First-principles Modeling of the Absorption spectrum of Crystal Violet in Solution - The Importance of Environmentally-driven Symmetry Breaking

Dayana Bashirova and Tim J. Zuehlsdorff[a)]
Department of Chemistry, Oregon State University, Corvallis, Oregon 97331, USA

(Dated: 1 May 2024)

Theoretical spectroscopy plays a crucial role in understanding the properties of materials and molecules. One of the most promising methods for computing optical spectra of chromophores embedded in complex environments from the first principles is the cumulant approach, where both (generally anharmonic) vibrational degrees of freedom and environmental interactions are explicitly accounted for. In this work, we verify the capabilities of the cumulant approach in describing the effect of complex environmental interactions on linear absorption spectra by studying Crystal Violet (CV) in different solvents. The experimental absorption spectrum of CV strongly depends on the nature of the solvent, indicating strong coupling to the condensed-phase environment. We demonstrate that these changes in absorption lineshape are driven by an increased splitting between absorption bands of two low-lying excited states that is caused by a breaking of the D_3 symmetry of the molecule, and that in polar solvents this symmetry-breaking is mainly driven by electrostatic interactions with the condensed phase environment, rather than distortion of the structure of the molecule, in contrast with conclusions reached in a number of previous studies. Our results reveal the importance explicitly including a counterion in the calculations in nonpolar solvent due to electrostatic interactions between CV and the ion. In polar solvent, these interactions are strongly reduced due to solvent screening effects, thus minimizing the symmetry breaking. Computed spectra in methanol are found to be in reasonable agreement with experiment, demonstrating the strengths of the outlined approach in modeling strong environmental interactions.

I. INTRODUCTION

Crystal Violet (CV) is a triarylmethane dye with compelling optical properties. The low-energy absorption lineshape of CV contains two peaks, the splitting between which depends on the solvent and temperature.[1] Korppi-Tommola and Yip observed that the overlap between absorption bands strongly depends on the solvent polarity, so that in nonpolar solvents there are two distinct peaks, whereas in polar solvents the bands overlap strongly.[2] The nature of the higher energy peak has been debated for decades and remains uncertain.[1–8]

There are two main theories describing this phenomenon: 1) the ground state of CV is inhomogeneous and this causes the appearance of the second peak in absorption lineshape; or 2) solute-solvent interactions induce the D_3 symmetry of CV to break and cause the lifting of the degeneracy of the excited state. For instance, Lewis *et al.* suggested the idea that in a solvent there is a thermal equilibrium between two isomers of CV[1] one of those is symmetrical helical type (D_3 symmetry) and the other one is distorted helical type (C_2 symmetry). Even though the existence of the particular C_2 isomer was not supported by the experiment, the D_3 isomerization was confirmed by femtosecond pump-probe measurement.[4] Maruyama *et al.* suggested that the structural changes between the two ground states cannot be large, so that CV is present as a planar D_3 configuration and a pyramidal C_3 configuration. The calculated absorption bands of these isomers in ethanol agreed well with the experiment.[4] Korppi-Tommola and Yip hold a different point of view, suggesting that the degeneracy of the S_1 state is lifted due to solvation in polar solvents and interaction with a counterion in nonpolar ones, the last interaction being stronger, which leads to greater separation between the absorption bands.[2] This theory was supported by McHale *et al.* using polarized resonance Raman scattering, fluorescence emission data, and semiempirical molecular orbitals (MO) calculations,[3] the latter were also used to indicate the energy barrier for torsional isomerization, which was found to exceed k_BT, thus is was shown that rotation of the phenyl groups is unlikely. However, Lovell *et al.* later stated, based on the polarization and analysis of the previous studies, that the two theories are not mutually exclusive,[5] and that the ground state of CV exists as an equilibrium of two isomers, one of which has a non-degenerate S_1 state. Additionally, it was shown that the absorption lineshape depends on the solvatochromic parameters of the solvent, again arguing against the inhomogeneity of the ground state.[6] The most recent studies either do not support both theories,[7] or suggest that the center of the absorption band of CV corresponds to the excitation of the symmetric molecule, and the edges correspond to the S_1 and S_2 transitions of the symmetry-broken molecules according to QM calculations and time-resolved polarized spectroscopy data.[8] Thus, the dominant factor in the appearance of a shoulder in absorption spectrum of CV remains unclear, and can be ascribed to either the existence of two isomers of the ground state, torsional

[a)]Electronic mail: tim.zuehlsdorff@oregonstate.edu

disorder of the molecule, environmental interactions, or a combination of these factors.

The ongoing debate with regards to the origin of peak splittings in the lineshape of the molecule, as well as the strong dependence of the splitting on solvent interactions, make CV an ideal target for a detailed study using theoretical/computational spectroscopy predictions. First-principles computational predictions of optical properties of solvated dyes can serve as an excellent tool to provide insight and interpretation of experimental data.[9–11] However, to obtain truly predictive simulated spectra in real condensed phase systems, it is often necessary to explicitly describe the interaction of the chromophore with its complex environment, which may demand large-scale electronic structure calculations.[12–18] Very accurate calculations of a complicated system are resource-intensive, thus a good trade-off between the accuracy and computational cost is essential.[16,19–22]

To accurately model absorption lineshapes of molecules embedded in complex environments, vibronic couplings, direct solute-solvent interactions, temperature effects and environmental polarization effects all have to be accounted for.[17,23–27] In recent years, significant efforts have been dedicated to developing robust and computationally affordable approaches to capture these effects from first principles.[26,28–30] The most widely-used method, the Franck-Condon (FC)[24,31–33] approach, captures vibronic effects by approximating the ground and excited state potential energy surfaces (PESs) of the chromophore as harmonic around their respective minima. In addition to the approximations inherent in this method, the FC method often neglects direct solvent interactions, since the calculations of ground- and excited state vibrational frequencies require optimized molecular geometries that are only straightforwardly accessible when representing the solvent environment through polarizable continuum models.[34,35] The ensemble method,[13,36–41] although capable of capturing anharmonicity and solvent effects due to a molecular dynamics (MD) sampling of solute-solvent configurations, treats nuclear degrees of freedom purely classically and thus cannot describe vibronic effects. A number of hybrid approaches combining features of both methods[42–45] have been introduced recently, but they often require an arbitrary timescale separation between slow solvent and fast chromophore motions that might not be applicable in systems like CV that undergo collective chromophore-environment motion.

The shortcomings of both the ensemble and the Franck-Condon method in describing vibronic effects and condensed phase interactions can be overcome by the cumulant approach.[28,46–50] The method constructs the linear absorption spectrum directly from equilibrium time-correlation functions of fluctuations of the electronic excitation energy, which can be conveniently sampled along ground-state molecular dynamics (MD) simulations of the system in its complex environment. While the method is computationally expensive, requiring the computation of tens of thousands of vertical excitation energies, it treats chromophore and environment fluctuations on the same footing, while the direct MD sampling approximately accounts for anharmonic effects.[27,45] Additionally, the computational cost of the approach can potentially be significantly reduced using machine-learning techniques,[51,52] making it a reliable and computationally competitive framework for modeling optical spectra in systems that strongly interact with their environments.[28,50,53]

In the present work, we investigate the effect of symmetry-breaking due to chromophore motion, solvent interactions and counterions on the lineshape of CV from first principles. We consider two solvents: toluene, a weakly interacting solvent; and methanol, a polar protic solvent expected to interact much stronger with the chromophore. The strong dependence of the experimentally observed spectral lineshape on solvent polarity indicates a significant coupling of CV to its condensed-phase environment, making it an ideal system to verify the accuracy of the cumulant method in this limit.

II. METHODS

A. Theoretical Background

For a two-level system consisting of an electronic ground- and excited state, the absorption lineshape $\sigma(\omega)$ can be expressed as[46,50]

$$\sigma(\omega) = \alpha(\omega) \int_{-\infty}^{\infty} \mathrm{d}t\, e^{\mathrm{i}\omega t} \chi(t) \tag{1}$$

where $\alpha(\omega)$ is a frequency-dependent prefactor[24] and $\chi(t)$ is the linear response function of the transition dipole moment given by:

$$\chi(t) = \mathrm{Tr}\left[\rho_g \hat{\boldsymbol{\mu}}_{ge}(t) \cdot \hat{\boldsymbol{\mu}}_{eg}(0)\right]. \tag{2}$$

Here, the subscripts g and e denote the electronic ground- and excited state respectively, ρ_g is the ground state density matrix of nuclear degrees of freedom, and the trace is carried out over all nuclear degrees of freedom.[46]

In practice, Eqn. 2 is impossible to evaluate for realistic condensed-phase systems, and suitable approximations have to be found. In the cumulant scheme,[46] the linear response function is defined as an infinite expansion with respect to cumulants of the energy gap fluctuation operator. Under the Condon approximation,[54] the electronic transition dipole moment is independent of nuclear degrees of freedom, and we can write:

$$\begin{aligned}\chi(t) &\approx |\boldsymbol{\mu}_{ge}|^2 e^{-\mathrm{i}\omega_{eg}^{\mathrm{av}} t} \mathrm{Tr}\left[\rho_g \exp_+\left[-\mathrm{i}\int_0^t \mathrm{d}\tau\, \delta U(\hat{\boldsymbol{q}}, \tau)\right]\right] \\ &= |\boldsymbol{\mu}_{ge}|^2 e^{-\mathrm{i}\omega_{eg}^{\mathrm{av}} t} \exp\left[-G_\infty(t)\right], \end{aligned} \tag{3}$$

where

$$G_m(t) = \sum_{n=2}^{m} g_n(t) \tag{4}$$

is the mth order approximation to the exact lineshape function $G_\infty(t)$ and $g_n(t)$ is the nth order cumulant contribution. The cumulants are explicit functionals of increasingly higher order quantum correlation functions of the energy gap fluctuation operator $\delta U(\hat{\boldsymbol{q}}, t)$ defined as

$$\begin{aligned}\delta U(\hat{\boldsymbol{q}}, t) &= U(\hat{\boldsymbol{q}}, t) - \omega_{eg}^{\mathrm{av}} \\ &= \hat{H}_e - \hat{H}_g - \omega_{eg}^{\mathrm{av}}, \qquad (5)\\ \omega_{eg}^{\mathrm{av}} &= \mathrm{Tr}\left[\rho_g\, U(\hat{\boldsymbol{q}})\right]. \qquad (6)\end{aligned}$$

