THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL

Letter

Extending the Time Scales of Nonadiabatic Molecular Dynamics via Machine Learning in the Time Domain

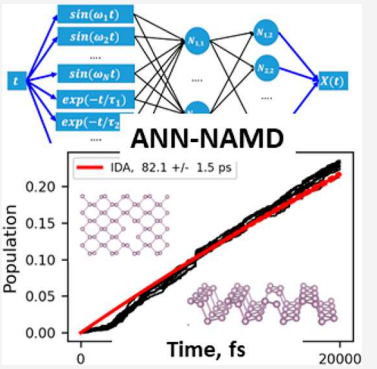
Alexey V. Akimov*

Cite This: J. Phys. Chem. Lett. 2021, 12, 12119−12128

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: A novel methodology for direct modeling of long-time scale nonadiabatic dynamics in extended nanoscale and solid-state systems is developed. The presented approach enables forecasting the vibronic Hamiltonians as a direct function of time via machine-learning models trained directly in the time domain. The use of periodic and aperiodic functions that transform time into effective input modes of the artificial neural network is demonstrated to be essential for such an approach to work for both abstract and atomistic models. The best strategies and possible limitations pertaining to the new methodology are explored and discussed. An exemplary direct simulation of unprecedentedly long 20 picosecond trajectories is conducted for a divacancy-containing monolayer black phosphorus system, and the importance of conducting such extended simulations is demonstrated. New insights into the excited states photophysics in this system are presented, including the role of decoherence and model definition.



Nonadiabatic molecular dynamics (NA-MD) is a powerful tool for modeling the evolution of excited states—the key process that defines the outcomes of many photoinduced processes, such as nonradiative electron−hole recombination, "hot" charge carrier cooling, charge transfer and carrier trapping, excitation energy transfer and relaxation, exciton dissociation, and photoinduced isomerization.[1−6] Modeling nonadiabatic processes in extended nanoscale and solid-state systems is a computationally demanding task and naturally requires various approximations in order to handle large-size systems and long-time scale processes. First, the quantum-classical trajectory surface hopping (TSH) technique[7,8] is the most widely used family of approximations. Second, the neglect of back-reaction approximation (NBRA) is often used in the TSH methods for extended systems.[9−11] The NBRA postulates that the nuclear dynamics is insensitive to the underlying electronic dynamics and helps reduce the number of trajectories by neglecting possible branching of the nuclear probability density as represented by the swarms of classical trajectories.

Despite the use of TSH and NBRA approaches, the modeling of slow nonadiabatic (NA) processes in extended systems remains a challenge. Indeed, if the state transition probabilities are on the order of 10^{-6}, many (on the order of 10^{4}−10^{6}) long (e.g., on the nanosecond scale) trajectories are required to obtain statistically converged results. Depending on the level of theory used and the size of the systems studied, such calculations may be quite expensive. Conventionally, few-picosecond trajectories are computed for systems of few hundreds of atoms while using Kohn−Sham (KS) orbitals and pure functionals such as the Perdew−Burke−Ernzerhof (PBE).[3,10,12−20] Recently, we have extended this procedure to go beyond the KS picture via the time-dependent density functional theory (TD-DFT) calculations of excited states and computing the corresponding couplings.[21,22] However, one major drawback of such an approach is that the TD-DFT still relies on the short-ranged exchange component of the PBE functional and hence is incapable of fully capturing excitonic effects. To capture such effects, the TD-DFT calculations with hybrid functionals are desirable. The calculations with hybrid functionals may be orders of magnitude more expensive than those with pure functionals,[23] especially when the TD-DFT level is leveraged. Jiang et al.[24] also developed an approach that leverages the GW plus Bete−Salpeter equation (GW-BSE) states, which may be even more time-consuming. In the end, the practical threshold is at a few hundred of femtoseconds of direct trajectory computed for a few-hundred atom systems. For the systems with over 1000 atoms, such calculations remain prohibitively expensive and require falling back to semiempirical and tight-binding theories.

Several approaches have been proposed or undertaken to extend the time scales of NA-MD simulations. Excluding the whole spectrum of methods that aim to accelerate the electronic structure calculations part, a number of techniques for sampling rare events in the context of NA-MD calculations have been proposed in the past.[25−27] However, they are formulated with

Received: November 21, 2021
Accepted: December 14, 2021
Published: December 16, 2021

the assumptions that go beyond the NBRA framework, such as the use of non-NBRA trajectories. In the context of the NBRA framework, only a handful of approaches have been used. Prezhdo and Long[28−30] have been using a simplistic approach in which the original short (few-picosecond) trajectory is repeated several times to yield arbitrary long time-series of the properties needed to conduct the TSH calculations. Wang et al.[31] utilized Fourier transform of the directly obtained Hamiltonian matrix elements to reconstruct indefinitely long time-series of such quantities in simulations of charge carrier dynamics in organic semiconductors with dynamic disorder. Because of the periodicity properties of the Fourier transform, the data forecasted by a Fourier series reconstructed from the short-time dynamics of the target property (e.g., energy) is exactly the repetition of the short-time time-series used to compute the Fourier transform of the target property. In this regard, the method of Wang et al.[31] is similar to the repetition approach of Long and Prezhdo.[28−30]

Several years ago, the author proposed the biased NAD[32] and quasi-stochastic Hamiltonian (QSH)[33] approaches to address the infrequent hops problem within the NBRA NA-MD context. Recognizing the quasi-periodicity of the time-only dependent vibronic Hamiltonian that appears in this context, the QSH approach constructs the Hamiltonian in a physically guided way such that the key frequencies, mean values, and fluctuation amplitudes of energy levels and scalar nonadiabatic couplings (NACs) time-series are captured. The resulting analytical expression is explicitly dependent on a single "effective coordinate"—time—and can be used to compute indefinitely long trajectories at a negligible cost. Unlike the Fourier transform approach of Wang et al.[31] or the short-trajectory-repetition technique of Long and Prezhdo, the QSH method yields the time-series of the desired properties that are not mere repetition of the properties known from the short-time dynamics, especially when a sufficiently large number of modes is included. Instead, the QSH generates qualitatively new trajectories that can better mimic stochastic effects in long-term simulations instead of repeating the noise present in the short-time trajectories.

The QSH can be regarded as a simplistic machine learning (ML) approach. It uses the initial data set to "train/reconstruct" the function that is used to produce the time-series with the statistical properties similar to those of the original data set, although QSH does not mean to provide the best fit to such a data set, only to capture its statistical properties. A logical next step is to apply the formal ML techniques to predict the vibronic Hamiltonians as the function of time, capturing the original data set as much as possible this time. The use of ML in NA-MD modeling is a vigorously spreading paradigm that has seen many early successes already,[34−41] although met several challenges as well.[42] For instance, Dral, Barbatti, and Thiel (DBT)[37] reported the kernel ridge regression (KRR) approach that fits energies and NACs to the "training" data points computed for multiple molecular geometries. Considering this technique represents the properties of interest in terms of internal coordinates of the studies systems, the dimensionality of the input space of the ML learning models increases with the number of classical degrees of freedom of the atomistic systems. As a result, such an approach needs a very large number of training points to achieve a good convergence. As an example, for a 33-dimensional model, one needs already about 10000 points (each requires expensive ab initio calculations). Such a large number of points may be prohibitive for nanoscale systems. Naturally, the nonlinear dimensionality reduction (NLDR) techniques have been suggested. Virshup et al.[43] utilized the diffusion map (DM) approach for the detailed analysis of ethylene's NA dynamics. Li et al.,[44,45] explored the utility of classical multidimensional scaling (MDS) and isometric feature mapping (ISOMAP) as meaningful descriptors of the photodynamics in several molecular and model systems. Hughes et al.,[46] utilized the NLDR to construct effective low-dimensional potentials for the use in wavepacket dynamics. Zhou et al.[38] developed unsupervised ML approaches to extract the "most important" coordinates that could be used in the NA-MD. Mangan et al.[47] explored the unsupervised ML approaches to establish correlations between the structural features of condensed-matter systems such as lead halide perovskites and their NACs and band gaps, although no use of ML in the NA-MD simulations has been reported. More recently, Wu et al.[48] utilized the KRR approach to conduct the NA-MD calculations of perovskite systems using several internal degrees of freedom as the inputs to the ML model to predict the NACs and energy gaps.

