JCTC Journal of Chemical Theory and Computation

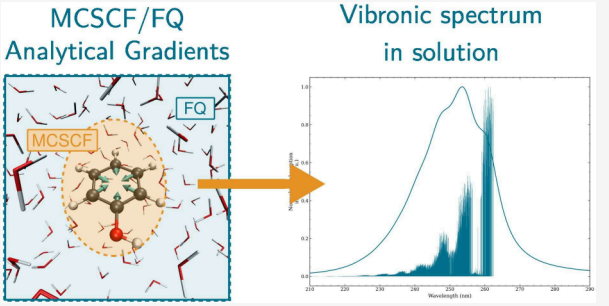
pubs.acs.org/JCTC Article

Analytical Nuclear Gradients for the Multiconfigurational Self-Consistent Field Method Coupled with the Polarizable Fluctuating Charges Model

Francesco Mazza, Marco Trinari, Chiara Sepali, and Chiara Cappelli*

Cite This: https://doi.org/10.1021/acs.jctc.5c01890 Read Online

ABSTRACT: The multiscale model combining the multiconfigurational self-consistent field (MCSCF) method with the fully atomistic polarizable Fluctuating Charges (FQ) force field (Sepali, C.; et al. *J. Chem. Theory Comput.* **2024**, *20*, 9954−9967) is here extended to the calculation of analytical nuclear gradients. The gradients are derived from first-principles, implemented in the OpenMolcas package, and validated against numerical references. The resulting MCSCF/FQ nuclear gradients are employed to simulate vibronic absorption spectra of aromatic molecules in aqueous solution, namely benzene and phenol. By integrating this approach with molecular dynamics simulations, both solute conformational flexibility and the dynamical aspects of solvation are properly captured. The computed spectra reproduce experimental profiles and relative band intensities with remarkable accuracy, demonstrating the capability of the MCSCF/FQ model to simultaneously describe the multireference character of the solute and its interaction with the solvent environment.



1. INTRODUCTION

There is an ongoing effort in quantum chemistry to investigate molecular systems or processes that exhibit a significant degree of static (or nondynamic) electron correlation. In such cases, the wave function cannot be adequately represented by a single electronic configuration, but rather requires the simultaneous consideration of multiple electronic configurations.[1] These are referred to as multireference systems or processes. Typical examples include transition states, highly conjugated organic molecules (such as polyenes and aromatic systems), and transition metal complexes.[2] Processes such as bond dissociation or photoinduced dynamics also fall into this category.[3,4] For such cases, qualitatively correct results can only be achieved using multiconfigurational methods. Among these, the multiconfigurational self-consistent field (MCSCF) method[5] is one of the most widely employed, as it optimizes both the molecular orbitals and the configuration interaction coefficients simultaneously.

Despite the high accuracy achievable with multireference methods, they alone cannot adequately describe complex systems subject to numerous intermolecular interactions, such as molecules embedded in condensed-phase environments, including solutions or biological matrices. To model such systems, multireference approaches must be combined with multiscale strategies that partition the system into multiple regions.[6] This approach relies on the assumption that only a limited portion of the system is directly responsible for the properties of interest (e.g., a spectral signal), while the surrounding environment mainly acts as a perturbation. Accordingly, the primary region includes the part of the system that gives rise to the property under investigation and is described using a high-level quantum mechanical method. The secondary region, which represents the environment, is modeled at a more approximate level, for instance, by resorting to classical physics.

A widely adopted implementation of this concept is the hybrid quantum mechanics/molecular mechanics (QM/MM) approach, where the primary region is described quantum mechanically and the environment classically.[7−9] The classical region can be treated as a dielectric continuum, as in the Polarizable Continuum Model (PCM),[10,11] or by retaining an atomistic representation, as in molecular mechanics (MM).[8] While continuum models intrinsically provide a statistical average of the environment and accurately describe long-range electrostatic effects, atomistic models are essential for capturing short-range specific interactions, such as hydrogen bonding.[11]

Among the atomistic polarizable models,[12−22] the Fluctuating Charges (FQ) approach represents an optimal compromise

Received: November 12, 2025
Revised: December 23, 2026
Accepted: January 16, 2026

between accuracy and computational cost. In this model, each atom is endowed with a charge that is not fixed but adjusts to account for the presence of the other charges and the QM molecular potential.[17,20,23] The FQ approach belongs to the class of polarizable embedding models and thus accounts for mutual solute (QM) − solvent (FQ) polarization effects.[17]

For some applications, such as estimating solvation free energies, it may be sufficient to couple the QM and classical regions only at the total energy level. However, for the calculation of molecular properties, it is necessary to extend the formulation to energy derivatives. In particular, the nuclear gradient, which is defined as the first derivative of the energy with respect to the nuclear coordinates, constitutes a key quantity for accessing ground (and excited-state) properties. Due to the scaling of nuclear gradients with the number of atoms (i.e., with the molecule's dimension) analytical formulations are especially desirable, so to enable fast geometry optimizations, the computation of vibronic absorption and emission spectra, and the simulation of nonadiabatic processes in condensed-phase systems.[24−26]

Based on the aforementioned considerations, this work builds upon the previously developed MCSCF/FQ method[27] implemented in the OpenMolcas package,[28] extending it to the calculation of analytical nuclear gradients.

The paper is organized as follows. Section 2 provides a brief theoretical overview of the FQ approach, its coupling with an MCSCF wave function, together with the derivation of the novel MCSCF/FQ gradient equations. After a brief section focusing on the computational protocols exploited in the study, Section 4 showcases the potentialities of the MCSCF/FQ approach to simulate vibronic absorption spectra of benzene and phenol in aqueous solution, for which experimental spectra have been reported in the literature.[29,30] A final section summarizes the most relevant findings of the paper and proposes some perspectives for future developments.

2. THEORY

This section describes the development and implementation of analytical nuclear gradients within the MCSCF/FQ framework. The MCSCF/FQ approach[27] is briefly summarized in Section 2.1, while the derivation of the corresponding analytical gradient equations is presented in Section 2.2. The theoretical formulation follows the work reported in ref 27, adopting the notation introduced by Roos in ref 5 and employing the Einstein summation convention throughout.

2.1. Multiscale MCSCF/FQ Approach

Generally, in a QM/MM multiscale approach, the total energy of the system can be written as a sum of three terms[7]

$$E^{tot} = E^{QM} + E^{MM} + E^{QM/MM} \tag{1}$$

where E^{QM} and E^{MM} are the energies of the isolated QM and MM regions, respectively, and $E^{QM/MM}$ represents the interaction term between the two moieties. Within the MCSCF/FQ approach,[27] the QM region is described at the MCSCF level. The corresponding energy reads[5]

$$E^{QM} = \langle \Psi | \hat{H} | \Psi \rangle = h_{pq} D_{pq} + g_{pqrs} P_{pqrs} + V^{nn} \tag{2}$$

where h_{pq} and g_{pqrs} are the one-electron and two-electron integrals, respectively, and V^{nn} is the nucleus−nucleus repulsion term. $(\boldsymbol{D})$ and $(\boldsymbol{P})$ indicate the first- and second-order reduced density matrices. The electrical charges of the MM region are described using the Fluctuating Charges (FQ)[17,20,23] polarizable force field. Within this approach, each atom in the classical - FQ - portion is endowed with a charge that can dynamically change to fulfill Sanderson's electronegativity equalization (EE) principle.[31] This principle states that, at equilibrium, all atoms in a system share the same electronegativity. Charges can therefore be obtained through an equivalent reformulation of the EE principle, which involves minimizing the energy functional derived by truncating the Taylor expansion of the energy with respect to the charges up to second order.[17,23] To prevent unphysical charge transfer between distant molecules, the total charge Q_α of each molecule α is constrained to remain constant by introducing a set of Lagrange multipliers λ_α. Therefore, the functional to be minimized, for the isolated FQ region, can be written as[17]

$$E^{FQ} = \sum_{i\alpha} \chi^0_{i\alpha} q_{i\alpha} + \frac{1}{2} \sum_{i\alpha, j\beta} q_{i\alpha} T_{i\alpha j\beta} q_{j\beta} + \sum_\alpha \lambda_\alpha \left[\sum_i (q_{i\alpha}) - Q_\alpha \right] \tag{3}$$

where i and j run over FQ atoms within each molecule, $\chi^0_{i\alpha}$ is the electronegativity of the isolated atom, and $T_{i\alpha,j\beta}$ is the charge−charge interaction kernel. The diagonal terms of the kernel, $T_{i\alpha,i\alpha}$, account for the contribution $\frac{1}{2}\eta_{i\alpha} q^2_{i\alpha}$ due to the chemical harnesses $\eta_{i\alpha}$. To avoid the so-called "polarization catastrophe",[13] the Ohno kernel is used,[17,32] with the diagonal elements expressed in terms of atomic chemical hardnesses $\eta_{i\alpha}$. The FQ force field thus depends on only two atomic parameters: the electronegativity $\chi^0_{i\alpha}$ and the chemical hardness $\eta_{i\alpha}$.