Here, $\omega_{eg}^{\mathrm{av}}$ is the thermal average of the energy gap operator and $\hat{H}_g$ and $\hat{H}_e$ denote the nuclear Hamiltonians in the electronic ground- and excited state respectively. While Eqn. 3 is exact for a two-level system within the Condon approximation, no closed form expressions for $G_\infty(t)$ exist except in model systems with simplified potential energy surfaces.[50] Thus in practice, to get a suitable closed-form expression for realistic systems, the infinite order cumulant expansion is truncated at finite order. We denote the approximate response function due to an mth order cumulant expansion, which in turn can be represented by the mth order approximation of the line shape function, as

$$\chi_m(t) = |\boldsymbol{\mu}_{eg}|^2\, e^{-\mathrm{i}\omega_{eg}^{\mathrm{av}} t - G_m(t)}. \qquad (7)$$

Most commonly, the cumulant expansion is truncated at second order, such that $G_2(t) = g_2(t)$, which yields the exact lineshape within the Condon approximation for a two-level system where the energy gap fluctuations follow Gaussian statistics. For non-Gaussian energy gap fluctuations, which can be induced by anharmonic PESs and mismatches in ground and excited state PES curvatures, higher order cumulants can become important.[50,55,56] While it is possible to compute these higher order correction terms to the lineshape function in both model potentials and realistic systems,[50,55] for the purpose of this work we limit ourselves to the second order cumulant approximation only. The second order cumulant $g_2(t)$ has the closed-form expression

$$\begin{aligned}g_2(t) = \frac{1}{\pi}\int_0^\infty d\omega \frac{J(\omega)}{\omega^2}\Bigg[\coth\left(\frac{\beta\omega}{2}\right)\left[1-\cos(\omega t)\right] + \\ \mathrm{i}\left[\sin(\omega t) - \omega t\right]\Bigg],\end{aligned} \qquad (8)$$

with

$$J(\omega) = \mathrm{i}\theta(\omega)\int \mathrm{d}t\, e^{\mathrm{i}\omega t}\,\mathrm{Im}\left[C_{\delta U}.(t)\right], \qquad (9)$$

Here, $J(\omega)$ is the spectral density of system-bath coupling, $\theta(\omega)$ is the Heaviside function and $C_{\delta U}$ is the energy gap fluctuation quantum autocorrelation function given by:

$$C_{\delta U} = \mathrm{Tr}\left[\rho_g\, \delta U\left(\hat{\boldsymbol{q}}, t\right) \delta U\left(\hat{\boldsymbol{q}}, 0\right)\right]. \qquad (10)$$

In the condensed phase systems that form the focus of this work, an additional approximation has to be introduced, as the exact quantum autocorrelation function is generally inaccessible. A common strategy is based on using the purely classical autocorrelation function, which is generally much easier to calculate, to approximately reconstruct its quantum counterpart using quantum correction factors.[57–61] While the choice of correction factor is not unique, in this work we make use of the *harmonic* quantum correction factor throughout. The spectral density can then be written in terms of the classical autocorrelation function as:

$$J(\omega) \approx \theta(\omega)\frac{\beta\omega}{2}\int \mathrm{d}t\; e^{\mathrm{i}\omega t} C_{\delta U}^{\mathrm{cl}}(t). \qquad (11)$$

This quantum correction factor has been extensively utilized for computing linear spectra of chromophores in complex environments and has demonstrated good performance in a variety of systems.[28,39,62] Using Eqn. 11 the spectral density $J(\omega)$ can be constructed directly from *classical* excitation energy fluctuations sampled along a ground-state MD trajectory.

The cumulant approach outlined here relies on two major approximations: That the system of interest can be described as a two-level system or a combination of two-level systems that are non-interacting, and that the Condon approximation is valid. The CV molecule exhibits two bright low-energy transitions that contribute to the visible spectrum, from HOMO to LUMO (S_1) and from HOMO-1 to LUMO (S_2) (see Sec. III A). These two excitations are relatively close in energy, meaning that nonadiabatic interactions[53,63] between the two excited states that go beyond a two-level treatment of the system cannot necessarily be ignored. However, for CV, we find that the two excited states stay sufficiently well separated along the MD trajectory, with transition dipole moments associated with S_1 and as S_2 remaining largely orthogonal, thus indicating only weak nonadiabatic coupling between the two excited states. It is therefore assumed that the optical spectrum of CV can be computed by summing the contributions of independent S_1 and S_2 spectra calculated using the cumulant approach.

Additionally, the Condon approximation neglects the effects of nuclear motion on the electronic transition dipole moment between the electronic ground- and excited state. Non-Condon contributions[64] to the optical spectra, such as Herzberg-Teller effects,[23,24] are often considered to be negligible for bright transitions with a large average transition dipole moment. For CV, the mean S_1 transition dipole moment along the QM/MM trajectories 3.66 D (with a standard deviation of 0.18 D) in toluene and 3.63 D (with a standard deviation of 0.18 D) in methanol. Similarly, the mean S_2 transition dipole moment is 3.00 D (with a standard deviation of 0.24 D) in toluene and 3.16 D (with a standard deviation of 0.29 D) in methanol. The relatively small standard deviation and large mean value of both transition dipole moments indicate that the Condon approximation likely provides an accurate description of the lineshape of CV.

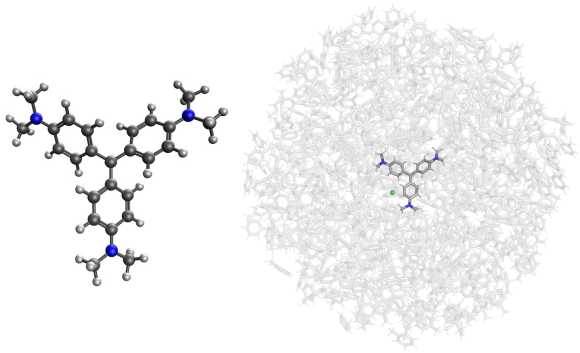

FIG. 1. The optimized geometry of CV displaying D_3 symmetry, and the CV molecule solvated in a 32 Å sphere of toluene. The counterion is depicted as green.

B. Computational Details

1. Static calculations

To assess the importance of symmetry breaking in understanding the optical properties CV, Kohn-Sham molecular orbitals, excited state energies and transition densities were computed for selected static configurations of the molecule. Optimized geometries and molecular orbitals were computed for the chromophore in vacuum using Gaussian,[65] excited state energies and transition densities were computed using TeraChem.[66,67] The ωb97x[68] functional with the 6-31+G*[69] basis set were used throughout for results presented in the main text, and detailed analysis of the dependence of low-lying excitation energies on DFT functional choice can be found in SI Sec. 5. Since CV is a cationic organic dye, calculations were performed both on the isolated chromophore and on the chromophore in the presence of a chloride counterion. In each case, the ground state geometry was optimized using density-functional theory (DFT) and the excited state energies were computed using full time-dependent DFT (TDDFT) in the linear-response formalism.[70] Transition densities for S_1 and S_2 were computed using TDDFT within the Tamm-Dancoff approximation.[71,72]

2. Molecular Dynamics

To study the effects of solvent interactions and chromophore vibrations on the absorption spectrum of CV, mixed quantum mechanical/molecular mechanical (QM/MM) dynamics[73] were carried out using the TeraChem[66,67] (QM) and AMBER[74] (MM) software interface.[75] Two different condensed phase environments were considered: The nonpolar solvent toluene and the polar protic solvent methanol. This choice was motivated by the significant differences in the experimentally reported lineshapes of CV in these solvents (see Ref. 2 and Fig. 3). Force-field parameters for CV and toluene were generated using the generalized AMBER force field (GAFF)[76] and antechamber.[77] For methanol, a standard non-polarizable AMBER FF was used. An MD timestep of 0.5 fs was used throughout, both for initial FF equilibrations and QM/MM dynamics.

For CV in methanol, the chromophore was solvated in a pre-equilibrated solvent sphere of 32 Å, containing 1887 methanol molecules. The system was heated to 300 K using a Langevin thermostat with a collision frequency of 1 ps^{-1}, followed by a 100 ps equilibration under open boundary conditions. A 26 ps QM/MM trajectory was then generated, with the temperature held at 300 K using the Langevin thermostat. CV was treated at the QM level using the ωb97x[68] functional with the 6-31+G* basis set.[69] The first 2 ps of the trajectory were discarded to allow for additional equilibration after switching the chromophore from the MM to the QM Hamiltonian. This procedure was repeated three times to obtain three independent trajectories. The 100 ps equilibration was carried out each time with different random number seeds, such that the QM/MM dynamics was started with different initial configurations of the system.

Since CV is a cationic organic dye, the influence of a chloride counterion on the system dynamics and computed spectra was also considered. For this purpose, CV together with a Cl^- ion was solvated in a sphere, containing 1887 methanol molecules, with the counterion initialized as a contact ion pair with CV. A static FF for CV likely does not yield a correct description of the electrostatic interactions with the Cl^- ion, as DFT calculations in vacuum reveal a significant redistribution of electronic charge on the N atoms to stabilize a contact ion pairing (see SI Sec. 7). Thus, the MM equilibration step was skipped and a pure 77 ps QM/MM simulation was performed for CV with the ion in methanol, with only the CV molecule treated at the QM level. The first 41 ps were discarded to allow for an equilibration of the chromophore-ion configuration (see SI Sec. 4), generating 36 ps of usable trajectory. Since interactions between CV and the Cl^- were expected to be weak due to the polar nature of methanol solvent environment, only a single trajectory was generated explicitly including the counterion.

A cubic toluene box with 700 toluene molecules and a side length of 40 Å was generated using packmol.[78] The energy of the box was minimized, it was heated from 0 to 300 K using Langevin dynamics for 300 ps, and then equilibrated at constant temperature and pressure (NPT ensemble) for 400 ps. Since toluene is nonpolar, interactions between CV and the Cl^- counterion were expected to be stronger than in methanol. For this reason, CV and Cl^- in a contact ion pair configuration were solvated in a 32 Å sphere, containing 588 toluene molecules (see Fig. 1), with the toluene positions taken from the equilibrated box configuration. A 48.5 ps QM/MM trajectory was generated, using the same settings as for CV

in methanol. The first 12.5 ps were discarded for equilibration of the ion position, yielding 36 ps of usable dynamics (see SI Sec. 4). To probe the effect of the Cl^- on the computed dynamics, a single 24 ps trajectory for CV in toluene without a counterion was also produced, following the same computational protocol as outlined for CV in methanol. First, the chromophore was solvated in a toluene sphere of 32 Å, which was pre-heated to 300 K using Langevin thermostat for 300 ps and pre-equilibrated for 400 ps NPT ensemble. Then the system was equilibrated in the MM regime under open boundary conditions, followed by 26 ps of QM/MM dynamics, where the first 2 ps of the trajectory was discarded to allow for equilibration after switching to the QM/MM Hamiltonian.

To understand the effect of the condensed phase environment on the computed spectra, we also generated pure ab-initio MD (AIMD) trajectories of CV with and without the Cl^- counterion in vacuum using TeraChem. Both trajectories were carried out at 300 K using a Langevin thermostat and were 10 ps in length, with the first 2 ps being discarded for equilibration. For CV with the Cl^- counterion, the initial structure was taken from the optimized geometry in vacuum, and both Cl^- and CV were treated fully quantum mechanically during the dynamics.

3. Vertical Excitation Energies

Vertical excitation energies were calculated along the generated QM/MM MD and AIMD trajectories using full linear-response TDDFT as implemented in the TeraChem code.[66,67] The trajectories were sampled every 2 fs, yielding 12,000 snapshots per trajectory for each of the 24 ps MD trajectories generated for CV in methanol and toluene respectively, 18,000 snapshots per MD trajectory generated for CV in the presence of a counterion in the two solvents, and 4,000 snapshots per each MD trajectory of CV and CV with the Cl^- in vacuum. In all solvated calculations, only CV was treated at the QM level and the solvent environment and the counterion were included in the calculation as classical fixed point charges taken directly from the AMBER force fields. For the vacuum trajectory including the chloride ion, CV was treated at the QM level and the counterion was included as a classical point charge when computing vertical excitation energies.

All calculations were performed using the ωb97x density functional and the 6-31+G* basis set to be fully consistent with the QM/MM dynamics. This was done to avoid well-known issues in the computation of spectral densities that arise from a mismatch between the Hamiltonian used to propagate the system and the Hamiltonian used to compute vertical excitations.[79] The effect of DFT functional choice in both QM/MM dynamics and vertical excitation energies on resulting optical spectra was analyzed in detail, with results reported in SI Sec. 5.

	Splitting	Exp. splitting	$\Delta E_{S_1/S_2}$	E_{S_1}		E_{S_2}	
GS opt CV	-	-	0.002	2.998	[0.890]	3.000	[0.885]
GS opt CV with Cl	-	-	0.521	2.863	[1.014]	3.384	[0.730]
CV in vacuum (8 ps MD)	0.070	-	0.096	2.897	[0.894]	2.993	[0.779]
CV with Cl in vacuum (8 ps MD)	0.559	-	0.559	2.820	[0.883]	3.379	[0.725]
CV in methanol (24 ps MD)	0.135	0.126	0.177	2.870	[0.907]	3.047	[0.786]
CV with Cl in methanol (36 ps MD)	0.141	0.126	0.189	2.854	[0.912]	3.043	[0.771]
CV in toluene (24 ps MD)	0.119	0.256	0.145	2.858	[0.930]	3.003	[0.780]
CV with Cl in toluene (36 ps MD)	0.475	0.256	0.512	2.797	[0.920]	3.309	[0.734]

TABLE I. Computed excitation energies (eV) and oscillator strengths (unitless, in brackets) for CV in different environments. *Splitting* denotes the energy difference (eV) between the two maxima in the computed absorption spectrum, whereas $\Delta E_{S_1/S_2}$ denotes the difference between the TDDFT excitation energies E_{S_1} and E_{S_2}. For all MD-based calculations, E_{S_1} and E_{S_2} are computed from the average excitation energies along the MD trajectory. The experimental splitting[1,2] is computed directly from data presented in the references.

4. Calculation of spectra

All spectra presented in this work were calculated using the cumulant approach as implemented in our open-source spectroscopy Python package.[80] The vertical excitation energies computed along the MD trajectories were used to construct classical autocorrelation functions, that, following Eqns. 11 and 8, yield the linear response function and thus the absorption spectrum for a given electronic transition. Nonadiabatic effects[53] were assumed to be negligible (see Sec. II A), such that the full absorption spectrum was computed as the sum of the S_1 and the S_2 spectrum respectively.

All spectra reported in this work were normalized to have an equal maximum intensity with the experimental spectra. Furthermore, spectra were shifted such that the position of the maximum intensity peak in the computed spectrum agrees with the experimental lineshape.

III. RESULTS AND DISCUSSION

A. Crystal Violet: symmetry and molecular orbital degeneracy

The ground state optimized structure of CV possesses D_3 symmetry, as shown in Fig. 1. This symmetry causes a degeneracy in the two lowest lying excitations, both of which are dipole-allowed. Static (TD)DFT calculations of the optimized chromophore in vacuum reveal that this degeneracy is caused by the degeneracy of two specific canonical molecular orbitals (MOs), the HOMO and the HOMO-1 orbitals, with S_1 being dominated by a HOMO to LUMO (S_1) and S_2 a HOMO-1 to LUMO (S_2) transition (see Fig. 2). When the HOMO and HOMO-1 orbitals are degenerate, the S_1 and S_2 excitations are found to be equal in energy, respectively, with an oscillator strength of 0.89. Thus without the presence of symmetry-breaking chromophore motion or environmental interactions lifting the degeneracy of the low-lying excited states, one would expect the optical absorption

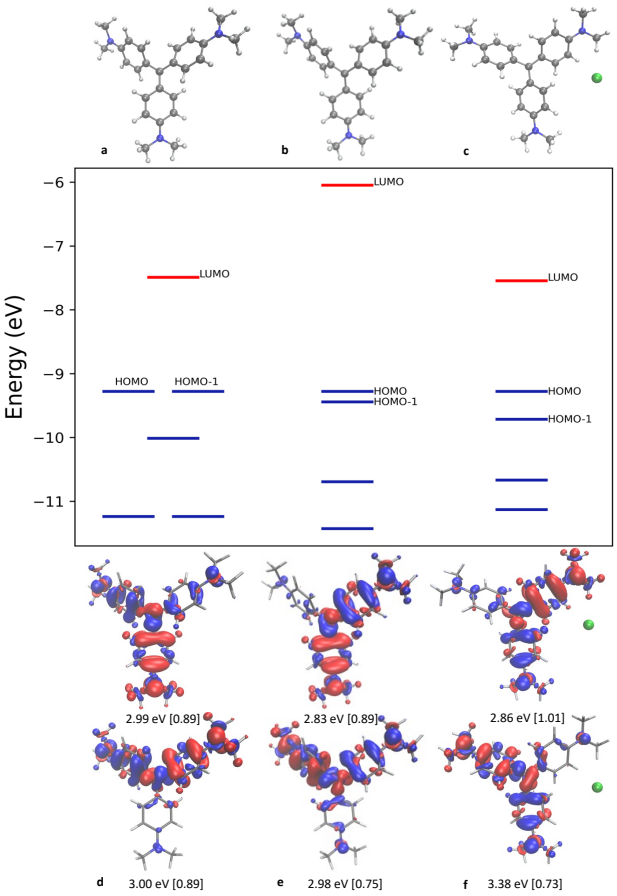

FIG. 2. The relationship between the symmetry and degeneracy of MOs in CV. *(a)* optimized geometry of CV and corresponding MO energy diagram. CV has D_3 symmetry, and the HOMO and HOMO-1 MOs are degenerate; *(b)* Configuration taken from an MD trajectory of CV in methanol; *(c)* optimized geometry of CV in vacuum in the presence of a chloride ion. In structures *(b)* and *(c)*, CV no longer has D_3 symmetry, as a result of which the MOs are no longer degenerate. All MOs were shifted for the structures *(b)* and *(c)* (by 0.74 eV and -2.67 eV respectively) in order to align all HOMO energies. The LUMO was shifted by -3 eV for better visualization for all structures. *(d)*, *(e)*, and *(f)* represent the S_1 and S_2 transition densities for the structures in *(a)*, *(b)*, and *(c)* respectively. The corresponding excitation energies and oscillator strengths are displayed beneath the structures. An isosurface value of 2.5×10^{-7} $e^-/Å^3$ was used for all structures.

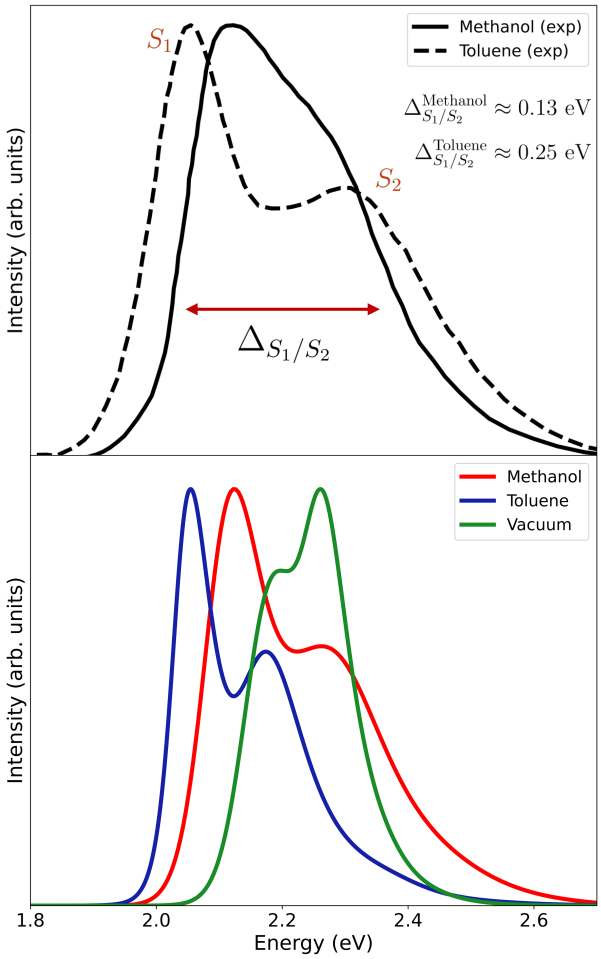

FIG. 3. Top part: The experimental spectra of CV taken from Ref. 2. The solid black line is the experimental spectrum of CV in methanol, the dashed black line in toluene. Bottom part: Calculated spectra for CV within the cumulant approach. The solid red line corresponds to the calculated spectrum of CV in methanol based on a 24 ps of QM/MM trajectory, blue is the spectrum in toluene based on a 24 ps QM/MM trajectory, and green is the spectrum in vacuum based on an 8 ps of trajectory. All calculations are performed without the presence of a counterion. The methanol, toluene, and vacuum spectra are shifted by -0.724 eV, -0.778 eV, and -0.700 eV, respectively, to ease comparison with the experimental spectra.

spectrum to be dominated by a single spectral feature.

