This is a repository copy of *Ab initio multiple cloning algorithm for quantum nonadiabatic molecular dynamics*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/88795/

Version: Accepted Version

Article:

Ab initio Multiple Cloning algorithm for quantum nonadiabatic molecular dynamics

Dmitry Makhov,[1] William J. Glover,[2,3] Todd J. Martinez,[2,3] Dmitrii Shalashilin[1]

[1] Department of Chemistry, University of Leeds

[2] Department of Chemistry and The PULSE Institute, Stanford University, Stanford, CA 94305

[3] SLAC National Accelerator Laboratory, Menlo Park, CA

Abstract: We present a new algorithm for *ab initio* quantum nonadiabatic molecular dynamics that combines the best features of *ab initio* Multiple Spawning (AIMS) and Multiconfigurational Ehrenfest (MCE) methods. In this new method, *ab initio* multiple cloning (AIMC), the individual trajectory basis functions (TBFs) follow Ehrenfest equations of motion (as in MCE). However, the basis set is expanded (as in AIMS) when these TBFs become sufficiently mixed, preventing prolonged evolution on an averaged potential energy surface. We refer to the expansion of the basis set as "cloning," in analogy to the "spawning" procedure in AIMS. This synthesis of AIMS and MCE allows us to leverage the benefits of mean-field evolution during periods of strong nonadiabatic coupling while simultaneously avoiding mean-field artifacts in Ehrenfest dynamics. We explore the use of time-displaced basis sets, "trains," as a means of expanding the basis set for little cost. We also introduce a new bra-ket averaged Taylor expansion (BAT) to approximate the necessary potential energy and nonadiabatic coupling matrix elements. The BAT approximation avoids the necessity of computing electronic structure information at intermediate points between TBFs, as is usually done in saddle-point approximations used in AIMS. The efficiency of AIMC is demonstrated on the nonradiative decay of the first excited state of ethylene. The AIMC method has been implemented within the AIMS-MOLPRO package, which was extended to include Ehrenfest basis functions.

I. Introduction

The dynamics of electronically nonadiabatic transitions is fundamental to many photoinduced processes in chemistry and biochemistry. Examples include photosynthetic light harvesting, the first events in visual reception in rhodopsin, and fluorescent sensors such as the Green Fluorescent Protein (GFP). As the dynamical events in these problems are initiated by electronic excitation, multiple electronic states and nonadiabatic transitions between these states are invariably involved. Furthermore, both nuclear and electronic coherence effects can be crucial, as recently shown in the context of photosynthetic light harvesting.[1-5] An acceptable model for nonadiabatic processes must be able to describe coherent phenomena, including also the loss of coherence, i.e. decoherence. Ideally, this is achieved by a quantum mechanical description of the nuclear degrees of freedom. At the same time, such models must be able to describe the potential energy surfaces and their nonadiabatic couplings near the conical intersections[6-7] (true degeneracies of two or more electronic states) that often promote electronic transitions. This can be difficult using traditional analytic force fields and provides impetus (even more so than for ground state processes) for *ab initio* molecular dynamics (AIMD) methods[8-10] where the electronic wavefunctions are solved simultaneously with the nuclear dynamics.

Trajectory-based mixed quantum-classical methods are perhaps the simplest approach for modeling nonadiabatic phenomena in the context of *ab initio* molecular dynamics. There are two extremes here, according to the form of the potential energy surface used to guide the trajectories. This can be chosen as an average over the populated electronic states (Ehrenfest mean-field method),[11-13] in which case a single trajectory represents evolution on multiple electronic states. Mean-field methods can often be quite accurate for very short times, but fail when the wavefunction splits into multiple parts (one on each of the involved electronic states). This wavepacket splitting cannot be described because of the chosen ansatz (which demands that

the trajectory describe electronic superposition while maintaining a single phase space center). Indeed, this was well recognized by Meyer and Miller, who proposed incorporating a binning procedure to resolve the resulting dynamics into pure electronic states.[11] The other extreme is the surface hopping method,[14-15] where each trajectory evolves on a pure electronic state and nonadiabatic transitions are described as stochastic hops between different electronic states. The widely used "fewest-switches" surface-hopping[15] leverages an Ehrenfest-like dynamics for the quantum amplitudes (but not the phase space centers of the trajectory) in order to determine when to stochastically switch between electronic states. The description of electronic coherence with a single trajectory (even one which always evolves on a pure electronic state) leads to difficulties in describing these coherences correctly. Indeed, the method sometimes seems to enhance coherence (introducing interferences where none should be present) and other times seems to do the opposite (averaging over interferences that should be present), and these characteristics of the method were recognized at the outset.[15] On the other hand, surface-hopping provides a much better (but still incorrect) description of the expected equilibration behavior compared to the Ehrenfest method that often leads to an unphysical heating of the quantum mechanical subsystem.[16-18]

A major difficulty of the mixed quantum-classical trajectory methods is that there is no simple way to improve their description of quantum coherence. Neither the Ehrenfest nor surface-hopping schemes contain any limit where they are guaranteed to converge to the correct quantum mechanical wavepacket evolution. This can be quite unsettling when considering nonadiabatic AIMD methods, which can be quite computationally time-consuming. Ideally, it would be possible to improve the description of quantum mechanical coherence and verify the robustness of the results. Indeed, this is exactly why the first implementations[19-22] of

nonadiabatic AIMD eschewed Ehrenfest or surface-hopping methods for the dynamics. Instead, the *ab initio* multiple spawning (AIMS) method was developed,[10,23-24] exploiting the physical intuition underlying the surface-hopping method[15] but augmenting it with a basis set expansion perspective. The basis set was expressed in terms of traveling Gaussian wavepackets (trajectory basis functions or TBFs), due to their well-known connections to classical mechanics.[25-26] Because of its formulation as a basis set expansion, the AIMS method is guaranteed to converge to the exact solution of the Schrodinger equation for a large enough basis set and this has been demonstrated for model problems.[27-28]

The development of AIMS occurred during a revival of interest in Gaussian wavepacket methods (motivated largely by their suitability for AIMD), and was quickly followed by the development of the coupled coherent states[29-30] (CCS) and variational multiconfigurational Gaussian wavepacket[31-32] (vMCG) methods. The vMCG method is perhaps the most flexible of all of these, using a variational procedure to determine the equations of motion for the phase space centers of the TBFs. Unfortunately, this also introduces strong couplings between the equations of motion for the phase space centers and the quantum amplitudes, leading to stiff differential equations which are numerically difficult to integrate. Another issue with vMCG in the context of AIMD is that it is not easy to employ adiabatic representations of the electronic states (which are the natural representation provided by electronic structure methods) and a "diabatization" procedure has to be developed. Nevertheless an *ab initio* on-the-fly direct dynamics version of vMCG does exist.[33]

The *ab initio* Multiple Spawning (AIMS)[10,24,34-36] method uses a much simpler choice for the evolution of the TBFs: their phase space centers evolve classically on a specific electronic state and the basis set is expanded adaptively as needed, in a process called spawning. The

spawning becomes particularly important near the intersection with another electronic state, where more Gaussians are spawned on the second PES. The growing adaptive basis follows the dynamics more accurately than a basis of simple classical trajectories.

Recently the Multiconfigurational Ehrenfest (MCE)[37-40] dynamics was introduced. This method also employs predetermined trajectories to guide the basis set, but the MCE recipe is different from that of AIMS: the MCE technique uses Ehrenfest mean field trajectories with the Hamiltonian averaged (with proper weights) between two or more electronic PESs. The MCE approach is a fully coupled basis set expansion of TBFs (like AIMS) and thus is also guaranteed to converge to the correct solution of the Schrodinger equation for a sufficiently large number of TBFs. It has the advantage that TBFs do not separate as quickly as they would in AIMS, which could aid in convergence at short time. Unfortunately, the average potential guiding the Ehrenfest trajectories is likely to hamper convergence at longer times. An intermediate representation would be desirable and might lead to faster convergence of the underlying dynamics.

In this paper we propose a new algorithm that combines the advantages of AIMS and MCE methods. The new *ab initio* multiple cloning (AIMC) algorithm introduces Ehrenfest Configuration Cloning (ECC), which expands the basis set of TBFs representing the wavefunction. The cloning procedure is introduced when any of the TBFs becomes strongly mixed (corresponding to a coherent superposition of multiple electronic states with similar amplitudes) and the forces on these electronic states are considerably different (such that representing the evolution of the wavepacket by its average will be a poor approximation). This will typically happen after a TBF passes through/near a conical intersection or region of strong nonadiabatic coupling. Cloning reexpresses the "mixed state" Ehrenfest configuration as a

superposition of distinct TBFs, each dominated by a single electronic state. By cloning, we sidestep the prolonged mixed state propagation which is known to lead to erroneous retention of coherence[41-42] and violations of detailed balance.[15-17] The cloning procedure is analogous to spawning in AIMS, and the AIMS and AIMC methods can be viewed as complementary approaches.

In AIMD methods, the computational bottleneck is invariably the solution of the electronic structure problem to obtain potential energy surfaces, forces, and nonadiabatic couplings. Here, we use the idea of time-displaced basis functions[43] or "trains"[44] in order to increase the size of the TBF basis set without increasing the number of electronic structure calculations. Time-displaced basis functions follow each other along the same trajectory, i.e. are displaced from each other by a time shift, allowing multiple reuse of electronic structure results. This means of increasing the basis set is particularly efficient in AIMD, where the trains greatly increase the basis set size at almost no extra cost.