In the MCSCF/FQ framework, the interaction energy $E^{QM/MM}$ is the electrostatic interaction between the QM and MM regions. It can be written as[27]

$$E^{QM/MM} = E^{QM/FQ} = q_{i\alpha} V^{i\alpha}_{pq} D_{pq} + q_{i\alpha} V^{i\alpha}_A Z_A \tag{4}$$

where $V^{i\alpha}_{pq} D_{pq}$ is the electrostatic potential evaluated at the $i\alpha$-th FQ charge due to the QM electronic density, and $V^{i\alpha}_A Z_A$ is the analogous potential due to the QM nuclear charges. In this work, a purely state-specific (SS) approach is considered, meaning that both the wave function and the FQs are optimized with respect to a single target state. Consequently, the density D_{pq} used to polarize the solvent corresponds to that state. The interaction kernels $V^{i\alpha}_{pq}$ and $V^{i\alpha}_A$ are expressed as[27]

$$V^{i\alpha}_{pq} = -\left\langle \phi_p \left| \frac{1}{|\boldsymbol{r}_{i\alpha} - \boldsymbol{r}|} \right| \phi_q \right\rangle \tag{5}$$

$$V^{i\alpha}_A = \frac{1}{|\boldsymbol{r}_{i\alpha} - \boldsymbol{R}_A|} \tag{6}$$

The total functional to be variationally minimized is therefore[27]

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

$$E(\boldsymbol{D}, \boldsymbol{P}, \boldsymbol{q}, \boldsymbol{\lambda}) = V^{nn} + h_{pq}D_{pq} + g_{pqrs}P_{pqrs} + \chi^{0}_{i\alpha}q_{i\alpha} + \frac{1}{2}q_{i\alpha}T_{i\alpha j\beta}q_{j\beta} + \lambda_{\alpha}\left[\sum_{i}(q_{i\alpha}) - Q_{\alpha}\right] + q_{i\alpha}V^{i\alpha}_{pq}D_{pq} + q_{i\alpha}V^{i\alpha}_{A}Z_{A} \tag{7}$$

The first three terms correspond to the QM energy, obtained as the expectation value of the molecular Hamiltonian (see eq 2), while the next three terms describe the FQ subsystem (see eq 3). Finally, the last two terms represent the electrostatic interaction between the QM and FQ portions (see eq 4).

Within this scheme, the effect of the FQ charges enters as a perturbation to the MCSCF Hamiltonian, which can be written as follows[27]

$$\hat{H}^{eff} = V^{nn} + q_{i\alpha}V^{i\alpha}_{A}Z_{A} + [h_{pq} + q_{i\alpha}V^{i\alpha}_{pq}]\hat{E}_{pq} + \frac{1}{2}g_{pqrs}(\hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps}) \tag{8}$$

in which the quantity $q_{i\alpha}V^{i\alpha}_{pq}$ is added to the monoelectronic term and $q_{i\alpha}V^{i\alpha}_{A}Z_{A}$ to the nuclear repulsion term.[27]

The total MCSCF/FQ energy (eq 1) is minimized with respect to all parameters (see ref 27 for more details) by exploiting the usual optimization techniques until self-consistency is achieved. This approach ensures that the mutual polarization between the MCSCF portion and the FQ layer is recovered.

2.2. Analytical MCSCF/FQ Nuclear Gradient

This section introduces a set of novel equations for calculating SS-MCSCF/FQ analytical nuclear gradients. The derivation is given by resorting to a "fully-focused" approach. This means that gradients (and related properties) are computed only for the MCSCF portion of the multilayer systems, whereas the explicit terms related to the FQ layer are discarded. The formulation of these latter terms, corresponding to the first derivatives with respect to the nuclear coordinates of the atoms in the FQ region, is taken from ref 33 and is recalled in Section S1 of the Supporting Information (SI). This procedure has been proposed many times in the context of QM/classical approaches and is in line with the so-called Partial Hessian Vibrational Approach (PHVA)[34−36] for vibrational analysis.

The nuclear gradient is the first derivative of the energy with respect to the nuclear coordinates ξ:

$$\frac{dE}{d\xi} = \frac{d}{d\xi}\langle\Psi|\hat{H}|\Psi\rangle \tag{9}$$

Using the chain rule, this derivative can be expanded as

$$\frac{dE(\boldsymbol{D}, \boldsymbol{P}, \boldsymbol{q}, \boldsymbol{\lambda})}{d\xi} = \frac{\partial E}{\partial \xi} + \frac{dE}{d\boldsymbol{D}}\frac{\partial \boldsymbol{D}}{\partial \xi} + \frac{dE}{d\boldsymbol{P}}\frac{\partial \boldsymbol{P}}{\partial \xi} + \frac{dE}{d\boldsymbol{q}}\frac{\partial \boldsymbol{q}}{\partial \xi} + \frac{dE}{d\boldsymbol{\lambda}}\frac{\partial \boldsymbol{\lambda}}{\partial \xi} \tag{10}$$

The MCSCF/FQ energy depends on four sets of parameters: the MOs ($\boldsymbol{\kappa}$), the CI coefficients ($\boldsymbol{C}$), the FQ charges ($\boldsymbol{q}$), and the FQ Lagrangian multipliers ($\boldsymbol{\lambda}$). In an SS-MCSCF/FQ calculation, all these parameters are optimized for a specific electronic state, and the corresponding energy derivatives vanish:

$$\frac{dE}{d\kappa} = \frac{dE}{dC} = \frac{dE}{dq} = \frac{dE}{d\lambda} = 0 \tag{11}$$

This means that the last two contributions to eq 10 vanish. For an isolated SS-MCSCF system, using the first two conditions in eq 11 to obtain the derivative of the energy of eq 2 gives[37−40]

$$\frac{dE}{d\xi} = \frac{dV^{nn}}{d\xi} + \frac{dh_{pq}}{d\xi}D_{pq} + \frac{dg_{pqrs}}{d\xi}P_{pqrs} + F_{pq}\frac{dS_{pq}}{d\xi} \tag{12}$$

where the last contribution $\left(F_{pq}\frac{dS_{pq}}{d\xi}\right)$ is related to the orthonormality constraint on the MOs.[37−40] This term can be calculated from the derivative of the overlap matrix $\boldsymbol{S}$ and the generalized Fock matrix $\boldsymbol{F}$:[41]

$$F_{pq} = D_{pr}h_{qr} + 2P_{prst}g_{qrst} \tag{13}$$

Proceeding analogously for the SS-MCSCF/FQ energy, eq 7 can be rearranged by grouping terms with the same dependence on the density matrices as follows:

$$E(\boldsymbol{D}, \boldsymbol{P}, \boldsymbol{q}, \boldsymbol{\lambda}) = [V^{nn} + q_{i\alpha}V^{i\alpha}_{A}Z_{A}] + [h_{pq} + q_{i\alpha}V^{i\alpha}_{pq}]D_{pq} + g_{pqrs}P_{pqrs} + \chi^{0}_{i\alpha}q_{i\alpha} + \frac{1}{2}q_{i\alpha}T_{i\alpha j\beta}q_{j\beta} + \lambda_{\alpha}\left[\sum_{i}(q_{i\alpha}) - Q_{\alpha}\right] \tag{14}$$

The last three terms do not contribute to the nuclear gradient because of the conditions shown in eq 11 and because $\boldsymbol{\chi}^{0}$, $\boldsymbol{T}$, and $\boldsymbol{Q}$ do not depend on the nuclear coordinates of the QM portion of the system. To evaluate the remaining terms, the same approach as for isolated systems can be followed, resulting in the following equation

$$\frac{dE}{d\xi} = \left[\frac{dV^{nn}}{d\xi} + q_{i\alpha}\frac{dV^{i\alpha}_{A}}{d\xi}Z_{A}\right] + \left[\frac{dh_{pq}}{d\xi} + q_{i\alpha}\frac{dV^{i\alpha}_{pq}}{d\xi}\right]D_{pq} + \frac{dg_{pqrs}}{d\xi}P_{pqrs} + \tilde{F}_{pq}\frac{dS_{pq}}{d\xi} \tag{15}$$

which includes the FQ contributions. The generalized Fock matrix in the last term is replaced by an effective Fock matrix which accounts for FQ contributions, i.e.,

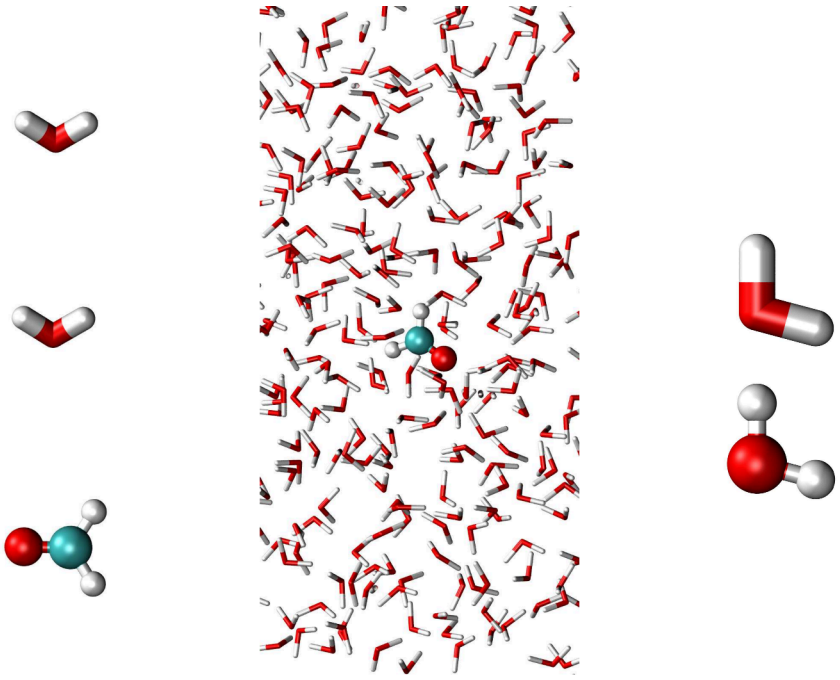
$$\tilde{F}_{pq} = D_{pr}[h_{qr} + q_{i\alpha}V^{i\alpha}_{qr}] + 2P_{prst}g_{qrst} \tag{16}$$

