text{mod}} = 2^{1-\zeta} \sum_{j,k \neq i}^{\text{atoms}} [1 + \cos(\theta_{ijk} - \theta_s)]^\zeta \times e^{\left[-\eta \left(\frac{R_{ij} + R_{ik}}{2} - R_s\right)^2\right]} f_C(R_{ij}) f_C(R_{ik}) \quad (1)$$

where the index i refers to the atom to which the function is applied, θ_{ijk} refers to the angle centered on atom i , while R_{ij} refers to the distance between atoms i and j . Angular dependence is incorporated via the cosine terms, while radial information is represented via the Gaussian terms. θ_s and ζ are parameters for tuning the center of the angular term and the width of the angular environment, while R_s and η do the same for the Gaussian term. The ζ and η parameters are fixed at 1 and 0.15, respectively, such that the radial and angular terms in eq 1 are kept at similar magnitudes, and both radial and angular features are properly represented. R_s and θ_s are set to the average minimum distance and maximum angles, respectively, of each pair of elements in the crystal lattice, totaling to three R_s and six θ_s values overall. As a result, each atom is represented by 18 different features, for every combination of R_s and θ_s , summing to 360 features in total, with 72 for the cesium and lead atoms and 216 for the iodine atoms in the simulation cell (Figure 1). $f_C(R)$ represents a piecewise cutoff function that defines the size of the approximate local atomic environment and is defined as follows:

$$f_C(R) = \begin{cases} 0.5 \times \cos\left(\frac{\pi R_{ij}}{R_C}\right) + 0.5 & \text{for } R_{ij} \leq R_C \\ 0.0 & \text{for } R_{ij} > R_C \end{cases} \quad (2)$$

In this work, the cutoff radius, R_C , is set of 9.1 Å, corresponding to the distance between the center of the rectangular simulation cell and its vertex.

The pairwise MI, $I(X, Y)$, between each feature and the target variable was calculated on the basis of the estimation of entropy from k -nearest neighbor distances, using $k = 3$, which has been demonstrated to be the optimal parameter for halide perovskites.^{38,41,42}

$$I(X, Y) = \iint dx dy p(x, y) \log \left[\frac{p(x, y)}{p(x)p(y)} \right] \quad (3)$$

where $p(x, y)$ represents the joint probability density function of both variables and $p(x)$ and $p(y)$ represent the marginal probability density function for each variable.

Feature generation, calculation of MI, training, and evaluation of the model are done using the Scikit-learn package⁵⁵ using the Python programming language. All of the models used are kernel ridge regressors using the Laplacian kernel with the L2 penalty set to 0.0001. Additional details are provided in the [Supporting Information](#).

The geometric structure of the perovskite and the projected density of states (DOS), shown in [Figure 1](#), demonstrate that the valence band maximum (VBM) and the conduction band minimum (CBM) are separated by a bandgap of 1.67 eV. The VBM is supported by Pb and I atoms, while the CBM is mainly localized on the Pb atoms. [Table 1](#) presents the results for the

mean MI for each element with the bandgap and NAC. The mean MI was calculated by summing the MI of all features associated with the particular atom type and averaging over the number of atoms of each type. The values agree with the fact that the VBM and CBM are supported by lead and iodine atoms as observed in the PDOS in [Figure 1](#); thus, the MI of lead and iodine is expected to be higher than that of cesium for both the bandgap and NAC between the VBM and CBM. It is also expected that the MI with NAC is smaller than the MI with the bandgap, because the bandgap is calculated from nuclear positions, which are represented by the features, while NAC also depends on velocity, which is not included in the symmetry function. In addition, NAC is a more complex function of geometry than the bandgap, and its correlation with the features should be weaker.

The results of using the KRR model to predict both the bandgap and NAC based on all 18 features, defined by three R_s times six θ_s values (eq 1), are shown in [Figure 2](#). The predictions based on just one element are compared to predictions based on all three elements. Although it is expected that the performance of the model trained on only one element follows the same ranking as that for MI, it is rather surprising that using the features representing the chemical environment of cesium alone ([Figure 2a,b](#)) can give a good fit for both the bandgap and NAC, even though cesium itself does not contribute to the VBM and CBM that determine the bandgap and NAC.

