JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

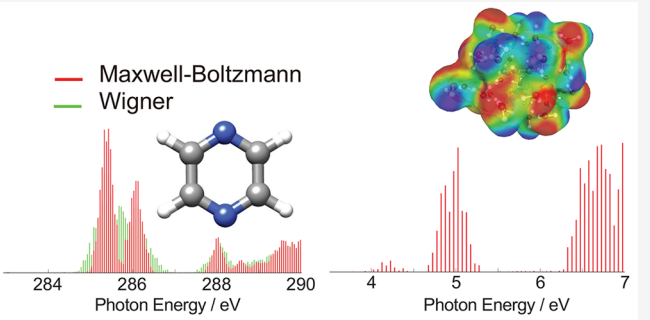
Structural Sampling and Solvation Models for the Simulation of Electronic Spectra: Pyrazine as a Case Study

Shota Tsuru, Bikramjit Sharma, Dominik Marx, and Christof Hättig*

Cite This: https://doi.org/10.1021/acs.jctc.2c01129 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: The impact of sampling methods on spectral broadening in the gas phase and on the convergence of spectrum in the aqueous solution when using microsolvation, continuum solvation, and hybrid models is studied using pyrazine as a test case. For the sake of comparing classical Maxwell−Boltzmann and Wigner samplings in the gas phase, static and time-resolved X-ray absorption spectra after photoexcitation to the lowest $^1B_{2u}(\pi\pi^*)$ state, as well as the static UV−vis absorption spectrum, are considered. In addition, the UV−vis absorption spectrum of pyrazine in the aqueous solution is also computed in order to systematically investigate its convergence with the number of explicitly included solvent shells with and without taking bulk solvation effects into account with the conductor-like screening model to represent implicit water beyond such explicit solute complexes. Concerning the static and time-resolved X-ray absorption spectra of pyrazine at the carbon K-edge as well as its UV−vis absorption spectrum in the gas phase, we find that these spectra obtained with Wigner and Maxwell−Boltzmann samplings substantially agree. For the UV−vis absorption spectrum in the aqueous solution, only the first two energetically low-lying bands converge quickly with the size of the explicitly included solvation shells, either with or without an additional continuum solvation taken into account. In stark contrast, calculations of the higher-lying excitations relying on finite microsolvated clusters without additional continuum solvation severely suffer from unphysical charge-transfer excitations into Rydberg-like orbitals at the cluster/vacuum interface. This finding indicates that computational UV−vis absorption spectra covering sufficiently high-lying states converge only if continuum solvation of the explicitly microsolvated solutes is included in the models.



1. INTRODUCTION

Fundamental studies in the field of physical chemistry often focus on quasi-rigid molecules such as aromatic compounds, nucleobases, and transition metal complexes to name but a few. For these kinds of systems in the gas phase, zero-point vibration barely affects the physical properties, and the vibrational broadening of various kinds of static spectra can usually be included via Franck−Condon (FC) factors. The photoinduced dynamics of such essentially rigid systems are mostly simulated by semiclassical trajectory surface hopping at finite temperature with initial conditions sampled from the Wigner distribution corresponding to the harmonic vibrational quantum ground state, thus at zero temperature. However, molecules of practical interest have often a flexible structure due to intramolecular large-amplitude motion or are embedded in fluctuating environments such as molecular liquids or soft matter matrices at ambient conditions providing intermolecular couplings. For example, linear polyenes are of major importance in many biological systems, such as carotenoids, for the process of vision and in materials science. The torsional motion of linear polyenes makes the reproduction of the UV−vis absorption spectra or excited-state dynamics simulations challenging, even for the simplest ones such as 1,3-butadiene and 1,3,5-hexatriene.[1−4] To the best of the authors' knowledge, nuclear dynamics simulations based on Tully's fewest-switches surface hopping (FSSH) algorithm[5] have never been conducted for polyenes, despite the advantages of on-the-fly sampling.

For molecules in solution, even if they are quasi-rigid, the vibrational wave functions of the solute/solvent system are practically unavailable, and Wigner sampling can hence not be applied except in various approximate ways, e.g. in the local harmonic or Gaussian approximations among others.[6] Moreover, solvent effects need to be included not only for the ground state but also for excited states for interpretation of spectra recorded for UV−vis absorption, fluorescence, or

Received: November 10, 2022

ACS Publications © XXXX American Chemical Society

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

phosphorescence emission, etc. Thus, proper inclusion of solvent effects at finite temperature is crucial for meaningful excited-state dynamics simulations of chromophores in solution, using either all-QM on-the-fly FSSH dynamics of explicit aqueous bulk solutions[7] (This pioneering work on photoinduced chemical processes in a liquid that seamlessly takes into account the full electronic structure of the corresponding periodic bulk solution is based on a realization of nonadiabatic Car−Parrinello molecular dynamics using Tully's surface hopping dynamics.[8,9]) or a similar QM/MM nonadiabatic surface hopping scheme in case of more complex condensed bulk phase environments.[10,11] About one to two decades later, either pure QM of finite (nonperiodic) systems or hybrid QM/MM electronic structure techniques of various levels of sophistication combined with Tully's FSSH dynamics are not only broadly accepted in the field but also readily accessible thanks to powerful and easy-to-use workflow software such as SHARC/COBRAMM[12] or INAQS.[13]

In this light, effects of sampling[14−17] on absorption spectra and excited-state dynamics as well as methods for modeling spectra of molecules in solution[18−24] have been intensively investigated recently. It has been shown that classical Maxwell−Boltzmann sampling, abbreviated herein by MB, does not always give distributions of coordinates and momenta that are comparable to those obtained with Wigner sampling.[14] If the initial conditions generated by different sampling methods in the ground state are transferred to the potential energy surface of a certain excited state, the difference of the initial conditions may affect the results obtained for the excited-state dynamics after a photoexcitation.[17] Wigner sampling can even provide extremely broad such phase space distributions such that they cover regions which are irrelevant in practical simulations.[16] Moreover, it has been observed a case where neither Wigner sampling nor sampling a classical distribution function from an ab initio molecular dynamics (AIMD) trajectory works well: For 2-nitronaphthalene, the shifts of the peaks in the UV−vis absorption spectrum due to displacement from the FC geometry were better reproduced in the computational spectrum based on the Wigner sampling, whereas the effects from the torsional dynamics of the nitro group, leading to intramolecular large-amplitude motion, were better reproduced with the sampling based on the AIMD simulation.[15]

For taking solvent effects on ground and excited states into account with low computational cost, implicit continuum models such as the polarizable continuum model (PCM)[25] or the conductor-like screening model (COSMO)[26] are efficient, albeit these models do not describe well strong noncovalent intermolecular interactions like hydrogen bonds between solute and solvent molecules. They also do not account for the spectral broadening due to fluctuations in the solvation shell. These drawbacks of the implicit continuum models are likely to be overcome with microsolvation models, consisting of the solute and several explicit solvent molecules, either used as such or embedded in continuum models such as PCM and COSMO or in an electrostatic MM environment in the QM/MM sense. The performance of such combinations has been studied in the literature for hybrid microsolvation plus continuum models and also for QM/MM models.[18,19,23,24] However, convergence of the solvent effects for absorption spectra in different energy regions with respect to the number of explicit solvent molecules (solvation shells) in the QM region has not been thoroughly discussed so far.

The present study is aimed at (i) the evaluation of the performance of rigorous MB sampling of classical nuclei at finite temperature in comparison to Wigner sampling of quantum harmonic oscillator modes at zero temperature in practical spectral simulations including time-resolved (TR) ones and (ii) the investigation of the convergence of the solvation effects on absorption spectra with respect to the size of the explicitly included solvation shell. We selected pyrazine as a test case, mainly due to availability of the static[27,28] and TR-XAS[28] spectra as well as the static UV−vis absorption spectrum[29,30] in the gas phase as references for aim (i). This selection also has an advantage concerning aim (ii) because of the prominent effect of hydrogen bonds between pyrazine and water molecules on the low-lying $n\pi^*$ excitations in the UV−vis spectrum. Both need to be described correctly for modeling TR-XAS spectra after photoexcitation into the lowest valence-excited states not only of pyrazine in water but also of many other chromophores in the aqueous solution for which the current study lays cornerstones.

2. METHODS

2.1. Ab Initio Molecular Dynamics Simulations for Maxwell−Boltzmann Sampling in the Gas and Aqueous Phases. For thermal MB sampling of the gas and aqueous phase systems under investigation at 300 K, we performed constant temperature (NVT) AIMD simulations using the CP2K simulation package[31,32] and its Quickstep module[33] for electronic structure calculations. The gas phase AIMD simulation system contained a single pyrazine molecule, while the aqueous phase system contained one pyrazine solvated with 112 water molecules using the same cubic supercell of length 15.0414 Å subject to periodic boundary conditions (also in the gas phase for consistency). This cell dimension was determined based on force field molecular dynamics (FFMD) simulations of the aqueous solution as detailed in the next paragraph. In order to ensure that ergodic MB sampling of all degrees of freedom was achieved, in particular in the difficult case of isolated molecules in vacuum, we thermostated the system at the given temperature with the massive Nosé−Hoover chain thermostat method.[34] This particular approach, where each and every Cartesian degree of freedom gets individually thermostated by its own Nosé−Hoover chain, is known to produce the canonical distribution even for molecular systems with stiff vibrational modes that are largely decoupled from other degrees of freedom. The revised version of the PBE functional, namely RPBE[35] along with the D3 dispersion correction[36] was used in the AIMD simulations. We performed the calculations with the atom centered Gaussian basis set TZV2P in conjunction with Goedecker-Teter-Hutter[37−39] pseudopotentials in order to replace the core electrons.

