Trajectory Surface Hopping for a Polarizable Embedding QM/MM

Mattia Bondanza,† Baptiste Demoulin,‡ Filippo Lipparini,† Mario Barbatti,*‡¶ and Benedetta Mennucci*,†

†Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via G. Moruzzi 13, 56124 Pisa, Italy
‡Aix Marseille University, CNRS, ICR, Marseille, France
¶Institut Universitaire de France, 75231 Paris, France

E-mail: mario.barbatti@univ-amu.fr; benedetta.mennucci@unipi.it
Abstract

We present the implementation of trajectory surface hopping nonadiabatic dynamics for a polarizable embedding QM/MM formulation. The time-dependent density functional theory has been used as the quantum mechanical level of theory, whereas the molecular mechanics description has involved the polarizable AMOEBA force field. This implementation has been obtained by integrating the surface-hopping program Newton-X NS with an interface between the Gaussian 16 and Tinker suites of codes to calculate QM/AMOEBA energies and forces. The implementation has been tested on a photoinduced electron-driven proton-transfer reaction involving pyrimidine and a hydrogen-bonded water surrounded by a small cluster of water molecules and within a large water droplet.

1 Introduction

Computational photochemistry is a mature research field, but its application to complex systems remains challenging, especially when tackled with theoretically accurate methods. The problem is, in fact, twofold: from one side, the essence of photochemistry is connected to the quantum behavior of nuclear wave packets, which is something challenging to reconstruct computationally; on the other side, the motion of such wave packets is determined by the electronic structure of the molecule itself, which can be accurately computed only with techniques having a very unfavorable scaling. Finally, both the electronic and the nuclear degrees of freedom of the system undergoing the photochemical process are coupled to the embedding environment, which can enormously enhance both the difficulty of the calculations and their computational cost.

An effective strategy to make the computational simulation feasible without losing too much accuracy is introducing a mixed quantum-classical description for modeling both dynamics and potential energy surfaces (PES). The mixed quantum-classical reformulation of dynamics allows introducing the concept of nuclear trajectories in a way that closely resembles Newton’s dynamics, thus drastically reducing the computational effort needed to simulate a wave packet propagation. On the other hand, the mixed quantum-classical calculation of energy, forces, and state couplings
allows one to keep an accurate quantum mechanical (QM) description for the most relevant part of the system while introducing the effect of the rest through classical interactions.

A successful example of this strategy is the coupling between the trajectory surface hopping (TSH)\textsuperscript{5} nonadiabatic method with a hybrid QM/MM (quantum mechanics/molecular mechanics) description.\textsuperscript{6–9} This approach has often been applied to solvated systems and molecular systems embedded in more complex environments such as biological macrostructures or solid matrices.\textsuperscript{10,11} In those applications, the usual QM/MM description adopted is an electrostatic type of embedding where the effect of the MM atoms on the QM subsystem is represented in terms of the electrostatic potential generated by the fixed atomic charges used in MM force fields to describe electrostatic interactions.\textsuperscript{12} In the last years, however, polarizable QM/MM formulations have rapidly diffused,\textsuperscript{13–20} especially for simulating light-induced processes of (multi)chromophoric systems.\textsuperscript{21–23}

Extending polarizable models to nonadiabatic dynamics presents an inherent challenge due to the non-linearity introduced by the polarizable environment into the system’s Hamiltonian. In fact, in a polarizable description, the interaction with the environment is proportional to the polarization degrees of freedom of the classical environment, which in turn depend linearly on the QM density of charge.\textsuperscript{22–24} Consequently, in a state-specific (SS) formulation, where the environment responds to the density of a given electronic state, each state is the eigenfunction of a different non-linear Hamiltonian, which leads to electronic states that are no longer orthogonal. While this non-linearity is a consequence of the physics of the system (the polarization mimics the electronic density of the environment), the practical realization of TSH using such a state-specific environment model faces some major difficulties.\textsuperscript{22} In particular, some quantities essential for TSH are considerably more difficult to compute on a non-orthogonal basis. Furthermore, having to deal with non-orthogonal states blurs the distinction between the states themselves, which can introduce difficulties in the interpretation of the results.