In contrast, the experimental absorption spectrum for CV (see Ref. 2 and Fig. 3) consists of two separate peaks, and the splitting between them is found to be highly dependent on the nature of the solvent. For instance, the splitting in water, methanol, acetone, chloroform, benzene, and toluene is 0.145 eV, 0.126 eV, 0.113 eV, 0.102 eV, 0.255 eV, and 0.253 eV, respectively.[2] The results indicate that CV undergoes solvent-dependent symmetry breaking, which can be caused by a number of different factors, such as the torsional rotation of the three aryl rings and specific interactions with the condensed phase environment. To capture the explicit coupling between the low-frequency torsion of the rings and the solvent environment, and assess their effect on the electronic structure of CV, it is necessary to go beyond a static description of the molecule based on a single optimized geometry.

Ground state MD simulations reveal that a breaking

of the D_3 symmetry due to torsion of the aryl rings occurs both in vacuum and in solvent environments. If the dihedral angles defining the orientation of the rings (see Table S1) are not equal, the symmetry of the molecule is broken. As shown in Fig. 2, the degeneracy of the HOMO and HOMO-1 orbitals is preserved in the optimized geometry with equal dihedral angles *(a)*, but is lifted in the geometry taken from an MD snapshot in methanol *(b)*, with a corresponding lifting of the degeneracy of the S_1 and S_2 excitation energies. Furthermore, it is found that the amount of torsional motion of the aryl rings varies in different environments. The MD trajectory in vacuum shows almost identical average values of the three dihedral angles to those in the optimized vacuum structure, but with a large standard deviation indicating relatively unrestricted torsional motion of the rings (see Table S1). This dihedral motion causes a 0.08 eV energy splitting in the broad S_1 and S_2 absorption bands computed along the vacuum trajectory using the cumulant approach (see SI Sec. 1). The large amount of spectral broadening observed in vacuum is a result of the unconstrained rotation of the aryl rings, resulting in the combined spectrum due to both S_1 and S_2 excitations only exhibiting a very minor peak splitting of 0.07 eV (see Fig. 3).

The average values of the aryl ring dihedral angles in solution, computed along the MD trajectories in methanol and toluene, show a larger deviation from the values of the optimized ground state geometry when compared with the vacuum trajectory, indicating that some amount of torsion and thus a broken symmetry of the molecule is permanently stabilized by the solvent environment. Specifically, the recorded mean dihedral angles for two aryl rings in toluene are found to differ by almost 6°, indicating a sustained symmetry-breaking of the chromophore structure along the MD trajectory (see Table S1). This observation may be related to the bulkiness of the toluene molecule and π-stacking interactions with the aryl rings, which prevents free rotation observed in vacuum. Additionally, the standard deviations of the dihedral angles are found to be higher in vacuum than in toluene, further suggesting that the rotation of the aryl rings in toluene is significantly constrained by the condensed phase environment. In contrast, the inclusion of a Cl^- counterion in the MD dynamics for both methanol and toluene condensed phase environments was found to have only a negligible effect on the computed mean and standard deviation of the dihedral angles of the aryl rings (see SI Sec. 1), as compared with trajectories generated by ignoring the presence of the counterion. However, from Fig. 2 *(c)* it is clear that a close association of CV with Cl^-, as is observed in the optimized structure in vacuum, leads to a strong breaking of the molecular orbital degeneracy, with a corresponding S_1-S_2 splitting that is significantly larger than for an MD snapshot taken from the trajectory in methanol (Fig. 2 *(b)*).

The symmetry breaking of CV and the resulting splitting of the S_1 and S_2 absorption band found in experiment thus arises from a combination of two main factors: the rotation of the aryl rings and the interaction with the condensed-phase environment (such as solvent and the Cl^- counterion). To probe the effect of the solvent environment on the spectral lineshape, we compute cumulant spectra for MD trajectories for CV in methanol and toluene in the absence of a counterion. The results are displayed in Fig. 3, in comparison with the experimental CV spectra in toluene and methanol taken from Ref. 1 and 2, respectively. The calculated spectrum in methanol is found to be good agreement with the experimental spectrum in terms of the lineshape and peak positions, showing an S_1-S_2 splitting that is larger than in vacuum (see Table I). The origins of this good agreement will be discussed in detail in Sec. III B. However, the same method, when applied to calculate the absorption spectrum in toluene, yields a poor agreement with experiment, as the splitting between the peaks is is severely underestimated compared to what is observed experimentally. Indeed, the observed S_1-S_2 splitting is only slightly larger than that observed for CV in vacuum and smaller than in methanol, in contrast to the experimentally observed trend of increased splitting in nonpolar solvent. This apparent failure of the MD-based cumulant approach to qualitatively describe the absorption lineshape of CV in toluene will be investigated in section III C.

B. The absorption spectrum of CV in methanol

Methanol, being a polar solvent, can interact with CV via electrostatic interactions. Due to the solvent polarity, it is assumed that CV and its counterion exist in a fully solvated form in methanol,[2,5] and that solvent molecules prevent the direct interaction of CV and its counterion due to screening effects.[81] To investigate the influence of the Cl^- on the lineshape, a 36 ps QM/MM trajectory of CV together with a chloride ion in methanol was collected, on the basis of which an absorption spectrum was computed using the cumulant approach calculated (see Fig. 4 for the resulting spectrum). Analyzing the counterion position with respect to CV during the simulation (see Fig. S5 (*c*), (*d*) and Fig. S14), we find that, after an initial equilibration time, the ion is mostly screened by the solvent environment, minimizing electrostatic interactions with CV.

We also examine radial distribution functions (RDFs) between the three nitrogen atoms of CV and the and oxygen atoms of the surrounding methanol molecules, as well as the counterion along the QM/MM trajectory in methanol (see Fig. S7). Along the 24 ps trajectory of CV with the counterion in methanol, the oxygen atoms are located closer to the nitrogen atoms of the CV molecule compared to the N-Cl^- distance (see Fig. S7 *(b)* and *(c)*), which indicates a mostly screened interaction between CV with the counterion. There is also a noticeable imbalance between the number of O atoms of nearest neighbor methanol molecules with respect to different N atoms

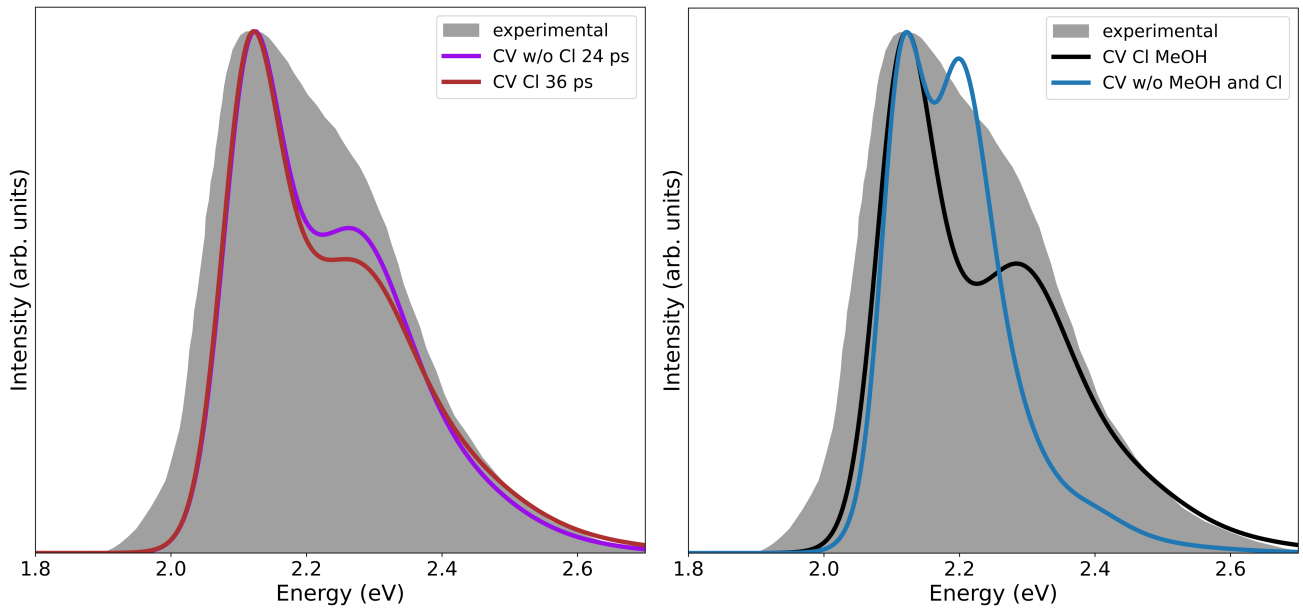

FIG. 4. Computed absorption spectra of CV in Methanol. Left: The purple line corresponds to the average spectrum of CV without a counterion calculated based on three individual 24 ps (12,000 snapshots) QM/MM trajectories. The red line corresponds to the cumulant spectrum of CV in the presence of a Cl^- counterion based on a single 36 ps (18,000 snapshots) QM/MM trajectory. The lines are shifted to match the experimental data by -0.724 eV and -0.708 eV, respectively. Right: The black line shows the cumulant spectrum based on an 8 ps QM/MM trajectory where the solvent and counterion are included as classical point charges in the TDDFT calculation. The blue line is same solvated trajectory but with the solvent and the counterion stripped away when computing vertical excitation energies with TDDFT. The lines are shifted to match the experimental data by -0.717 eV and -0.746 eV, respectively. The experimental spectrum[2] is shown in gray.

of CV. One of the nitrogen atoms is found to have on average more neighboring oxygen atoms along the entire MD trajectory, meaning that the electrostatic interaction stabilizes more of a positive charge located around that particular N atom. An analogous solvation of the CV molecule occurs for the trajectory without the counterion (Fig. S7 *(a)*), with a symmetry-broken local solvent environment being stabilized over the entire trajectory length of 24 ps. The RDFs reveal a relatively long-lived symmetry-broken configuration for CV in methanol, that is caused by an imbalance in local solvent environment around the three aryl rings.

Comparing the spectral lineshape based on trajectories with and without the Cl^- explicitly included in the dynamics, we observe only minor differences (see Fig. 4). Both spectra are in reasonable agreement with the experimental lineshape. Since both approaches, with and without the explicit inclusion of a Cl^- in the dynamics, produce almost identical lineshapes, all further analysis of the optical spectrum of CV in methanol is based on the system in the absence of a counterion.

To obtain a converged calculated spectrum of CV in methanol, three independent 24 ps MD simulations were performed without the presence of a counterion. The computed absorption spectra based on these individual trajectories produce very similar lineshapes (Fig. S13), indicating that slow collective solute-solvent motions are sufficiently sampled even by a small number of trajectories. The average spectrum obtained from these three trajectories correlates well with the experimental data (Fig. 4), with the exception of a slight underestimation of the second peak intensity and the overall width of the spectrum. Analysis performed in SI Sec. 5 suggests that the discrepancy in peak heights can be partially ascribed to DFT functional choice, whereas the overall lack in broadening is likely due to neglecting dynamic polarization effects in the spectral density due to representing the solvent environment as fixed classical point charges, as shown by some of our previous work.[27] The results show that the explicit inclusion of the chloride ion in the calculation is not critical for modeling the absorption lineshape of CV in methanol. The splitting between the S_1 and S_2 peaks in the calculated absorption spectrum with and without the counterion is 0.137 eV and 0.141 eV, respectively, compared to the the experimental value of 0.126 eV. The shape of the spectral line indicates that the degeneracy of the HOMO and HOMO-1 orbitals is broken by explicit interactions between CV and its polar solvent environment.

To analyze the origin of the obtained S_1-S_2 splitting of CV in methanol, we aim to decompose it into two distinct contributions: 1) from torsional rotation of the

aryl rings stabilized by the solvent environment; and 2) from non-covalent interactions between CV and its solvent environment. To quantify the amount of splitting that can be directly ascribed to solute-solvent interactions, we compute cumulant spectra along an 8 ps part of a single trajectory of CV in methanol in the presence of a counterion, but where the electrostatic environment is stripped away when computing vertical excitation energies with TDDFT. The results can be found in the right panel of Fig. 4. The stripped solvent calculations show only a very minor splitting between S_1 and S_2 peaks of about 80 meV. This energy splitting for CV with its solvent environment stripped away is even smaller than the splitting observed for the trajectory performed in vacuum (see Fig. 3 and Table I), indicating that the solvent environment restricts the motion of CV and thus diminishes the dihedral contribution to the symmetry breaking. Additionally, we notice a significant dampening in the peak intensities for the high-frequency modes of the spectral density when the solvent is stripped from the TDDFT calculations (Fig. S11 *(a)*, *(b)*), due to a strong coupling between localized chromophore modes coupling to the energy gap fluctuations and collective electrostatic effects due to the polar solvent environment.

The results presented in Fig. 4 demonstrate that the S_1-S_2 peak splitting of the absorption spectrum of CV in methanol is largely driven by symmetry breaking induced through electrostatic interactions with the polar solvent environment around the molecule, rather than symmetry breaking due to solvent-induced slow chromophore motion driven by the twisting of the aryl rings, in contrast with the results presented in a recent study.[8]

C. The absorption spectrum of CV in toluene

In toluene, CV interacts with solvent molecules through non-covalent interaction such as π-stacking. The torsion angle of the aryl rings of CV in toluene has a smaller standard deviation compared to vacuum (Table S1), suggesting that bulky solvent molecules hinders rotation. When constructing the cumulant spectrum, we find that the predicted spectral lineshape for CV in pure toluene solvent has an incorrect splitting between the peaks with a value of 0.119 eV compared to the experimental value 0.256 eV considering only the above contributions (see Fig. 3 and Table I). Being a nonpolar solvent, toluene cannot screen the interactions between the cationic CV and its Cl^- counterion, in contrast to the polar solvent methanol. This interaction, especially a close association between CV and the ion through the formation of a contact ion pair, can break the D_3 symmetry of CV. According to the ground state optimized structure of CV in the presence of a counterion in vacuum, the distance between the closest nitrogen atom to the ion and the chloride ion is 4.4 Å and this configuration yields a substantial energy difference between S_1 and S_2 transition of 0.52 eV (see Fig. 2). To investigate the influence of thermal fluctuations on the position of the counterion, we perform a 8 ps AIMD trajectory of CV with the Cl^- ion in vacuum. We find that the average distance between the closest nitrogen and the counterion is 4 Å along the trajectory, even less than for the optimized geometry in vacuum. The dynamic calculations again reveal a strong direct electrostatic interaction between CV and the chloride ion, leading to a large S_1 and S_2 splitting in the computed absorption lineshape in the cumulant approach, with the value of 0.57 eV (Fig. S1).

To investigate the influence of the counterion on CV in a nonpolar solvent, three individual MD simulations are performed for the chromophore in toluene. All three trajectories are initialized in the identical contact ion pair configuration, but have different solvent configurations. We find that all three trajectories reach the same steady-state counterion configuration after different equilibration times (see Fig. S15), meaning that solvent-induced disorder has only a limited influence on CV-ion configurations. Since all trajectories reach the same steady state, we focus on only a single trajectory with the fastest equilibration time for the chloride ion to compute optical spectra in the cumulant approach.

Once the trajectory is equilibrated, the ion is located between two nitrogen atoms in a stable configuration that lasts for the full 36 ps of computed QM/MM trajectory (Fig. S8), with an average distance to the two nearest nitrogens of 5.2 Å and 5.3 Å respectively. The close association with the N atoms throughout the entire trajectory suggests that the chloride ion forms a stable contact ion pair with CV. Interestingly, while the optimized geometry for CV in vacuum has the counterion mainly associated with a single N, in toluene the ion is found to have a higher tendency to be equidistant to both N (see Fig. S10). The equidistant positioning of the ion leads to less symmetry breaking of the CV electronic structure compared to the optimized vacuum position, leading to an overall smaller energy splitting between the S_1 and S_2 excitations compared to the vacuum configuration. An example of an equidistant is shown in Fig. S5 *(a)*. The greater the difference in distance to the two nearest nitrogens, the greater the symmetry breaking of the CV molecule, resulting in a larger S_1-S_2 energy difference (see Fig. S5 *(b)*, as well as Fig. S8). Along the 36 ps trajectory of CV in toluene, the average difference in distance to the two N atoms is 0.75 Å and the average S_1-S_2 energy difference is 0.51 eV, resulting in a wide splitting of 0.475 eV between the peaks in the simulated absorption spectrum (Fig. 5), that is nevertheless smaller than the splitting for CV in the presence of a counterion in vacuum (see Table I and Fig. S1). To further analyze the origin of the large S_1-S_2 splitting in nonpolar solvents, we performed vertical excitation energy calculations along the same MD trajectory snapshots in toluene, but with the counterion omitted from the calculation. The resulting spectrum is shown in the right panel of Fig. 5, confirming that the main contribution to the symmetry breaking and strong S_1-S_2 splitting in the absorption lineshape is

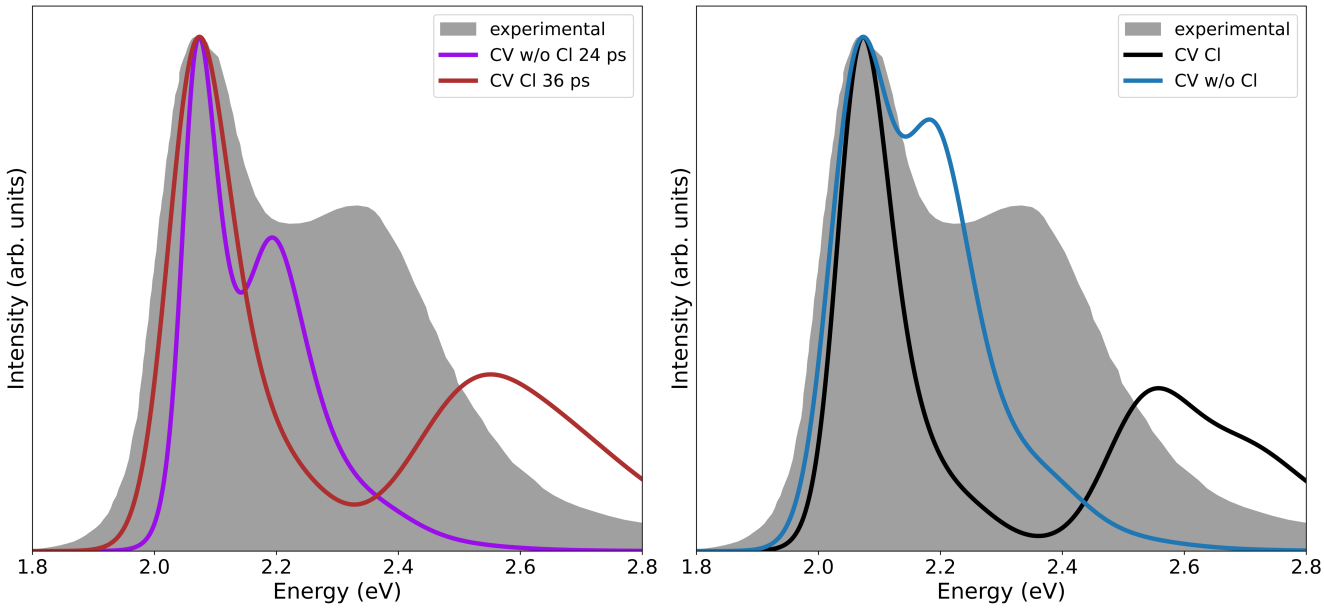

FIG. 5. Computed absorption spectra of CV in toluene. Left: The purple line is the cumulant spectrum of CV without a counterion calculated based on a single 24 ps (12,000 snapshots) QM/MM trajectory. The red line corresponds to the cumulant spectrum of CV with the counterion calculated based on a 36 ps (18,000 snapshots) QM/MM trajectory. The lines are shifted to match the experimental data by -0.759 eV and -0.692 eV, respectively. Right: The black line is the cumulant spectrum computed based on an 8 ps QM/MM trajectory where the solvent is treated as classical point charges in the TDDFT calculation. The blue line is same solvated trajectory but with the counterion stripped away when computing vertical excitation energies with TDDFT. The lines are shifted to match the experimental data by -0.662 eV and -0.738 eV, respectively. The experimental spectrum[1] is shown in gray.

due to the electrostatic interaction between CV and the chloride ion. When removing the ion, the absorption lineshape exhibits a strongly reduced splitting of 0.108 eV, which is significantly smaller than the experimentally observed splitting.

We note that the splitting in the calculated spectrum in toluene is considerably overestimated compared to the experimental one (see Fig. 5). There are several possible reasons that could explain this discrepancy: 1) an oversampling of contact ion configurations that cause large symmetry-breaking and S_1-S_2 splitting; 2) errors in the CV-Cl^- dynamics and computation of vertical excitation energies due to the treatment of the counterion as a classical point charge; 3) errors in the computed S_1-S_2 splittings for the strongly symmetry-broken configuration due to the choice of DFT functional.

First, while all QM/MM MD calculations are initialized with Cl^- in a contact ion pair configuration taken from CV in vacuum, all trajectories equilibrate to the same configuration distinct from the optimized vacuum configuration, that remains stable over 10s of ps, suggesting that the contact-ion pairing is indeed highly stable. We have tested other initial configurations for Cl^-, such as the ideal position chosen by AMBER to minimize the electrostatic potential, yielding similar final spectra (see SI Sec. 2). While the position of the counterion in the real solution might be combination of contact ion pairings and configurations with the ions further separated, the strong-electrostatic interactions in nonpolar solvents and the long-timescale stability of the contact-ion pairing in our simulation means we likely sufficiently sample the configurations that dominate the experimental lineshape. Second, we have carried out calculations where Cl^- is included in the QM region for both the QM/MM dynamics and the calculation of vertical excitation energies (see SI Sec. 5), with only minor influence on the resulting spectra, confirming that a FF treatment of Cl^- is not a dominant factor in the discrepancy with experimental lineshapes. Finally, we note that for the strongly symmetry-broken configuration caused by direct contact ion pairing, the S_1 state of CV has significantly less charge-transfer (CT) character than the S_2 state. This causes functionals with different treatment of long-range exchange to predict a range of S_1-S_2 splittings for the same CV-Cl^- configurations (see SI Sec. 5), with the CAM-B3LYP functional resulting in a much closer agreement with experiment for a short QM/MM trajectory of CV in toluene (Fig. S16). However, the CAM-B3LYP functional does not yield the correct description of CV in Methanol, significantly overestimating the intensity of the second spectral peak (Fig. S18), suggesting that due to the different strengths of environmental symmetry

breaking, standard DFT functionals do not describe CV in both condensed phase environments equally well. We conclude that the dominant factor causing the discrepancy between computed spectra of CV in toluene and experiment is likely errors in description of the CT character of the symmetry-broken S_2 state in the (TD)DFT calculations of vertical excitation energies.

D. The importance of counterions

While the inclusion of a counterion has no significant effect on the absorption lineshape of CV in polar solvent (Fig. 4), its presence becomes crucial in nonpolar solvent. Fig. 6 illustrates the distribution of the counterion along the 36 ps trajectory in toluene, where *(a)* shows the dependence of the S_1-S_2 energy splitting on the counterion position. We quantify the counterion position relative to CV through the mean distance to the two nearest nitrogen atoms, as well as the dimensionless difference in the distances to both atoms, defined through the unsigned difference in distance divided by the mean distance. The purple region in Fig. 6 *(a)* corresponds to almost zero splitting (i.e. degenerate S_1 and S_2 states), and for those configurations the counterion is almost equidistant from both N atoms. The maximum splitting of the excitation energies occurs when the counterion is much closer to the one of the nitrogen atoms than the other, represented by yellow to red regions on the heat map. Hence, for the contact ion pair configuration with CV, the main contribution to the energy splitting is due to the difference in the distance between the counterion and the two nitrogen atoms, since an increase in the difference in distance leads to an increase in the mean distance.

In Fig. 6 *(b)*), the probability distribution of the counterion positions, as computed along the 36 ps MD trajectory of CV in toluene, is displayed. We find that in toluene, the counterion is preferentially located equidistant between the two nearest nitrogens, a configuration that is associated with a comparatively lower S_1-S_2 splitting than the fully symmetry-broken configuration where the Cl^- ion is mainly associated with a single N. This is in contrast to the behavior of CV in vacuum (see Fig. S9), where the difference in the distance to the two nearest N atoms is significantly larger. Fig. S9 shows not only a clear relationship between the distance difference and the S_1-S_2 energy splitting, but also the relative position of the ion with respect to the nearest atom N and the energy difference. The average dimensionless difference in distance is 0.44, the average mean distance is 5.13 Å, and the dimensionless difference in energy is 0.18 along the vacuum trajectory, compared to the same quantities of 0.14, 5.24 Å, and 0.16 along the toluene trajectory. This suggests different counterion positions are sampled along these trajectories, with a greater S_1-S_2 splitting along the vacuum trajectory. Consequently, we can conclude the solvent environment restrains the motion of the counterion, explaining the origin of the larger S_1-S_2 splitting observed in vacuum as compared to CV in toluene.

A similar heatmap of ion configurations can also be constructed for CV in methanol (see Fig. S19). We note that in the polar solvent, the mean distance to the N is significantly increased to approximately 8.2 Å along the trajectory, suggesting that the ion is mostly screened by a solvation shell of methanol molecules. Additionally there is much less of a correlation between the ion position and the S_1-S_2 splitting, again confirming that the influence of the counter ion is effectively screened by the polar solvent as soon as Cl^- is not forming a contact ion pair with CV.

We conclude that the location of the counterion along the trajectory plays a crucial role in determining the splitting of the peaks in the absorption spectral lineshape in nonpolar solvents. Long QM/MM trajectories of the scale of 10s to 100s of ps are required to efficiently sample these configurations in the condensed phase, highlighting the challenge of modeling complex environmental interactions between chromophores and counter ions from first principles.

IV. CONCLUSIONS

In this study, we have calculated and analyzed the absorption spectra of CV in vacuum, methanol, and toluene using the cumulant approach. We confirmed that the degeneracy of the S_1 and the S_2 state is lifted through a symmetry-breaking of the molecule and its condensed phase surroundings, by either the presence of solvent and counterion interactions, or aryl ring torsion.

In polar solvents, we find that the presence of a counterion in the calculation has only minor influence on the predicted spectral lineshapes, with the calculated absorption spectrum of CV in methanol in absence of a counterion agreeing well with both the experimental spectrum and the spectrum computed by explicitly including a Cl^- in the dynamics. An analysis of the ion dynamics reveals that Cl^- is screened by a solvation shell of methanol, thus diminishing its effect on the optical spectrum. Additionally, our analysis demonstrates that the dominant factor responsible for the symmetry breaking of CV and the resulting splitting of S_1 and S_2 peaks in methanol is the non-covalent interaction between the CV and the solvent environment rather than torsional disorder, in contrast to the results presented in a recent study.[8]

In nonpolar solvents, we find that CV preferentially forms a contact ion configuration with Cl^-, leading to strong symmetry breaking and a substantially larger S_1-S_2 splitting than in methanol. Interestingly, our results demonstrate that the chromophore-ion configurations sampled during the MD in toluene, where the Cl^- is preferentially located equidistant between two N atoms, are qualitatively different to the configurations sampled in vacuum, where the Cl^- is more preferentially associated with a single N. This highlights that an explicit treatment of the condensed phase environment, as provided by MD sampling, is crucial in describing the com-

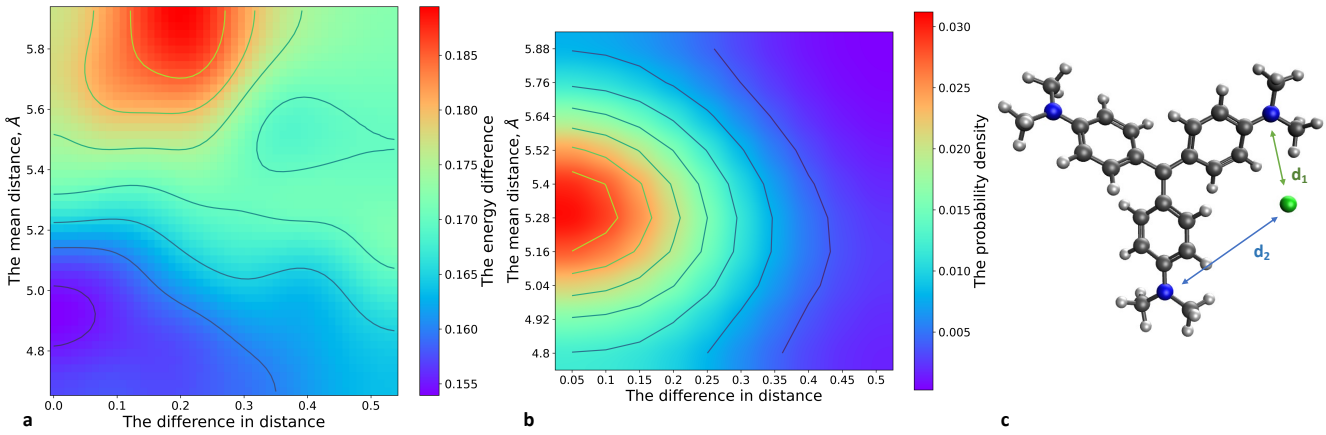

FIG. 6. *(a)* Relationship between the difference in distance between one nitrogen atom in the CV molecule to the counterion and another nitrogen atom to the counterion $\frac{|d_1-d_2|}{\frac{d_1+d_2}{2}}$, the mean distance between the counterion and two nitrogen atoms $\frac{d_1+d_2}{2}$, and the energy difference between the S_1 and S_2 excitations $\frac{|E_{S_1}-E_{S_2}|}{\frac{E_{S_1}+E_{S_2}}{2}}$. The 2D heat map was calculated based on the 36 ps trajectory in toluene. *(b)* Distribution of different counterion positions along the 36 ps trajectory in toluene. The color bar represents the number of occurrences of a certain difference in the distance between the two nitrogen atoms in the CV molecule and the counterion, and the average distance of the counterion to these nitrogen atoms, which determines which solvation shell the ion is in.

plex interplay between chromophore, counterion and solvent dynamics that shapes the optical spectrum in this system. Additionally, given that the position of the counterion breaks the symmetry of the electron density on the CV molecule (Table S2), static force fields are likely insufficient in describing their interactions, highlighting the necessity of performing long QM/MM dynamics, which can be computationally costly. The computed spectra for CV in methanol are in good agreement with the experimental lineshape. In toluene, our calculations overestimate the experimentally observed S_1-S_2 splitting. However, given the sensitivity of the splitting on the choice of DFT functional for the contact ion pair configurations due to the significant amount of CT character in the S_2 state (see Fig. S16), we can likely ascribe this discrepancy to an error in the chosen DFT functional, rather than a shortcoming of the cumulant method.

Our results highlight the strengths of the cumulant approach in modelling chromophores in the condensed phase. While more computationally demanding than the widely used Franck-Condon approach, generally requiring QM/MM dynamics and vertical energy calculations on 10s of thousands of MD snapshot, the method can describe environmental polarization effects, direct solute solvent interactions and slow collective chromophore-environment interactions from first principles. These types of interactions are ubiquitous in a wide variety of systems, ranging from semi-flexible solvated dyes to pigment-protein complexes.

V. ACKNOWLEDGEMENTS

T. J. Z. acknowledges startup funding provided by Oregon State University.

VI. SUPPORTING INFORMATION

See supporting information for additional analysis of geometric effects and solvent and counter ion effects on the computed spectra, the sampling of ion configurations along MD trajectories and the influence of DFT functional choice and choice of QM region on the computed spectra.

VII. DATA AVAILABILITY

The underlying data of this publication, including input files for MD trajectories and electronic structure calculations, force fields, and raw data of energy gap fluctuations that form the basis of the cumulant spectra reported in this work, can be found under the following persistent DOI: 10.5281/zenodo.10526781. All cumulant spectra have been generated with MolSpeckPy, a spectroscopy python code freely available on GitHub.[80]

[1] G. N. Lewis, T. T. Magel, and D. Lipkin, "Isomers of crystal violet ion. their absorption and re-emission of light," J. Am. Chem. Soc. **64**, 1774–1782 (1942).

[2] J. Korppi-Tommola and R. W. Yip, "Solvent effects on the visible absorption spectrum of crystal violet," Can. J. Chem. **59**, 191–194 (1981).

[3]H. B. Lueck, J. L. McHale, and W. D. Edwards, "Symmetry-breaking solvent effects on the electronic structure and spectra of a series of triphenylmethane dyes," J. Am. Chem. Soc. **114**, 2342–2348 (1992).
[4]Y. Maruyama, M. Ishikawa, and H. Satozono, "Femtosecond isomerization of crystal violet in alcohols," J. Am. Chem. Soc. **118**, 6257–6263 (1996).
[5]S. Lovell, B. J. Marquardt, and B. Kahr, "Crystal violet's shoulder," J. Chem. Soc., Perkin Trans. 2 , 2241–2247 (1999).
[6]L. M. Lewis and G. L. Indig, "Solvent effects on the spectroscopic properties of triarylmethane dyes," Dyes Pigm. **46**, 145–154 (2000).
[7]E. C. Wu, Q. Ge, E. A. Arsenault, N. H. C. Lewis, N. L. Gruenke, M. J. Head-Gordon, and G. R. Fleming, "Two-dimensional electronic-vibrational spectroscopic study of conical intersection dynamics: an experimental and electronic structure study," Phys. Chem. Chem. Phys. **21**, 14153–14163 (2019).
[8]J. Sissaoui, D. S. Budkina, and E. Vauthey, "Torsional disorder, symmetry breaking, and the crystal violet shoulder controversy," J. Phys. Chem. Lett. **14**, 5602–5606 (2023), pMID: 37307294.
[9]C. Daday, C. Curutchet, A. Sinicropi, B. Mennucci, and C. Filippi, "Chromophore-protein coupling beyond nonpolarizable models: Understanding absorption in green fluorescent protein," J. Chem. Theory Comput. **11**, 4825–4839 (2015).
[10]S. Aydinoglu, A. Pasti, T. Biver, and B. Mennucci, "Auramine o interaction with dna: a combined spectroscopic and td-dft analysis," Phys. Chem. Chem. Phys. **21**, 20606–20612 (2019).
[11]K. T. Smith, C. A. Ramsperger, K. E. Hunter, T. J. Zuehlsdorff, and K. C. Stylianou, "Colorimetric detection of acidic pesticides in water," Chem. Commun. **58**, 953–956 (2022).
[12]C. Curutchet, J. Kongsted, A. Muñoz Losa, H. Hossein-Nejad, G. D. Scholes, and B. Mennucci, "Photosynthetic light-harvesting is tuned by the heterogeneous polarizable environment of the protein," J. Am. Chem. Soc. **133**, 3078–3084 (2011), pMID: 21322565.
[13]C. M. Isborn, A. W. Götz, M. A. Clark, R. C. Walker, and T. J. Martínez, "Electronic absorption spectra from mm and ab initio qm/mm molecular dynamics: Environmental effects on the absorption spectrum of photoactive yellow protein," J. Chem. Theory Comput. **8**, 5092–5106 (2012).
[14]C. M. Isborn, B. D. Mar, B. F. E. Curchod, I. Tavernelli, and T. J. Martnez, "The charge transfer problem in density functional theory calculations of aqueously solvated molecules," J. Phys. Chem. B **117**, 12189–12201 (2013).
[15]E. Rivera, D. Montemayor, M. Masia, and D. F. Coker, "Influence of site-dependent pigment–protein interactions on excitation energy transfer in photosynthetic light harvesting," J. Phys. Chem. B **117**, 5510–5521 (2013), pMID: 23597258.
[16]M. R. Provorse, T. Peev, C. Xiong, and C. M. Isborn, "Convergence of excitation energies in mixed quantum and classical solvent: Comparison of continuum and point charge models," J. Phys. Chem. B **120**, 12148–12159 (2016).
[17]T. J. Zuehlsdorff, P. D. Haynes, F. Hanke, M. C. Payne, and N. D. M. Hine, "Solvent effects on electronic excitations of an organic chromophore," J. Chem. Theory Comput. **12**, 1853–1861 (2016).
[18]J. M. Milanese, M. R. Provorse, E. Alameda, and C. M. Isborn, "Convergence of computed aqueous absorption spectra with explicit quantum mechanical solvent," J. Chem. Theory Comput. **13**, 2159–2171 (2017).
[19]R. B. Murphy, D. M. Philipp, and R. A. Friesner, "A mixed quantum mechanics/molecular mechanics (qm/mm) method for large-scale modeling of chemistry in protein environments," J. Comput. Chem. **21**, 1442–1457 (2000).
[20]H. Lin and D. G. Truhlar, "Qm/mm: what have we learned, where are we, and where do we go from here?" Theor. Chem. Acc. **117**, 185–199 (2007).
[21]B. Mennucci, "Modeling absorption and fluorescence solvatochromism with qm/classical approaches," Int. J. Quantum Chem. **115**, 1202–1208 (2015).
[22]H. J. Kulik, J. Zhang, J. P. Klinman, and T. J. Martinez, "How Large Should the QM Region Be in QM/MM Calculations? The Case of Catechol *O*-Methyltransferase," J. Phys. Chem. B **120**, 11381–11394 (2016).
[23]F. Santoro, A. Lami, R. Improta, J. Bloino, and V. Barone, "Effective method for the computation of optical spectra of large molecules at finite temperature including the duschinsky and herzberg-teller effect: The qx band of porphyrin as a case study," J. Chem. Phys. **128**, 224311 (2008).
[24]A. Baiardi, J. Bloino, and V. Barone, "General time dependent approach to vibronic spectroscopy including franck–condon, herzberg–teller, and duschinky effects," J. Chem. Theory Comput. **9**, 4097–4115 (2013).
[25]J. Bloino, M. Biczysko, F. Santoro, and V. Barone, "General approach to compute vibrationally resolved one-photon electronic spectra," J. Chem. Theory Comput. **6**, 1256–1274 (2010).
[26]F. J. A. Ferrer, M. D. Davari, D. Morozov, G. Groenhof, and F. Santoro, "The lineshape of the electronic spectrum of the green fluorescent protein chromophore, part ii: Solution phase," ChemPhysChem **15**, 3246–3257 (2014).
[27]T. J. Zuehlsdorff, H. Hong, L. Shi, and C. M. Isborn, "Influence of Electronic Polarization on the Spectral Density," J. Phys. Chem. B **124**, 531–543 (2020).
[28]D. Loco, S. Jurinovich, L. Cupellini, M. F. S. J. Menger, and B. Mennucci, "The modeling of the absorption lineshape for embedded molecules through a polarizable QM/MM approach," Photochem. Photobiol. Sci. **17**, 552–560 (2018).
[29]T. J. Zuehlsdorff, J. A. Napoli, J. M. Milanese, T. E. Markland, and C. M. Isborn, "Unraveling electronic absorption spectra using nuclear quantum effects: Photoactive yellow protein and green fluorescent protein chromophores in water," J. Chem. Phys. **149**, 024107 (2018).
[30]J. Cerezo, D. Aranda, F. J. Avila Ferrer, G. Prampolini, and F. Santoro, "Adiabatic-molecular dynamics generalized vertical hessian approach: A mixed quantum classical method to compute electronic spectra of flexible molecules in the condensed phase," J. Chem. Theory Comput. **16**, 1215–1231 (2020), pMID: 31855424.
[31]J. Franck, "Elementary processes of photochemical reactions," J. Trans. Faraday Soc. **21**, 536–542 (1925).
[32]B. de Souza, F. Neese, and R. Izsák, "On the theoretical prediction of fluorescence rates from first principles using the path integral approach," J. Chem. Phys. **148**, 034104 (2018).
[33]M. I. Sorour, A. H. Marcus, and S. Matsika, "Unravelling the origin of the vibronic spectral signatures in an excitonically coupled indocarbocyanine cy3 dimer," J. Phys. Chem. A **127**, 9530–9540 (2023), pMID: 37934679.
[34]J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models," Chem. Rev. **105**, 2999–3094 (2005).
[35]B. Mennucci, "Polarizable continuum model," Wiley Interdiscip. Rev. Comput. Mol. Sci. **2**, 386–404 (2012).
[36]R. Crespo-Otero and M. Barbatti, "Spectrum simulation and decomposition with nuclear ensemble: formal derivation and application to benzene, furan and 2-phenylfuran," Theor. Chem. Acc. **131**, 1237 (2012).
[37]X. Ge, I. Timrov, S. Binnie, A. Biancardi, A. Calzolari, and S. Baroni, "Accurate and inexpensive prediction of the color optical properties of anthocyanins in solution," J. Phys. Chem. A **119**, 3816–3822 (2015).
[38]T. J. Zuehlsdorff, P. D. Haynes, M. C. Payne, and N. D. M. Hine, "Predicting solvatochromic shifts and colours of a solvated organic dye: The example of nile red," J. Chem. Phys. **146**, 124504 (2017).
[39]D. Loco and L. Cupellini, "Modeling the absorption lineshape of embedded systems from molecular dynamics: A tutorial review," Int. J. Quantum Chem. **119**, e25726 (2018).
[40]A. Kumar, R. Cardia, and G. Cappellini, "Electronic and optical properties of chromophores from bacterial cellulose," Cellulose **25**, 2191–2203 (2018).

[41]S. Di Grande, I. Ciofini, C. Adamo, M. Pagliai, and G. Cardini, "Absorption spectra of flexible fluorescent probes by a combined computational approach: Molecular dynamics simulations and time-dependent density functional theory," J. Phys. Chem. A **126**, 8809–8817 (2022), pMID: 36383687.
[42]A. Petrone, J. Cerezo, F. J. A. Ferrer, G. Donati, R. Improta, N. Rega, and F. Santoro, "Absorption and emission spectral shapes of a prototype dye in water by combining classical/dynamical and quantum/static approaches," J. Phys. Chem. A **119**, 5426–5438 (2015), pMID: 25699575.
[43]J. Cerezo, F. J. Avila Ferrer, G. Prampolini, and F. Santoro, "Modeling solvent broadening on the vibronic spectra of a series of coumarin dyes. from implicit to explicit solvent models," J. Chem. Theory Comput. **11**, 5810–5825 (2015).
[44]T. J. Zuehlsdorff and C. M. Isborn, "Modeling absorption spectra of molecules in solution," Int. J. Quantum Chem. **119**, e25719 (2019).
[45]T. J. Zuehlsdorff, S. V. Shedge, S.-Y. Lu, H. Hong, V. P. Aguirre, L. Shi, and C. M. Isborn, "Vibronic and environmental effects in simulations of optical spectroscopy," Annu. Rev. Phys. Chem. **72**, 165–188 (2021), pMID: 33395546, https://doi.org/10.1146/annurev-physchem-090419-051350.
[46]S. Mukamel, *Principles of Nonlinear Optical Spectroscopy* (Oxford University Press, New York, 1995).
[47]S. Mukamel, "Fluorescence and absorption of large anharmonic molecules. Spectroscopy without eigenstates," J. Phys. Chem. **89**, 1077 (1985).
[48]Y. Georgievskii, C.-P. Hsu, and R. A. Marcus, "Linear response in theory of electron transfer reactions as an alternative to the molecular harmonic oscillator model," J. Chem. Phys. **110**, 5307–5317 (1999).
[49]R. Borrelli and A. Peluso, "Quantum Dynamics of Radiationless Electronic Transitions including normal modes displacement and Duschinsky Rotations: A Second Order Cumulant Approach," J. Chem. Theory Comput. **11**, 415–422 (2015).
[50]T. J. Zuehlsdorff, A. Montoya-Castillo, J. A. Napoli, T. E. Markland, and C. M. Isborn, "Optical spectra in the condensed phase: Capturing anharmonic and vibronic features using dynamic and static approaches," J. Chem. Phys **151**, 074111 (2019).
[51]M. S. Chen, T. J. Zuehlsdorff, T. Morawietz, C. M. Isborn, and T. E. Markland, "Exploiting machine learning to efficiently predict multidimensional optical spectra in complex environments," J. Phys. Chem. Lett. **11**, 7559–7568 (2020).
[52]M. S. Chen, Y. Mao, A. Snider, P. Gupta, A. Montoya-Castillo, T. J. Zuehlsdorff, C. M. Isborn, and T. E. Markland, "Elucidating the role of hydrogen bonding in the optical spectroscopy of the solvated green fluorescent protein chromophore: Using machine learning to establish the importance of high-level electronic structure," J. Phys. Chem. Lett. **14**, 6610–6619 (2023), pMID: 37459252.
[53]A. J. Dunnett, D. Gowland, C. M. Isborn, A. W. Chin, and T. J. Zuehlsdorff, "Influence of non-adiabatic effects on linear absorption spectra in the condensed phase: Methylene blue," J. Chem. Phys. **155**, 144112 (2021).
[54]E. Condon, "Nuclear Motion Associated with Electron Transitions in Diatomic molecles," Phys. Rev. **32**, 858–872 (1928).
[55]A. Anda, L. De Vico, T. Hansen, and Abramavičius, "Absorption and fluorescence lineshape theory for polynomial potentials," J. Chem. Theory Comput. **12**, 5979–5989 (2016).
[56]A. F. Fidler and G. S. Engel, "Nonlinear spectroscopy theory of displaced harmonic oscillators with differing curvatures: A correlation function approach," J. Phys. Chem. A **117**, 9444–9453 (2013).
[57]J. S. Bader and B. J. Berne, "Quantum and classical relaxation rates from classical simulations," J. Chem. Phys. **100**, 8359 (1994).
[58]S. A. Egorov, K. F. Everitt, and J. L. Skinner, "Quantum Dynamics and Vibrational Relaxation," J. Phys. Chem. A **103**, 9494–9499 (1999).
[59]H. Kim and P. J. Rossky, "Evaluation of Quantum Correlation Functions from Classical Data," J. Phys. Chem. B **106**, 8240 (2002).
[60]I. R. Craig and D. E. Manolopoulos, "Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics," J. Chem. Phys. **121**, 3368 (2004).
[61]R. Ramirez, T. Lopez-Ciudad, P. K. P, and D. Marx, "Quantum corrections to classical time-correlation functions: Hydrogen bonding and anharmonic floppy modes," J. Chem. Phys. **121**, 3973 (2004).
[62]S. Valleau, A. Eisfeld, and A. Aspuru-Guzik, "On the alternatives for bath correlators and spectral densities from mixed quantum-classical simulations," J. Chem. Phys. **137**, 224103 (2012).
[63]B. F. E. Curchod and T. J. Martínez, "Ab initio nonadiabatic quantum molecular dynamics," Chem. Rev. **118**, 3305–3336 (2018).
[64]Z. R. Wiethorn, K. E. Hunter, T. J. Zuehlsdorff, and A. Montoya-Castillo, "Beyond the Condon limit: Condensed phase optical spectra from atomistic simulations," J. Chem. Phys. **159**, 244114 (2023).
[65]M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, "Gaussian Development Version, Revision I.09," (2016), Gaussian Inc. Wallingford CT.
[66]I. S. Ufimtsev and T. J. Martinez, "Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients and First Principles Molecular Dynamics," J. Chem. Theory Comput. **5**, 2619–2628 (2009).
[67]A. V. Titov, I. S. Ufimtsev, N. Luehr, and T. J. Martinez, "Generating Efficient Quantum Chemistry Codes for Novel Architectures," J. Chem. Theory Comput. **9**, 2213–221 (2013).
[68]J.-D. Chaia and M. Head-Gordon, "Systematic optimization of long-range corrected hybrid density functionals," J. Chem. Phys. **128**, 084106 (2008).
[69]T. H. Dunning Jr., "Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen," J. Chem. Phys. **90**, 1007 (1989).
[70]M. E. Casida, "Time-dependent density functional response theory for molecules," World Scientic, Singapore , 155–192 (1995).
[71]S. Hirata and M. Head-Gordon, "Time-dependent density functional theory within the Tamm–Dancoff approximation," Chem. Phys. Lett. **314**, 291–299 (1999).
[72]C. M. Isborn, N. Luehr, I. S. Ufimtsev, and T. J. Martínez, "Excited-State Electronic Structure with Configuration Interaction Singles and Tamm–Dancoff Time-Dependent Density Fucntional Theory on Graphical Processing Units," J. Chem. Theory Comput. **7**, 1814–1823 (2011).
[73]A. Warshel and M. Levitt, "Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme," J. Mol. Biol. **103**, 227–249 (1976).
[74]D. A. Case, T. A. Darden, I. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossvari, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, T. G. H. C. Q. Brozell, S. R.; Stein-

brecher, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. A. Kollman, "Amber 12," (2012), university of California: San Francisco, CA.

[75]R. C. Walker, M. F. Crowley, and D. A. Case, "The implementation of a fast and accurate QM/MM potential method in Amber," J. Comput. Chem. **29**, 1019–1031 (2007).

[76]J. Wang, R. M. Wolf, P. A. Caldwell, J. W. an Kollman, and D. A. Case, "Developing and testing of a general amber force field," J. Comput. Chem. **25**, 1157–1174 (2004).

[77]J. Wang, W. Wang, P. A. Kollman, and D. A. Case, "Automatic atom type and bond type perception in molecular mechanical calculations," J. Mol. Graph. Model. **25**, 247–260 (2006).

[78]J. M. Martínez and L. Martínez, "Packing optimization for automated generation of complex system's initial configurations for molecular dynamics and docking," J. Comput. Chem. **24**, 819–825 (2003).

[79]M. K. Lee, P. Huo, and D. F. Coker, "Semiclassical Path Integral Dynamics: Photosynthetic Energy Transfer with Realistic Environment Interactions," Annu. Rev. Phys. Chem. **67**, 639–668 (2016).

[80]T. J. Zuehlsdorff, "Molspeckpy: Spectroscopy python code, available on github: https://github.com/tjz21/spectroscopy_python_code," (2021), accessed on 02/10/2023.

[81]S. V. Shedge, T. J. Zuehlsdorff, M. J. Servis, A. E. Clark, and C. M. Isborn, "Effect of ions on the optical absorption spectra of aqueously solvated chromophores," J. Phys. Chem. A **123**, 6175–6184 (2019).