Recalling the QM/MM energy partition of eq 1, the gradient contributions in eq 15 can be mapped as follows:

$$\begin{aligned}
&\frac{dE^{QM}}{d\xi} \quad \Rightarrow \quad \frac{dV^{nn}}{d\xi} + \frac{dh_{pq}}{d\xi}D_{pq} + \frac{dg_{pqrs}}{d\xi}P_{pqrs} + F_{pq}\frac{dS_{pq}}{d\xi}\\
&\frac{dE^{QM/MM}}{d\xi} \quad \Rightarrow \quad q_{i\alpha}\frac{dV^{i\alpha}_{A}}{d\xi}Z_{A} + q_{i\alpha}\frac{dV^{i\alpha}_{pq}}{d\xi}D_{pq} + q_{i\alpha}V^{i\alpha}_{qr}D_{pr}\frac{dS_{pq}}{d\xi}\\
&\frac{dE^{MM}}{d\xi} \quad \Rightarrow \quad \emptyset
\end{aligned} \tag{17}$$

Notice that the approach presented above could in principle be extended to a SA-MCSCF/FQ framework by adopting

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX



(a) System **A** (b) System **B** (c) System **C**

Figure 1. Test systems employed for validating and debugging the CASSCF/FQ analytical gradient implementation: a) formaldehyde (QM) and two water molecules (MM) (system **A**), b) formaldehyde (QM) surrounded by 505 water molecules (MM) (system **B**), c) two water molecules, one described at the QM level and the other at the MM level (system **C**).

similar strategies as those developed for alternative polarizable embedding approaches.[42−44] However, in that case, since multiple states are included, the evaluation of analytical gradients requires solving coupled-perturbed equations,[45,46] which significantly increases the complexity of the derivation. For this reason, in this work, we focus exclusively on SS-MCSCF/FQ gradients.

2.3. Implementation

SS-CASSCF/FQ analytical nuclear gradients are implemented in a local version of OpenMolcas,[28] expanding on prior developments of the CASSCF/FQ model.[27] The implementation is fully integrated into the *Alaska* package[28] and involves three distinct terms that capture the interaction between the MCSCF and FQ regions, as detailed in eq 17.

The quantities added to the nuclear gradient arise from the interaction energy $E^{QM/MM}$, which in MCSCF/FQ is purely electrostatic. However, nonelectrostatic interactions may also play an important role, because they contain the dispersion and the repulsion components. To recover these contributions, it is possible to resort to a Lennard-Jones potential V(r),[47] which for two atoms separated by a distance r reads

$$V(r) = 4\varepsilon_{ab}\cdot\left[\left(\frac{\sigma_{ab}}{r}\right)^{12} - \left(\frac{\sigma_{ab}}{r}\right)^{6}\right] \tag{18}$$

where the parameters ε_{ab} and σ_{ab} depend on the pair of atom types. To automatically add this potential to the nuclear gradient, the integration between OpenMolcas and Tinker[48] is extended to support the CASSCF/FQ model.

Analytical gradients with respect to MM coordinates have also been developed and implemented (see Section S1 in the SI). Notice that, by exploiting this integration, geometry optimizations of the entire system can be performed.

3. COMPUTATIONAL DETAILS

3.1. Validation Step: Analytical vs Numerical Gradients

The validation of the analytical gradients is performed through comparison with numerical gradients for three model systems (see Figure 1):

A. Formaldehyde (QM) with two water molecules (FQ).
B. Formaldehyde (QM) surrounded by 505 water molecules (FQ).
C. Two water molecules, one treated as QM and the other as FQ.

Numerical gradients are computed using the finite difference method up to the eighth order, with a step size of $\Delta = 0.0025$Å to minimize numerical inaccuracies. Nonequilibrium geometries of the QM molecules are used to amplify gradient components, thereby reducing numerical errors. The nonequilibrium geometry of formaldehyde is obtained by perturbing by a small value (0.1Å) some coordinates of the positions taken from an optimized structure.

For System A, CASSCF/FQ(12,10) gradients are computed using the basis sets CC-PVDZ, CC-PVTZ, aug-cc-PVDZ, and aug-cc-PVTZ. Initial guess orbitals are generated with the GuessOrb program in OpenMolcas.[28] FQ parameters, referred to as FQ^a, are taken from ref 49. For System B, gradients are computed at the CASSCF/FQ(12,10)/CC-PVTZ level of theory, with two sets of FQ parameters being tested: FQ^a[49] and FQ^b.[50] In System C, gradients are evaluated at the CASSCF/FQ^a(8,6)/CC-PVDZ level of theory, starting from SCF orbitals. The distance between the QM hydrogen and the FQ oxygen atom is varied from 1.0 to 2.5 Å to sample typical hydrogen bond geometries (see Figure 2). As a reference, a single water molecule *in vacuo* is analyzed using SS-CASSCF, which represents an infinite hydrogen bond distance. Only x

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

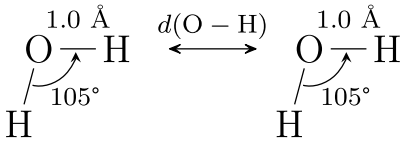


Figure 2. System **C**. One water molecule is treated at the CASSCF level, while the other at the FQ level.

and y gradient components are considered, as the z-components vanish due to symmetry.

To quantify the agreement between analytical and numerical gradients, three quantities are computed, namely the root-mean-square deviation (RMSD), the relative root-mean-square error (RRMSE) and the maximum discrepancy between the components of analytical and numerical gradients (Δ_{max}):

$$\sigma_a = \mathrm{RMSD} = \sqrt{\frac{1}{3N}\sum_{i=1}^{3N}(g_i^{analytical} - g_i^{numerical})^2} \tag{19}$$

$$\sigma_r = \mathrm{RRMSE} = \sqrt{\frac{1}{3N}\sum_{i=1}^{3N}\left(\frac{g_i^{analytical} - g_i^{numerical}}{g_i^{numerical}}\right)^2} \tag{20}$$

$$\Delta_{max} = \max_i(|g_i^{analytical} - g_i^{numerical}|) \tag{21}$$

3.2. Vibronic Spectra Calculation

Vibronic spectra of benzene and phenol in aqueous solution are computed through a multistep procedure adapted from protocols developed by some of us for simulating molecular spectral signals.[8,51] The multistep procedure consists of:[8,51]

1. *Definition of the QM/FQ partition:* The solute (QM subsystem) is treated at the CASSCF level, whereas the solvent is modeled with the FQ force field.
2. *Conformational sampling:* Classical MD simulations are performed over nanosecond time scales to sample the solute−solvent phase space. More details on MD settings for both systems are given in Section S2 of the SI.
3. *Extraction of representative structures:* 500 uncorrelated frames are extracted from MD trajectories, and a solvent droplet of sufficient radius is cut around the solute to include long-range interactions between the solute and the solvent. Representative snapshots of benzene (left) and phenol (right) in aqueous solution are shown in Figure 3.
4. *CASSCF/FQ calculations:* First, the solutes' geometry is optimized both in the ground (GS) and excited (ES) states. On optimized geometries, SS-CASSCF/FQ calculations are performed to extract all quantities that are required to simulate vibronic spectra (vide infra).
5. *Spectra extraction and analysis:* Final vibronic spectra are obtained by averaging computed signals for all snapshots, followed by a convolution with Lorentzian functions. Results are analyzed and compared with experimental data.

The calculation of vibronic spectra is perfomed in the following way. For each of the 500 snapshots, the solute's geometry is first optimized in both the GS and ES with SS-CASSCF/FQ-CC-PVDZ while keeping the FQ solvent molecules frozen, and the Hessians are computed numerically by finite differences of the nuclear gradients using the program *Slapaf* in OpenMolcas.[28] The ES optimized geometries resulted in structures where the ring was planar, both for benzene and phenol, and this is consistent with previous in vacuo calculations.[52,53] Then, the geometries of the solutes are optimized again at the SS-CASSCF/FQ-aug-cc-PVTZ level of theory, and the quantities needed for the subsequent vibronic spectra calculation are computed in both geometries and for both electronic states. These quantities include: energies, nuclear gradients and the transition dipole moment. For aqueous benzene, two sets of FQ parameters are employed and compared, namely FQ[b] taken from ref 50 and FQ[c] taken from ref 20. For the aqueous solution of phenol only the FQ[c] parameters are employed. For benzene, the active space comprises six π orbitals with six active electrons, while for phenol it includes seven π orbitals with eight active electrons.

Five snapshots are used to perform preliminary tests on CASSCF/FQ parameters, such as the selection of starting orbitals. Specifically, different starting orbitals are tested: Restricted Hartree−Fock (RHF), Unrestricted Hartree−Fock Natural Orbitals (UNO),[2,54,55] and Guess orbitals generated by the *GuessOrb* program in OpenMolcas.[28] As a result, Guess orbitals are selected as starting orbitals for the benzene solution. For the active space selection of solvated phenol, which presents more challenges, a protocol based on the maximum MO overlap, similar to the one presented by

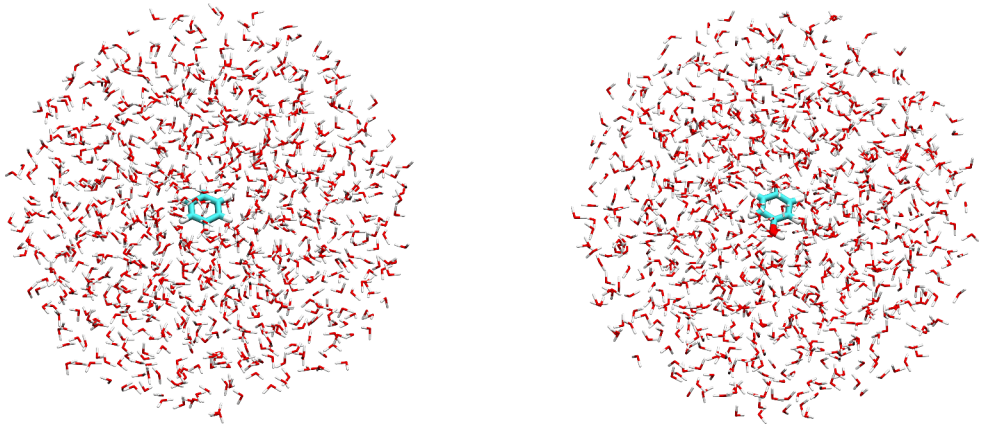


Figure 3. Representative snapshots of benzene (left) and phenol (right) in aqueous solution, extracted from MD simulations. The radius of the solvation spheres is 18 Å for both solutes and includes ∼850 water molecules.

