

Analysis of bath motion in MM-SQC dynamics via dimensionality reduction approach: Principal component analysis

Cite as: J. Chem. Phys. 154, 094122 (2021); doi: 10.1063/5.0039743

Submitted: 6 December 2020 • Accepted: 11 February 2021 •

Published Online: 5 March 2021



Jiawei Peng,¹ Yu Xie,¹ Deping Hu,¹ and Zhenggang Lan^{1,2,a)}

AFFILIATIONS

¹SCNU Environmental Research Institute, Guangdong Provincial Key Laboratory of Chemical Pollution and Environmental Safety & MOE Key Laboratory of Theoretical Chemistry of Environment, South China Normal University, Guangzhou 510006, China

²School of Environment, South China Normal University, University Town, Guangzhou 510006, China

^{a)}Author to whom correspondence should be addressed: zhenggang.lan@m.scnu.edu.cn and zhenggang.lan@gmail.com

ABSTRACT

The system-plus-bath model is an important tool to understand the nonadiabatic dynamics of large molecular systems. Understanding the collective motion of a large number of bath modes is essential for revealing their key roles in the overall dynamics. Here, we applied principal component analysis (PCA) to investigate the bath motion in the basis of a large dataset generated from the symmetrical quasi-classical dynamics method based on the Meyer–Miller mapping Hamiltonian nonadiabatic dynamics for the excited-state energy transfer in the Frenkel-exciton model. The PCA method clearly elucidated that two types of bath modes, which either display strong vibronic coupling or have frequencies close to that of the electronic transition, are important to the nonadiabatic dynamics. These observations were fully consistent with the physical insights. The conclusions were based on the PCA of the trajectory data and did not involve significant pre-defined physical knowledge. The results show that the PCA approach, which is one of the simplest unsupervised machine learning dimensionality reduction methods, is a powerful one for analyzing complicated nonadiabatic dynamics in the condensed phase with many degrees of freedom.

Published under license by AIP Publishing. <https://doi.org/10.1063/5.0039743>

I. INTRODUCTION

The theoretical investigation of nonadiabatic dynamics in the condensed phase is a challenging research topic. In addition to the existence of strong coupling between the nuclear and electronic motions,^{1–3} it is also difficult to treat a huge number of nuclear degrees of freedom (DOFs). In many situations, chemical dynamics may occur in a very small region, referred to as the reaction center, while a large number of surrounding molecules forms the reaction environment. This naturally leads to the widely used system-plus-bath model² in which the system consists of a small number of DOFs in the reaction center and all the other environmental DOFs are treated collectively as a thermal bath. Generally, when the system-plus-bath model is employed for the treatment of nonadiabatic dynamics in condensed phases, the dynamics of the reduced system is described in detail. In contrast, much less attention has been paid to address the bath motion explicitly.

In several theoretical approaches based on the reduced density operator, such as the Redfield⁴ theory and the hierarchical equations of motion (HEOM),^{3–5} a large number of bath DOFs are often traced out. Within the HEOM framework, some theoretical works have proposed to the so-called “dissipation” -based density matrix approach when discussing the bath motion.⁶ In the stochastic framework, the influence of bath DOFs on the system dynamics is described by the stochastic noise,^{7–11} and all intrinsic details of bath motions are omitted. Alternatively, various theoretical approaches, such as the multiconfigurational time-dependent Hartree (MCTDH),¹² multilayer multiconfigurational time-dependent Hartree (ML-MCTDH),^{12–15} time-dependent tensor-network,^{16–22} Gaussian-wavepacket-based methods,^{23–26} and mixed-quantum-classical/semiclassical approaches,^{27–44} employ a set of discrete vibrational modes to represent a bath composed of a huge number of DOFs. Within these approaches, the intrinsic bath motion can be examined explicitly.

In trajectory-based mixed-quantum-classical or semiclassical simulations, the dynamics evolution takes place in a very high dimensional space that is spanned by the electronic variables and a large number of bath coordinates. In addition, trajectory-based simulations usually generate a large amount of data when monitoring the evolution of a large number of the bath DOFs. Hence, analysis of the bath motion dynamics becomes extremely difficult because of the large amount of high-dimensional data.

With the rapid development of machine learning algorithms and progress in computational capacity, it is possible to analyze large amounts of high-dimensional data via dimensionality reduction approaches.^{45–50} As a group of unsupervised machine learning algorithms, dimensionality reduction methods try to transform from the original high-dimensional space spanned by the data to a reduced low-dimensional one spanned by a few effective coordinates, which preserves the primary pattern of the data distribution. As a consequence, the complex information of the data pattern distribution in the original high-dimensional space can be well analyzed within the reduced low-dimensional space. Over the decades, several groups have made substantial efforts to analyze dynamics evolution of different chemical systems via dimensionality reduction approaches, such as principal component analysis (PCA),^{51–56} isometric feature mapping (ISOMAP)/multidimensional scaling (MDS),^{55–59} diffusion map,^{55,56,60,61} auto-encoders,^{62,63} and other approaches.^{64–67}

PCA^{48,51,68–70} is a widely used, highly efficient dimensionality reduction approach, which constructs a linear mapping transformation from the original high-dimensional space to a reduced low-dimensional one. This linear mapping defines the transformation from the original coordinates to reduced ones. Because of its simple implementation, clear mathematical insights, and the explicit construction of reduced coordinates, PCA is widely used for analysis of the trajectory information in chemical dynamics.^{51–56}

In this work, we employed PCA to analyze collective bath motion in theoretical simulations of the condensed-phase nonadiabatic dynamics. Here, we were particularly interested in a group of important condensed-phase reactions, i.e., the photoinduced excited-state energy-transfer dynamics that plays an essential role in solar energy conversion. The exciton dynamics can often be described by the Frenkel site-exciton model,^{2,3,71} which is composed of a system that includes several coupled electronic states and a thermal bath that includes a large number vibrational modes. This model provides us with a typical example to analyze the role of bath modes in the exciton dynamics. We wanted to consider a practical trajectory-based dynamics approach that may give a proper description of photoinduced exciton dynamics; at minimum, this theoretical approach should give the correct dynamics when the parameters of the site-exciton Hamiltonian model fall into some particular domains. For instance, when the bath motion is slow enough and the classical treatment of the bath motion is allowed, we hope that the selected trajectory-based dynamics approaches can describe the nonadiabatic dynamics well. Finally, we selected the symmetrical quasi-classical dynamics method based on the Meyer–Miller mapping Hamiltonian (MM-SQC).^{71–85} Within this framework, the MM mapping Hamiltonian is constructed via the mapping from discrete electronic states to coupled harmonic oscillators. The quasi-classical dynamics simulations are carried out based on the MM

mapping Hamiltonian in which the window function is employed for initial sampling and final assignment of the quantum states. Previous benchmark calculations^{71,73,75,77,86,87} showed that the MM-SQC method provides correct treatment of the nonadiabatic dynamics of condensed phases, including the photoinduced exciton dynamics. In this work, we used pure classical propagation to treat all the bath modes in the MM-SQC dynamics for simplicity. Thus, we only considered two types of system-plus-bath models in which the time scale of the bath motion was slower or comparable to that of the electronic motions (adiabatic or intermediate bath).^{2,38} It should be pointed out that the MM-SQC simulation results are meaningful under these conditions.⁷¹ For PCA of the bath motion in the MM-SQC dynamics, we selected two types of descriptors, i.e., the action variables and the normal coordinates of the bath modes. The results clearly showed that PCA can provide physical insights directly from the trajectory data generated from the MM-SQC dynamics simulation. The current work demonstrates that dimensionality reduction is a very powerful tool to analyze multidimensional bath motion in the nonadiabatic dynamics of the condensed phase without the need for significant existing knowledge.

This paper is organized in the following structure: Sec. II focuses on theory and computational details, including the multi-mode site-exciton Hamiltonian, the MM-SQC method, and PCA; Sec. III discusses the PCA results with different descriptors; and Sec. IV provides a summary of this work.

II. THEORY AND COMPUTATIONAL DETAILS

A. Site-exciton Hamiltonian

The site-exciton^{2,3,71,73,89} model is considered to be composed of two electronic states, and the total Hamiltonian H is given by

$$H = H_s + H_b + H_{sb}, \quad (1)$$

where the electronic part H_s , the vibrational part H_b , and the electronic-phonon interaction H_{sb} are further expanded in the following forms:

$$\begin{aligned} H_s &= \sum_{k=1}^2 |\phi_k\rangle V_{kk} \langle \phi_k| + \sum_{k \neq l} |\phi_k\rangle V_{kl} \langle \phi_l|, \\ H_b &= \sum_{k=1}^2 \sum_j^{N_b} \frac{1}{2} \omega_{kj} [Q_{kj}^2 + P_{kj}^2], \\ H_{sb} &= \sum_{k=1}^2 |\phi_k\rangle \left(\sum_j^{N_b} \kappa_{kj} Q_{kj} \right) \langle \phi_k|, \end{aligned} \quad (2)$$

where V_{kk} ($k = 1, 2$) represents the energy of the k th local excited state ϕ_k (LE1 or LE2) and V_{kl} ($k \neq l$) defines the coupling between the different states; N_b is the total number of bath modes; for the j th vibrational bath mode in the k th quantum state index that couples with the k th LE state ϕ_k , ω_{kj} is the corresponding frequency, and the quantities Q_{kj} and P_{kj} denote the corresponding coordinate and momentum, respectively; and κ_{kj} denotes the electron–phonon coupling term of each mode; thus, κ/ω can be used to characterize the corresponding electron–phonon coupling strength.

In this work, the Debye-type spectral density² was used,

$$J(\omega) = \frac{2\lambda\omega\omega_c}{\omega^2 + \omega_c^2}, \quad (3)$$

where ω_c is the characteristic frequency of the bath mode and λ is the reorganization energy.

A set of discrete bath modes were generated based on the above spectral density. In principle, the spectral density can be written as the sum of the contribution of each mode² and is given by

$$J_k(\omega) = \frac{\pi}{2} \sum_{i=1}^N \kappa_{ki}^2 \delta(\omega - \omega_{ki}), \quad (4)$$

where κ_{ki} is the coupling coefficient representing the coupling strength of each mode. Once the sampling interval $\Delta\omega$ is given, the coupling coefficient κ_{ki} can be evaluated by the following equation:^{71,90}

$$\kappa_{ki} = \sqrt{\frac{2}{\pi} J_k(\omega_{ki}) \Delta\omega}. \quad (5)$$

The time scales of the bath motion and electronic motion are determined by the characteristic bath frequency and the Rabi-type oscillation frequency, respectively. When the characteristic frequency of the bath is much smaller than the electronic transition frequency

(Rabi frequency), the overall bath motion is much slower than the electronic motion and the bath is the adiabatic one. When the above two frequencies are comparable and two motions show the similar time scales, the intermediate bath is obtained. When the bath motion is much faster than the electronic motion, the bath becomes the nonadiabatic one.

