

Increasing Efficiency of Nonadiabatic Molecular Dynamics by Hamiltonian Interpolation with Kernel Ridge Regression

Yifan Wu, Natalie Prezhdo, and Weibin Chu*



Cite This: *J. Phys. Chem. A* 2021, 125, 9191–9200



Read Online

ACCESS |



Metrics & More

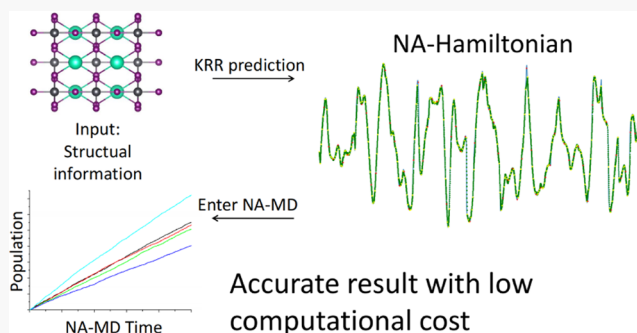


Article Recommendations



Supporting Information

ABSTRACT: Nonadiabatic (NA) molecular dynamics (MD) goes beyond the adiabatic Born–Oppenheimer approximation to account for transitions between electronic states. Such processes are common in molecules and materials used in solar energy, optoelectronics, sensing, and many other fields. NA-MD simulations are much more expensive compared to adiabatic MD due to the need to compute excited state properties and NA couplings (NACs). Similarly, application of machine learning (ML) to NA-MD is more challenging compared with adiabatic MD. We develop an NA-MD simulation strategy in which an adiabatic MD trajectory, which can be generated with a ML force field, is used to sample excitation energies and NACs for a small fraction of geometries, while the properties for the remaining geometries are interpolated with kernel ridge regression (KRR). This ML strategy allows for one to perform NA-MD under the classical path approximation, increasing the computational efficiency by over an order of magnitude. Compared to neural networks, KRR requires little parameter tuning, saving efforts on model building. The developed strategy is demonstrated with two metal halide perovskites that exhibit complicated MD and are actively studied for various applications.



1. INTRODUCTION

Machine learning (ML) techniques have experienced rapid and extensive developments in recent decades, stimulated by growing computational power and large amounts of experimental and theoretical data accumulated in all fields, including science, technology, and daily life. Many efforts are devoted to applying ML to chemistry and materials science to predict relevant properties,^{1–5} identify promising candidates for various applications,^{6–9} and guide and design ongoing experiments.^{9,10} With the help of ab initio-calculated data, ML is successfully practiced for predicting molecular properties.^{11–16} Prediction of time-series data, such as forces on atoms in molecular dynamics (MD) simulation, can alleviate the burden of the computational cost of quantum mechanical calculations.¹⁷ The key to the success stems from the fact that ML models can quantitatively estimate the behavior in an unknown realm by learning the pattern from an existing data set. The accumulated results of calculations and experiments provide an abundance of data that serve as training sets necessary for ML model prediction, giving sufficient resources for learning.¹⁸ While many time-series studies focus on extrapolation of the properties of interest, our strategy is to extrapolate the simpler properties, such as ground state forces, and to interpolate the more complex properties, such as the nonadiabatic coupling (NAC).

Ab initio calculations suffer from high computational cost if applied to relatively big systems and at high levels of theory.

This is particularly true of NA MD (NA-MD) simulations that require thousands of ab initio electronic structure calculations, including not only ground state electronic properties but also excited state properties and NAC.^{19–24} NA-MD is capable of modeling quantum dynamics of molecular, condensed matter, and nanoscale systems that undergo transitions between electronic states. NA-MD simulations imitate closely time-resolved spectroscopy experiments at the atomistic level and in the time domain in systems that form the basis for modern technologies, such as solar cells, photocatalysts, light-emitting diodes, optoelectronic devices, and so forth. NA-MD simulations generate MD trajectories along which one needs to calculate excited state energies and NAC between the states, sometimes including hundreds of thousands of states.^{25,26} Depending on the processes under consideration, one may also need to compute the Coulomb matrix elements²⁷ and spin–orbit coupling.^{28,29} Such calculations are time-consuming, especially for large, nanoscale systems involving long processes, such as formation of large polarons³⁰ and defect diffusion.³¹

Received: June 9, 2021

Revised: September 23, 2021

Published: October 12, 2021



ACS Publications

© 2021 American Chemical Society

9191

<https://doi.org/10.1021/acs.jpca.1c05105>
J. Phys. Chem. A 2021, 125, 9191–9200

Bypassing some of the computationally demanding electronic structure calculations can provide great benefits for NA-MD.

Modeling of charge carrier dynamics in metal halide perovskites (MHPs) presents an important and modern example of NA-MD simulation.^{24,32–37} MHPs, such as CsPbI₃, constitute an actively investigated class of solar energy and optoelectronic materials, due to high power conversion efficiency.^{38–41} MHPs possess many favorable attributes, such as high carrier mobility and lifetime, low exciton binding energy, tunable band gap, easy fabrication, and low cost.^{41–43} Both hybrid organic–inorganic and all-inorganic MHPs have been synthesized and extensively studied. Pb-based MHPs generally show higher power conversion efficiencies compared to non-lead MHPs, while non-lead perovskites avoid the toxicity of lead.^{44,45} MHPs exhibit many unusual physical and chemical properties, and intricate MD with large-amplitude anharmonic motions, slow diffusion of atoms and defects, long-range electric ordering and disordering, phase transitions, and so forth. Modeling excited state dynamics in MHPs is a challenging task for NA-MD, requiring further methodological developments and providing a strong motivation for the current work.

Development of ML force fields (FF) is becoming a routine task and has been achieved for many classes of molecules and materials.^{1,3,11} It is more challenging to develop ML models for NAC.^{46–49} The adiabatic force and the NAC arise from the same quantum mechanical nuclear derivative operator. However, the force is its diagonal matrix element, while the NAC is the off-diagonal element. The NAC exhibits a more complex dependence on atomic geometry than the force and can fluctuate wildly, diverging to positive or negative infinity as the energy gap between two states approaches zero. In addition, the number of NAC matrix elements scales quadratically with the number of electronic states, while the number of forces scales linearly.

The classical path approximation^{50,51} (CPA) provides significant computational savings for NA-MD because it allows for one to replace multiple excited state MD trajectories with a single, typically ground state trajectory. The CPA is valid in the cases when MD is driven by thermal fluctuations, and the differences in the ground and excited state geometries are insignificant compared to the fluctuation amplitude. Many classes of systems obey such conditions.^{24,35,52–57} Under the CPA, one can use a ML FF to generate the trajectory, use a small fraction of geometries from the trajectory to calculate the excitation energies and NACs, and develop a ML model that interpolates the excitation energies and NACs between the sampled points. This task is notably easier than time extrapolation because sampling along the trajectory ensures that relevant geometries are present in the training set. In contrast, sampling from a short MD trajectory may not represent all geometries needed for extrapolation of the trajectory to longer times. Already challenging for development of ground state ML FFs, the task becomes much harder for excited state forces and NACs.