Subsequently, MI was used to select for the optimal R_s and θ_s values for feature selection. This step gives a significant dimensionality reduction and establishes structural properties that are most important for the prediction of the bandgap and

Table 1. Mean Mutual Information (MI) Values for Each Element in CsPbI₃

element	MI with bandgap	MI with NAC
cesium	27.33	26.23
lead	28.15	27.10
iodine	28.33	27.69

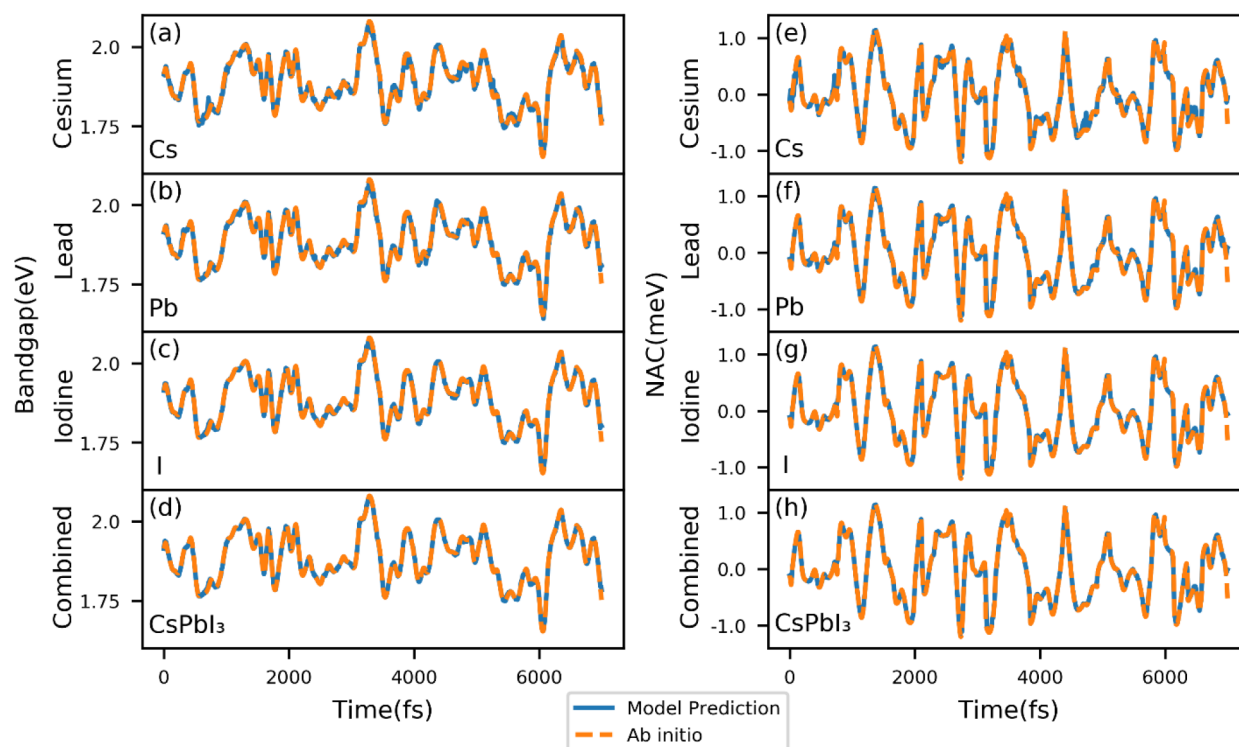
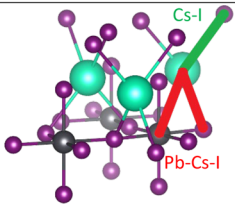
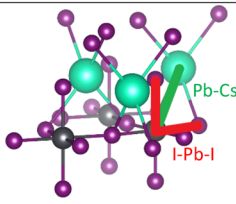
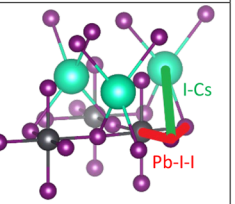
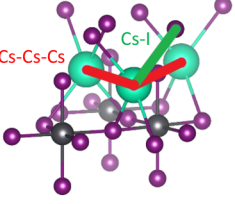
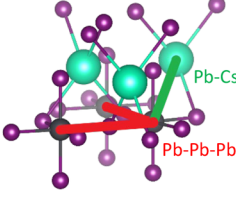
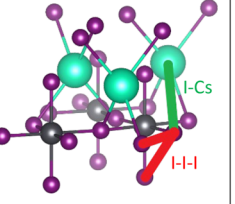