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Cárdenas and Nogueira,[56] is used to preserve its consistency among the snapshots, starting from preliminary CASSCF/FQ-CC-PVDZ calculations. The details of this procedure are reported in Section S3 in the SI.

Subsequently, for each snapshot, vibronic spectra are calculated with both vertical and adiabatic harmonic approximations.[26,57] Vertical approximations describe both the GS and ES PESs by performing the calculation only at the GS optimized geometry. Vertical approximations include Vertical Gradient (VG),[57,58] where the description of the ES PES is obtained with just the ES nuclear gradients, and Vertical Hessian (VH),[59] which requires the computation of the ES Hessian. Adiabatic approximations instead describe the ES PES performing the calculations at the ES optimized geometry. Adiabatic approximations include Adiabatic Shift (AS),[57] which describes the ES PES only with the position of its minimum, and Adiabatic Hessian (AH),[57] which requires the computation of the ES Hessian. Among these methods, AS and VG are considerably simpler than AH and VH, because they do not require the calculation of ES Hessian and rely on the approximation that the GS and ES Hessians are equal. All spectral calculations employ the time-independent (TI) approach,[57] the Franck−Condon approximation, at a temperature of 298.15K. The TI approach was chosen over the time-dependent (TD) approach[60] because, by directly computing stick spectra for each snapshot, it allows us to clearly distinguish between homogeneous and inhomogeneous broadening contributions. For all calculations, it was checked that the recovered fraction of spectra was ≥99.9%. Spectral calculations are performed using internal coordinates, which have proved to be more robust than Cartesian coordinates, especially employing vertical models.[61] The final vibronic spectrum is obtained by averaging all spectra of the individual snapshots with a convolution with a Lorentzian function with full width at half-maximum (fwhm) of 0.04 eV for benzene, and a Lorentzian function with fwhm of 0.15 eV for phenol

The OPLS-AA force field[62] is used to model van der Waals interactions in gradient calculations with Tinker.[48] All CASSCF/FQ calculations are performed with a locally modified version of the OpenMolcas software,[28,63] whereas vibronic contributions to electronic spectra are computed with FCclasses3.[64]

4. RESULTS AND DISCUSSION

In this section, we first validate CASSCF/FQ gradients for the model systems depicted in Figure 1. To this end, beyond checking analytical/numerical consistency, we evaluate the effect of the basis set, FQ parameters, and the size of the water droplet (i.e., the number of water molecules). We further analyze how the distance between the QM and FQ fragments influences CASSCF/FQ nuclear gradients, which also allows for a direct comparison with the analytical−numerical error observed in isolated systems.

Analytical gradients are then employed to simulate vibronic absorption spectra of benzene and phenol in aqueous solution.

4.1. Analytical Gradient Validation

For system **A**, consisting of formaldehyde (QM) and two water molecules (FQ), numerical and analytical CASSCF/FQ^a(12,10) gradients are compared by varying the basis set (CC-PVDZ, CC-PVTZ, aug-cc-PVDZ, and aug-cc-PVTZ). The corresponding results are summarized in Table 1. Values obtained with or without augmenting the basis set are remarkably similar, thus indicating that augmentation does not improve the results for this simple system.

Table 1. Maximum Deviation among Analytical and Numerical Gradients (Δ_{max}), RMSD (σ_a), and RRMSE (σ_r) for System A and Different Basis Sets[a]

Basis set	Δ_{max}	σ_a	σ_r
CC-PVDZ	1.6×10^{-6}	6.7×10^{-7}	2.3×10^{-5}
CC-PVTZ	2.6×10^{-6}	1.1×10^{-6}	5.8×10^{-5}
aug-cc-PVDZ	1.4×10^{-6}	5.6×10^{-7}	2.7×10^{-5}
aug-cc-PVTZ	2.2×10^{-6}	1.1×10^{-6}	6.1×10^{-5}

[a]The FQ^a paramtrization is employed.

System **B** (formaldehyde (QM) surrounded by 505 water molecules (FQ)) allows us to assess the impact of both the number of FQ atoms and the choice of the FQ parametrization (FQ^a,[49] FQ^b[50]) on nuclear gradients. CASSCF/FQ$^{(a,b)}$(12,10)-CC-PVTZ results are reported in Table 2. The comparison

Table 2. CASSCF/FQ$^{(a,b)}$(12,10)-CC-PVTZ Δ_{max}, σ_a, and σ_r Values for System B, as Computed by Employing FQ^a and FQ^b Parameter Sets

System	Δ_{max}	σ_a	σ_r
B (FQ^a)	1.7×10^{-5}	6.9×10^{-6}	2.4×10^{-4}
B (FQ^b)	2.0×10^{-6}	9.5×10^{-7}	5.1×10^{-5}

with system **A** values obtained with the same basis set ad parametrization (Table 1) shows that increasing the number of water molecules (system **B**) yields deviations that are about 1 order of magnitude larger, while the results obtained by changing the FQ parametrization from FQ^a to FQ^b differ by roughly the same scale. This shows that both the size of the solvent droplet and the choice of the FQ parameter sets can influence nuclear gradients to the same extent; nevertheless, deviations remain small, thus indicating that these factors marginally affect the accuracy of analytical gradients with respect to their numerical counterparts.

System **C** is designed to probe the effect of the QM−FQ distance on computed gradients. The distance between the QM hydrogen atom and the FQ oxygen atom of the aligned O− − H bonds (see Figure 2) is varied from 1.0 to 2.5 Å, and extended to infinity (i.e., to a pair of isolated water molecules). The results are summarized in Table 3. They show that the

Table 3. Δ_{max}, σ_a, and σ_r as a Function of the H−O Distance in System C

d(O−H)	Δ_{max}	σ_a	σ_r
1.0 Å	1.4×10^{-7}	8.6×10^{-8}	1.0×10^{-5}
1.5 Å	1.5×10^{-7}	8.7×10^{-8}	4.2×10^{-6}
2.0 Å	1.5×10^{-7}	8.7×10^{-8}	4.4×10^{-6}
2.5 Å	1.4×10^{-7}	8.6×10^{-8}	4.5×10^{-6}
∞	1.4×10^{-7}	8.8×10^{-8}	4.7×10^{-6}

QM−FQ distance does not affect the quality of the computed gradients, even in the hydrogen-bonding region. Most importantly, the deviation parameters for the isolated molecule are fully consistent with those obtained using the SS-CASSCF/FQ model, thus demonstrating that the SS-CASSCF/FQ analytical nuclear gradient is as robust as standard CASSCF gradients computed in OpenMolcas.

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

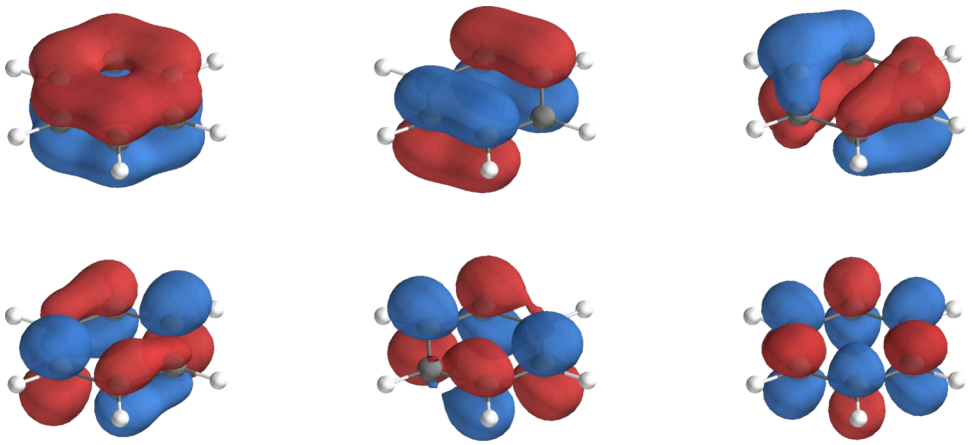


Figure 4. Active orbitals selected for CASSCF(6,6)/FQ calculations. They are obtained with the *GuessOrb* program in OpenMolcas.[28]

4.2. Vibronic Spectrum of Benzene in Aqueous Solution

In this section, CASSCF/FQ analytical gradient implementation is applied to simulate the virbonic absorption spectrum of aqueous benzene. Calculations focus on the $S_0 \rightarrow S_1$ electronic transition, and spectra are evaluated using both vertical (VG and VH) and adiabatic (AS and AH) approaches.

Before showing final spectra, preliminary tests are carried out on a limited number of snapshots. The simplest way of obtaining the starting MOs for CASSCF calculations consists of performing a preliminary RHF calculation.[5] However, RHF orbitals can mix relevant and less relevant contributions, which complicates the identification of the proper active space. Moreover, different snapshots may yield different MOs, hindering the possibility of defining a consistent active space across all geometries. Alternative strategies include the use of Natural Orbitals obtained from a UHF calculation (UNO)[2,54,55] or directly utilizing the guess orbitals generated by the *GuessOrb* program in OpenMolcas.[28] The results obtained by employing all three choices of starting orbitals discussed above are compared for five snapshots. All techniques yield the same converged CASSCF MOs and result in the same energies for benzene. Therefore, guess orbitals are selected in subsequent calculations. As already suggested in the literature,[65] the active space is select so to include the 6π orbitals displayed in Figure 4.