In this paper, we employ kernel ridge regression (KRR) to predict excitation energies and NACs for NA-MD simulations under the CPA. We assume that a MD trajectory can be obtained efficiently using a ML or other FF and demonstrate an over an order-of-magnitude computational speed up by computing the NA-MD Hamiltonian using density functional theory (DFT) only for a small fraction of geometries along the trajectory and interpolating the remaining excitation energies

and NACs with KRR. The strategy is illustrated with two MHP systems with and without a common defect. The data show that KRR is capable of representing accurately the full NA-MD Hamiltonian with only 4% of the sampled points, and even 0.8% sampling gives reasonable results. Importantly, the NA-MD simulations based on the KRR NA-MD Hamiltonian converge systematically to the DFT data. Generally, the KRR-based simulation underestimates the NAC and transition time because KRR misses some of the NAC peaks that contribute strongly to the transition rate. Therefore, further progress can be achieved either by improving sampling of the NAC peaks, for example, by considering velocity of the atoms that contribute to NAC or by extrapolating NAC and rate values obtained at different levels of sampling. The over an order-of-magnitude speedup provided by the developed strategy allows for one to extend NA-MD simulations to larger systems and longer timescales, accessing new classes of systems and processes.

2. THEORY

Combining the simplicity of the least-squares solution of linear regression and the kernel handling of nonlinear problems gives birth to KRR, making KRR a simple but powerful model.⁵⁸ Requiring tuning of only the kernel function and a regularization term, KRR has notably fewer hyperparameters compared to neural networks (NNs), which need tuning of the number of hidden layers, neurons per layer, the way how neurons connect, learning rate, and so forth. Further, KRR only needs a simple matrix operation to find the optimal solution of the model, while NNs involve gradient descent, which may lead to a local rather than global minimum. KRR enables a simple ML-assisted NA-MD simulation. Given a trajectory, we sample excitation energies and NACs for a small proportion (0.8, 2, 4, and 8%) of points along the trajectory and employ KRR to predict the remaining missing values.

The DFT calculations, including geometry optimization and adiabatic MD, are performed with the Vienna ab initio Simulation Package (VASP).⁵⁹ The NACs are calculated from the overlap of the two wavefunctions at adjacent time steps using the CA-NAC package^{60,61}

$$\begin{aligned} d_{ji} &= -i\hbar\varphi_j(r, R(t))|\nabla_R|\varphi_i(r, R(t))\frac{dR}{dt} \\ &= -i\hbar\frac{\varphi_j(r, R(t))|\nabla_R H(R(t))|\varphi_i(r, R(t))}{E_i - E_j}\frac{dR}{dt} \\ &= -i\hbar\varphi_j(r, R(t))\left|\frac{\partial}{\partial t}\right|\varphi_i(r, R(t)) \\ &\approx -\frac{i\hbar}{2\Delta t}\{\varphi_j(r, R(t))|\varphi_i(r, R(t + \Delta t)) \\ &\quad - \varphi_j(r, R(t + \Delta t))|\varphi_i(r, R(t))\} \end{aligned} \quad (1)$$

While the abovementioned expression is written in terms of the Kohn–Sham (KS) orbitals, the calculations are performed using the many-particle Slater determinant basis, and the NAC between Slater determinants is reduced to the NAC between KS orbitals, as detailed in ref 50.

The Perdew–Burke–Ernzerhof exchange–correlation functional⁶² is used. The projected-augmented wave method is employed to handle interactions between electrons and ion cores.⁶³ After geometry relaxation at 0 K, the pristine and

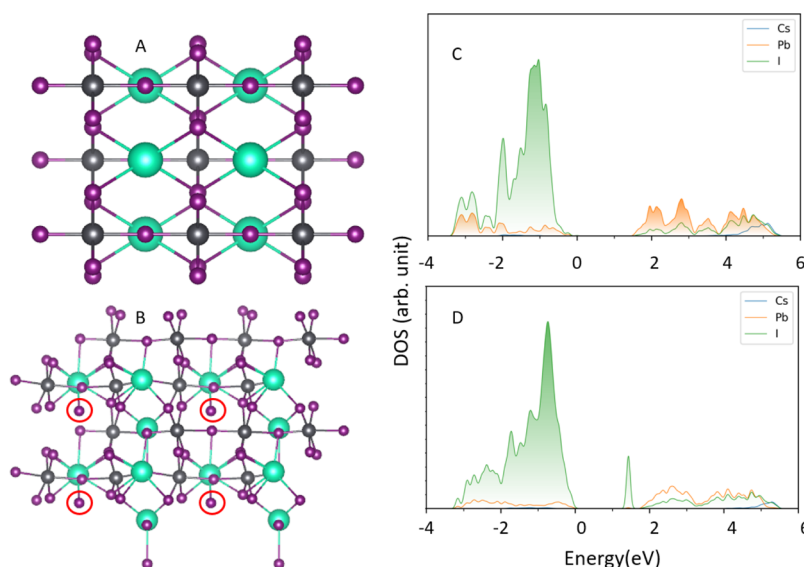


Figure 1. Geometric structures of (A) pristine CsPbI_3 and (B) CsPbI_3 with the I_{Cs} defect. The I atom replacing Cs is marked by the red circle. (C,D) Corresponding projected DOS, with the zero-energy set at the VBM.

defective MHPs are heated to 300 K, reaching thermal equilibrium. Then, 7000 fs adiabatic MD trajectories are produced in a microcanonical ensemble with a 1 fs timestep. We employ the DISH method²³ to simulate the nonradiative charge trapping and recombination processes in the MHPs. The pure-dephasing times used in DISH are obtained via the second-order cumulant approximation of the optical response theory.^{64,65} The NA-MD simulations are carried out with the Pyxaid software.^{50,51}

The calculations are performed with pristine CsPbI_3 and CsPbI_3 containing a I_{Cs} defect, in which a cesium atom is replaced by an iodine atom, as shown in Figure 1. The tetragonal phase is represented using a $(\sqrt{2} \times \sqrt{2} \times 2)$ simulation cell of the cubic phase with the lattice constants $a = 9.02$ Å and $c = 12.76$ Å. Aiming at accomplishing a good prediction precision with a minimum amount of training data, we train and evaluate the KRR models with 0.8, 2, 4, and 8% of the DFT data serving as the training set, and, respectively, 0.2, 0.5, 1, and 2% of the DFT data used as the validation set, with the rest of the data used for testing. The chosen training/validation data set are equally spaced along the trajectory. We demonstrate the results based on the 4% training set in detail, while the remaining results are presented in the [Supporting Information](#). Specifically, 5% of the 7000 data points are chosen and split into 4% (280 data points) as the training set and 1% (70 data points) as the validation set, while the remaining 95% (6750 data points) is used as the testing set. The 280 + 70 points employed for training and validation are equally spaced every 20 fs. Every fifth of the chosen points is assigned to validation, and the remaining points are the training set. The DFT data are provided in the [Supporting Information](#).

All KRR training, validation, and testing are carried out with the Scikit-learn package.⁶⁶ The kernel is used to map the original data to a higher-dimensional space. Therefore, the linear model can fit nonlinear data in a more complicated space. Laplacian and cosine similarity kernels have been tested. The Laplacian kernel has been found to give a better overall performance, while the cosine similarity kernel gives better NAC peaks when training is performed with very small

sampling, 0.8%. Though we value the prediction of the NAC peaks because the rate of an electronic transition increases when the NAC is large, and according to Fermi's golden rule the rate depends on NAC squared,⁶⁷ we still wish to achieve a good overall precision. Therefore, the Laplacian kernel is selected to build the KRR model. The Laplacian kernel is defined as

$$K(x, y) = e^{-\gamma \|x - y\|_1} \quad (2)$$

where x and y are the input vectors and γ is a hyperparameter that needs to be tuned. Tuning the hyperparameter is a common procedure of training any ML model because the hyperparameters cannot be acquired from training data. After tuning the hyperparameter γ for a few different values, we realized that $\gamma = 1/(\text{number of features})$ shows relatively better results for most of the models, compared to other tested hyperparameters. For simplicity and consistency, we set $\gamma = 1/(\text{number of features})$. $\|x - y\|_1$ is the Manhattan distance between the input vectors.