In this work, I present a new time-domain ML (TD-ML) approach for long-time NA-MD, specifically designed for the NBRA-type of calculations. Unlike most of the existing NLDR-ML techniques, the key "reaction" coordinate is time itself. As was alluded to above, within the NBRA framework, the energy gaps and NACs are explicit functions of time. Thus, from the mathematical or computer science standpoint the problem is essentially a forecast problem—the prediction of the time-series. The solution to such a problem can be found via either convolution neural networks or long short-term memory networks.[49] However, such techniques are prone to potential error accumulation problems and lack physical insights. Instead, in the current TD-ML a multilayer perceptron (MLP) artificial neural network (ANN) with a specially designed input layer structure is used (Figure 1).

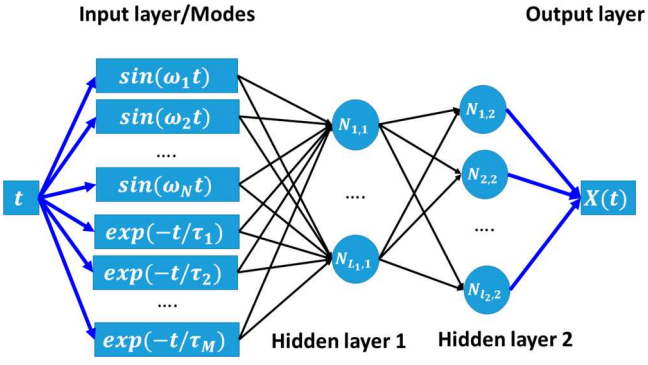


Figure 1. Conceptual schematic of the TD-ML model for NBRA-based NA-MD simulations.

Considering that the training of the MLPs with the hyperbolic tangent transfer function is the most efficient only for input values around zero, where the transfer function has the largest slope, using the time variable as the direct input to the MLP is doomed to fail. At the same time, we recognize the quasi-periodicity of the energy gap and NAC time-series that appear in NBRA NA-MD calculations. Thus, it is reasonable and essential to utilize a nonlinear transformation of time. In this work, a series of periodic and bound basis functions (hereafter referred to as the modes) of time, $\sin(\omega_i t)$, is used as such a transformation. Although only one effective reaction coor-

https://doi.org/10.1021/acs.jpclett.1c03823
J. Phys. Chem. Lett. 2021, 12, 12119−12128

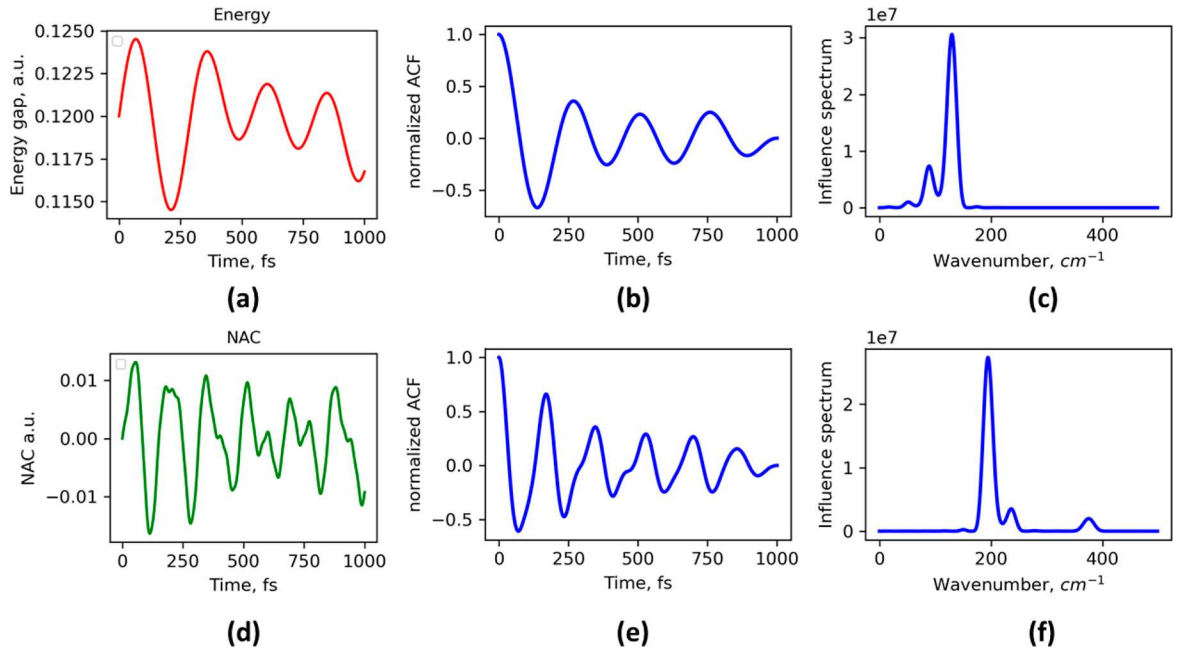


Figure 2. Definition of frequencies of the ANN input modes' frequencies in the TD-ML method.

dinate—time—is used as a nominal input to the MLP, it generates a series of linearly independent input basis functions that are propagated to the MLP (Figure 1) and undergo further nonlinear transformations through the hidden neurons. The realistic time-series of electronic energy gaps and NACs obtained from ab initio calculations are almost never ideal periodic functions, since the thermal effects (bath, thermostat) are included. Therefore, it is desirable to break the exact periodicity of the output, $X(t)$, as the function of the input t. For this reason, the input layer of the TD-ML scheme is extended to include nonperiodic but bound (for instance, to the $[0, 1]$ interval) basis functions (modes) of the form, $\exp(-t/\tau_i)$, Figure 1. The input from the basis functions of the two types is then propagated to multiple hidden layers of the MLP to yield a single output value, $X(t)$, of the desired property. Thus, the overall effect of the constructed ANNs is to obtain the mapping of time, t, to a quasiperiodic time-only dependent property of interest, $X(t)$: MLP:$t \rightarrow X(t)$.

The TD-ML method can be regarded as a simple yet quite dramatic NL-DR approach: the dynamics of an arbitrary system (model or atomistic, finite or periodic, of any size and composition, with or without external bath modes), is described by a single coordinate, time, which nonlinearly modulates the system's properties of interest. The time variable replaces the explicit consideration of the system's and bath's degrees of freedom, which all are nonlinear functions of time. However, in order to use the time-variable to predict the properties of interest at its arbitrary values via MLP, the nonlinear transformation of this variable to a set of auxiliary linearly independent basis functions is necessary.

Although the dimensionality reduction in the TD-ML model can be quite dramatic (from hundreds and thousands of internal coordinates to just one), it is important to remember its main limitation—the NBRA. In other words, the approach is expected to work best when the dynamics of the nuclear degrees of freedom is confined to a limited region of the phase space, which is usually the case for systems that do not undergo significant structural reorganization on the time scale of simulations. Having said that, the approach may still be applicable to systems with a greater flexibility of nuclear degrees of freedom (e.g., photoinduced isomerization situations), but the corresponding MLP models may be harder to train and would require a significant amount of the training data. The TD-ML approach also assumes the regularity of the dynamics and cannot describe the processes that are not included in the training dynamics. For instance, if the system can undergo multiple reactive steps not seen in the training data, the TD-ML approach is expected to break down. If the training data contains information on reactive processes, the TD-ML would predict such processes (implicitly, via the energy gaps and NACs) in the forecasted time-series with the regularity seen in the training data, which may or may not be adequate in the long-time simulations.

In the present work, the aim is to predict adiabatic energy gaps between all adjacent pairs of states, $X(t) = \Delta E_{i+1} = E_{i+1}(t) - E_i(t)$ and scalar NACs for all distinct pairs of states,

$$X(t) = d_{ij} = -d_{ji} \equiv \left\langle \psi_i \middle| \frac{\partial}{\partial t} \middle| \psi_j \right\rangle, \; i < j$$

Separate ANNs are trained for each property. Thus, for a 2-level system, 2 ANNs are trained, one for the energy gap $\Delta E_1 = E_1 - E_0$ and one for the NAC, d_{01}; for a five-state system, a total of 14 ANNs are trained: 4 for energy gaps and $\frac{5\times 4}{2} = 10$ for distinct NACs. Although it is possible to train the joint ANNs for all states and all NACs, such an approach is not explored in this work and will be studied separately. Furthermore, the primary goal of this work is to explore the viability of the ANN based approach rather than provide a comprehensive evaluation of various possible strategies.

