The PYXAID Program for Non-Adiabatic Molecular Dynamics in Condensed Matter Systems

Alexey V. Akimov†‡ and Oleg V. Prezhdo*†

†Department of Chemistry, University of Rochester, Rochester, New York 14627, United States
‡Chemistry Department, Brookhaven National Laboratory, Upton, New York 11973, United States

Supporting Information

ABSTRACT: This work introduces the PYXAID program, developed for non-adiabatic molecular dynamics simulations in condensed matter systems. By applying the classical path approximation to the fewest switches surface hopping approach, we have developed an efficient computational tool that can be applied to study photoinduced dynamics at the ab initio level in systems composed of hundreds of atoms and involving thousands of electronic states. The technique is used to study in detail the ultrafast relaxation of hot electrons in crystalline pentacene. The simulated relaxation occurs on a 500 fs time scale, in excellent agreement with experiment, and is driven by molecular lattice vibrations in the 200–250 cm⁻¹ frequency range. The PYXAID program is organized as a Python extension module and can be easily combined with other Python-driven modules, enhancing user-friendliness and flexibility of the software. The source code and additional information are available on the Web at the address http://gdriv.es/pyxaid. The program is released under the GNU General Public License.

1. INTRODUCTION

An increased interest in photochemical processes, taking place in a wide variety of molecular and condensed phase systems, has stimulated a large number of theoretical and experimental studies. Investigations performed in the time domain present a particularly useful perspective, since they probe the non-equilibrium phenomena as they occur in nature, eliminating steady-state, quasi-equilibrium, linear response, and similar assumptions needed to introduce rate constants, transition states, transport coefficients, and other commonly used parameters. We refer the reader to recent reviews on these topics for more details.¹−¹³ Various groups perform simulations of photoinduced dynamics. Examples include charge-carrier transfer and recombination in photovoltaic and photocatalytic applications,¹⁴−²⁶ multiple exciton generation,²⁵−²⁹ and singlet fission,³⁰−³⁵ in inorganic and organic semiconductors, elastic and inelastic electron–phonon scattering in quantum dots,³³−³⁶ energy transfer in nanoscale materials,³⁷−⁴⁰ photoisomerization in conjugated systems,⁴¹−⁴⁵ and nuclear motions in photo-switches,⁴⁶−⁴⁸ and molecular machines.⁴⁹ However, the number of computational investigations is significantly smaller than the amount of experimental studies. This is partially because simulation of real-time dynamics of photochemical processes remains at the state-of-the-art level, requiring significant efforts and resolution of multiple conceptual problems. As a result, tools for theoretical modeling of excited state dynamics are not generally available or are restricted by severe limitations on system size.

A number of programs capable of performing non-adiabatic molecular dynamics (NA-MD) simulations have been reported, including Newton-X,⁴⁸,⁴⁹ MOLPRO,⁵⁰,⁵¹ MNDO,⁴²,⁵² CPMD,⁵³,⁵⁴ and Octopus.⁵⁵ The programs are designed primarily for relatively small systems, since they employ advanced electronic structure techniques required for accurate description of excited state properties. They are hard to apply to large-scale condensed matter systems, which can be treated at a less sophisticated electronic structure level. Further, these codes typically employ the most basic NA-MD approaches, such as the Ehrenfest method,⁵⁸−⁶³ and surface hopping.⁶⁴,⁶⁵ The nuclear motions are treated classically, excluding quantum effects, such as decoherence,⁶⁶−⁶⁸ which are particularly important in condensed phases.

As an alternative, the present work reports the PYXAID program (PYthon eXtension for Ab Initio Dynamics)—an open-source, flexible, and computationally efficient implementation of the NA-MD methodology in the framework of Kohn–Sham (KS) density functional theory (DFT).⁶⁹ It implements a number of basic and more advanced functionalities, including fewest-switches surface hopping (FSSH),⁶⁴ decoherence-induced surface hopping (DISH),⁷⁰ multielectron adiabatic representation of the time-dependent KS (TD-KS) equations, and explicit field–matter interactions. The classical path approximation (CPA) achieves additional and considerable...
computational savings. The package is interfaced with the Quantum Espresso (QE)\textsuperscript{71} distribution, which is used as an efficient driver for ab initio adiabatic electronic structure and MD calculations. Particular attention is dedicated to stability, efficiency, and accuracy of solutions to the TD-KS equations in the adiabatic representation.

2. THEORY AND METHODS

2.1. Non-Adiabatic Molecular Dynamics. A rigorous solution of the time-dependent (TD) Schrödinger equation (SE):

\[
i\hbar \frac{\partial \Psi(r_n^N, R^N, t)}{\partial t} = H(r_n^N, R^N, t)\Psi(r_n^N, R^N, t)
\]

where \( r_n^N \) and \( R^N \) are the coordinates of \( n \) electrons and \( N \) nuclei, respectively. For simplicity of notation, in further discussion, these quantities will be referred to as \( r \) and \( R \), respectively. The Hamiltonian \( H(r, R, t) \) can be described by linear combinations of products of electronic wave functions for different adiabatic states, \( \Phi(r; R(t)) \), and nuclear wave functions, \( \chi(t, R(t)) \), representing vibrational levels associated with the electronic states:

\[
\Psi(r, R, t) = \sum_i \chi(t, R(t))\Phi(r; R(t))
\]

The dynamics of nuclear wavepackets correlated with different electronic states can then be obtained by projecting the solution, eq 2, onto the corresponding electronic states:

\[
i\hbar \frac{\partial \chi_j(t, R(t))}{\partial t} = \sum_j \left( T_{\text{nucl}} + \tilde{E}_j(t, R) \right)\delta_{ij} - \hbar^2 \frac{\nabla \Phi_j(r)}{M} \tilde{\nabla}_j
\]

where \( d^{(0)}_{ij} \) are the \( n \)th order NA couplings between electronic states \( i \) and \( j \):

\[
\tilde{d}^{(1)}_j = \langle \Phi_i(r; R(t)) | \nabla \Phi_j(r; R(t)) \rangle
\]

\[
\tilde{d}^{(2)}_j = \langle \Phi_i(r; R(t)) | \nabla^2 \Phi_j(r; R(t)) \rangle
\]

In eq 3 and subsequent discussion, the product \( \tilde{d}^{(1)}_j \tilde{\nabla} \) is understood as a scalar product of two 3N-dimensional vectors \( \tilde{d}^{(1)}_j \) and \( \tilde{\nabla} \).