Figure 2. Individual element model prediction vs ab initio values for the bandgap and NAC. (a–c) Models trained using only cesium, lead, and iodine features, respectively, for the prediction of the bandgap. (e–g) Models trained using only cesium, lead, and iodine features, respectively, for the prediction of NAC. (d and h) Models trained using all cesium, lead, and iodine features to predict the bandgap and NAC, respectively.

Table 2. Sets of θ_s and R_s Values with the Highest Mutual Information with the Target Variable

Target	Cs (θ_s , R_s)	Pb (θ_s , R_s)	I (θ_s , R_s)
Bandgap			
NAC			

NAC. The set of optimal R_s and θ_s values for each element is shown in Table 2, and the respective mean MI value for features using that set of R_s and θ_s values is shown in Table 3.

Table 3. Mean Mutual Information (MI) Values for the Optimal Set of θ_s and R_s for Each Element in CsPbI₃

element	MI with the bandgap	MI with NAC
cesium	1.55	1.49
lead	1.61	1.54
iodine	1.59	1.55

As each atom type is represented by only one feature instead of 18, the mean MI values in Table 2 are noticeably smaller than those in Table 1. The ranking of the MI values between each chemical element and NAC is the same as that obtained with all 18 sets of R_s and θ_s values (Table 1); however, lead shares a higher MI with the bandgap than iodine. As described above, cesium has the lowest MI with both the bandgap and NAC, because cesium does not contribute to the VBM and CBM (Figure 1). The MI value for NAC is lower than that of the bandgap (Table 3), similar to that in Table 1. Inspection of the