However, if one adopts a linear-response (LR) QM formulation, of which time-dependent density functional theory (TD-DFT) represents the most well-known and successful example, the cou-
pling to a polarizable description becomes much simpler. Within this theoretical framework, the response of the polarizable environment can be recast in such a way that it does not depend on a specific state but only on transition properties, thus making the computation analogous to a calculation for an isolated QM system. In principle, as the LR and the SS responses describe two different physical interactions, they should be simultaneously taken into account to get a complete description of the environment effect. Still missing within an LR QM method, the SS correction can be approximately recovered. One of the most common strategies is the so-called corrected linear-response method (cLR), initially developed for continuum solvation models, but successively extended to polarizable QM/MM descriptions.

Here we investigate the possibilities TSH offers when combined with TD-DFT and an induced point dipole (IDP) formulation of the polarizable embedding by coupling the surface-hopping engine Newton-X with the polarizable QM/AMOEBA approach developed interfacing the Tinker and Gaussian suites of codes.

We tested this implementation on the photodynamics of pyrimidine-water clusters and a watersolvated pyrimidine droplet. These systems are tailored for these tests because, first, we can count on previous experimental and theoretical benchmark results; second, their ultrafast dynamics and the chromophore’s small size limit the computational costs, allowing the simulation of multiple data sets.

2 Methods and Implementation

2.1 Ground and excited states with QM/AMOEBA

In the following, we assume a Kohn-Sham (KS) density functional theory (DFT) description of the molecular system within an environment described with the AMOEBA polarizable force field. The QM/AMOEBA self-consistent polarization problem can be derived starting from the following
polarization Lagrangian\textsuperscript{24}

$$L(P, \mu_d, \mu_p) = E^\text{QM}(P) + \varepsilon^\text{self}(M) + \langle q, V(P) \rangle - \langle \mu_s, E(P) \rangle + \langle \Theta, G(P) \rangle$$

$$- \frac{1}{2} \langle \mu_d, E(P) + E_p(M) \rangle + \frac{1}{2} \langle \mu_p, T \mu_d - E(P) - E_d(M) \rangle \quad (1)$$

In eq. 1, the brackets denote a dot product; $M$ is the AMOEBA multipolar distribution of static charges $q$, dipoles $\mu_s$, and quadrupoles $\Theta$; $V$, $E$, and $G$ are the electrostatic potential, field, and field gradient whereas $T$ is the dipole interaction tensor. $E_p(M)$ and $E_d(M)$ are fields produced by the multipolar distribution at the polarizable sites. They differ among themselves because of different exclusion rules used to assemble them, as from the definition of the AMOEBA force field.\textsuperscript{35,36} As a result, two sets of induced dipoles are generated $\mu_d$ and $\mu_p$.

By imposing the stationarity of the Lagrangian with respect to the density matrix, subject to the usual idempotency constraints, to the induced dipoles $\mu_d$ and $\mu_p$, we get an effective Kohn-Sham equation and the following equations for the induced dipoles:

$$T \mu_d = E_d(M) + E(P)$$

$$T \mu_p = E_p(M) + E(P). \quad (2)$$

In the effective Kohn-Sham equation, the operator contains additional terms due to the presence of the distribution $M$ and a polarization term which can be recast in the following form

$$- \frac{1}{2} \langle \mu_d + \mu_p, E_{\mu \nu} \rangle, \quad (3)$$

where $E_{\mu \nu}$ are electric field one-electron integrals. As the induced dipoles depend on the QM density matrix, the QM and polarization equations are coupled, and by solving them together, mutual polarization of the QM density and AMOEBA-induced dipoles is achieved. From a numerical point of view, this non-linearity is not an issue for standard self-consistent field (SCF) implementations, as they already self-consistently solve a non-linear eigenvalue problem. When a polarizable
embedding is added, one has to modify the SCF algorithm by including, at each SCF iteration, the
calculation of the induced dipoles so that one can add the polarization contribution to the effective
Kohn-Sham matrix. Thus, polarizable embedding models are considerably more computationally demanding than electrostatic embedding QM/MM. In our implementation, the polarization linear systems in eq. 2 are solved iteratively, using a preconditioned conjugate gradient strategy and the Fast Multipole Method (FMM) to compute the required matrix-vector products in a linear-scaling fashion.

In an LR TD-DFT scheme, excitation frequencies \( \omega \) are obtained by solving the generalized eigenvalue problem,

\[
\begin{pmatrix}
A & B \\
B^* & A^*
\end{pmatrix}
\begin{pmatrix}
X \\
Y
\end{pmatrix}
= \omega
\begin{pmatrix}
1 & 0 \\
0 & -1
\end{pmatrix}
\begin{pmatrix}
X \\
Y
\end{pmatrix},
\]

where the \( A \) and \( B \) matrices are components of the orbital rotation Hessian, and \( X \) and \( Y \) are the off-diagonal blocks of the transition density matrix. It is important to underline that solving eq. 4 does not provide excited-state densities. Nevertheless, they can be reconstructed by first solving a set of coupled-perturbed equations, usually called Z-vector equations.

The effect of a polarizable environment on the description of excitation processes is far from trivial. The self-consistent polarization interaction gives rise to an additional term in the definition of the \( A \) and \( B \) matrices, thus modifying the molecular response function. In particular, both matrices are augmented by

\[
\gamma^\text{pol}_{iajb} = - \sum_k \langle \phi_i | \hat{\mathbf{e}}(\mathbf{r}_k) | \phi_a \rangle \mu_k(\phi_j, \phi_b),
\]

where \( \hat{\mathbf{e}}(\mathbf{R}) \) is the electric field operator at \( \mathbf{R} \) and \( \mu_k(\phi_j, \phi_b) \) is the dipole at site \( k \) induced by the density element \( \phi_j \phi_b \). At convergence, this LR environment term gives rise to an excitation energy contribution determined by the electric field generated by the transition density \( (X + Y) \) at the polarizable sites, and the AMOEBA dipoles induced by the same field, namely \( -\langle \mu^{X+Y} | E^{X+Y} \rangle \).
Such LR term can be interpreted as the instantaneous response of the AMOEBA polarizable sites to the transition density of the electronic excitation and, in the literature, has also been described as a dispersion-like interaction.\textsuperscript{25}

The same computational considerations made for the SCF apply here as well. In particular, a polarization linear system needs to be solved at each iteration of the solution of the TD-DFT equations, which is again done using our FMM-based implementation.\textsuperscript{40,44} From eq. 5, we can see that, in the LR formulation, the environment’s dipoles are not relaxed on the excited state density, but for each state, their response to the state transition density is computed within the solution of the TD-DFT equations. This formulation makes the calculation of analytical gradients feasible using precisely the same approach used for an isolated QM system as described in detail in ref. 44.

As reported in the Introduction, it is possible to recover the additional type of response due to the polarizable environment, namely the one due to the change in the electronic charge densities of the ground and excited states (the SS effect) through the corrected LR (cLR). Within AMOEBA, the cLR correction to the excitation energy reads:\textsuperscript{17}

\[ \Delta \omega^{cLR} = -\frac{1}{2} \langle \mu^\Lambda, E^\Delta \rangle, \quad (6) \]

where \( E_k^\Delta \) is the electric field generated by the change in the electronic charge densities of the ground and excited states at the polarizable site \( k \) and \( \mu^\Lambda \) are the induced dipoles generated by the same field.

To simultaneously take into account LR and SS responses, a simple but effective protocol has been proposed for continuum solvation models and called cLR\textsuperscript{2,27} but it can equivalently be applied to a QM/AMOEBA description. The only problem is that analytical gradients of the corrected energies are no longer possible; for this reason, in the following, only an LR formulation will be used in the TSH dynamics. SS effects will be exclusively considered as \textit{a posteriori} correction of the LR trajectories, precisely as proposed in the original cLR\textsuperscript{2} protocol.\textsuperscript{27}
2.2 Nonadiabatic dynamics with QM/AMOEBA

Figure 1: Schematic representation of Newton-X-Tinker-Gaussian implementation. Newton-X acts as a driver providing at the beginning of each step the coordinates of the system to Tinker, that in turn calls Gaussian to compute energies and gradients with TD-DFT/AMOEBA PES. Then, to evaluate the transition probabilities between states, the overlap of wave functions at two subsequent time steps are computed by a program written on purpose (ci ovlp in the scheme, it is provided together with Newton-X) that uses the orbital overlap matrix computed using Gaussian.

In the present work, we focus on the coupling of TSH with LR-TD-DFT performed with polarizable embedding in the IPD formulation. To implement such a methodology, we developed an interface between the TSH software Newton-X and a modified version of Tinker that can, together with a modified version of the Gaussian 16 suite of codes, compute energies and geometrical gradients of the TD-DFT/AMOEBA Hamiltonian.

In our implementation (Figure 1), Newton-X NS (new series) handles the time propagation and the surface hopping algorithm. At each step of the dynamics, it calls the Tinker/Gaussian interface to compute the required electronic-structure quantities: (a) the potential energies, (b) the potential energy gradients, (c) the single-excitation coefficients (corresponding to the elements of the X + Y matrix), and (d) the molecular to atomic orbitals transformation matrix. The latter two are required to evaluate the overlap between Casida’s wavefunctions, used to estimate time-derivative nonadiabatic couplings.45,46
In particular, auxiliary wavefunctions for each electronic excited state \( I \) are written as

\[
\lvert \Psi^I \rangle = N_I^{1/2} \sum_{ia} (X + Y)^I_{ia} \lvert \Theta^a \rangle,
\]

(7)

where \( N_I = \langle (X + Y) \lvert (X + Y) \rangle \) is the normalization factor and \( \lvert \Theta^a \rangle \) is a Slater determinant with an electron promoted from Kohn-Sham spin-orbital orbital \( i \) to \( a \). With this wavefunction definition, we can compute the time-derivative nonadiabatic coupling,

\[
\sigma_{IJ}(t) = \left\langle \Psi^I \lvert \frac{\partial \Psi^J}{\partial t} \right\rangle, \quad I \neq J,
\]

(8)

employing the Hammes-Schiffer and Tully (HST) approach

\[
\sigma_{IJ}(t) \approx \frac{1}{2\Delta t} \left[ S_{IJ} \left( t - \frac{\Delta t}{2}, t + \frac{\Delta t}{2} \right) - S_{IJ} \left( t + \frac{\Delta t}{2}, t - \frac{\Delta t}{2} \right) \right],
\]

(9)

where \( S_{IJ} \) are the overlap terms

\[
S_{IJ}(t', t) \equiv \langle \Psi^I (t') \lvert \Psi^J (t) \rangle.
\]

(10)

Either adopting the determinant-derivative or the orbital-derivative approach, calculating these overlaps reduces to determining the atomic-orbital overlap matrix between sequential time steps. Within this formalism, the presence of a polarizable MM environment does not pose any further complications because the overlap calculation is not explicitly affected by the environment.

Gaussian is used to compute ground and excited-state energies and gradients, considering the electrostatic and polarization part of the AMOEBA force field, the TD-DFT, and their coupling. Energies and gradients of the bonding components of the force field, together with the ones of van der Waals non-bonding interactions, are evaluated by Tinker, which adds them to the terms computed by Gaussian. To keep this framework unchanged, given the system’s geometry, we wrote an interface tool within the Tinker package, which allows computing single-point energies.
and gradients for the LR-TD-DFT/AMOEBA potential. When such a calculation is performed, Tinker computes bonding and van der Waals non-bonding interactions for the system, creates the input for Gaussian, runs it, and finally collects all the results together, printing energies and forces on a formatted interface file.

All the implementations described up to now are self-contained in a Fortran module and a Perl script used to interface Newton-X NS and Tinker/Gaussian, providing the data needed to the core code in the correct format (and units). Since the TSH core code was not ready to deal with QM/MM calculations, some minor modifications have also been performed. In particular, manipulation of the system’s kinetic energy is needed to handle frustrated hoppings, velocity rescaling after hopping, and decoherence corrections. When a QM/MM potential is used, it is more sound to exclude atoms from the environment and to only use the kinetic energy of the QM part.

3 A test application

As a test case of the TSH with TD-DFT/AMOEBA implementation, we focused on a pyrimidine (Pm) water cluster.

The photoinduced process in nitrogen aromatic heterocycles interacting with water molecules have been accurately investigated, combining laser spectroscopy, mass spectrometry, and QM calculations. These investigations show that the mechanism involves a photoinduced electron-driven proton-transfer (EDPT) reaction, representing a subcase of the proton-coupled electron-transfer (PCET) mechanism. Namely, after the photo-excitation of the aromatic heterocycle, an electron is transferred from the lone pair of an H-bonded water to its π* orbital (see Figure 2). The resulting charge transfer (CT) state is quickly neutralized by the displacement of the H-bonding proton, forming a new bond with the aromatic nitrogen. After the EDPT, the system is still excited, but it quickly relaxes to the ground state (GS), which is a biradical at this geometry.

Huang et al. also showed that a TSH algorithm combined with the algebraic diagrammatic construction to the second-order [ADC(2)], a single-reference method like TD-DFT, could ade-
Figure 2: Graphical representation of the clusters Pm(H$_2$O)$_4$ (left); hydrogen bonding is represented with dashed yellow lines, molecules in MM region are represented with ball-and-sticks while the one in QM region are represented as licorice. Relevant transitions for the EDPT process represented as single excitations between molecular orbitals at the equilibrium geometry of Pm(H$_2$O)$_4$. MOs are computed with QM/AMOEBA Hamiltonian and plotted as isosurfaces at +0.01 and -0.01 in orange and blue, respectively.

quately describe the excited-state process until the system relaxed to the ground state, which has a diradical character. Here we followed a similar computational protocol for the Pm(H$_2$O)$_4$ system using TD-DFT as the QM method for Pm and the water molecule H-bonded to a Pm nitrogen atom and AMOEBA for the three "spectator" water molecules (see Figure 2).

### 3.1 Computational Details

Due to the importance of accurately describing CT states, we used the long-range corrected functional CAM-B3LYP, which should provide a balanced representation of the different electronic states involved in the process. The calculations were done with the 6-31+G(d) basis set.

All TSH simulations were performed applying the same protocol. We started from the optimized structure of the cluster; then, we computed the Hessian with respect to the nuclear coordinates. Geometry optimizations for QM/AMOEBA and QM/TIP3P were performed with the program `minimize.x` from the Tinker package using a threshold of 0.1 kcal mol$^{-1}$ Å$^{-1}$ on the gradients' root-mean-square. Since analytical second derivatives are not implemented in our software for QM/AMOEBA Hamiltonians, we used numerical differences of analytical gradients with a modified version of `vibrate.x` program from the Tinker package. Ground-state equilibrium
geometries and Hessian were used to sample 512 structures from a harmonic-oscillator Wigner distribution\textsuperscript{61} of the clusters, using Newton-X.\textsuperscript{30,62} On each structure, we computed the lowest five excitation energies at the same level of theory. Finally, we selected initial conditions from each state according to a stochastic algorithm based on the state’s oscillator strength within a 0.5 eV energy width around the equilibrium excitation energy of $S_4$.

Preparation of initial conditions for the pyrimidine-water droplet was performed differently: we took 128 configurations extracted from the Wigner distribution of Pm(H$_2$O)$_4$ and, using PACK-MOL,\textsuperscript{63} we added 1148 water molecules inside a sphere of 20-Å radius around the original cluster. Then, we performed a short Born-Oppenheimer dynamics simulation in the ground state to equilibrate the solvent. This simulation kept the QM region (the Pm molecule and the water H-bound to the nitrogen atom) rigid while all the classical waters were free to move. The Berendsen thermostat was applied with a $\tau$ constant of 0.05 ps$^{-1}$ and a reference temperature of 300 K. A harmonic wall (as implemented in Tinker) kept the molecules within 20 Å from the system’s center. The dynamics was performed with a time step of 0.5 fs for a total length of 1 ps. We noticed that the temperature is stabilized after about 500 fs. So, for each trajectory, we extracted two configurations spaced by 250 fs from the second part of the equilibration. On those structures, we computed excitation energies and performed the sampling of initial conditions, as described above for the Pm(H$_2$O)$_4$ cluster.

Each initial condition was used for a single trajectory. TSH was run with the decoherence-corrected\textsuperscript{53} fewest-switches surface hopping\textsuperscript{5} (DC-FSSH) approach. Time-derivative nonadiabatic couplings were calculated with the orbital-derivative approach\textsuperscript{52} as described in the Methods section. Each trajectory was propagated as a microcanonical ensemble for a maximum of 100 fs with a time step of 0.5 fs for the Newton’s equation integration; the time step for integration of the time-dependent Schrödinger equation was 0.025 fs with electronic properties interpolated between classical steps. The decoherence correction by Granucci and Persico\textsuperscript{53} was applied with a parameter of 0.1 Hartree. After a frustrated hopping, the velocity was kept in the original direction. After a hopping, the velocity was rescaled in the momentum direction. As mentioned in Methods, only
QM kinetic energies were considered for decoherence, frustrated hoppings, and velocity rescaling. Due to limitations of TD-DFT to describe the crossing region between the ground and the first excited state, only hoppings between excited states were evaluated. Whenever the $S_1/S_0$ energy gap dropped below 0.2 eV, the trajectory was terminated and the corresponding time was taken as the time for the internal conversion to the ground state. Trajectories were considered valid only if they were longer than 5 fs. They were terminated if they underwent a total energy drift superior to 0.5 eV.

The trajectories were analyzed using a Python code based on NumPy, SciPy, and MD-Analysis libraries. Data visualization was performed using Matplotlib, while molecular visualization was performed using PyMOL. Automatic analysis of the TD transition density matrices was performed with TheoDORE.

### 3.2 Results

We ran three different sets of TSH simulations to establish a robust dataset to test the implementation of TD-DFT/AMOEBA. The first reference set focused on Pm(H$_2$O)$_4$ with pyrimidine and one water molecule in the QM (TD-DFT) and three water molecules in the MM (AMOEBA) regions (Figure 2). The second set adopted the same QM and MM regions but with regular electrostatic embedding, using the TIP3P force field. The third set striped the spectator water molecules and simulated only Pm(H$_2$O)$_1$ with TD-DFT.

As a first analysis, we compare the EDPT probabilities in the three trajectory sets. The results are shown in Figure 3, where we report the Pm(N)-H(water) distance computed along TSH trajectories of QM/AMOEBA, QM/TIP3P, and Pm(H$_2$O)$_1$. As it can be seen, for all three sets of trajectories, the H transfer occurs in the first 100 fs of the excited state radiationless decay. When we compare the EDPT events we see only slight differences among the three sets: 5 in 222 trajectories (2.2 %) for QM/AMOEBA, 3 in 220 (1.4 %) for QM/TIP3P, and 3 in 190 for Pm(H$_2$O)$_1$ (1.6 %). For a 95% confidence interval, the margin of error is $\pm$2% in the three cases, meaning these results agree with each other within the statistical uncertainty.
Figure 3 also shows that in Pm(H$_2$O)$_1$ many trajectories lead to a breaking of the H-bond (34±7%) while this behavior is largely reduced in QM/AMOEBA (7±3% of trajectories with N-H distance reaching over 2.5 Å) and almost canceled in QM/TIP3P (1±1%). The difference between Pm(H$_2$O)$_1$ and Pm(H$_2$O)$_4$ reflects the chemical intuition that an H-bonding network contributes to the system’s stability. On the other hand, the larger propensity of the QM/AMOEBA cluster to break apart the N(Pm)-H(water) hydrogen bond with respect to QM/TIP3P can be interpreted as the result of a softer nature of the polarizable AMOEBA force field with respect to TIP3P.

![Figure 3: H-N distance computed along TSH trajectories of QM/TIP3P (left), QM/AMOEBA (center), and Pm(H$_2$O)$_1$ (right). Lines corresponding to EDPT events are highlighted; the final cross indicates that the trajectory is terminated before 100 fs.](image)

We have also performed a fourth set of TSH trajectories using a full-QM Pm(H$_2$O)$_4$ cluster. In this case, the EDPT probability is significantly increased, with 26 out of 167 trajectories (16±6%) showing the H transfer (see Figure S1 of the SI). We recall that in the previous ADC(2) simulations of same cluster,$^{33}$ the EDPT event was seen with a probability of 6%. Analyzing a reactive trajectory and evaluating the character of the different excited states in time (see Figure S2 of the SI), we have seen low-lying charge-transfer states involving orbitals delocalized over all water molecules. As a result, the H-transfer mechanism significantly changes as the EDPT is no longer localized on the water directly H-bonded to the Pm nitrogen. Because of that, the comparison with QM/MM descriptions is no longer meaningful and the full-QM Pm(H$_2$O)$_4$ model will no longer be considered.

To have a more direct analysis of the process, we compared two representative examples of reactive and nonreactive QM/AMOEBA trajectories. For each frame of the two trajectories,
we computed the transition density matrices for the lowest-lying states and analyzed them with TheoDORE. In Figure 4, we show oscillator strengths \((f)\) and the CT character of the different electronic states along the two trajectories; the N-H distance is also reported for reference. To compute the CT character of a state from the transition density matrices, the system has to be partitioned into molecular fragments; since the molecular entities do change upon EDPT, we adopted a dynamic definition of the molecular fragments based on the N-H bond distance. Using such a definition, during the first part of a reactive trajectory, the fragments are Pm and water, while they change into PmH and OH after EDPT. These graphs show a straightforward behavior for the nonreactive trajectory: the bright \(\pi\pi^*\) state populated by light absorption quickly relaxes to a dark state \((n\pi^*\) in character), which evolves in time, substantially lowering its energy from the initial \(\sim 5.5\ eV\) to less than 3 eV. During the dynamics, we observe the presence of some states with a significant CT character but always well separated (much more than 1 eV) from the current state.

The reactive trajectory still starts on a bright state in a situation that does not seem very different from the previous one. However, the dynamics now follows a different path: for the first \(\sim 50\) fs, the bright state remains populated until it jumps to a low-lying (dark) CT state. When the CT state is initially populated, the N–H distance is still above 1.5 Å, indicating that the electron transfer (populating the CT state) causes the subsequent proton transfer, as previously found in the ADC(2) simulations. To better interpret the graphs reported in Fig. 4, we recall that the definition of CT character of each excited state depends on the reference ground state. Since we are using a closed-shell representation, after EDPT the GS is described as an ion pair \((\text{PmH}^+ \cdots \text{OH}^-)\) and, as a consequence, the diradical excited states \((\text{PmH}^\bullet \cdots \text{OH}^\bullet)\) present a high CT character with respect to this reference.

We further investigated the energy and the character of the states in time. To do so, we considered the same QM/AMOEBA reactive trajectory already analyzed before (see Figure 4), and we recomputed the excited states using different models for the spectator waters, namely, neglecting them (no env), using TIP3P, and correcting QM/AMOEBA for a state-specific effect using the cLR\(^2\) approximation commented in the Methods section. The graphs reported in Figure 5 show...
Figure 4: Comparison of a nonreactive (left) and a reactive QM/AMOEBA trajectory (right). Top panel: excitation energies of the low-lying states colored in terms of the intensity of their oscillator strength (see color bar inside the picture); the state with green edges is the current state of TSH, vertical lines indicate the hopping events, and the gray shadow the region where the simulation is stopped and considered as relaxed on GS. Central panel: same as the top one, but the color maps the CT \( \text{CT}_{\text{Pm} \rightarrow \text{H}_2\text{O}} \) character of the states. Bottom panel: distance of Pm N atom from the hydrogen atom of the initially H-bonded water molecule.

these results in comparison to the original linear-response QM/AMOEBA (LR) for reference.

The presence of the additional water molecules destabilizes the CT states, increasing their energies with respect to the model with no environment. Such an effect, slightly more marked for QM/AMOEBA than for QM/TIP3P, can be explained in terms of stabilization of the lone pair orbital when in the presence of the additional water molecules. The introduction of the state-specific correction \( (c\text{LR}^2) \) inverts this behavior, stabilizing the CT states, while leaving the other states unperturbed. The global effect, however, is small due to the small number of polarizable water molecules.

Following these findings, we decided to further investigate how state-specific effects could
Figure 5: Energies and CT character of the lowest excited states computed along the QM/AMOEBA reactive trajectory shown in Figure 4 with different representation of environment water molecules: omitted (upper left), LR QM/AMOEBA (upper right), QM/TIP3P (bottom left), and cLR² QM/AMOEBA for each state (bottom right). In the latter case only the lowest 10 excited state have been analyzed.

impact the description in a large (bulk-like) environment. To check this point, we ran another set of TSH trajectories for Pm(H₂O)₁ within a 20-Å droplet of AMOEBA water molecules (Figure 6). We found that the EDPT probability for the droplet remains close to the one found for Pm(H₂O)₄ (2 reactive out of 81 trajectories, or 2.5 ± 3 %). By analyzing a reactive trajectory, we saw that the behavior also closely resembles the one already discussed for Pm(H₂O)₄, with a transition from a bright to a dark, CT state, finally leading to the H transfer. However, when we calculated the QM/AMOEBA-cLR² energies of the low-lying states using the frames from the first 30 fs of the QM/AMOEBA droplet trajectory, we found that the CT states are stabilized by a much larger amount than what was observed for Pm(H₂O)₄ (about 0.4 eV). One would expect that this stabilization had a not negligible effect on the EDPT probability. Unfortunately, the lack of a state-specific formulation of TSH due to the difficulties of such an implementation discussed in the Methods section, prevents a quantitative estimate of the final effect on the EDPT yield.
Figure 6: Pyrimidine water droplet (top-right). At the left, analysis of a reactive trajectory for the droplet. The format is the same as adopted in Figure 4. The bottom right shows the CT character of the first 30 fs of the same trajectory but considering cLR² AMOEBA.

4 Conclusions

We reported the implementation of trajectory surface hopping nonadiabatic dynamics performed with a QM method (TD-DFT) coupled with an atomistic polarizable embedding (AMOEBA). This implementation has been done by interfacing a development version, soon to be released, of the surface-hopping program Newton-X NS with the polarizable QM/AMOEBA approach from an interface between the Gaussian 16 and Tinker suites of codes. Our NX/TINKER/G16 interface will be released under the GPL3 license, freely available to anyone with access to Gaussian 16 and the Tinker version as modified in this work.

We tested our approach on photoexcited pyrimidine-water clusters, a challenging system because of the relevant role played by CT states. Our analysis led us to conclude that the performance of AMOEBA and TIP3P water models in this kind of problem are very similar if AMOEBA
is treated in a linear-response framework. Instead, significant differences in the charge-transfer states’ energies are obtained when QM/AMOEBA is treated within a state-specific framework. These findings push for a reliable and efficient implementation of analytical gradients of state-specific formulations of polarizable QM/MM approaches.

Supporting Information Available

Results for the set of TSH trajectories calculated using a full-QM Pm(H₂O)₄ cluster.

Acknowledgement

M. Bo. and B. M. acknowledge funding by the European Research Council under grant ERC-AdG-786714 (LIFETimeS). B. D. and M. Ba. acknowledge funding by the European Research Council under grant ERC-AdG-832237 (SubNano).

Data Availability

All the raw data supporting this study’s findings (complete trajectories and transition density matrices analysis) are available on figshare (10.6084/m9.figshare.20236674).

References


21


(43) Scalmani, G.; Frisch, M. J.; Mennucci, B.; Tomasi, J.; Cammi, R.; Barone, V. Geometries
and properties of excited states in the gas phase and in solution: Theory and application of a
time-dependent density functional theory polarizable continuum model. J. Chem. Phys. 2006,
124, 094107.

(44) Nottoli, M.; Mennucci, B.; Lipparini, F. Excited state Born–Oppenheimer molecular dynamics
through coupling between time dependent DFT and AMOEBA. Phys. Chem. Chem. Phys. 2020,
22, 19532–19541.

(45) Barbatti, M.; Pittner, J.; Pederzoli, M.; Werner, U.; Mitrić, R.; Bonačić-Koutecký, V.; Lis-
chka, H. Non-adiabatic dynamics of pyrrole: Dependence of deactivation mechanisms on the

(46) Huix-Rotllant, M.; Ferré, N.; Barbatti, M. In Quantum Chemistry and Dynamics of Excited

(47) Casida, M. In Recent advances in density functional methods, Part I; Chong, D., Ed.; World

(48) Hammes-Schiffer, S.; Tully, J. C. Proton-Transfer in Solution - Molecular-Dynamics with

(49) Mitrić, R.; Werner, U.; Bonačić-Koutecký, V. Nonadiabatic Dynamics and Simulation of
Time Resolved Photoelectron Spectra within Time-Dependent Density Functional Theory:

(50) Tapavicza, E.; Tavernelli, I.; Rothlisberger, U. Trajectory Surface Hopping within Linear

(51) Pittner, J.; Lischka, H.; Barbatti, M. Optimization of Mixed Quantum-Classical Dynamics:
Time-Derivative Coupling Terms and Selected Couplings. Chemical Physics 2009, 356, 147–
152.

25


(64) Oliphant, T. E. *A guide to NumPy*; Trelgol Publishing USA, 2006; Vol. 1.