The lengths of the gas and aqueous phase trajectories used to sample statistically independent snapshots were 200 and 100 ps, respectively, after equilibration. Based on these two AIMD simulations, phase space snapshots were sampled every 1 ps in both gas and aqueous phases, thus providing 200 initial conditions to compute the static X-ray absorption spectrum and to launch FSSH simulations for computing the time-resolved X-ray absorption spectrum in the gas phase and 100 explicit solvation environments to compute the UV−vis absorption spectra in the aqueous solution as described in the following.

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

In order to determine the length of the periodic simulation box for the aqueous solution and its subsequent pre-equilibration in the NVT ensemble, we first performed FFMD simulations of one pyrazine molecule in 112 water molecules in the NpT ensemble using Gromacs.[40−43] Water was represented by the TIP4P/2005 water model[44] which scored best in reproducing a large set of experimental condensed phase properties of H_2O among many rigid, nonpolarizable force fields of water.[45] The force field for pyrazine was generated using the ACPYPE tool.[46]

2.2. Surface Hopping Simulations. We conducted the excited-state molecular dynamics simulations for pyrazine in the gas phase in a manner corresponding to that of ref 28. The FSSH method was employed as implemented in SHARC 2.1.1,[47−49] with TDDFT/B3LYP/cc-pVDZ as the electronic structure method adopting the RI-J and COSX approximations, which was provided by the ORCA[50,51] backend. Grid 3 was used for the DFT calculation. As initial conditions, 200 sets of coordinates and momenta generated by the AIMD simulation in the electronic ground state of pyrazine were used to carry out FSSH simulations corresponding to an UV laser field at resonance with the $^1B_{2u}$ state of pyrazine in vacuum that was enveloped by a Gaussian of 1.0 TW cm^{-2} intensity and 80 fs fwhm, up to the delay time of 340 fs (time zero was considered as the instant of the laser pulse maximum). The reference energy was −264.17706067897 hartree, the SCF energy at the DFT/B3LYP/cc-pVDZ level calculated at the FC geometry, which was optimized using ORCA at the DFT/RPBE-D3/def2-TZVPP level for consistency with the AIMD simulation. The vertical excitation energy of the $^1B_{2u}$ state, 5.4255 eV, was obtained as that calculated at the TDDFT/B3LYP/cc-pVDZ level at the FC geometry. An energy-based decoherence scheme was applied to the diagonal states.[52] A standard (nonrelativistic) electronic Hamiltonian, called molecular Coulomb Hamiltonian, without the inclusion of spin−orbit coupling (SOC) was utilized. The coefficients of the wave functions were propagated by the local diabatization technique using wave function overlaps.[53]

Applying this approach to the 200 phase space snapshots obtained from the AIMD simulations of pyrazine in the gas phase at 300 K, thus sampling the thermal MB distribution function as described in Sec. 2.1, a total of 34 out of the generated 200 FSSH trajectories underwent excitation from the ground state to the $^1B_{2u}$ state. The excited state configurations of these 34 trajectories at a propagation time of 220 fs, corresponding to the pump−probe time delay employed in the relevant TR-XAS experiment,[28] were used to compute the classical MB averaged TR-XAS spectrum of pyrazine in vacuum that is presented below.

The resulting TR-XAS spectrum obtained from MB sampling can be compared to that based on our previously published data obtained from Wigner sampling of the initial phase space conditions for the FSSH simulations, where 59 trajectories ended up in the excited state. We recapitulate only the salient aspects of that work here and refer the interested reader to the original publication[28] for more details. The default settings of the `wigner.py` module of SHARC 2.1.1 were used,[49] thus the vibrational quantum ground state (formally corresponding to a temperature of 0 K) of pyrazine was sampled in the harmonic approximation of its vibrational motion (as described by the B3LYP/6-311++G** electronic structure). In particular, Wigner's version of the phase space distribution function of a set of uncoupled quantum-mechanical harmonic oscillators was generated by starting from the equilibrium structure of the molecule in its electronic ground state and zero velocities in order to stochastically sample the coordinates and velocities along all normal modes of the molecule one after the other, which provided one initial condition in classical phase space to be used in the subsequent FSSH simulations as described above.

2.3. Simulations of Absorption Spectra. *2.3.1. X-ray Absorption in the Gas Phase.* We calculated the carbon 1*s* core excitation energies from the electronic ground state of pyrazine in the gas phase using all 200 configurations generated by the AIMD gas phase simulation and another 200 coordinates which were taken from the Zenodo database of ref 28. To compute the excited state TR-XAS spectrum, we then launched FSSH simulations and carried out spectral simulations for the excited pyrazine molecules in the gas phase at 220 fs in the same manner as in ref 28 and described in Sec. 2.2. Thus, excitation energies and oscillator strengths of the core excitations from the valence-excited states were calculated at the coordinates given by those 34 FSSH trajectories that were in the excited states at 220 fs. For the sake of consistency, these calculations were performed as in ref 28 using TDDFT/B3LYP in conjunction with the 6-311++G** basis set. The core−valence separation scheme was used to access the core-excited states,[54] and the maximum overlap method[55] was employed to obtain reference wave functions corresponding to the initial valence-excited states exactly as ref 28 to stay consistent.

The X-ray absorption spectra were obtained by plotting the oscillator strengths in histogram style and by shifting the theoretical values by +10.7 eV before comparison with the experimental data. The +10.7 eV constant shift was applied to correct a systematic error of TDDFT, attributable to relaxation of the electrons caused by the core hole and by self-interaction errors as already determined in ref 28. All spectral simulations for the core-to-valence transitions were performed with the Q-Chem 5.4 electronic structure package.[56]

2.3.2. UV−Vis Absorption in the Gas Phase. We also calculated the valence excitation energies from the electronic ground state of pyrazine in the gas phase for the same set of coordinates as used for the core excitation energies in Sec. 2.3.1 with the algebraic−diagrammatic construction through second order method, ADC(2),[57] in combination with the resolution-of-identity (RI) approximation, adopting the aug-cc-pVTZ basis set. All the calculations of the valence excitations of the present work were performed with the TURBOMOLE V7.6 electronic structure package.[58,59] In each excited-state calculation, 12 roots were computed to include the third absorption band. The UV−vis absorption spectra based on the MB and Wigner samplings were obtained by plotting the oscillator strengths in histogram style and compared with two experimental spectra: one from ref 29 covering the second and third and some higher absorption bands and one from ref 30 covering the first and second bands. These UV−vis spectral simulations in direct comparison to experiments were done not only to investigate the effects due to using different sampling methods but also to confirm the accuracy of RI-ADC(2) as electronic structure method. Importantly, the well-known self-association of pyrazine dissolved in water prevents such a confirmation by comparison with the experimental UV−vis absorption spectra of aqueous solutions of pyrazine.[60,61]

2.3.3. UV−Vis Absorption in the Aqueous Solution. For the investigation of the solvent effects on the UV−vis spectra

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

of pyrazine dissolved in water at ambient conditions, we calculated the valence excitation energies also with the electronic structure model RI-ADC(2) but now in combination with the continuum solvation model COSMO, in the following referred to as COSMO-RI-ADC(2).[62] To systematically study the effect of solvation on UV−vis spectrum, two solvation shells, namely the first and second solvation shells, were identified around the pyrazine nitrogens based on the radial distribution function of water oxygen (OW) around these two atoms calculated from the AIMD trajectory. The first and second minima at 3.15 and 5.20 Å of that radial distribution function were used to demarcate the first and second solvation shells, respectively. Using this approach, the average coordination number of each nitrogen site turned out to be ∼0.99 (the minimum and maximum numbers encountered during the AIMD simulations being zero and two whereas three is found to be negligible). The entire second solvation shell around the two nitrogen atoms contained on average about 23.8 water molecules in total (the minimum and maximum numbers being 10 and 32, respectively). From the 100 configuration snapshots sampled from the AIMD trajectory for pyrazine in the aqueous solution, we extracted coordinates for (i) only pyrazine, (ii) pyrazine and water molecules in the first solvation shells of both pyrazine nitrogens, and (iii) pyrazine and water molecules in the first and second solvation shells of both pyrazine nitrogens, in the following denoted as (i) "bare pyrazine", (ii) "pyrazine+1st shell", and (iii) "pyrazine+2nd shell", respectively. One snapshot of pyrazine+second shell together with the surface charge induced on the COSMO cavity is displayed in Figure 1.