Previous work⁷¹ has demonstrated that the MM-SQC method gives a roughly correct description of the nonadiabatic dynamics when the adiabatic bath or intermediate bath is considered, particularly for the short-time-scale dynamics. However, the MM-SQC treatment of the model with the nonadiabatic bath requires great caution. It is not recommended to employ the purely classical dynamics to treat the motion of high-frequency modes because of their quantum nature. Thus, our current study considered situations that can be well described by the MM-SQC dynamics. Two models with different energy gaps $\Delta_E = V_{11} - V_{22}$ and interstate couplings V_{12} were selected in this work, which corresponded with the typical situations of the intermediate and adiabatic baths. All parameters in the models are given in the below discussion. In this sense, we believe that the dynamics of these models can be treated well by the MM-SQC dynamics method according to the previous benchmark calculations on the similar models.

B. MM-SQC dynamics

In the MM-SQC dynamics,^{80,91} the MM Hamiltonian is constructed by mapping a set of discrete quantum states to a set of coupled harmonic oscillators. If all quantum operators are replaced by their classical correspondences, the classical MM Hamiltonian is given as

$$H_{MM}(x, p, Q, P) = \sum_k \left[\frac{1}{2} (x_k^2 + p_k^2) - \gamma \right] H_{kk}(Q, P) + \frac{1}{2} \sum_{k \neq l} (x_k x_l + p_k p_l) H_{kl}(Q, P), \quad (6)$$

where (x, p) and (Q, P) are the coordinates and momentum for the electronic mapping variables and nuclear DOFs, respectively; H_{kk} and H_{kl} denote the diagonal and off-diagonal elements of the Hamiltonian matrix, respectively; and the quantity γ represents the zero-point energy (ZPE) correction term.^{34,86,92–94}

In the MM-SQC method, both the initial-state sampling and final-state assignment employ the “bin” (or “window function”) approach.⁸⁰ In this work, the triangle window function was used,⁷⁴ which defines the quantity $\gamma = 1/3$ automatically. For a two-state model, the joint window functions W_1 and W_2 are defined as

$$W_1(n_1, n_2) = 2 \cdot h(n_1 + \gamma - 1)h(n_2 + \gamma)h(2 - 2\gamma - n_1 - n_2), \quad (7)$$

$$W_2(n_1, n_2) = 2 \cdot h(n_1 + \gamma)h(n_2 + \gamma - 1)h(2 - 2\gamma - n_1 - n_2),$$

where $h(z)$ is the Heaviside function. The action variable n_k ($k = 1, 2$) of the electronic mapping variables is evaluated by

$$n_k = \frac{1}{2} [(x_k)^2 + (p_k)^2] - \gamma. \quad (8)$$

In the MM-SQC dynamics simulations, the initial state was determined by placing the lowest vibrational level of the ground electronic state into the LE1 state. For the electronic variables, a triangle window function was employed in the initial sampling and final assignment. The action-angle sampling method was used for the nuclear DOFs. In the MM-SQC dynamics, the nuclear and electronic integration time step sizes were 0.01 fs and 0.001 fs, respectively, and total 10 000 trajectories were used to obtain the average results in the simulations.

C. PCA method

As a powerful unsupervised learning approach, PCA^{48,51,68–70} constructs a linear mapping transformation from the original high-dimensional space to a reduced low-dimensional one, which can preserve the primary pattern features of the data distribution. As a consequence, the complex information of the data pattern distribution in the high-dimensional space can be well analyzed within a reduced low-dimensional space.

In PCA, the covariance matrix or the correlation matrix M_{co} is constructed as

$$M_{co} = (\mathbf{X} - \langle \mathbf{X} \rangle)^T (\mathbf{X} - \langle \mathbf{X} \rangle), \quad (9)$$

where \mathbf{X} is the descriptor matrix and superscript T denotes the matrix transpose operation.

Diagonalization of the covariance matrix gives the eigenvalues and eigenvectors, namely,

$$M_{co} = \mathbf{U}^T \mathbf{E} \mathbf{U}, \quad (10)$$

where the variance ratio E_i is the ratio of the i th eigenvalue (the diagonal element of \mathbf{E}) and the sum of all eigenvalues, which indicates the proportion of the independent variance that lies along each principal component; the corresponding eigenvector \mathbf{U}_i , which is the column vector of \mathbf{U} , defines the transformation from the original coordinates in the high-dimensional space to the reduced coordinates in the low dimensional space.

As a linear dimensionality reduction method, PCA cannot describe nonlinear data pattern distributions, which can be

found in many on-the-fly nonadiabatic dynamics.^{57,60} For example, when the nonadiabatic dynamics show either the stretching/isomerization/pyramidalization mixed motion or aromatic-ring deformation, the reactive coordinates become highly nonlinear in Cartesian coordinate space and the trajectory data in the nonadiabatic dynamics show a highly nonlinear distribution. To deal with this issue, the employment of various nonlinear dimensionality reduction approaches was proposed, such as ISOMAP^{45,46,55–58,95} and diffusion map.^{49,50,55,56,60,61,96} Compared with these more sophisticated nonlinear dimensionality reduction methods, PCA is more convenient for data analysis, particularly because of its simple mathematical insights and ability to treat large amounts of data.^{51–56,95} In addition, PCA displays a good ability to treat the problem being studied in the current work, possibly because the site-exciton Hamiltonian defines the bilinear coupling terms between the electronic and vibrational parts. Importantly, PCA directly provides the reduced coordinates from the transformation between the original space and the reduced space. This gives the possibility to build the low-dimensional model with reduced coordinates. We can justify our analyses based on further dynamics simulations using the reduced models (see the below discussions). Therefore, PCA was selected in this work to perform dimensionality reduction.

The PCA procedure was carried out using the following steps:

1. *Geometry data collection.* The original MM-SQC dynamics simulation employed 10 000 trajectories. The ultrafast dynamics within 200 fs almost converged on randomly selecting 2500 trajectories (Appendix A). Therefore, we used 2500 trajectories in the subsequent analysis and truncated their propagation time to 200 fs. For each trajectory, the geometries were selected every 1 fs, and nearly 500 000 geometries were collected. Although the sampling number is large, PCA can treat them without the need for a large amount of computational time.
2. *Descriptor construction.* One natural choice is the normal coordinate Q_j of the bath mode, which can provide direct information on the evolution of polyatomic systems or model systems.^{52,61,97} Another suitable choice is the action variable N_j of the bath mode, which is given by $N_j = \frac{1}{2}Q_j^2 + \frac{1}{2}P_j^2$. In the normal-mode description, N_j corresponds with the vibrational level of the harmonic oscillator, and it is relevant for the vibrational energy of the system. Thus, N_j is a suitable parameter to measure mode excitation/de-excitation for harmonic oscillators. Direct employment of the energy of each mode as a descriptor is not a suitable choice because their values are quite different even when remaining on the same vibrational level. In MM-SQC dynamics, we used the window function to bin the electronic action variable and give the discrete quantum occupation number. In principle, the same quantities can be obtained for the bath modes as well. However, it is not suitable to employ the PCA method to analyze such discrete quantum occupation numbers because PCA requires a continuous data distribution pattern. Thus, N_j and Q_j were employed to represent the difference (or distance) between sampled geometries.
3. *Dimensionality reduction.* After \mathbf{X} was constructed, we calculated the leading eigenvalues and the corresponding eigenvectors of the covariance matrix via singular value decomposition

(SVD).⁹⁸ The dominant eigenvalues or the variance ratios give us the principle component, and the corresponding eigenvectors define the transformation of the data distribution from a high-dimensional space to the low-dimensional one.

III. RESULTS AND DISCUSSION

A. Intermediate bath

We first considered the dynamics with the intermediate bath region in which the characteristic frequency of the bath is close to the Rabi frequency. A model with $\Delta_E = 0.03$ eV and $V_{12} = 0.03$ eV was constructed, and the Rabi frequency was ~ 0.067 eV. The characteristic bath frequency ω_c was chosen as 500 cm^{-1} (~ 0.06 eV) and the reorganization energy λ was chosen as 62.5 cm^{-1} (~ 0.0077 eV). In the current model, the ratio between the Rabi frequency and the characteristic frequency of the bath was close to 1. The sampling interval $\Delta\omega$ was chosen as 12 cm^{-1} (~ 0.0015 eV), and 200 bath modes were employed to represent the continuous bath coupled with a single electronic state. Thus, a total of 400 modes were used in the final simulation of the nonadiabatic dynamics, i.e., modes 1–200 coupled to the LE1 state and modes 201–400 coupled to the LE2 state.

1. PCA results with the N_j descriptor

When action variable N_j is chosen as the descriptor in the PCA, the variance ratio E_i along each principle component and the transformation matrix elements are given in the left panel of Figs. 1(a)–1(c). The value of E_i decays dramatically with increasing dimensions and goes to nearly zero, as shown in Fig. 1(a). The value of E_i decays quickly in the first ~ 110 dimensions. After ~ 200 dimensions, the value of E_i remains basically close to 0, which suggests that the later modes do not play any role in the original dynamics data. Overall, the first ~ 150 dimensions, or at least the first ~ 50 dimensions, are important for recovering the data distribution in the PCA. Based on the decay feature, the current nonadiabatic dynamics were not recovered via the reduced models with only a few DOFs or simple collective modes re-constructed by PCA. This clearly reflects the complexity of the nonadiabatic dynamics of complex systems in which strong electron–phonon coupling exists among many DOFs.

In the current work, two sets of bath modes are considered that couple to either the LE1 or LE2 state individually with the same electronic–phonon coupling strengths. Thus, a similar distribution pattern for $|U_{ij}|^2$ in Fig. 1(b) can be observed for both sets of vibrational modes. It is possible that because the initial preparation puts the lowest vibrational level of the ground state into the LE1 state, the two parts of the $|U_{ij}|^2$ distribution (the upper and lower part) are not exactly the same. The $|U_{ij}|^2$ distribution exhibits high localization, indicating that each reduced dimension can be decomposed into only a few of original modes. As shown in Fig. 1(c), the first several dimensions are clearly dominated by the low-frequency modes and modes with frequencies close to 0.0669 eV.