The quantities \( \tilde{E}_j(t, R) \) are the adiabatic potential energy surfaces (PES), corresponding to electronic states \( \Phi(r; R(t)) \). They are the eigenvalues of stationary electronic SE, parametrically dependent on the nuclear coordinates \( R(t) \):

\[
H_0(r, t; R)\Phi_i(r, t; R(t)) = \tilde{E}_j(t, R)\Phi_j(r, t; R(t))
\]

Similarly to \( \tilde{E}_j(t, R) \), the term \(-\hbar^2\sum_j d^{(2)}_{ij}/2M \) in eq 3 is effectively a number, parametrically dependent on the nuclear coordinates. The diagonal terms can be considered as a correction to the PES:

\[
E_i(R) = \tilde{E}_i(R) - \hbar^2 \sum_j \frac{d^{(2)}_{ij}}{2M}
\]

The off-diagonal terms are often assumed to be of higher order than the \( \hbar^2\sum_j d^{(2)}_{ij}/M \) term and small, and are, therefore, neglected\textsuperscript{72}. In some cases, however, these terms may become particularly important.\textsuperscript{72} With such redefinition of PES, eq 3 can be rewritten in a simplified form:

\[
i\hbar \frac{\partial \chi_j(t, R(t))}{\partial t} = \sum_j \left( T_{\text{nucl}} + E_j(t, R) \right)\delta_{ij} - \hbar^2 \frac{\nabla \Phi_j(r)}{M} \tilde{\nabla}_j
\]

A correlated evolution of the electronic and nuclear degrees of freedom can be described by solving eqs 5 and 7 that provide a solution of the general TD-SE, eq 1. The solution of eq 5 constitutes the stationary SE problem, involving well-developed machinery with decades of history. The solution of eq 7 is less straightforward and can be obtained in a variety of ways, depending on the approximations used\textsuperscript{60,63,76–89}. Among them, the trajectory surface hopping techniques, exemplified by fewest-switches surface hopping\textsuperscript{82,84,90,91}, is a popular family of approximations, due to their conceptual simplicity, ease of implementation, high efficiency, and reasonable accuracy. The next section summarizes FSSH, introduces some notation and terminology used in the program implementation and interface, and highlights the distinctive features of our implementation.

Equation 7 deserves further analysis. In the adiabatic limit, the variation of the electronic wave function with time is slow, the NA couplings vanish, the electronic states are decoupled, and transitions are prohibited. In such case, the system evolves exclusively in the state in which it starts, most typically, in the ground electronic state. The evolution of the nuclei is adequately described within the framework of adiabatic MD. Adiabatic MD is often treated classically, with the PES defined by computationally efficient analytic functions, giving rise to the popular molecular mechanics methods.

The operators \( T_{\text{nucl}} \) and \(-\hbar^2\sum_j (\tilde{d}^{(1)}_{ij}/M) \tilde{\nabla} \) in eq 7 are nonlocal in the position representation. They act on the nuclear wave functions, \( \chi(t, R(t)) \), and give rise to nuclear quantum effects, such as tunneling and zero-point energy. They lead to the uncertainty principle, effectively turning the classical phase space point, \( \chi(t, R(t)) = \{ R, P \} = \delta(R - \tilde{R})\delta(P - \tilde{P}) \), into a wavepacket, in the simplest case, a Gaussian, \( \chi(t, R(t)) = G(R, \tilde{R}, P, \tilde{P}) \). Often, an opposite limit is considered—as follows from the representation of the delta functions by a limit of Gaussian distributions with infinitesimal width. In the limit of \( \hbar \to 0 \), the Gaussian wavepacket becomes a delta function representing a classical particle\textsuperscript{85,92,93}. Depending on the desired approximation for the adiabatic or non-adiabatic dynamics, one can choose the fully quantum description, utilizing the standard quantum-mechanical definition of the operators, the fully classical description, in which both operators are approximated using classical momenta: \( T_{\text{nucl}} = \langle P^2/2M \rangle \) and \(-\hbar^2\sum_j (\tilde{d}^{(1)}_{ij}/M) \tilde{\nabla} = -i\hbar\tilde{d}^{(1)}_{ij}(P/M) \), or a semiclassical scheme, such as QHD\textsuperscript{94,97} in which the quantum nature of the operators is captured via higher order correlators of classical-like momentum and position variables. The fully classical limit straightforwardly follows from standard quantum-classical correspondence rules. The fully classical limit lacks electron–nuclear correlation and other quantum effects that may lead to complex quantum-interference phenomena, including loss of electronic coherence. As an efficient alternative to the fully quantum description of coupled
2.2. Semiclassical NA-MD within the DFT Framework

**Time-Dependent Schrödinger Equation along the MD Trajectory.** The PYXAID program focuses on NA-MD, in which nuclear degrees of freedom are treated classically, and one uses the following substitutions in eq 7:

\[ T_{\text{nucl}} \equiv -\frac{\hbar^2}{2M} \nabla^2 + \frac{P^2}{2M} \]

\[ -\hbar^2 \frac{d}{dt} \left( \begin{array}{c} \Phi_i \end{array} \right) = -i\hbar \frac{\nabla^2}{M} \]

The solution of the TD-SE is expressed on the basis of the solutions of the stationary SE, \( \{ \Phi_i \} \), where the basis size:

\[ \Psi(r_i, R, t) = \sum_{i=0}^{N_b-1} c_i(t) \Phi_i(r_i) \]