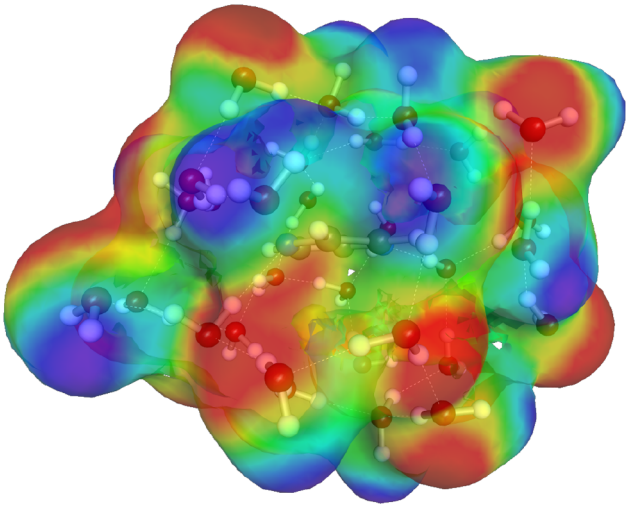


Figure 1. Representative configuration snapshot of pyrazine together with the first and second solvation shells around its two nitrogen sites sampled from the AIMD trajectory of pyrazine in explicit bulk water at 300 K together with the surface of the COSMO cavity. The color of the surface indicates the surface charge induced by the quantum system (positive charge in red, negative charge in blue).

Using this approach, we calculated the valence excitation energies up to 7 eV using ADC(2) for all three structural models in conjunction with the COSMO continuum solvation model to account for bulk solvation effects, thereby considering implicit solvation beyond bare pyrazine and pyrazine with explicit first/second solvation shell. We used the post self-consistent field reaction field scheme as described in refs 62 and 63 with the dielectric constant set to 80.4, the refractive index set to 1.33, and default parameters for the cavity construction. We adopted the aug-cc-pVTZ and aug-cc-pVDZ basis sets to the pyrazine and water molecules, respectively, as a compromise between accuracy and computational costs. Test calculations for pyrazine+first shell with COSMO indicated that for the lowest six excitations the results obtained with this mixed basis set deviate by less than 0.02 eV from calculations using the aug-cc-pVTZ basis at all atoms.

To test for the importance of the bulk solvation effects included by COSMO, we did similar calculations for the three structural models without COSMO. For the bare pyrazine and the pyrazine+first shell models, we computed again all states up to 7 eV. For the pyrazine+second shell model, however, the appearance of a huge number of unphysical surface excitations (vide infra) made this technically unfeasible. Therefore, for this setup, we limited the calculations to the 12 lowest excitations, the maximum needed in any of the other setups to cover the energy range up to 7 eV.

We also conducted single-point calculations of the vertical excitation energies of pyrazine at the FC geometry using (COSMO-)RI-ADC(2)/aug-cc-pVTZ, aiming to demonstrate the configuration characters with the natural transition orbitals (NTOs)[64−67] and the squares of the corresponding singular values σ_K which indicate the weight of the respective configuration in the excited-state eigenvector.

Note that reproduction of the well-known self-association of pyrazine dissolved in water reported in ref 61 is not intended in the present work, and therefore, the calculated spectra are not compared to the experimental UV−vis absorption spectra of the pyrazine aqueous solution.

3. RESULTS AND DISCUSSION

3.1. Comparison of the Sampling Methods for XAS Spectra in the Gas Phase. We first show the comparison of the ground state XAS spectra at the carbon K-edge (Figure 2) calculated for the 200 configurations obtained from (a) the MB sampling generated by the present AIMD simulation and (b) the Wigner sampling taken from the Zenodo database of ref 28 plotted in histogram style, as well as the experimental spectrum taken from Figure 2 of ref 28, which is in good agreement with previous data obtained at a synchrotron facility.[27] The computational XAS spectra obtained from both sampling methods reproduce the positions and intensities not only of the main peaks but also their shoulders as seen in the experimental XAS spectrum in reasonably good agreement. The contrast between the spectra from MB and Wigner samplings is seen only in the width of the computed spectra. Although the lifetime of core-excited states is shorter than that of valence-excited states, the lifetime width observed in the carbon 1*s* photoelectron spectra is about only 0.1 eV.[68] Since the observed lifetime width is sufficiently smaller than the fwhm of this experimental XAS spectrum, namely 0.4 eV, the geometry effects are dominant also in XAS spectra. Therefore, this performance of MB sampling supports future application to other molecules or molecules dissolved in solvents for assignments of the peaks in their XAS spectra.

Before comparing the TR-XAS spectra, it should be emphasized that the two sampling methods differ in their philosophy in several important ways. AIMD based MB sampling of classical nuclear motion at finite temperature T rigorously samples the classical rotational and vibrational

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

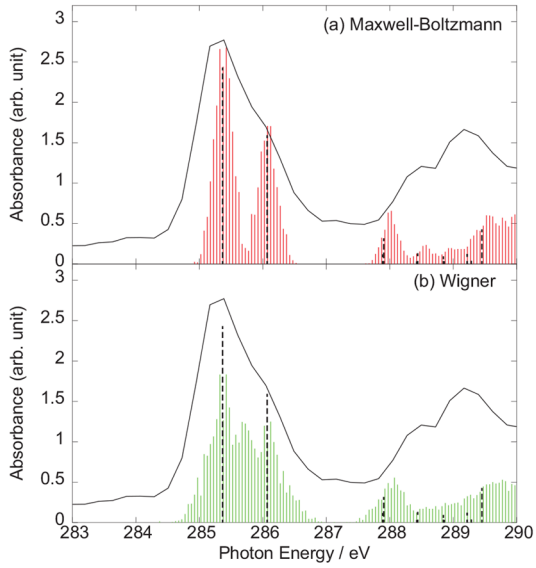


Figure 2. Carbon K-edge core excitation energies from the electronic ground state of pyrazine calculated for the (a) Maxwell−Boltzmann and (b) Wigner ensembles of pyrazine in the gas phase. The experimental spectrum, shown as black solid lines, and the calculated vertical excitation energies, black dashed vertical lines, included in both panels are taken with permission from Figure 2 of ref 28. Copyright 2021 The Author(s).

degrees of freedom of an isolated pyrazine molecule. The corresponding average total energy is equally distributed according to the equipartition theorem where each quadratic degree of freedom contains on average a thermal energy of 1/2 $k_B T$. Thus, high and low frequency vibrational modes get the same amount of kinetic energy in the limit of purely harmonic vibrations, which is distinctly different for the corresponding zero-point motion in quantum mechanics which is frequency dependent. Moreover, classical MB sampling also accounts for ro-vibrational couplings and full anharmonicities according to classical nuclear motion on the dynamically sampled potential energy surface described by the particular electronic structure method that is used in the AIMD simulations. So-called Wigner sampling, which is sometimes grossly referred to as quantum sampling,[14,15] on the other hand, maps the quantum-mechanical vibrational ground state in the harmonic approximation to the phase space distribution according to the Wigner phase space function to generate initial configurations and velocities; note that there are many different ways to carry out this mapping such as the Husimi distribution function[69] that is a proper phase space probability distribution function which is nowhere negative, in stark contrast to the Wigner function which is only normalized. In the harmonic oscillator limit employed here, each vibrational nuclear degree of freedom is assigned its harmonic oscillator zero-point energy, which is larger (smaller) for higher (lower) vibrational frequencies. Thus, using such initial phase space conditions to launch dynamical simulations using classical nuclei, such as FSSH, leads to the well-known zero-point energy leakage problem, which implies systematic kinetic energy flow from stiff to soft vibrational modes to eventually establish energy equipartition. This artifact is absent when such initial conditions for FSSH dynamics are consistently sampled from the classical MB distribution function as generated here from AIMD simulations. Moreover, the presently used Wigner initial conditions include neither finite temperature effects nor anharmonic and ro-vibrational couplings as customarily done in the present context, while noting that applying Wigner sampling that rigorously includes finite temperature, full anharmonicity, and ro-vibrational coupling effects would be a daunting task.

Figure 3 shows the excited state XAS spectra calculated at a FSSH propagation time of 220 fs for (a) MB distribution and

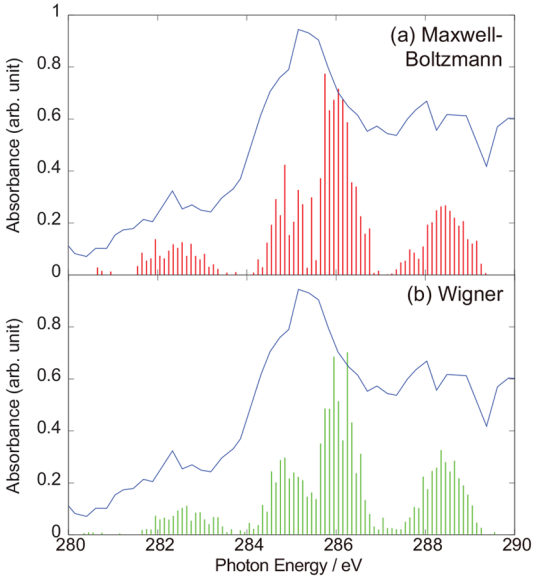


Figure 3. Carbon K-edge core excitation energies from the valence excited states of pyrazine at a FSSH propagation time of 220 fs calculated for the (a) Maxwell−Boltzmann and (b) Wigner ensembles (using the same initial states as those underlying Figure 6(b) of ref 28) of pyrazine in the gas phase. The blue lines in both panels are the experimental spectrum measured at a pump−probe time delay of 220 fs taken with permission from Figure 6(a) of ref 28. Copyright 2021 The Author(s).

(b) Wigner distribution as well as the experimental spectrum taken from Figure 6(a) of ref 28. Note that the histogram in Figure 3(b) is computed here using the same initial states as those underlying Figure 6(b) of ref 28. Satisfactory agreement between the MB and Wigner sampled spectra is obtained, despite the differences in widths seen in the ground state spectra depicted in Figure 2. This agreement of the excited state spectra indicates that the excitation generated by the explicit pump process of a pump−probe scheme experiment usually occurs in a region around the FC geometry, which many of the trajectories pass during the pump process regardless of slight differences in the initial geometries. One might wonder whether the close resemblance of the excited-state XAS spectra obtained using MB and Wigner initial conditions for FSSH propagation is consistent with several precedent studies where differences in the population dynamics were observed.[14,17] Concerning this comparison, it should be emphasized that the explicit laser field of the pump process has been taken into account only in the FSSH simulations of the present work. In addition, massive Nosé−Hoover chain thermostatting as used in the present AIMD simulations to sample the phase space initial conditions according to the MB distribution function is a method that ergodically thermalizes all vibrational modes even for molecular systems with stiff vibrational modes with weak

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

couplings to other degrees of freedom. As demonstrated in the next section, we also find overall agreement between our MB and Wigner samplings when it comes to computing the UV−vis spectrum of pyrazine in the gas phase subject to molecular vibrations. Thus, properly generating the initial conditions in the framework of MB sampling, i.e. sampling the snapshot configurations in phase space, in conjunction with including the pump process might be a decisive combination. Clearly, a definite answer would require systematic investigations of several representative systems thus leading beyond this study on pyrazine.

Considering the results of the comparisons above, our particular approach of FSSH propagation based on classical MB initial conditions obtained from canonical AIMD simulations at finite temperature is found to work with reasonable accuracy compared to experiment; it also compares satisfactorily to using Wigner initial conditions obtained from harmonic normal modes in the vibrational ground state. Therefore, this FSSH approach combined with MB sampling may be applied to other molecules in the gas phase or molecules in solution not only for calculations of static ground state spectra but also for simulations of time-resolved experiments provided that the potential energy surfaces are sufficiently accurate. Yet, the fundamental limitations of FSSH propagation to approximate the exact nonadiabatic quantum dynamics must be kept in mind when it comes to intricate nuclear quantum effects occurring in the excited state dynamics coupled to the ground state.

3.2. Comparison of the Sampling Methods for UV−Vis Spectrum in the Gas Phase. We show UV−vis absorption spectra of pyrazine in the gas phase calculated at the RI-ADC(2)/aug-cc-pVTZ level for the 200 snapshots of (a) MB and (b) Wigner samplings in Figure 4. The black dashed bars indicate the valence excitation energies calculated at the FC geometry optimized with DFT/RPBE-D3 adopting the def2-TZVPP basis set. The experimental data of refs 29 and 30 are shown for evaluating the accuracy of the computational first, second, and third absorption bands. Although the excitation energies calculated at the FC geometry are 0.2−0.3 eV higher than the centers of the absorption bands due to an intrinsic error of ADC(2), the distances between the absorption bands and the relative intensities are well reproduced even in the vertical excitation energies. This agreement indicates the accuracy of RI-ADC(2) for practical UV−vis spectral simulations.

Qualitatively, the difference between the peaks of the experimental and the computed spectra is slightly smaller when using the 200 sampled configurations instead of the single configuration at the FC point. In addition, the envelope of the first absorption band is also well reproduced using the two sampling methods. The second absorption band in the sampled spectra covers almost the entire energy range of that in the experimental spectrum, subject to the aforementioned slight ADC(2)-shift to higher energies, and a bit more broadening when using Wigner sampling. When it comes to the third band, the MB-based spectrum also reproduces the shape of the low-frequency flank of the corresponding experimental peak, again subject to the systematic overall shift to higher energies due to using ADC(2), whereas Wigner sampling generates a different peak structure with a rather high intensity plateau to the left of the sharp peak maximum. Yet, both sampling methods qualitatively reproduce fairly consistently the systematic broadening of the three peaks due to molecular vibrations as seen in experiment, thus providing information beyond the single-point calculation at the FC point.

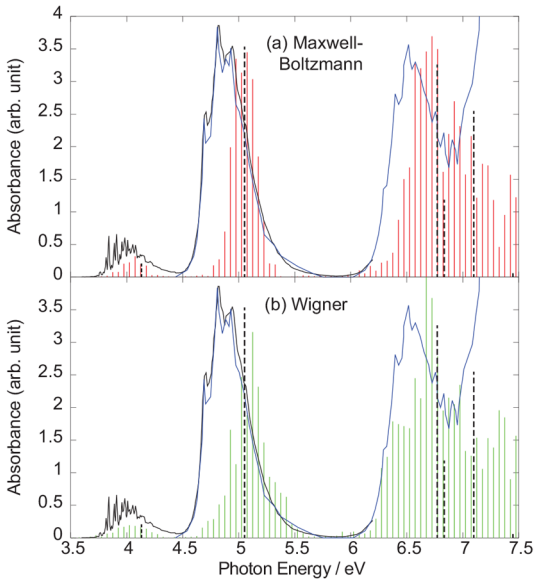


Figure 4. UV−vis spectra of pyrazine in the gas phase calculated for the (a) Maxwell−Boltzmann and (b) Wigner ensembles at the RI-ADC(2)/aug-cc-pVTZ level. The black dashed bars indicate the excitation energies calculated at the optimized geometry imposing the D_{2h} symmetry. The experimental spectra, shown as black and blue solid lines, are taken with permission from Figure 1 of ref 30 (energy window: 3.6−6.2 eV, Copyright 2020 Elsevier B.V.) and Figure 4 of ref 29 (energy window: 4.5 eV, Copyright 2011 IOP Publishing Ltd.), respectively.

Relying on the confirmed accuracy of the electronic states calculated with RI-ADC(2) and the performance of explicit configurational sampling for spectral simulations taking molecular geometry effects into account, we evaluate another important aspect of the simulation of electronic spectra of molecules dissolved in a solvent, namely the computational description of solvation effects depending on the solvation model, in the subsequent part of this investigation.

3.3. Solvent Effects on UV−Vis Spectrum. In the next step, we address the effects of implicit, explicit, and mixed solvent models on the spectrum in solution. We concentrate on the three absorption bands in the energy window from 3.5 to 7.0 eV. Here, we denote the first, second, third, etc. excited states with $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, ..., respectively, because the nomenclature S_1, S_2, ... for pyrazine may be confusing here since it conventionally skips optically dark states.

In order to set the stage for understanding solvation effects on the UV−vis spectrum of pyrazine in water, we determined the characters of the excited states which contribute to the three absorption bands of pyrazine based on NTO analysis. Table 1 shows the energies, the oscillator strengths, and NTOs of excitations from the ground state ($\tilde{X}$) at the optimized equilibrium structure to the $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{E}$, and $\tilde{F}$ states of pyrazine dissolved in water, calculated at the COSMO-RI-ADC(2)/aug-cc-pVTZ level, thus taking the solvent fully implicitly into account using the COSMO continuum solvation model. Excitation to the $\tilde{D}$ state is not discussed in the present work, because the $\tilde{D}$ state is optically dark and does not contribute to any absorption band. The $\tilde{A}$ state has a $n\pi^*$ character, where the n orbital is the symmetric linear

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Energies, Oscillator Strengths, and NTOs of Valence Excitations at the Optimized Ground-State Equilibrium Structure Calculated with COSMO

	Ex. energy (eV)	Osc. strength	Hole	σ_K^2	Particle
$\tilde{A}$	4.2168	0.0072		0.99	
$\tilde{B}$	4.9250	0.1432		0.89	
				0.10	
$\tilde{C}$	5.0063	-		1.00	
$\tilde{E}$	6.4760	-		0.99	
$\tilde{F}$	6.6834	0.1708		0.87	
				0.09	
				0.02	

combination of the nitrogen lone pairs (see "Hole" in line $\tilde{A}$ of Table 1). Excitation to this state forms the first absorption band. The $\tilde{B}$ state has (at this level and geometry) a $\pi\pi^*$ character and dominantly contributes to the second absorption band. The $\tilde{C}$ state at this geometry with COSMO has also $n\pi^*$ character, where the n hole is common with that of $\tilde{A}$ (see "Hole" in lines $\tilde{A}$ and $\tilde{C}$ of Table 1). Although the excitation to $\tilde{C}$ is optically dark when the pyrazine molecule has the D_{2h} symmetry, the characters of $\tilde{B}$ and $\tilde{C}$ may mix or even switch with slight conformational changes since they are rather close in energy. Therefore, as we will disclose below when considering the MB ensemble of configurations extracted from the AIMD ground state trajectory instead of just using the optimized FC structure, the second absorption band has also contributions from the $\tilde{C}$ state in the UV−vis spectrum of pyrazine. The excitation to the $\tilde{E}$ state is found to be a Rydberg transition in this COSMO-RI-ADC(2) calculation (see "Particle" of line $\tilde{E}$ of Table 1). It may, however, mix with the $\tilde{F}$ state at the distorted geometries sampled from the AIMD simulation due to the small energetic separation of 0.21 eV and also with further states. The transition to the $\tilde{F}$ state is dominantly of the $\pi\pi^*$ type, albeit the NTO analysis also shows a small mixture with a Rydberg transition as shown in "Particle" of the third configuration.

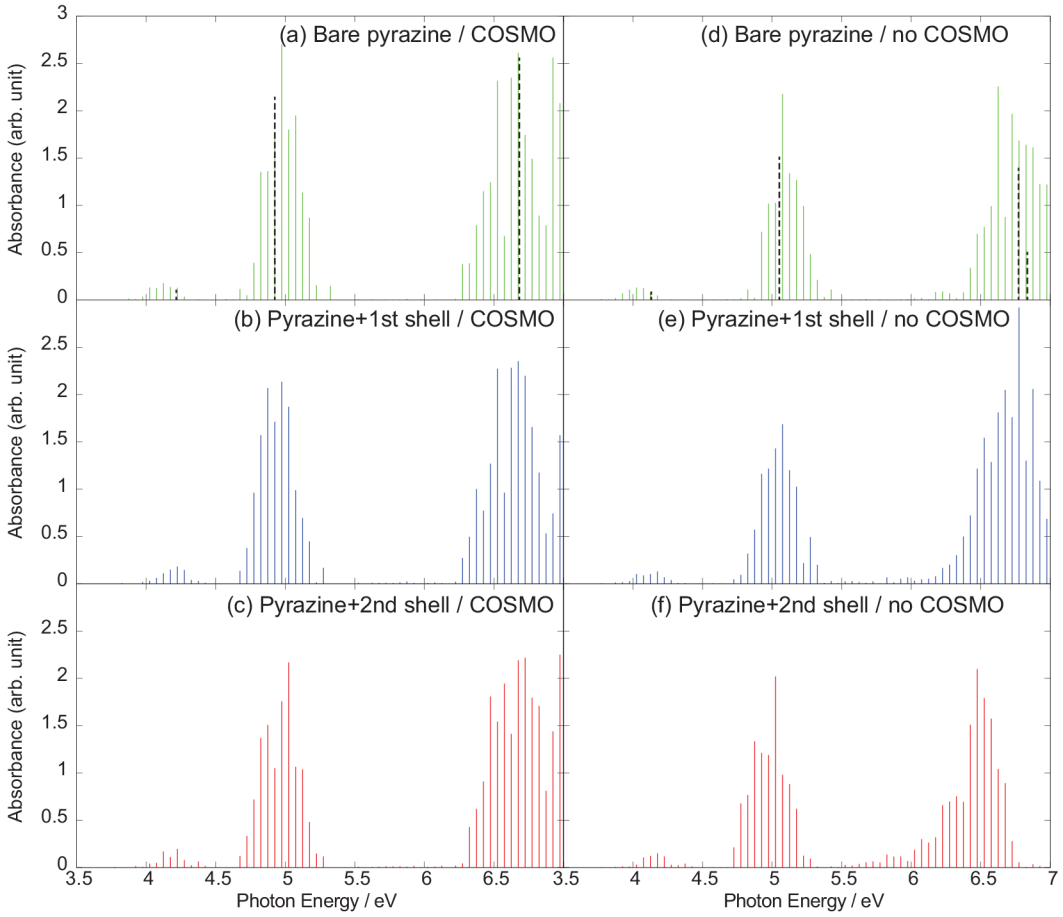


Figure 5. UV−vis spectra of pyrazine in the aqueous solution calculated for (a,d) bare pyrazine, (b,e) pyrazine+first shell, and (c,f) pyrazine +second shell. Calculations for panels (a,b,c) and (e,f,g) are conducted with and without COSMO, respectively. In panels (a) and (d), excitation energies calculated at the FC geometries are shown in black dashed bars with oscillator strength multiplied by 15. Note that the spectrum in panel (f) is only complete up to 6 eV due to limiting the calculations to the lowest 12 excitations (see Sec. 2.3.3 for details).

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Based on this knowledge, we can now proceed with a comparison of the UV−vis spectra obtained with the different solvent models using the configurational snapshots sampled from the AIMD simulation of pyrazine in water, thus providing MB sampling of both, the structural distortions of pyrazine itself as well as of the fluctuations of its first (first) and second (second) solvation shells in the aqueous phase at 300 K. We refer to Sec. 2.3.3 for the details on how these solvation shells have been defined in terms of the coordination numbers of the two nitrogen atoms but recall here that the first solvation shell of each nitrogen atom includes one water molecule on average, whereas their combined second solvation shell contains in total from 10 to 32 water molecules and on average about 24. The f5 left column of Figure 5 shows the UV−vis spectra of pyrazine in the aqueous solution calculated for (a) bare pyrazine, (b) pyrazine+first shell, and (c) pyrazine+second shell, including COSMO to take bulk solvation effects into account beyond the respective finite structural models, which includes explicit microsolvation in variants (b) and (c). The peak maxima of the three absorption bands in the spectrum calculated for bare pyrazine agree with the vertical excitation energies shown in Table 1, as expected from the rigidity of the pyrazine molecule. The changes of the absorption bands with the inclusion of explicit water molecules are mostly minor. As the number of explicit water molecules is increased, the center of the first absorption band shifts a bit to the blue, while those of the second and third absorption bands slightly red-shift. Moreover, all three bands get a bit broader as a result of solvation. The right column of Figure 5 shows the results of the simulation with exactly the same structures as those for the left column apart from the absence of COSMO continuum solvation beyond the corresponding finite structural models. As one compares panels (a) and (d), the first absorption band is found to be blue-shifted when COSMO is included. This difference of the position of the first absorption band is barely discernible between panels (b) and (e). In contrast, looking at the second absorption band in panels (d), (e), and (f), a significant red shift of the band in response to the inclusion of the explicit water molecules is observed. Only after inclusion of explicit water molecules in the second solvation shell, the position of the second absorption band without and with COSMO are practically the same as seen in panels (c) and (f). The third absorption band, however, looks very different if the calculations are done without COSMO embedding of the respective microsolvation complexes. It exhibits a tail structure at 5.5−6.3 eV, which is absent in the calculations with COSMO, but most strikingly, it shows no convergence with the size of the explicitly included microsolvation shell which indicates that this band is not adequately described by these models and that, instead of the physical correct states, transitions to artificial surface states are obtained.

To check the convergence of the results with the number of snapshots, we divided the complete set into two subsets and t2 compared the results. Table 2 shows the excitation energies averaged over the 50 snapshots sampled at even or odd times (in ps after equilibration) as well as the total average of all 100 snapshots. Since the difference of the excitation energy of each state between the even and odd sets is less than 0.1 eV, which is lower than the accuracy expected from ADC(2) for valence transitions in closed-shell molecules dominated by single excitations, ensemble convergence of the excitation energies can be assumed already with 50 configuration that have been equidistantly sampled from the 100 ps AIMD simulation of pyrazine in water.

Table 2. Excitation Energies Averaged over 50 Snapshots with Even Numbers, Those with Odd Numbers, and All 100 Snapshots, of the Pyrazine+2nd Shell Model with COSMO

	even set	odd set	total set
$\tilde{A}$	4.1686	4.1870	4.1778
$\tilde{B}$	4.8361	4.8434	4.8397
$\tilde{C}$	5.0984	5.0635	5.0809
$\tilde{D}$	5.8225	5.9337	5.8781
$\tilde{E}$	6.5336	6.5579	6.5458
$\tilde{F}$	6.6590	6.7185	6.6887

To further analyze the solvent effects on the excited states and how they are captured by the six models, we did a comparison of the excitation energies and the transition characters for the six models using the 50th snapshot, where the NTOs are close to those at the optimized ground-state equilibrium structure. Table 3 t3 shows the energies and the oscillator strengths of the excitations from the ground state to the six lowest states calculated for the bare model, the first-shell model, and the second-shell model, of the 50th snapshot computed without and with COSMO. Tables 4 and 5 t4t5 show the NTOs of these excitations (except for that to $\tilde{D}$ which has very low intensity compared to all six models) obtained in the calculations without and with COSMO, respectively. In the calculation for the pyrazine+second shell model without COSMO, the $\tilde{E}$ and $\tilde{F}$ states are quasi-degenerate and have similar configuration characters. Therefore in Tables 3 and 4, the data of the $\tilde{F}$ and $\tilde{G}$ states are shown instead of $\tilde{E}$ and $\tilde{F}$, respectively.