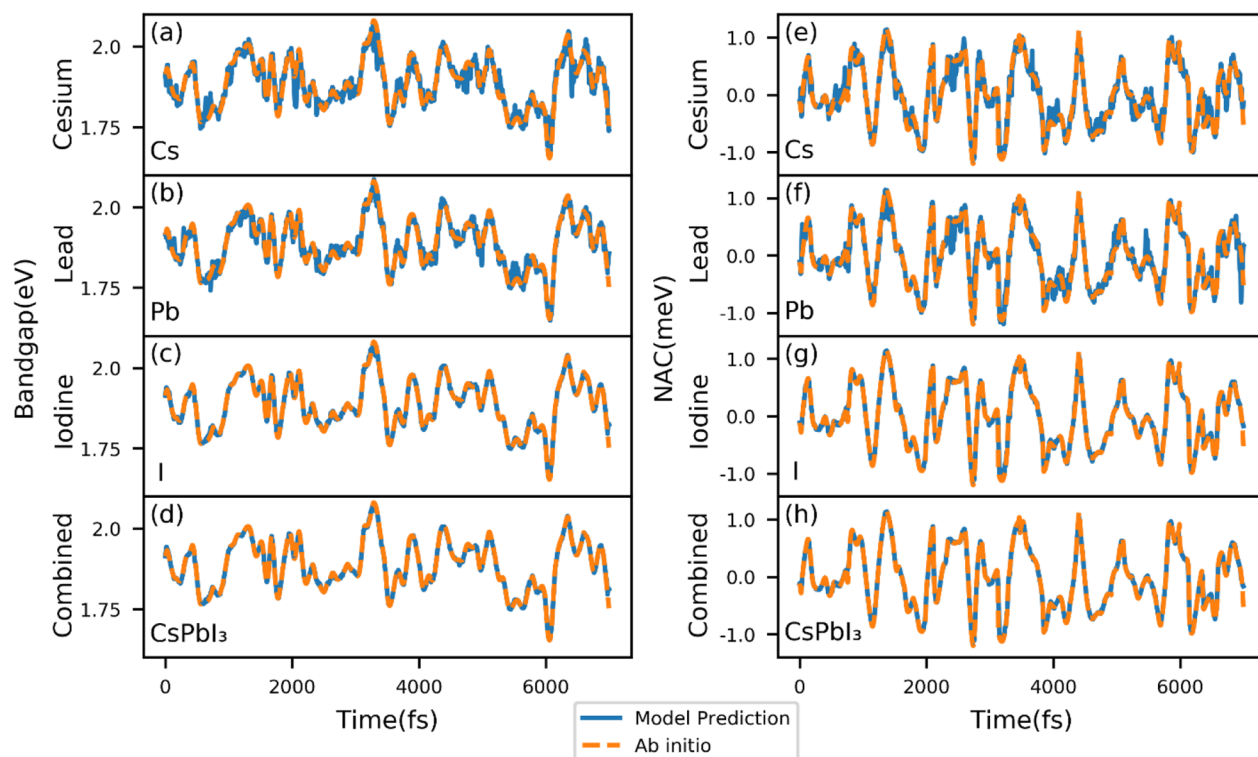


Figure 3. Individual element model prediction using only the set of θ_s and R_s values with the highest mutual information in Table 3 vs ab initio values for bandgap and NAC. (a–c) Models trained using only cesium, lead, and iodine optimal features, respectively, for the prediction of the bandgap. (e–g) Models trained using only cesium, lead, and iodine optimal features, respectively, for the prediction of NAC. (d and h) Models trained using all cesium, lead, and iodine optimal features to predict the bandgap and NAC, respectively.

single most important features for each element (Table 2) indicates that information about atomic distances between cesium and either iodine or lead is the most important for the prediction of the bandgap and NAC. Such distances appear in all six features shown in Table 2. Although surprising, because cesium does not contribute to the VBM or CBM, this observation can be rationalized by the fact that distances involving cesium represent the degree of PbI_6 octahedral tilt and the strong electrostatic interactions between cesium cations and the negatively charged PbI_6 octahedra.

The KRR model predictions of the bandgap and NAC, obtained on the basis of only the features with the optimal R_s and θ_s values, are shown in Figure 3. Panels a–c and e–g show the bandgap and NAC, respectively, predicted using just one feature of one atom type, while panels d and h show predictions based on one feature for all three atom types. Although the performance of the model has decreased compared to that of the model based on all 18 sets of features, the optimal iodine features and the combined optimal features of the three atom types perform very well, with an accuracy comparable to that when all 18 sets of features are used (Figure 2). The model containing 12 parameters, i.e., the optimal feature of the 12 iodine atoms, has managed to perform almost as well as the model containing 360 parameters, involving 18 features from each of the 20 atoms in the lattice (Figure 1). This finding demonstrates a significant dimensionality reduction without much loss of model performance. Although the models using the optimal features of either cesium or lead perform worse than that of iodine, the models can still learn the general trends in the data, and the predictions have a rather good fit with the ab initio values.

Subsequently, we evaluated the performance of the model predictions on NA-MD simulations. The raw results are shown in Figure S3, and the fit parameters are summarized in Table 4.

Table 4. NA-MD Results Obtained Using either the Ab Initio or Predicted Bandgap and NAC^a

NA-MD data	gradient (ns)	intercept ($\times 10^{-3}$)
ab initio	1.12×10^{-1}	−0.47
combined (18 features)	1.02×10^{-1}	1.11
iodine (1 feature)	1.01×10^{-1}	1.95
lead (18 features)	1.00×10^{-1}	1.03
combined (one feature)	0.98×10^{-1}	1.77
iodine (18 features)	0.97×10^{-1}	1.34
cesium (18 features)	0.94×10^{-1}	0.03
lead (one feature)	0.91×10^{-1}	0.28
cesium (one feature)	1.69×10^{-1}	4.90

^aThe simulation calculates the ground state population over 10 ns (Figure S3), and the results are represented via the gradient and intercept of the line of the best fit. The data are ordered from closest to farthest from the ab initio line.

The NA-MD simulations based on the NA Hamiltonians from all KRR models with all sets of features show rather good agreement with the ab initio results, with the exception of the model based on the single most important feature of cesium, in which case the deviation in the gradient, representing the electron–hole recombination rate, is around 50%.

To recapitulate, we have shown that consideration of the chemical environment around one element in CsPbI_3 is sufficient for a satisfactory prediction of the bandgap and NAC values that can lead to accurate NA-MD simulations. Any of

the three elements present in the material can be used for this purpose, including not only Pb and I that support the relevant electronic states but also Cs, which is a spectator ion. This result has allowed a significant reduction in the dimensionality of the ML model of the NA Hamiltonian, decreasing the number of features from 360 that would be used in a typical ML force-field model to just 12, while maintaining the NA-MD simulation quality. We have also demonstrated the use of MI to evaluate the relative importance of the features of different elements, to predict the performance of different ML models, and to determine the optimal structural information to encode in the symmetry function used to define the features. The findings from MI corroborate the theoretical knowledge of the CsPbI_3 electronic properties, further lending credibility to the findings. The dimensionality reduction provided by the analysis helps the application of ML to overcome the high computational cost of ab initio NA-MD and to extend NA-MD simulations to larger, more complex systems and longer time scales.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpclett.1c03469>.

Simulation details, kernel ridge regressor, errors in predicted gaps and nonadiabatic coupling, full mutual information lists, and results of nonadiabatic molecular dynamics simulations (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Oleg V. Prezhdo – Department of Chemical Engineering, Department of Chemistry, and Department of Physics and Astronomy, University of Southern California, Los Angeles, California 90089, United States; orcid.org/0000-0002-5140-7500; Email: prezhdo@usc.edu

Authors

Wei Bin How – Division of Chemistry and Biological Chemistry, School of Physical and Mathematical Sciences, Nanyang Technological University, 637371, Singapore

Bipeng Wang – Department of Chemical Engineering, University of Southern California, Los Angeles, California 90089, United States; orcid.org/0000-0003-0924-5867

Weibin Chu – Department of Chemistry, University of Southern California, Los Angeles, California 90089, United States; orcid.org/0000-0001-5951-0337

Alexandre Tkatchenko – Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg City, Luxembourg; orcid.org/0000-0002-1012-4854

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpclett.1c03469>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work was supported by U.S. National Science Foundation Grant CHE-1900510.

REFERENCES

- (1) Prezhdo, O. V. Advancing Physical Chemistry with Machine Learning. *J. Phys. Chem. Lett.* **2020**, *11* (22), 9656–9658.
- (2) Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C. Machine Learning for Molecular Simulation. *Annu. Rev. Phys. Chem.* **2020**, *71* (1), 361–390.
- (3) Glielmo, A.; Husic, B. E.; Rodriguez, A.; Clementi, C.; Noé, F.; Laio, A. Unsupervised Learning Methods for Molecular Simulation Data. *Chem. Rev.* **2021**, *121* (16), 9722–9758.
- (4) Westermayr, J.; Marquetand, P. Machine Learning for Electronically Excited States of Molecules. *Chem. Rev.* **2021**, *121* (16), 9873–9926.
- (5) Jain, D.; Chaube, S.; Khullar, P.; Goverapet Srinivasan, S.; Rai, B. Bulk and Surface DFT Investigations of Inorganic Halide Perovskites Screened Using Machine Learning and Materials Property Databases. *Phys. Chem. Chem. Phys.* **2019**, *21* (35), 19423–19436.
- (6) Li, J.; Pradhan, B.; Gaur, S.; Thomas, J. Predictions and Strategies Learned from Machine Learning to Develop High-Performing Perovskite Solar Cells. *Adv. Energy Mater.* **2019**, *9* (46), 1901891.
- (7) Moreno, J. R.; Flick, J.; Georges, A. Machine Learning Band Gaps from the Electron Density. *Phys. Rev. Mater.* **2021**, *5* (8), 083802.
- (8) Xu, Q.; Li, Z.; Liu, M.; Yin, W.-J. Rationalizing Perovskite Data for Machine Learning and Materials Design. *J. Phys. Chem. Lett.* **2018**, *9* (24), 6948–6954.
- (9) Yao, K.; Herr, J. E.; Brown, S. N.; Parkhill, J. Intrinsic Bond Energies from a Bonds-in-Molecules Neural Network. *J. Phys. Chem. Lett.* **2017**, *8* (12), 2689–2694.
- (10) Janet, J. P.; Kulik, H. J. Resolving Transition Metal Chemical Space: Feature Selection for Machine Learning and Structure–Property Relationships. *J. Phys. Chem. A* **2017**, *121* (46), 8939–8954.
- (11) Kim, C.; Pilania, G.; Ramprasad, R. Machine Learning Assisted Predictions of Intrinsic Dielectric Breakdown Strength of ABX₃ Perovskites. *J. Phys. Chem. C* **2016**, *120* (27), 14575–14580.
- (12) Botu, V.; Batra, R.; Chapman, J.; Ramprasad, R. Machine Learning Force Fields: Construction, Validation, and Outlook. *J. Phys. Chem. C* **2017**, *121* (1), 511–522.
- (13) Sun, Y. T.; Bai, H. Y.; Li, M. Z.; Wang, W. H. Machine Learning Approach for Prediction and Understanding of Glass-Forming Ability. *J. Phys. Chem. Lett.* **2017**, *8* (14), 3434–3439.
- (14) Keith, J. A.; Vassilev-Galindo, V.; Cheng, B.; Chmiela, S.; Gastegger, M.; Müller, K.-R.; Tkatchenko, A. Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems. *Chem. Rev.* **2021**, *121* (16), 9816–9872.
- (15) Poltavsky, I.; Tkatchenko, A. Machine Learning Force Fields: Recent Advances and Remaining Challenges. *J. Phys. Chem. Lett.* **2021**, *12* (28), 6551–6564.
- (16) Dral, P. O.; Barbatti, M.; Thiel, W. Nonadiabatic Excited-State Dynamics with Machine Learning. *J. Phys. Chem. Lett.* **2018**, *9* (19), 5660–5663.
- (17) Wang, B.; Chu, W.; Tkatchenko, A.; Prezhdo, O. V. Interpolating Nonadiabatic Molecular Dynamics Hamiltonian with Artificial Neural Networks. *J. Phys. Chem. Lett.* **2021**, *12* (26), 6070–6077.
- (18) Chen, W.-K.; Liu, X.-Y.; Fang, W.-H.; Dral, P. O.; Cui, G. Deep Learning for Nonadiabatic Excited-State Dynamics. *J. Phys. Chem. Lett.* **2018**, *9* (23), 6702–6708.
- (19) Hu, D.; Xie, Y.; Li, X.; Li, L.; Lan, Z. Inclusion of Machine Learning Kernel Ridge Regression Potential Energy Surfaces in On-the-Fly Nonadiabatic Molecular Dynamics Simulation. *J. Phys. Chem. Lett.* **2018**, *9* (11), 2725–2732.
- (20) Zhang, Z.; Zhang, Y.; Wang, J.; Xu, J.; Long, R. Doping-Induced Charge Localization Suppresses Electron–Hole Recombination in Copper Zinc Tin Sulfide: Quantum Dynamics Combined with Deep Neural Networks Analysis. *J. Phys. Chem. Lett.* **2021**, *12* (2), 835–842.
- (21) Akimov, A. V.; Prezhdo, O. V. The PYXAID Program for Non-Adiabatic Molecular Dynamics in Condensed Matter Systems. *J. Chem. Theory Comput.* **2013**, *9* (11), 4959–4972.
- (22) Akimov, A. V.; Prezhdo, O. V. Advanced Capabilities of the PYXAID Program: Integration Schemes, Decoherence Effects, Multiexcitonic States, and Field-Matter Interaction. *J. Chem. Theory Comput.* **2014**, *10* (2), 789–804.
- (23) Nijjar, P.; Jankowska, J.; Prezhdo, O. V. Ehrenfest and Classical Path Dynamics with Decoherence and Detailed Balance. *J. Chem. Phys.* **2019**, *150* (20), 204124.
- (24) Zhou, G.; Lu, G.; Prezhdo, O. V. Modeling Auger Processes with Nonadiabatic Molecular Dynamics. *Nano Lett.* **2021**, *21* (1), 756–761.
- (25) Tapavicza, E.; Bellchambers, G. D.; Vincent, J. C.; Furche, F. Ab Initio Non-Adiabatic Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2013**, *15* (42), 18336.
- (26) Long, R.; Prezhdo, O. V.; Fang, W. Nonadiabatic Charge Dynamics in Novel Solar Cell Materials. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2017**, *7* (3), e1305.
- (27) Li, W.; She, Y.; Vasenko, A. S.; Prezhdo, O. V. Ab Initio Nonadiabatic Molecular Dynamics of Charge Carriers in Metal Halide Perovskites. *Nanoscale* **2021**, *13* (23), 10239–10265.
- (28) Stier, W.; Prezhdo, O. V. Nonadiabatic Molecular Dynamics Simulation of Light-Induced Electron Transfer from an Anchored Molecular Electron Donor to a Semiconductor Acceptor[†]. *J. Phys. Chem. B* **2002**, *106* (33), 8047–8054.
- (29) Sarkar, R.; Habib, M.; Kovalenko, S. M.; Pal, S.; Prezhdo, O. V. Mixed Metals Slow Down Nonradiative Recombination in Saddle-Shaped Porphyrin Nanorings: A Time-Domain Atomistic Simulation. *J. Phys. Chem. C* **2021**, *125* (30), 16620–16628.
- (30) Zhang, L.; Chu, W.; Zhao, C.; Zheng, Q.; Prezhdo, O. V.; Zhao, J. Dynamics of Photoexcited Small Polarons in Transition-Metal Oxides. *J. Phys. Chem. Lett.* **2021**, *12* (9), 2191–2198.
- (31) Olson, D. H.; Sales, M. G.; Tomko, J. A.; Lu, T.-F.; Prezhdo, O. V.; McDonnell, S. J.; Hopkins, P. E. Band Alignment and Defects Influence the Electron–Phonon Heat Transport Mechanisms across Metal Interfaces. *Appl. Phys. Lett.* **2021**, *118* (16), 163503.
- (32) Su, J.; Zheng, Q.; Shi, Y.; Zhao, J. Interlayer Polarization Explains Slow Charge Recombination in Two-Dimensional Halide Perovskites by Nonadiabatic Molecular Dynamics Simulation. *J. Phys. Chem. Lett.* **2020**, *11* (21), 9032–9037.
- (33) Banerjee, S.; Kang, J.; Zhang, X.; Wang, L.-W. The Effects of Interstitial Iodine in Hybrid Perovskite Hot Carrier Cooling: A Non-Adiabatic Molecular Dynamics Study. *J. Chem. Phys.* **2020**, *152* (9), 091102.
- (34) Guan, Y.; Yarkony, D. R. Accurate Neural Network Representation of the Ab Initio Determined Spin–Orbit Interaction in the Diabatic Representation Including the Effects of Conical Intersections. *J. Phys. Chem. Lett.* **2020**, *11* (5), 1848–1858.
- (35) Westermayr, J.; Gastegger, M.; Marquetand, P. Combining SchNet and SHARC: The SchNarc Machine Learning Approach for Excited-State Dynamics. *J. Phys. Chem. Lett.* **2020**, *11* (10), 3828–3834.
- (36) Posenitskiy, E.; Spiegelman, F.; Lemoine, D. On Application of Deep Learning to Simplified Quantum-Classical Dynamics in Electronically Excited States. *Mach. Learn. Sci. Technol.* **2021**, *2* (3), 035039.
- (37) Westermayr, J.; Gastegger, M.; Menger, M. F. S. J.; Mai, S.; González, L.; Marquetand, P. Machine Learning Enables Long Time Scale Molecular Photodynamics Simulations. *Chem. Sci.* **2019**, *10* (35), 8100–8107.
- (38) Zhou, G.; Chu, W.; Prezhdo, O. V. Structural Deformation Controls Charge Losses in MAPbI₃: Unsupervised Machine Learning of Nonadiabatic Molecular Dynamics. *ACS Energy Lett.* **2020**, *5* (6), 1930–1938.
- (39) Tavazde, P.; Avendaño Franco, G.; Ren, P.; Wen, X.; Li, Y.; Lewis, J. P. A Machine-Driven Hunt for Global Reaction Coordinates of Azobenzene Photoisomerization. *J. Am. Chem. Soc.* **2018**, *140* (1), 285–290.

- (40) Mangan, S. M.; Zhou, G.; Chu, W.; Prezhdov, O. V. Dependence between Structural and Electronic Properties of CsPbI_3 : Unsupervised Machine Learning of Nonadiabatic Molecular Dynamics. *J. Phys. Chem. Lett.* **2021**, *12* (35), 8672–8678.
- (41) Ross, B. C. Mutual Information between Discrete and Continuous Data Sets. *PLoS One* **2014**, *9* (2), e87357.
- (42) Kraskov, A.; Stögbauer, H.; Grassberger, P. Estimating Mutual Information. *Phys. Rev. E* **2004**, *69* (6), 066138.
- (43) Xiang, W.; Liu, S.; Tress, W. A Review on the Stability of Inorganic Metal Halide Perovskites: Challenges and Opportunities for Stable Solar Cells. *Energy Environ. Sci.* **2021**, *14* (4), 2090–2113.
- (44) Grätzel, M. The Light and Shade of Perovskite Solar Cells. *Nat. Mater.* **2014**, *13* (9), 838–842.
- (45) Green, M. A.; Ho-Baillie, A.; Snaith, H. J. The Emergence of Perovskite Solar Cells. *Nat. Photonics* **2014**, *8* (7), 506–514.
- (46) Ahn, N.; Son, D.-Y.; Jang, I.-H.; Kang, S. M.; Choi, M.; Park, N.-G. Highly Reproducible Perovskite Solar Cells with Average Efficiency of 18.3% and Best Efficiency of 19.7% Fabricated via Lewis Base Adduct of Lead(II) Iodide. *J. Am. Chem. Soc.* **2015**, *137* (27), 8696–8699.
- (47) Wei, Z.; Xing, J. The Rise of Perovskite Light-Emitting Diodes. *J. Phys. Chem. Lett.* **2019**, *10* (11), 3035–3042.
- (48) Cai, T.; Shi, W.; Hwang, S.; Kobbekaduwa, K.; Nagaoka, Y.; Yang, H.; Hills-Kimball, K.; Zhu, H.; Wang, J.; Wang, Z.; Liu, Y.; Su, D.; Gao, J.; Chen, O. Lead-Free $\text{Cs}_4\text{CuSb}_2\text{Cl}_{12}$ Layered Double Perovskite Nanocrystals. *J. Am. Chem. Soc.* **2020**, *142* (27), 11927–11936.
- (49) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1999**, *59* (3), 1758–1775.
- (50) Kresse, G.; Hafner, J. *Ab Initio* Molecular Dynamics for Liquid Metals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *47* (1), 558–561.
- (51) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54* (16), 11169–11186.
- (52) Kresse, G.; Hafner, J. *Ab Initio* Molecular-Dynamics Simulation of the Liquid-Metal–Amorphous-Semiconductor Transition in Germanium. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1994**, *49* (20), 14251–14269.
- (53) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost. *Chem. Sci.* **2017**, *8* (4), 3192–3203.
- (54) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 146401.
- (55) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.