Finally, we introduce a bra-ket averaged Taylor expansion method (BAT) for calculating the matrix elements of the potential energy surface and nonadiabatic couplings. The BAT approximation can be computed using only the electronic structure information along each TBF and does not require any further electronic structure calculations at intermediate molecular geometries. A form of this approximation has been previously used for AIMD with MCE dynamics,[37] but here we show that it can easily take advantage of *both* potential energies and gradients along the trajectories. Our AIMC algorithm has been implemented within the AIMS-MOLPRO package.[36]

II. Theory

II.I Working equations

We first provide the equations of the Multiconfigurational Ehrenfest (MCE) approach. The Hamiltonian of a system of electrons and nuclei is given as:

$$\hat{H} = \hat{H}_{el}(r, R) - \frac{\hbar^2}{2}\nabla_R \mathbf{M}^{-1}\nabla_R \ ,$$

(1)

where $\hat{H}_{el}$ is the fixed-nucleus electronic Hamiltonian and $\mathbf{M}^{-1}$ is a diagonal matrix of inverse masses of the atoms. In MCE dynamics, the wave-function is represented as:

$$|\Psi(\mathbf{R},\mathbf{r},t)\rangle = \sum_n c_n(t)|\psi_n(\mathbf{R},\mathbf{r},t)\rangle \ ,$$

(2)

where ψ_n are trajectory basis functions (TBFs) composed of nuclear and electronic parts:

$$|\psi_n(\mathbf{R},\mathbf{r},t)\rangle = |\chi_n(\mathbf{R},t)\rangle \left(\sum_I a_I^n(t)|\varphi_I(\mathbf{r};\mathbf{R})\rangle \right)$$

(3)

where φ_I are electronic wavefunctions (assumed to be orthonormal) and χ_n are Gaussian nuclear basis functions:

$$\chi_n\left(\mathbf{R},\overline{\mathbf{R}}_n(t),\overline{\mathbf{P}}_n(t)\right) = \left(\frac{2\alpha}{\pi}\right)^{N_{dof}/4} \exp\left(-\alpha\left(\mathbf{R}-\overline{\mathbf{R}}_n\right)^2 + \frac{i}{\hbar}\overline{\mathbf{P}}_n\left(\mathbf{R}-\overline{\mathbf{R}}_n\right) + \frac{i}{\hbar}\gamma_n(t)\right),$$

(4)

where $\overline{\mathbf{R}}_n$ and $\overline{\mathbf{P}}_n$ are the phase space center of the nth basis function and N_{dof} is the number of degrees of freedom, i.e. the length of the vector $\overline{\mathbf{R}}_n$. The parameter α determines the width of the Gaussian in each degree of freedom, and we choose it to be atom dependent and constant, i.e. these are frozen Gaussian basis functions.[26] Note that there are three formally redundant sets of variables in Eq. 2: the TBF amplitudes (c_n), the Ehrenfest amplitudes for a given TBF $\left(a_I^n\right)$, and the phase of the frozen Gaussian basis function $\left(\gamma_n\right)$. The redundancy of the phase is resolved by choosing it semiclassically such that

$$\frac{d\gamma_n}{dt} = \frac{\overline{\mathbf{P}}_n \dot{\overline{\mathbf{R}}}_n}{2}.$$
(5)

As in AIMS,[10,23-24,27] the equations of motion for the phase space centers are not chosen variationally,[45] but rather prescribed. Unlike AIMS, we do not use Hamilton's equations for these basis function parameters, but instead choose to use Ehrenfest equations of motion. The electronic representation is taken to be adiabatic in this work (although this is not necessary and a diabatic representation is also possible). The Ehrenfest equations of motion for the phase space centers of a trajectory in the adiabatic representation are given as[46] (here shown for the nth TBF):

$$\dot{\overline{\mathbf{R}}}_n = \mathbf{M}^{-1}\overline{\mathbf{P}}_n$$

$$\dot{\overline{\mathbf{P}}}_n = \overline{\mathbf{F}}_n$$

$$\overline{\mathbf{F}}_n = \sum_I \left|a_I^n\right|^2 \nabla_R V_I\left(\overline{\mathbf{R}}_n\right) + \sum_I \sum_{J \neq I} \left(a_I^n\right)^* a_J^n \mathbf{d}_{IJ}\left(\overline{\mathbf{R}}_n\right)\left[V_I\left(\overline{\mathbf{R}}_n\right) - V_J\left(\overline{\mathbf{R}}_n\right)\right]$$
(6)

where $V_I(\mathbf{R})$ is the adiabatic electronic energy of the Ith electronic state for molecular coordinates $\mathbf{R}$, and $\mathbf{d}_{IJ}$ is the nonadiabatic coupling vector connecting the Ith and Jth electronic states:

$$\mathbf{d}_{IJ}\left(\overline{\mathbf{R}}_n\right) = \left\langle \varphi_I\left(\mathbf{r};\overline{\mathbf{R}}_n\right)\middle| \nabla_R \middle| \varphi_J\left(\mathbf{r};\overline{\mathbf{R}}_n\right)\right\rangle_r$$
(7)

As is commonly done, we assume here and throughout that the second derivative of the electronic wavefunction can be neglected, i.e. that

$$\left\langle \varphi_I \middle| \nabla_R \mathbf{M}^{-1} \nabla_R \middle| \varphi_J \right\rangle \approx 0$$
(8)

The equations of motion for the Ehrenfest amplitudes of each TBF are

$$\dot{a}_I^n = -\frac{i}{\hbar} \sum_J H_{IJ}^{el}\left(\overline{\mathbf{R}}_n, \overline{\mathbf{P}}_n\right) a_J^n,$$
(9)

where the electronic Hamiltonian matrix is defined as:

$$H_{IJ}^{el}\left(\overline{\mathbf{R}}_n,\overline{\mathbf{P}}_n\right) = \begin{cases} V_I\left(\overline{\mathbf{R}}_n\right), & I = J \\ -i\hbar\,\overline{\mathbf{P}}_n\mathbf{M}^{-1}\mathbf{d}_{IJ}\left(\overline{\mathbf{R}}_n\right), & I \neq J \end{cases} \quad. \tag{10}$$

Both $V_I\left(\overline{\mathbf{R}}_n\right)$ and $\mathbf{d}_{IJ}\left(\overline{\mathbf{R}}_n\right)$ can be obtained from standard electronic structure codes for many methods, including complete active space self-consistent field (CASSCF)[47-48] and perturbation corrected variants (CASPT2).[49-51]

With the above prescribed equations for the evolution of the phase and Ehrenfest amplitudes, we can now obtain the equations of motion for the TBF amplitudes c_n by inserting the ansatz into the time-dependent Schrödinger equation, giving:

$$\sum_n \left\langle \psi_m(t) \middle| \psi_n(t) \right\rangle \dot{c}_n(t) = -\frac{i}{\hbar}\sum_n \left(H_{mn} - i\hbar \left\langle \psi_m(t) \middle| \frac{d}{dt} \middle| \psi_n(t) \right\rangle \right) c_n(t), \tag{11}$$

where

$$\left\langle \psi_m \middle| \psi_n \right\rangle = \sum_I \left\langle a_I^m \chi_m \middle| a_I^n \chi_n \right\rangle \equiv S_{mn} \tag{12}$$

is an overlap matrix and

$$H_{mn} = \sum_{IJ} \left\langle a_I^m \chi_m \varphi_I \middle| \hat{H} \middle| a_J^n \chi_n \varphi_J \right\rangle \tag{13}$$

are the matrix elements of the full (electronic and nuclear) Hamiltonian from Eq. 1. The term $\left\langle \psi_m(t) \middle| \frac{d}{dt} \middle| \psi_n(t) \right\rangle$, which originates from the time dependence of the basis functions, can easily be calculated analytically:

$$\left\langle \psi_m \middle| \frac{d\psi_n}{dt} \right\rangle = \left\langle \chi_m \middle| \frac{d\chi_n}{dt} \right\rangle \sum_I \left(a_I^m\right)^* a_I^n + \left\langle \chi_m \middle| \chi_n \right\rangle \sum_I \left(a_I^m\right)^* \dot{a}_I^n , \tag{14}$$

where,

$$\left\langle \chi_m \middle| \frac{d\chi_n}{dt} \right\rangle = \dot{\overline{\mathbf{R}}}_n \left\langle \chi_m \middle| \frac{d}{d\overline{\mathbf{R}}_n} \middle| \chi_n \right\rangle + \dot{\overline{\mathbf{P}}}_n \left\langle \chi_m \middle| \frac{d}{d\overline{\mathbf{P}}_n} \middle| \chi_n \right\rangle + \frac{i}{\hbar}\dot{\gamma}_n \left\langle \chi_m \middle| \chi_n \right\rangle. \tag{15}$$

Detailed formulas for the integrals in Eqs 14 and 15 are given in the Appendix.

It remains to deal with the integrals in Eq. 13. In general, these cannot be evaluated completely analytically, at least when the electronic structure is being solved "on the fly." Neglecting the second derivative coupling term (see above), the matrix element becomes:

$$\langle \chi_m \varphi_I | \hat{H} | \chi_n \varphi_J \rangle =$$
$$= \delta_{IJ} \langle \chi_m | -\frac{\hbar^2}{2} \nabla_R \mathbf{M}^{-1} \nabla_R | \chi_n \rangle + \delta_{IJ} \langle \chi_m | V_I(\mathbf{R}) | \chi_n \rangle - \hbar^2 \langle \chi_m | \mathbf{d}_{IJ}(\mathbf{R}) \mathbf{M}^{-1} \nabla_R | \chi_n \rangle \qquad (16)$$

The first term (kinetic energy) is easily calculated analytically (see Appendix). Approximations are needed for the two remaining terms. A saddle point approximation (SPA) has been used previously in AIMS, expanding V($\mathbf{R}$) or $\mathbf{d}_{IJ}$($\mathbf{R}$) in a Taylor series centered about the maximum of the $\chi_m \chi_n$ product (the "centroid" of the two basis functions). This has the drawback of requiring electronic structure calculations at the centroids, implying $O(N_{TBF}^2)$ electronic structure calculations at each time step. In practice, many of these can be (and are) neglected because the basis functions are far apart. The practical scaling in terms of number of electronic structure calculations per time step is thus intermediate between $O(N_{TBF})$ and $O(N_{TBF}^2)$. Here, we show that a slightly different (bra-ket averaged Taylor expansion or BAT) approximation can be much more efficient. Instead of a single Taylor expansion, we average the result of two Taylor expansions, each one centered about the maximum of one of the basis functions involved in the matrix element. For the second term in Eq. 16, expansion to first order yields:

$$\langle \chi_m | V_I(\mathbf{R}) | \chi_n \rangle \approx \langle \chi_m | \chi_n \rangle \left(\frac{V_I(\bar{\mathbf{R}}_n) + V_I(\bar{\mathbf{R}}_m)}{2} \right)$$
$$+ \left(\frac{\langle \chi_m | (\mathbf{R} - \bar{\mathbf{R}}_n) | \chi_n \rangle \nabla_R V_I(\bar{\mathbf{R}}_n) + \langle \chi_m | (\mathbf{R} - \bar{\mathbf{R}}_m) | \chi_n \rangle \nabla_R V_I(\bar{\mathbf{R}}_m)}{2} \right) \qquad (17)$$