The notation \( \Phi_i(r_i) \) indicates functional dependence of the electronic wave functions, \( \Phi_i \), on the electronic coordinates, \( r_i \), and its parametric dependence on the nuclear coordinates, \( R \). The ansatz, eq 10, simplifies the TD-SE, eq 7, to the equations of motion for the amplitudes, \( c_i(t) \):

\[ i\hbar \frac{dc_i}{dt} = \sum_{j=0}^{N_b-1} (c_j(t) \delta_{ij} - i\hbar d_{ij} c_j) \]

where

\[ d_{ij} = \left( \Phi_i \right| \alpha \frac{\partial \Phi_j}{\partial \alpha} ) = d_{ij}(\alpha) \frac{\tilde{p}}{M} \]

Equation 11 can be written in a matrix-vector form:

\[ i\hbar \frac{d\vec{c}}{dt} = \vec{H}\vec{c} \]

with the matrix elements of the Hamiltonian \( \vec{H} \) defined as

\[ H_{ii} = \epsilon_i \]  

\[ H_{ij} = H_{ij}^{\text{NAC}} = -i\hbar d_{ij} \]

In a more general case, the matrix elements, eq 14, can include other effects, such as electro-magnetic field\[^{53,24,101}\] or spin–orbit coupling:\[^{101,102}\]

\[ H_{ij} = H_{ij}^{\text{NAC}} + H_{ij}^{\text{field}} + H_{ij}^{\text{SO}} \]

Spin–orbit coupling can be included by solving the relativistic version of the Kohn–Sham equations within the framework of non-collinear magnetism.\[^{103–105}\] In such an approach, the conventional “spin-up” and “spin-down” orbitals are substituted with the two-component spinor orbitals \( \Phi_\alpha = (\begin{array}{c} \alpha_1 \\ \alpha_2 \end{array} ) \) and \( \Phi_\beta = (\begin{array}{c} \beta_1 \\ \beta_2 \end{array} ) \). These new wave functions are coupled to each other, since in general, they both have nonzero values for each component. In the simplest case, when the non-collinearity is not included, the above spinors reduce to the well-known spin–orbitals: \( \Phi_\alpha = (\begin{array}{c} \alpha_1 \\ 0 \end{array} ) \) and \( \Phi_\beta = (\begin{array}{c} 0 \\ \beta_2 \end{array} ) \) are the spin basis vectors and \( \alpha_1 \) and \( \beta_2 \) are the spatial parts of the corresponding spin–orbitals. It is important to note that the entire machinery of NA-MD and related calculations will not change, and the above expressions for the NA coupling will remain the same. The only difference will be in the nature of the basis in which the orbitals are expressed. At this point, we only focus on the non-spin-polarized calculations, and the generalization to spin–orbit couplings will be reported in our future works.

Solution of eq 13 describes coherent time evolution of electronic states, coupled to evolution of the nuclear subsystem, and is expressed as a superposition of the adiabatic states. Such a superposition is a convenient mathematical construction but is not a physical reality. The laws of quantum mechanics and measurement theory dictate that the system can only exist in a given pure state, once the measurement has been performed. To overcome this conceptual difficulty, a stochastic fewest switches surface hopping (FSSH) has been developed, pioneered by Tully.\[^{64}\] The detailed discussion of original FSSH algorithm is presented in section A of the Supporting Information.

The standard FSSH algorithm determines the fate of the system stochastically on-the-fly, producing different nuclear trajectories depending on realization of the stochastic process. Therefore, the electronic structure calculations must be done at each nuclear time step and for all trajectories, increasing the computational expenses linearly with the total trajectory time, determined as the sum of the lengths of all trajectories in the ensemble. The on-the-fly approach is virtually unavoidable in molecular systems in which photoexcitation may induce substantial changes of the nuclear geometry. Such a requirement is dictated by the strong back-reaction of the electronic degrees of freedom on the nuclear ones, which directly affects future dynamics. In contrast, the excitation has little effect on the structure of large, condensed matter systems, and variations in the nuclear geometry for different electronic states may be smaller than thermal fluctuations of atoms, indicating negligible back-reaction. Then, one can apply a more efficient computational technique, known as the CPA. Presently, the CPA is, perhaps, the only way to study electron–nuclear dynamics in sufficiently large systems at the \textit{ab initio} level of theory.

**The FSSH Method within the Classical Path Approximation (FSSH-CPA).** The CPA is based on the assumption that nuclear dynamics of the system remains unaffected by the dynamics of the electronic degrees of freedom. The electronic dynamics remains driven by the nuclear dynamics. The electronic Hamiltonian depends parametrically on the classical nuclear variables, which evolve along a MD trajectory. The MD trajectory is predetermined and independent of the electronic evolution. For example, one can use a trajectory computed for the ground electronic state at a finite temperature. Under such an assumption, only one or few MD trajectories are required, and a large number of stochastic realizations of the FSSH algorithm can...
be computed along the same trajectory. The CPA drastically limits the number of expensive electronic structure calculations, thus greatly enhancing the computational efficiency of the method.

The CPA is valid when kinetic energy is the primary source of nuclear dynamics. It works when the electronic excitation does not result in significant structural changes, such as reorganization, fragmentation, isomerization, etc., or at relatively large temperatures, when thermal energy is sufficient to drive such changes. The required condition is often met in condensed matter systems, such as inorganic and organic crystals, interfaces, and solutions. Even if the relevant nuclear rearrangements are impossible. Then, the matrix elements of the Hamiltonian in eq 4q are modified analytically, since its numerical evaluation requires a large number of electronic structure calculations. The latter is a single number and can be easily obtained numerically along the MD trajectory.109

If light–matter interactions are included explicitly, eq 17 should be modified

$$b_{i\rightarrow j}(t) = \begin{cases} \exp\left(-\frac{E_j - E_i - E_{ho}}{kT}\right) & E_j > E_i + E_{ho} \\ 1 & E_j \leq E_i + E_{ho} \end{cases}$$

(18)

where $E_{ho} = \hbar\omega$ is the energy of the absorbed photon. The modification stems from energy conservation in the combined light–matter–photon system. In particular, absorption of photons with energies below the minimal energy gap becomes practically impossible. Then, the matrix elements of the Hamiltonian in eq 13 include contributions from the transition dipole moments, eq 15. This affects both solution of the TD-SE and the quantity $g_{i\rightarrow j}(t)$, and hence the probabilities of electronic transitions.