Although we observed good performance of the bare solute and COSMO model for the description of the three lowest excited states of pyridazine in our previous work,[23] it does not fully describe even the $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ states of the very similar pyrazine molecule in all details as seen by the results for snapshot number 50: For the $n\pi^*$ transitions, the excitation energies of $\tilde{A}$ and $\tilde{B}$ in bare pyrazine with the COSMO model are lower than those calculated for the pyrazine+second shell models by 0.2 and 0.3 eV (note that in the first-shell and second-shell models the second $n\pi^*$ transition is $\tilde{C}$). This significant additional blue shift due to the explicit water molecules originates from the stabilization of the nitrogen lone pairs by hydrogen bonds in the case of pyrazine, which is not captured by COSMO. Moreover, the switch of the energy ordering of the second and third singlet states caused by the water molecules (compare Tables 3 and 5) in the first solvation shell for snapshot number 50 is not observed for bare pyrazine when only implicitly solvated with the COSMO model. We note, however, that due to the small energy separation, the ordering of these two states does also depend on small fluctuations in the geometry: At the FC geometry, the $\pi\pi^*$ transition calculated for the bare pyrazine with COSMO is already below the second $n\pi^*$ transition (compare Table 1). This difference in the performance of the continuum solvation model between pyridazine and pyrazine is most likely related to the fact that pyrazine is apolar contrary to pyridazine, yet the two nitrogens of pyrazine have strong hydrogen bond with water molecules, and that this makes the effects of explicit inclusion of the hydrogen bonds essential: In pyridazine, the excitation energies for the lowest two $n\pi^*$ transitions obtained for the average of the sampled MB ensemble describing

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 3. Excitation Energy in eV (and Oscillator Strength in Parentheses) of Each Valence-Excited State Calculated for the 50th Snapshot Modeled by Bare Pyrazine, Pyrazine with the First Solvation Shell, Pyrazine with the First and Second Solvation Shells, and These Three Combined with COSMO[a]

	bare		+first shell		+second shell	
	no-COSMO	COSMO	no-COSMO	COSMO	no-COSMO	COSMO
$\tilde{A}$	3.8287	3.9190	3.9615	4.0714	4.1209	4.1147
	(0.0061)	(0.0081)	(0.0052)	(0.0076)	(0.0069)	(0.0078)
$\tilde{B}$	4.5689	4.7590	4.8363	4.8702	4.8880	4.8630
	(0.0007)	(0.0060)	(0.0084)	(0.1155)	(0.0780)	(0.0969)
$\tilde{C}$	5.0670	4.9654	5.0229	5.0560	5.0750	5.0827
	(0.0882)	(0.1258)	(0.0897)	(0.0220)	(0.0140)	(0.0125)
$\tilde{D}$	5.3312	5.5393	5.4050	5.6457	5.6843	5.6754
	(0.0005)	(0.0008)	(0.0045)	(0.0044)	(0.0029)	(0.0032)
$\tilde{E}$	6.0448	6.2245	5.9072	6.6287	6.2918	6.7066
	(0.0024)	(0.0022)	(0.0007)	(0.0010)	(0.0275)	(0.1049)
$\tilde{F}$	6.1908	6.5050	6.2617	6.7077	6.3119	6.7326
	(0.0049)	(0.0103)	(0.0017)	(0.0433)	(0.0014)	(0.0159)

[a]For the pyrazine+2nd shell model without COSMO, the data of the $\tilde{F}$ and $\tilde{G}$ states are shown instead of $\tilde{E}$ and $\tilde{F}$, respectively.

Table 4. NTOs of Excitations from the Ground State to the $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{E}$, and $\tilde{F}$ States Depicted for the Models *without* COSMO[a]

	Bare		+1st shell		+2nd shell	
	Hole	Particle	Hole	Particle	Hole	Particle
$\tilde{A}$						
$\tilde{B}$						
$\tilde{C}$						
$\tilde{E}$						
$\tilde{F}$						

[a]For the pyrazine+2nd shell model, the data of the $\tilde{F}$ and $\tilde{G}$ states are shown instead of $\tilde{E}$ and $\tilde{F}$, respectively.

Table 5. NTOs of Excitations from the Ground State to the $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{E}$, and $\tilde{F}$ States Depicted for the Models *with* COSMO

	Bare		+1st shell		+2nd shell	
	Hole	Particle	Hole	Particle	Hole	Particle
$\tilde{A}$						
$\tilde{B}$						
$\tilde{C}$						
$\tilde{E}$						
$\tilde{F}$						

solvation in the aqueous solution with the first solvent shell and COSMO included are blue-shifted by 0.4 and 0.5 eV. The COSMO calculation at the FC point without explicit waters captured ≈0.3 and 0.4 eV of this shift.[23] In the apolar pyrazine, the blue shifts of the first two $n\pi^*$ states are smaller, but with only ≈0.1 eV, also a smaller fraction of it is captured by the implicit solvent model.

From these results, one can conclude that simple replacement of all explicit solvent molecules by a continuum solvation model such as COSMO does not satisfactorily describe the solvent effects for pyrazine in the aqueous solution even for these low-lying valence excitations. Despite these changes, the differences between the absorption spectra obtained with the bare pyrazine, pyrazine+first shell, and pyrazine+second shell models that include COSMO continu-

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

um solvation (panels (a), (b), and (c) in Figure 5) are minor, mainly due to the fact that the $n\pi^*$ transitions have low intensities and that the effects of explicit water molecules on the bright $\pi\pi^*$ transitions are smaller: A red shift of about 0.1 eV is found for the first $\pi\pi^*$ state, which is obtained at 4.97 eV as $\tilde{C}$ in the bare pyrazine with the COSMO model and at 4.86 eV as $\tilde{B}$ in the pyrazine+second shell with the COSMO model.

Agreement between the excitation energies of the $\tilde{A}$, $\tilde{B}$, and $\tilde{C}$ states calculated for the second-shell models without and with COSMO indicates convergence of the solvent effects in the second solvation shell for these lowest excited states. Thus, the second solvation shell can practically be replaced by COSMO continuum solvation for this spectral region.

The situation turns out to be quite different for the higher-lying $\tilde{E}$ and $\tilde{F}$ states. With the first solvation shell included in the microsolvation complex but without additional implicit COSMO solvation beyond, a charge transfer from the water molecules to pyrazine is found. With the second solvation shell included, these transitions change into a Rydberg transition in the outer solvent shell and a charge-transfer excitation from pyrazine to the outer solvent shell. These are the unphysical states mentioned in Sec. 2.3.3 which made it unfeasible to compute all excited states for the pyrazine+second shell model up to 7 eV because of the sheer number of such states between 6 and 7 eV for this size of the microsolvated structure model. In the calculations for the same models but with COSMO solvation, excitations of these types are not observed in this energy region, and this observation accords with general expectations for such chromophores in water. This qualitative change of the configuration characters is explicable by the presence of those hydrogen atoms of water molecules that are not involved in hydrogen bonding, so-called dangling or free OH bonds, at the cluster/vacuum interface of the micro-solvation models. If COSMO is included, the excess of positive charge that accumulates at these hydrogens is balanced by negative charges on the cavity surface that delimits the conductor-like continuum. In the models without such additional continuum embedding, however, these regions at the interface to vacuum lead to artificially low-lying charge-transfer and Rydberg excitations. Meanwhile, also different from what is observed for the lowest excitations, a combination of the explicit first solvation shell and COSMO is still insufficient for a full description of the solvent effects: The Rydberg-type excitation ($\tilde{E}$ in the first two columns of Table 5) is only quenched when the second solvent shell is included as explicit water molecules in the QM system. Based on this analysis, we conclude that explicit water molecules in the first and second solvation shells should be taken into account for any meaningful description of the third absorption band without unphysical excited states, replacing the rest of the solvent water molecules with COSMO, although the effects on the bright $\pi\pi^*$ transitions that dominate the absorption spectrum are small.

4. CONCLUSIONS AND OUTLOOK

In the present work, we have evaluated the effects of structure sampling on static and time-resolved X-ray absorption spectra and on the UV−vis absorption spectrum in the gas phase, as well as the convergence of the solvent effects on UV−vis absorption spectrum with respect to the number of explicit solvent molecules, selecting pyrazine in water as a test case.

In the static spectral simulations for X-ray absorption in the gas phase, we computed the XAS spectrum based on sampling pyrazine configurations from the classical Maxwell−Boltzmann distribution at 300 K generated by ab initio molecular dynamics simulations in the canonical ensemble and based on sampling the quantum Wigner distribution of the harmonic vibrations of pyrazine at 0 K. The computational XAS spectra obtained from both sampling methods reproduce the positions and intensities not only of the main peaks but also their shoulders as seen in the corresponding experimental spectrum in reasonably good agreement. In the simulations of the corresponding time-resolved XAS spectra in the gas phase, where the laser field of the pump process has been explicitly taken into account to dynamically populate the $^1B_{2u}$ excited state, the spectrum obtained at the experimental delay time of 220 fs based on the MB sampling of classical nuclei at 300 K agrees well with that based on the Wigner sampling of the harmonic quantum ground state in order to generate the initial phase space conditions for the subsequent nonadiabatic surface hopping simulations. The consistency of the two approaches together with their agreement with the experimental TR-XAS spectrum of pyrazine suggests the use of classical MB sampling at finite temperature in those cases where appropriate Wigner sampling is difficult or impossible to achieve, for instance when flexible (rather than quasi-rigid) or solvated (rather than isolated) molecules are of interest. In the latter case, it is particularly important to sample the proper thermal solvent fluctuations, notably including noncovalent hydrogen bonds to the solute, at the relevant thermodynamic conditions of the solution, which is straightforwardly and accurately accessible in the framework of AIMD simulations in the appropriate ensemble.

For the UV−vis absorption spectrum in the gas phase, the results obtained with both MB and Wigner samplings show good agreement in the positions and the relative intensity of all three absorption bands with those in the experimental spectrum of pyrazine. This agreement suggests that spectral simulations based on ADC(2) in combination with MB sampling might be sufficiently reliable for the evaluation of solvent effects on excited states of such molecules in solution if described by using appropriate solvation models as follows.