Note that the values of the potential energy and its gradient at these two positions $\left(\bar{\mathbf{R}}_n \text{ and } \bar{\mathbf{R}}_m \right)$ were already required in order to propagate the phase space centers in Eq. 6. Thus, there are no

extra electronic structure calculations associated with the BAT approximation (in contrast to the SPA). A zeroth order BAT approximation is used for the third term in Eq. 16:

$$\left\langle \chi_m \middle| \mathbf{d}_{IJ}(\mathbf{R})\mathbf{M}^{-1}\nabla_R \middle| \chi_n \right\rangle \approx \frac{i}{2\hbar}\left\langle \chi_m \middle| \chi_n \right\rangle\left(\overline{\mathbf{P}}_m\mathbf{M}^{-1}\mathbf{d}_{IJ}\left(\overline{\mathbf{R}}_m\right)+\overline{\mathbf{P}}_n\mathbf{M}^{-1}\mathbf{d}_{IJ}\left(\overline{\mathbf{R}}_n\right)\right). \tag{18}$$

An additional advantage of the BAT approximation for the matrix elements in Eq. 13 is that one can separate the evolution of TBFs from the solution of the TBF amplitudes c_n. Thus, it is possible to *independently* evolve all of the desired TBFs using Eqs. 5, 6, and 9. Provided all of the electronic structure information at each time step for each of the TBFs was saved, one then has all the needed information to solve Eq. 11 which determines the TBF amplitudes c_n and thus provides the time-evolved nuclear wavefunction. This is not possible with the SPA used previously in AIMS.

We have previously used a zeroth order BAT approximation for the potential energy matrix elements in MCE.[37] The first-order BAT approximation should be better and, as noted above, incurs no extra cost. Just as in the SPA,[10] it is of course possible to go to second order, but this requires the evaluation of second derivatives of the potential energy surfaces, which is usually quite costly.

A rather subtle point is worthy of some discussion here. The electronic wave function φ_I in Eq. 3 depends parametrically on the nuclear coordinates $\mathbf{R}$. However, because of the BAT approximation for the matrix elements, it is only ever calculated at the center of the wave packet, i.e. for $\mathbf{R}=\overline{\mathbf{R}}_\mathbf{n}(t)$. One can thus reformulate the above MCE approach with the alternative ansatz (compare Eqs. 2 and 3):

$$\left|\Psi(\mathbf{R},\mathbf{r},t)\right\rangle = \sum_n c_n(t)\left|\tilde{\psi}_n(\mathbf{R},\mathbf{r},t)\right\rangle, \tag{19}$$

$$\left|\tilde{\psi}_n(\mathbf{R},\mathbf{r},t)\right\rangle = \left|\chi_n(\mathbf{R},t)\right\rangle\left(\sum_I a_I^n(t)\left|\varphi_I^{\,n}(\mathbf{r};\overline{\mathbf{R}}_\mathbf{n}(t))\right\rangle\right). \tag{20}$$

In this case, the nonadiabatic coupling no longer originates from the kinetic energy operator, since the electronic wavefunction does not depend on the nuclear coordinates (but only on the basis function parameters). Instead, nonadiabatic coupling terms will appear from the time derivative of $\tilde{\psi}_n$, i.e. the analog of Eq. 14 will have new terms involving the nonadiabatic coupling. Fortunately, the final equations are unchanged as long as the electronic wave function does not depend too strongly on the nuclear coordinates. A detailed derivation is provided in the Appendix.

II.II Basis improvement

II.II.I Coherent state trains

Apart from the BAT approximation for matrix elements, the only approximation in MCE is the finite size of the basis set, i.e. we only use a limited number of TBFs. Thus, it would be useful to expand the basis set, if this were possible at little or no computational expense. Since most of the expense in *ab initio* molecular dynamics will be in the computation of the electronic structure (energies, gradients, and nonadiabatic couplings), one might thus consider adding time-displaced variants of each TBF into the simulation. This has been suggested previously, and is known as time-displaced basis sets[43] or trains.[44] The basic idea is illustrated in Figure 1. The upper panel (Figure 1a) depicts an evolving set of TBFs, with each TBF shown as a circle evolving in time along a dashed line. We now place additional TBFs along the *same* trajectory (dashed line) corresponding to each of the existing TBFs, as shown in Figure 1b. The basis set has been expanded, but no new electronic structure information is needed since the new TBFs all evolve exactly the same as the TBFs in the smaller basis set.

This idea is straightforwardly implemented within the context of a sequential procedure where all TBF evolution is calculated prior to the solution of Eq. 11 for the TBF amplitudes. As

an example, consider that we want each "train" to consist of $N_{train}+1$ time-displaced TBFs, with $N_{initial}$ distinct initial conditions. Our calculation will thus comprise $N_{initial}(N_{train}+1)$ TBFs, while the calculation without time-displacement would have a much smaller basis set of only $N_{initial}$ TBFs. We define a time shift Δt_{train} corresponding to the temporal spacing between the "cars" in the "train" (this should be an integer multiple of the time step used to integrate the trajectories). We calculate the complete time evolution up to $t_{sim}+\Delta t_{train}N_{train}/2$ of the $N_{initial}$ TBFs and then back-propagate each of the initial conditions by $\Delta t_{train}N_{train}/2$. For each initial condition, we thus now have all data required to specify any of the time-displaced TBFs for the entire simulation time t_{sim}. Each of the "cars" in the "train" is now specified by the initial condition TBF it corresponds to and also a time shift. We choose Δt_{train} to be small enough that neighboring basis functions have significant overlap (and thus significant coupling). However, one should not choose it so small that the expansion of the basis is meaningless or that unnecessary linear dependence is introduced (numerically complicating the solution of Eq. 11). A basis set of several trains, as shown in Fig 1b, can be made quite large and dense at almost no additional cost (in the context of *ab initio* molecular dynamics, where the electronic structure is much more costly than the solution of Eq. 11) because electronic structure information is reused within each train many times.

II.II.II Basis-function cloning

As in any mean field approach, the fundamental difficulty with Ehrenfest propagation occurs when the average is no longer a faithful representation of the whole. In the context of nonadiabatic dynamics, this occurs when an Ehrenfest trajectory has significant population on multiple electronic states with widely varying forces. The force that governs the Ehrenfest trajectory becomes an unphysical average of the different forces on each electronic state and the

Ehrenfest trajectory can follow a path that does not resemble the evolution predicted by the Schrodinger equation. In MCE, there are multiple interacting Ehrenfest trajectories. Thus, the failure of any of these Ehrenfest trajectories could be ameliorated in principle by the solution of Eq. 11, *if* there were a large enough number of TBFs. Unfortunately, experience suggests that practical calculations are far from the number of TBFs needed to remedy this defect. The solution we propose here follows the spawning idea introduced in AIMS, i.e. adaptive expansion of the basis set of TBFs. We expand the basis set by "cloning" a new TBF when the Ehrenfest force differs significantly from the force that would be felt on one of the electronic states that the Ehrenfest amplitudes predict to have significant population. The difference between the force on the *I*th state and the averaged force is given by:

$$\Delta \mathbf{F}_{I,n} = \nabla_R V_I - \sum_J \left(a_J^n \right)^* a_J^n \ \nabla_R V_J \tag{21}$$

We also want to consider the population on the *I*th state, since there is no point in cloning to the *I*th state if the corresponding Ehrenfest amplitudes indicate it is insignificantly populated. Thus, we define the "breaking force" which will trigger cloning as:

$$\mathbf{F}_{I,n}^{br} = \left(a_I^n \right)^* a_I^n \Delta \mathbf{F}_{I,n} \tag{22}$$

Multiplication by the inverse mass matrix gives an acceleration, and we use the norm of this acceleration as the trigger which signals the need to clone, i.e. when $\left| \mathbf{M}^{-1} \mathbf{F}_{I,n}^{br} \right| > \delta_{clone}$, a new TBF will be cloned from the *n*th TBF onto the *I*th electronic state. The cloning threshold δ_{clone} needs to be determined empirically. Choosing a value which is too small will lead to a large rate of basis set expansion with no benefit, while a value which is too large will lead back to the original MCE method.

As in spawning, we choose the basis set expansion such that it does not alter the wavefunction at the time of spawning/cloning. In AIMS, this simply means the newly spawned

TBF enters the calculation with a vanishing amplitude. For AIMC, the procedure is slightly different. The Ehrenfest TBF corresponds to a superposition of multiple electronic states, evolving on an averaged potential energy surface. When we clone, we create two copies of the same TBF, but we then adjust the Ehrenfest amplitudes such that one of the clones corresponds to a pure electronic state and the other contains all the remaining electronic states (in the case of two electronic states which we consider here, both clones correspond to pure electronic states). The sum of the two cloned TBFs is then exactly equivalent to the original "parent" TBF. A schematic description of the cloning procedure is shown in Figure 1c.

Prior to cloning, the TBF is represented as an Ehrenfest configuration ψ_n as in Eq. 3. After cloning, this TBF is replaced by two TBFs, ψ_n' and ψ_n'':

$$\left| \psi_n' \right\rangle = \left| \chi_n \right\rangle \left(\frac{a_I^n}{\left| a_I^n \right|} \times \left| \varphi_I \right\rangle + \sum_{J \neq I} 0 \times \left| \varphi_J \right\rangle \right) \tag{23}$$

$$\left| \psi_n'' \right\rangle = \left| \chi_n \right\rangle \left(0 \times \left| \varphi_I \right\rangle + \frac{1}{\sqrt{1 - \left| a_I^n \right|^2}} \sum_{J \neq I} a_J^n \left| \varphi_J \right\rangle \right). \tag{24}$$

The TBF amplitudes for the two new TBFs are set to:

$$c_n' = c_n \left| a_I^n \right|, \tag{25}$$

$$c_n'' = c_n \sqrt{1 - \left| a_I^n \right|^2} \tag{26}$$

As can be easily verified, the sum of the two clone TBFs is precisely the same as the parent TBF:

$$c_n \left| \psi_n \right\rangle = c_n' \left| \psi_n' \right\rangle + c_n'' \left| \psi_n'' \right\rangle. \tag{27}$$

However, the ensuing dynamics will be different – the positions and momenta of the ψ_n' clone will evolve on the Ith pure electronic state and those of the ψ_n'' clone will evolve on a mixed electronic state that has no contribution from the Ith electronic state. When the nonadiabatic

coupling increases, both of these TBFs will become mixed states again, and they may be subject to further cloning.

In the presently coded version of the AIMC algorithm, we also demand that the nonadiabatic coupling be small before cloning can occur, i.e. $\left|\mathbf{d}_{IJ}\right| < \delta_{nac}$. This ensures that cloning happens outside the region of strong transitions, limiting the rate of basis set growth. This is not strictly necessary and future work will investigate the effects of lifting this restriction. In the specific case of two electronic states, the breaking force that triggers cloning (Eq. 22) can be written as:

$$\mathbf{F}_{1,n}^{br} = -\mathbf{F}_{2,n}^{br} = \left|a_1^n a_2^n\right|^2 \nabla\left(V_1 - V_2\right).$$
(27)

This expression makes it clear that $\left|\mathbf{F}_1^{br}\right|$ is largest when both states have equal amplitudes, vanishes when only one state is populated, and increases as the difference between the forces on the two states increases. These are precisely the cases where the Ehrenfest dynamics of the TBF becomes suspect. Of course, the algorithm for cloning can be further refined and this will be investigated in future work.

III. Computational details and results

In order to benchmark the AIMC method, we performed calculations of the photodynamics of ethylene after $\pi \rightarrow \pi^*$ excitation. As ethylene is the simplest molecule with a C=C double bond, it has been extensively studied both experimentally and computationally.[22,50-58] The AIMC and AIMS simulations were carried out with a modified version of AIMS-MOLPRO,[36] which was extended to include Ehrenfest dynamics for the TBFs. Electronic structure calculations were performed with the complete active space self-consistent field (CASSCF) method using an active space of two electrons in two orbitals and equally-weighted

state-averaging over the three lowest singlet states, i.e. SA-3-CAS(2/2). The electronic wavefunction is represented with the 6-31G** basis set. The width α of the Gaussian trajectory functions was taken to be 4.7 Bohr^{-2} for hydrogen and 22.7 Bohr^{-2} for carbon atoms, as used in the AIMS-MOLPRO package.[59] The simulations are restricted to two electronic states – S_0 and S_1.

We compare the standard MCE method (without cloning) to the AIMC and AIMS methods. For MCE and AIMC, we also compare the influence of trains on the results. Matrix elements for AIMC were evaluated using the BAT approximation as discussed above, while the AIMS calculations use the SPA for this purpose. For the AIMC simulations, we set the cloning thresholds δ_{clone} and δ_{nac} to 5×10^{-6} and 2×10^{-3} atomic units, respectively. The similar "spawning threshold" in AIMS was set to 4.0 atomic units. When the norm of the nonadiabatic coupling vector exceeds the spawning threshold, the AIMS calculations will spawn a new TBF (see previous work[10] for details). In the AIMS calculations, spawning is prohibited for TBFs that are insignificantly populated (norm of TBF amplitude less than 0.1). A similar procedure controls the growth of the basis set in the AIMC calculations, where we allow each TBF to clone at most 3 times.

In all cases, the simulations are initiated with 200 distinct initial TBFs, sampled randomly from a Wigner distribution corresponding to the ground vibrational state in the harmonic approximation, and then projected onto the S1 state. For the MCE and AIMC calculations with train basis sets, we use a time-shift Δt_{train} of 25 atomic time units ($\approx$0.6fs), which corresponds to a nearest neighbor overlap of $\approx$0.6. The number of basis functions in each train, N_{train}, is set to 100 for MCE/AIMC calculations. Initially each train is independent from all other trains. For calculations without cloning, each of the 200 trains is propagated independently so that the size

of basis set (2) remains equal to 100 and does not change during the simulations. In contrast, the size of basis set increases every time cloning occurs in AIMC. Moreover, each event of cloning increases the number of trains, which are coupled with each other. Figure 2 shows the growth of the average number of TBFs in the AIMC simulations. Each of the initial TBFs gives rise to an average of ≈ 4 further TBFs through the cloning process. In the AIMS calculations, each initial condition spawns an average of 4.1 new TBFs, giving a similar basis set growth rate compared to AIMC.

An example of wavefunction spreading over a train basis set with $N_{train}=200$ for ethylene is shown in Figure 3. The initial population is localized on the middle "car" of the set of TBFs in a "train" corresponding to the same initial condition. As time progresses, the window of TBFs along the trajectory that are being included in the calculation also shifts forward. Thus, the corresponding population of the "cars" stays localized in the middle TBF, but spreads over the cars in the train. As Figure 3 shows, for this case, the wavefunction remains largely localized on the 100 TBFs in the middle of the train. On the basis of this and similar preliminary studies, we concluded that $N_{train}=100$ would be sufficient for the MCE and AIMC calculations in this case.

Figure 4 compares the average ground-state population as a function of time given by MCE and AIMC calculations, with and without basis set expansion using trains. The population was averaged over 200 non-interacting trajectories with initial conditions randomly generated from the ground-state vibrational Wigner distribution. The ground-state population evolution predicted by calculations , with and without train basis functions, is very similar. The results are somewhat smoother when the train basis set is used, but the difference is within the error bars. One can conclude that, at least for this problem where the conical intersection is highly peaked (and thus nonadiabatic transitions are both ultrafast and ultra-efficient), the benefits of the train

basis set expansion are not pronounced. Of course, one should also recall that this basis set expansion carries no cost (in the context of *ab initio* molecular dynamics). Future work will investigate the degree to which the representation of the nuclear wavefunction is improved by the use of time-displaced basis set expansion.

The initial population dynamics predicted by MCE and AIMC are similar. However, as the ground state population increases the predictions begin to deviate, and by 100fs they are quite different. The relaxation rate predicted by AIMC is significantly faster. This behavior is as expected – as the population on the ground and excited states becomes nearly equal, the Ehrenfest dynamics of the TBFs will become that of the average of the two electronic states. This will tend to keep the TBFs in the region near the conical intersection longer and population transfer in *both* directions (to the upper state and to the lower state) will be equally probable. This is a manifestation of the violations of detailed balance which are a well known difficulty in Ehrenfest dynamics.[16-17] The cloning procedure in AIMC solves this problem by allowing the TBFs to separate and evolve on adiabatic states.

In Figure 5, we compare the results from AIMS and AIMC calculations. The AIMS results for ethylene have been previously verified by direct comparison to experiment using CASPT2 for the electronic structure.[52] Gratifyingly, the AIMC and AIMS methods agree almost quantitatively on the population dynamics in this case, and are certainly indistinguishable within the error bars shown.

IV Summary

We propose a new algorithm for quantum molecular dynamics (*ab initio* Multiple Cloning or AIMC) which combines the best features of two existing methods: MCE and AIMS. The algorithm uses an ensemble of Ehrenfest trajectories to guide the basis of nuclear Gaussian

coherent states in place of the classical equations of motion used in AIMS. Thus, the trajectories move on an averaged potential energy surface. As is well-known,[15] this can lead to problematic behavior when the electronic states in the average are very different. We alleviate this by expanding the basis set, similar to spawning, in a procedure we call cloning. As for spawning in AIMS, cloning does not change the wavefunction – it merely reexpresses it in an expanded representation. However, it *does* lift the restrictions of the ansatz, i.e. the newly cloned trajectory basis functions now feel the force appropriate to an adiabatic state and are free to separate.

We have also introduced a new approximation (BAT) for the required matrix elements that is an alternative to the SPA used in AIMS. The BAT approximation averages truncated Taylor expansions centered on the two TBFs involved in the matrix element. By comparison, the SPA evaluates the matrix element by a truncated Taylor series expansion about a single point intermediate between the position centers of the two trajectory basis functions. The advantage of the BAT approximation for the matrix element is that it can be computed without any additional information beyond the energies, gradients, and nonadiabatic couplings at the position centers of each TBF. There is no need for evaluation of these quantities at intermediate positions. Hence the number of electronic structure calculations needed scales linearly with the number of TBFs, as compared to a formally quadratic scaling when the SPA is used.

We have implemented the AIMC algorithm in the AIMS module of the MOLPRO package and tested it on the benchmark case of photodynamics in ethylene. We showed that the AIMS and AIMC algorithms give comparable results for the excited state lifetime in this molecule. In this test case, there does not appear to be any benefit from the use of time-displacement, i.e. trains, to increase the size of the nuclear basis set. However, there may be

advantages to time-displacement in other cases or for more sensitive properties than the electronic population (e.g. the absorption spectrum[60-61] or multi-dimensional spectroscopies).

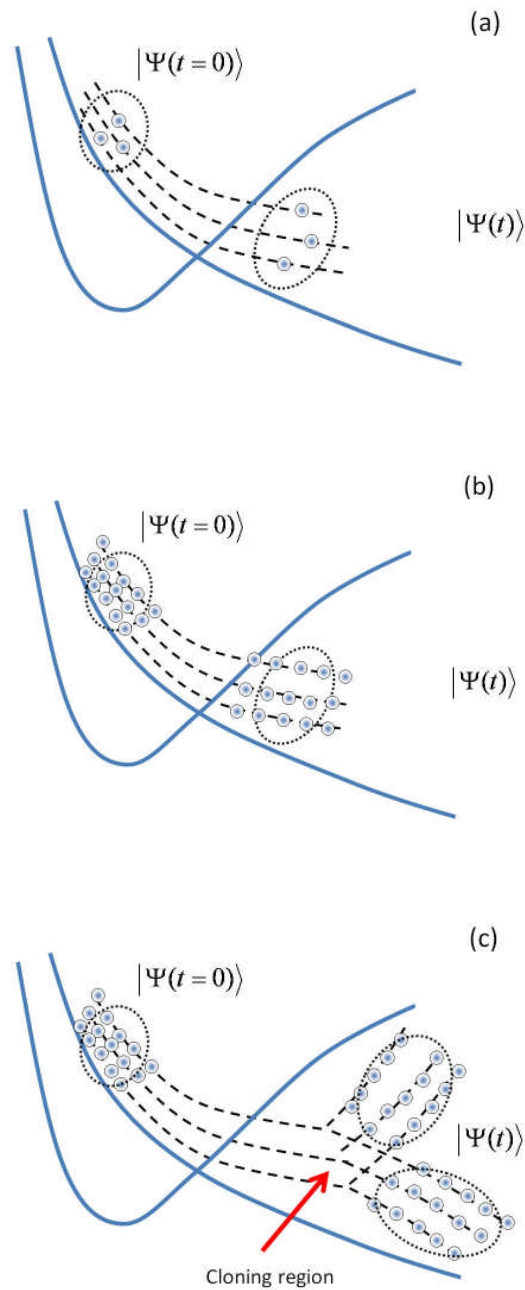**Acknowledgement**

Figures

Figure 1. Graphical illustration of the propagation on a small Ehrenfest trajectory guided basis (a), a basis of Ehrenfest trajectory guided trains (b) and trains with cloning (c)

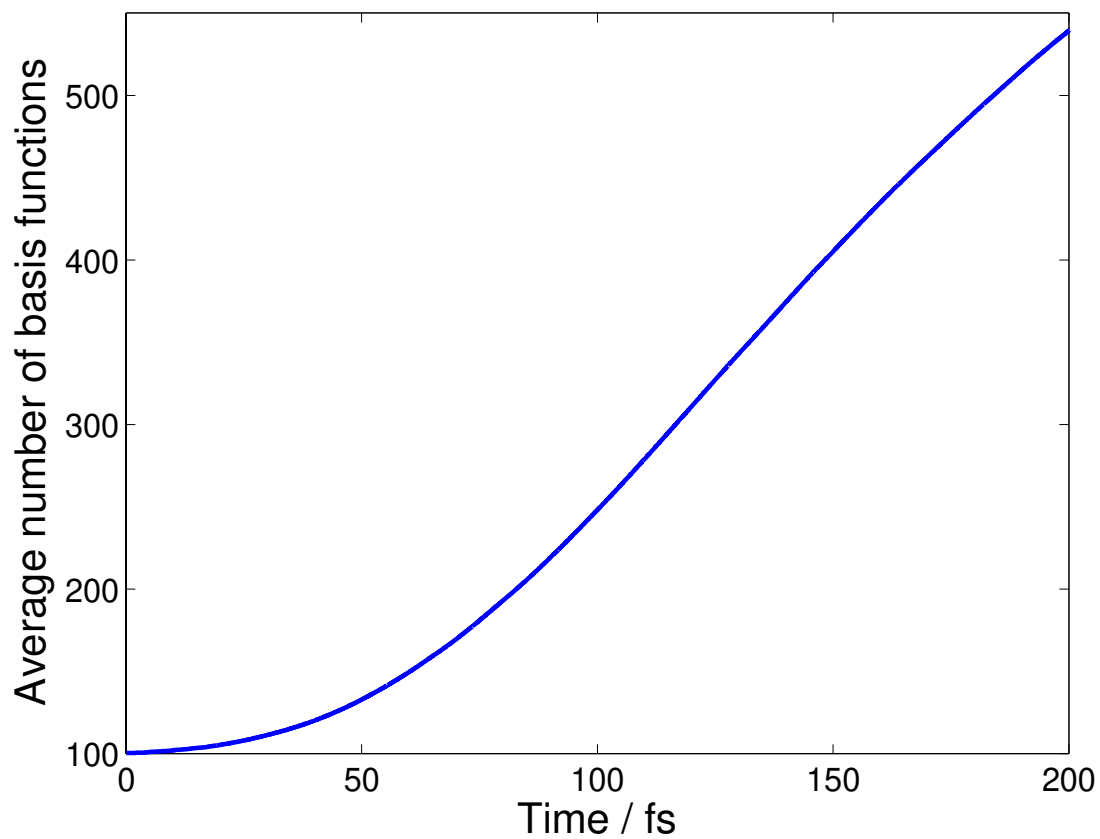**Figure 2.** Growth of the average number of TBFs in an AIMC calculation (with 200 initial conditions) using trains with N_{train}=100. On average, each TBF produces an additional 4 TBFs through cloning.
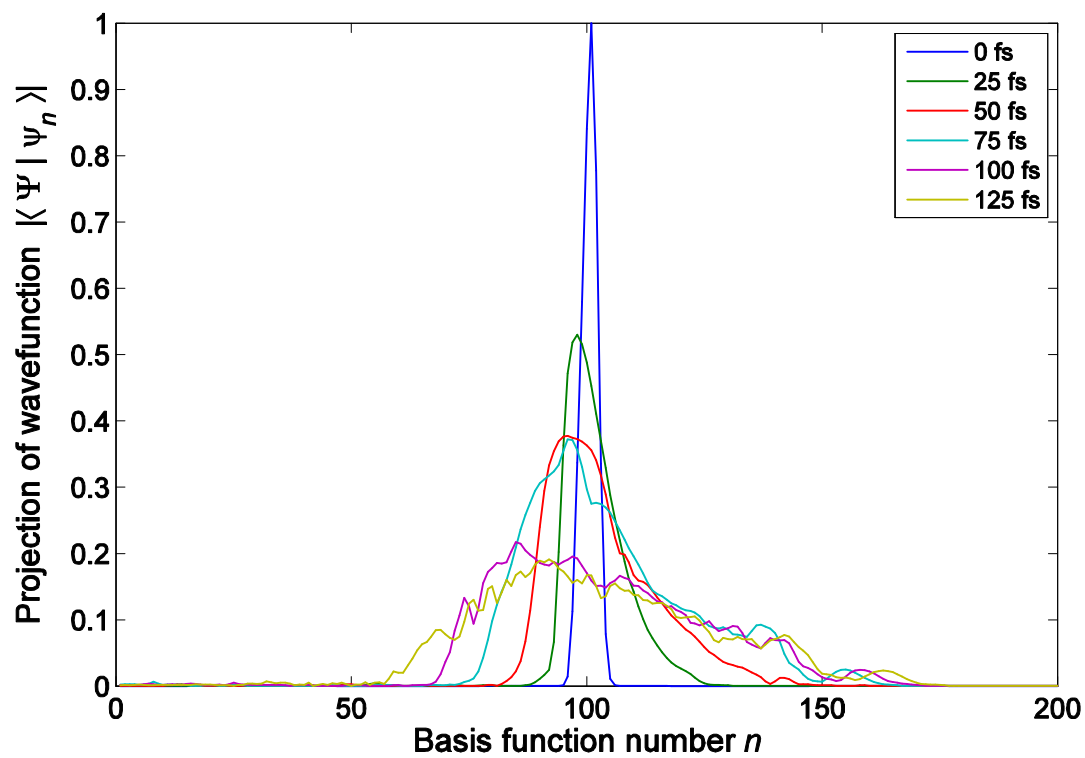
Figure 3. Time evolution of ethylene's wave function in a train basis set. The basis is moving along a quasi-classical trajectory so that the maximum of amplitude remains in the middle of the train.
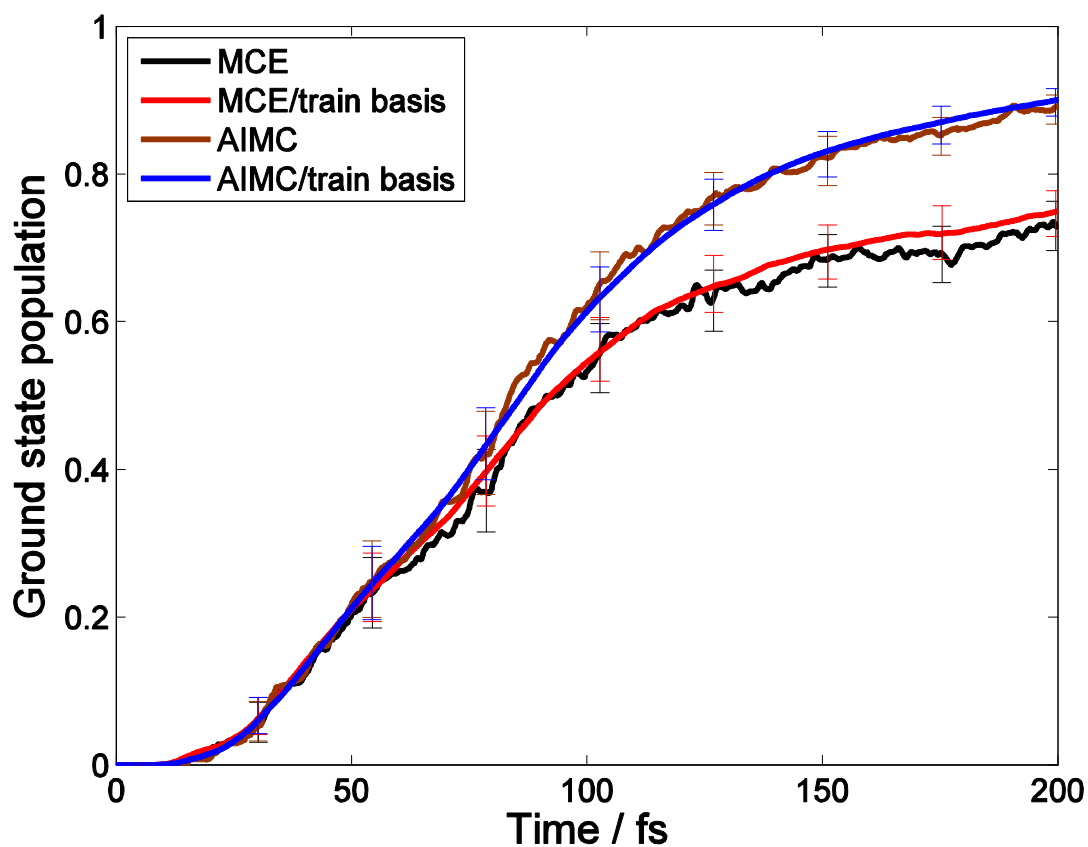
Figure 4. Ground-state population dynamics of ethylene following excitation to the spectroscopic ^1B_u ($\pi{\to}\pi^*$) state. . We compare calculations using Ehrenfest dynamics with a basis of one Gaussian function (MCE), coherent state trains (MCE/train, N_{train}=100), one Gaussian function with basis function cloning (AIMC), and coherent state trains with basis function cloning (AIMC/train, N_{train}=100). Results are averaged over 200 initial conditions and error bars (95% confidence interval) are shown.