**Basis Set.** Construction of the set of functions, $\{\Phi_i\}$, $i = 1, \ldots, N_0 - 1$, used as a basis for solving the TD-SE deserves special attention. The TD-KS theory employs a Slater determinant composed of TD single-particle KS orbitals in order to represent the electron density. Following the analogy with the wave function methods, one can represent the TD Slater determinant in the basis of Slater determinants composed of the adiabatic one-electron KS orbitals, as proposed by Craig, Duncan, and Prezhdo.110,111

$$\Phi_1(1, 2, \ldots, N) = \hat{A}[\phi_{i_1}(1)\phi_{i_2}(2)\ldots\phi_{i_N}(N)]$$

(19)

where $\phi_i(n)$ is the adiabatic KS spin–orbital ($i = 1, \ldots, N$) dependent on the coordinates of the $n$th electron, $N \equiv N_{KS}$ is the number of the KS orbitals included in the NA-MD model, and $\hat{A}$ denotes the antisymmetrization operator. In our current
implementation, we assume that the energy levels for spin-orbitals \( \omega_i = \sigma \omega_0 \) do not depend on spin-component \( \sigma \), eliminating explicit spin dependence of the Hamiltonian in eq 13. As so, the current implementation is limited to electronic transitions conserving spin symmetry. However, from the fundamental point of view, the methodology described in this paper is still applicable to more general electronic transitions involving spin flips. This capability will be implemented and reported in our later works.

In many practical applications, the electron excitation and relaxation dynamics involves a relatively small manifold of KS orbitals, whose range is determined by the excitation energies accessible experimentally. For example, under irradiation by visible light, an electron is typically promoted in a bulk or nanoscale material, from KS orbitals near the top of the valence band (VB) to KS orbitals near the bottom of the conduction band (CB). Deep VB and CB states can be important for UV excitation or high-energy ionization processes. Thus, all KS orbitals obtained from the electronic structure calculations may not be necessary for the TD-KS calculations, and one may wish to use a restricted subset. For this reason, our code uses an active space, similar in spirit to the one used in the complete active space (CAS) methodologies.

A subset of \( N_{KS} \) KS orbitals to be used in construction of the basis configurations forms the active space and needs to be defined before NA-MD can be executed (Figure 1a). Once the active space is selected, one can consider the ground and various excited state configurations by promoting a given number of electrons from the occupied orbitals included in the active space to the virtual KS levels (Figure 1b).

A hypothetical four-electron system is considered in Figure 1. One can obtain a set of 5 KS orbitals and their energies, \( (\varphi_i, \varepsilon_i) \), \( i = 1, ..., 5 \), with the standard electronic structure calculations. This number 5 is specified by the number of molecular orbitals to be computed (e.g., by the nbnd keyword in the QE code). One then decides to construct a basis set, in which the excitations are allowed only among HOMO, LUMO, and LUMO+1, defining an active space \( (\varphi_i, \varepsilon_i) \) with \( i = 2, 3, \) and 4. The basis state configurations are defined by promoting zero (ground state configuration \( \Phi_0 \)), one (singly excited configuration \( \Phi_1 \)), or two (doubly excited configuration \( \Phi_2 \)) electrons from the HOMO to either LUMO or LUMO+1 or both.

The Hamiltonian matrix elements for the multielectron basis configurations are determined on the basis of the KS orbitals. The excitation energies are represented by the differences in the energies of the KS orbitals occupied in the ground \( \{ \varepsilon_i; i \in GS \} \) and excited state \( \{ \varepsilon_i; i \in ES \} \) configurations. The diagonal elements of the Hamiltonian are

\[
E_i \equiv H_{ii} = \sum_{i \in ES} \varepsilon_i - \sum_{i \in GS} \varepsilon_i
\]

The NA couplings determine the off-diagonal elements of the many-electron Hamiltonian. In particular, the coupling between the many-electronic states \( i \) and \( j \), represented by Slater determinants \( \Phi_0 \) and \( \Phi_1 \) respectively, is given by

\[
H_{ij} \equiv -i\hbar \left( \Phi_0 \frac{\partial}{\partial t} \Phi_j \right)
= -i\hbar \left( \prod_{k=1}^{N} \varphi_k(k) \frac{\partial}{\partial t} \prod_{k=1}^{N} \varphi_k(k') \right)
= -i\hbar \left( \prod_{k=1}^{N} \varphi_k(k) \sum_{k=1}^{N} \frac{\partial}{\partial t} \varphi_k(k') \prod_{k'=1, k' \neq k}^{N} \varphi_k(k') \right)
= \sum_{k=1}^{N} -i\hbar \left( \varphi_k(k') \frac{\partial}{\partial t} \varphi_k(k') \right) \prod_{k'=1, k' \neq k}^{N} \varphi_k(k')
= \sum_{k=1}^{N} h_{ik/k'} \prod_{k'=1, k' \neq k}^{N} \delta_{ik/k'}
\]

where \( h_{ik/k'} = -i\hbar (\varphi_i(k')(\partial/\partial t)\varphi_k(k')) = -i\hbar d_{ik/k'} \) is an off-diagonal matrix element of the one-electron Hamiltonian, that arises from the NA coupling between the KS orbitals \( \varphi_i(k') \) and \( \varphi_k(k') \) depending on the coordinates of the \( k' \)-th electron.