Similar to QSH formulations, the frequencies, ω_i, used in the mode definition are taken as the peak positions of the Fourier transform of the target property's autocorrelation functions (ACF), $I(\omega) = |FT[\langle X(0)X(\tau)\rangle]|^2$ (Figure 2). This choice is motivated by the line shape theory,[50,51] which associates such frequencies with the phonon modes that are most strongly coupled to the corresponding electronic state transitions.

https://doi.org/10.1021/acs.jpclett.1c03823
J. Phys. Chem. Lett. 2021, 12, 12119−12128

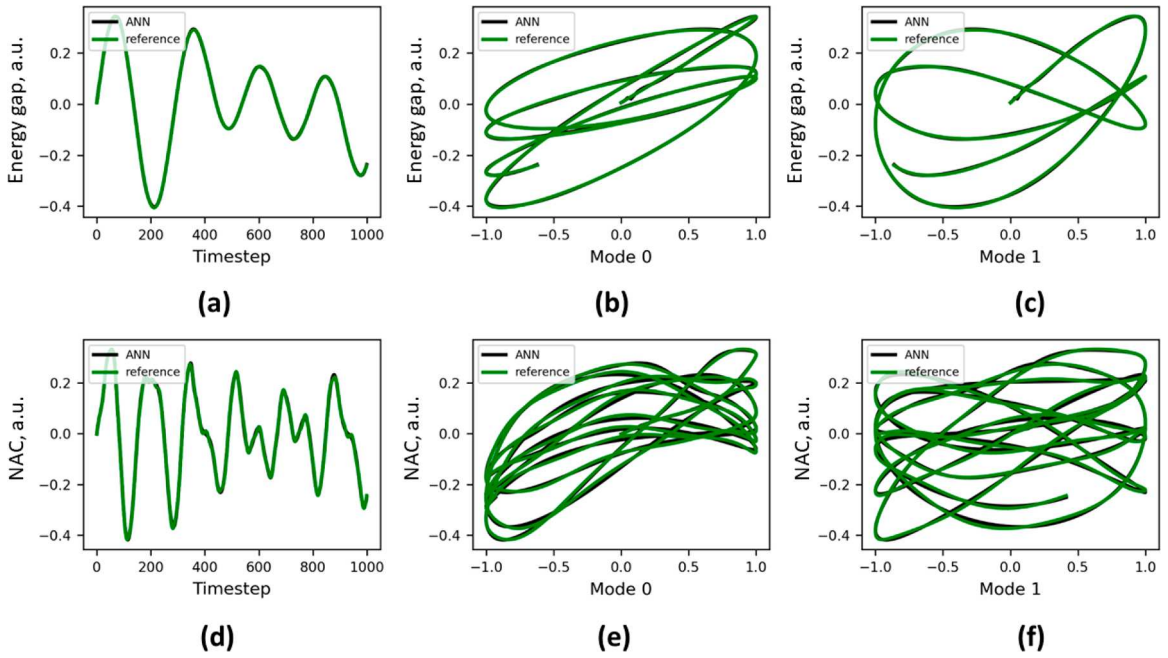


Figure 3. Energy gap (a, b, c) and NAC (d, e, f) reproduced by the ANNs trained on the data set of 1000 points. The panels a and d show the desired time-dependent functions, while the actual training is conducted in terms of the input basis functions (modes). The target functions as the functions of the first modes (different for energy gap and NAC) are shown in panels b, c, e, and f.

Therefore, such modes may be regarded as the key internal degrees of freedom that parametrize the vibronic Hamiltonian. The number of the maxima in the ACF for a given property X may be large, so only a fraction of the dominant frequencies is selected to define the modes. Specifically, only the frequencies with the amplitudes larger than 0.01% of the maximal peak amplitude: ω:$I(\omega) \geq 10^{-4} \times I(\omega_{\max})$, are used in this work. At this point, it is unclear what the best choice of the τ_i parameters is, so they are selected to span a range of time scales up to the order of magnitude of the "training" trajectory. Following the common ANN training practices, the time-series of the target properties $X(t)$ used to train the MLPs are first centered at zero and then rescaled to the $[-0.75, 0.75]$ interval before they are used in the ANN training. The corresponding transformations are saved and stored in the .json files for later reuse in NA-MD calculations. In all cases, the resilient propagation (RProp) algorithm[52−54] is used for the ANN training. Unlike the commonly used backpropagation algorithm (back-prop) which can use a variable number of training patterns at every training epoch, the RProp algorithm uses all the presented training samples at every training step (epoch). This makes it more memory-consuming and often slower than the back-prop, but more robust and faster to converge. In all the simulations presented in this work, the ANNs are trained with the RProp algorithm for 250 epochs with 250 training iterations per each epoch. Tracking the error function of the trained ANNs shows that this is sufficient to achieve the convergence of the trained MLPs. The training data preparation, ANN training algorithms, and the NA-MD computations, both direct and ANN-based, are conducted with the Libra[55] (version 5.1.0[56]) software package in which all these capabilities are implemented.

The typical output of the trained ANNs is illustrated in Figure 3, which shows how the time-series for the energy gap (Figure 3a) and NAC (Figure 3d) can be reconstructed as functions of a continuous time variable. As mentioned above, the training is not conducted in terms of the direct time variable, but rather in terms of periodic and aperiodic basis functions. The desired properties as the functions of two leading periodic modes are shown in Figure 3b,c for gap, and Figure 3e,f for NAC. Note that in these coordinates, the functions appear nonsingle-valued since for each value of the basis function there may be several values of the target function. However, one should remember that these panels show only the projections of the actual function on lower-dimensional subspaces. The actual function is single-valued in terms of the effective coordinate and time, as is apparent from Figure 3, panels a and d.

To demonstrate the performance of the TD-ML NA-MD approach, we consider several types of calculations: (a) a model two-state time-dependent Hamiltonian; (b) a two-state and (c) a five-state Hamiltonian based on the TD-DFT calculations for the divacancy-containing monolayer black phosphorus, studied recently by the author.[57] Since the main focus of this work is on the TD-ML, the NA-MD is conducted only at the Tully's fewest switches surface hopping (FSSH)[7] level, for systems (b), and (c) at the instantaneous decoherence at the attempted hops (IDA) of Tretiak.[58]

For all models, the ANN training data consist of the time-dependent Hamiltonian matrix elements, obtained either analytically (as in model I) or via atomistic calculations (models II and III). Such a time-dependence of the energy gaps and NACs should be understood in a general way: it may reflect either the evolution of the system alone or can also incorporate system-bath interactions with an indefinite number of implicit thermal bath modes. From the mathematical/algorithmic point of view, the TD-ML approach handles all types of data equally, but the way the training data are produced determines the kind of physics/interactions that are accounted for in the resulting ANN model. Another important consideration for NA-MD simulations is whether a phase correction is applied to the wave functions and the corresponding properties[59] along the precomputed trajectory and whether the state identities are tracked along the trajectories.[60−63] For model I below, one can

https://doi.org/10.1021/acs.jpclett.1c03823
J. Phys. Chem. Lett. 2021, 12, 12119−12128

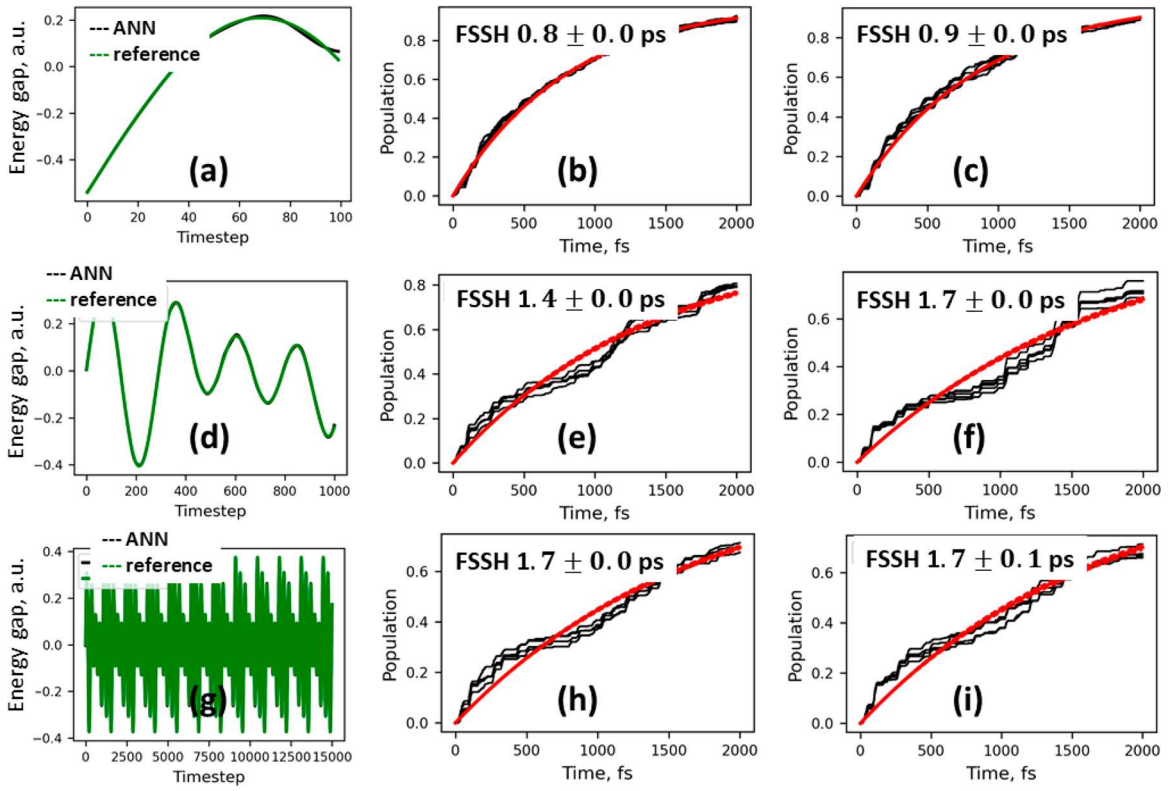


Figure 4. Dependence of the computed dynamics on the size of the ANN data set and the selection of the τ parameters for a fixed ANN architecture with one layer of five hidden neurons.

assume without the loss of generality that both effects are included by the construction. For models II and III, such corrections have been applied explicitly, when the properties were obtained, as explained elsewhere.[57]

Model I. First, we consider a model vibronic Hamiltonian of the form

$$H(t) = \begin{pmatrix} 0 & i\hbar d_{01}(t) \\ -i\hbar d_{01}(t) & H_{11}(t) \end{pmatrix} \tag{1a}$$

$$H_{11}(t) = 0.12 + 0.002\sin(\omega_{11}t) + 0.0032\sin(\omega_{12}t) - 0.001\sin(\omega_{13}t) \tag{1b}$$

$$d_{01}(t) = \frac{0.2\sin(\omega_{21}t) + 0.12\sin(\omega_{22}t) - 0.1\sin(\omega_{23}t) + 0.01\sin(\omega_{24}t)}{H_{11}(t)} \tag{1c}$$

Here, $\omega_{11} = 100$ cm^{-1}, $\omega_{12} = 125$ cm^{-1}, $\omega_{13} = 50$ cm^{-1}, $\omega_{21} = 200$ cm^{-1}, $\omega_{21} = 225$ cm^{-1}, $\omega_{23} = 375$ cm^{-1}, and $\omega_{24} = 1175$ cm^{-1}. The plots of the first 1 ps time-series for these components of this Hamiltonian are shown in Figure 2a,d. The model Hamiltonian given by eq 1a can be regarded as obtained within the NVE or NVT ensemble sampling. It is convenient to study several important methodological questions related to the formulation of the TD-ML method. The first question being, How does the quality of the trained ANNs depend on the amount of the training patterns used? The key aspiration here is to be able to use a small subset of the available data to construct ANNs that could accurately capture the dynamics (referred to as the ANN-based dynamics) computed with the full data set used directly (referred to as the direct dynamics). The FSSH dynamics computed using the 2 ps trajectories with the Hamiltonians given directly by eq 1 yields the ground state repopulation time scale of 1.7 ps. This time scale is considered a reference result for this model.

The ANNs with fixed architecture (with five hidden and one output neurons) are trained using different sets of the training data, ranging from as little as 100 fs (100 data points), Figure 4, panels a, b, c to as much as 15 ps (15 000 data points), Figure 4, panels g, h, and i. As Figure 4 suggests, the use of a small data sets usually leads to faster dynamics −0.8−0.9 ps (Figure 4, panels b and c) for the smallest training set, while increasing the number of the training points to 15 000 helps recover the reference ground state repopulation time scale of 1.7 ps (Figure 4, panels h and i. The excessive amount of the training data points is not necessary though: reasonable time scales of 1.4−1.7 ps are already recovered using the half of the data points used in the direct calculations (Figure 4, panels d−f). Despite the general trends observed, the outliers are also possible—just increasing the amount of the data points may sometimes lead to larger deviations of the ANN-based results from the reference dynamics (e.g., see Supporting Information, Figures S1 and S2). This is likely due to overtraining the ANN.

The choice of the τ_i parameters in the nonperiodic basis functions is not obvious from the beginning. For this reason, two sets of calculations have been conducted−one with three τ parameters: 1000, 3000, and 5000 au (Figure 4, panels b, e, h) and the other with eight τ parameters spanning several orders of magnitude: 1, 10, 25, 100, 500, 1000, 3000, and 5000 au (Figure 4, panels c, f, i). A rather general observation is that such an extended set increases the flexibility of the resulting ANNs and leads to a more favorable agreement with the target reference (e.g., compare second and third columns in Figure 4, also see Supporting Information, Figures S1 and S2). So far, there is no

https://doi.org/10.1021/acs.jpclett.1c03823
J. Phys. Chem. Lett. 2021, 12, 12119−12128

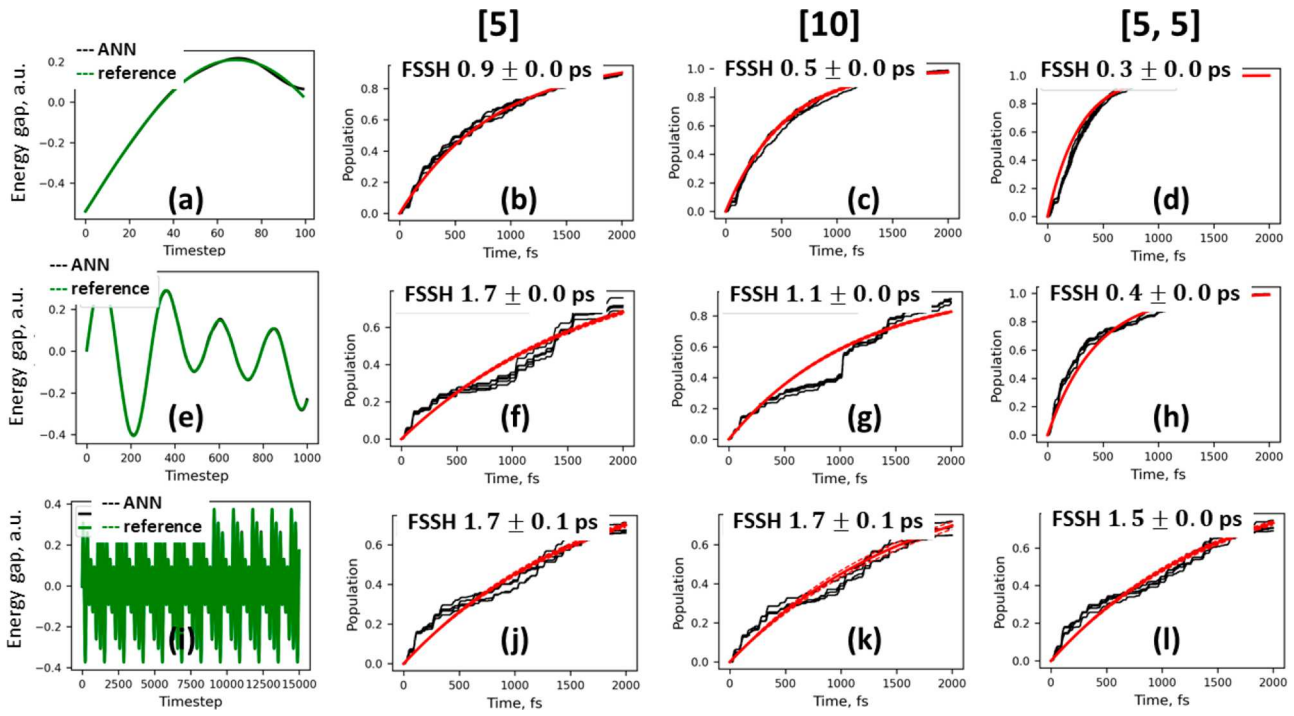


Figure 5. Comparison of the direct and ANN-based dynamics for the model I Hamiltonian with the ANNs of different architectures (5 hidden neurons, b, f, j; 10 hidden neurons, c, g, k; 5 and 5 hidden neurons in 2 hidden layers, d, h, l) trained using different sets of the training data (100 fs, a–d, 1000 fs, e–h, 15 000 fs, i–l). Only the hidden layers' architectures are shown. Panels a, e, and i show the original (green) and ANN-based (black) energy gap as a function of time. The ANN reproduces the gap almost identically so the lines superimpose.

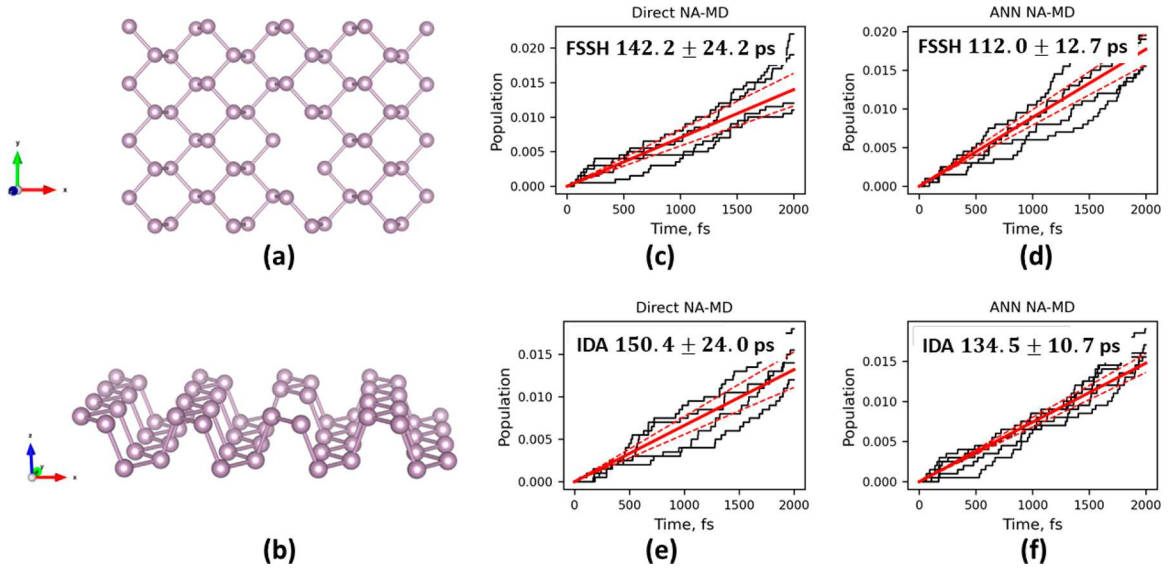


Figure 6. Nonadiabatic dynamics in a 2-state model of a divacancy-containing ML-BP (model II): (a) top and (b) side views of the supercell. The ground state population recovery kinetics computed using (c, d) FSSH and (e, f) IDA methods with the (c, e) direct and (d, f) ANN-based dynamics.

clear way to determine these parameters a priori. The choice here is to use a set of tau parameters that span several orders of magnitude. Such an approach allows for a higher flexibility of the ANN to "select" the relevant parameters during the training procedure.

Another logical question regarding the formulation of the optimal TD-ML model is the choice of the MLP architecture. We observe that using the larger number of neurons in the hidden layers as well as the larger number of such layers tends to accelerate the computed dynamics (Figure 5). This effect tends to fade away for larger training data sets (e.g., Figure 5, panels j–l). It may be attributed to the quality of each neuron connection (converged weights and biases) in the trained ANNs. Increasing the number of neurons and hidden layers increases the number of free ANN parameters to be optimized. Using insufficient data to train the ANN means a smaller amount of training information is used to optimize the corresponding parameters (Figure 5, panels b–d). This explanation agrees with the weaker deterioration of the ANN results for more complex ANN architectures when sufficient or even excessive amount of training information is used (e.g., Figure 5, panels j–l). Therefore, it is preferable to use minimalistic ANN architec-

https://doi.org/10.1021/acs.jpclett.1c03823
J. Phys. Chem. Lett. 2021, 12, 12119–12128

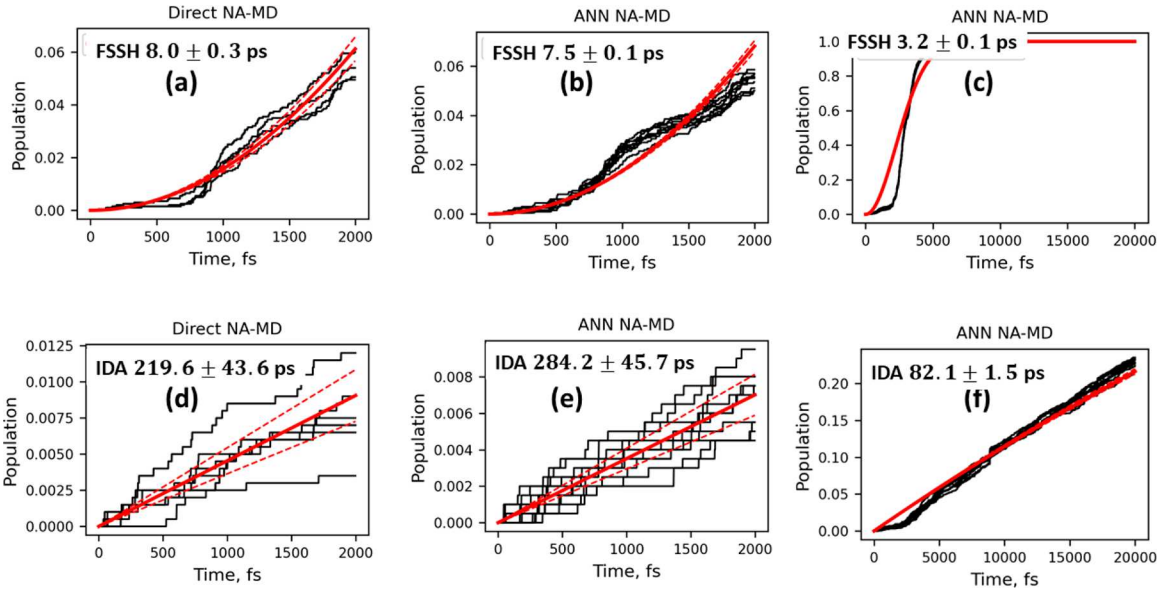


Figure 7. Comparison of the ground state population recovery kinetics in the direct (a, d) and ANN-based (b, c, e, f) dynamics for the five-state divacancy ML-BP model (model III). Panels c and f demonstrate the explicit 20 ps dynamics computed with the ANN-based Hamiltonian.

tures, especially when the amount of the training data points is limited. Stated differently, the time scales predicted by the ANNs are the lower bounds to the true time scales obtained in the direct dynamics. This statement should be taken mainly as a phenomenological observation (although rather consistent in many calculations, e.g. see Supporting Information, Figures S1 and S2), and deviations from this rule are possible. Furthermore, this conclusion follows from a single model Hamiltonian and the atomistic calculations that are described below. Further exploration of this rule may be needed with a more extensive set of models, although this goal goes outside the scope of this work.

Second, the developed TD-ML NA-MD approach is applied to the atomistic problem. Specifically, to the divacancy-containing monolayer black phosphorus (ML-BP, Figure 6, panels a and b). Here, the "ML" abbreviation stands for monolayer and is not to be confused with the "ML" which stands for "machine learning" in other contexts. In the atomistic model, the vibronic Hamiltonians are computed at the TD-DFT/PBE level of theory for the periodic supercell of ML-BP along the precomputed ground-state MD trajectory and are stored in files. The details of all calculations are given in a recent work of the author,[57] so are not repeated here. The same set of files with vibronic Hamiltonians as used in that prior work is used here. A total of 2000 files are used for a direct dynamics of 2 ps.

Model II. First, a minimalistic ML-BP model of only two states (ground and the first excited) is considered. The vibronic Hamiltonians for all 2000 steps are used to train the two ANNs—one for the band gap and the other one for the scalar NAC coupling these two states. Following the conclusions from the model Hamiltonian (model I), 10 τ parameters are used to define the nonperiodic basis functions for the ANN input: 1, 5, 10, 25, 50, 100, 250, 500, 1000, and 2000 au. The automatic procedure determines an additional 18 modes for the energy gap and 36 modes for the NAC. Following the observation for model I, only one hidden layer is used. The number of neurons in the hidden layer is chosen to be roughly half of the number of input neurons. Hence, the architectures of the gap and NAC ANNs are [28, 15, 1] and [46, 25, 1], respectively.

The direct dynamics computed with both the FSSH (Figure 6c) and IDA (Figure 6e) yield the exponential ground state recovery kinetics with the time scales of 142−150 ps. The time scales are comparable to each other for both the FSSH and IDA methods, suggesting the relative unimportance of decoherence effects in this model. This effect is attributed to the model properties—having only two states makes the decoherence events less important. This hypothesis is confirmed below using a five-state model. The ANN-based dynamics yields slightly shorter time scales of 112−134 ps (Figure 6d,f), but still within the error bars for each of the FSSH and IDA methods separately. Thus, the TD-ML approach works reasonably well although tends to give the lower bound for the time scales, consistent with the observations for model I.

Model III. Although the ANN-based and direct dynamics are consistent with each other, the computed time scales are much larger than a few picoseconds time scale reported earlier.[57] This difference may be attributed to the difference in the number of excited states included in modeling. The model II only uses one, whereas the prior work utilizes 10 excited states. To test this hypothesis, the NA-MD calculations are repeated for the ML-BP but with five electronic states (model III, Figure 7). Indeed, the ground state recovery kinetics changes to the Gaussian one and the time scales for the FSSH calculations decrease dramatically to 7−8 ps (Figure 7, panels a and b), whereas the IDA time scales increase to 220−280 ps (Figure 7, panels d and e). These numbers are in much better agreement with the results reported previously,[57] even though only four excited states are used instead of 10. These results confirm the above-mentioned attribution of the increased FSSH time scales in model I to the number of excited states used in simulations. Indeed, as shown earlier,[57] there are several closely spaced excited states nearby the S_1 state in the divacancy ML-BP. Thus, a strong population transfer to those upper states is facilitated in the early dynamics. The rest of the dynamics can be regarded as the one started from a band of states. Such a many-state delocalization of the initial state leads to the Gaussian dynamics, which is characterized by much shorter time scales than those derived from the exponential kinetics.

https://doi.org/10.1021/acs.jpclett.1c03823
J. Phys. Chem. Lett. 2021, 12, 12119−12128

The five-state model III also highlights the importance of decoherence effects as compared to the two-state model II (e.g., compare Figure 7, panels a and d). As noted above, the five-state model favors the fast coherent population transfer to all the excited states from the starting S_1 state since there are several closely positioned excited states. The increased wave function amplitudes on all such states favor the hopping in and out of these states. The IDA decoherence algorithm prevents the accumulation of the wave function amplitudes on the inactive states and slows down the dynamics involving such states. Furthermore, by virtue of the wave function collapse, the IDA method "resets" the dynamics to an effective 2-model situation, favoring the exponential kinetics instead of the Gaussian. In the 2-state model II, no such coherent population transfer on states above S_1 is possible by definition, so the IDA algorithm does not change the qualitative picture, leading to both FSSH and IDA results appearing similar (Figure 6, panels c−f).

It should be reminded that the observations regarding the importance of decoherence effects in models with different number of electronic states are obtained within the NBRA framework. Outside the NBRA, decoherence effects may manifest themselves already in two-state models, as is known from the Fermi-golden rule types of approaches. Furthermore, the decoherence effects discussed here are explored only with one of the simplest decoherence correction schemes—the IDA method. The observed insensitivity of the dynamics to the inclusion of decoherence via the IDA scheme may be its intrinsic deficiency. Other, more sophisticated decoherence correction schemes such as decoherence-induced surface hopping (DISH),[57,64] modified simplified decay of mixing (mSDM),[65] or an augmented FSSH (A-FSSH)[66,67] could be used to further explore these effects. However, such a discussion is outside the scope of the present work, which focuses on the TD-ML methodology.

As with the two-state model, the dynamics computed with the ANN-based vibronic Hamiltonian yields the time scales (Figure 7, panels b and e) that agree with the values obtained from the direct dynamics (Figure 7, panels a and d) within the error bars. The agreement is observed for both the FSSH and IDA methods, suggesting that the TD-ML approach remains robust for more complex problems such as the one defined by the model III.

Finally, we are in a position to explore the applicability of the TD-ML to direct long-time NA-MD simulations. Using the ANN-based vibronic Hamiltonian developed for model III, explicit 20 ps NA-MD trajectories are computed using the FSSH (Figure 7c) and IDA (Figure 7f) methods. As a reminder, each simulation consists of 10 batches of 2000 stochastic trajectories, yet requires only a couple dozen hours of the wall-time. The wall-time may likely be further reduced by reducing the number of the input basis functions. In the current calculations, the NAC ANNs used 70−90 input modes. The performed long-time simulations reveal several interesting effects. First, the FSSH-based ground state population recovery kinetics is not Gaussian in the long-time limit, although rather close. To be able to compare the time scale with other simulations, the population curve is fitted to the best Gaussian fit (Figure 7c). The corresponding time scale is twice shorter than those obtained using the fit to shorter 2 ps trajectories, 3.2 ps. A similar effect holds for the IDA calculations (Figure 7f), for which the time scale falls into the 82 ps range instead of the extrapolation-based 284 ps range. A close examination of Figure 7f shows that the initial slope of the population recovery curves is smaller than the one at the later times. This effect may be rationalized as follows: at longer simulation times, larger average amplitudes on all states develop, even with the decoherence effects in place. Thus, the state transitions are more likely than at shorter simulation times when the average amplitudes on all states are relatively small. Alternatively, this could be considered a locally nonlinear kinetics. A possible reason for such a nonlinearity could be the methodological deficiency of the ad hoc IDA method. A conjecture is that a proper decoherence method should preserve the kinetic order of the state transition rates, thus making the time scales derived from short trajectories potentially more consistent with those obtained from longer simulations. However, testing this hypothesis goes outside the scope of the present work.

In summary, this work reports the time-domain machine-learning methodology for direct NA-MD with dozens, if not hundreds, of picoseconds to model slow nonadiabatic processes. This work also demonstrates the application of this technique to the explicit modeling of 20 ps dynamics of excited state relaxation in divacancy-containing ML-BP. The developed TD-ML approach is rooted in the idea of using the MLP ANN with periodic and aperiodic basis functions of time as the input modes. The ANN quality improves when an extensive set of training data can be used, but a reasonable performance can be obtained with rather small training sets already. The ANN quality can also be improved by increasing the number of aperiodic basis functions. The ANN architecture is not too critical to the quality of the derived ANNs—only one hidden layer with a modest number of neurons is sufficient to lead to a good performance. Increasing the complexity of the ANN architecture may in fact lead to poorer results due to the increased number of parameters that need to be optimized. In this case, larger training data sets are required. The ANN-based dynamics tend to accelerate the dynamics as compared to the direct dynamics, as observed both in model and atomistic simulations.

The simulations conducted in this work, suggest that including higher-lying excited states in simulations can significantly affect the computed time scales for the ground state population recovery, even if the system starts in the S_1 state. In FSSH simulations, such additional states accelerate the kinetics (to about 7−8 ps) and make it Gaussian, whereas including only one excited state in the model leads to slower kinetics of the exponential type (140−150 ps). The presence of extra states in the model leads to a notable difference in the ground state recovery time scales as computed with the overcoherent FSSH (7−8 ps) and decoherence-accounting IDA methods (220−280 ps). In the model problem with only two states, there is almost no difference between the FSSH and IDA results (140−150 ps). Finally, the direct long-trajectory simulations conducted with the TD-ML method reveal that direct simulation of the slow dynamics may lead to smaller population transfer time scales (e.g., 3.5 ps for FSSH and 82 ps for IDA) as compared to those obtained from the short-time dynamics due to stronger population-dependence of the nonadiabatic transfer rates at such later times. It is also conjectured that such an acceleration may be due to intrinsic deficiencies of ad hoc decoherence correction schemes such as the IDA methodology. The computational protocols used in this work (Python and SLURM scripts), the key input and output files, as well as important computational data are available online.[68]

https://doi.org/10.1021/acs.jpclett.1c03823
J. Phys. Chem. Lett. 2021, 12, 12119−12128

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.1c03823.

Comprehensive study of ANN-based dynamics for model I (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Alexey V. Akimov − *Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, New York 14260-3000, United States;* orcid.org/0000-0002-7815-3731; Email: alexeyak@buffalo.edu

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.1c03823

Notes

The author declares no competing financial interest.

■ ACKNOWLEDGMENTS

The author acknowledges financial support of the U.S. National Science Foundation (Grant NSF-CHE-2045204). The computations were performed at the Center for Computational Research at the University at Buffalo, SUNY.

■ REFERENCES

(1) Smith, B.; Akimov, A. V. Modeling Nonadiabatic Dynamics in Condensed Matter Materials: Some Recent Advances and Applications. *J. Phys.: Condens. Matter* **2020**, *32*, 073001.

(2) Prezhdo, O. V. Modeling Non-Adiabatic Dynamics in Nanoscale and Condensed Matter Systems. *Acc. Chem. Res.* **2021**, *54*, 4239−4249.

(3) Zheng, Q.; Chu, W.; Zhao, C.; Zhang, L.; Guo, H.; Wang, Y.; Jiang, X.; Zhao, J. Ab Initio Nonadiabatic Molecular Dynamics Investigations on the Excited Carriers in Condensed Matter Systems. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2019**, *9*, e1411.

(4) Curchod, B. F. E.; Martínez, T. J. Ab Initio Nonadiabatic Quantum Molecular Dynamics. *Chem. Rev.* **2018**, *118*, 3305−3336.

(5) Agostini, F.; Curchod, B. F. E. Different Flavors of Nonadiabatic Molecular Dynamics. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2019**, *9*, No. e1417.

(6) Crespo-Otero, R.; Barbatti, M. Recent Advances and Perspectives on Nonadiabatic Mixed Quantum−Classical Dynamics. *Chem. Rev.* **2018**, *118*, 7026−7068.

(7) Tully, J. C. Molecular Dynamics with Electronic Transitions. *J. Chem. Phys.* **1990**, *93*, 1061−1071.

(8) Nikitin, E. E. *Theory of Elementary Atomic and Molecular Processes in Gases*; Clarendon: Oxford, 1974.

(9) Prezhdo, O. V.; Duncan, W. R.; Prezhdo, V. V. Photoinduced Electron Dynamics at the Chromophore−Semiconductor Interface: A Time-Domain Ab Initio Perspective. *Prog. Surf. Sci.* **2009**, *84*, 30−68.

(10) Habenicht, B.; Craig, C.; Prezhdo, O. Time-Domain Ab Initio Simulation of Electron and Hole Relaxation Dynamics in a Single-Wall Semiconducting Carbon Nanotube. *Phys. Rev. Lett.* **2006**, *96*, 187401.

(11) Akimov, A. V.; Prezhdo, O. V. Theory of Solar Energy Materials. *J. Phys.: Condens. Matter* **2015**, *27*, 130301.

(12) Craig, C.; Duncan, W.; Prezhdo, O. Trajectory Surface Hopping in the Time-Dependent Kohn-Sham Approach for Electron-Nuclear Dynamics. *Phys. Rev. Lett.* **2005**, *95*, 163001.

(13) Wong, J. C.; Li, L.; Kanai, Y. Size Dependence and Role of Decoherence in Hot Electron Relaxation within Fluorinated Silicon Quantum Dots: A First-Principles Study. *J. Phys. Chem. C* **2018**, *122*, 29526−29536.

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

(15) Wang, L.; Long, R.; Prezhdo, O. V. Time-Domain Ab Initio Modeling of Photoinduced Dynamics at Nanoscale Interfaces. *Annu. Rev. Phys. Chem.* **2015**, *66*, 549−579.

(16) Ma, J.; Wang, L.-W. The Nature of Electron Mobility in Hybrid Perovskite CH3NH3PbI3. *Nano Lett.* **2017**, *17*, 3646−3654.

(17) Smith, B.; Akimov, A. V. Hot Electron Cooling in Silicon Nanoclusters via Landau−Zener Nonadiabatic Molecular Dynamics: Size Dependence and Role of Surface Termination. *J. Phys. Chem. Lett.* **2020**, *11*, 1456−1465.

(18) Lystrom, L.; Tamukong, P.; Mihaylov, D.; Kilina, S. Phonon-Driven Energy Relaxation in PbS/CdS and PbSe/CdSe Core/Shell Quantum Dots. *J. Phys. Chem. Lett.* **2020**, *11*, 4269−4278.

(19) Chen, S.; Ullah, N.; Zhao, Y.; Zhang, R. Nonradiative Excited-State Decay via Conical Intersection in Graphene Nanostructures. *ChemPhysChem* **2019**, *20*, 2754−2758.

(20) Agrawal, S.; Lin, W.; Prezhdo, O. V.; Trivedi, D. J. Ab Initio Quantum Dynamics of Charge Carriers in Graphitic Carbon Nitride Nanosheets. *J. Chem. Phys.* **2020**, *153*, 054701.

(21) Smith, B.; Shakiba, M.; Akimov, A. V. Nonadiabatic Dynamics in Si and CdSe Nanoclusters: Many-Body vs. Single-Particle Treatment of Excited States. *J. Chem. Theory Comput.* **2021**, *17*, 678−693.

(22) Smith, B.; Shakiba, M.; Akimov, A. V. Crystal Symmetry and Static Electron Correlation Greatly Accelerate Nonradiative Dynamics in Lead Halide Perovskites. *J. Phys. Chem. Lett.* **2021**, *12*, 2444−2453.

(23) Lin, Y.; Akimov, A. V. Dependence of Nonadiabatic Couplings with Kohn−Sham Orbitals on the Choice of Density Functional: Pure vs Hybrid. *J. Phys. Chem. A* **2016**, *120*, 9028−9041.

(24) Jiang, X.; Zheng, Q.; Lan, Z.; Saidi, W. A.; Ren, X.; Zhao, J. Real-Time GW-BSE Investigations on Spin-Valley Exciton Dynamics in Monolayer Transition Metal Dichalcogenide. *Sci. Adv.* **2021**, *7*, No. eabf3759.

(25) Hammes-Schiffer, S.; Tully, J. C. Nonadiabatic Transition State Theory and Multiple Potential Energy Surface Molecular Dynamics of Infrequent Events. *J. Chem. Phys.* **1995**, *103*, 8528−8537.

(26) Nangia, S.; Jasper, A. W.; Miller, T. F.; Truhlar, D. G. Army Ants Algorithm for Rare Event Sampling of Delocalized Nonadiabatic Transitions by Trajectory Surface Hopping and the Estimation of Sampling Errors by the Bootstrap Method. *J. Chem. Phys.* **2004**, *120*, 3586−3597.

(27) Lingerfelt, D. B.; Williams-Young, D. B.; Petrone, A.; Li, X. Direct Ab Initio (Meta-)Surface-Hopping Dynamics. *J. Chem. Theory Comput.* **2016**, *12*, 935−945.

(28) He, J.; Fang, W.-H.; Long, R.; Prezhdo, O. V. Superoxide/Peroxide Chemistry Extends Charge Carriers' Lifetime but Undermines Chemical Stability of CH3NH3PbI3 Exposed to Oxygen: Time-Domain Ab Initio Analysis. *J. Am. Chem. Soc.* **2019**, *141*, 5798−5807.

(29) Zhang, Z.; Long, R. Doping-Induced Rapid Decoherence Suppresses Charge Recombination in Mono/Divalent Cation Mixed Perovskites from Nonadiabatic Molecular Dynamics Simulation. *J. Phys. Chem. Lett.* **2019**, *10*, 3433−3439.

(30) Zhang, Z.; Fang, W.-H.; Tokina, M. V.; Long, R.; Prezhdo, O. V. Rapid Decoherence Suppresses Charge Recombination in Multi-Layer 2D Halide Perovskites: Time-Domain Ab Initio Analysis. *Nano Lett.* **2018**, *18*, 2459−2466.

(31) Wang, L.; Li, Q.; Shuai, Z.; Chen, L.; Shi, Q. Multiscale Study of Charge Mobility of Organic Semiconductor with Dynamic Disorders. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3309−3314.

(32) Nijamudheen, A.; Akimov, A. V. Excited-State Dynamics in Two-Dimensional Heterostructures: SiR/TiO2 and GeR/TiO2 (R = H, Me) as Promising Photocatalysts. *J. Phys. Chem. C* **2017**, *121*, 6520−6532.

(33) Akimov, A. V. Stochastic and Quasi-Stochastic Hamiltonians for Long-Time Nonadiabatic Molecular Dynamics. *J. Phys. Chem. Lett.* **2017**, *9*, 5190.

(34) Kulichenko, M.; Smith, J. S.; Nebgen, B.; Li, Y. W.; Fedik, N.; Boldyrev, A. I.; Lubbers, N.; Barros, K.; Tretiak, S. The Rise of Neural Networks for Materials and Chemical Dynamics. *J. Phys. Chem. Lett.* **2021**, *12*, 6227−6243.

(35) Zubatiuk, T.; Nebgen, B.; Lubbers, N.; Smith, J. S.; Zubatyuk, R.; Zhou, G.; Koh, C.; Barros, K.; Isayev, O.; Tretiak, S. Machine Learned

Hückel Theory: Interfacing Physics and Deep Neural Networks. *J. Chem. Phys.* **2021**, *154*, 244108.
(36) 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*, 3828−3834.
(37) Dral, P. O.; Barbatti, M.; Thiel, W. Nonadiabatic Excited-State Dynamics with Machine Learning. *J. Phys. Chem. Lett.* **2018**, *9*, 5660−5663.
(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*, 1930−1938.
(39) Xue, B.-X.; Barbatti, M.; Dral, P. O. Machine Learning for Absorption Cross Sections. *J. Phys. Chem. A* **2020**, *124*, 7199−7210.
(40) Dral, P. O.; Ge, F.; Xue, B.-X.; Hou, Y.-F.; Pinheiro, M.; Huang, J.; Barbatti, M. MLatom 2: An Integrative Platform for Atomistic Machine Learning. *Top. Curr. Chem.* **2021**, *379*, 27.
(41) Dral, P. O.; Barbatti, M. Molecular Excited States through a Machine Learning Lens. *Nat. Rev. Chem.* **2021**, *5*, 388−405.
(42) Westermayr, J.; Marquetand, P. Machine Learning and Excited-State Molecular Dynamics. *Mach. Learn.: Sci. Technol.* **2020**, *1*, 043001.
(43) Virshup, A. M.; Chen, J.; Martínez, T. J. Nonlinear Dimensionality Reduction for Nonadiabatic Dynamics: The Influence of Conical Intersection Topography on Population Transfer Rates. *J. Chem. Phys.* **2012**, *137*, 22A519.
(44) Li, X.; Hu, D.; Xie, Y.; Lan, Z. Analysis of Trajectory Similarity and Configuration Similarity in On-the-Fly Surface-Hopping Simulation on Multi-Channel Nonadiabatic Photoisomerization Dynamics. *J. Chem. Phys.* **2018**, *149*, 244104.
(45) Li, X.; Xie, Y.; Hu, D.; Lan, Z. Analysis of the Geometrical Evolution in On-the-Fly Surface-Hopping Nonadiabatic Dynamics with Machine Learning Dimensionality Reduction Approaches: Classical Multidimensional Scaling and Isometric Feature Mapping. *J. Chem. Theory Comput.* **2017**, *13*, 4611−4623.
(46) Hughes, K. H.; Christ, C. D.; Burghardt, I. Effective-Mode Representation of Non-Markovian Dynamics: A Hierarchical Approximation of the Spectral Density. II. Application to Environment-Induced Nonadiabatic Dynamics. *J. Chem. Phys.* **2009**, *131*, 124108.
(47) Mangan, S. M.; Zhou, G.; Chu, W.; Prezhdo, O. V. Dependence between Structural and Electronic Properties of $CsPbI_3$: Unsupervised Machine Learning of Nonadiabatic Molecular Dynamics. *J. Phys. Chem. Lett.* **2021**, *12*, 8672−8678.
(48) Wu, Y.; Prezhdo, N.; Chu, W. Increasing Efficiency of Nonadiabatic Molecular Dynamics by Hamiltonian Interpolation with Kernel Ridge Regression. *J. Phys. Chem. A* **2021**, *125*, 9191−9200.
(49) Brownlee, J. *Deep Learning for Time Series Forecasting. Predict the Future with MLPs, CNNs, and LSTMs in Python*; Machine Learning Mastery, 2021.
(50) Mukamel, S. *Principles of Nonlinear Optical Spectroscopy*; Oxford University Press: New York, 1995.
(51) Madrid, A. B.; Hyeon-Deuk, K.; Habenicht, B. F.; Prezhdo, O. V. Phonon-Induced Dephasing of Excitons in Semiconductor Quantum Dots: Multiple Exciton Generation, Fission, and Luminescence. *ACS Nano* **2009**, *3*, 2487−2494.
(52) Igel, C.; Hüsken, M. Improving the Rprop Learning Algorithm. *Proceedings of the Second ICSC International Symposium on Neural Computation*; ICSC Academic Press, 2000; pp 115−121.
(53) Rojas, R. Fast Learning Algorithms. In *Neural Networks*; Springer: Berlin, Heidelberg, 1996; pp 183−225. DOI: 10.1007/978-3-642-61068-4_8 (accessed 12-15-2021).
(54) Loshchilov, I.; Hutter, F. Decoupled Weight Decay Regularization. *arXiv (Computer Science, Machine Learning)*, November 14, 2017, 1711.05101 ver. 3, 2019, https://arxiv.org/abs/1711.05101 (accessed 2021).
(55) Akimov, A. V. Libra: An Open-Source "Methodology Discovery" Library for Quantum and Classical Dynamics Simulations. *J. Comput. Chem.* **2016**, *37*, 1626−1649.
(56) Akimov, A. V.; Smith, B.; Shakiba, M.; Sato, K.; Temen, S.; Li, W.; Sun, X.; Stippell, L.; Chan, M. *Libra*, ver. 5.1.0; Zenodo, 2021. DOI: 10.5281/zenodo.5708303.
(57) Akimov, A. V. Excited State Dynamics in Monolayer Black Phosphorus Revisited: Accounting for Many-Body Effects. *J. Chem. Phys.* **2021**, *155*, 134106.
(58) Nelson, T.; Fernandez-Alberti, S.; Roitberg, A. E.; Tretiak, S. Nonadiabatic Excited-State Molecular Dynamics: Treatment of Electronic Decoherence. *J. Chem. Phys.* **2013**, *138*, 224111.
(59) Akimov, A. V. A Simple Phase Correction Makes a Big Difference in Nonadiabatic Molecular Dynamics. *J. Phys. Chem. Lett.* **2018**, *9*, 6096−6102.
(60) Fernandez-Alberti, S.; Roitberg, A. E.; Nelson, T.; Tretiak, S. Identification of Unavoided Crossings in Nonadiabatic Photoexcited Dynamics Involving Multiple Electronic States in Polyatomic Conjugated Molecules. *J. Chem. Phys.* **2012**, *137*, 014512.
(61) Temen, S.; Akimov, A. V. A Simple Solution to Trivial Crossings: A Stochastic State Tracking Approach. *J. Phys. Chem. Lett.* **2021**, *12*, 850−860.
(62) Wang, L.; Prezhdo, O. V. A Simple Solution to the Trivial Crossing Problem in Surface Hopping. *J. Phys. Chem. Lett.* **2014**, *5*, 713−719.
(63) Qiu, J.; Bai, X.; Wang, L. Crossing Classified and Corrected Fewest Switches Surface Hopping. *J. Phys. Chem. Lett.* **2018**, *9*, 4319−4325.
(64) Jaeger, H. M.; Fischer, S.; Prezhdo, O. V. Decoherence-Induced Surface Hopping. *J. Chem. Phys.* **2012**, *137*, 22A545.
(65) Smith, B.; Akimov, A. V. A Comparative Analysis of Surface Hopping Acceptance and Decoherence Algorithms within the Neglect of Back-Reaction Approximation. *J. Chem. Phys.* **2019**, *151*, 124107.
(66) Subotnik, J. E. Augmented Ehrenfest Dynamics Yields a Rate for Surface Hopping. *J. Chem. Phys.* **2010**, *132*, 134112.
(67) Jain, A.; Alguire, E.; Subotnik, J. E. An Efficient, Augmented Surface Hopping Algorithm That Includes Decoherence for Use in Large-Scale Simulations. *J. Chem. Theory Comput.* **2016**, *12*, 5256−5268.
(68) Akimov, A. V. *AkimovLab/Project-NBRA-Machine-Learninig: NBRA-Based NAMD Dynamics with Machine-Learned Hamiltonians*; Zenodo, 2021. DOI: 10.5281/zenodo.5708673.

https://doi.org/10.1021/acs.jpclett.1c03823
J. Phys. Chem. Lett. 2021, 12, 12119−12128