Overall, there are four models for the NAC prediction and four models for the energy gap prediction for the pristine system, corresponding to 8, 4, 2, and 1% of the data used for the training. The defect system includes three NA transitions, and there are 12 models for the NAC prediction and 12 models for the energy gap prediction. All models share the same hyperparameter set. It is possible to tune the hyperparameters separately for each model. However, we have achieved good accuracy with 4% training data, and the improvement is not necessary. Therefore, we keep the hyperparameters the same for all models for simplicity and consistency.

KRR requires input vectors containing features. Currently, the features are obtained from atom positions, enabling portrayal of an accurate system structure. Many efforts have been made to extract the structural information, while balancing the accuracy of description of structural details and the redundancy.^{68–70} The modified symmetry function¹ with the following form is used to capture the individual atom's radial and angular information in a specific chemical environment and to record contributions from all atoms.

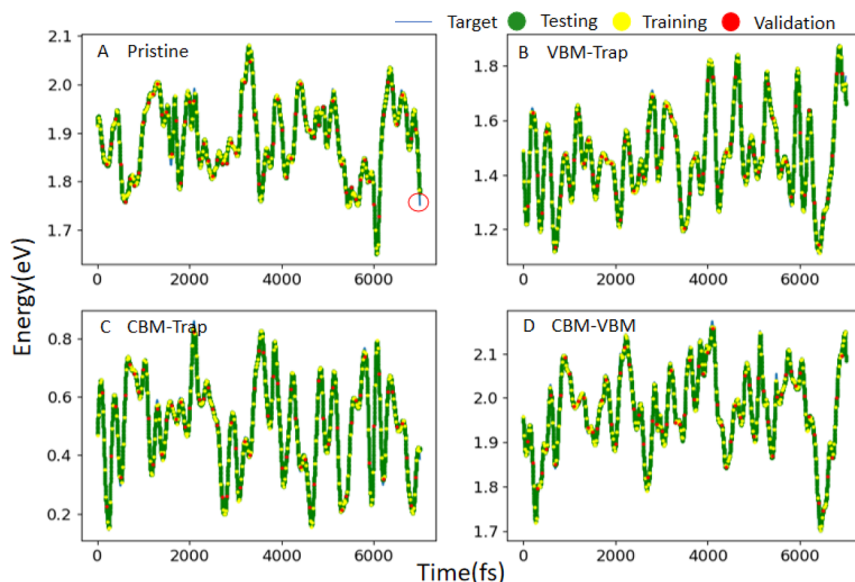


Figure 2. Kernel ridge regression predicted energy gaps for (A) pristine CsPbI₃ and (B–D) CsPbI₃ with the I_{Cs} defect. The training is performed with 4% of the data. The deviation at the end of the trajectory for the pristine system, indicated by the circle, is discussed in the text.

$$G_i^{\text{mod}} = 2^{1-\zeta} \sum_{j,k \neq i}^{\text{atoms}} (1 + \cos(\theta_{ijk} - \theta_s))^\zeta \times e^{\left[-\eta \left(\frac{R_{ij} + R_{ik}}{2} - R_s\right)^2\right]} \times f_C(R_{ij})f_C(R_{ik}) \quad (3)$$

The symmetry function consists of a cosine part and a Gaussian part. Considering one central atom, there are three different distances and six angles. θ_{ijk} is the angle between three atoms with the i atom standing in the center. R_{ij} is the distance between atoms i and j . R_s and θ_s are the average distances and angles for each pair ij and triplet ijk of elements, respectively. Taking Pb as an example, R_s are the Pb–Pb, Pb–Cs, and Pb–I distances averaged over all MD time steps, only considering the nearest neighbors. For θ_s , given Pb as the central atom, the angles are Cs–Pb–Pb, Cs–Pb–I, Cs–Pb–Cs, I–Pb–Pb, I–Pb–I, and Pb–Pb–Pb. Atoms of the same type are grouped together. The R_s and θ_s values are computed separately for the pristine and defect structures, and the defect atom is assigned its own atom type. The KRR models are invariant with respect to permutations of atoms within the same type.

Given that there are 4 Cs, 4 Pb, and 12 I atoms in the simulation cell of the pristine system, we obtain a total of 72 features for the Cs and Pb atoms ($4 \times 3 \times 6$) and 216 features for the I atoms ($12 \times 3 \times 6$). In the defective systems, the features associated with one Cs atom become the features of the defect I atom that replaces the Cs atom. The input vectors defined in eq 2 are the feature vectors. Each geometry from the MD trajectory gives one input vector, containing 360 features coming from the Cs, Pb, and I atoms.

The cutoff function f_C that enters the symmetry function, eq 3, ensures chemical locality and reduces the computational cost

$$f_C(R) = \begin{cases} 0.5 \times \cos\left(\frac{\pi R_{ij}}{R_C}\right) + 0.5 & \text{for } R_{ij} \leq R_C \\ 0.0 & \text{for } R_{ij} > R_C \end{cases} \quad (4)$$

ζ and η control the magnitudes. They are set to 1 and 0.15, respectively, making the angular and radial terms similar in magnitude. The cutoff radius $R_C = 9.1$ Å is set to be the distance between the center of the simulation cell and its vertex.

During the training, the algorithm searches for an optimal solution of the weights to best describe the provided training set. Then, the validation and testing sets are used to compute the following measure, demonstrating that the obtained optimal weights and bias provide a good prediction.

$$J_w = \min_w \{ \|Xw - y\|^2 + \alpha \|w\|^2 \} \quad (5)$$

Here, X are the feature vectors, w are the corresponding weights, and y are the input vectors of either the NAC or the energy gap. αw^2 is an L2 regularization term. KRR uses an L2-norm regularization term to prevent overfitting. The penalty for L2 is set to 0.0001 for the pristine and defective systems. The optimal weights can be expressed as

$$w^* = (X^T X + \alpha I)^{-1} X^T y \quad (6)$$

where I is the identity matrix. Application of the kernel technique simplifies computing the inner product $X^T X$ and fitting of the model. The hyperparameters γ and α used for the pristine and defect structures are the same.

Equations 3 and 4 generate the features that are used as input vectors to the KRR model. Introducing the L2 regularization term is used to prevent overfitting and to take care of the collinearity during multiplication of the matrix of the input vectors and their transpose, when solving for the optimal weights of the model. Introduction of the kernel function is needed to map the features to a higher dimensional space, such that the model is able to fit nonlinear data. Generally, KRR only requires tuning of the kernel type, the hyperparameter inside the kernel function, and the penalty for L2. This is quite straightforward and bypasses the complexity of building layers of NNs.

3. RESULTS AND DISCUSSION

To demonstrate the KRR-assisted NA-MD method, we utilize pristine CsPbI₃ and CsPbI₃ containing the I_{Cs} defect, as shown in Figure 1A,B. The calculated element projected densities of states (DOS) are shown in Figure 1C,D. The direct band gap of pristine CsPbI₃ is 1.67 eV. The defective system introduces a trap state inside the band gap, separated from the valence band minimum (VBM) and the conduction band maximum (CBM) by 1.27 and 0.15 eV, respectively. The trap state is empty and is much closer to the CBM than the VBM, indicating that it is an electron trap. The VBM in pristine CsPbI₃ arises from I atoms, while the CBM is localized on Pb atoms, as shown in Figure 1C. I and Pb support the VBM and the CBM, while the defect state is localized on I atoms. The charge densities shown in Figure S1 support these conclusions. Interestingly, the defect charge density is delocalized over many I atoms, as shown in Figure S1D.

Figure 2 shows the DFT and KRR predicted energy gaps between the VBM, the CBM, and the electron trap state in the pristine and defective systems. Using 4% of the data for training successfully captures the energy gap fluctuations, showing good agreement with the DFT results for both systems. Using 8% of the data reproduces all peaks and fluctuations, as shown in Figure S2. Training with 2% of the data fails to reproduce some peaks, though it exhibits a good overall precision, as shown in Figure S3. The errors in the calculated average energy gaps in pristine CsPbI₃ are 0.053 and 0.106% for the 4 and 2% training, respectively. The 8% training converges to the DFT result, as shown in Table 1.

Table 1. KRR and DFT RMS NAC, Average Absolute (Abs) NACs, Fundamental Band gap, Pure-Dephasing Times, and Charge Recombination Times in Pristine CsPbI₃ Obtained with 8, 4, 2, and 0.8% Training Sets

pristine	RMS-NAC (meV)	Abs-NAC (meV)	gap (eV)	dephasing (fs)	recombination (ns)
8%	0.534	0.445	1.888	8.53	93.5
4%	0.522	0.435	1.887	8.67	98.4
2%	0.490	0.410	1.886	8.73	123.7
0.8%	0.403	0.330	1.875	7.74	69.1
DFT	0.538	0.448	1.888	8.49	90.8

Note that the KRR results deviate from the DFT data at the very end of the 7000 fs trajectory, as shown in Figure 2A. It is because the last few femtoseconds are not included in the training set. This observation highlights the difficulty of time-extrapolating the NA-MD Hamiltonian. The 7000 fs trajectory is not sufficient to sample all geometries that are possible in CsPbI₃ at 300 K. In contrast, interpolating the gap works very well, demonstrating the efficiency of the current strategy. In applications, the starting and ending points of the MD trajectory can be included into the training set manually to obtain a better performance. The interpolation works very well in general. Even the 0.8% sampling captures most of the points, as shown in Figure S4. Time extrapolation requires much longer training trajectories and is likely to predict the gap values only for short times into the future. The interpolation strategy works well because KRR is able to interpolate complicated continuous functions, offering large computational savings.

Generally, the NAC has a more complicated dependence on geometry than the energy gap. Nevertheless, the KRR model

with the 4% sampling successfully obtains an accurate prediction of the NAC, capturing all the peaks and fluctuations, as shown in Figure 3. Predicting NAC peaks is important because electronic transitions are more likely to happen at these timesteps. The NAC may increase and decrease dramatically in a few timesteps, showing larger fluctuations than energy gaps. An example is seen with the CBM-trap NAC early in the trajectory in Figure 3C. The NAC increases because the corresponding energy gap approaches zero, as shown in Figure 2C, and the NAC is inversely proportional to the energy gap, eq 1. The NAC also depends on the overlap of the initial and final wavefunctions and is sensitive to atomic geometry. That is why the NAC is not as large at around 4500 fs, when the CBM-trap energy gap also becomes small.

The errors in the root-mean-square (RMS) NAC computed with KRR, relative to DFT, are 0.74, 2.97, and 8.92% for the 8, 4, and 2% sampling in pristine CsPbI₃, respectively, as shown in Table 1. The RMS NAC is calculated using the standard

expression, $\text{RMS} = \sqrt{\frac{\sum_{k=1}^N (d_{ji}^k)^2}{N}}$, where d_{ji}^k is the NAC, eq 1,

for the timestep k , and N is the total number of timesteps. Compared to the energy gap prediction, the NAC shows larger errors due to a more complex dependence on the system structure. KRR shows good agreement with the DFT calculation even when the NAC exhibits large fluctuations, indicating that the NA-MD simulations based on the KRR NAC should give good results.

The KRR results for the energy gaps and NACs in both pristine and defective systems are obtained with the same hyperparameters, reducing the hyperparameter tuning efforts, as discussed in the Theory section. The data are based on the Laplacian kernel, eq 2. The cosine similarity kernel, $K(x,y) = xy^T/xy$, was used as well, but it showed poor prediction for the CBM-trap NAC in defective CsPbI₃ owing to the large jump of the NAC at the early time, as shown in Figure 3C. To fit the jumps of the CBM-trap NAC, the weights of the model with the cosine similarity kernel should be generally larger than the overall optimal values. In such a case, the NAC after the jump is overestimated. In comparison, the Laplacian kernel works fine, both capturing the peaks and giving a reasonable NAC prediction after the peaks. Note that the deviation between the KRR and DFT NAC is at the very end of the 7000 fs trajectory for the pristine system, as shown in Figure 3A. The deviation is similar to that seen in Figure 2A. It arises because the last few femtoseconds are not included into the training set.

Next, we computed Fourier transforms (FTs) of the ACF of the energy gap and NACs, as shown in Figure 4. The ACF of the energy gap is defined as

$$C_{ij}(t) = \delta E_{ij}(t') \delta E_{ij}(t - t')_t \quad (7)$$

where δE_{ij} is the deviation of the energy gap E_{ij} between states i and j from its canonically averaged value. The FT, known as the influence spectrum or the spectral density, is defined as

$$I_{ij}(\omega) = \int_{-\infty}^{+\infty} dt e^{i\omega t} C_{ij}(t) \quad (8)$$

The NAC ACFs and FTs are computed the same way. The FTs characterize the phonon modes that couple to the electronic subsystem. The peak heights indicate the strength of the electron–phonon coupling for the phonon modes of the corresponding frequencies. The 4% KRR prediction is nearly identical to the DFT data, with the NAC FTs showing slightly

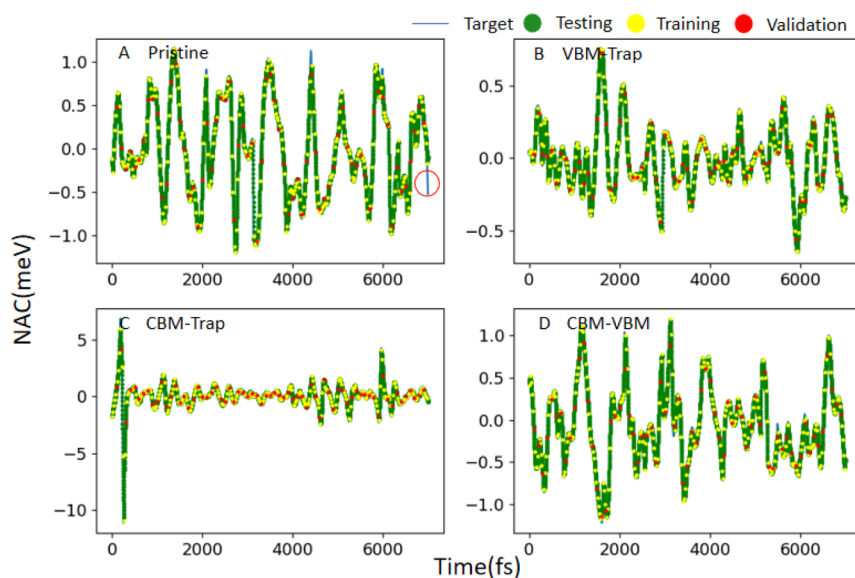


Figure 3. Kernel ridge regression predicted NAC for (A) pristine CsPbI₃ and (B–D) CsPbI₃ with the I_{Cs} defect. The training is performed with 4% of the data. The deviation at the end of the trajectory for the pristine system, indicated by the circle, is discussed in the text.

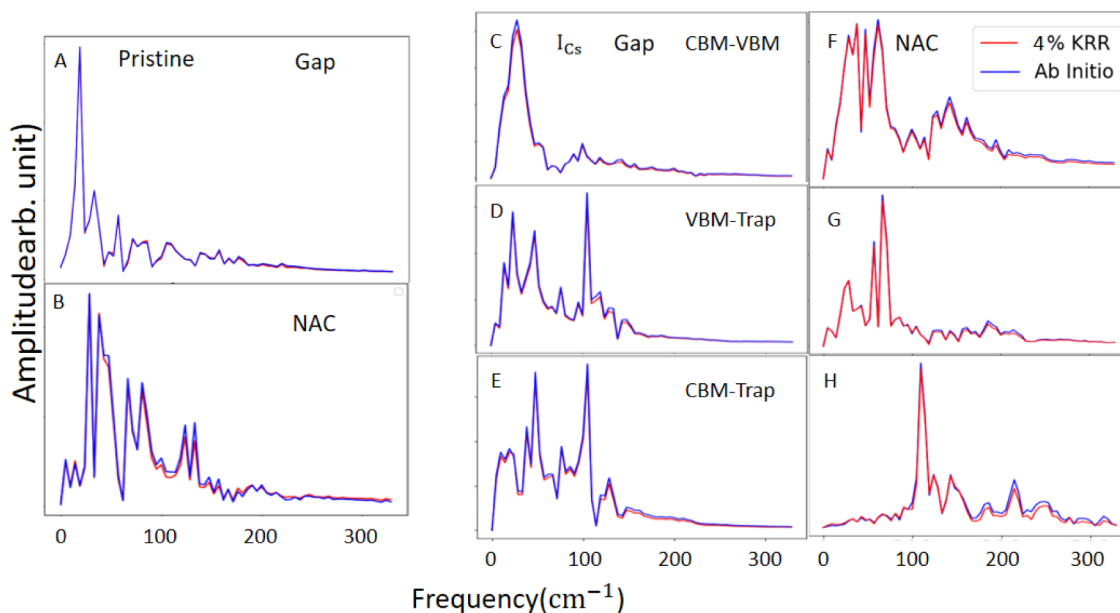


Figure 4. FTs, eq 8, of autocorrelation functions (ACF), eq 7, of the energy gaps and NACs in (A,B) pristine CsPbI₃ and (C–H) CsPbI₃ containing the I_{Cs} defect.

more noticeable differences than the gap FTs. The NAC FTs exhibit higher frequency signals than the gap FTs. This is because the NAC depends not only on the gap but also on other factors, eq 1. Generally, the NAC is a higher order property than the energy gap, and it is more sensitive to changes in the wavefunctions.

The NAC reflects inelastic electron–phonon scattering that leads to energy exchange between the two subsystems. Elastic scattering is quantified by the pure-dephasing time, which represents loss of coherence within the electronic subsystem due to coupling to phonons. The pure-dephasing times are calculated based on the energy gap fluctuations using the second-order cumulant approximation of the optical response theory.^{64,65} The pure-dephasing functions

$$D_{ij}(t) = \exp\left(-\frac{1}{\hbar^2} \int_0^t dt' \int_0^{t'} dt'' C_{ij}(t'')\right) \quad (9)$$

are computed based on the energy gap ACF, eq 7, and are fitted to Gaussians, $\exp[-0.5(-t/\tau)^2]$, to obtain the pure-dephasing times, τ . The KRR pure-dephasing times converge to the DFT values, similarly to the other properties, as shown in Tables 1 and S1.

The pure-dephasing time is an important property on its own right. It can be measured using photon-echo experiments and obtained as an inverse of the single chromophore optical linewidth.^{64,65} In addition, pure-dephasing provides an estimate of the decoherence process that should be included into NA-MD simulations^{23,71,72} when loss of coherence is

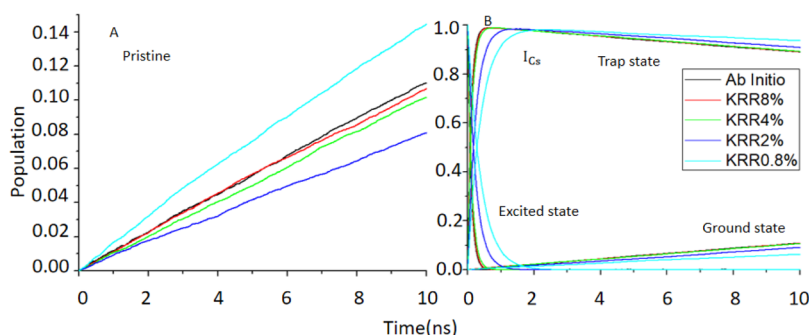


Figure 5. (A) Population of the ground state in pristine CsPbI₃ and (B) populations of the excited, ground, and trap states in CsPbI₃ containing the I_{Cs} defect, obtained by NA-MD simulation using DFT energy gaps and NACs, and corresponding KRR results for different training set sizes.

faster than the quantum transition, which is true in the present case.

The energy gaps and NACs predicted by KRR define the KRR Hamiltonian. It is used to perform the NA-MD simulations and to compare the results to those obtained with the NA-MD Hamiltonian from DFT. Figure 5 presents the NA-MD results for the pristine and defective systems, obtained with the DFT and KRR Hamiltonians. Shown are evolutions of populations the ground state in pristine CsPbI₃ and the excited, ground, and trap states in defective CsPbI₃. The charge recombination times, τ , are obtained by fitting the data to the short-time, linear approximation to the exponential evolution, $1 - \exp(-t/\tau) \approx t/\tau$. The KRR results converge to the DFT data with increasing sampling. KRR with the 4% sampling shows good agreement with the DFT data, while the 2 and 0.8% data deviate more. Even the 0.8% sampling gives meaningful estimates of the charge trapping and recombination timescales. The 4% sampling provides a good balance between the computational cost and accuracy.

Tables 1 and S1 summarize the DFT and KRR prediction-based properties for the pristine and defective systems, respectively. The KRR quantities are obtained the same way as the DFT results, using the NA Hamiltonians predicted by the KRR models. All KRR data converge systematically to the DFT results with increasing sampling, and the 4% sampling gives good results. The errors in the average energy gaps are the smallest among all properties, indicating that the excitation energies are easier to predict than the pure-dephasing times and the NACs. Both the pure-dephasing time and the NAC are higher-order properties than the gap. The pure-dephasing time depends on the amplitude of fluctuation of the gap away from its average value,⁶⁴ while the NAC depends on the gap and other properties, eq 1. The NACs are underestimated by the KRR models because they miss some of the NAC peaks. Similarly, the pure-dephasing times are overestimated because KRR misses some peaks in the energy gap fluctuation and therefore, underestimates the fluctuation amplitude. The relaxation times are overestimated as well, primarily because of the underestimated NAC.

Because the NAC error is generally larger than the error in the excitation energies, it is desirable to improve the NAC prediction. This can be achieved by considering not only system geometries but also velocities because the NAC depends explicitly on the nuclear velocity, as shown in eq 1.

All KRR data show systematic convergence to the DFT values with increased sampling, and therefore, one may use results from several sampling levels to extrapolate the DFT results. For example, one may use the fact that the transition

rate is proportional to the NAC squared, according to the Fermi golden rule.²⁵ The relationship holds for the transitions studied here, as can be verified using the data presented in Tables 1 and S1. Therefore, for example, one can first extrapolate the RMS-NAC to the DFT values with some function and then use the NAC to obtain a more accurate transition rate with the Fermi golden rule dependence of the rate on NAC squared.

It is important to understand why the 4% sampling provides good results while decreasing the sampling below 1% increases the errors. This fact can be rationalized by considering the frequencies in the FTs of the gaps and NACs, as shown in Figure 4. The 4% sampling protocol is based on 5% of the data, with 1% used for validation. The 5% data are spaced 20 fs apart, and the 20 fs period corresponds to the 1666 cm⁻¹ frequency. The FTs decay by about 400 cm⁻¹, as shown in Figure 4H, which corresponds to about 80 fs. Thus, one needs to sample at about four times the highest relevant frequency, to capture both maxima and minima in the NAC and gap fluctuations. Sampling at double frequency may capture the extrema, but may also fall between them. Quadrupling the frequency samples both the extrema and the points between them.

The electrons and holes in MHPs are supported by the metal and halide atoms, and the characteristic frequencies of the lattice are 95 cm⁻¹ for Pb–I stretching and 60 cm⁻¹ for tilting.⁷³ The FTs of the energy gap and NACs exhibit higher frequencies, as shown in Figure 4, corresponding to overtones. Therefore, one should not rely purely on the known vibrational frequencies of the system and obtain a short MD trajectory to identify the highest frequencies that contribute to the gap and NAC fluctuations.

The computational cost of KRR is trivial. The train/validate/predict cycle is finished as soon as the python script is run on a personal laptop. The time-consuming part is to obtain the DFT data for the training. The 4% training/validation requires 350 DFT data points, while the full DFT calculation requires 7000 data points. Because the KRR NA-MD calculation uses only 350 DFT points, and the KRR training part is essentially instantaneous, the computational advantage relative to the fully DFT NA-MD is 20-fold.

Finally, it is important to emphasize once again that the current ML methodology uses time interpolation rather than time extrapolation, and that the time interpolation strategy provides a major advantage over time extrapolation. A small fraction of DFT calculations of energy gaps and NACs are obtained along a precomputed MD trajectory. Then, KRR is employed to interpolate between these points. Such a strategy

allows us to use only 4% of 7000 geometries to obtain high quality NA-MD results. The interpolation strategy works much better than the extrapolation strategy, in which the training is performed using an early part of the trajectory, and the model is then used to extrapolate the energy gaps and NACs for later times. This is because a short 7000 fs trajectory is not sufficient to sample all structures that can be encountered in the future. Much larger sampling using significantly longer MD trajectories would be needed to achieve this goal. For example, if another 2000 fs of MD simulation is performed, it would be impossible to extrapolate the energy gaps and NAC based on the 7000 fs of the already performed MD. However, if 4% of the points from the 2000 fs of the additional MD are used to compute DFT energy gaps and NACs, then accurate interpolation between these points can be easily achieved. The strategy is based on the general idea that some properties, such as ground state energies, are easier to extrapolate in time than other properties, such as excitation energies and NACs. Then, a ML model can be used to time-extrapolate the simpler properties, and another ML model can be used to time-interpolate the more complex properties.

4. CONCLUSIONS

We have demonstrated a ML-assisted NA-MD methodology, which allows one to reduce the calculation cost by over an order of magnitude under the CPA. Using the fact that DFT-quality ML FFs can be obtained rather routinely already, we use a few percent of geometries from pre-computed trajectories to sample the NACs and excitation energies and generate the remaining NACs and excitation energies by interpolation with KRR. This strategy works very well, compared to strategies that aim to generate NACs and excitation energies for a complete set of geometries, similar to the generation of ML FFs, or to extrapolate NACs and excitation energies forward in time. Excited state energies, in particular NACs, are more complex functions of system geometries than the ground state energy and force, and they are harder to predict. The strategy combining time extrapolation of the simpler properties, that is, ground state trajectories, and interpolation of the more complex ones, that is, excitation energies and NACs, provides an efficient alternative. Compared to NNs, the KRR method requires a minimum amount of hyperparameter tuning, which is a time-consuming part of building ML models. The cost of the KRR calculation is negligible compared to the DFT calculations.

The method has been demonstrated with two MHPs, which exhibit intricate MD involving anharmonic motions and multiple timescales. KRR training with 4% of the DFT data generates reliable results. The analysis shows that it is notably easier to obtain the excitation energies than the NACs and the pure-dephasing times, the latter depending on the amplitude of the excitation energy fluctuation away from the average value. Sampling an insufficient amount of data can miss the maxima and minima in the fluctuating NACs and excitation energies. As a result, both inelastic and elastic electron–phonon scattering events become underrepresented. That is, the NACs, the pure-dephasing rates, and the relaxation rates are underestimated. Compared to NNs, the KRR results converge systematically to the exact answers with an increased amount of training data. Valuable on its own right, this property can be used to extrapolate the final answers to the DFT answers and to correct for the systematic underrepresentation of the electron–phonon scattering. The achieved increase in the

efficiency of the NA-MD simulations allows one to study larger systems, longer timescales, and more complicated processes at a significantly reduced computational cost.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpca.1c05105>.

Raw data, including system geometries, energy gaps, and NACs (ZIP)

Charge densities of the key electronic states, 8, 2, and 0.8% sampling prediction for the NACs and energy gaps; and KRR results for CsPbI₃ containing the I_{Cs} defect (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Weibin Chu – Department of Chemistry, University of Southern California, Los Angeles, California 90089, United States; orcid.org/0000-0001-5951-0337; Email: wc_086@usc.edu

Authors

Yifan Wu – Department of Chemistry, University of Southern California, Los Angeles, California 90089, United States

Natalie Prezhdo – Department of Chemistry, University of Southern California, Los Angeles, California 90089, United States

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.jpca.1c05105>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work was supported by the US National Science Foundation, grant no. CHE-1900510.

■ REFERENCES

- (1) 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*, 3192–3203.
- (2) Han, H.; Choi, S. Transfer Learning from Simulation to Experimental Data: Nmr Chemical Shift Predictions. *J. Phys. Chem. Lett.* **2021**, *12*, 3662–3668.
- (3) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine Learning Force Fields. *Chem. Rev.* **2021**, *121*, 10142–10186.
- (4) Legrain, F.; Carrete, J.; van Roekeghem, A.; Curtarolo, S.; Mingo, N. How Chemical Composition Alone Can Predict Vibrational Free Energies and Entropies of Solids. *Chem. Mater.* **2017**, *29*, 6220–6227.
- (5) Marchetti, F.; Moroni, E.; Pandini, A.; Colombo, G. Machine Learning Prediction of Allosteric Drug Activity from Molecular Dynamics. *J. Phys. Chem. Lett.* **2021**, *12*, 3724–3732.
- (6) Loftis, C.; Yuan, K.; Zhao, Y.; Hu, M.; Hu, J. Lattice Thermal Conductivity Prediction Using Symbolic Regression and Machine Learning. *J. Phys. Chem. A* **2021**, *125*, 435–450.
- (7) Faber, F. A.; Hutchison, L.; Huang, B.; Gilmer, J.; Schoenholz, S. S.; Dahl, G. E.; Vinyals, O.; Kearnes, S.; Riley, P. F.; von Lilienfeld, O. A. Prediction Errors of Molecular Machine Learning Models Lower Than Hybrid Dft Error. *J. Chem. Theory Comput.* **2017**, *13*, 5255–5264.

- (8) Gómez-Bombarelli, R.; et al. Design of Efficient Molecular Organic Light-Emitting Diodes by a High-Throughput Virtual Screening and Experimental Approach. *Nat. Mater.* **2016**, *15*, 1120–1127.
- (9) Raccuglia, P.; Elbert, K. C.; Adler, P. D. F.; Falk, C.; Wenny, M. B.; Mollo, A.; Zeller, M.; Friedler, S. A.; Schrier, J.; Norquist, A. J. Machine-Learning-Assisted Materials Discovery Using Failed Experiments. *Nature* **2016**, *533*, 73–76.
- (10) Wang, S.; Zhu, J.; Blackwell, R.; Fischer, F. R. Automated Tip Conditioning for Scanning Tunneling Spectroscopy. *J. Phys. Chem. A* **2021**, *125*, 1384–1390.
- (11) Atlas, S. R. Embedding Quantum Statistical Excitations in a Classical Force Field. *J. Phys. Chem. A* **2021**, *125*, 3760–3775.
- (12) Collins, C. R.; Gordon, G. J.; von Lilienfeld, O. A.; Yaron, D. J. Constant Size Descriptors for Accurate Machine Learning Models of Molecular Properties. *J. Chem. Phys.* **2018**, *148*, 241718.
- (13) Hansen, K.; Biegler, F.; Ramakrishnan, R.; Pronobis, W.; von Lilienfeld, O. A.; Müller, K.-R.; Tkatchenko, A. Machine Learning Predictions of Molecular Properties: Accurate Many-Body Potentials and Nonlocality in Chemical Space. *J. Phys. Chem. Lett.* **2015**, *6*, 2326–2331.
- (14) Browning, N. J.; Ramakrishnan, R.; von Lilienfeld, O. A.; Roethlisberger, U. Genetic Optimization of Training Sets for Improved Machine Learning Models of Molecular Properties. *J. Phys. Chem. Lett.* **2017**, *8*, 1351–1359.
- (15) He, Y.; Cubuk, E. D.; Allendorf, M. D.; Reed, E. J. Metallic Metal–Organic Frameworks Predicted by the Combination of Machine Learning Methods and Ab Initio Calculations. *J. Phys. Chem. Lett.* **2018**, *9*, 4562–4569.
- (16) Chen, H.-A.; Tang, P.-H.; Chen, G.-J.; Chang, C.-C.; Pao, C.-W. Microstructure Maps of Complex Perovskite Materials from Extensive Monte Carlo Sampling Using Machine Learning Enabled Energy Model. *J. Phys. Chem. Lett.* **2021**, *12*, 3591–3599.
- (17) Li, Z.; Kermode, J. R.; De Vita, A. Molecular Dynamics with on-the-Fly Machine Learning of Quantum-Mechanical Forces. *Phys. Rev. Lett.* **2015**, *114*, 096405.
- (18) Schleder, G. R.; Padilha, A. C. M.; Acosta, C. M.; Costa, M.; Fazzio, A. From Dft to Machine Learning: Recent Approaches to Materials Science—a Review. *J. Phys.: Mater.* **2019**, *2*, 032001.
- (19) Mai, S.; González, L. Molecular Photochemistry: Recent Developments in Theory. *Angew. Chem. Int. Ed.* **2020**, *59*, 16832–16846.
- (20) Popp, W.; Brey, D.; Binder, R.; Burghardt, I. Quantum Dynamics of Exciton Transport and Dissociation in Multichromophoric Systems. In *Annual Review of Physical Chemistry*; Johnson, M. A., Martinez, T. J., Eds., 2021; Vol. 72, pp 591–616.
- (21) Ben-Nun, M.; Quenneville, J.; Martínez, T. J. Ab Initio Multiple Spawning: Photochemistry from First Principles Quantum Molecular Dynamics. *J. Phys. Chem. A* **2000**, *104*, 5161–5175.
- (22) Miller, W. H. The Semiclassical Initial Value Representation: A Potentially Practical Way for Adding Quantum Effects to Classical Molecular Dynamics Simulations. *J. Phys. Chem. A* **2001**, *105*, 2942–2955.
- (23) Jaeger, H. M.; Fischer, S.; Prezhdo, O. V. Decoherence-Induced Surface Hopping. *J. Chem. Phys.* **2012**, *137*, 22A545.
- (24) 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*, 10239–10265.
- (25) Hyeon-Deuk, K.; Prezhdo, O. V. Multiple Exciton Generation and Recombination Dynamics in Small Si and Cdse Quantum Dots: An Ab Initio Time-Domain Study. *ACS Nano* **2012**, *6*, 1239–1250.
- (26) Hyeon-Deuk, K.; Prezhdo, O. V. Time-Domain Ab Initio Study of Auger and Phonon-Assisted Auger Processes in a Semiconductor Quantum Dot. *Nano Lett.* **2011**, *11*, 1845–1850.
- (27) Zhou, G.; Lu, G.; Prezhdo, O. V. Modeling Auger Processes with Nonadiabatic Molecular Dynamics. *Nano Lett.* **2021**, *21*, 756–761.
- (28) Li, W.; Zhou, L.; Prezhdo, O. V.; Akimov, A. V. Spin-Orbit Interactions Greatly Accelerate Nonradiative Dynamics in Lead Halide Perovskites. *ACS Energy Lett.* **2018**, *3*, 2159–2166.
- (29) Habenicht, B. F.; Prezhdo, O. V. Ab Initio Time-Domain Study of the Triplet State in a Semiconducting Carbon Nanotube: Intersystem Crossing, Phosphorescence Time, and Line Width. *J. Am. Chem. Soc.* **2012**, *134*, 15648–15651.
- (30) Zhu, H.; Miyata, K.; Fu, Y.; Wang, J.; Joshi, P. P.; Niesner, D.; Williams, K. W.; Jin, S.; Zhu, X.-Y. Screening in Crystalline Liquids Protects Energetic Carriers in Hybrid Perovskites. *Science* **2016**, *353*, 1409–1413.
- (31) Tong, C.-J.; Li, L.; Liu, L.-M.; Prezhdo, O. V. Synergy between Ion Migration and Charge Carrier Recombination in Metal-Halide Perovskites. *J. Am. Chem. Soc.* **2020**, *142*, 3060–3068.
- (32) Shi, R.; Vasenko, A. S.; Long, R.; Prezhdo, O. V. Edge Influence on Charge Carrier Localization and Lifetime in $\text{CH}_3\text{NH}_3\text{PbBr}_3$ Perovskite: Ab Initio Quantum Dynamics Simulation. *J. Phys. Chem. Lett.* **2020**, *11*, 9100–9109.
- (33) Li, W.; Zhan, J.; Liu, X.; Tang, J.; Yin, W.-J.; Prezhdo, O. V. Atomistic Mechanism of Passivation of Halide Vacancies in Lead Halide Perovskites by Alkali Ions. *Chem. Mater.* **2021**, *33*, 1285–1292.
- (34) Qiao, L.; Fang, W.-H.; Long, R.; Prezhdo, O. V. Atomic Model for Alkali Metal Passivation of Point Defects at Perovskite Grain Boundaries. *ACS Energy Lett.* **2020**, *5*, 3813–3820.
- (35) Qiao, L.; Fang, W.-H.; Long, R.; Prezhdo, O. V. Photoinduced Dynamics of Charge Carriers in Metal Halide Perovskites from an Atomistic Perspective. *J. Phys. Chem. Lett.* **2020**, *11*, 7066–7082.
- (36) He, J.; Vasenko, A. S.; Long, R.; Prezhdo, O. V. Halide Composition Controls Electron-Hole Recombination in Cesium-Lead Halide Perovskite Quantum Dots: A Time Domain Ab Initio Study. *J. Phys. Chem. Lett.* **2018**, *9*, 1872–1879.
- (37) Li, W.; Sun, Y.-Y.; Li, L.; Zhou, Z.; Tang, J.; Prezhdo, O. V. Control of Charge Recombination in Perovskites by Oxidation State of Halide Vacancy. *J. Am. Chem. Soc.* **2018**, *140*, 15753–15763.
- (38) Kojima, A.; Teshima, K.; Shirai, Y.; Miyasaka, T. Organometal Halide Perovskites as Visible-Light Sensitizers for Photovoltaic Cells. *J. Am. Chem. Soc.* **2009**, *131*, 6050–6051.
- (39) Jena, A. K.; Kulkarni, A.; Sanehira, Y.; Ikegami, M.; Miyasaka, T. Stabilization of α -CsPbI₃ in Ambient Room Temperature Conditions by Incorporating Eu into CsPbI₃. *Chem. Mater.* **2018**, *30*, 6668–6674.
- (40) Hu, Y.; Bai, F.; Liu, X.; Ji, Q.; Miao, X.; Qiu, T.; Zhang, S. Bismuth Incorporation Stabilized α -CsPbI₃ for Fully Inorganic Perovskite Solar Cells. *ACS Energy Lett.* **2017**, *2*, 2219–2227.
- (41) Eperon, G. E.; Paternò, G. M.; Sutton, R. J.; Zampetti, A.; Haghighirad, A. A.; Cacialli, F.; Snaith, H. J. Inorganic Cesium Lead Iodide Perovskite Solar Cells. *J. Mater. Chem. A* **2015**, *3*, 19688–19695.
- (42) Shum, K.; Tsatskina, A. Solar Cells: Stabilizing Tin-Based Perovskites. *Nat. Energy* **2016**, *1*, 16188.
- (43) Hodes, G. Perovskite-Based Solar Cells. *Science* **2013**, *342*, 317–318.
- (44) Noel, N. K.; et al. Lead-Free Organic–Inorganic Tin Halide Perovskites for Photovoltaic Applications. *Energy Environ. Sci.* **2014**, *7*, 3061–3068.
- (45) Giustino, F.; Snaith, H. J. Toward Lead-Free Perovskite Solar Cells. *ACS Energy Lett.* **2016**, *1*, 1233–1240.
- (46) 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*, 835–842.
- (47) Li, J.; Reiser, P.; Boswell, B. R.; Eberhard, A.; Burns, N. Z.; Friederich, P.; Lopez, S. A. Automatic Discovery of Photoisomerization Mechanisms with Nanosecond Machine Learning Photodynamics Simulations. *Chem. Sci.* **2021**, *12*, 5302–5314.
- (48) Richings, G. W.; Habershon, S. Direct Grid-Based Non-adiabatic Dynamics on Machine-Learned Potential Energy Surfaces:

Application to Spin-Forbidden Processes. *J. Phys. Chem. A* **2020**, *124*, 9299–9313.

(49) 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*, 6702–6708.

(50) Akimov, A. V.; Prezhdo, O. V. The Pyxaid Program for Non-Adiabatic Molecular Dynamics in Condensed Matter Systems. *J. Chem. Theory Comput.* **2013**, *9*, 4959–4972.

(51) Akimov, A. V.; Prezhdo, O. V. Advanced Capabilities of the Pyxaid Program: Integration Schemes, Decoherence Effects, Multi-excitonic States, and Field-Matter Interaction. *J. Chem. Theory Comput.* **2014**, *10*, 789–804.

(52) Long, R.; Prezhdo, O. V.; Fang, W. H. Nonadiabatic Charge Dynamics in Novel Solar Cell Materials. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2017**, *7*, No. e1305.

(53) Neukirch, A. J.; Hyeon-Deuk, K.; Prezhdo, O. V. Time-Domain Ab Initio Modeling of Excitation Dynamics in Quantum Dots. *Coord. Chem. Rev.* **2014**, *263–264*, 161–181.

(54) Chaban, V. V.; Prezhdo, V. V.; Prezhdo, O. V. Covalent Linking Greatly Enhances Photoinduced Electron Transfer in Fullerene-Quantum Dot Nanocomposites: Time-Domain Ab Initio Study. *J. Phys. Chem. Lett.* **2013**, *4*, 1–6.

(55) Long, R.; Casanova, D.; Fang, W.-H.; Prezhdo, O. V. Donor-Acceptor Interaction Determines the Mechanism of Photoinduced Electron Injection from Graphene Quantum Dots into TiO₂: π -Stacking Supersedes Covalent Bonding. *J. Am. Chem. Soc.* **2017**, *139*, 2619–2629.

(56) Zhou, Z.; Liu, J.; Long, R.; Li, L.; Guo, L.; Prezhdo, O. V. Control of Charge Carriers Trapping and Relaxation in Hematite by Oxygen Vacancy Charge: Ab Initio Non-Adiabatic Molecular Dynamics. *J. Am. Chem. Soc.* **2017**, *139*, 6707–6717.

(57) Akimov, A. V.; Asahi, R.; Jinnouchi, R.; Prezhdo, O. V. What Makes the Photocatalytic CO₂ Reduction on N-Doped Ta₂O₅ Efficient: Insights from Nonadiabatic Molecular Dynamics. *J. Am. Chem. Soc.* **2015**, *137*, 11517–11525.

(58) Witten, I. H.; Frank, E.; Hall, M. A.; Pal, C. J.; Witten, I. H. Chapter 4—Algorithms: The Basic Methods. In *Data Mining*, 4th ed.; Frank, E., Hall, M. A., Pal, C. J., Eds.; Morgan Kaufmann, 2017; pp 91–160.

(59) 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*, 11169–11186.

(60) Chu, W.; Prezhdo, O. V. Concentric Approximation for Fast and Accurate Numerical Evaluation of Nonadiabatic Coupling with Projector Augmented-Wave Pseudopotentials. *J. Phys. Chem. Lett.* **2021**, *12*, 3082–3089.

(61) Chu, W.; Zheng, Q.; Akimov, A. V.; Zhao, J.; Saidi, W. A.; Prezhdo, O. V. Accurate Computation of Nonadiabatic Coupling with Projector Augmented-Wave Pseudopotentials. *J. Phys. Chem. Lett.* **2020**, *11*, 10073–10080.

(62) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(63) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953–17979.

(64) Akimov, A. V.; Prezhdo, O. V. Persistent Electronic Coherence Despite Rapid Loss of Electron–Nuclear Correlation. *J. Phys. Chem. Lett.* **2013**, *4*, 3857–3864.

(65) Habenicht, B. F.; Kamisaka, H.; Yamashita, K.; Prezhdo, O. V. Ab Initio Study of Vibrational Dephasing of Electronic Excitations in Semiconducting Carbon Nanotubes. *Nano Lett.* **2007**, *7*, 3260–3265.

(66) 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. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

(67) Hyeon-Deuk, K.; Madrid, A. B.; Prezhdo, O. V. Symmetric Band Structures and Asymmetric Ultrafast Electron and Hole Relaxations in Silicon and Germanium Quantum Dots: Time-Domain Ab Initio Simulation. *Dalton Trans.* **2009**, 10069–10077.

(68) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754.

(69) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(70) Rupp, M.; Tkatchenko, A.; Müller, K. R.; von Lilienfeld, O. A. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys. Rev. Lett.* **2012**, *108*, 058301.

(71) Schwartz, B. J.; Bittner, E. R.; Prezhdo, O. V.; Rossky, P. J. Quantum Decoherence and the Isotope Effect in Condensed Phase Nonadiabatic Molecular Dynamics Simulations. *J. Chem. Phys.* **1996**, *104*, 5942–5955.

(72) Nijjar, P.; Jankowska, J.; Prezhdo, O. V. Ehrenfest and Classical Path Dynamics with Decoherence and Detailed Balance. *J. Chem. Phys.* **2019**, *150*, 204124.

(73) Xiang, S.; Fu, Z.; Li, W.; Wei, Y.; Liu, J.; Liu, H.; Zhu, L.; Zhang, R.; Chen, H. Highly Air-Stable Carbon-Based α -CsPbI₃ Perovskite Solar Cells with a Broadened Optical Spectrum. *ACS Energy Lett.* **2018**, *3*, 1824–1831.