The NA couplings are nonzero only for pairs of configurations that differ by no more than one KS spin-orbital. For example, the coupling between the configurations \( \Phi_0 \) and \( \Phi_1 \) or \( \Phi_1 \) and \( \Phi_2 \) shown in Figure 1b are nonzero, while the couplings for all other pairs are zero. The following equation

\[
H = \begin{pmatrix}
E_0 & -i\hbar \left( \Phi_0 \frac{\partial}{\partial t} \Phi_1 \right) & -i\hbar \left( \Phi_0 \frac{\partial}{\partial t} \Phi_2 \right) & -i\hbar \left( \Phi_0 \frac{\partial}{\partial t} \Phi_3 \right) \\
-i\hbar \left( \Phi_1 \frac{\partial}{\partial t} \Phi_0 \right) & E_1 & -i\hbar \left( \Phi_1 \frac{\partial}{\partial t} \Phi_2 \right) & -i\hbar \left( \Phi_1 \frac{\partial}{\partial t} \Phi_3 \right) \\
-i\hbar \left( \Phi_2 \frac{\partial}{\partial t} \Phi_0 \right) & -i\hbar \left( \Phi_2 \frac{\partial}{\partial t} \Phi_1 \right) & E_2 & -i\hbar \left( \Phi_2 \frac{\partial}{\partial t} \Phi_3 \right) \\
-i\hbar \left( \Phi_3 \frac{\partial}{\partial t} \Phi_0 \right) & -i\hbar \left( \Phi_3 \frac{\partial}{\partial t} \Phi_1 \right) & -i\hbar \left( \Phi_3 \frac{\partial}{\partial t} \Phi_2 \right) & E_3
\end{pmatrix}
\]
shows the correspondence between the many-electron Hamiltonian matrix elements, eq 22, and the one-electron properties, eq 23, for the model presented in Figure 1.

We should note that the approximation, eq 20, includes electron correlations implicitly via dependence of the KS eigenvalues on the exchange-correlation functionals used. However, the computationally efficient pure (LDA or GGA) density functionals, such as PW or PBE, suffer from the self-interaction error.112−115 In particular, the DFT band gaps and higher excitation energies obtained at such level of theory are often underestimated. Multixciton states, such as \( \Phi_k \) in Figure 1b, present additional challenges, in particular for strongly correlated materials. For instance, the KS orbital description of multiple excitons provides a good representation for semiconductor quantum dots28,29 but is overly simplistic for molecular crystals.116,117 Excitation energy corrections may be found with higher-level methods or semiempirically,113,118,119,122 or for systems in which the electron interaction is relatively weak, eq 20 gives good results, and energy corrections are not needed.

The PYXAID program provides the possibility of user-defined constant energy corrections, specific for each configuration. Similarly, there is a possibility to rescale the NA couplings to meet special needs of a user. These features of the program may also be useful for abstract model calculations, where one is not required to use accurate parameters for the Hamiltonian, or if one needs to study the effects of energy level shifts and NA coupling variations.

3. PROGRAM ORGANIZATION

3.1. Python Interface. Over the years, Python has emerged into a mature programming language, with an enormous variety of extensions and applications. Such popularity of Python is related to its simple and clear syntax, very similar to that of natural languages. In addition, the programs written in the Python language require no explicit compilation, allowing the user to take full control of the algorithms and data structures, to easily modify them for specific needs, and to add new features. Finally, Python-based programs can be run on different operating systems—a property that enhances their transferability and makes them easy to use on heterogeneous computational facilities. For all these reasons, we have chosen Python to be the interface language between the user and the program. It should be stressed that the whole code is not written exclusively in Python. For computational efficiency, we adopt a mixed scheme, where only relatively simple operations, which may need customization, are exposed to the user via Python scripts, while the computationally demanding routines are efficiently coded using the lower-level C/C++ language. The mixed Python/C/C++ scheme is realized as a Python extension module, created with the help of the Boost.Python libraries.134,135

The Python interface allows flexibility of input format and its execution. All simulation parameters that are normally supplied by input files with a strictly defined format can now be defined inside the Python script, which therefore may be considered as an interactive input file. This philosophy of the input files helps to avoid possible problems associated with the use of specific format and minimizes efforts in learning new syntax. In a Python script, the user only needs to know the standard Python language and a content of the extension module, i.e., which parameters to vary. With the help of such an “interactive input file”, one can easily handle complex data structures, occasionally needed for simulations, in the algorithmic way in the form of small and simple scripts.

3.2. Code Design. The PYXAID code is organized in a three-level structure, schematically shown in Figure 2. The code of the first level is written in a low-level language, namely, C/C++, and contains basic and time-critical parts of the code. It is hidden from the user, since it does not require adjustment or modifications for each specific purpose, unless new algorithms or methodologies are being developed.

The code of the second level is written in Python, which is a higher-level language than C/C++. It contains predefined
routines and execution scenarios. The routines of the first level are directly called from the code of the second level. The predefined scenarios constitute the workflow the user will normally want to utilize. Occasionally, routines of this level may need modifications, depending on specific needs of the user.

Finally, the code of the third level is a collection of dynamic Python/Bash/PBS scripts, containing task-specific information the user will need to modify according to his/her needs. Such scripts usually invoke predefined procedures of the second level of the PYXAID distribution. Knowing how the second- and first-level procedures work, the user can always start with the scripts of the third-level code. Merely modifying input parameters and small pieces of the Python code of the third level may be sufficient to get started with applied calculations.

We want to emphasize that the code of the third level should be mostly considered as a collection of the dynamic (interactive) input files, rather than an integral part of the PYXAID distribution. The word “dynamic” in this context means that, unlike the standard “keyword-value” list approach, the input parameters may be defined algorithmically and in a rather free format style, taking advantage of the Python language. This may greatly simplify input of complex parameters, such as definition of basis states or initial conditions.

3.3. General PYXAID Workflow and Analysis Scheme: a Simple Three-Step Procedure. Application of the PYXAID program to studying non-adiabatic processes in a given system comprises three major steps:

1. obtaining ground state molecular dynamics trajectory or trajectories
2. computing matrix elements of the NA-MD Hamiltonian in the KS basis, including adiabatic state energies and NA couplings between states, and, optionally, matrix elements of the perturbation Hamiltonian to describe the photo-excitation, i.e., transition dipole moments
3. executing the NA-MD simulations

At each step, a number of additional computations can be performed to analyze the electronic and nuclear dynamics and to gain additional insights into these processes. The schematics depicting the three-step procedure along with the types of analysis available at each step are presented in Figure 3.

In principle, steps 1 and 2 can be combined, because the NA couplings, state energies, and transition dipole moments can be computed as the MD trajectory is generated. We treat steps 1 and 2 independently, because the MD trajectory can be generated in step 1 using different computational methodology for efficient sampling of the ground state molecular dynamics trajectory. In particular, one can use a lower precision setup for standard ab initio Born–Oppenheimer molecular dynamics or Car–Parrinello method for efficient time propagation of joint electronic and nuclear degrees of freedom. Yet another alternative to efficient sampling of MD trajectories is utilization of semiempirical and tight-binding DFT approaches or even classical molecular mechanics force fields. In addition, one may end up discarding an initial part of the MD trajectory, if it is insufficiently equilibrated. Higher precision is needed in step 2 for the NA coupling calculation. These higher precision calculations can be obtained in a parallel way, as described below.

3.4. High-Throughput Methodology and Parallelization. The stage denoted “Hamiltonian” in Figure 3 is most computationally expensive, because it requires accurate electronic structure calculations for each configuration sampled along the MD trajectory. Such calculations are needed for numerical evaluation of the NA couplings (off-diagonal terms of the Hamiltonian matrix), energies of the eigenstates (diagonal terms), as well as the transition dipole moments (optionally). We use the scheme proposed by Hammes-Schiffer and Tully for numerical evaluation of the NA coupling:

$$H_f(t + \frac{dt}{2}) = \frac{\langle \Psi_f(t) | H(t + dt) | \Psi_f(t) \rangle - \langle \Psi(t + dt) | H(t) | \Psi(t) \rangle}{2 \cdot dt}$$  \hspace{1cm} (24)
For consistency, the energies are also evaluated at the midpoint between the adjacent nuclear steps:

\[ H_n(t) = \frac{E_n(t) + E_n(t + dt)}{2} \]  

(25)

Although the above listed quantities must be computed for many points, the computations are independent of each other, except for the adjacent time steps. Therefore, the calculations can be efficiently parallelized in blocks, as schematically shown in Figure 4.

The entire trajectory is split onto several pieces that are recomputed independently. For example, the length of each piece corresponds to five MD steps in Figure 4. The Hamiltonian matrix element computations involve results obtained for two adjacent MD points (Figure 4b). Thus, the number of quantum-mechanical, single-point calculations required in each job is one more than the number of time steps at which the Hamiltonian matrix elements are evaluated. Because the electronic structure (wave functions and eigenvalues) is required for adjacent MD configurations, the workflow within each job is sequential, as shown in Figure 4b. Several jobs may run in parallel at the same time (Figure 4a). Note that such a scheme is a realization of the high-throughput computing (HTC) paradigm, rather than high-performance computing (HPC), meaning that it can run efficiently on heterogeneous clusters and does not require fast communication between the nodes. In turn, each individual job may run on several processors (HPC paradigm), thus providing a way for additional speedup of calculations.

Once the calculations for all independent jobs in step 2 are finished, the results are collected in one place. The main results are the files containing the Hamiltonian matrix elements (Ham_ files) and, optionally, the transition dipole moments (Hprime_ files), computed along the ground state MD trajectories. These files are used as the main input data for the NA-MD calculations in step 3. In addition, they contain important information about the system under investigation and its dynamics. The tools available for analysis of the data stored in these files are described in section B of the Supporting Information.

4. APPLICATION

The present implementation partially follows the organization used in the scripts and subprograms developed in our lab previously. Those have been applied to a wide variety of electron and hole relaxation and transfer processes in different systems. An early development version of the PYXAID software was recently used to study the hole relaxation and transfer dynamics in the GaN/water system. To test our production-stage implementation of the NA-MD in the PYXAID program and to showcase some of the program capabilities, we study a non-radiative relaxation dynamics of the hot electrons in crystalline pentacene, which is known to happen on the ultrafast time scale.

Pentacene is a popular material for organic photovoltaic devices, due to its favorable light absorption properties, good conductivity, and ability to support the singlet fission process, leading to charge carrier multiplication. Photoinduced charge generation, including singlet fission, competes with and is preceded by hot electron relaxation. In order to illustrate the NA-MD methodology and its implementation and to verify the methodology against the experimentally measured time scales, we consider the initial hot electron relaxation process. We leave the complexities of the singlet fission and charge transport for the future works and point the reader to the recent theoretical studies on these topics, utilizing a variety of approaches that can be related to our methodology. Simulations of this type can be carried out with the help of the PYXAID program and will be the subject of our ongoing work. The simulations described in the sections below are organized in accord with the three-step scheme, presented in Figure 3.

4.1. Steps 1 and 2: Adiabatic MD and Electronic Structure of Solid Pentacene. The geometry optimization, electronic structure, and adiabatic (Born–Oppenheimer, BO) MD calculations are performed with the Quantum Espresso program, utilizing a converged plane wave basis and a pseudopotential representation of the core electrons. In particular, the ultrasoft pseudopotential generated within the Perdew–Burke–Ernzerhof (PBE) generalized gradient approximation (GGA) was used. Stipulated by the size of the molecular system and the pseudopotential, the simulation cell contains 204 explicit electrons. The size of the plane wave basis is chosen to satisfy the 40 Ry cutoff on energy and 400 Ry cutoff on charge density. All calculations are performed at a single gamma point.

During the optimization step, we allow the lattice parameters and atomic positions to relax. The resulting unit cell parameters...
are the following: \(a = 16.134 \text{ Å}, \ b = 4.615 \text{ Å}, \ c = 9.904 \text{ Å}, \ \alpha = 62.1^\circ, \beta = 93.6^\circ, \ \gamma = 93.2^\circ\). The optimized structure is used as the starting point for the adiabatic (Born–Oppenheimer) MD calculations. A 10 ps trajectory is computed using the Verlet integration algorithm with a 1 fs time step. The Andersen thermostat is employed to maintain ambient temperature. A representative structure of the molecular system sampled along the MD trajectory and the corresponding partial density of states (pDOS) are shown in Figure 5.

The electronic DOS, projected onto each of two pentacene molecules (Figure 5a), show that HOMO−1 and LUMO are primarily localized on one pentacene fragment (P2), while HOMO and LUMO+1 are mostly localized on the other fragment (P1). However, in both cases, a non-negligible fraction of the orbital delocalizes onto the adjacent pentacene molecule.

The computed HOMO−LUMO gap is about 2.4 eV. We found that, in order to obtain a reasonable band gap for the crystalline pentacene that agrees with the previous theoretical calculations (2.2 eV),\(^{148}\) it was crucial to use the van der Waals density functional (vdW-DF2).\(^{149}\) Other functionals, such as PBE\(^{144,145}\) and even vdw-DF,\(^{150,151}\) misrepresented the intermolecular interaction between nearby pentacene molecules. As a result, the geometry optimization converged to a quasi-2D sheet formed by nearby pentacene molecules melded along their edges, rather than to a geometry composed of individual molecules packed together. As a consequence, some carbon atoms showed erroneous sp\(^3\)-like hybridization, and the computed band gaps were significantly overestimated. Such artifacts were small with the vDW-DF2 functional: the geometry of the pentacene system was nearly planar, and the band gap was overestimated by only 0.1−0.2 eV in comparison to the previous calculations.

### 4.2. Step 3: Non-Adiabatic Dynamics of Hot Electrons in Solid Pentacene

To study the dynamics of non-adiabatic relaxation of hot electrons, we use the Ham_\textmd{ files obtained in step 2 (Figure 3). The choice of the electronic state basis and initial conditions is an important part of the computation. The size of the basis set, the character of included configurations, e.g., singly or doubly excited, and the energy range spanned by the basis states define a model for NA dynamics and convergence of the calculations. Thus, the choice of the basis set is a matter of physical justification and computational convenience. The choice of the initial conditions can be made with one of the two methods. A specific state or a range of states can be chosen on the basis of the excitation energy and transition dipole moments. Excitation of this type represents interaction of the system with continuous wave radiation. Alternatively, one can simulate explicitly photoexcitation by a laser pulse. The transition dipole moments needed to describe interaction with light are stored in the Hprime_\textmd{ files.

Once the basis of electronic states and the initial conditions are defined, the time-dependent population of each state can be computed. Further, one can compute the average electronic energy as a function of time:

\[
E(t) = \langle \sum_i E_i(t) \rangle_{\text{ens}}
\]

where \(\langle \rangle_{\text{ens}}\) denotes ensemble average (over initial configurations and realizations of the surface hopping trajectories), \(E_i(t)\) is the energy of the basis state \(i\), and \(p_i(t)\) is the time-dependent population of the state. The latter quantity can be defined in either SE or SH way, namely,

\[
p_i^{\text{SE}}(t) = |\phi_i(t)|^2
\]

\[
p_i^{\text{SH}}(t) = \frac{N_i(t)}{N}
\]

where \(N_i(t)\) is the number of SH trajectories in electronic state \(i\) at time \(t\) and \(N\) is the total number of trajectories in the ensemble.

As an example, Figure 6 shows the evolution of the average energy of the pentacene system after its initial excitation to a state at \(\sim 1.5\) eV above the energy of the lowest excited state configuration. One can observe a two-exponential character of the relaxation kinetics. The fast phase completes on the 500 fs
time scale and describes the intraband relaxation of the hot electron to the CB edge. The second phase corresponds to the relaxation across the bandgap, happening on a significantly larger time scale. The relaxation time scale for the first phase, describing de-excitation of the high-energy singlet excited states $S_n$ to the lowest excited state $S_1$, is in good agreement with the experimental results.\textsuperscript{138}

Figure 6 is very useful for obtaining the energy relaxation rates and comparing theoretical predictions with existing experimental measurements for a given material. One may be interested to obtain more detailed information about the states and transitions involved in the relaxation process. For instance, one may visualize and analyze the probabilities with which the system evolves into states in a group of the basis configurations $I = \{i_1, i_2, \ldots, i_n\}$, namely,

$$ P_I(t) = \sum_{i \in I} P_i(t) $$

where both $P_i(t)$ and $P_I(t)$ can be defined in either the SE or SH sense, as in eq 27. In the simplest case, each $P_i(t)$ corresponds to one basis state. A detailed analysis of the electron relaxation kinetics for the seven-state model is presented in Figure 7. It shows the time evolution of the populations of each of the seven basis states.

It should be noted that, normally, the SH-based probabilities provide better results, because the FSSH algorithm properly describes the quantum-mechanical branching of the trajectories and the energy exchange between the electronic and nuclear subsystems. The SE-based probabilities correspond to a mean-field description of the dynamics of the excited states. In particular, as follows from Figure 7, the evolution of the SH-based probabilities (panel a) indicates that the initially prepared excited state SE$_6$ eventually relaxes to the SE$_1$ state, in agreement with the experimentally proposed $S_n \rightarrow S_1$ relaxation mechanism.\textsuperscript{138} On the contrary, the SE-based populations indicate that the main reason of the fast relaxation phase is the relaxation from the SE$_6$ state to the next-lowest state SE$_5$. The lowest excited state SE$_1$ is never fully populated. As a consequence of the differences in the SE and SH probabilities, the energy of the basis states weighted by the SH-based populations predicts a notably faster decay, in good agreement with the experimental data (Figure 6). The analysis of this type can be performed with the module average.

The plots, such as those shown in Figure 7, provide important information about the relaxation mechanisms and the involved states. When the number of the states becomes large, plotting and analyzing many transient probabilities becomes a tedious task. Time−energy 2D maps provide an alternative way of representing the data. Here, one plots the probability (z-axis, color, or contour isovalue) of finding the system at a given energy level (y-axis) at a given time (x-axis). Such a representation resembles that often used by experimentalists,\textsuperscript{139} thus providing direct comparison of theory with experiment. An example of such a map is given in Figure 8. The presented information is equivalent to that in Figure 7. Panel a represents the SH-based populations and shows clearly the fast relaxation on the $\sim$0.5 ps time scale is followed by a significantly slower decay, which appears as a practically not-changing population of the CB edge state. The mean-field picture (panel b) shows a strong delocalization of the populations over the higher energy states.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure7.png}
\caption{Time evolution of the populations of the seven basis states used to describe hot electron relaxation in crystalline pentacene. (a) SH-based probabilities; (b) SE-based probabilities. Averaging is taken over 20 initial configurations and 1000 SH trajectories for each configuration.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8.png}
\caption{Relaxation time−energy map. The relaxation kinetics is represented in the form of a 2D map, showing the SH-based (a) or SE-based (b) populations of all basis states as a function of their energy and time.}
\end{figure}
4.3. Decoherence Times and Phonon Influence Spectra. To compute the decoherence (pure-dephasing) times $\tau_{ij}$ for each pair of electronic states $i$ and $j$, we use the linear response formalism, as discussed elsewhere. The details of the formulation are summarized in section C of the Supporting Information. The decoherence function computed for the SE$_1$ and SE$_6$ pair of states is shown in Figure 9a. The Gaussian fit gives the estimate for the dephasing time of ca. 10 fs.

Along with the decoherence functions, we compute the phonon influence spectrum (Figure 9b). It is obtained as the Fourier transform of the energy gap autocorrelation function for a pair of electronic states, as explained in section C of Supporting Information. The influence spectrum shows an intense peak in the region of $\sim 200-250$ cm$^{-1}$, which corresponds to vibrations of individual molecules in the lattice sites, generating acoustic modes in bulk pentacene. This type of vibration is, therefore, responsible for the energy dissipation during the SE$_6$ $\rightarrow$ SE$_1$ transition. No high-frequency modes are observed in the influence spectrum, implying that the C-C and C-H vibrations are irrelevant to the energy dissipation mechanism.

5. CONCLUSIONS

In this paper, we have presented the PYXAID program, which focuses on non-adiabatic molecular dynamics simulations in large-scale systems. We have described the basic theoretical foundations of the methods used in the program, as well as the computational strategies and the code organization. A particular focus has been given to a detailed explanation and justification of the theoretical background of the non-adiabatic molecular dynamics method in general and associated approximations. The capabilities, advantages, and limitations of the PYXAID package, as well as the conditions under which the methodologies and approximations are valid, are also thoroughly discussed.

In particular, the basic capabilities of the program include FSSH within the classical path approximation, calculations of phonon spectral densities (influence spectra) and dephasing times characterizing electron–phonon interactions, and a variety of pre- and postprocessing tools. The advanced capabilities of the program include the use of efficient time-propagation techniques and decoherence effects, incorporation of multielectronic Slater determinant adiabatic states, and explicit treatment of field–matter interactions. These topics will be discussed in our later works. Description of electronic transitions that involve change of spin symmetry (e.g., singlet to triplet) is of great interest to a wide community of chemists but also is a challenging task. While this capability is not yet added in our present implementation, the theoretical framework will be similar to the one already present. The main difference will arise due to the use of a spinor basis to express the electronic states. This goal will be pursued in our future studies.

The program extensively relies on the classical path approximation, which leads to great computational savings, allowing one to study photoinduced dynamics at the $ab$ initio level in systems composed of hundreds of atoms and involving hundreds or thousands of electronic states. The classical path approximation and its implementation within FSSH are valid if the differences in the nuclear dynamics associated with different electronic states are insignificant, for instance, relative to thermal nuclear motions. In particular, the approximation is expected to work very well for the systems, which exhibit no substantial reorganization of the nuclear subsystem, such as fragmentation, isomerization, bond breaking, and formation, on the time scale of electronic dynamics. Although these conditions may seem restrictive, they are satisfied in many nanoscale, solid-state, gas-phase, and soft-matter systems.

The developed techniques have been applied to characterize the excited state dynamics in crystalline pentacene. In particular, we have found that the relaxation of hot electrons occurs on a 500 fs time scale, in excellent agreement with the experimental results. Application of a FSSH-type scheme is crucial for the correct description of the relaxation mechanisms: The unitary dynamics given by solution of the TD-SE fails to describe the relaxation. The electron–phonon dephasing and relaxation processes are driven by lattice vibrations of the molecular species in the crystalline environment, with the characteristic frequencies in the 200–250 cm$^{-1}$ range.

We hope that the developed computational tool will further stimulate and facilitate research in the field of photochemistry and related disciplines.

ASSOCIATED CONTENT

Supporting Information

Detailed description of the FSSH method, the classical path approximation for FSSH, the analysis modules of the PYXAID package, and the linear response formulation for the dephasing time and phonon influence spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

ACKNOWLEDGMENTS

The authors are grateful to Dhara Trivedi for comments on the manuscript. A.V.A. was funded by the Computational Materials and Chemical Sciences Network (CMCSN) project at Brookhaven National Laboratory under contract DE-AC02-98CH10886 with the U.S. Department of Energy and supported by its Division of Chemical Sciences, Geosciences & Biosciences, Office of Basic Energy Sciences. O.V.P. acknowledges financial support of the U.S. Department of Energy, grant DE-SC0006527.

REFERENCES