Figure 5. Comparison of time evolution of ground state population after $\pi \rightarrow \pi^*$ photoexcitation from AIMS (red) and AIMC (using trains, N_{train}=100, blue) calculations with SA-3-CAS(2/2)/6-31G**. Results are averaged over 200 initial conditions and error bars (95% confidence interval) are shown.

Appendix 1

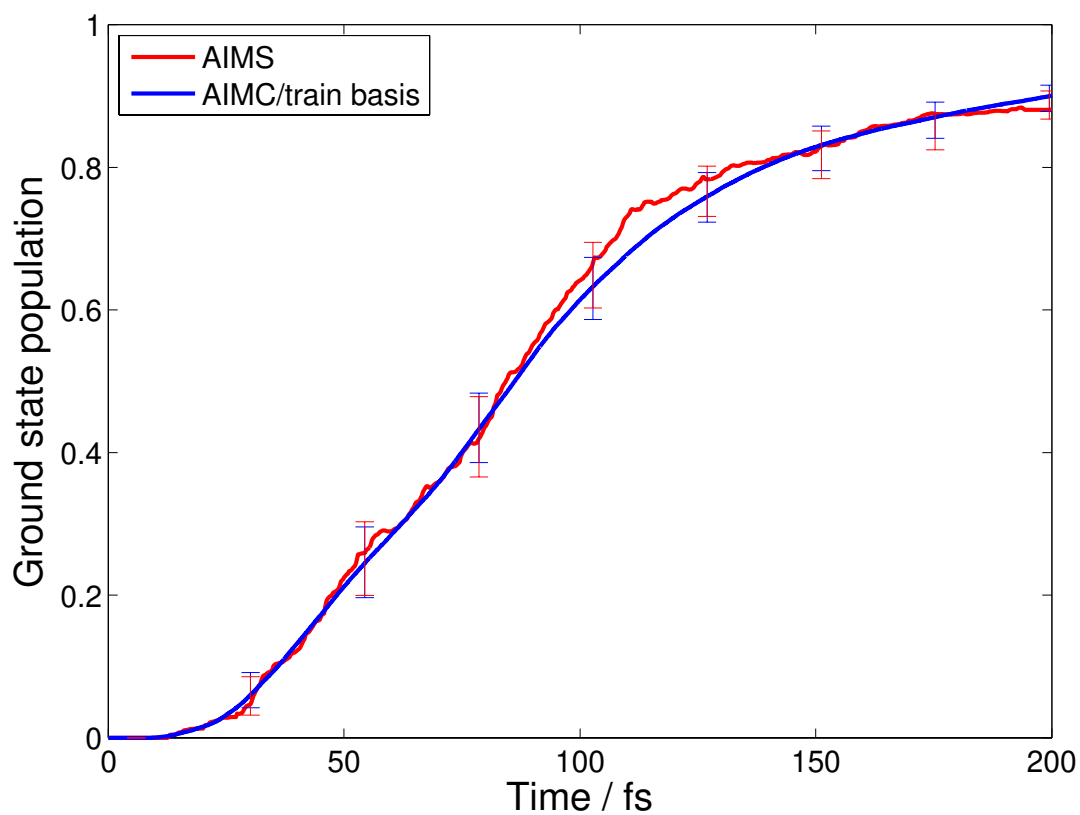The equations for integrals involving complex Gaussian basis functions, which appear in our formalism, are as follows:

$$\left\langle \chi_m \middle| \chi_n \right\rangle = \exp\left(-\frac{\alpha}{2}\Delta\mathbf{R}^2 - \frac{1}{8\alpha\hbar^2}\Delta\mathbf{P}^2 + \frac{i}{\hbar}\left(\bar{\mathbf{P}}_m\bar{\mathbf{R}}_m - \bar{\mathbf{P}}_n\bar{\mathbf{R}}_n + \bar{\mathbf{R}}\Delta\mathbf{P} \right) + \frac{i}{\hbar}\Delta\gamma \right) \tag{A1.1}$$

$$\left\langle \chi_m \middle| \mathbf{R} \middle| \chi_n \right\rangle = \left[\frac{i}{4\alpha\hbar}\Delta\mathbf{P} + \bar{\mathbf{R}} \right]\left\langle \chi_m \middle| \chi_n \right\rangle \tag{A1.2}$$

$$\left\langle \chi_m \middle| \nabla_R \middle| \chi_n \right\rangle = \left[-\frac{i}{\hbar}\bar{\mathbf{P}} - \alpha\Delta\mathbf{R} \right]\left\langle \chi_m \middle| \chi_n \right\rangle \tag{A1.3}$$

$$\left\langle \chi_m \middle| \nabla_R^{\,2} \middle| \chi_n \right\rangle = \left[\frac{i}{\hbar}2\alpha\Delta\mathbf{R}\bar{\mathbf{P}} - \alpha + \alpha^2\Delta\mathbf{R}^2 - \frac{\bar{\mathbf{P}}^2}{\hbar^2} \right]\left\langle \chi_m \middle| \chi_n \right\rangle \tag{A1.4}$$

$$\left\langle \chi_m \middle| \frac{d}{d\bar{\mathbf{R}}_n} \middle| \chi_n \right\rangle = \left[\frac{i}{\hbar}\bar{\mathbf{P}} + \alpha\Delta\mathbf{R} \right]\left\langle \chi_m \middle| \chi_n \right\rangle \tag{A1.5}$$

$$\left\langle \chi_m \middle| \frac{d}{d\bar{\mathbf{P}}_n} \middle| \chi_n \right\rangle = \left[-\frac{1}{4\alpha\hbar^2}\Delta\mathbf{P} - \frac{i}{2\hbar}\Delta\mathbf{R} \right]\left\langle \chi_m \middle| \chi_n \right\rangle \tag{A1.6}$$

where $\Delta\mathbf{R} = \bar{\mathbf{R}}_n - \bar{\mathbf{R}}_m$, $\Delta\mathbf{P} = \bar{\mathbf{P}}_n - \bar{\mathbf{P}}_m$, $\Delta\gamma = \gamma_n - \gamma_m$, $\bar{\mathbf{R}} = \dfrac{\bar{\mathbf{R}}_n + \bar{\mathbf{R}}_m}{2}$, $\bar{\mathbf{P}} = \dfrac{\bar{\mathbf{P}}_n + \bar{\mathbf{P}}_m}{2}$.

Appendix 2

Let us consider the ansatz of Eqs. 19-20 for the full wave function

$$\left| \Psi(t) \right\rangle = \sum_n c_n(t)\left| \tilde{\psi}_n(t) \right\rangle, \tag{A2.1}$$

where the basis functions are the Ehrenfest configurations composed, as before, of nuclear and electronic parts:

$$\left| \tilde{\psi}_n(t) \right\rangle = \left| \chi_n(t) \right\rangle \sum_I a_I^n(t)\left| \tilde{\varphi}_I^n \right\rangle, \tag{A2.2}$$

but unlike the standard ansatz in Eq. 2, the electronic part here is "attached" to the center of the nuclear wave packet:

$$\left| \tilde{\varphi}_I^n \right\rangle = \left| \varphi_I \left(\mathbf{r}; \overline{\mathbf{R}}_n(t) \right) \right\rangle. \tag{A2.3}$$

Substituting the ansatz (A2.1) into the Schrodinger equation $i\hbar \dfrac{\partial \Psi}{\partial t} = H\Psi$ one obtains:

$$i\hbar \sum_n \frac{\partial c_n}{\partial t} \left\langle \tilde{\psi}_m \middle| \tilde{\psi}_n \right\rangle = \sum_n \left\langle \tilde{\psi}_m \middle| \hat{H} \middle| \tilde{\psi}_n \right\rangle c_n - i\hbar \sum_n \left\langle \tilde{\psi}_m \middle| \frac{\partial}{\partial t} \middle| \tilde{\psi}_n \right\rangle c_n. \tag{A2.4}$$

The matrix elements are as follows:

$$\left\langle \tilde{\psi}_m \middle| \tilde{\psi}_n \right\rangle = \sum_{IJ} \left(a_I^m \right)^* a_J^n \left\langle \tilde{\varphi}_I^m \middle| \tilde{\varphi}_J^n \right\rangle \left\langle \chi_m \middle| \chi_n \right\rangle, \tag{A2.5}$$

$$\begin{aligned}
\left\langle \tilde{\psi}_m \middle| \frac{\partial}{\partial t} \middle| \tilde{\psi}_n \right\rangle &= \sum_{IJ} \left(a_I^m \right)^* \dot{a}_J^n \left\langle \tilde{\varphi}_I^m \middle| \tilde{\varphi}_J^n \right\rangle \left\langle \chi_m \middle| \chi_n \right\rangle \\
&+ \sum_{IJ} \left(a_I^m \right)^* a_J^n \left\langle \tilde{\varphi}_I^m \middle| \frac{\partial}{\partial t} \middle| \tilde{\varphi}_J^n \right\rangle \left\langle \chi_m \middle| \chi_n \right\rangle + \sum_{IJ} \left(a_I^m \right)^* a_J^n \left\langle \tilde{\varphi}_I^m \middle| \tilde{\varphi}_J^n \right\rangle \left\langle \chi_m \middle| \frac{\partial}{\partial t} \middle| \chi_n \right\rangle
\end{aligned} \tag{A2.6}$$

$$\begin{aligned}
\left\langle \tilde{\psi}_m \middle| \hat{H} \middle| \tilde{\psi}_n \right\rangle &= \sum_{IJ} \left(a_I^m \right)^* a_J^n \left\langle \chi_m \middle| \left\langle \tilde{\varphi}_I^m \middle| \left(\hat{H}_{el}(r,R) - \frac{\hbar^2}{2M} \nabla_R^2 \right) \middle| \tilde{\varphi}_J^n \right\rangle \middle| \chi_n \right\rangle = \\
&= \sum_{IJ} \left(a_I^m \right)^* a_J^n \left(\left\langle \chi_m \middle| \left\langle \tilde{\varphi}_I^m \middle| \hat{H}_{el}(r,R) \middle| \tilde{\varphi}_J^n \right\rangle \middle| \chi_n \right\rangle - T_{mn} \left\langle \tilde{\varphi}_I^m \middle| \tilde{\varphi}_J^n \right\rangle \right)
\end{aligned} \tag{A2.7}$$

where $T_{mn} = -\hbar^2/2M \left\langle \chi_m \middle| \nabla_R^2 \middle| \chi_n \right\rangle$ are the kinetic energy matrix elements.

Comparing to Eq. 14, Eq.(A2.6) contains an extra term which includes $\left\langle \tilde{\varphi}_I^m \middle| \frac{\partial}{\partial t} \middle| \tilde{\varphi}_J^n \right\rangle$; likewise, comparing to Eq. 16, Eq.(A2.7) does not have the term containing $\left\langle \varphi_N \middle| \nabla_R \middle| \varphi_M \right\rangle$ and $\left\langle \varphi_N \middle| \nabla_R^2 \middle| \varphi_M \right\rangle$. The difference is due to the fact that $\left| \varphi_I \right\rangle = \left| \varphi_I(\mathbf{r}; \mathbf{R}) \right\rangle$ is a function of $\mathbf{R}$ and, therefore, is acted upon by the nuclear kinetic energy operator. On the other hand, the function $\left| \tilde{\varphi}_I^n \right\rangle = \left| \tilde{\varphi}_I^n(\mathbf{r}; \overline{\mathbf{R}}_n) \right\rangle$ depends on the time dependent $\overline{\mathbf{R}}_n$ and therefore is subject to time differentiation.

Approximations must again be introduced in order to evaluate the matrix elements. As in the previous case, we assume that on the scale of the Gaussian width, the electronic wave

function depends weakly on $\overline{\mathbf{R}}_k$, the position of the nuclei. Then for pairs of basis functions with sufficient nuclear overlap we can write:

$$\left\langle \tilde{\varphi}_I^m \middle| \tilde{\varphi}_J^n \right\rangle \approx \delta_{IJ} , \tag{A2.8}$$

$$\left\langle \tilde{\varphi}_I^m \middle| \frac{\partial}{\partial t} \middle| \tilde{\varphi}_J^n \right\rangle \approx \frac{\left\langle \tilde{\varphi}_I^m \middle| \frac{\partial}{\partial t} \middle| \tilde{\varphi}_I^m \right\rangle + \left\langle \tilde{\varphi}_J^n \middle| \frac{\partial}{\partial t} \middle| \tilde{\varphi}_J^n \right\rangle}{2} = \frac{\dot{\mathbf{R}}_m \mathbf{d}_{IJ}\left(\overline{\mathbf{R}}_m\right) + \dot{\mathbf{R}}_n \mathbf{d}_{IJ}\left(\overline{\mathbf{R}}_n\right)}{2} , \tag{A2.9}$$

i.e. the nonadiabatic coupling matrix element between two wave functions localized at points $\overline{\mathbf{R}}_m$ and $\overline{\mathbf{R}}_n$ is an average between the two "local" nonadiabatic couplings. For $m=n$ it is not an approximation.

We assume the following approximation for the Born-Oppenheimer potential energy matrix elements:

$$\left\langle \chi_m \middle| \left\langle \tilde{\varphi}_I^m \middle| \hat{H}_{el}(\mathbf{r},\mathbf{R}) \middle| \tilde{\varphi}_j^n \right\rangle \middle| \chi_n \right\rangle \approx$$
$$\left\langle \chi_m \middle| \left\langle \varphi_I^m(\mathbf{r};\mathbf{R}) + (\mathbf{R}_m - \mathbf{R})\nabla_\mathbf{R}\varphi_I^m(\mathbf{r};\mathbf{R}) \middle| \hat{H}_{el}(\mathbf{r},\mathbf{R}) \right.$$
$$\left. \middle| \varphi_J^n(\mathbf{r};\mathbf{R}) + (\mathbf{R}_n - \mathbf{R})\nabla_\mathbf{R}\varphi_J^n(\mathbf{r};\mathbf{R}) \right\rangle \middle| \chi_n \right\rangle \approx$$
$$\left\langle \chi_m \middle| \left\langle \varphi_I^m(\mathbf{r},\mathbf{R}) \middle| \hat{H}_{el}(\mathbf{r},\mathbf{R}) \middle| \varphi_I^n(\mathbf{r},\mathbf{R}) \right\rangle \middle| \chi_n \right\rangle +$$
$$\left\langle \chi_m \middle| \left\langle \varphi_J^m(\mathbf{r},\mathbf{R}) \middle| \hat{H}_{el}(\mathbf{r},\mathbf{R}) \middle| (\mathbf{R}_n - \mathbf{R})\nabla_\mathbf{R}\varphi_J^n(\mathbf{r},\mathbf{R}) \right\rangle \middle| \chi_n \right\rangle +$$
$$\left\langle \chi_m \middle| \left\langle (\mathbf{R}_m - \mathbf{R})\nabla_\mathbf{R}\varphi_I^m(\mathbf{r},\mathbf{R}) \middle| \hat{H}_{el}(\mathbf{r},\mathbf{R}) \middle| \varphi_J^n(\mathbf{r},\mathbf{R}) \right\rangle \middle| \chi_n \right\rangle = \tag{A2.10}$$
$$\delta_{IJ}\left\langle \chi_m \middle| V_I(\mathbf{R}) \middle| \chi_n \right\rangle +$$
$$\left\langle \chi_m \middle| V_J(\mathbf{R})(\mathbf{R}_n - \mathbf{R})\left\langle \varphi_I^m(\mathbf{r},\mathbf{R}) \middle| \nabla_\mathbf{R}\varphi_J^n(\mathbf{r},\mathbf{R}) \right\rangle \middle| \chi_n \right\rangle +$$
$$\left\langle \chi_m \middle| V_I(\mathbf{R})(\mathbf{R}_m - \mathbf{R})\left\langle \nabla_\mathbf{R}\varphi_I^m(\mathbf{r},\mathbf{R}) \middle| \varphi_J^n(\mathbf{r},\mathbf{R}) \right\rangle \middle| \chi_n \right\rangle$$

The last two terms can be disregarded because far from intersection the nonadiabatic coupling terms are small and near the intersection where $V_I \approx V_J$ the two terms cancel out. Then using the BAT approximation of Eq. 17 for the first term in (A2.10) we get:

$$\langle \chi_m | \langle \tilde{\phi}_I^{(m)} | \hat{H}_{el}(\nabla_r, \mathbf{r}, \mathbf{R}) | \tilde{\phi}_J^{(n)} \rangle | \chi_n \rangle \approx \delta_{IJ} \langle \chi_m | \chi_n \rangle \left(\frac{V_I(\overline{\mathbf{R}}_m) + V_I(\overline{\mathbf{R}}_n)}{2} \right)$$

$$+ \delta_{IJ} \left(\frac{\langle \chi_m | (\mathbf{R} - \overline{\mathbf{R}}_m) \chi_n \rangle \nabla_{\mathbf{R}} V_I(\overline{\mathbf{R}}_m) + \langle \chi_m | (\mathbf{R} - \overline{\mathbf{R}}_n) \chi_n \rangle \nabla_{\mathbf{R}} V_I(\overline{\mathbf{R}}_n)}{2} \right). \quad \text{(A2.11)}$$

As before

$$\langle \chi_m | \frac{\partial}{\partial t} | \chi_n \rangle = \left(\dot{\overline{\mathbf{R}}}_n \langle \chi_m | \frac{\partial}{\partial \overline{\mathbf{R}}_n} | \chi_n \rangle + \dot{\overline{\mathbf{P}}}_n \langle \chi_m | \frac{\partial}{\partial \overline{\mathbf{P}}_n} | \chi_n \rangle \right). \quad \text{(A2.12)}$$

Altogether, under the same assumption of smoothness of the electronic wave function and its weak dependence on the nuclear coordinates, the equations for the amplitudes c_n become the same as for those in the ansatz of Eqs. 2-3. The only difference is that the nonadiabatic coupling

term $\sum_{IJ} \langle a_I^m \chi_m | a_J^n \chi_m \rangle \dfrac{\dot{\overline{\mathbf{R}}}_m \mathbf{d}_{IJ}(\overline{\mathbf{R}}_m) + \dot{\overline{\mathbf{R}}}_n \mathbf{d}_{IJ}(\overline{\mathbf{R}}_n)}{2}$ originates now from the time dependence of the

electronic basis functions but not the kinetic energy operator. Although currently both forms of the total wave function given by Eqs. 2-3 or Eqs. 19-20 obey the same equation, it is not clear which of the two approaches will be more convenient for generalization in future, and for the sake of future reference we present here both of them.

References

1. Cheng, Y.-C.; Fleming, G. R. Dynamics of Light Harvesting in Photosynthesis. *Ann. Rev. Phys. Chem.* **2009**, *60*, 241.

2. Blanchet, V.; Zgierski, M. Z.; Seideman, T.; Stolow, A. Discerning vibronic molecular dynamics using time-resolved photoelectron spectroscopy. *Nature* **1999**, *401*, 52.

3. Collini, E.; Wong, C. Y.; Wilk, K. E.; Curmi, P. M. G.; Brumer, P.; Scholes, G. D. Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature. *Nature* **2010**, *463*, 644.

4. Herek, J. L.; Wohlleben, W.; Cogdell, R. J.; Zeidler, D.; Motzkus, M. Quantum control of energy flow in light harvesting. *Nature* **2002**, *417*, 533.

5. Kuroda, D. G.; Singh, C. P.; Peng, Z.; Kleiman, V. D. Mapping Excited-State Dynamics by Coherent Control of a Dendrimers Photoemission Efficiency. *Science* **2009**, *326*, 263.

6. Yarkony, D. R. Diabolical conical intersections. *Rev. Mod. Phys.* **1996**, *68*, 985.

7. Lasorne, B.; Worth, G. A.; Robb, M. A. Excited-state dynamics. *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2011**, *1*, 460.

8. Car, R.; Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **1985**, *55*, 2471.

9. Tuckerman, M. E.; Ungar, P. J.; von Rosenvinge, T.; Klein, M. L. Ab initio molecular dynamics simulations. *J. Phys. Chem.* **1996**, *100*, 12878.

10. Ben-Nun, M.; Martínez, T. J. Ab Initio Quantum Molecular Dynamics. *Adv. Chem. Phys.* **2002**, *121*, 439.

11. Meyer, H.-D.; Miller, W. H. A classical analog for electronic degrees of freedom in nonadiabatic collision processes. *J. Chem. Phys.* **1979**, *70*, 3214.

12. Billing, G. D. On the use of Ehrenfest's theorem in molecular scattering. *Chem. Phys. Lett.* **1983**, *100*, 535.

13. Negele, J. W. The mean-field theory of nuclear structure and dynamics. *Rev. Mod. Phys.* **1982**, *54*, 913.

14. Tully, J. C.; Preston, R. K. Trajectory surface hopping approach to nonadiabatic molecular collisions: The reaction of H+ with D2. *J. Chem. Phys.* **1971**, *55*, 562.

15. Tully, J. C. Molecular dynamics with electronic transitions. *J. Chem. Phys.* **1990**, *93*, 1061.

16. Parandekar, P. V.; Tully, J. C. Mixed quantum-classical equilibrium. *J. Chem. Phys.* **2005**, *122*, 094102.

17. Parandekar, P. V.; Tully, J. C. Detailed Balance in Ehrenfest Mixed Quantum-Classical Dynamics. *J. Chem. Theo. Comp.* **2006**, *2*, 229.

18. Schmidt, J. R.; Parandekar, P. V.; Tully, J. C. Mixed quantum-classical equilibrium: Surface-hopping. *J. Chem. Phys.* **2008**, *129*, 044104.

19. Martinez, T. J.; Levine, R. D. First-principles molecular dynamics on multiple electronic states: A case study of NaI. *J. Chem. Phys.* **1996**, *105*, 6334.

20. Martínez, T. J.; Levine, R. D. Dynamics of the collisional electron transfer and femtosecond photodissociation of NaI on ab initio electronic energy curves. *Chem. Phys. Lett.* **1996**, *259*, 252.

21. Martinez, T. J. Ab initio molecular dynamics around a conical intersection: Li(2p)+H2. *Chem. Phys. Lett.* **1997**, *272*, 139.

22. Ben-Nun, M.; Martinez, T. J. Ab Initio Molecular Dynamics Study of cis-trans Photisomerization in Ethylene. *Chem. Phys. Lett.* **1998**, *298*, 57.

23. Martinez, T. J.; Ben-Nun, M.; Ashkenazi, G. Classical/quantal method for multistate dynamics: A computational study. *J. Chem. Phys.* **1996**, *104*, 2847.

24. Martinez, T. J.; Ben-Nun, M.; Levine, R. D. Multi-Electronic-State Molecular Dynamics: A Wave Function Approach with Applications. *J. Phys. Chem.* **1996**, *100*, 7884.

25. Heller, E. J. Time-dependent approach to semiclassical dynamics. *J. Chem. Phys.* **1975**, *62*, 1544.

26. Heller, E. J. Frozen Gaussians: A very simple semiclassical approximation. *J. Chem. Phys.* **1981**, *75*, 2923.

27. Ben-Nun, M.; Martinez, T. J. Nonadiabatic Molecular Dynamics: Validation of the multiple spawning method for a multidimensional problem. *J. Chem. Phys.* **1998**, *108*, 7244.

28. Ben-Nun, M.; Martinez, T. J. A Continuous Spawning Method for Nonadiabatic Dynamics and Validation for the Zero-Temperature Spin-Boson Problem. *Isr. J. Chem.* **2007**, *47*, 75.

29. Shalashilin, D. V.; Child, M. S. Time dependent propagation in phase space. *J. Chem. Phys.* **2000**, *113*, 10028.

30. Shalashilin, D. V.; Child, M. S. The phase space CCS approach to quantum and semiclassical molecular dynamics for high-dimensional systems. *Chem. Phys.* **2004**, *304*, 103.

31. Burghardt, I.; Meyer, H. D.; Cederbaum, L. S. Approaches to the approximate treatment of complex molecular systems by the multiconfiguration time-dependent Hartree method. *J. Chem. Phys.* **1999**, *111*, 2927.

32. Worth, G. A.; Burghardt, I. Full quantum mechanical molecular dynamics using Gaussian wavepackets. *Chem. Phys. Lett.* **2003**, *368*, 502.

33. Worth, G. A.; Robb, M. A.; Lasorne, B. Solving the time-dependent Schrodinger equation for nuclear motion in one step: direct dynamics of non-adiabatic systems. *Mol. Phys.* **2008**, *106*, 2077.

34. Ben-Nun, M.; Martinez, T. J. Photodynamics of ethylene: ab initio studies of conical intersections. *Chem. Phys.* **2000**, *259*, 237.

35. Yang, S.; Coe, J. D.; Kaduk, B.; Martinez, T. J. An ``optimal'' spawning algorithm for adaptive basis set expansion in nonadiabatic dynamics. *J. Chem. Phys.* **2009**, *130*, 134113.

36. Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martiinez, T. J. Implementation of ab initio multiple spawning in the Molpro quantum chemistry package. *Chem. Phys.* **2008**, *347*, 3.

37. Saita, K.; Shalashilin, D. V. On-the-fly ab initio molecular dynamics with multiconfigurational Ehrenfest method. *J. Chem. Phys.* **2012**, *137*, 8.

38. Shalashilin, D. V. Multiconfigurational Ehrenfest approach to quantum coherent dynamics in large molecular systems. *Faraday Disc.* **2011**, *153*, 105.

39. Shalashilin, D. V. Nonadiabatic dynamics with the help of multiconfigurational Ehrenfest method: Improved theory and fully quantum 24D simulation of pyrazine. *J. Chem. Phys.* **2010**, *132*.

40. Shalashilin, D. V. Quantum mechanics with the basis set guided by Ehrenfest trajectories: Theory and application to spin-boson model. *J. Chem. Phys.* **2009**, *130*.

41. Bittner, E. R.; Rossky, P. J. Quantum decoherence in mixed quantum-classical systems: Nonadiabatic processes. *J. Chem. Phys.* **1995**, *103*, 8130.

42. Prezhdo, O. V.; Rossky, P. J. Mean-field molecular dynamics with surface hopping. *J. Chem. Phys.* **1997**, *107*, 825.

43. Ben-Nun, M.; Martinez, T. J. Exploiting temporal nonlocality to remove scaling bottlenecks in nonadiabatic quantum dynamics. *J. Chem. Phys.* **1999**, *110*, 4134.

44. Shalashilin, D. V.; Child, M. S. Basis set sampling in the method of coupled coherent states: Coherent state swarms, trains and pancakes. *J. Chem. Phys.* **2008**, *128*, 054102.

45. Ronto, M.; Shalashilin, D. V. Numerical Implementation and Test of the Modified Variational Multiconfigurational Gaussian Method for High-Dimensional Quantum Dynamics. *J. Phys. Chem. A*.

46. Doltsinis, N. L.; Marx, D. First Principles Molecular Dynamics Involving Excited States and Nonadiabatic Transitions. *J. Theo. Comp. Chem.* **2002**, *01*, 319.

47. Roos, B. O. The Complete Active Space Self-Consistent Field Method and Its Applications in Electronic Structure Calculations. *Adv. Chem. Phys.* **1987**, *69*, 399.

48. Lengsfield, B. H., III; Yarkony, D. R. Nonadiabatic interactions between potential energy surfaces: Theory and applications. *Adv. Chem. Phys.* **1992**, *82*, 1.

49. Roos, B. O. Theoretical Studies of Electronically Excited States Using Multiconfigurational Perturbation Theory. *Acc. Chem. Res.* **1999**, *32*, 137.

50. Mori, T.; Glover, W. J.; Schuurman, M. S.; Martinez, T. J. Role of Rydberg States in the Photochemical Dynamics of Ethylene. *J. Phys. Chem. A* **2012**, *116*, 2808.

51. Tao, H.; Levine, B. G.; Martinez, T. J. Ab initio multiple spawning dynamics using multi-state second-order perturbation theory. *J. Phys. Chem. A* **2009**, *113*, 13656.

52. Tao, H.; Allison, T. K.; Wright, T. W.; Stooke, A. M.; Khurmi, C.; van Tilborg, J.; Liu, Y.; Falcone, R. W.; Belkacem, A.; Martinez, T. J. Ultrafast internal conversion in ethylene. I. The excited state lifetime. *J. Chem. Phys.* **2011**, *134*, 244306.

53. Sension, R. J.; Hudson, B. S. Vacuum ultraviolet resonance Raman studies of the excited electronic states of ethylene. *J. Chem. Phys.* **1989**, *90*, 1377.

54. Cromwell, E. F.; Stolow, A.; Vrakking, M. J. J.; Lee, Y. T. Dynamics of ethylene photodissociation from rovibrational and translational energy distributions of H2 products. *J. Chem. Phys.* **1992**, *97*, 4029.

55. Stert, V.; Lippert, H.; Ritze, H. H.; Radloff, W. Femtosecond time-resolved dynamics of the electronically excited ethylene molecule. *Chem. Phys. Lett.* **2004**, *388*, 144.

56. Viel, A.; Krawczyk, R. P.; Manthe, U.; Domcke, W. Photinduced dynamics of ethene in the N, V, and Z valence states: A six-dimensional nonadiabatic quantum dynamics investigation. *J. Chem. Phys.* **2004**, *120*, 11000.

57. Viel, A.; Krawczyk, R. P.; Manthe, U.; Domcke, W. The sudden-polarization effect and its role in the ultrafast photochemistry of ethene. *Ang. Chem. Int. Ed.* **2003**, *42*, 3434.

58. Kosma, K.; Trushin, S. A.; Fuss, W.; Schmid, W. E. Ultrafast dynamics and coherent oscillations in ethylene and ethylene-d(4) excited at 162nm. *J. Phys. Chem. A* **2008**, *112*, 7514.

59. Thompson, A. L.; Punwong, C.; Martinez, T. J. Optimization of width parameters for quantum dynamics with frozen Gaussian basis sets. *Chem. Phys.* **2010**, *370*, 70.

60. Ben-Nun, M.; Martinez, T. J. Electronic absorption and resonance Raman spectroscopy from Ab Initio Quantum Molecular Dynamics. *J. Phys. Chem. A* **1999**, *103*, 10517.

61. Sulc, M.; Hernandez, H.; Martinez, T. J.; Vanicek, J. Relation of exact Gaussian basis methods to the dephasing representation: Theory and application to time-resolved electronic spectra. *J. Chem. Phys.* **2013**, *139*, 034112.