Another relevant point, which necessitates checking, is the use of an SS-CASSCF approach. When two distinct states, namely the GS and the ES, are obtained through SS-CASSCF calculations, they are optimized independently and, in principle, are neither orthogonal nor noninteracting. However, according to the theory, the true states should be two eigenstates of the same Hamiltonian, which are orthogonal to each other. In SS-CASSCF/FQ calculations the Hamiltonians for the GS and the ES are different, because the FQ charges are equilibrated to a specific state of interest (the GS or ES). The impact of nonorthogonality on computed excitation energies is evaluated using Löwdin symmetric orthogonalization,[66,67] which minimizes the distance between the original and orthogonalized states. The results show that the correction to the excitation energy is negligible (of the order of 10^{-7} a.u. - see Section S4 in the SI).

Stick and averaged CASSCF/FQ$^{(b,c)}$ vibronic spectra of benzene, obtained from 200 snapshots and computed with vertical (VG and VH) and adiabatic (AS and AH) approximations,[26] are reported in Figure 5, together with the experimental spectrum.[29] Sticks represent the set of vibronic transitions computed for each snapshot, with their distribution reflecting the different solvent environments sampled during the MD. In this framework, inhomogeneous broadening arises naturally from conformational sampling, while homogeneous broadening is introduced by convoluting averaged stick spectra with a Lorentzian function with fwhm of 0.04 eV. With FQ^c (Figure 5, right), the stick spectra are considerably narrower than with FQ^b (Figure 5, left), leading to sharper convoluted bands. In both parametrizations, however, the AH stick spectra remain significantly broader than those obtained with the other vibronic approximations. Note that 200 snapshots are sufficient to ensure convergence for both FQ parametrizations, as increasing the number of snapshots from 150 to 200 does not alter the results (see Figures S3 and S4 in the SI).

To evaluate the quality of the convoluted spectra shown in Figure 5, three key quantities are considered: the spectral profile, the position of the peaks, and their relative intensities.

The CASSCF/FQ$^{(b,c)}$ spectral profiles exhibit a vibronic structure with clearly distinguishable peaks. At the CASSCF/FQ^b level of theory (see Figure 5, left), VG, VH, and AS approximations yield about six distinct vibronic peaks, while the AH spectrum shows fewer peaks, as excessive broadening prevents a clear resolution of the vibronic structure. Concerning relative intensities, the vertical models predict the three strongest peaks at longer wavelengths, with the central peak being the most intense. In the adiabatic models, instead, the longest-wavelength 0−0 transition is surpassed by the bands at approximately 236 nm (AS) and 244 nm (AH), and therefore does not appear among the three dominant peaks. Regarding absolute peak positions, the AS, VG, and VH bands appear at comparable wavelengths, with the 0−0 transition around 254 nm (AS and VG) and 256 nm (VH). The AH bands are instead red-shifted by about 8 nm relative to AS/VG and by 6 nm relative to VH.

Moving to CASSCF/FQ^c spectra (Figure 5, right), the line shapes are simpler and closely resemble each other, with six distinct and well-separated peaks. The broadening is reduced compared to the CASSCF/FQ^b spectra and remains consistent across all cases, being slightly larger for the AH model. Focusing on relative intensities, the three most intense peaks appear at longer wavelengths also when adiabatic approaches are employed, differently from what we have observed for FQ^b calculations. The peak positions follow the same trend as previously discussed: AS and VG are nearly identical, with the

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

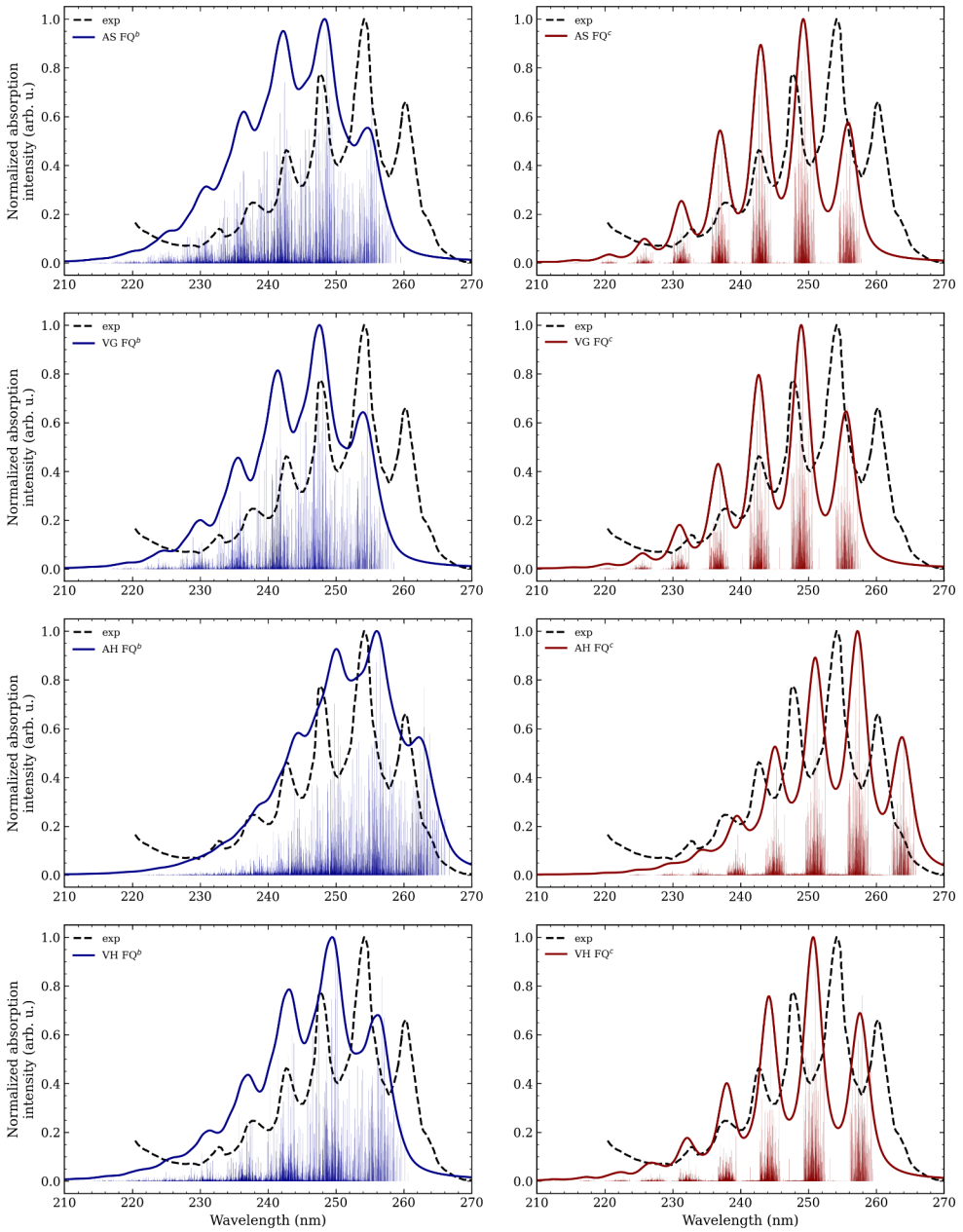


Figure 5. CASSCF(6,6)/FQ$^{(b,c)}$ vibronic spectra of benzene in aqueous solution, obtained with vertical (VG and VH) and adiabatic (AS and AH) approximations. In blue (left), spectra obtained with FQ^b parametrization. In red (right), spectra obtained with FQ^c parametrization. The experimental spectrum (black, dotted) adapted from ref 29 Copyright 1976, American Chemical Society, is superimposed.

0−0 transition around 256 nm, while VH gives 258 nm, and the AH spectrum is red-shifted to 264 nm.

We now turn to the comparison with the experimental spectrum. The experimental profile displays a vibronic structure consisting of six main peaks. With the FQ^c parametrization, all experimental bands are well reproduced, and the computed broadening closely matches the experiment. In contrast, FQ^b parametrization yields less resolved peaks, due to a larger broadening. Nevertheless, computed spectra successfully capture the change in slope near the minimum between the two most intense experimental peaks, particularly with VG and VH approaches.

If FQ^b parametrization is employed, AS, VG, and VH show a blue shift of about 4−6 nm for the 0−0 transition (experimentally at ∼ 260 nm), while the AH spectrum is red-shifted by about 2 nm. A similar behavior is observed with FQ^c, for which AS, VG, and VH give blue-shifted values by about 2−4 nm and AH values are red-shifted by approximately 4 nm. Overall, the more sophisticated models (AH and VH) provide peak positions in closer agreement with the experimental spectrum.

Turning to relative intensities, all CASSCF/FQ^b spectra reproduce the experimental feature, where the most intense peak is accompanied by a 0−0 transition at roughly 60% of its intensity. Differences emerge in the secondary peaks: in the adiabatic models (AS and AH), the peak to the left of the most intense one (236 nm for AS and 244 nm for AH) is systematically overestimated, whereas in the vertical models (VG and VH) the intensities show very good agreement with experiment. The experimental ordering of intensities is

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

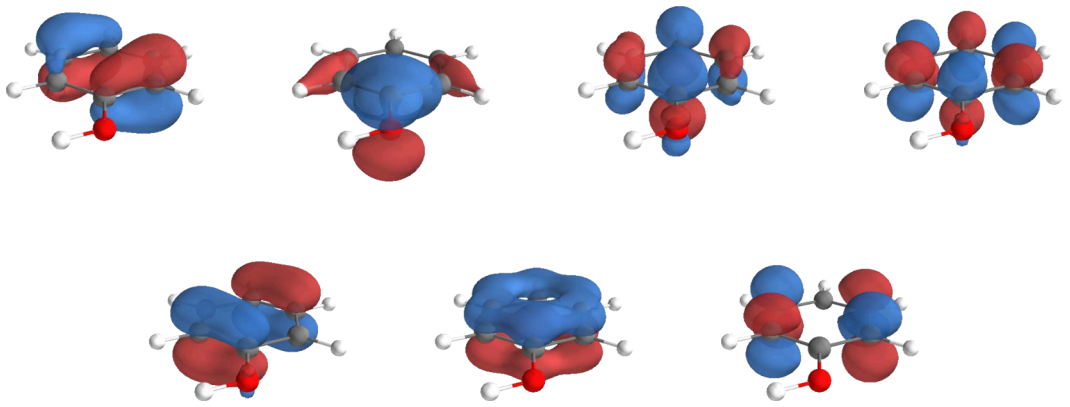


Figure 6. Active orbitals selected for CASSCF(8,7)/FQ calculations of phenol in aqueous solution. This active space is composed of the 7 phenol π orbitals, in which the orbital localized mainly on the oxygen can appear distorted as a consequence of the OH bond out-of-plane rotation.

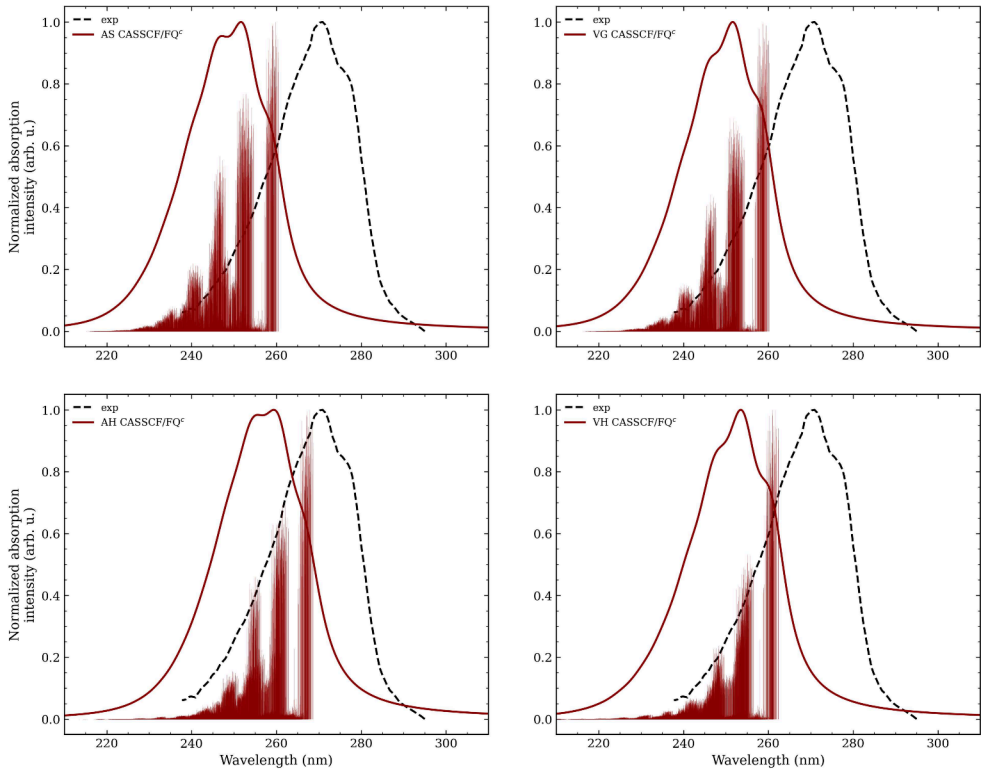


Figure 7. Stick and convoluted CASSCF(8,7)/FQ^c vibronic spectra of phenol in aqueous solution, obtained using the vertical (VG and VH) and adiabatic (AS and AH) approaches. The experimental spectrum (black, dotted) adapted from ref 30 available under a CC-BY license. Copyright 2018 the Authors, is superimposed.

preserved in all cases when FQ^c is employed, with VG and VH providing the closest match, while AS and AH tend to slightly overestimate the peaks to the left and underestimate the 0−0 band. The difference between the spectra obtained with the two FQ parametrizations is consistent with their intrinsic characteristics. The broader profiles associated with FQ^b likely stem from the fact that FQ^c parameters are optimized to reproduce the properties of bulk water,[20] whereas FQ^b parameters are explicitly designed to describe electrostatic and polarization interactions between solutes and the aqueous solvent.[50] Nonetheless, both parametrizations, combined with the different vibronic approaches, can capture the main experimental features. Overall, the VG and VH spectra provide a more accurate reproduction of the experimental profile than AS and AH. This discrepancy can be rationalized by noting that, while adiabatic models offer a more accurate description of the final PES, vertical approaches better represent the vertical region of the PES, which dominates the most intense peak.[68]

4.3. Vibronic Spectrum of Phenol in Aqueous Solution

We now move to a more challenging case: aqueous phenol. Unlike benzene, phenol features an explicit site for hydrogen bonding with the surrounding water, making an accurate treatment of the explicit solvent around the solute essential. In fact, it is reasonable to assume that, in the case of phenol, a reliable reproduction of the vibronic spectrum requires a subtle interplay among conformational analysis, the characterization of the solvent arrangement around the solute, correlation effects, and the level of treatment of vibronic couplings.

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Similar to benzene, VG, VH, AS, and AH approaches are challenged, coupled with CASSCF(8,7)/FQ^c-aug-cc-PVTZ calculations. The employed active space includes the seven π orbitals of phenol, in agreement with previous studies.[53] The presence of the OH group introduces additional complexity, since during the classical MD (step 2 of the computational protocol reported in Section 4) the OH bond may rotate out of the ring plane. This rotation complicates the definition of the active space compared to benzene. For most snapshots, the selected active orbitals correspond to those reported in Figure 6. To generate suitable starting orbitals, RHF orbitals, UNO, GuessOrb-generated orbitals, and preliminary CASSCF(8,7)/FQ^c calculations are compared. Only the latter consistently provides a comparable set of MOs across all snapshots and is therefore adopted in the subsequent calculations. Consistency in the choice of the active space among different sets is ensured through the protocol detailed in Section S3 in the SI.

In Figure 7, CASSCF/FQ^c vibronic spectra are shown for all four harmonic approximations, including both convoluted spectra and stick transitions. The experimental spectrum[30] is also reported for the sake of comparison. Similar to benzene, the variability of the vibronic signals across 150 snapshots accounts for the inhomogeneous broadening, while the Lorentzian convolution with a fwhm of 0.15 eV recovers the homogeneous broadening. Notice that 150 snapshots are enough to reach convergence, as shown in Figure S5 in the SI. Stick spectra show a vibronic structure with three main peaks. At longer wavelengths, peaks appear more isolated, while in the central region, a denser distribution of signals is observed. This leads to broad convoluted spectra, of which the intensity at higher wavelengths decreases steeply. All vibronic approaches provide a similar shape for the rightmost region, but AS and AH display higher intensities for the other features compared to the VG and VH, leading to clear differences in the overall spectral shape.

Convoluted spectra display a broad main band with two vibronic shoulders (see Figure 7). The central peak is the most intense across all vibronic models, while the red-most peak is the weakest. However, the relative intensities differ between vertical and adiabatic approaches: unlike vertical models, adiabatic approaches predict the low-wavelength peak to be nearly as intense as the central band. The positions of the peaks in AS and VG spectra are almost perfectly aligned, with the main peak at about 250 nm. The VH spectrum is slightly blue-shifted (main peak at 253 nm), while the AH spectrum shows a more pronounced blue shift of about 6 nm with respect to vertical approaches. Note that a similar trend was observed in the previous section for benzene in aqueous solution. The experimental spectrum (see Figure 7) presents a single broad peak that features two clearly visible vibronic shoulders, one on the left and one on the right at the top of the peak. Overall, these features are well reproduced by calculations. In both the computed and experimental VG and VH spectra, the shoulders reach about 80% of the main peak intensity. AS and AH overestimate the intensity of the left shoulder, and this results in a line shape that is less in accordance with the experimental profile. The position of the main peak presents a large blue shift of about 15 to 20 nm for the AS, VG and VH spectra and about 10 nm for the AH spectrum. These shifts, which are much larger than those observed for benzene, are probably caused by the explicit hydrogen bonding interaction between phenol and water, which requires an accurate description of both the flexibility of the phenolic OH group, and the specific arrangement of water around it. The importance of the description of the hydrogen bonding pattern is confirmed by additional calculations performed using the same computational protocol but resorting to the electrostatic embedding (EE) approach (with the OPLS-AA force-field),[62] see Section S6 in the SI. The similarity between the EE and FQ spectra suggests that, in this case, the atomistic representation of the environment plays a far more significant role than the mutual polarization between solute and solvent. Another concurrent factor that can explain this larger shift is the lack of dynamical correlation in CASSCF calculations.

In conclusion, as in the case of the benzene aqueous solution, vertical approaches produced spectra more in accordance with the experimental absorption spectrum than the adiabatic approximations, with little differences between the VG and the VH models. This suggests that, also in this case, an accurate description of the PES around the Franck–Condon region, as achieved with the vertical models, is more critical than an overall accurate representation of the excited-state PES obtained with the adiabatic approximations.[68]

5. SUMMARY, CONCLUSIONS, AND FUTURE PERSPECTIVES

We have presented the development and implementation of multiscale SS-MCSCF/FQ analytical nuclear gradients. The model has been validated through comparison between analytical and numerical gradients. Subsequently, analytical CASSCF/FQ nuclear gradients were employed to simulate the vibronic spectra of aqueous benzene and phenol, using four different approximations to describe vibronic progressions.[26,68]

All tested vibronic schemes, combined with the various FQ parametrizations, successfully reproduce the main experimental features, including spectral profiles, peak positions, and relative intensities. Overall, vertical approximations yield spectra in closer agreement with experiments. In particular, the VG model proved nearly as effective as the more elaborate VH approach, providing a computationally cheaper yet reliable alternative.

The consistency observed in the vibronic spectra obtained from different harmonic approximations can also serve as a metric to assess the reliability and accuracy of the current implementation, since each approximation requires the calculation of distinct and independent quantities (e.g., GS and ES optimized geometries, Hessians at different geometries, etc.).

The good agreement between computed and experimental spectra also confirms that the CASSCF/FQ approach effectively captures both the multireference character of benzene and phenol and their interactions with the aqueous environment.

Future developments will focus on achieving quantitatively accurate results by including dynamical correlation effects. Starting from a multiconfiguration reference, these could be incorporated via a perturbative treatment,[69] for instance, through the development of a CASPT2/FQ scheme.[44]

Moreover, the generality of the present formulation makes it suitable for extension to treat other solvents[70] or more complex environments.[71] Another important perspective concerns the inclusion of nonelectrostatic solute–solvent interactions, which are not accounted for in the current electrostatic-only FQ model. In particular, quantum repulsion and other short-range effects are expected to play a significant

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

role and will be essential to achieve quantitatively accurate descriptions of solvation.[72−75]

Finally, the extension of the analytical nuclear gradient to a state-average formalism[44] would open the way to the study of photochemical processes[3,4] involving multiple excited states.[76−80]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.5c01890.

Nuclear gradients with respect to MM coordinates. Computational details and analysis of MD simulations of benzene and phenol in aqueous solution. An algorithm to select active spaces among different structures of the same system, using MO overlap. Estimates of the effects of Non-Orthonormality of SS wave functions. Convergence tests on the number of snapshots used for computing vibronic spectra. Vibronic spectra with electrostatic embedding. (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Chiara Cappelli − *Scuola Normale Superiore, I-56126 Pisa, Italy*; orcid.org/0000-0002-4872-4505; Email: chiara.cappelli@sns.it

Authors

Francesco Mazza − *Scuola Normale Superiore, I-56126 Pisa, Italy*

Marco Trinari − *Scuola Normale Superiore, I-56126 Pisa, Italy*

Chiara Sepali − *Scuola Normale Superiore, I-56126 Pisa, Italy*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.5c01890

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge funding from MUR-FARE Ricerca in Italia: Framework per l'attrazione ed il rafforzamento delle eccellenze per la Ricerca in Italia - III edizione. Prot. R20YTA2BKZ and the European Union's Horizon Europe research and innovation programme under the project HORIZON-MSCA-2023-DN-01 - LUMIÈRE G.A. No 101169312. The Center for High-Performance Computing (CHPC) at SNS is also acknowledged for providing the computational infrastructure.

■ REFERENCES

(1) Schmidt, M. W.; Gordon, M. S. The construction and interpretation of MCSCF wavefunctions. *Annu. Rev. Phys. Chem.* **1998**, *49*, 233−266.

(2) Keller, S.; Boguslawski, K.; Janowski, T.; Reiher, M.; Pulay, P. Selection of active spaces for multiconfigurational wavefunctions. *J. Chem. Phys.* **2015**, *142*, 142.

(3) Choudhury, A.; Santra, S.; Ghosh, D. Understanding the Photoprocesses in Biological Systems: Need for Accurate Multi-reference Treatment. *J. Chem. Theory Comput.* **2024**, *20*, 4951.

(4) Curchod, B. F.; Martínez, T. J. Ab initio nonadiabatic quantum molecular dynamics. *Chem. Rev.* **2018**, *118*, 3305−3336.

(5) Roos, B. O. In *Lecture Notes in Quantum Chemistry: European Summer School in Quantum Chemistry*; Roos, B. O., Ed.; Springer: Berlin, 1992; pp 177−254.

(6) Visscher, L.; Bolhuis, P.; Bickelhaupt, F. M. Multiscale modelling. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10399−10400.

(7) Lin, H.; Truhlar, D. G. QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2007**, *117*, 185−199.

(8) Giovannini, T.; Egidi, F.; Cappelli, C. Molecular spectroscopy of aqueous solutions: a theoretical perspective. *Chem. Soc. Rev.* **2020**, *49*, 5664−5677.

(9) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

(10) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999−3094.

(11) Giovannini, T.; Cappelli, C. Continuum vs. atomistic approaches to computational spectroscopy of solvated systems. *Chem. Commun.* **2023**, *59*, 5644−5660.

(12) Loco, D.; Polack, É.; Caprasecca, S.; Lagardere, L.; Lipparini, F.; Piquemal, J.-P.; Mennucci, B. A QM/MM approach using the AMOEBA polarizable embedding: from ground state energies to electronic excitations. *J. Chem. Theory Comput.* **2016**, *12*, 3654−3661.

(13) Thole, B. T. Molecular polarizabilities calculated with a modified dipole interaction. *Chem. Phys.* **1981**, *59*, 341−350.

(14) Curutchet, C.; Munoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. Electronic energy transfer in condensed phase studied by a polarizable QM/MM model. *J. Chem. Theory Comput.* **2009**, *5*, 1838−1848.

(15) Olsen, J. M. H.; Kongsted, J. *Advances in Quantum Chemistry*; Elsevier, 2011; Vol. *61*; pp 107−143.

(16) Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. Excitation energies in solution: the fully polarizable QM/MM/PCM method. *J. Phys. Chem. B* **2011**, *115*, 3027−3037.

(17) Cappelli, C. Integrated QM/polarizable MM/continuum approaches to model chiroptical properties of strongly interacting solute−solvent systems. *Int. J. Quantum Chem.* **2016**, *116*, 1532−1542.

(18) Giovannini, T.; Puglisi, A.; Ambrosetti, M.; Cappelli, C. Polarizable QM/MM approach with fluctuating charges and fluctuating dipoles: the QM/FQFμ model. *J. Chem. Theory Comput.* **2019**, *15*, 2233−2245.

(19) Boulanger, E.; Thiel, W. Solvent boundary potentials for hybrid QM/MM computations using classical drude oscillators: a fully polarizable model. *J. Chem. Theory Comput.* **2012**, *8*, 4527−4538.

(20) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical fluctuating charge force fields: Application to liquid water. *J. Chem. Phys.* **1994**, *101*, 6141−6156.

(21) Thompson, M. A.; Schenter, G. K. Excited states of the bacteriochlorophyll b dimer of Rhodopseudomonas viridis: a QM/MM study of the photosynthetic reaction center that includes MM polarization. *J. Phys. Chem.* **1995**, *99*, 6374−6386.

(22) Dziedzic, J.; Mao, Y.; Shao, Y.; Ponder, J.; Head-Gordon, T.; Head-Gordon, M.; Skylaris, C.-K. TINKTEP: A fully self-consistent, mutually polarizable QM/MM approach based on the AMOEBA force field. *J. Chem. Phys.* **2016**, *145*, 145.

(23) Rick, S. W.; Berne, B. Dynamical fluctuating charge force fields: the aqueous solvation of amides. *J. Am. Chem. Soc.* **1996**, *118*, 672−679.

(24) Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; Wiley: United Kingdom, 2006.

(25) List, N. H.; Jensen, H. J. A.; Kongsted, J.; Hedegård, E. D. A unified framework for the polarizable embedding and continuum methods within multiconfigurational self-consistent field theory. *Adv. Quantum Chem.* **2013**, *66*, 195−238.

(26) Ferrer, F. J. A.; Santoro, F. Comparison of vertical and adiabatic harmonic approaches for the calculation of the vibrational structure of electronic spectra. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13549−13563.

(27) Sepali, C.; Goletto, L.; Lafiosca, P.; Rinaldi, M.; Giovannini, T.; Cappelli, C. Fully Polarizable Multiconfigurational Self-Consistent

Field/Fluctuating Charges Approach. *J. Chem. Theory Comput.* **2024**, *20*, 9954−9967.
(28) Li Manni, G.; et al. The OpenMolcas Web: A Community-Driven Approach to Advancing Computational Chemistry. *J. Chem. Theory Comput.* **2023**, *19*, 6933.
(29) Ilan, Y.; Luria, M.; Stein, G. Photochemistry of benzene in aerated aqueous solutions in the range of 214 to 265 nm. *J. Phys. Chem.* **1976**, *80*, 584−587.
(30) Riley, J. W.; Wang, B.; Woodhouse, J. L.; Assmann, M.; Worth, G. A.; Fielding, H. H. Unravelling the role of an aqueous environment on the electronic structure and ionization of phenol using photoelectron spectroscopy. *J. Phys. Chem. Lett.* **2018**, *9*, 678−682.
(31) Sanderson, R. An interpretation of bond lengths and a classification of bonds. *Science* **1951**, *114*, 670−672.
(32) Ohno, K. Some remarks on the Pariser-Parr-Pople method. *Theor. Chem. Acc.* **1964**, *2*, 219−227.
(33) Lipparini, F.; Cappelli, C.; Scalmani, G.; De Mitri, N.; Barone, V. Analytical first and second derivatives for a fully polarizable QM/classical hamiltonian. *J. Chem. Theory Comput.* **2012**, *8*, 4270−4278.
(34) Giovannini, T.; Grazioli, L.; Ambrosetti, M.; Cappelli, C. Calculation of ir spectra with a fully polarizable qm/mm approach based on fluctuating charges and fluctuating dipoles. *J. Chem. Theory Comput.* **2019**, *15*, 5495−5507.
(35) Besley, N. A.; Bryan, J. A. Partial Hessian vibrational analysis of organic molecules adsorbed on Si (100). *J. Phys. Chem.C* **2008**, *112*, 4308−4314.
(36) Vester, J.; Olsen, J. M. H. Assessing the Partial Hessian Approximation in QM/MM-Based Vibrational Analysis. *J. Chem. Theory Comput.* **2024**, *20*, 9533−9546.
(37) Kato, S.; Morokuma, K. Energy gradient in a multiconfigurational SCF formalism and its application to geometry optimization of trimethylene diradicals. *Chem. Phys. Lett.* **1979**, *65*, 19−25.
(38) Schlegel, H. B.; Robb, M. A. MC SCF gradient optimization of the H2CO → H2 + CO transition structure. *Chem. Phys. Lett.* **1982**, *93*, 43−46.
(39) Yamamoto, N.; Vreven, T.; Robb, M. A.; Frisch, M. J.; Schlegel, H. B. A direct derivative MC-SCF procedure. *Chem. Phys. Lett.* **1996**, *250*, 373−378.
(40) Almlöf, J.; Taylor, P. R. Molecular properties from perturbation theory: A unified treatment of energy derivatives. *Int. J. Quantum Chem.* **1985**, *27*, 743−768.
(41) Siegbahn, P. E.; Almlöf, J.; Heiberg, A.; Roos, B. O. The complete active space SCF (CASSCF) method in a Newton−Raphson formulation with application to the HNO molecule. *J. Chem. Phys.* **1981**, *74*, 2384−2396.
(42) Song, C. State averaged CASSCF in AMOEBA polarizable water model for simulating nonadiabatic molecular dynamics with nonequilibrium solvation effects. *J. Chem. Phys.* **2023**, *158*, 014101.
(43) Song, C.; Wang, L.-P. A polarizable QM/MM model that combines the state-averaged CASSCF and AMOEBA force field for photoreactions in proteins. *J. Chem. Theory Comput.* **2024**, *20*, 6632−6651.
(44) Nishimoto, Y. Analytic First-Order Derivatives of CASPT2 Combined with the Polarizable Continuum Model. *J. Chem. Theory Comput.* **2025**, *21*, 730−746.
(45) Stålring, J.; Bernhardsson, A.; Lindh, R. Analytical gradients of a state average MCSCF state and a state average diagnostic. *Mol. Phys.* **2001**, *99*, 103−114.
(46) Osamura, Y.; Yamaguchi, Y.; Schaefer III, H. F. Generalization of analytic configuration interaction (CI) gradient techniques for potential energy hypersurfaces, including a solution to the coupled perturbed Hartree−Fock equations for multiconfiguration SCF molecular wave functions. *J. Chem. Phys.* **1982**, *77*, 383−390.
(47) Lu, T.; Chen, Q. van der Waals potential: an important complement to molecular electrostatic potential in studying intermolecular interactions. *J. Mol. Model.* **2020**, *26*, 315.
(48) Rackers, J. A.; Wang, Z.; Lu, C.; Laury, M. L.; Lagardère, L.; Schnieders, M. J.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. Tinker 8: software tools for molecular design. *J. Chem. Theory Comput.* **2018**, *14*, 5273−5289.
(49) Carnimeo, I.; Cappelli, C.; Barone, V. Analytical gradients for MP 2, double hybrid functionals, and TD-DFT with polarizable embedding described by fluctuating charges. *J. Comput. Chem.* **2015**, *36*, 2271−2290.
(50) Giovannini, T.; Lafiosca, P.; Chandramouli, B.; Barone, V.; Cappelli, C. Effective yet reliable computation of hyperfine coupling constants in solution by a QM/MM approach: Interplay between electrostatics and non-electrostatic effects. *J. Chem. Phys.* **2019**, *150*, 124102.
(51) Gómez, S.; Giovannini, T.; Cappelli, C. Multiple facets of modeling electronic absorption spectra of systems in solution. *ACS Phys. Chem. Au* **2023**, *3*, 1−16.
(52) Palmer, I. J.; Ragazos, I. N.; Bernardi, F.; Olivucci, M.; Robb, M. A. An MC-SCF study of the S1 and S2 photochemical reactions of benzene. *J. Am. Chem. Soc.* **1993**, *115*, 673−682.
(53) Schumm, S.; Gerhards, M.; Roth, W.; Gier, H.; Kleinermanns, K. A CASSCF study of the S0 and S1 states of phenol. *Chem. Phys. Lett.* **1996**, *263*, 126−132.
(54) Bao, J. J.; Truhlar, D. G. Automatic active space selection for calculating electronic excitation energies based on high-spin unrestricted hartree−fock orbitals. *J. Chem. Theory Comput.* **2019**, *15*, 5308−5318.
(55) Tóth, Z.; Pulay, P. Finding symmetry breaking Hartree-Fock solutions: The case of triplet instability. *J. Chem. Phys.* **2016**, *145*, 164102.
(56) Cardenas, G.; Nogueira, J. J. An algorithm to correct for the CASSCF active space in multiscale QM/MM calculations based on geometry ensembles. *Int. J. Quantum Chem.* **2021**, *121*, No. e26533.
(57) Biczysko, M.; Bloino, J.; Santoro, F.; Barone, V. *Computational Strategies for Spectroscopy*; John Wiley & Sons, Ltd, 2011; Chapter 8, pp 361−443.
(58) Macak, P.; Luo, Y.; Ågren, H. Simulations of vibronic profiles in two-photon absorption. *Chem. Phys. Lett.* **2000**, *330*, 447−456.
(59) Hazra, A.; Chang, H. H.; Nooijen, M. First principles simulation of the UV absorption spectrum of ethylene using the vertical Franck-Condon approach. *J. Chem. Phys.* **2004**, *121*, 2125−2136.
(60) Baiardi, A.; Bloino, J.; Barone, V. General time dependent approach to vibronic spectroscopy including Franck−Condon, Herzberg−Teller, and Duschinsky effects. *J. Chem. Theory Comput.* **2013**, *9*, 4097−4115.
(61) Cerezo, J.; Santoro, F. Revisiting vertical models to simulate the line shape of electronic spectra adopting Cartesian and internal coordinates. *J. Chem. Theory Comput.* **2016**, *12*, 4970−4985.
(62) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(63) Tange, O. *GNU Parallel 20240522*; Tbilisi, 2024.
(64) Cerezo, J.; Santoro, F. FCclasses3: Vibrationally-resolved spectra simulated at the edge of the harmonic approximation. *J. Comput. Chem.* **2023**, *44*, 626−643.
(65) Bernhardsson, A.; Forsberg, N.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. A theoretical study of the 1B2u and 1B1u vibronic bands in benzene. *J. Chem. Phys.* **2000**, *112*, 2798−2809.
(66) Löwdin, P.-O. *Advances in Quantum Chemistry*; Elsevier, 1970; Vol. *5*; pp 185−199.
(67) Aiken, J. G.; Erdos, J. A.; Goldstein, J. A. On löwdin orthogonalization. *Int. J. Quantum Chem.* **1980**, *18*, 1101−1108.
(68) Bloino, J.; Baiardi, A.; Biczysko, M. Aiming at an accurate prediction of vibrational and electronic spectra for medium-to-large molecules: an overview. *Int. J. Quantum Chem.* **2016**, *116*, 1543−1574.
(69) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483−5488.

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(70) Ambrosetti, M.; Skoko, S.; Giovannini, T.; Cappelli, C. Quantum mechanics/fluctuating charge protocol to compute solvatochromic shifts. *J. Chem. Theory Comput.* **2021**, *17*, 7146−7156.
(71) Gómez, S.; Lafiosca, P.; Egidi, F.; Giovannini, T.; Cappelli, C. Uv-resonance Raman spectra of systems in complex environments: A multiscale modeling applied to doxorubicin intercalated into dna. *J. Chem. Info. Model.* **2023**, *63*, 1208−1217.
(72) Giovannini, T.; Ambrosetti, M.; Cappelli, C. Quantum confinement effects on solvatochromic shifts of molecular solutes. *J. Phys. Chem. Lett.* **2019**, *10*, 5823−5829.
(73) Giovannini, T.; Lafiosca, P.; Cappelli, C. A general route to include Pauli repulsion and quantum dispersion effects in QM/MM approaches. *J. Chem. Theory Comput.* **2017**, *13*, 4854−4870.
(74) Jensen, J. H.; Gordon, M. S. An approximate formula for the intermolecular Pauli repulsion between closed shell molecules. II. Application to the effective fragment potential method. *J. Chem. Phys.* **1998**, *108*, 4772−4782.
(75) Amovilli, C.; Mennucci, B. Self-consistent-field calculation of Pauli repulsion and dispersion contributions to the solvation free energy in the polarizable continuum model. *J. Phys. Chem. B* **1997**, *101*, 1051−1057.
(76) Ruckenbauer, M.; Mai, S.; Marquetand, P.; González, L. Revealing deactivation pathways hidden in time-resolved photoelectron spectra. *Sci. Rep.* **2016**, *6*, 35522.
(77) Crespo-Otero, R.; Barbatti, M. Recent advances and perspectives on nonadiabatic mixed quantum−classical dynamics. *Chem. Rev.* **2018**, *118*, 7026−7068.
(78) Lischka, H.; Nachtigallova, D.; Aquino, A. J.; Szalay, P. G.; Plasser, F.; Machado, F. B.; Barbatti, M. Multireference approaches for excited states of molecules. *Chem. Rev.* **2018**, *118*, 7293−7361.
(79) Plasser, F.; Gómez, S.; Menger, M. F.; Mai, S.; González, L. Highly efficient surface hopping dynamics using a linear vibronic coupling model. *Phys. Chem. Chem. Phys.* **2019**, *21*, 57−69.
(80) Cigrang, L. L. E.; et al. Roadmap for Molecular Benchmarks in Nonadiabatic Dynamics. *J. Phys. Chem. A* **2025**, *129*, 7023−7050.

https://doi.org/10.1021/acs.jctc.5c01890
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX