The Journal of Chemical Physics

RESEARCH ARTICLE | MAY 07 2024

Non-adiabatic direct quantum dynamics using force fields: Toward solvation

L. L. E. Cigrang ; J. A. Green ; S. Gómez ; J. Cerezo ; R. Improta ; G. Prampolini ; F. Santoro ; G. A. Worth

J. Chem. Phys. 160, 174120 (2024)
https://doi.org/10.1063/5.0204911

AIP Publishing

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Non-adiabatic direct quantum dynamics using force fields: Toward solvation

Cite as: J. Chem. Phys. 160, 174120 (2024); doi: 10.1063/5.0204911
Submitted: 23 February 2024 • Accepted: 17 April 2024 •
Published Online: 7 May 2024

L. L. E. Cigrang,[1] J. A. Green,[2] S. Gómez,[3] J. Cerezo,[4] R. Improta,[5] G. Prampolini,[6] F. Santoro,[6] and G. A. Worth[1,a)]

AFFILIATIONS

[1] Department of Chemistry, University College London, 20 Gordon St., WC1H 0AJ London, United Kingdom
[2] Institut für Physikalische Theoretische Chemie, Goethe-Universität, Max-von-Laue-Str. 7, 60438 Frankfurt am Main, Germany
[3] Departamento de Química Física, Universidad de Salamanca, Salamanca 37008, Spain
[4] Departamento de Química and Institute for Advanced Research in Chemical Sciences (IAdChem), Universidad Autónoma de Madrid, 28049 Madrid, Spain
[5] Istituto di Biostrutture e Bioimmagini-CNR, Via De Amicis 95, I-80145 Napoli, Italy
[6] Istituto di Chimica dei Composti Organometallici (ICCOM-CNR), Area della Ricerca del CNR, Via Moruzzi 1, I-56124 Pisa, Italy

[a)] **Author to whom correspondence should be addressed:** g.a.worth@ucl.ac.uk

ABSTRACT

Quantum dynamics simulations are becoming a powerful tool for understanding photo-excited molecules. Their poor scaling, however, means that it is hard to study molecules with more than a few atoms accurately, and a major challenge at the moment is the inclusion of the molecular environment. Here, we present a proof of principle for a way to break the two bottlenecks preventing large but accurate simulations. First, the problem of providing the potential energy surfaces for a general system is addressed by parameterizing a standard force field to reproduce the potential surfaces of the molecule's excited-states, including the all-important vibronic coupling. While not shown here, this would trivially enable the use of an explicit solvent. Second, to help the scaling of the nuclear dynamics propagation, a hierarchy of approximations is introduced to the variational multi-configurational Gaussian method that retains the variational quantum wavepacket description of the key quantum degrees of freedom and uses classical trajectories for the remaining in a quantum mechanics/molecular mechanics like approach. The method is referred to as force field quantum dynamics (FF-QD), and a two-state $\pi\pi^*/n\pi^*$ model of uracil, excited to its lowest bright $\pi\pi^*$ state, is used as a test case.

I. INTRODUCTION

Quantum dynamics (QD) simulations, which solve the time-dependent Schrödinger equation (TDSE), are able to simulate molecular systems correctly accounting for quantum effects such as non-adiabaticity and tunneling. They can thus provide the dynamical data required to understand the behavior of photo-excited molecules. These simulations do, however, require significant computer resources and scale exponentially with increasing system size. In addition, the environment is known to play a significant role. For example, polar solvents shift the energies and relative ordering of states, and hydrogen-bonding from solvent molecules can further change the relevant potential energy surfaces (PESs). For state-of-the-art simulations, this presents a major challenge as the many molecules involved make standard QD simulations prohibitively expensive.[1]

Over recent years, a number of QD methods tailored to systems with many vibrational degrees of freedom (DoFs) have been developed with differing levels of accuracy. They fall into three main classes. The most accurate method, exemplified by the multi-configuration time-dependent Hartree (MCTDH)[2–6] method, uses a grid-based representation of the Hamiltonian and wavefunction to propagate a wavepacket. They can provide a complete solution of the TDSE for systems of tens of atoms but require global PESs to

be calculated before a simulation can be performed. The simplest techniques are based on classical mechanics and approximate the evolving wavepacket as a swarm of independent trajectories. The most widespread of these for non-adiabatic simulations is surface hopping,[7–11] which uses a stochastic algorithm to treat the population transfer between electronic states. These simulations are cheap to run and can calculate the potential surfaces *on-the-fly* using standard quantum chemistry codes. They, however, are expected to perform poorly in cases where the coupling between the electronic and nuclear DoFs is very strong. In this case, a full quantum treatment of the evolving wavepacket is required.

With this in mind, a third class of QD methods was developed based on expanding the wavepacket in a basis set of time-dependent Gaussian functions known as Gaussian wavepackets (GWPs). These GWPs are localized and follow trajectories, bringing some of the simplicity of trajectory based methods and allowing the use of *on-the-fly* potentials. They can also, at least in principle, converge on the full TDSE solution. Methods in this class include *ab initio* multiple spawning (AIMS),[12,13] coupled coherent states (CCS),[14–16] and the variational multiconfigurational Gaussian (vMCG) method.[17,18] The latter differentiates itself by obeying a set of equations of motion derived from the time-dependent Dirac–Frenkel variational principle, allowing the propagated GWPs to follow fully variational trajectories. This is an important distinction from AIMS and CCS, in which the GWPs follow classical trajectories and have difficulties reaching classically forbidden regions of the PES. They may also rapidly become uncoupled from one another during a simulation, requiring a large number of trajectories to reach convergence. Variational trajectories, however, can easily account for the quantum behavior of nuclei and require fewer GWPs for convergence. The vMCG method thus promises to provide flexible and accurate non-adiabatic simulations, albeit at a higher computational cost than its classical trajectory based alternatives. Despite this, a recent benchmark study indicates that vMCG may, in fact, be more efficient in direct dynamics calculations as it uses a database to build up the potential surfaces rather than calculate new energies and gradients at each time-step during a propagation.[19] Yet, to reach the many DoFs in a solvated molecular system while retaining the accuracy of the vMCG method, further developments are needed. In this paper, a hierarchy of methods is investigated in which different equations of motion (EOMs) are used for the GWP basis, from fully quantum (variational) to fully classical. It will be shown that by treating different parts of the system at different levels of theory in a similar way to quantum mechanics (QM)/molecular mechanics (MM) approaches, the quantum features that need to be described correctly can be combined with a cheaper treatment of the environment modes.

An additional major computational resource required is related to the quantum chemistry calculations needed when using *ab initio* potentials calculated on-the-fly in so-called direct quantum dynamics simulations. A second aim of this work is to provide a methodology for efficient, yet accurate, coupled excited-state potentials that require lower computational resources so that it can be applied to much larger systems. Recent methodological advances have been made toward this goal. One such development is the semi-automatic parameterization of quantum-mechanically derived force fields (QMD-FFs),[20,21] able to reproduce equilibrium structures, normal modes, and frequencies and, at the same time, accurately describe the molecular flexibility for both ground and excited states.[22,23] The adoption of QMD-FFs in direct quantum dynamics simulations will enormously speed up the propagations, providing cheap yet reliable estimates of energies, gradients, and second derivatives, with a gain even larger for vMCG than for classical MD, where Hessians are not needed. In order to exploit this potential, however, it is necessary to interface codes for QD computations with those able to efficiently manage the FFs. To that end, in this contribution, we introduce an interface between the quantum dynamics suite QUANTICS and the classical MD engine GROMACS,[24] making it possible to perform QD with GWPs driven by QMD-FFs. The expected huge speed up of the computations will also make it possible to push the convergence of direct QD simulations much further than is usually possible, allowing new aspects of the performance of the methods themselves to be examined.

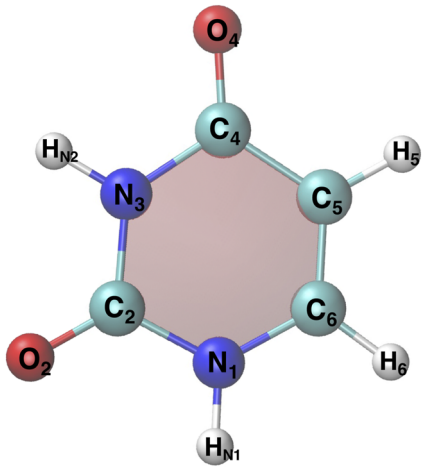

FIG. 1. Schematic structures and atom labeling of uracil. Color code: carbon (green), nitrogen (blue), oxygen (red), and hydrogen (white).

We shall use the photoactivated dynamics of uracil (see Fig. 1) as a test case for our method development. Absorption of UV light by DNA can indeed trigger many photochemical events, which lead to the damage of the genetic code and, hence, to many harmful processes for living beings.[25] This makes the study of the excited state dynamics of the constituents of DNA, the nucleobases, a hot topic for time-resolved spectroscopy and computational studies.[26–29] In particular, the photophysics of uracil has been extensively studied, and many experimental and computational results are available, making this molecule, beyond its intrinsic biochemical interest, an ideal playground for methodological development. Moreover, uracil and uridine (its corresponding nucleoside), despite their sub-ps excited state lifetimes, exhibit extremely short and quite rich photophysics, and many features have not yet been completely understood.[26–28,30–32] Among them, one of the most debated issues concerns the possible involvement of a "dark" electronic transition in the excited state decay of the spectroscopic state of uracil and the other pyrimidines.

In these molecules, the lowest energy bright state corresponds to a $\pi\pi^*$ transition with a predominant HOMO → LUMO character,[26,28,29] hereafter labeled $S\pi\pi^*$. In addition, independently of the electronic structure method, all the recent computational

studies agree in predicting that in the Franck–Condon (FC) region, one (for uracil and thymine) or two (for cytosine) $n\pi^*$ transitions (hereafter S$n\pi^*$) fall within ±0.5 eV of S$\pi\pi^*$.[26,28,29,32] These excited states can be described as resulting from the excitation from the lone pair of the carbonyl oxygen and (for cytosine) of the nitrogen of the ring toward the π^* LUMO. Many experimental and computational studies indicate that depending on the pyrimidine and the solvent, a significant part of the population of S$\pi\pi^*$ (even ≥90%) can be transferred to one S$n\pi^*$.[26,28,29] These studies indicate that the amount of population transferred to the lowest energy, dark $n\pi^*$ state depends critically on the nature of the embedding medium. For example, for thymine in the gas phase, the S$\pi\pi^* \rightarrow$ S$n\pi^*$ transfer is a major deactivation channel, according to both experiments[33] and calculations,[34–36] whereas in water, the population transfer is negligible (≤10%).[31,37] For uracil (see Ref. 32 for a review of the theoretical studies in the gas phase), both quantum dynamical computations on reduced dimensionality models describing the solvent effect with continuum models and trajectory surface hopping simulations in explicit solvents indicate that the S$\pi\pi^* \rightarrow$ S$n\pi^*$ is much more favored in the gas phase than in polar solvents.[38,39]

In the following, the hierarchy of vMCG methods will be described, along with the parameterization of the QMD-FF for simulating photo-excited uracil via the semi-automatic JOYCE procedure.[20,21] The interface between the QUANTICS program that runs direct dynamics (DD)-vMCG simulations and the classical MD code GROMACS is also introduced as a tool to perform FF-QD calculations. These are necessary prerequisites for any further exploration. Therefore, although in this contribution we do not explicitly run computations in the solvent, we use our QUANTICS/GROMACS interface to run a number of tests on the expansion of the total WP in GWPs that are preparatory to such propagations. The final addition of solvent will then be tackled in a subsequent dedicated contribution.

II. METHODOLOGY

A. Grid-based quantum dynamics

In order to provide a reliable benchmark for our study, we used the Multi-Configuration Time-Dependent Hartree (MCTDH) method,[2,3] which is able to provide a numerically exact solution to the time-dependent Schrödinger equation. In this method, the nuclear wavefunction is expanded as a full direct-product of time-dependent basis functions,

$$\Psi(\mathbf{q},t) = \sum_s \sum_{j_1\dots j_p} A_{j_1\dots j_p s} \varphi^{(1)}_{j_1}(q_1,t)\cdots\varphi^{(p)}_{j_p}(q_p,t)|s\rangle, \tag{1}$$

where the "single-particle functions" (SPFs), φ, are expanded in turn on a time-independent basis,

$$\varphi_j(q_\kappa,t) = \sum_a c_{aj}(t)\chi_a(q_\kappa), \tag{2}$$

and $|s\rangle$ is a vector to provide the associated electronic state basis. The Dirac–Frenkel variational principle is then used together with this ansatz to provide equations of motion for the expansion coefficients and basis functions to optimally describe the evolving wavepacket.

MCTDH is well established and described in detail in the literature. In the following, the multi-layer (ML-MCTDH) form will be used,[4,5,40] which is required for systems with large numbers of vibrational modes, such as uracil. In this approach, the single particle functions of Eq. (1) are multi-dimensional functions that are first expanded in an MCTDH form. This is repeated, forming layers of functions, until the lowest layer is reached, which is described by a set of time-independent grids.

The disadvantage of the MCTDH scheme is that it requires the potential surfaces to be known analytically in advance and to be in a particular "sum of product" form for efficiency. Fortunately, the vibronic coupling model on which the QMD-FF to be developed is based has exactly this form, whereas the QMD-FF potentials do not. An ML-MCTDH calculation can thus provide benchmarks for the time-evolution of the state populations. This is done by projecting the time-dependent wavepacket onto an electronic state,

$$P_s = \langle\Psi(t)|s\rangle\langle s|\Psi(t)\rangle. \tag{3}$$

B. vMCG: Partitioning and approximation hierarchy

A different, potentially more flexible way of solving the TDSE is to expand the wavefunction using a set of time-dependent Gaussian wavepackets (GWPs), G_j,

$$\Psi(\mathbf{q},t) = \sum_{js} A_{js}(t)G_j(\mathbf{q},t)|s\rangle, \tag{4}$$

where $|s\rangle$ is again the electronic state basis. The Gaussian wavepackets themselves are separable functions of one-dimensional Gaussian functions,

$$G_j(\mathbf{q},t) = g^{(1)}_j(q_1,t)\cdots g^{(f)}_j(q_f,t), \tag{5}$$

with the form

$$g^{(\kappa)}_i(q_\kappa,t) = \exp\left(-\alpha^{(\kappa)}_i q^2_\kappa + \xi^{(\kappa)}_i(t)q_\kappa + \eta^{(\kappa)}_i(t)\right). \tag{6}$$

In these functions, the widths, α, of the functions are time independent (frozen Gaussian approximation). The scalar, η, parameters are defined by keeping the functions normalized with no phase, allowing the phase to be carried by the expansion coefficient. This leaves the linear, ξ_i, parameters to carry the time-dependence of the functions.

The GWP basis functions can be compared with the Gaussian wavepackets pioneered by Heller,[41,42]

$$\tilde{g}^{(\kappa)}_i(q_\kappa,t) = \exp\left(-\alpha(q_\kappa - q_{i\kappa}(t))^2 + ip_{i\kappa}(t)(q_\kappa - q_{i\kappa}(t)) + i\gamma_i(t)\right), \tag{7}$$

where $p_{i\kappa}, q_{i\kappa}$ are the momentum and coordinate of the center of the wavepacket and γ_i the phase and normalization. This leads to a relationship between the linear parameter, the momentum and the coordinate,

$$\xi_i(t) = -2\alpha q_{i\kappa}(t) + ip_{i\kappa}(t). \tag{8}$$

Using the Dirac–Frenkel variational principle to solve the TDSE with the ansatz Eq. (4) leads to equations of motion for the expansion coefficients,

$$i\dot{A}_{is} = \sum_{jks'} S^{-1}_{ij}\left(H^{(ss')}_{jk} - i\tau_{jk}\delta_{ss'}\right)A_{ks'}, \tag{9}$$

with

$$S_{ij} = \langle G_i | G_j \rangle, \quad \text{GWP overlap,} \tag{10}$$

$$H_{ij}^{(ss')} = \langle sG_i | H | G_j s' \rangle, \quad \text{Hamiltonian matrix,} \tag{11}$$

$$\tau_{ij} = \langle G_i | \dot{G}_j \rangle, \quad \text{overlap time – derivative,} \tag{12}$$

and for the Gaussian parameters for each set of functions,

$$i\dot{\xi}_{\kappa i} = X_{\kappa i} + \sum_{\mu j} C^{-1}_{\kappa i,\mu j} Y_{\mu j}, \tag{13}$$

with

$$C_{\kappa i,\mu j} = \rho_{ij}\left(S_{ij}^{\kappa\mu} - S_{ik}^{(\kappa 0)} S_{kl}^{-1} S_{lj}^{(0\mu)}\right), \tag{14}$$

$$Y_{\kappa i} = \sum_{\mu j}\left(i\rho_{ij} H_{ij}^{0\mu} - \rho_{im} S_{ik}^{(\kappa 0)} S_{kl}^{-1} S_{lj}^{(0\mu)} H_{jm}\right), \tag{15}$$

$$H_{ij} = \langle G_i | H | G_j \rangle, \quad \text{Hamiltonian matrix,} \tag{16}$$

$$X_{\kappa i} = -2\alpha \frac{\partial V}{\partial q_\kappa} + i\frac{p_{\kappa i}}{m_\kappa}, \quad \text{classical terms,} \tag{17}$$

$$\rho_{ij} = A_i^* A_j, \quad \text{density matrix,} \tag{18}$$

$$S_{ij}^{(\alpha\beta)} = \langle \frac{\partial G_i}{\partial \lambda_{i\alpha}} | \frac{\partial G_j}{\partial \lambda_{j\beta}} \rangle, \tag{19}$$

$$H_{ij}^{(\alpha\beta)} = \langle \frac{\partial G_i}{\partial \lambda_{i\alpha}} | H | \frac{\partial G_j}{\partial \lambda_{j\beta}} \rangle, \tag{20}$$

where the last two lines define the overlaps and Hamiltonian matrix elements in terms of the derivatives of the GWPs with respect to the linear parameters.

This is the variational Multiconfigurational Gaussian (vMCG) ansatz.[18,43] It has the property that the centers of the GWP functions follow trajectories that have a classical part and an additional "variational coupling" ($C^{-1}Y_R$) that allows the GWPs to move optimally as a set to describe the evolving wavepacket. In addition, if the integrals involving the Hamiltonian are evaluated exactly, the result will converge on the full numerical solution of the TDSE with conservation of energy and norm by construction.

The major part of the effort of solving the vMCG equations is inverting the C-matrix of Eq. (14). In a calculation with N DoFs and n GWPs, this matrix has dimensions $(nN)^2$, and the effort of inverting a matrix scales with the cube of the dimension. It is possible to reduce this effort by partitioning the GWPs into products of lower-dimensional GWPs, i.e.,

$$G_j(\mathbf{q}, t) = G_{j_1}^{(1)}(\mathbf{q}_1, t) G_{j_2}^{(2)}(\mathbf{q}_2, t) \cdots. \tag{21}$$

Using these lower-dimensional GWPs as the basis sets, the overall wavefunction ansatz is

$$\Psi(\mathbf{q}, t) = \sum_{j_1 j_2 \ldots s} A_{j_1 j_2 \ldots s}(t) G_{j_1}^{(1)}(\mathbf{q}_1, t) G_{j_2}^{(2)}(\mathbf{q}_2, t) \cdots |s\rangle. \tag{22}$$

This is the G-MCTDH method of Burghardt *et al.*,[44] and it can be viewed as partitioning a wavefunction into κ parts, described by sets of GWPs, $G_j^{(\kappa)}(\mathbf{q}_\kappa, t)$. The equations of motion for the expansion coefficients, Eq. (9), are unchanged, with G_j referring to a configuration of GWPs [Eq. (21)] rather than a single basis function. The equations for the GWP function parameters also retain the same form as Eq. (13), but now the H terms in Eq. (15) become mean field operators linking the different parts of the system. The full details of the G-MCTDH equations of motion are in the literature.[44] This partitioning reduces the effort by providing sets of lower dimensional GWPs with associated smaller C-matrices to invert. The cost is that there are now mean-field operators that need to be built.

Ignoring the second term in Eq. (13) further saves effort by removing the need for the matrix inversion. If this is done, the GWPs follow classical trajectories, and the coupling that provides the variational character of the basis is lost. This has the effect of slowing the convergence of the method, and energy conservation is no longer ensured. Following earlier work,[45] this is referred to as the classical multiconfigurational Gaussian (clMCG) approximation.

A major part of the effort is also required to calculate the Hamiltonian matrix elements $\langle G_i | H | G_j \rangle$. If rectilinear coordinates are used, then the kinetic energy operator is a sum of one-dimensional terms. For example, in the mass-frequency scaled normal modes used in this work,

$$T = \sum_\kappa -\frac{1}{2}\frac{\partial^2}{\partial q_\kappa^2}, \tag{23}$$

the integrals over this operator thus become simple one-dimensional integrals that can be evaluated analytically. In contrast, integrals over the potential operator are multi-dimensional integrals over the GWPs. For these, in vMCG, it is typical to use a local harmonic approximation (LHA) in which the surfaces are expanded to second order around the coordinate at the center of the GWP on the right-hand side of the matrix element,

$$V(\mathbf{q}) \approx V(\mathbf{q}_j) + \sum_\kappa \left.\frac{\partial V}{\partial q_\kappa}\right|_{\mathbf{q}_j} (q_\kappa - q_{\kappa j}) + \frac{1}{2}\sum_{\kappa\kappa'} \left.\frac{\partial^2 V}{\partial q_\kappa q_{\kappa'}}\right|_{\mathbf{q}_j} (q_\kappa - q_{\kappa j})(q_{\kappa'} - q_{\kappa' j}), \tag{24}$$

where $\mathbf{q}_j$ is the center coordinate of GWP G_j. This allows the integral to be calculated analytically. It should be noted that this method leads to a matrix that needs to be made Hermitian by averaging H_{ij} and H_{ji}.

The number of second-order terms, however, grows quadratically with system size and soon becomes a major part of the effort. A simpler approximation is to take just the zero-order term and approximate the integral as

$$\langle G_i | V | G_j \rangle \approx \langle G_i | G_j \rangle V(\mathbf{q}_j). \tag{25}$$

This *saddle point approximation* is regularly used in non-adiabatic simulations.[46] It cannot be used in the vMCG approximation, as the variational coupling would be lost. In the clMCG approximation, however, this is not the case, and it is merely an approximation to the matrices used in the expansion coefficient propagation. The use of this approximation will be referred to as saddleMCG.

It should be noted that different partitions of a wavefunction can be treated with different levels of approximation, leading to various possibilities for "hybrid" calculations, as will be detailed further in Sec. II F.

C. DD-vMCG: *On-the-fly* non-adiabatic dynamics

The vMCG method, as described above, has the property that the potential given in Eq. (24) can be calculated *on-the-fly*, requiring only information about the potential around the center of a GWP. This means that regions of the PES are only calculated if and when the wavepacket moves into them during the propagation. The three terms in Eq. (24) are readily obtained from standard quantum chemistry packages in the form of the adiabatic potential energies, the nuclear gradients, and the Hessian matrices, respectively. By also calculating the non-adiabatic coupling vectors, one can now solve the nuclear and electronic parts of the TDSE and account for the coupling between them, resulting in the DD-vMCG method, which is well suited for the study of photo-induced processes in systems with complex PESs.[17,47,48]

In practice, DD-vMCG makes use of a database containing all the information from the individual quantum chemistry calculations.[49] Instead of calculating new points of the PES at every time-step, the database is checked, and a comparison is made between the previously stored molecular geometries and the new ones. This comparison is based on the evaluation of the norms of the atomic difference vectors between the structures.[50] If the lowest norm is higher than a fixed threshold, a new electronic structure calculation is launched, the result of which is then added to the database. If, however, the lowest norm is lower than the threshold, a Shepard interpolation[51,52] is done to obtain the LHA parameters at the new geometry.

In the calculations presented later, rather than using the potential surfaces from quantum chemistry calculations, the potentials will be provided by the QMD-FF to be detailed in Sec. II D. Importantly, the LHA in vMCG simulations requires the Hamiltonian to be in the diabatic picture to ensure smoothness under each GWP. The JOYCE QMD-FF is in the diabatic picture, so it can be directly used without any transformations.

A final point worth mentioning is that it is common practice when using DD-vMCG to approximate the Hessian at every point rather than calculating it explicitly. This saves an enormous amount of computational effort since the full Hessian of every electronic state only needs to be calculated once at the FC point and is then updated based on this (so-called Hessian updating).[18,53] However, since the new approach described in this work uses molecular mechanics rather than quantum chemistry to obtain the necessary properties, explicitly calculating the Hessian at each new point is feasible without it becoming a computational bottleneck.

D. Model Hamiltonian

In order to use QMD-FFs for vMCG simulations, it is necessary to have a reliable protocol to parameterize a FF for every ground and excited diabatic state involved in the process (diagonal FFs) along with the diabatic couplings between them (off-diagonal FFs). The JOYCE methodology developed by some of us intrinsically provides the diagonal diabatic FFs since it is based on a local expansion of the adiabatic potentials (for rigid systems). In order to obtain the diabatic couplings, in this contribution we propose a protocol to merge Joyce FFs with a Vibronic Coupling (VC) model. Both Joyce and VC models can be parameterized at the same level of electronic structure theory [here density functional theory (DFT) and time-dependent (TD)-DFT], which will then describe the same potentials.

The Joyce FF has the advantage of providing greater flexibility than a VC model for the molecular motions and, crucially for our aims, can be combined with a solvent FF, as we discuss in the supplementary material, Sec. S4.1. This allows us to exploit the knowledge acquired in the last few decades in the field of molecular mechanics (MM) to obtain straightforward FFs for systems embedded in solvents or more complex explicit environments. As we mentioned in the introduction, we leave this final addition of solvent to a subsequent work, and in the following, we describe how the intramolecular force-fields for a molecule may be incorporated into a QD simulation in the gas phase.

A specific QMD-FF is separately parameterized, according to the JOYCE protocol, for the uracil ground state (S_0) and each of the first two excited states, $S n\pi^*$ and $S \pi\pi^*$, respectively. A brief description of the protocol will be given below, and full details of the parameterization procedure can be found in the original papers.[20,21,54] Considering the overall stiffness expected for the uracil ring, such an intra-molecular force field, $E^{intra}_{QMD-FF|s}$ takes the simplified expression,

$$E^{intra(\mathbf{r}^{ric},\mathbf{R}^{ric})}_{QMD-FF|s} = E^{s}_{str}(\mathbf{r}^{ric}) + E^{s}_{bnd}(\mathbf{r}^{ric}) + E^{s}_{tors}(\mathbf{r}^{ric}), \tag{26}$$

where, for each state s, the three terms refer to the bond stretching (E^s_{str}), angle bending (E^s_{bnd}) and dihedral torsion (E^s_{tors}) energies, which depend on stiff redundant internal coordinates (RICs, $\mathbf{r}^{\mathrm{RIC}}$) of uracil and are all modeled with harmonic potentials,

$$\begin{aligned} E^{s}_{str} &= \frac{1}{2}\sum_{i}^{N_{bonds}} ks^{s}_{i}(r_i - r^{S}_{0,i})^2, \\ E^{s}_{bnd} &= \frac{1}{2}\sum_{i}^{N_{angles}} kb^{s}_{i}(\theta_i - \theta^{s}_{0,i})^2, \\ E^{s}_{tors} &= \frac{1}{2}\sum_{i}^{N_{dihedrals}} kt^{s}_{i}(\phi_i - \phi^{s}_{0,i})^2. \end{aligned} \tag{27}$$

The parameters defining $E^{intra}_{QMD-FF|s}$ for each state s are obtained through the JOYCE code[55] by minimizing the difference between selected QM and FF descriptors through the objective function,

$$I^{s} = \sum_{K\le L}^{3N-6} \frac{2W_{KL}}{(3N-6)(3N-5)}\left[H^{s}_{KL} - \left(\frac{\partial^2 E^{QMD-FF|s}_{intra}}{\partial Q_K \partial Q_L}\right)\right]^2_{g_s}, \tag{28}$$

where, for each state s, the (double) sum runs over the $3N-6$ QM normal modes (N being the number of uracil atoms) obtained at its own optimized geometry (g_s) by diagonalizing the QM Hessian matrix $\mathbf{H}^s$, and the W_{KL} terms are tunable parameters, which were set to weight the normalized diagonal elements of the matrix $\mathbf{W}$ to twice the value of the off diagonal ones. In brief, the excited state FFs of the two lowest electronic excited states are expanded locally at their optimized geometries in C_S symmetry, where they have a pure

$n\pi^*$ or $\pi\pi^*$ character. Therefore, the obtained excited state FFs can be taken as diabatic $n\pi^*$ or $\pi\pi^*$ FFs.

Up to now, we have defined a diagonal FF for each state. In order to describe the non-adiabatic transition between $n\pi^*$ and $\pi\pi^*$, we need to account for the coupling between the states. The simplest strategy is to build a quadratic vibronic Hamiltonian (QVC) by taking as coordinates the normal modes of the S_0 state and taking an LHA of the FFs for the different states.

The Hamiltonian thus takes the following form:

$$H = \sum_i \left(T + V^{(i)}(\boldsymbol{q})\right)|i\rangle\langle i| + \sum_{i,j>i} V^{(i,j)}(\boldsymbol{q})\left(|i\rangle\langle j| + |j\rangle\langle i|\right), \quad (29)$$

where $\boldsymbol{q} = \{q_1, q_2, \ldots, q_\nu\}$ are the dimensionless normal mode coordinates. The kinetic energy operator is defined in Eq. (23), and the potential terms have the form

$$V^{(i)}(\boldsymbol{q}) = E^{(i)}(t) + \sum_\alpha^\nu \kappa_\alpha^{(i)}(t) q_\alpha + \frac{1}{2}\omega_\alpha q_\alpha^2 + \sum_{\alpha,\beta}^\nu \gamma_{\alpha,\beta}^{(i)}(t) q_\alpha q_\beta, \quad (30)$$

$$V^{(i,j)}(\boldsymbol{q}) = \sum_\alpha^\nu \lambda_\alpha^{(i,j)}(0) q_\alpha, \quad (31)$$

where ω_α are the ground state normal mode frequencies, $\kappa_\alpha^{(i)}$ the gradient of electronic state i with respect to normal mode α, $\gamma_{\alpha,\beta}^{(i)}$ the bi-linear coupling between normal modes α and β on state i to account for the fact that each state has different normal modes and frequencies (Duschinsky rotation), and $\lambda_\alpha^{(i,j)}$ the vibronic coupling between states i and j with respect to normal mode α. Notice that a parenthesis (t) has been added to indicate that the corresponding terms are updated during the dynamics *on the fly*.

At time $t = 0$ of the dynamics, where the wavepacket is centered at the S_0 equilibrium geometry, the only ingredient missing to fully define the Hamiltonian are the vibronic couplings $\lambda^{(i,j)}$. They can be easily computed through a diabatization technique[56] and are not updated further during the dynamics. At later times, the other parameters of the Hamiltonian are straightforwardly obtained by computing the gradients and Hessians of the excited state FFs at the molecular structures visited by the GWPs. This protocol is able to account for the effect of possible anharmonicities in the FFs that originate from the redundancies of the set of internal coordinates adopted to parameterize the FFs, which, therefore, are not strictly speaking harmonic. In more general cases, when dealing with flexible systems, the QMD-FFs can contain explicitly anharmonic terms.

Before concluding this section, we note that, in principle, the diabatic couplings, $\lambda^{(i,j)}$, make the diabatic and adiabatic Hessians differ along A″ non-total symmetric coordinates (the only ones for which $\lambda^{(i,j)}$ are non-vanishing). As they are fitted to TD-DFT data, the Hessians of the FFs are actually adiabatic in nature. They should thus be corrected by differentiation of the analytical adiabatic PES arising from the QVC model so as to obtain truly "diabatic" Hessians.[57] Such a refinement has not been taken here due to the exploratory nature of the current work and the fact that the FFs' Hessians feature all positive frequencies and, therefore, do not introduce any problematic effects into the dynamics (see also the supplementary material).

In addition, by not updating the parameters of the Hamiltonian in Eq. (30), i.e., taking their values at the S_0 geometry ($t = 0$), one obtains a standard QVC model, which is useful since a benchmark ML-MCTDH calculation can be run to compare to those obtained with the hierarchy of approximated methods discussed in Sec. II B. In the following, we refer to this Hamiltonian as QVC_0.

E. Workflow

The workflow of the proposed FF-QD approach is illustrated in Fig. 2. The first two steps (*Vibronic Coupling Model* and *Generate Force Field*) generate the force field for each electronic state i. This will provide the terms contained in Eqs. (30) and (31) during the propagation. In our implementation, we have generated these intramolecular force fields via the JOYCE protocol, outlined above. Since these QMD-FFs are diabatic, the vibronic couplings $\lambda^{(i,j)}$ between the states are a separate component, and we have evaluated them using the diabatization scheme implemented in the *Overdia* code.[56,58,59] In principle, any diabatization scheme could be employed, and these vibronic coupling terms are set at the beginning of the calculation based on the reference geometry of the molecule and not updated.

With this in place, one can set up a calculation using the QUANTICS program, setting the desired initial conditions and simulation parameters. As illustrated by Fig. 2, the initial geometry is used by GROMACS to calculate the single point potential energies, atomic forces, and second derivatives (Hessian) for each electronic state (i.e., using the appropriate QMD-FF for each state). These data are stored in a database and provides the information necessary to propagate the GWPs in QUANTICS. When necessary, a new dataset is generated using GROMACS and stored in the database. This is the standard *on-the-fly* direct dynamics procedure, but instead of using electronic structure calculations, the energies, forces, and Hessians are generated by molecular mechanics calculations using the QMD-FFs. The output of the quantum dynamics simulations using this method is, however, identical to any other calculation using the DD-vMCG method and can be analyzed in the same way (e.g., diabatic state populations, PESs, and wavepacket trajectories).[19]

F. Details of calculations

1. Note on nomenclature

This work will compare a number of different QD calculations and, hence, it is important to define a general nomenclature to ensure clarity in the reported results. The following convention provides a general way to describe each calculation: $M//(T_1 f_1/N_1, T_2 f_2/N_2, \ldots, T_n f_n/N_n)$, where

- M = QD method (vMCG or DD-vMCG),
- T = Type of GWP trajectory (V = variational, C = classical, and S = classical and using the saddle point approximation),
- f = Number of DoFs (normal modes),
- N = Size of the Gaussian basis (nr. of GWPs),
- n = Number of sets of combined DoFs in a partitioned calculation.

2. Benchmarking of the QMD-FF derived operator

To test the validity of using a Gaussian basis in simulations with the QMD-FF, different vMCG calculations were benchmarked

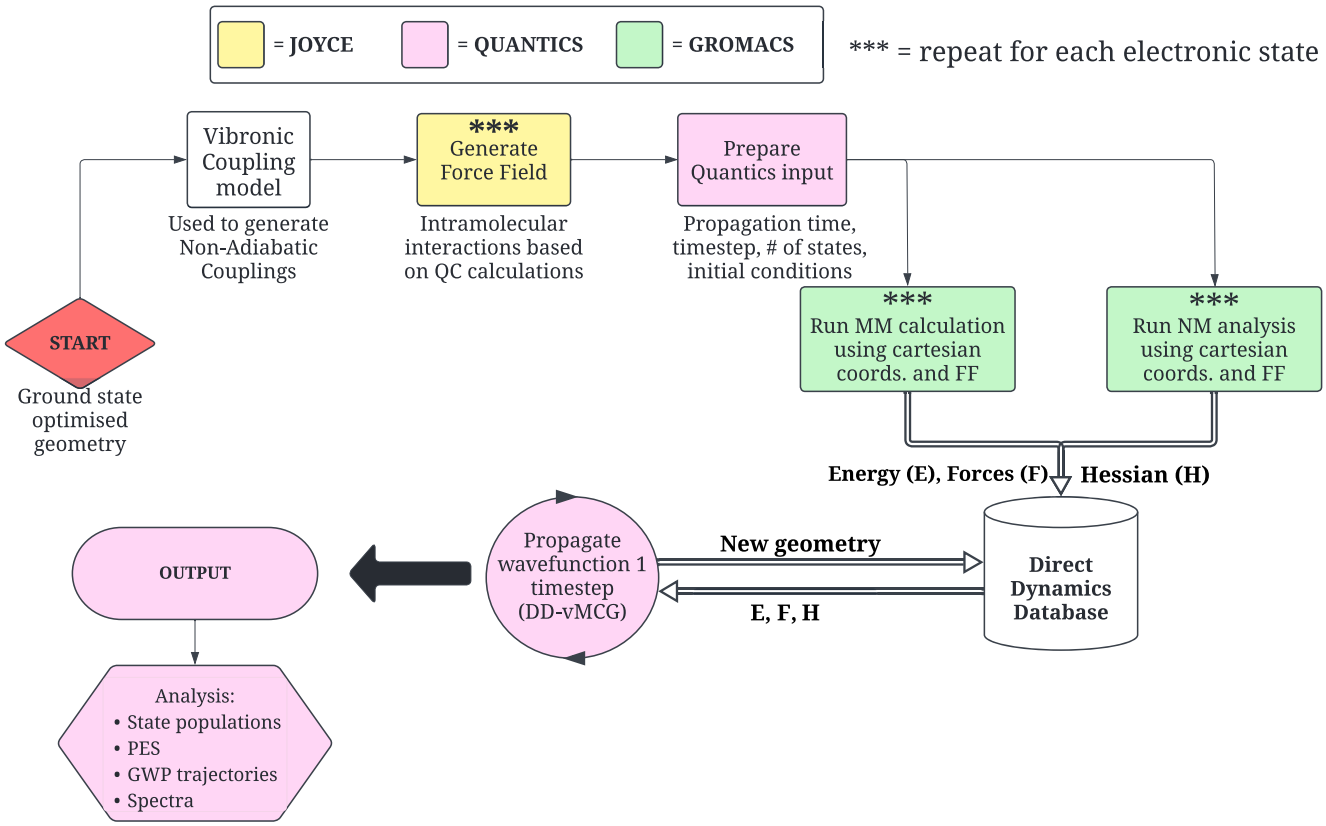

FIG. 2. Flowchart outlining the steps needed to set up the calculations, as well as a schematic representation of the propagation using a database. (FF = Force Field, MM = Molecular Mechanics, and NM = Normal Mode).

against a Multi-Layer (ML)-MCTDH result using the QVC_0 Hamiltonian. This was obtained by generating a Hamiltonian operator [Eq. (29)] at the initial time (i.e., before any on-the-fly updating takes place). As this Hamiltonian only includes terms up to second-order, ML-MCTDH and vMCG calculations converge to the same result when the LHA is used.[18] The linear parameters for this Hamiltonian are given in the supplementary material, and the quadratic parameters are given in the available datasets in the form of a QUANTICS operator file.

Since convergence depends on the size of the basis used, it is important to consider the computational effort of the simulations. Uracil has 30 normal modes, all of which will be included in the calculations, and this can become computationally expensive since the scaling is strongly dependent on the number of DoFs. While this number of modes is on the limit of what the standard MCTDH method is able to handle, ML-MCTDH potentially offers a significant reduction in the central processing unit (CPU) time for systems of this size. In this approach, the system is partitioned into branches and layers, leading to a contraction of the representation of the wavefunction.[40] For efficiency, it is important that modes that are strongly coupled are kept close in the branching scheme to treat the correlation as directly as possible. If this is not done, many basis functions may be required.

As described in Sec. II B, a similar approach can be taken with vMCG calculations, combining normal modes into a single mode, described by a set of GWPs. Three vMCG calculations will be compared to the ML-MCTDH benchmark in Sec. III, each of which has been partitioned differently.

The choice of which modes to combine into a set when partitioning the system will affect the speed and convergence of a calculation significantly. It is advisable to group together the modes that lead to large vibronic coupling (so-called coupling modes, characterized by large λ parameters) and the modes that cause a large shift in the minimum of the excited states away from the Franck–Condon point (called tuning modes, characterized by large κ parameters). These modes will be the most important during the dynamics. The remaining modes (spectator modes) can then be grouped together.

From an analysis of the parameters of the QVC_0 Hamiltonian, ten modes were identified as the strongest coupling and tuning modes (see the supplementary material, Sec. S3.2, for details). These modes were grouped together in the first partition. The remaining 20 spectator modes were either divided into two groups of 10 or taken together. Three calculations were thus carried out:

- vMCG//(V30/61): No partitioning, 61 30-dimensional GWPs.
- vMCG//(V10/61, V10/61, V10/61): Partition into three sets of 61 10-dimensional GWPs.

- vMCG//(V10/61, V20/41): Partition into two sets: 61 10-dimensional and 41 20-dimensional GWPs.

This partition is analogous to the separation of modes into different branches of the ML-MCTDH tree, and indeed, the partitioning in this calculation aimed to mimic the structure of the ML-MCTDH tree, which puts the key modes on one branch (see Fig. S1 in the supplementary material). One can think of this type of partition as a separation between the "system," consisting of the modes that are important during the propagation, and a "bath," which contains the modes that have a weaker effect on the dynamics and, hence, need fewer basis functions. To check that the key ten modes were the correct choice, a final vMCG calculation was run, but only including these modes:

- vMCG//(V10/61): No partition, 61 10-dimensional GWPs.

In Sec. II B, the clMCG and saddleMCG approximations were introduced. With these approximations, instead of following the variational trajectories of vMCG, the GWPs follow classical trajectories, and in the saddleMCG approximation, the integrals are only done to zeroth order (i.e., considering only the potential energies). It is also possible that different partitions of the total wavepacket can be treated at different levels of approximation. Thus, it is possible to treat the key modes with a variational vMCG basis, and the less important modes can be treated with the clMCG or saddleMCG approximations. The use of different approximations for different sets of normal modes gives rise to a hybrid between a full quantum vMCG and a (zeroth order) classical propagation aiming to retain the accuracy of the full vMCG picture with increased efficiency.

The effect of using different approximations for the propagation of the GWPs was thus investigated with the following simulations, performed using the two set partition:

- vMCG//(C10/61, C20/41): All GWPs propagated on classical trajectories.
- vMCG//(S10/61, S20/41): All GWPs propagated on classical trajectories using, additionally, the saddle point approximation.
- vMCG//(V10/61, C20/41): A set of 10-D GWPs following variational trajectories and a set of 20-D GWPs following classical trajectories.
- vMCG//(V10/61, S20/41): A set of 10-D GWPs following variational trajectories and a set of 20-D GWPs following classical trajectories and employing the saddle point approximation.

3. DD-vMCG calculations using Quantics–Gromacs interface

The new interface between QUANTICS and GROMACS was tested by benchmarking the results against the most converged vMCG calculation [i.e., vMCG//(V10, V20), see Sec. III in the following]. However, it is not expected that the DD-vMCG results should converge to the vMCG result since the propagation of the GWPs includes the *on-the-fly* calculation of the diabatic potentials, the atomic gradients, and the excited state Hessian matrices. Essentially, this allows for anharmonicity to be introduced in the calculations thanks to the FFs being expressed in a redundant set of curvilinear coordinates (see Sec. II D). Several different DD-vMCG calculations were run, using the same types of variations as the previously described vMCG calculations:

- DD-vMCG//(V10/61, V20/41).
- DD-vMCG//(C10/61, C20/41).
- DD-vMCG//(S10/61, S20/41).
- DD-vMCG//(V10/61, C20/41).
- DD-vMCG//(V10/61, S20/41).

Importantly, the wavepackets in these simulations are all propagated on the exact same potential surfaces, with the same gradients and Hessians at every point to ensure a proper comparison. This is achieved by first performing the different calculations in a typical *on-the-fly* manner (meaning the data defining the potentials is generated during the propagation itself) and, subsequently, collecting these data in one database. Ultimately, the calculations are run without generating any more points but solely using this single database, which contains just under 10 000 points, which defines the potential energy landscape.

III. RESULTS AND DISCUSSION

A. Force field

To validate the parameterized QMD-FF intra-molecular terms, a number of tests were devised. First, the QMD-FF Hessian matrix, $\mathbf{H}^{s}_{\mathrm{QMD-FF}} = \frac{\partial^2 E^{\mathrm{QMD-FF}|s}_{intra}}{\partial q^2}$, can be exploited for each state s to perform a MM geometry optimization.

In the top panel of Fig. 3, the optimized geometries obtained at the QM and MM levels for each state s are compared, and it is evident that the QM and MM structures overlap well. The visual analysis was confirmed by computing the standard deviation between QM and MM optimized structures, which amounts for all states to less than 10^{-3} Å and less than 0.1° for bond lengths and bond angles, or dihedrals, respectively. Next, the QMD-FF normal modes can be computed together with the associated vibrational frequencies, and the results are eventually compared to the QM parent data employed in the JOYCE procedure. The comparison is displayed in the bottom panels of Fig. 3 for the three states. Again, the FF can reproduce the QM data with only minor deviations for the low frequency modes. The complete set of intra-molecular parameters defining the FF (i.e., bond-stretching, angle-bending, and torsional parameters) can be found in the supplementary material.

B. Benchmarking vMCG

1. Effect of partitioning

Simulations using ML-MCTDH and vMCG with different partitionings were run as a benchmark to validate the proposed approach, i.e., using parameterized FFs in combination with the vMCG method. The comparison will be done based on the electronic state population transfer from the initially excited $\pi\pi^*$ state to the $n\pi^*$ state of uracil (note: only the decay of the population of the $\pi\pi^*$ state will be plotted, with all the transfer going to the $n\pi^*$ state, as the S_0 state is not included in the dynamics). All simulations were run for 100 fs, using the exact same Hamiltonian operator derived from the FF (QVC_0). The results of this 2-state system are summarized below. Figure 4 compares the ML-MCTDH calculation and the four vMCG calculations using different partitions, as outlined in Sec. II F 2.

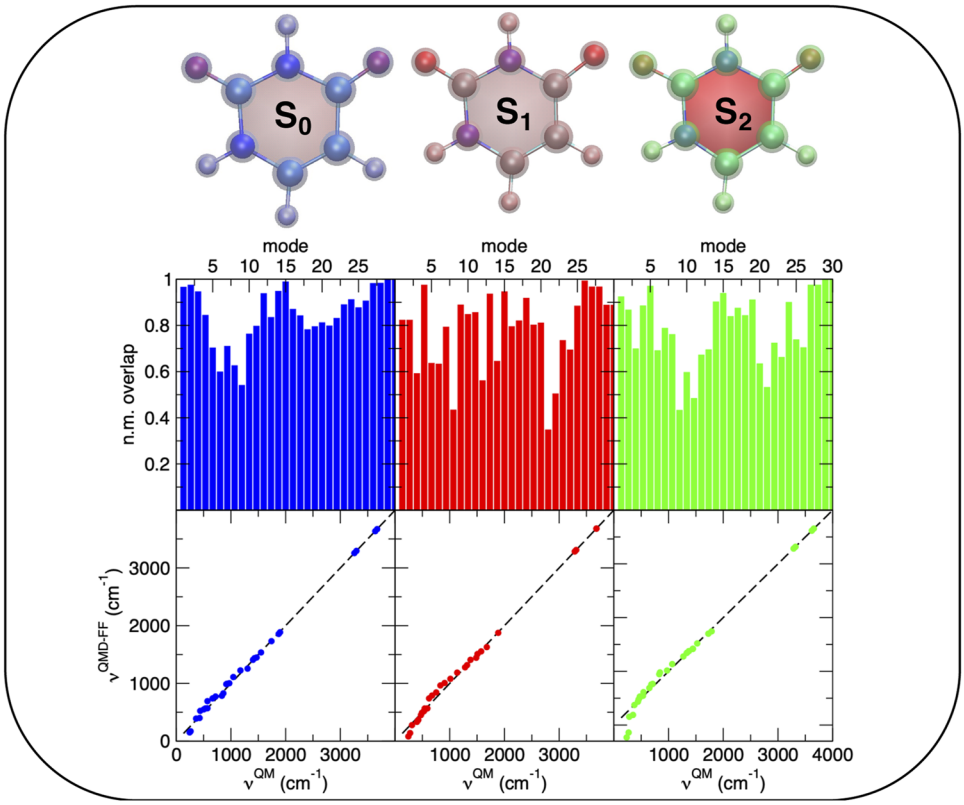

FIG. 3. Top panels: overlap of the QM (solid colored spheres) and QMD-FF (transparent blue, red, and green spheres for S_0, S_1, and S_2) optimized geometries. Middle panels: overlap of QM and MM normal modes. Bottom panels: correlation plot between QM and MM vibrational frequencies computed by JOYCE. From left to right, results are displayed for S_0 (blue), S_1 (red), and S_2 (green), respectively.

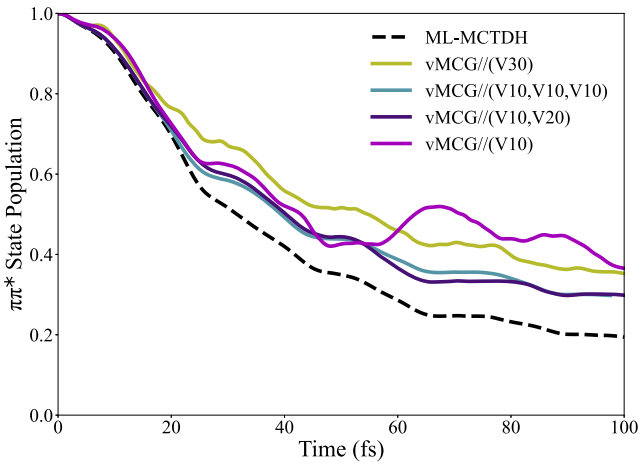

FIG. 4. Calculated population transfer for uracil: ML-MCTDH benchmark (black, dashed), vMCG with 61 GWPs (green), vMCG with three sets of GWPs (ten modes and 61 GWPs in each) (blue), vMCG with two sets of GWPs (ten modes with 61 GWPs and 20 modes with 41 GWPs) (purple), and vMCG with only the ten most important normal modes included (magenta).

Using the population transfer as a measure of convergence, we conclude that the vMCG calculations are not fully converged but approach the ML-MCTDH result as the number of basis functions increases, as expected. By comparing an unconverged ML-MCTDH calculation with the fully converged result, we confirmed that the lack of convergence in a simulation does indeed lead to a slower population transfer (see the supplementary material, Fig. S2). The QVC_0 operator does not have very strong non-adiabatic coupling, and the coupling is spread over five modes (Q_1, Q_{10}, Q_{11}, Q_{13}, and Q_{15}) (see parameters in the supplementary material). This makes it a system that is hard to converge as there is no dominant set of modes with a strong conical intersection, but rather a number of modes need to be described well to capture the weak coupling.

For the vMCG calculations, the aim is to provide the largest basis set at the lowest computational cost. Here, various factors play off against each other: low dimensional GWPs are cheap, but if many partitions are used, the need to calculate mean-field operators becomes expensive. The vMCG//V30 calculation with only 61 basis functions is the smallest. It has no mean-fields, but the 30-dimensional basis functions are expensive and restrict the basis size. The vMCG//V10/V10/V10 calculation has small (10-dimensional) GWPs and a large basis with 61^3 configurations, making it the largest calculation but also the most expensive due to the many mean-fields.

The vMCG//V10/V20 calculation is a compromise with 61 × 41 configurations but larger 20-dimensional GWPs to propagate.

It is noteworthy that even though the calculation with three sets of 61 10-dimensional GWPs [vMCG//(V10, V10, V10)] contains many more functions than the one with only two sets of GWPs [vMCG//(V10, V20)], the results are essentially the same. It is, of course, desirable to use as few basis functions as possible since the computational effort will be reduced. In this case, the smaller number of basis functions in the 2-set calculation compared to the 3-set calculation led to a massive reduction in simulation time (Table I). The single partition V30 calculation is also seen to be much slower than the V10/V20 calculation, despite the smaller basis due to the large GWPs. Evidently, this makes it clear that partitioning the system in an optimal way can achieve the accuracy of a much larger set of basis functions but at a significantly lower computational cost. The fact that the vMCG//(V10, V10, V10) simulation is very similar to the vMCG(V10, V20) calculation shows that the latter basis set describes the 20 bath modes well. The vMCG(V10, V20) will now be taken as the benchmark to test other (DD)-vMCG results.

Another obvious way to reduce the cost of calculations is to reduce the dimensionality of the system that is being simulated. When only the ten most important normal modes were included (i.e., a threefold reduction in dimensionality from 30-D to 10-D), the CPU time was reduced by a factor of around 7, compared to the vMCG//(V10, V20) calculation as shown in Table I. However, while the state populations of this simulation agree well with the benchmark in the first ~40 fs, the agreement subsequently worsens. This deviation has to be attributed to the missing 20 normal modes in the calculation, hence providing evidence that these do indeed play some role in the dynamics and, if possible, should be included in all calculations.

In summary, these results establish that the FFs used to describe the excited states of uracil are suitable for use in the vMCG method and that, even though the calculations are not fully converged, they can reproduce the accuracy of the ML-MCTDH results reasonably well. It can be seen that the full-dimensionality vMCG calculations are all slower than ML-MCTDH. This is due to the difference in scaling between the two methods, but a significant computational advantage is expected when using vMCG-based methods on larger systems due to the more realistic description of the potential.[18]

TABLE I. Computational effort of the reported quantum dynamics simulations in terms of CPU time.

Calculation type	CPU time (h)
ML-MCTDH	12.7
vMCG calculations (vMCG//):	
(V30)	494
(V10, V10, V10)	>1000
(V10, V20)	30
(V10)	4.5
vMCG with classical approx., compare to (V10, V20):	
(C10, C20)	11
(S10, S20)	10
(V10, C20)	18
(V10, S20)	16
Direct dynamics calculations (DD-vMCG//):	
(V10, V20)	140
(C10, C20)	57
(S10, S20)	24
(V10, C20)	88
(V10, S20)	36

2. Effect of classical approximations

Next, the effect of using different levels of approximation (outlined in Sec. II F 2) during the propagation is investigated. The simulations, benchmarked against the vMCG//(V10, V20) result in terms of state population transfer, are plotted in Fig. 5. Using only classical GWPs results in slower and reduced population transfer than variational GWPs. This is due to the slower convergence compared to the variational GWPs of vMCG. The saddle point approximation clearly reduces the quality of the description further. The reduction in accuracy is, however, balanced out by a speed-up of the calculations, requiring fewer computational resources (see Table I).

Figure 5 also shows that the hybrid calculations vMCG//(V10, C20) and vMCG//(V10, S20) reproduce the full vMCG//(V10, V20) simulation, indicating that although the role of the 20 spectator modes is not marginal, they can be properly described on a classical GWP basis. Table I clearly shows how the computational cost decreases when using the more approximate approaches, with the vMCG//(V10, S20) taking half the time of the full vMCG calculation. This finding provides a justification for the use of this approach to significantly speed up calculations. This procedure can, therefore, be

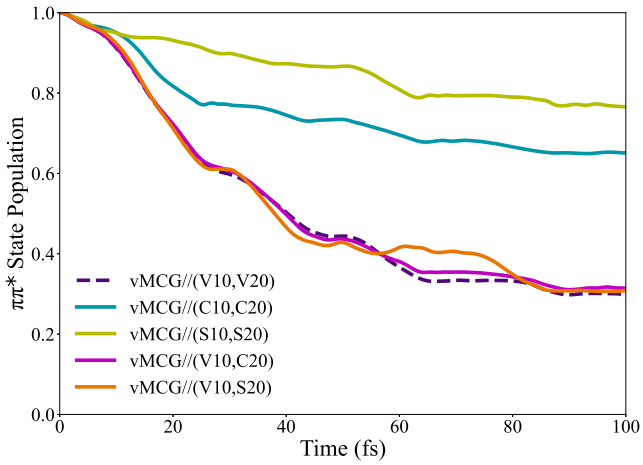

FIG. 5. Calculated population transfer for uracil: vMCG benchmark (purple, dashed), vMCG with classical trajectories (blue), vMCG with classical trajectories and using the saddle point approximation (green), hybrid calculations: vMCG with 61 GWPs following variational trajectories and 41 GWPs following classical ones (magenta), and vMCG with 61 GWPs following variational trajectories and 41 GWPs following classical ones and using the saddle point approximation (orange).

applied to much larger molecular systems, although it is clear that the accuracy of the results will require a good choice of partitioning of the normal modes.

C. Benchmarking direct dynamics

Here, the benchmarking of the *on-the-fly* vMCG results (i.e., DD-vMCG) is presented. The population transfer over 100 fs is again compared to the vMCG//(V10, V20) result. A trivial error-checking comparison for the implementation of the QUANTICS/GROMACS interface between this previously presented vMCG result and a DD-vMCG simulation where the database contains information only on the FC point and is not subsequently updated *on-the-fly* is shown in the supplementary material (Fig. S4). As expected, these calculations are identical since the QVC_0 parameters are derived from this single-point database. Now we move on to discuss the effect of updating the PES.

1. Non-updated vs updated potential surfaces

Thanks to the way the FF is set up (i.e., redundant, curvilinear coordinates), anharmonicity is included in the potentials. As the surfaces are updated, normal modes can deviate from their initial harmonicity as the energies, gradients, and Hessians are calculated at different points during the propagation of the GWPs. This is shown in Fig. 6, where cuts along the PESs of two normal modes are compared before and after updating, as well as a comparison of the Minimum Energy Conical Intersection (MECI). This intersection is presented here as a cut along the multidimensional potential energy landscape, from the FC point (at zero) in a direction that can be characterized by a combination of the following modes: in-plane N_1–H_{N1} and $C_{5/6}$–$H_{5/6}$ wagging modes, the ring breathing mode, and the C_5–C_6 stretching mode (see the supplementary material for full description and coordinates). It is also worth pointing out that the MECI is located at a significant distance away from the FC point, resulting in an initially slow rate of population transfer in the first 10 fs.

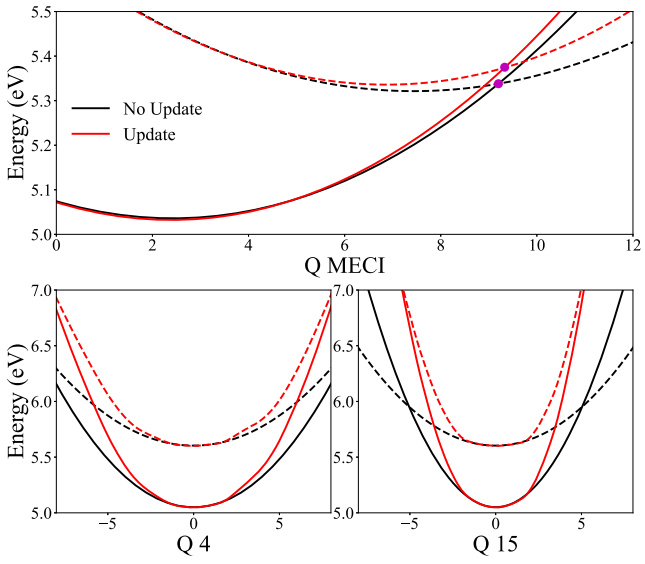

FIG. 6. Potentials of the $n\pi^*$ (solid) and $\pi\pi^*$ (dashed) states of uracil after DD-vMCG simulations both with and without *on-the-fly* updating, FC point located at Q = 0. Top: cut showing the MECI (marked by a purple dot). Bottom: non totally symmetric modes 4 and 15 (see the supplementary material for description).

Looking now at the shape of the surfaces in Fig. 6, a general trend is observed: the potential wells become steeper when updating is used. For example, the crossings of the surfaces along mode 15 are shifted up in energy. Figure 6 also shows the MECI being shifted slightly higher in energy by about 0.05 eV when updating the potential. As intersections between states become energetically less accessible, it is of course expected that the population transfer between these states will become less efficient. A second notable difference between the non-updated and updated surfaces can be seen in the shape of the potentials along mode 4, for example. The wells have become more quartic in nature, hinting at the importance of bi-linear vibronic coupling terms ($\gamma^{(i)}_{\alpha,\beta}$ in Sec. II D).

2. Effect of updated surfaces and classical approximations on dynamics

We now examine how the FF-QD method proposed in this work performs. By again looking at the state population dynamics, the effect of updating the surfaces will be apparent, and more importantly, the relative performance of the different hierarchies of approximations will be compared (as was done for the vMCG calculations, Sec. III B 2). The data are plotted in Fig. 7. A comparison with the reference vMCG//(V10, V20) computations in the upper panel shows that including anharmonicities through the updating of the surfaces leads to a lower population transfer within the simulation time. This result was expected since, as previously stated, the points of intersection between the $\pi\pi^*$ and $n\pi^*$ states are shifted higher in energy.

The bottom panel in Fig. 7 compares the performance of the different classical approximations. In this case, the reference result is the fully variational computation, DD-vMCG//(V10, V20). As for vMCG, pure classical trajectories perform poorly and underestimate the population transfer, especially when the saddle-point approximation is also used. Hybrid methods work much better, with (V10, C20) being practically equivalent to (V10, V20). When the dynamics of the 20 bath modes use the saddle-point approximation (V10, S20), however, the agreement is worse, predicting, in this case, a slightly accelerated population transfer. Compared to the vMCG results, this indicates an increased sensitivity to this approximation when surfaces are updated. Interestingly, the observation that the hybrid calculations overestimate the population transfer indicates that the 20 bath modes are, at least partly, responsible for the slower rate of transfer of the fully variational result, once again hinting at their relative importance for accurately describing the dynamics of uracil.

The CPU times for the simulations are given in Table I. The updating procedure means that the full DD-vMCG//(V10, V20) calculation takes four times as long as the comparable calculation on the analytic QVC_0 potential. This is the overhead due to the use of the Shepard interpolation of the database points. The speed up using the classical GWPs is also clear, with the DD-vMCG//(V10, S20) calculation taking 25% of the time of the full calculation.

In summary, the various DD-vMCG results presented here show the validity of the described FF-QD methodology, with *on-the-fly* calculations showing the effect of updating the PESs and also the

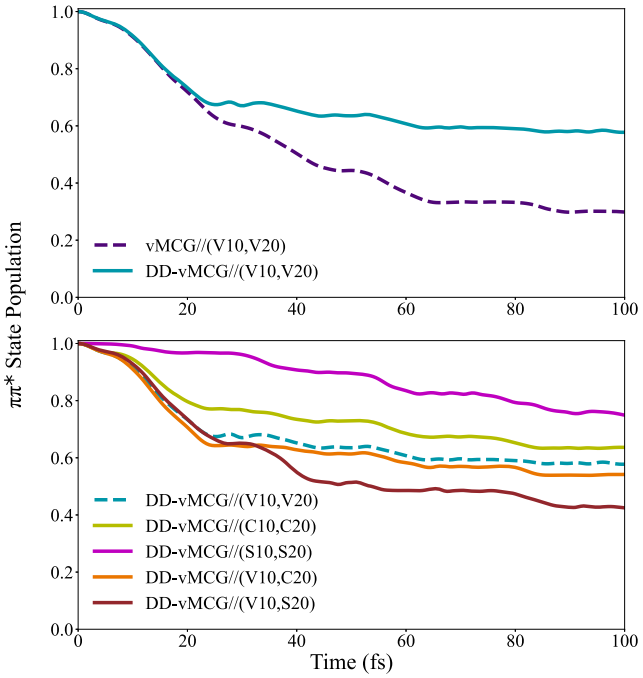

FIG. 7. Calculated population transfer for uracil: Top: vMCG benchmark (purple, dashed) and full variational DD-vMCG (blue). Bottom: classical approximations, compared to the full variational DD-vMCG benchmark (blue, dashed), all classical (green), all classical and using the saddle point approximation (magenta), hybrid variational + classical (orange), and hybrid variational + classical and using the saddle point approximation (brown).

performance of using different classical approximations is evaluated. Applied to the population transfer dynamics of a simple two-state model of uracil, all results obtained by the FF-QD approach can be rationalized both in terms of the chemistry of the molecule as well as the computational performance of the calculations. The quality of the state populations along with savings in computer time demonstrates the potential utility of hybrid simulations.

IV. CONCLUSION AND OUTLOOK

This work is a proof of principle for a route to the quantum dynamics simulation of large and solvated photo-excited molecules, and we have introduced two new developments to achieve this aim. First, we have addressed the problem of providing excited state PESs that can be interfaced with the treatment of solvent via QMD-FFs. The JOYCE parameterized QMD-FF is shown to reproduce the *ab initio* parameterized vibronic coupling model potential, allowing FF-QD simulations to have the accuracy of quantum chemistry.

This opens up a route to QD simulations in such scenarios going beyond the present state-of-the-art, in which the effect of the environment is usually added to a QD simulation through simple baths of harmonic oscillators.[1] The extension to adding a solvent is straightforward. However, the choice of solvent potentials is crucial. For instance, it is known that standard water models, such as TIP3P, do not provide a good description of the solvent for quantum dynamics simulations.[60] This exploration is for future work.

Second, we have introduced a hierarchy of approximations to the vMCG method to allow efficient partitioning of the wavefunction and permit different dynamical treatments for portions that have greater or lesser influence on the dynamics. The nuclear wavefunction partitioning and approximation hierarchy are also shown to be able to speed up calculations. Recognizing the key modes for the quantum process of internal conversion and treating them using variationally coupled GWPs keeps the basis set small and keeps the system in the correct parts of phase space so that the less important "bath" modes can follow classical trajectories with little loss of accuracy and with a saving of effort. When moving to larger systems, the saddle-point approximation will also prove useful. While it is clear that using the saddle-point approximation leads to a loss of accuracy, it results in a speed up of the calculations, which will be even more significant in the context of extending this approach to much larger systems. The fact that Hessians are not required for the modes treated in this way is important, as requiring full Hessians for hundreds of atoms would be prohibitive. As a result, the loss of accuracy will be potentially justified as the correct qualitative behavior is retained. A 100 fs full quantum dynamics simulation for a 30 mode molecule with flexible potentials that are sampled *on-the-fly* in 36 h with a single-core computation is a good starting point for future developments.

In the context of other approaches to including solvent effects in dynamics, it is interesting to note that very recently, in the framework of trajectory surface hopping simulations, it has been proposed to couple a linear vibronic coupling (LVC) Hamiltonian with a MM description of the solvent.[61] In that work, the solute/solvent electrostatic interactions were described with multipole expansions, beyond simple point charge models. The surface hopping ansatz allows for very efficient simulation, and although a much smaller system was considered in that work than this (single state, six normal modes), nanosecond simulation time was reached. As we have highlighted in this work, however, a quantum treatment of important normal modes can be significant in the accuracy of the simulations.

Our simulations confirm that for uracil in the gas phase, the population transfer from the bright S$\pi\pi^*$ to the dark S$n\pi^*$ state is fast and efficient. Within 100 fs, the population on S$n\pi^*$ is $\geq$0.4. Changing the conical intersection by including anharmonic effects appears to significantly modulate the population of S$n\pi^*$, providing another hint of the importance of proper treatment of quantum nuclear effects. The fast population transfer to S$n\pi^*$ is consistent with that provided by a recent surface hopping study based on extended multi-state (XMS)-CASPT2 calculations.[32] We note, however, that since the possible decay to S_0 is not considered in our study, it cannot provide a complete picture of the uracil photophysics in the gas phase. In future studies, it will thus be important to explore how the population dynamics of uracil depend on the description of the coupled potential energy surfaces as well as the dynamical method used.

A property of the uracil system that becomes clear when one analyzes the surfaces from the simulations presented here is that there is no obvious single coordinate that leads to efficient population transfer between the S$\pi\pi^*$ and S$n\pi^*$ states. The intersection is instead an extended seam, with different possible paths contributing to the overall transfer, depending on their energetic accessibility and the strength of the non-adiabatic coupling. Since many of the

intersections are high in energy (see, for example, the crossing point along mode 15 in Fig. 6) and the coupling between the states is generally not very strong, the mechanism of population transfer proceeds via many different routes. It is this fact that makes uracil a rather hard system to model since a full dimensional description of the potential is desirable to correctly capture the long time dynamics as well as the accurate inclusion of coupling terms. This example hence emphasizes the need for efficient and flexible non-adiabatic QD methods that can deal with the subtle interplay between the interactions that govern the overall dynamics of the system.

In summary, we believe that the advances in this work set the stage for the treatment of solvent effects in quantum dynamics simulations, with work ongoing that will be detailed in a future publication. We also believe in the importance of including nuclear vibrational quantum effects to give an accurate picture of the photoactivated dynamics, and that by pairing this with classical approximations for less important modes, we can open up the route to accurate simulation of large molecules in realistic environments.

SUPPLEMENTARY MATERIAL

Information supporting the work presented is available in the form of the supplementary material, which includes the ML-MCTDH tree structures along with a convergence study, the QMD-FF parameters, LVC Hamiltonian parameters, and a section on how solvents will be included in future studies.

ACKNOWLEDGMENTS

The research in this paper was a result of an exchange program funded by the Royal Society (Grant No. IEC/R2/202236) and the CNR (Bilaterale Grant No. CNR-RSC 2021). R.I., G.P., and F.S. also thank the CNR program "Progetti di Ricerca @ cnr," Project No. UCATG4, and R.I. also acknowledges the financial support through Grant No. NUTRAGE FOE-2021 DBA.AD005.225. F.S. and G.P. also thank the ICSC - Centro Nazionale di Richercha in High Performance Computing, Big Data and Quantum Computing, funded by the Next Generation EU - PNRR, Missione 4 Componente 2 Investimento 1.4. S.G. acknowledges the María Zambrano grant (NextGenEU funds), the USAL grant "Programa Propio C1," and the funding by the Spanish Ministry of Science and Innovation (MCIN/AEI/10.13039/501100011033), Grant No. PID2020-113147GA-I00. The work has also been supported in the UK by the EPSRC under Grant No. EP/V026690/1. J.A.G. gratefully acknowledges the current funding of a research fellowship by the Alexander von Humboldt Foundation.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

L. L. E. Cigrang: Investigation (equal); Methodology (equal); Software (equal); Writing – original draft (equal); Writing – review & editing (equal). **J. A. Green**: Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Resources (equal); Software (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **S. Gómez**: Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **J. Cerezo**: Investigation (equal); Methodology (equal); Software (equal); Writing – original draft (equal); Writing – review & editing (equal). **R. Improta**: Investigation (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **G. Prampolini**: Investigation (equal); Methodology (equal); Software (equal); Writing – original draft (equal); Writing – review & editing (equal). **F. Santoro**: Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Resources (equal); Software (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **G. A. Worth**: Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data from the simulations in the form of Quantics input and output files are available as zipped tar files from the UCL Research Data Repository at http://doi.org/10.5522/04/25236955.

REFERENCES

[1] F. Santoro, J. A. Green, L. Martinez-Fernandez, J. Cerezo, and R. Improta, Phys. Chem. Chem. Phys. **23**, 8181 (2021).
[2] M. Beck, A. Jäckle, G. Worth, and H.-D. Meyer, Phys. Rep. **324**, 1 (2000).
[3] *Multidimensional Quantum Dynamics: MCTDH Theory and Applications*, edited by H.-D. Meyer, F. Gatti, and G. A. Worth (Wiley-VCH, Weinheim, 2009).
[4] H. Wang and M. Thoss, J. Chem. Phys. **119**, 1289 (2003).
[5] H. Wang, J. Phys. Chem. A **119**, 7951 (2015).
[6] G. Worth, Comput. Phys. Commun. **248**, 107040 (2020).
[7] J. C. Tully, J. Chem. Phys. **93**, 1061 (1990).
[8] T. R. Nelson, A. J. White, J. A. Bjorgaard, A. E. Sifain, Y. Zhang, B. Nebgen, S. Fernandez-Alberti, D. Mozyrsky, A. E. Roitberg, and S. Tretiak, Chem. Rev. **120**, 2215 (2020).
[9] S. Mai, P. Marquetand, and L. González, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **8**, e1370 (2018).
[10] R. Crespo-Otero and M. Barbatti, Chem. Rev. **118**, 7026 (2018).
[11] M. Barbatti, M. Ruckenbauer, F. Plasser, J. Pittner, G. Granucci, M. Persico, and H. Lischka, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **4**, 26 (2014).
[12] M. Ben-Nun, J. Quenneville, and T. J. Martínez, J. Phys. Chem. A **104**, 5161 (2000).
[13] B. F. E. Curchod and T. J. Martínez, Chem. Rev. **118**, 3305 (2018).
[14] D. V. Shalashilin and M. S. Child, Chem. Phys. **304**, 103 (2004).
[15] J. A. Green, A. Grigolo, M. Ronto, and D. V. Shalashilin, J. Chem. Phys. **144**, 024111 (2016).
[16] D. V. Makhov, C. Symonds, S. Fernandez-Alberti, and D. V. Shalashilin, Chem. Phys. **493**, 200 (2017).
[17] G. A. Worth, M. A. Robb, and I. Burghardt, Faraday Discuss. **127**, 307 (2004).
[18] G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt, and B. Lasorne, Int. Rev. Phys. Chem. **34**, 269 (2015).
[19] S. Gómez, E. Spinlove, and G. A. Worth, Phys. Chem. Chem. Phys. **26**, 1829 (2024).
[20] I. Cacelli and G. Prampolini, J. Chem. Theory Comput. **3**, 1803 (2007).

[21] J. Cerezo, G. Prampolini, and I. Cacelli, Theor. Chem. Acc. **137**, 80 (2018).
[22] N. De Mitri, S. Monti, G. Prampolini, and V. Barone, J. Chem. Theory Comput. **9**, 4507 (2013).
[23] G. Prampolini, F. Ingrosso, A. Segalina, S. Caramori, P. Foggi, and M. Pastore, J. Chem. Theory Comput. **15**, 529 (2019).
[24] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, SoftwareX **1–2**, 19 (2015).
[25] R. Improta and T. Douki, *DNA Photodamage: From Light Absorption to Cellular Responses and Skin Cancer* (Royal Society of Chemistry, 2021).
[26] M. Barbatti, C. A. Borin, and S. Ullrich, *Photoinduced Phenomena in Nucleic Acids I: Nucleobases in the Gas Phase and in Solvents*, *Topics in Current Chemistry Vol. 355* (Springer International Publishing, Cham, Switzerland, 2015).
[27] C. T. Middleton, K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernández, and B. Kohler, Annu. Rev. Phys. Chem. **60**, 217 (2009).
[28] R. Improta, F. Santoro, and L. Blancafort, Chem. Rev. **116**, 3540 (2016).
[29] L. Martínez Fernández, F. Santoro, and R. Improta, Acc. Chem. Res. **55**, 2077 (2022).
[30] M. M. Brister and C. E. Crespo-Hernández, J. Phys. Chem. Lett. **10**, 2156 (2019).
[31] R. Borrego-Varillas, A. Nenov, P. Kabaciński, I. Conti, L. Ganzer, A. Oriana, V. K. Jaiswal, I. Delfino, O. Weingart, C. Manzoni *et al.*, Nat. Commun. **12**, 7285 (2021).
[32] P. Chakraborty, Y. Liu, T. Weinacht, and S. Matsika, Faraday Discuss. **228**, 266 (2021).
[33] T. J. A. Wolf, R. H. Myhre, J. P. Cryan, S. Coriani, R. J. Squibb, A. Battistoni, N. Berrah, C. Bostedt, P. Bucksbaum, G. Coslovich *et al.*, Nat. Commun. **8**, 29 (2017).
[34] D. Picconi, V. Barone, A. Lami, F. Santoro, and R. Improta, ChemPhysChem **12**, 1957 (2011).
[35] D. Picconi, A. Lami, and F. Santoro, J. Chem. Phys. **136**, 244104 (2012).
[36] J. A. Green, M. Y. Jouybari, D. Aranda, R. Improta, and F. Santoro, Molecules **26**, 1743 (2021).
[37] J. Cerezo, Y. Liu, N. Lin, X. Zhao, R. Improta, and F. Santoro, J. Chem. Theory Comput. **14**, 820 (2018).
[38] R. Improta, V. Barone, A. Lami, and F. Santoro, J. Phys. Chem. B **113**, 14491 (2009).
[39] M. S. Bezabih, D. S. Kaliakin, A. Blanco-González, L. Barneschi, A. N. Tarnovsky, and M. Olivucci, J. Phys. Chem. B **127**, 10871 (2023).
[40] O. Vendrell and H.-D. Meyer, J. Chem. Phys. **134**, 44135 (2011).
[41] E. J. Heller, J. Chem. Phys. **62**, 1544 (1975).
[42] E. J. Heller, J. Chem. Phys. **75**, 2923 (1981).
[43] G. A. Worth and I. Burghardt, Chem. Phys. Lett. **368**, 502 (2003).
[44] I. Burghardt, H.-D. Meyer, and L. S. Cederbaum, J. Chem. Phys. **111**, 2927 (1999).
[45] K. E. Spinlove, M. Vacher, M. Bearpark, M. A. Robb, and G. A. Worth, Chem. Phys. **482**, 52 (2017).
[46] M. Ben-Nun and T. J. Martínez, Adv. Chem. Phys. **121**, 439 (2002).
[47] C. S. M. Allan, B. Lasorne, G. A. Worth, and M. A. Robb, J. Phys. Chem. A **114**, 8713 (2010).
[48] B. Lasorne, M. A. Robb, and G. A. Worth, Phys. Chem. Chem. Phys. **9**, 3210 (2007).
[49] G. Christopoulou, A. Freibert, and G. A. Worth, J. Chem. Phys. **154**, 124127 (2021).
[50] G. W. Richings and G. A. Worth, J. Phys. Chem. A **119**, 12457 (2015).
[51] K. C. Thompson and M. A. Collins, J. Chem. Soc., Faraday Trans. **93**, 871 (1997).
[52] M. A. Collins, Theor. Chem. Acc. **108**, 313 (2002).
[53] T. J. Frankcombe, J. Chem. Phys. **140**, 114108 (2014).
[54] V. Barone, I. Cacelli, N. De Mitri, D. Licari, S. Monti, and G. Prampolini, Phys. Chem. Chem. Phys. **15**, 3736 (2013).
[55] I. Cacelli, J. Cerezo, N. De Mitri, and G. Prampolini, Joyce2.10, a Fortran 77 code for intra-molecular force field parameterization, available free of charge at http://www.iccom.cnr.it/en/joyce-2/ (last accessed September 2022).
[56] J. A. Green, M. Yaghoubi Jouybari, H. Asha, F. Santoro, and R. Improta, J. Chem. Theory Comput. **17**, 4660 (2021).
[57] D. Picconi, F. J. Avila Ferrer, R. Improta, A. Lami, and F. Santoro, Faraday Discuss. **163**, 223 (2013).
[58] M. Yaghoubi Jouybari, Y. Liu, R. Improta, and F. Santoro, J. Chem. Theory Comput. **16**, 5792 (2020).
[59] F. Santoro and J. A. Green, Overdia 01, a Fortran 90 code for parametrization of model Hamiltonians based on a maximum-overlap diabatisation, 2022, available at: http://www.iccom.cnr.it/en/overdia-en (last accessed 30 August 2022).
[60] S. Habershon, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **131**, 24501 (2009).
[61] S. Polonius, O. Zhuravel, B. Bachmair, and S. Mai, J. Chem. Theory Comput. **19**, 7171 (2023).