When it comes to studying the impact of solvation on the UV−vis absorption spectrum of pyrazine in liquid water at ambient conditions, we first of all note that the two nitrogen sites of this apolar molecule are strongly interacting with water molecules by accepting mostly one hydrogen bond each in strongly fluctuating local arrangements as generated by the AIMD simulation of the aqueous solution. Thus, we sampled the configurations of this first solvation shell around the nitrogens (first) as well as their combined second solvation shell (second) that contains on average 24 water molecules. This provides us with two sets of microsolvation complexes for the spectroscopic calculations that have been rigorously sampled from the thermal MB distribution which describes the aqueous phase at ambient conditions, notably including the hydrogen bond fluctuations in addition to the ro-vibrational motion of the solute species and its coupling to the solvent molecule.

With respect to the solvent effects on the excited states of pyrazine, the present results indicate that such effects converge only slowly with respect to the size of the explicit solvent shell even for the lowest bands, unless such microsolvation models are embedded in an implicit (continuum) solvent model in an effort to avoid surface artifacts. For the models without implicit solvent, absorption bands break down at about 6 eV (≈1 eV

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

below the first absorption band of the solvent) due to the appearance of artificial Rydberg-like states on the outermost boundary of the model. Due to its nature, this problem will persist irrespective of how many solvent shells are included as explicit water molecules in the QM system as long as it is not embedded in an additional model potential which quenches these unphysical states by avoiding the interface with vacuum characterized by its low dielectric constant. If embedded in a continuum solvation model, here COSMO, that creates instead an interface with a structureless medium characterized by the high dielectric constant of water, the excitation energies of the valence-excited states and the UV−vis absorption spectrum converge quickly with the size of the explicit solvation shell. For the lowest-lying states, only the first shell shifts the absorption bands by several hundreds of meV, whereas the second explicit solvation shell is needed to quench Rydberg excitations around 6 eV.

We thus conclude that for pyrazine in water the valence-excited states of practical interest, which lie below the lowest excited state of a solvent molecule, can be described with reasonable accuracy by microsolvating pyrazine using its first shell water molecules and embedding these complexes in their turn in COSMO continuum water. Among these excited states, the lowest ones, which do not have Rydberg character, are affected by only solvent molecules in the first and second solvation shells, and those in the second solvation shell can already be replaced with COSMO. Therefore, static UV−vis absorption spectra of solvated such chromophores in water can be faithfully analyzed with a combination of MB sampling of the thermal fluctuations of the solution and subsequent computation of the desired spectrum for each sampled configurational snapshot using the pyrazine+first shell or pyrazine+second shell model in conjunction with COSMO, in most of the practical cases.

Beyond the specific case, the insights obtained in the present study are also major steps toward accurate simulations of the photoinduced nonadiabatic dynamics of chromophores in the aqueous solution in future work. Moreover, they help one to understand the limitations of both microsolvation and implicit solvent models when studying electronic excitations of chromophores in aqueous solutions and at the same time delineate pragmatic approaches to overcome such failures by combining these models in the sense of hybrid "explicit microsolvation plus implicit continuum solvation" approaches.

AUTHOR INFORMATION

Corresponding Author

Christof Hättig − *Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany;* orcid.org/0000-0002-5752-2710; Phone: +49 (0)234 32 28082; Email: Christof.Haettig@ruhr-uni-bochum.de

Authors

Shota Tsuru − *Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany;* orcid.org/0000-0003-0014-7873

Bikramjit Sharma − *Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany;* orcid.org/0000-0002-8844-1796

Dominik Marx − *Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c01129

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was funded by the Alexander von Humboldt Foundation through a Research Fellowship to S.T. (JPN 1201668 HFST-P) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy − EXC 2033−390677874 − RESOLV. The calculations described in Secs. 2.2 and 2.3.1 have benefited from the DTU Computing Center resources as in ref 70, whereas those described in Secs. 2.3.2 and 2.3.3 have been performed using local HPC resources at RUB. All AIMD simulations have been carried out at RUB using HPC@ZEMOS, HPC-RESOLV, and BOVILAB@RUB.

REFERENCES

(1) Garavelli, M.; Smith, B. R.; Bearpark, M. J.; Bernardi, F.; Olivucci, M.; Robb, M. A. Relaxation Paths and Dynamics of Photoexcited Polyene Chains: Evidence for Creation and Annihilation of Neutral Soliton Pairs. *J. Am. Chem. Soc.* **2000**, *122*, 5568−5581.

(2) Komainda, A.; Ostojić, B.; Köppel, H. Ab Initio Quantum Study of Nonadiabatic S_1−S_2 Photodynamics of *s-trans*-Butadiene. *J. Phys. Chem. A* **2013**, *117*, 8782−8793.

(3) Komainda, A.; Lyskov, I.; Marian, C. M.; Köppel, H. Ab Initio Benchmark Study of Nonadiabatic S_1−S_2 Photodynamics of *cis*- and *trans*-Hexatriene. *J. Phys. Chem. A* **2016**, *120*, 6541−6556.

(4) Glover, W. J.; Mori, T.; Schuurman, M. S.; Boguslavskiy, A. E.; Schalk, O.; Stolow, A.; Martínez, T. J. Excited state non-adiabatic dynamics of the smallest polyene, *trans*1,3-butadiene. II. *Ab initio* multiple spawning simulations. *J. Chem. Phys.* **2018**, *148*, 164303.

(5) Tully, J. C. Molecular dynamics with electronic transitions. *J. Chem. Phys.* **1990**, *93*, 1061−1071.

(6) Liu, J. Recent advances in the linearized semiclassical initial value representation/classical Wigner model for the thermal correlation function. *Int. J. Quantum Chem.* **2015**, *115*, 657−670.

(7) Langer, H.; Doltsinis, N. L.; Marx, D. Excited-State Dynamics and Coupled Proton−Electron Transfer of Guanine. *ChemPhysChem* **2005**, *6*, 1734−1737.

(8) Doltsinis, N. L.; Marx, D. Nonadiabatic Car-Parrinello Molecular Dynamics. *Phys. Rev. Lett.* **2002**, *88*, 166402.

(9) Doltsinis, N. L.; Marx, D. First Principles Molecular Dynamics Involving Excited States and Nonadiabatic Transitions. *J. Theor. Comput. Chem.* **2002**, *01*, 319−349.

(10) Böckmann, M.; Doltsinis, N. L.; Marx, D. Azobenzene photoswitches in bulk materials. *Phys. Rev. E* **2008**, *78*, 036101.

(11) Böckmann, M.; Doltsinis, N. L.; Marx, D. Nonadiabatic Hybrid Quantum and Molecular Mechanic Simulations of Azobenzene Photoswitching in Bulk Liquid Environment. *J. Phys. Chem. A* **2010**, *114*, 745−754.

(12) Avagliano, D.; Bonfanti, M.; Garavelli, M.; González, L. QM/MM Nonadiabatic Dynamics: the SHARC/COBRAMM Approach. *J. Chem. Theory Comput.* **2021**, *17*, 4639−4647.

(13) Cofer-Shabica, D. V.; Menger, M. F. S. J.; Ou, Q.; Shao, Y.; Subotnik, J. E.; Faraji, S. INAQS, a Generic Interface for Nonadiabatic QM/MM Dynamics: Design, Implementation, and Validation for GROMACS/Q-CHEM simulations. *J. Chem. Theory Comput.* **2022**, *18*, 4601−4614.

(14) Barbatti, M.; Sen, K. Effects of different initial condition samplings on photodynamics and spectrum of pyrrole. *Int. J. Quantum Chem.* **2016**, *116*, 762−771.

(15) Zobel, J. P.; Heindl, M.; Nogueira, J. J.; González, L. Vibrational Sampling and Solvent Effects on the Electronic Structure of the Absorption Spectrum of 2-Nitronaphthalene. *J. Chem. Theory Comput.* **2018**, *14*, 3205−3217.