The involvement of low-frequency modes in Figs. 1(b) and 1(c) is because of their larger electronic–phonon coupling. Here, we should point out that the vibronic coupling strength is determined by κ/ω . Because the Debye-type spectral density is used, κ/ω decreases with increasing frequency. Hence, the low-frequency

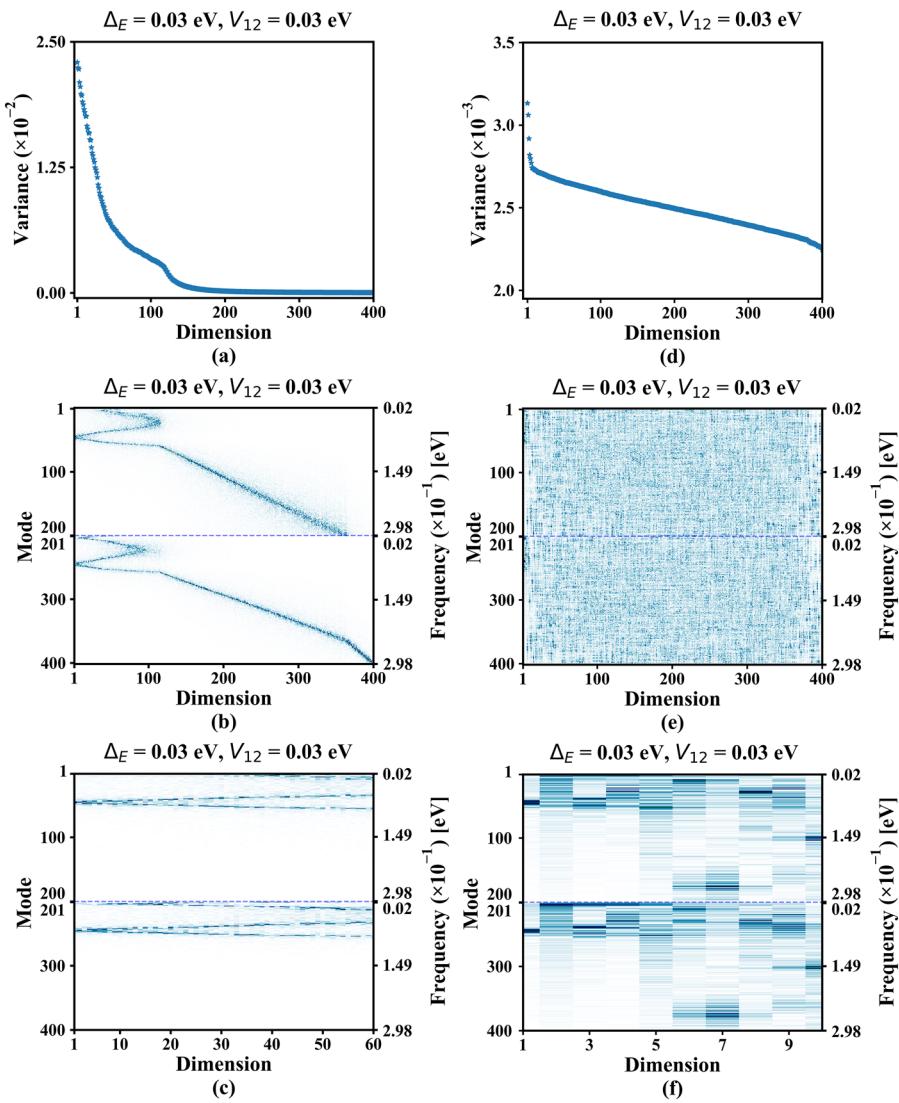


FIG. 1. PCA results of the original 400-mode model (intermediate bath region) with $\Delta_E = 0.03$ eV and $V_{12} = 0.03$ eV using different descriptors, i.e., the action variable N_j (left panel) or the bath coordinate Q_j (right panel). (a) and (d) The variance ratio E_i along each principal component; (b) and (e) the heatmap of $|U_{ij}|^2$: the abscissa is the principal component and the ordinate represents the bath mode, the darker color indicates a larger $|U_{ij}|^2$; and (c) and (f) the heatmap of the first few dimensions $|U_{ij}|^2$.

modes display significant electron–phonon coupling, which leads to their important roles in the first leading dimensions.

The importance of modes with vibrational frequencies close to 0.0669 eV can be attributed to resonance effects in the electronic transitions, which is because their frequencies are very close to the pure electronic transition with the Rabi frequency (0.067 eV). As expected, the current model (falling in the intermediate bath region) leads to a non-Markov process in which the nuclear vibrational DOFs display similar time scales as the electron dynamics during exciton transfer.^{2,3,75,88,99,100}

In the current PCA results, all high-frequency modes were only relevant to the later dimensions that are not important to the dynamics. This is explained by the fact that the time scale of their vibrational motions is far from that of electronic motion.

PCA of the MM-SQC dynamics carried out in this work could provide proper results that were fully consistent with the

physical insights when using the action variable N_j as the descriptor in the intermediate bath region. More importantly, PCA correctly identified these features purely based on automatic analysis of the MM-SQC simulation results without significant involvement of pre-defined physical knowledge. This shows that PCA is a powerful unsupervised machine learning tool for analysis of the bath motion in site-exciton dynamics.

To further analyze the above PCA results, we investigated the contribution of the bath mode again. Additional dynamics calculations were carried out based on several reduced models composed of bath modes that played important roles in the leading components of the PCA results. To build such reduced models, a two-step protocol was taken in which we first determined how many dimensions were chosen and then decided which modes were taken into account.

First, we need to determine which leading principal components (or dimensions) should be considered in the reduced model.

This is based on the fact that the sum of E_i reaches a threshold value N^D , namely,

$$\sum_{i=1}^{I+1} E_i > N^D \geq \sum_{i=1}^I E_i, \quad (11)$$

where N denotes that the descriptor is the action variable N_j and the superscript D denotes that this parameter is used to obtain reasonable dimensions. The summation here is performed according to the descending order of E_i until the cutoff value N^D is reached, leading to the important dimensions (from the first dimension to the I th dimension) on which further analysis is carried out.

Second, we need to decide which modes should be included for each chosen dimension. For the i th chosen dimension ($0 < i \leq I$), a cutoff value N_i^M is defined by

$$\sum_j^{J+1} |U_{ij}|^2 > N_i^M \geq \sum_j^J |U_{ij}|^2, \quad (12)$$

where the subscript i indicates that the cutoff value is for the i th dimension and the superscript M denotes that this parameter is used to obtain the important modes. The summation here is also performed in the descending order of $|U_{ij}|$ until the summation is greater than N_i^M . This provides us with the relatively important J modes in each dimension. In practice, we used the same N_i^M for all chosen dimensions and we later used the label N^M to define the cutoff values in the below discussions. A more detailed discussion for the case of $N^M = 0.1$ and $N^M = 0.6$ can be found in Appendix B.

The dynamics data within the first 200 fs were selected to perform PCA. To confirm that the obtained PCA results are meaningful, the time-dependent electronic populations of the full-dimension model and reduced models within the first 400 fs are compared in the following discussions. The long-time-scale propagation provides additional validation for the PCA results.

The first 43 dimensions were selected for the case of $N^D = 0.6$, which covered the dominant variance ratios in Fig. 1(a). Here, the choice of $N^M = 0.6$ defines a reduced model with 58 modes and their electronic-phonon coupling strengths (κ/ω) are characterized in Fig. 2(a). As shown in Figs. 1(b) and 2(a), the reduced model includes low-frequency modes with strong vibronic couplings (large κ/ω) and modes with frequencies close to that of the electronic transitions ($\omega \sim \omega_{\text{Rabi}}$). As a consequence, it basically captured the dynamics features of the original 400-mode model, as shown in Fig. 3(a). However, the early stage electronic coherence and late-stage asymptotical feature are different, and the reasons for this can be explained. The reduced model has fewer vibrational modes, resulting in weaker damping of the coherent electronic motion in the early dynamics. This was further confirmed by the significantly improved description of the early stage coherence dynamics when more bath modes are added in the reduced models by increasing N^D or N^M , as shown in Figs. 2(b)-2(d) and 3(b)-3(d).

In the case of $N^D = 0.6$ and $N^M = 0.8$, the reduced model includes 108 modes in total. Among them, the first group (39 modes) belongs to the low-frequency modes and the second group (69 modes) displays vibrational frequencies close to the Rabi frequency, as shown in Fig. 2(b). To clarify their individual roles, we performed additional calculations based on two reduced models that only included one group. When only the second group of modes (69 modes) was taken, the dynamics essentially captured the most important features of the nonadiabatic dynamics of the reduced

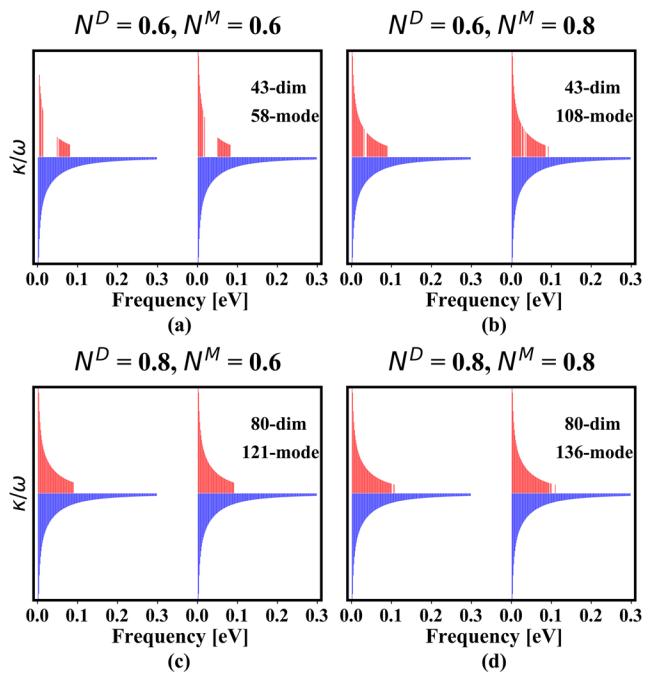


FIG. 2. The κ/ω value of the bath mode in the original 400-mode model ($\Delta_E = 0.03$ eV and $V_{12} = 0.03$ eV and the intermediate bath region) and the reduced models with different cutoff values of N^D [Eq. (11)] or N^M [Eq. (12)]. In each inset, the upper part (red lines) corresponds with κ/ω in the reduced model, while the lower part (blue lines) denotes κ/ω in the original 400-mode model.

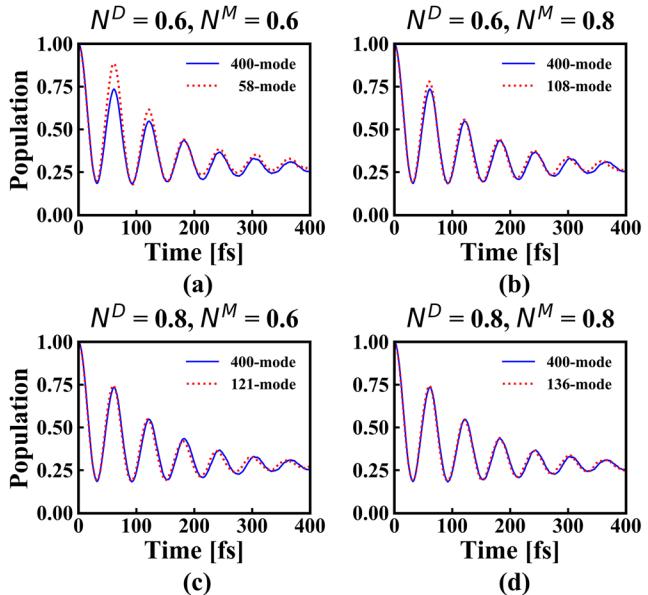


FIG. 3. Time-dependent electronic population of the LE1 state for the dynamics of the models ($\Delta_E = 0.03$ eV, $V_{12} = 0.03$ eV, and the intermediate bath) with different cutoff values of N^D [Eq. (11)] or N^M [Eq. (12)]. Blue solid lines denote the results of the original model with the full 400 bath modes. Red dotted lines denote the results of the reduced models, and the number of bath modes included in the models is given in each inset.

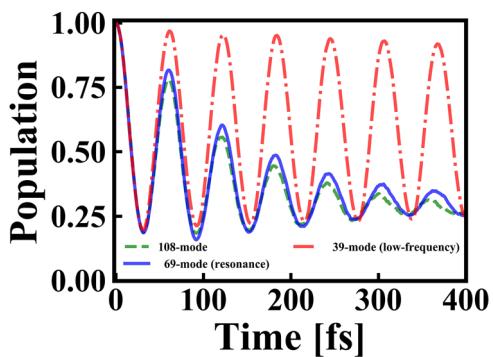


FIG. 4. Time-dependent electronic population of the LE1 state in the dynamics of the reduced 108-mode model based on the case with $N^D = 0.6$ and $N^M = 0.8$. The green dashed line denotes the result of the 108-mode reduced model. The blue solid line denotes the result of the reduced model including 69 modes belonging to the resonance region. The red dashed-dotted line denotes the result of the reduced model including 39 low-frequency modes.

model with 108 modes, as shown in Fig. 4. In particular, two models (69 modes and 108 modes) predict similar features in the early stage dynamics governed by the electronic coherent motion. This confirms the important role of the modes with frequencies close to the Rabi frequency. The low-frequency modes with a strong vibronic coupling do not play a significant role in the short-time dynamics. However, these modes should be included for proper description of the long-time-scale dynamics.

In the case of $N^D = 0.8$ and $N^M = 0.8$, the first 80 dimensions were selected and a total of 136 modes were included in the reduced model. As shown in Fig. 3(d), there is almost no difference in the early stage dynamics between the reduced model and the original model; however, we still observe some very minor differences in the long-time-scale dynamics, which indicates the existence of rather small contributions from a large number of other residual modes. In such a case, most modes involved in the reduced model still fall in the region between the low-frequency modes with strong electron-phonon coupling and modes with a frequency close to the electronic transition. All the high-frequency modes are not presented in these reduced models at all because the time scales of their motion are significantly faster than that of the electronic transitions.

2. PCA results with the Q_j descriptor

The PCA results using the bath coordinate Q_j as the descriptor are summarized in the right panel of Figs. 1(d)–1(f). Although the dominant variance ratio E_i along each principal component decays monotonically with increasing dimension, the overall decay feature shown in Fig. 1(d) is different to that in Fig. 1(a). The value of E_i drops dramatically in the first seven dimensions, while it decays very slowly afterward and its value does not become zero even at the last dimension. In Fig. 1(e), the distribution of $|U_{ij}|^2$ is highly delocalized, indicating that many bath modes are highly correlated. When we only focus on the first seven dimensions, the $|U_{ij}|^2$ distribution exhibits specific localization features and the relevant modes are mainly located either in the low-frequency domain or in the resonance domain, as shown in Fig. 1(f).

The effect of the bath mode was analyzed by running the dynamics in several reduced models. First, the number of reduced

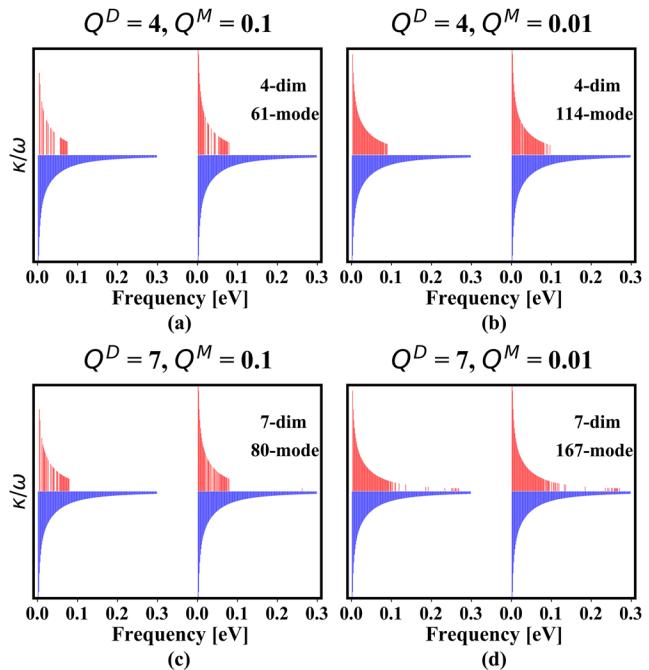


FIG. 5. The κ/ω value of the bath mode in the original 400-mode model ($\Delta_E = 0.03$ eV and $V_{12} = 0.03$ eV, the intermediate bath region) and the reduced models with different cutoff values of Q^D and Q^M [Eq. (13)]. In each inset, the upper part (red lines) is κ/ω in the reduced model, while the lower part (blue lines) denotes κ/ω in the original 400-mode model.

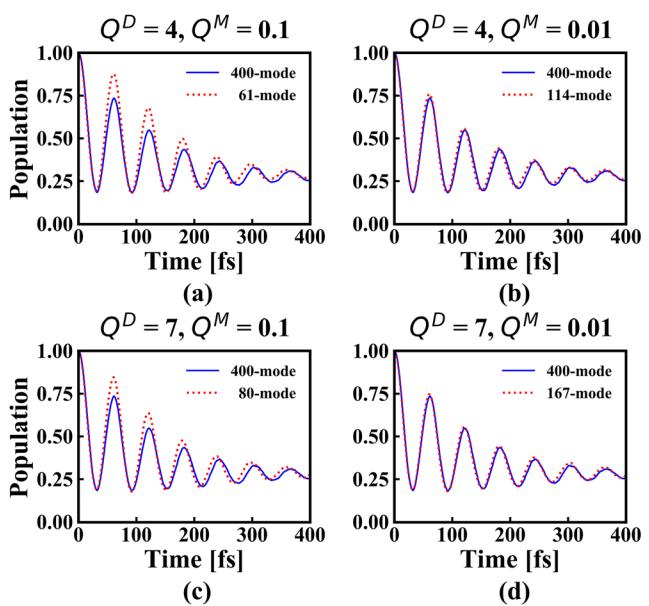
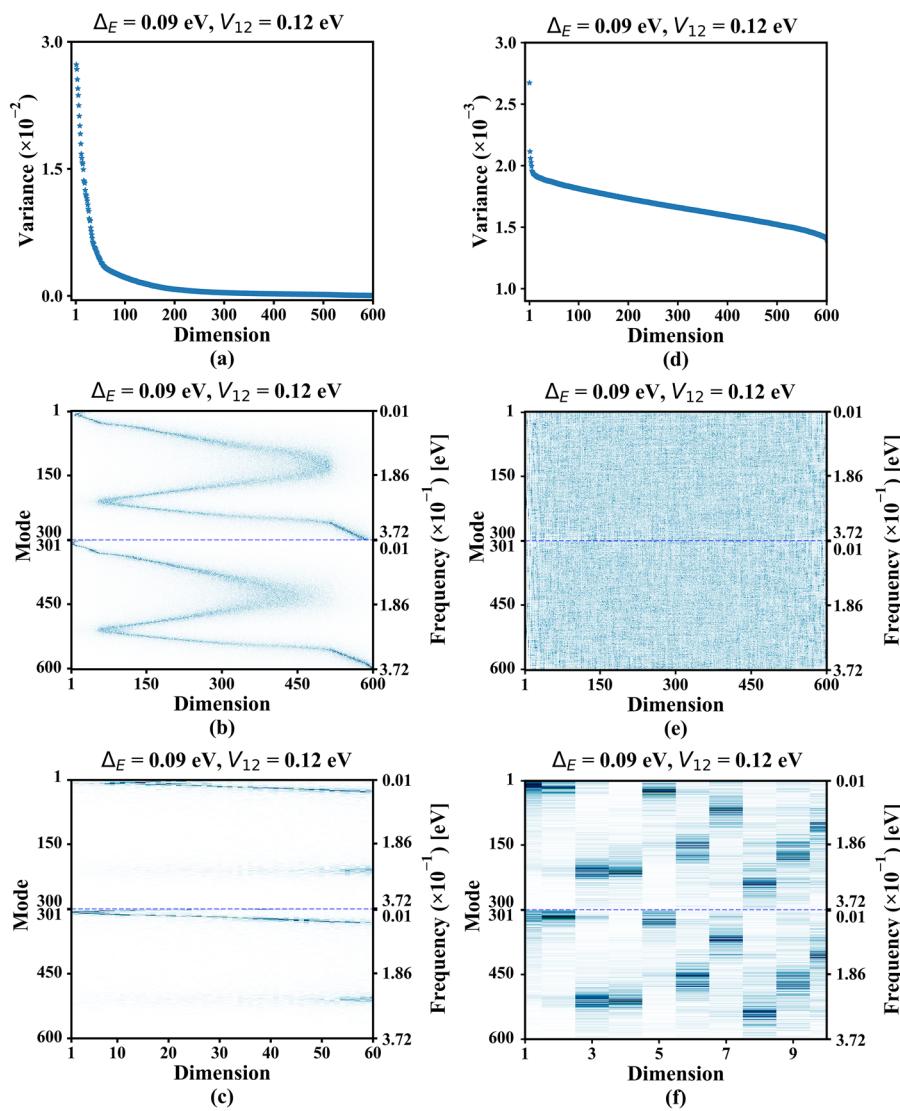


FIG. 6. Time-dependent electronic population of the LE1 state in the dynamics of the models ($\Delta_E = 0.03$ eV, $V_{12} = 0.03$ eV, and the intermediate bath) with different cutoff values of Q^D and Q^M [Eq. (13)]. Blue solid lines denote the results including the full 400 bath modes. Red dotted lines denote the results of the reduced models. The number of bath modes included in the models is given in each inset.

dimensions was selected. Unlike the case with the N_j descriptor [Fig. 1(a)], the decay of E_i with the Q_j descriptor [Fig. 1(d)] is rather different. A long tail in the decay curve was observed in the PCA results [Fig. 1(d)], and the first few dimensions included a large number of bath modes. When the strategy defined by Eqs. (11) and (12) was used to select the involved dimensions and modes, most modes become important ones even for small thresholds. Thus, we proposed another standard for dimension and mode selection. The important dimensions were selected according to the value of E_i directly. For illustration, we labeled the number of the selected dimensionality as Q^D . Here, we considered two cases: the leading four dimensions ($Q^D = 4$) and the leading seven dimensions ($Q^D = 7$). To select the important modes, we defined a cutoff value Q_i^M for the i th chosen dimension as

$$Q_i^M \leq \frac{|U_{ij}|^2}{\max(|U_{i1}|^2, |U_{i2}|^2, \dots, |U_{iN_b}|^2)}, \quad (13)$$



where Q denotes that the descriptor is the bath coordinate and the subscript i refers to the i th dimension. We use the label Q^M in the following discussion because the same Q_i^M was used for all dimensions.

When the first four leading dimensions were selected ($Q^D = 4$), the fastest E_i decay portion in Fig. 1(d) was taken into account. The selection of $Q^D = 4$ and $Q^M = 0.1$ defines a reduced model composed of 61 modes distributed in the low-frequency and resonant region, as shown in Fig. 5(a). In this case, the reduced model can reproduce the dynamics evolution qualitatively; however, the accuracy is still far from satisfactory [Fig. 6(a)]. For the $Q^D = 7$ and $Q^M = 0.1$ case, 80 modes were included in the reduced model [Fig. 5(c)] and no obvious improvement in the dynamics description was observed [Fig. 6(c)]. This indicates that the modes included in the reduced model were not enough. When $Q^M = 0.01$ was considered to add a sufficient number of bath modes [Figs. 5(b) and 5(d)], a better description of the dynamics could be realized, except

FIG. 7. PCA results of the original 600-mode model (adiabatic bath region) with $\Delta_E = 0.09$ eV and $V_{12} = 0.12$ eV using different descriptors, i.e., the action variable N_j (left panel) or the bath coordinate Q_j (right panel). (a) and (d) The variance ratio E_i along each principal component; (b) and (e) the heatmap of $|U_j|^2$: the abscissa is the principal component and the ordinate represents the bath mode, the darker color represents the larger $|U_j|^2$; and (c) and (f) the heatmap of the first 10-dimension $|U_j|^2$.

for the minor deviations in the long-time propagation [Figs. 6(b) and 6(d)].

The PCA results based on different descriptors basically provided a consistent physical picture on the role of the bath motion in the nonadiabatic dynamics. Both results predicted that there were two important types of modes, which either display strong vibronic couplings or have frequencies close to the Rabi frequency of the electronic transition. In addition, when a similar number of modes were chosen according to the PCA results of the two different descriptors, the dynamics based on these reduced models were also similar. For example, both the 58-mode model [N_j descriptor, Fig. 3(a)] and 61-mode model [Q_j descriptor, Fig. 3(a)] provided qualitative descriptions of the population dynamics of the original 400-mode model; however, they could not give a precise description of the short-time electronic coherence and the long-time-scale dynamics. When the reduced model contained more than 100 modes [Figs. 3(b) and 6(b)], the overall feature of the nonadiabatic dynamics with the full 400 modes may be well described, e.g., the oscillation pattern in the population dynamics, although some minor differences may exist in the long-time-scale dynamics.

All the above PCA results were obtained based on data from 2500 trajectories. We also performed PCA based on the data from 10 000 trajectories, as shown in Appendix A. Although the important bath modes chosen by the PCA may be different, they belong to the same frequency region. Thus, the above discussion on the role of bath motion is still reasonable.

B. Adiabatic bath

Next, we considered the model falling in the adiabatic bath region in which the characteristic frequency of the bath is much smaller than the Rabi frequency. For this, the parameters $\Delta_E = 0.09$ eV, $V_{12} = 0.12$ eV, $\omega_c = 250$ cm⁻¹ (~ 0.06 eV), and $\lambda = 125$ cm⁻¹ (~ 0.0155 eV) were used. The Rabi frequency of the electronic part was ~ 0.256 eV. In the current model, the Rabi frequency was more than four times larger than the characteristic frequency of the bath. The sampling interval $\Delta\omega$ was selected as 10 cm⁻¹ (~ 0.0012 eV); thus, a total of 600 modes were used, i.e., modes 1–300 coupled to the LE1 state and modes 301–600 coupled to the LE2 state.

1. PCA results with the N_j descriptor

When the action variable N_j was used as the descriptor, the value of E_i decayed quicker in the first ~ 250 dimensions, as shown in Fig. 7(a), and the rest of the dimensions had minor contributions to the dynamics.

As shown by the heatmap of $|U_{ij}|^2$ in Figs. 7(b) and 7(c), the bath modes, which are in the low-frequency region or close to the resonance region, show rather significant contributions in the leading dimensions of the PCA. The overall bath motion is clearly much slower than electronic motion in the adiabatic model. As a result, the modes in the low-frequency region with strong vibronic coupling become more important. However, because the Debye spectral density shows a rather broad distribution over the vibrational frequency, the modes with frequencies close to the Rabi oscillation frequency ~ 0.256 eV still display visible vibronic coupling strength, and these results demonstrate their important roles in the dynamics.

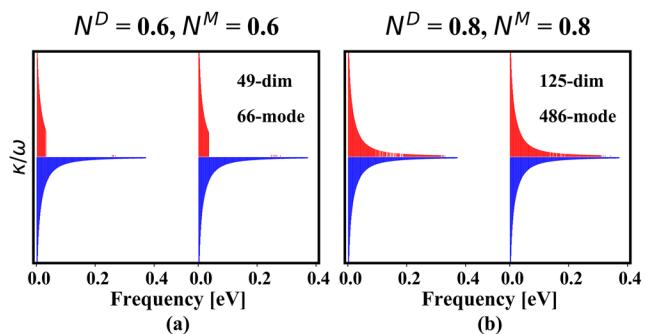


FIG. 8. The κ/ω value of the bath mode in the original 600-mode model ($\Delta_E = 0.09$ eV and $V_{12} = 0.12$ eV, the adiabatic bath region) and the reduced models with different cutoff values of N^D [Eq. (11)] or N^M [Eq. (12)]. In each inset, the upper part (red lines) is κ/ω in the reduced model, while the lower part (blue lines) denotes κ/ω in the original 600-mode model.

To understand the above PCA results, we used the proposed strategies [Eqs. (11) and (12)] to determine the contributions of the bath modes, as shown in Fig. 8. In the case of $N^D = 0.6$ and $N^M = 0.6$, the first 47 dimensions were selected and a total of 66 modes were included in the dynamics [Fig. 8(a)]. As shown in Fig. 9(a), there is a significant difference between the dynamics of the reduced model and the original model with the full 600 bath modes. When more modes were added by using $N^D = 0.8$ and $N^M = 0.8$ [Fig. 8(b)], more accurate dynamics could be obtained [Fig. 9(b)]. One of the major differences between the above two models is the inclusion of an increased number of modes that belong to the resonance region. This means that the inclusion of the modes in the resonance region is essential to capture the early time dynamics governed by the electronic coherence.

2. PCA results with the Q_j descriptor

The PCA results using coordinate Q_j as the descriptor are summarized in the right panel of Figs. 7(d)–7(f). Fast and slow decays of

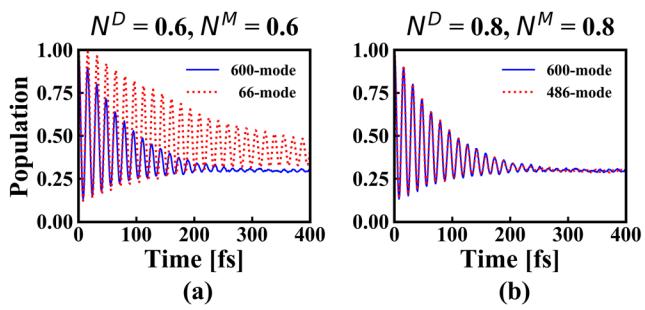


FIG. 9. Time-dependent electronic population of the LE1 state in the dynamics of the models ($\Delta_E = 0.09$ eV, $V_{12} = 0.12$ eV, and the adiabatic bath) with different cutoff values of N^D [Eq. (11)] or N^M [Eq. (12)]. Blue solid lines denote the results of the full model with the full 600 bath modes. Red dotted lines denote the results of the reduced models, and the number of the bath modes included in the models is given in each inset.

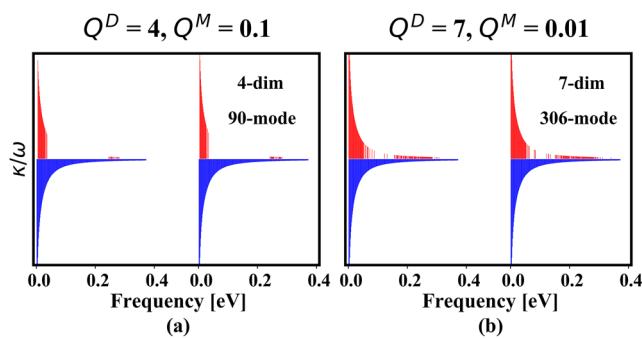


FIG. 10. The κ/ω value of the bath mode in the original 600-mode model ($\Delta_E = 0.09$ eV and $V_{12} = 0.12$ eV, the adiabatic bath region) and the reduced models with different cutoff values of Q^D and Q^M [Eq. (13)]. In each inset, the upper part (red lines) is κ/ω in the reduced model, while the lower part (blue lines) denotes κ/ω in the original 600-mode model.

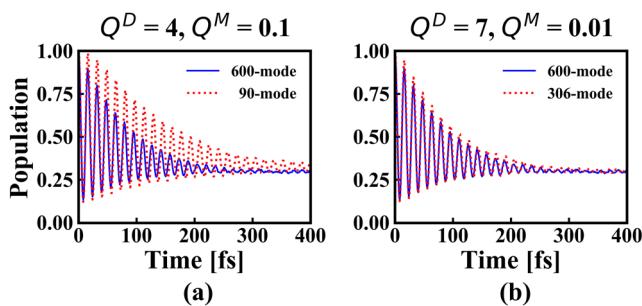


FIG. 11. Time-dependent electronic population of the LE1 state in the dynamics of the model ($\Delta_E = 0.09$ eV and $V_{12} = 0.12$ eV, and the adiabatic bath) with different cutoff values of Q^D and Q^M [Eq. (13)]. Blue solid lines denote the results including the full 600 bath modes with the triangle window. Red dotted lines denote the results of the different reduced models.

E_i are observed with increasing dimensionality, and the distribution of $|U_{ij}|^2$ is highly delocalized. Similarly, we used a few of the reduced models to perform dynamics calculations, and the relevant results are summarized in Figs. 10(a), 10(b), 11(a), and 11(b). As expected, the bath modes in the low-frequency region or close to the resonance region played important roles.

3. “Purely” adiabatic bath model

Because of the broad distribution of the Debye-type spectral density, the employment of such a bath always results in a model in which modes belonging to the resonance region show weak electron–phonon coupling strengths. As a consequence, these near-resonance modes always play important roles in the nonadiabatic dynamics. To remove their contributions, we built a reduced model from the original 600-bath model by keeping all the modes with frequencies less than 0.2 eV. This led to a purely “adiabatic”-bath model in which only the low-frequency modes were included. The dynamics of this reduced model were simulated, and the PCA results are summarized in Appendix C. In this situation,

only the low-frequency modes with strong vibronic couplings are significant.

IV. CONCLUSION

In this work, PCA was employed to explore the roles of the bath modes in the MM-SQC dynamics of the site-exciton models. Two descriptors were used, i.e., the action variable N_j and the normal coordinate Q_j of the bath mode. Two different types of site-exciton models were considered that corresponded with the intermediate-bath and adiabatic-bath models in which the characteristic frequency of the bath is close to or much smaller than the frequency of the electronic transition (Rabi frequency). The MM-SQC dynamics were obtained for both cases. The use of PCA to analyze the trajectory data enabled clear identification of the roles of the different modes.

The bath modes with a frequency close to the Rabi frequency of the electronic transitions play key roles in the nonadiabatic dynamics; when the nuclear vibrational motion displayed similar time scales with respect to the electron dynamics, the inclusion of these modes was essential for proper description of the dynamics, particularly for the early time dynamics governed by coherent electronic motion. The modes displaying strong vibronic coupling should also play important roles. In the current model with the Debye bath spectral density, the low-frequency modes always display significant vibronic couplings; therefore, their contributions can be easily identified via PCA. For the above two models (intermediate and adiabatic bath), the high-frequency modes did not display visible vibronic couplings and their frequencies were also far from the electronic transition, and they play almost no role here.

The above conclusions were consistent with the physical insights obtained based on PCA of the MM-SQC dynamics results, instead of the standard analysis procedures based on the pre-known dynamics procedure. This indicates that dimensionality reduction approaches can be powerful tools for analyzing the simulation results of nonadiabatic dynamics in the condensed phase.

ACKNOWLEDGMENTS

This work was supported by the NSFC projects (Grant Nos. 21873112, 21933011, and 21673266). The authors thank the Supercomputing Center, Computer Network Information Center, CAS, and National Supercomputing Center in Shenzhen for providing computational resources.

APPENDIX A: CONVERGENCE TEST

To test convergence, we performed the PCA based on the data from the MM-SQC dynamics with 10 000 trajectories. Figure 12 shows that the time-dependent electronic populations of the LE1 state based on different trajectory numbers (2500 and 10 000) give the same population dynamics within 400 fs. Figure 13 shows PCA results based on the data from 10 000 trajectories. Figure 14 shows the κ/ω value of the bath mode in the original 400-mode model and the reduced models built with different cutoff

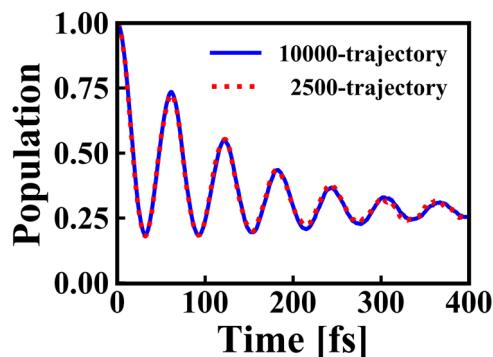


FIG. 12. Time-dependent electronic population of the LE1 state in the dynamics of the original 400-mode model using different trajectory numbers. The blue solid line denotes the result using 10 000 trajectories. The red dotted line denotes the result using 2500 trajectories.

values. Figure 15 shows the time-dependent electronic population of the LE1 state in the reduced models. Although slightly different modes may be chosen in the reduced model with 10 000 trajectories compared to the results based on the MM-SQC dynamics with 2500 trajectories, the selected modes belong to the same frequency region.

APPENDIX B: REDUCED MODEL WITH $N^D = 0.1$ AND $N^M = 0.6$

The PCA results (N_j descriptor) of the MM-SQC dynamics for the 400-mode model ($\Delta_E = 0.03$ eV and $V_{12} = 0.03$ eV and the intermediate bath) are given in Fig. 1. We can build the reduced model based on Eqs. (11) and (12). Here, we showed a simple model including only a few modes.

For $N^D = 0.1$, the first five dimensions were selected because the sum of the ratio of E_i over the first six dimensions is larger than 0.1.

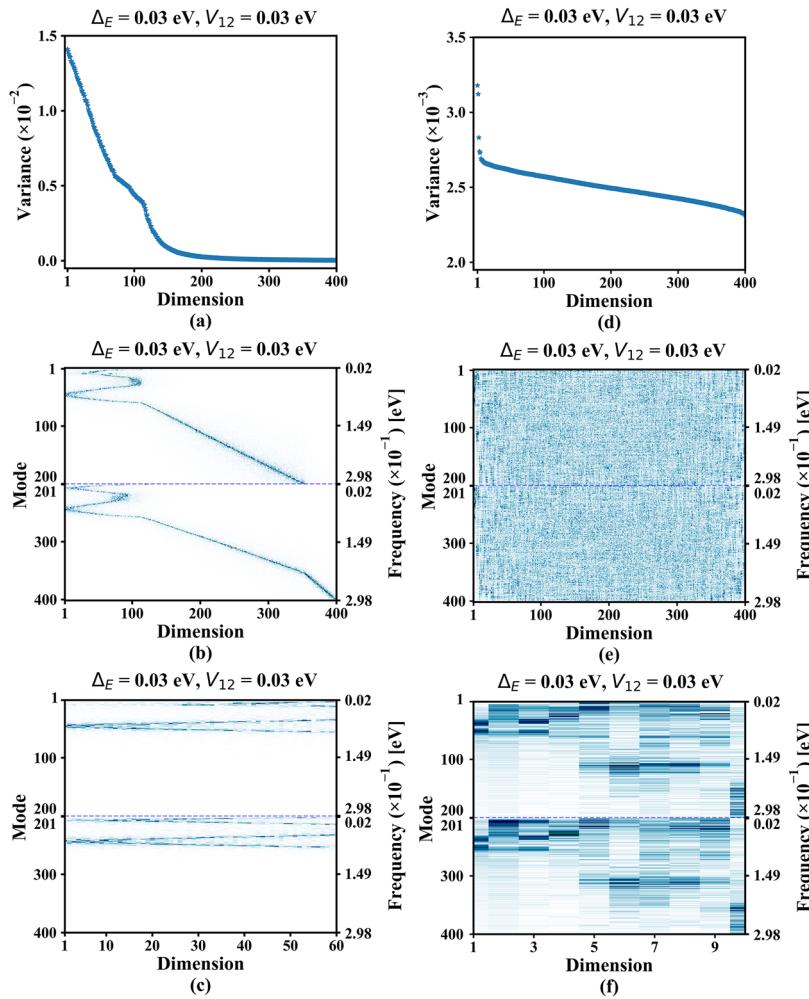
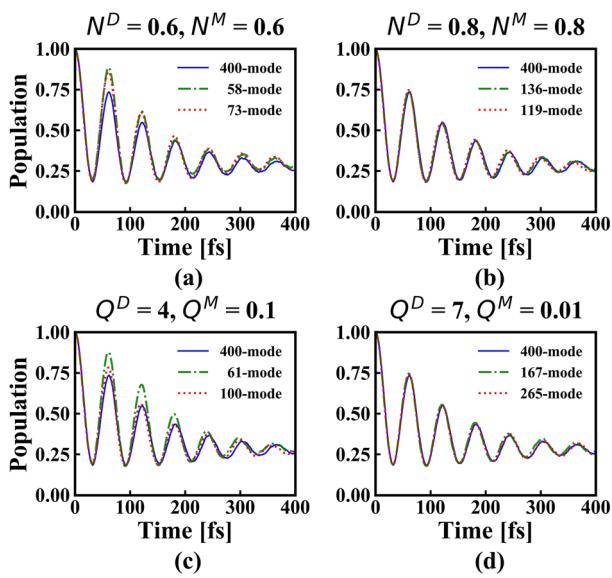
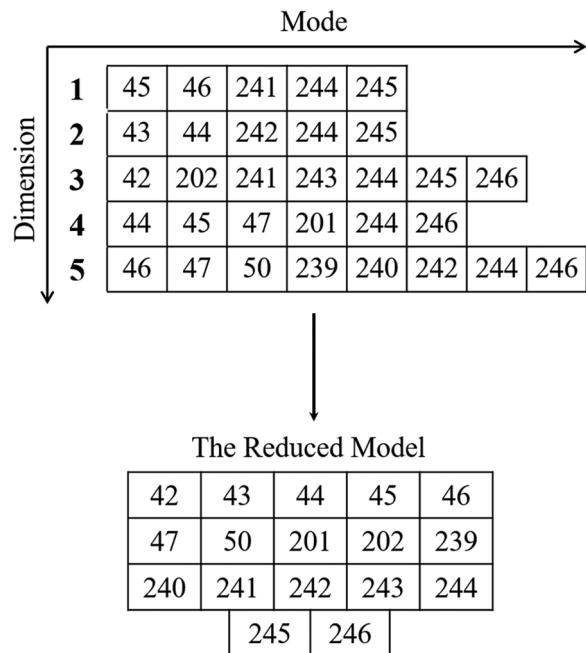
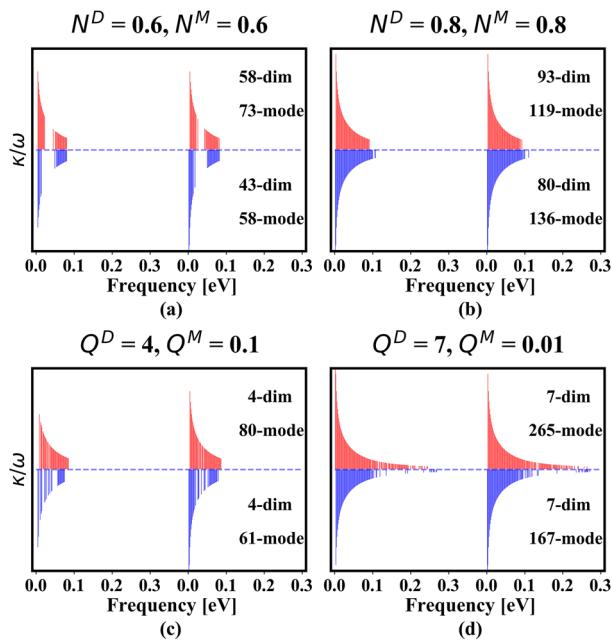
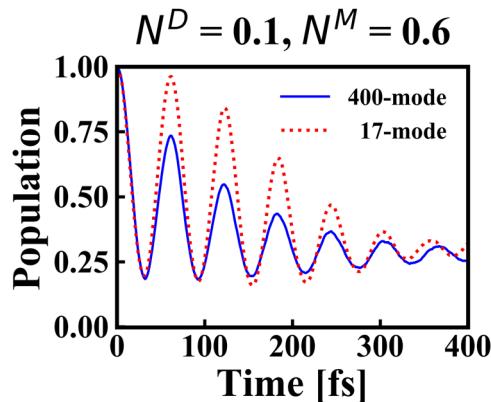


FIG. 13. PCA results of the original 400-mode model (intermediate bath region) with $\Delta_E = 0.03$ eV and $V_{12} = 0.03$ eV based on the data from 10 000 trajectories using different descriptors, namely, the action variable N_j (left panel) or the bath coordinate Q_j (right panel). (a) The variance ratio E_i along each principal component; (b) the heatmap of $|U_{ij}|^2$: the abscissa is the principal component and the ordinate represents the bath mode, the darker color represents the larger $|U_{ij}|^2$; and (c) the heatmap of the first 10-dimension $|U_{ij}|^2$.



The threshold $N^M = 0.6$ was used to determine which modes were selected for the first five dimensions. In this case, modes 45, 46, 241, 244, and 245 were selected in the first dimension and modes 43, 44, 242, 244, and 245 were selected in the second dimension. The chosen modes for each selected dimension are summarized in Fig. 16.



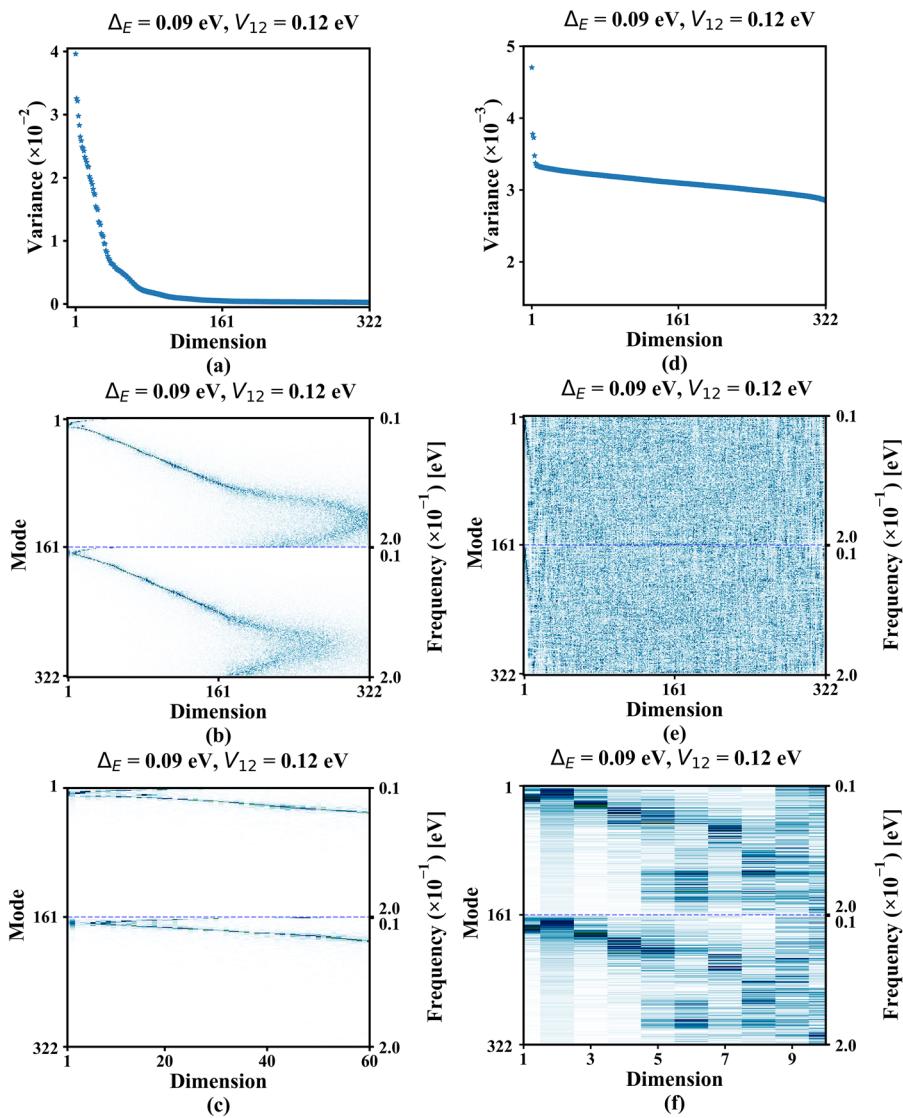


FIG. 18. PCA results of the purely “adiabatic”-bath model with $\Delta_E = 0.09$ eV and $V_{12} = 0.12$ eV using different descriptors, namely, the action variable N_j (left panel) or the bath coordinate Q_j (right panel). (a) The variance ratio E_i along each principal component; (b) the heatmap of $|U_{ij}|^2$: the abscissa is the principal component and the ordinate represents the bath mode, the darker color represents a larger $|U_{ij}|^2$; and (c) the heatmap of the first 10-dimension $|U_{ij}|^2$.

This defines a reduced model with 17 modes, and the corresponding dynamics is given in Fig. 17.

APPENDIX C: PURELY “ADIABATIC”-BATH MODEL

To remove the contributions of those near-resonance modes, we simply built a model from the original 600-bath model by keeping all the modes with frequency less than 0.2 eV. The corresponding PCA results are shown in Fig. 18. It can be seen that only the low-frequency modes with strong vibronic couplings become important.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- ¹Conical Intersections: Electronic Structure, Dynamics and Spectroscopy, edited by W. Domcke, D. R. Yorkony, and H. Köppel (World Scientific, Singapore, 2004).
- ²V. May and O. Kühn, *Charge and Energy Transfer Dynamics in Molecular Systems*, 3rd ed. (Wiley-VCH, Weinheim, Germany, 2011).
- ³M. Schröter, S. D. Ivanov, J. Schulze, S. P. Polyutov, Y. Yan, T. Pullerits, and O. Kühn, *Phys. Rep.* **567**, 1–78 (2015).
- ⁴Y. Tanimura, *J. Phys. Soc. Jpn.* **75**(8), 082001 (2006).
- ⁵Y. Tanimura, *J. Chem. Phys.* **153**(2), 020901 (2020).
- ⁶Y. Yan, *J. Chem. Phys.* **141**(5), 054105 (2014).
- ⁷Y. Tanimura and R. Kubo, *J. Phys. Soc. Jpn.* **58**(1), 101–114 (1989).
- ⁸X. Zhong and Y. Zhao, *J. Chem. Phys.* **138**(1), 014111 (2013).
- ⁹J. Shao, *J. Chem. Phys.* **120**(11), 5053–5056 (2004).
- ¹⁰L. Diósi and W. T. Strunz, *Phys. Lett. A* **235**(6), 569–573 (1997).
- ¹¹G. Ritschel, W. T. Strunz, and A. Eisfeld, *J. Chem. Phys.* **147**(6), 064113 (2017).
- ¹²M. H. Beck, A. Jäckle, G. A. Worth, and H.-D. Meyer, *Phys. Rep.* **324**(1), 1–105 (2000).

- ¹³H. Wang and M. Thoss, *J. Chem. Phys.* **119**(3), 1289–1299 (2003).
- ¹⁴H. Wang, *J. Phys. Chem. A* **119**(29), 7951–7965 (2015).
- ¹⁵*Multidimensional Quantum Dynamics: MCTDH Theory and Applications*, edited by H.-D. Meyer, F. Gatti, and G. Worth (Wiley-VCH, Weinheim, Germany, 2009).
- ¹⁶X. Xie, Y. Liu, Y. Yao, U. Schollwöck, C. Liu, and H. Ma, *J. Chem. Phys.* **151**(22), 224101 (2019).
- ¹⁷J. Ren, Z. Shuai, and G. Kin-Lic Chan, *J. Chem. Theory Comput.* **14**(10), 5027–5039 (2018).
- ¹⁸W. Li, J. Ren, and Z. Shuai, *J. Chem. Phys.* **152**(2), 024127 (2020).
- ¹⁹S. M. Greene and V. S. Batista, *J. Chem. Theory Comput.* **13**(9), 4034–4042 (2017).
- ²⁰Y. Yan, T. Xing, and Q. Shi, *J. Chem. Phys.* **153**(20), 204109 (2020).
- ²¹R. Borrelli and M. F. Gelin, *J. Chem. Phys.* **145**(22), 224101 (2016).
- ²²A. Baiardi and M. Reiher, *J. Chem. Theory Comput.* **15**(6), 3481–3498 (2019).
- ²³B. F. E. Curchod and T. J. Martínez, *Chem. Rev.* **118**(7), 3305–3336 (2018).
- ²⁴D. V. Makarov, C. Symonds, S. Fernandez-Alberti, and D. V. Shalashilin, *Chem. Phys.* **493**, 200–218 (2017).
- ²⁵M. Ben-Nun and T. J. Martínez, *Adv. Chem. Phys.* **121**, 439–512 (2002).
- ²⁶G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt, and B. Lasorne, *Int. Rev. Phys. Chem.* **34**(2), 269–308 (2015).
- ²⁷X. Li, J. C. Tully, H. B. Schlegel, and M. J. Frisch, *J. Chem. Phys.* **123**(8), 084106 (2005).
- ²⁸C. Zhu, S. Nangia, A. W. Jasper, and D. G. Truhlar, *J. Chem. Phys.* **121**(16), 7658–7670 (2004).
- ²⁹J. C. Tully, *J. Chem. Phys.* **93**(2), 1061–1071 (1990).
- ³⁰R. Crespo-Otero and M. Barbatti, *Chem. Rev.* **118**(15), 7026–7068 (2018).
- ³¹J. E. Subotnik, A. Jain, B. Landry, A. Petit, W. Ouyang, and N. Bellonzi, *Annu. Rev. Phys. Chem.* **67**(1), 387–417 (2016).
- ³²A. V. Akimov, A. J. Neukirch, and O. V. Prezhdo, *Chem. Rev.* **113**(6), 4496–4565 (2013).
- ³³S. Mai, P. Marquetand, and L. González, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **8**(6), e1370 (2018).
- ³⁴G. Stock and M. Thoss, *Adv. Chem. Phys.* **131**, 243–375 (2005).
- ³⁵H.-T. Chen and D. R. Reichman, *J. Chem. Phys.* **144**(9), 094104 (2016).
- ³⁶U. Müller and G. Stock, *J. Chem. Phys.* **107**(16), 6230–6245 (1997).
- ³⁷*Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. Berne, G. Ciccotti, and D. Coker (World Scientific, Singapore, 1998).
- ³⁸M. Thoss and H. Wang, *Annu. Rev. Phys. Chem.* **55**(1), 299–332 (2004).
- ³⁹S. Bonella and D. F. Coker, *Chem. Phys.* **268**(1), 189–200 (2001).
- ⁴⁰S. Bonella and D. F. Coker, *J. Chem. Phys.* **114**(18), 7778–7789 (2001).
- ⁴¹R. Kapral and G. Ciccotti, *J. Chem. Phys.* **110**(18), 8919–8929 (1999).
- ⁴²H. Kim, A. Nassimi, and R. Kapral, *J. Chem. Phys.* **129**(8), 084102 (2008).
- ⁴³X. Sun, H. Wang, and W. H. Miller, *J. Chem. Phys.* **109**(17), 7064–7074 (1998).
- ⁴⁴Q. Shi and E. Geva, *J. Phys. Chem. A* **107**(43), 9070–9078 (2003).
- ⁴⁵I. Borg and P. J. F. Groenen, *Modern Multidimensional Scaling: Theory and Applications*, 2nd ed. (Springer, New York, 2005).
- ⁴⁶W. Härdle and L. Simar, *Applied Multivariate Statistical Analysis*, 3rd ed. (Springer, Heidelberg, Berlin, 2012).
- ⁴⁷S. G. Sireci, *J. Educ. Meas.* **40**(3), 277–280 (2003).
- ⁴⁸C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006).
- ⁴⁹R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker, *Proc. Natl. Acad. Sci. U. S. A.* **102**(21), 7426–7431 (2005).
- ⁵⁰R. R. Coifman and S. Lafon, *Appl. Comput. Harmonic Anal.* **21**(1), 5–30 (2006).
- ⁵¹A. Amadei, A. B. M. Linssen, and H. J. C. Berendsen, *Proteins* **17**(4), 412–425 (1993).
- ⁵²G. Capano, T. J. Penfold, M. Chergui, and I. Tavernelli, *Phys. Chem. Chem. Phys.* **19**(30), 19590–19600 (2017).
- ⁵³J. P. Zauleck, S. Thallmair, M. Loipersberger, and R. de Vivie-Riedle, *J. Chem. Theory Comput.* **12**(12), 5698–5708 (2016).
- ⁵⁴S. R. Hare, L. A. Bratholm, D. R. Glowacki, and B. K. Carpenter, *Chem. Sci.* **10**(43), 9954–9968 (2019).
- ⁵⁵P. Das, M. Moll, H. Stamati, L. E. Kavraki, and C. Clementi, *Proc. Natl. Acad. Sci. U. S. A.* **103**(26), 9885–9890 (2006).
- ⁵⁶M. A. Rohrdanz, W. Zheng, and C. Clementi, *Annu. Rev. Phys. Chem.* **64**(1), 295–316 (2013).
- ⁵⁷X. Li, Y. Xie, D. Hu, and Z. Lan, *J. Chem. Theory Comput.* **13**(10), 4611–4623 (2017).
- ⁵⁸W. Shi, T. Jia, and A. Li, *Phys. Chem. Chem. Phys.* **22**(31), 17460–17471 (2020).
- ⁵⁹T. Tsutsumi, Y. Ono, Z. Arai, and T. Taketsugu, *J. Chem. Theory Comput.* **16**(7), 4029–4037 (2020).
- ⁶⁰A. M. Virshup, J. Chen, and T. J. Martínez, *J. Chem. Phys.* **137**(22), 22A519 (2012).
- ⁶¹A. K. Belyaev, W. Domcke, C. Lasser, and G. Trigila, *J. Chem. Phys.* **142**(10), 104307 (2015).
- ⁶²S. Reiter, T. Schnappinger, and R. de Vivie-Riedle, paper presented at the Artificial Neural Networks and Machine Learning-ICANN 2019: Workshop and Special Sessions, Cham, 2019.
- ⁶³J. P. P. Zauleck and R. de Vivie-Riedle, *J. Chem. Theory Comput.* **14**(1), 55–62 (2018).
- ⁶⁴S. Gómez, M. Heindl, A. Szabadi, and L. González, *J. Phys. Chem. A* **123**(38), 8321–8332 (2019).
- ⁶⁵S. Mai, M. F. S. J. Menger, M. Marazzi, D. L. Stolba, A. Monari, and L. González, *Theor. Chem. Acc.* **139**(3), 65 (2020).
- ⁶⁶F. Liu, L. Du, D. Zhang, and J. Gao, *Sci. Rep.* **7**(1), 8737 (2017).
- ⁶⁷F. Häse, I. Fdez. Galván, A. Aspuru-Guzik, R. Lindh, and M. Vacher, *Chem. Sci.* **10**(8), 2298–2307 (2019).
- ⁶⁸S. Wold, K. Esbensen, and P. Geladi, *Chemometr. Intell. Lab.* **2**(1–3), 37–52 (1987).
- ⁶⁹H. Abdi and L. J. Williams, *Wiley Interdiscip. Rev.: Comput. Stat.* **2**(4), 433–459 (2010).
- ⁷⁰I. T. Jolliffe and J. Cadima, *Philos. Trans. R. Soc., A* **374**, 2065 (2016).
- ⁷¹Y. Xie, J. Zheng, and Z. Lan, *J. Chem. Phys.* **149**(17), 174105 (2018).
- ⁷²H. D. Meyer and W. H. Miller, *J. Chem. Phys.* **70**(7), 3214–3223 (1979).
- ⁷³S. J. Cotton and W. H. Miller, *J. Chem. Theory Comput.* **12**(3), 983–991 (2016).
- ⁷⁴S. J. Cotton and W. H. Miller, *J. Chem. Phys.* **145**(14), 144108 (2016).
- ⁷⁵R. Liang, S. J. Cotton, R. Binder, R. Hegger, I. Burghardt, and W. H. Miller, *J. Chem. Phys.* **149**(4), 044101 (2018).
- ⁷⁶J. S. Sandoval C., A. Mandal, and P. Huo, *J. Chem. Phys.* **149**(4), 044115 (2018).
- ⁷⁷X. Gao, M. A. C. Saller, Y. Liu, A. Kelly, J. O. Richardson, and E. Geva, *J. Chem. Theory Comput.* **16**(5), 2883–2895 (2020).
- ⁷⁸A. A. Kananenka, C.-Y. Hsieh, J. Cao, and E. Geva, *J. Phys. Chem. Lett.* **9**(2), 319–326 (2018).
- ⁷⁹X. He and J. Liu, *J. Chem. Phys.* **151**(2), 024105 (2019).
- ⁸⁰W. H. Miller and S. J. Cotton, *Faraday Discuss.* **195**, 9–30 (2016).
- ⁸¹N. Bellonzi, A. Jain, and J. E. Subotnik, *J. Chem. Phys.* **144**(15), 154110 (2016).
- ⁸²A. Jain and J. E. Subotnik, *J. Phys. Chem. A* **122**(1), 16–27 (2018).
- ⁸³G. Tao, *J. Chem. Theory Comput.* **11**(1), 28–36 (2015).
- ⁸⁴G. Tao, *J. Phys. Chem. C* **118**(31), 17299–17305 (2014).
- ⁸⁵W. Zhou, A. Mandal, and P. Huo, *J. Phys. Chem. Lett.* **10**(22), 7062–7070 (2019).
- ⁸⁶S. J. Cotton and W. H. Miller, *J. Chem. Phys.* **139**(23), 234112 (2013).
- ⁸⁷J. Provazza and D. F. Coker, *J. Chem. Phys.* **148**(18), 181102 (2018).
- ⁸⁸M. Thoss, H. Wang, and W. H. Miller, *J. Chem. Phys.* **115**(7), 2991–3005 (2001).
- ⁸⁹J. Zheng, J. Peng, Y. Xie, Y. Long, X. Ning, and Z. Lan, *Phys. Chem. Chem. Phys.* **22**(32), 18192–18204 (2020).
- ⁹⁰H. Tamura, R. Martinazzo, M. Ruckebauer, and I. Burghardt, *J. Chem. Phys.* **137**(22), 22A540 (2012).
- ⁹¹G. Stock and M. Thoss, *Phys. Rev. Lett.* **78**(4), 578–581 (1997).

- ⁹²G. Stock and U. Müller, *J. Chem. Phys.* **111**(1), 65–76 (1999).
- ⁹³U. Müller and G. Stock, *J. Chem. Phys.* **111**(1), 77–88 (1999).
- ⁹⁴A. A. Golosov and D. R. Reichman, *J. Chem. Phys.* **114**(3), 1065–1074 (2001).
- ⁹⁵V. De Silva and J. B. Tenenbaum, “Sparse multidimensional scaling using landmark points,” Technical Report, Stanford University, 2004.
- ⁹⁶B. Nadler, S. Lafon, R. R. Coifman, and I. G. Kevrekidis, *Appl. Comput. Harmonic Anal.* **21**(1), 113–127 (2006).
- ⁹⁷S. Mai and L. González, *J. Chem. Phys.* **151**(24), 244115 (2019).
- ⁹⁸M. E. Wall, A. Rechtsteiner, and L. M. Rocha, in *A Practical Approach to Microarray Data Analysis*, edited by D. P. Berrar, W. Dubitzky, and M. Granzow (Springer, Boston, MA, 2003), pp. 91–109.
- ⁹⁹Y. Xie, J. Zheng, and Z. Lan, *J. Chem. Phys.* **142**(8), 084706 (2015).
- ¹⁰⁰J. Zheng, Y. Xie, S. Jiang, and Z. Lan, *J. Phys. Chem. C* **120**(3), 1375–1389 (2016).