(16) Yao, Y.; Hase, W. L.; Granucci, G.; Persico, M. Sampling initial positions and momenta for nuclear trajectories from quantum mechanical distributions. *J. Chem. Phys.* **2021**, *154*, 074115.
(17) Avagliano, D.; Lorini, E.; González, L. Sampling effects in quantum mechanical/molecular mechanics trajectory surface hopping non-adiabatic dynamics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2022**, *380*, 20200381.
(18) Loco, D.; Gelfand, N.; Jurinovich, S.; Protti, S.; Mezzetti, A.; Mennucci, B. Polarizable QM/Classical Approaches for the Modeling of Solvation Effects on UV−Vis and Fluorescence Spectra: An Integrated Strategy. *J. Phys. Chem. A* **2018**, *122*, 390−397.
(19) Zuehlsdorff, T. J.; Isborn, C. M. Modeling absorption spectra of molecules in solution. *Int. J. Quantum Chem.* **2019**, *119*, e25719.
(20) Chen, J.; Chan, B.; Shao, Y.; Ho, J. How accurate are approximate quantum chemical methods at modelling solute−solvent interactions in solvated clusters? *Phys. Chem. Chem. Phys.* **2020**, *22*, 3855−3866.
(21) Gómez, S.; Giovannini, T.; Cappelli, C. Absorption spectra of xanthines in aqueous solution: a computational study. *Phys. Chem. Chem. Phys.* **2020**, *22*, 5929−5941.
(22) Abou-Hatab, S.; Carnevale, V.; Matsika, S. Modeling solvation effects on absorption and fluorescence spectra of indole in aqueous solution. *J. Chem. Phys.* **2021**, *154*, 064104.
(23) Tsuru, S.; Sharma, B.; Nagasaka, M.; Hättig, C. Solvent Effects in the Ultraviolet and X-ray Absorption Spectra of Pyridazine in Aqueous Solution. *J. Phys. Chem. A* **2021**, *125*, 7198−7206.
(24) Manian, A.; Shaw, R. A.; Lyskov, I.; Russo, S. P. The quantum chemical solvation of indole: accounting for strong solute—solvent interactions using implicit/explicit models. *Phys. Chem. Chem. Phys.* **2022**, *24*, 3357−3369.
(25) Mennucci, B. Polarizable continuum model. *WIREs Comput. Mol. Sci.* **2012**, *2*, 386−404.
(26) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans.* **1993**, *2*, 799−805.
(27) Vall-llosera, G.; Gao, B.; Kivimäki, A.; Coreno, M.; Álvarez Ruiz, J.; de Simone, M.; Ågren, H.; Rachlew, E. The C 1*s* and N 1*s* near edge x-ray absorption fine structure spectra of five azabenzenes in the gas phase. *J. Chem. Phys.* **2008**, *128*, 044316.
(28) Scutelnic, V.; Tsuru, S.; Pápai, M.; Yang, Z.; Epshtein, M.; Xue, T.; Haugen, E.; Kobayashi, Y.; Krylov, A. I.; Møller, K. B.; et al. X-ray transient absorption reveals the $^1A_u(n\pi^*)$ state of pyrazine in electronic relaxation. *Nat. Commun.* **2021**, *12*, 5003.
(29) Stener, M.; Decleva, P.; Holland, D. M. P.; Shaw, D. A. A study of the valence shell electronic states of pyrimidine and pyrazine by photoabsorption spectroscopy and time-dependent density functional theory calculations. *J. Phys. B* **2011**, *44*, 075203.
(30) Samir, B.; Kalalian, C.; Roth, E.; Salghi, R.; Chakir, A. Gas-phase UV absorption spectra of pyrazine, pyrimidine and pyridazine. *Chem. Phys. Lett.* **2020**, *751*, 137469.
(31) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. CP2K: atomistic simulations of condensed matter systems. *WIREs Comput. Mol. Sci.* **2014**, *4*, 15−25.
(32) Kühne, T. D.; Iannuzzi, M.; Del Ben, M.; Rybkin, V. V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R. Z.; Schütt, O.; Schiffmann, F.; et al. CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **2020**, *152*, 194103.
(33) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. QUICKSTEP: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **2005**, *167*, 103−128.
(34) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé−Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635−2643.
(35) Hammer, B.; Hansen, L. B.; Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **1999**, *59*, 7413−7421.
(36) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(37) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703−1710.
(38) Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* **1998**, *58*, 3641−3662.
(39) Krack, M. Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals. *Theor. Chem. Acc.* **2005**, *114*, 145−152.
(40) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(41) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(42) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845−854.
(43) Abraham, M. J.; van der Spoel, D.; Lindahl, E.; Hess, B.; the G development team *GROMACS User Manual version 2019*. Available from http://www.gromacs.org (accessed 2023-01-25).
(44) Abascal, J. L. F.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.
(45) Vega, C.; Abascal, J. L. F. Simulating water with rigid non-polarizable models: a general perspective. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19663−19688.
(46) Sousa da Silva, A. W.; Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes* **2012**, *5*, 367.
(47) Richter, M.; Marquetand, P.; González-Vázquez, J.; Sola, I.; González, L. SHARC: ab Initio Molecular Dynamics with Surface Hopping in the Adiabatic Representation Including Arbitrary Couplings. *J. Chem. Theory Comput.* **2011**, *7*, 1253−1258.
(48) Mai, S.; Marquetand, P.; González, L. Nonadiabatic dynamics: The SHARC approach. *WIREs Computational Molecular Science* **2018**, *8*, e1370.
(49) Mai, S.; Richter, M.; Heindl, M.; Menger, M. F. S. J.; Atkins, A. J.; Ruckenbauer, M.; Plasser, F.; Ibele, L. M.; Kropf, S.; Oppel, M. et al. *SHARC2.1: Surface Hopping Including Arbitrary Couplings − Program Package for Non-Adiabatic Dynamics*. Available from sharc-md.org (accessed 2023-01-25).
(50) Neese, F. Software update: the ORCA program system version 4.0. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1327.
(51) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **2020**, *152*, 224108.
(52) Granucci, G.; Persico, M.; Zoccante, A. Including quantum decoherence in surface hopping. *J. Chem. Phys.* **2010**, *133*, 134111.
(53) Plasser, F.; Granucci, G.; Pittner, J.; Barbatti, M.; Persico, M.; Lischka, H. Surface hopping dynamics using a locally diabatic formalism: Charge transfer in the ethylene dimer cation and excited state dynamics in the 2-pyridone dimer. *J. Chem. Phys.* **2012**, *137*, 22A514.
(54) Besley, N. A.; Asmuruf, F. A. Time-dependent density functional theory calculations of the spectroscopy of core electrons. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024−12039.
(55) Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. Self-Consistent Field Calculations of Excited States Using the Maximum Overlap Method (MOM). *J. Phys. Chem. A* **2008**, *112*, 13164−13171.
(56) Epifanovsky, E.; Gilbert, A. T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.;

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

1108 Dempwolff, A. L.; et al. Software for the frontiers of quantum
1109 chemistry: An overview of developments in the Q-Chem 5 package. *J.*
1110 *Chem. Phys.* **2021**, *155*, 084801.
1111 (57) Dreuw, A.; Wormit, M. The algebraic diagrammatic
1112 construction scheme for the polarization propagator for the
1113 calculation of excited states. *WIREs Comput. Mol. Sci.* **2015**, *5*, 82−95.
1114 (58) TURBOMOLE V7.6 2021, a development of University of
1115 Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007,
1116 TURBOMOLE GmbH, since 2007. Available from https://www.
1117 turbomole.org (accessed 2023-01-25).
1118 (59) Balasubramani, S. G.; Chen, G. P.; Coriani, S.; Diedenhofen,
1119 M.; Frank, M. S.; Franzke, Y. J.; Furche, F.; Grotjahn, R.; Harding, M.
1120 E.; Hättig, C.; et al. TURBOMOLE: Modular program suite for *ab*
1121 *initio* quantum-chemical and condensed-matter simulations. *J. Chem.*
1122 *Phys.* **2020**, *152*, 184107.
1123 (60) Halverson, F.; Hirt, R. C. Near Ultraviolet Solution Spectra of
1124 the Diazines. *J. Chem. Phys.* **1951**, *19*, 711−718.
1125 (61) Peral, F.; Gallego, E. A study by ultraviolet spectroscopy on the
1126 self-association of diazines in aqueous solution. *Spectrochim. Acta Part*
1127 *A* **2003**, *59*, 1223−1237.
1128 (62) Karbalaei Khani, S.; Marefat Khah, A.; Hättig, C. COSMO-RI-
1129 ADC(2) excitation energies and excited state gradients. *Phys. Chem.*
1130 *Chem. Phys.* **2018**, *20*, 16354−16363.
1131 (63) Schwabe, T.; Sneskov, K.; Haugaard Olsen, J. M.; Kongsted, J.;
1132 Christiansen, O.; Hättig, C. PERI−CC2: A Polarizable Embedded RI-
1133 CC2Method. *J. Chem. Theory Comput.* **2012**, *8*, 3274−3283.
1134 (64) Luzanov, A. V.; Sukhorukov, A. A.; Umanskii, V. É. Application
1135 of transition density matrix for analysis of excited states. *Theor. Exp.*
1136 *Chem.* **1976**, *10*, 354−361.
1137 (65) Luzanov, A. V.; Zhikol, O. A. In *Practical Aspects of*
1138 *Computational Chemistry*; Leszczynski, J., Shukla, M. K., Eds.;
1139 Springer: Heidelberg, Germany, 2012; Vol. *I*, pp 415−449,
1140 DOI: 10.1007/978-94-007-0919-5_14.
1141 (66) Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **2003**,
1142 *118*, 4775−4777.
1143 (67) Mayer, I. Using singular value decomposition for a compact
1144 presentation and improved interpretation of the CIS wave functions.
1145 *Chem. Phys. Lett.* **2007**, *437*, 284−286.
1146 (68) Carroll, T. X.; Hahne, J.; Thomas, T. D.; Sæthre, L. J.; Berrah,
1147 N.; Bozek, J.; Kukk, E. Carbon 1*s* core-hole lifetime in CO_2. *Phys. Rev.*
1148 *A* **2000**, *61*, 042503.
1149 (69) Husimi, K. Some Formal Properties of the Density Matrix.
1150 *Phys. Math. Soc. Jpn. 3rd Series* 1940; Vol. *22*, pp 264−314,
1151 DOI: 10.11429/ppmsj1919.22.4_264.
1152 (70) DTU Computing Center, *DTU Computing Center resources*.
1153 https://doi.org/10.48714/DTU.HPC.0001 (accessed 2023-01-25),
1154 DOI: 10.48714/DTU.HPC.0001.

https://doi.org/10.1021/acs.jctc.2c01129
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX