Taming The Third Order Cumulant Approximation to Linear Optical Spectroscopy

Lucas Allan[1] and Tim J. Zuehlsdorff[1, a)]
Department of Chemistry, Oregon State University, Corvallis, Oregon 97331, USA

The second order cumulant method offers a promising pathway to predicting optical properties in condensed phase systems. It allows for the computation of linear absorption spectra from excitation energy fluctuations sampled along molecular dynamics (MD) trajectories, fully accounting for vibronic effects, direct solute-solvent interactions, and environmental polarization effects. However, the second order cumulant approximation only guarantees accurate lineshapes for energy gap fluctuations obeying Gaussian statistics. A third order correction has been derived recently, but often yields unphysical spectra or divergent lineshapes for moderately non-Gaussian fluctuations, due to the neglect of higher order terms in the cumulant expansion. In this work, we develop a corrected cumulant approach, where the collective effect of neglected higher order contributions is approximately accounted for through a dampening factor applied to the third order cumulant term. We show that this dampening factor can be expressed as a function of the skewness and kurtosis of the energy gap fluctuations and can be parameterized from a large set of randomly sampled model Hamiltonians for which exact spectral lineshapes are known. The approach is shown to systematically remove unphysical contributions in the form of negative absorbances from cumulant spectra in both model Hamiltonians and condensed phase systems sampled from MD, and dramatically improves over the second order cumulant method in describing systems exhibiting Duschinsky mode mixing effects. We successfully apply the approach to the coumarin-153 dye in toluene, obtaining an excellent agreement with experiment.

I. INTRODUCTION

First-principles predictions of linear optical spectra of molecules in the condensed phase remains a challenging problem in computational chemistry.[1–6] The linear absorption and fluorescence lineshapes of chromophores often have vibronic contributions, making an explicit quantum mechanical treatment of the nuclei necessary. Additionally, the condensed phase environment can strongly influence optical spectra, both through environmental polarization effects and direct chromophore-environment interactions, such as hydrogen bonding in protic solvents.[6–11] Slow, generally anharmonic, collective chromophore-environment motion also poses a challenge and is often of importance in complex biological systems.[3,12,13] Robust computational approaches capable of capturing these effects from first principles are desirable, as experimental condensed phase optical spectra are often highly congested, making it challenging to establish structure-property relationships from experimental data alone. Additionally, accurately modeling linear spectroscopy often forms the first step in trying to interpret more complicated non-linear experiments used to probe excited state relaxation dynamics in complex systems.[8,14–27]

A commonly used framework to model optical spectra of molecules is the Franck-Condon (FC) approach.[28–37] It relies on approximating the ground- and excited state potential energy surfaces (PESs) as harmonic around their respective minima. Nuclear wavefunctions are then harmonic oscillator wavefunctions of vibrational modes and the intensity of vibronic peaks is directly related to ground- and excited state wavefunction overlaps following Fermi's golden rule.[35] The methodology accounts for changes in curvature between the ground- and excited state PES, as well as Duschinsky[38] mode-mixing effects. The underlying harmonic nuclear Hamiltonian (referred to as a Generalized Brownian Oscillator Model or GBOM,[39] see Appendix A 1) can be directly parameterized from ground- and excited state geometry optimizations and frequency calculations implemented in many electronic structure methods. Additionally, the exact optical spectrum for the GBOM Hamiltonian can be computed analytically.[36,37] The method performs very well in predicting linear spectra for small, rigid molecules in non-polar solvents, where the harmonic approximation is expected to hold. However, the approach can struggle in semi-flexible molecules undergoing anharmonic motion,[40–45] and cannot account for strong solute-solvent interactions, as the condensed phase environment is generally represented collectively through polarizable continuum models (PCMs).[46–48] To compare directly to experiments, solvent broadening effects have to be accounted for through approximate broadening parameters,[49,50] or by invoking timescale separation arguments.[4,5,51] These shortcomings limit the applicability of the FC approach in large semi-flexible molecules and systems where the chromophore and the condensed phase environment undergo slow, coupled motion.

An alternative approach to constructing linear spectra in the condensed phase is the cumulant method,[52–54] where the linear response function is directly constructed from the fluctuations of the excitation energy sampled along ground-state molecular dynamics (MD)

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

simulations in thermal equilibrium.[3,4,10,39,55,56] Direct solute-solvent interactions, anharmonic effects, collective chromophore-environment motion and environmental polarization effects[10] are all accounted for in the MD sampling, making it a highly promising, albeit computationally expensive, approach in biological systems and pigment-protein complexes.[6,16–19,21–26,57] Additionally, the computational cost of sampling energy gap fluctuations can be significantly reduced using machine learning (ML) techniques.[58,59] However, in practical calculations, the cumulant expansion is generally truncated at second order, corresponding to mapping energy gap fluctuations onto a bath of linearly coupled harmonic oscillators (Brownian Oscillator Model or BOM). This truncation is only exact for systems where energy-gap fluctuations follow Gaussian statistics, and already leads to errors in the harmonic GBOM Hamiltonian including changes in PES curvature and Duschinsky mode mixing effects,[39,60] that are captured exactly by the FC approach.[36,37]

The inclusion of higher order cumulants is vital in computing the correct lineshapes many systems,[39,60,61] and we have recently demonstrated that a third order correction term can be constructed directly from MD, yielding improvements of spectral shapes in systems with small non-Gaussian contributions to the energy gap fluctuations. However, the approach has the significant shortcoming of being numerically unstable, yielding unphysical spectra with negative absorbances and divergent lineshapes for moderate to strong non-Gaussian fluctuations.[39,61] These divergences can be traced to the discarding of higher order terms in the cumulant expansion and have been observed previously in 1D anharmonic model potentials.[61] Since it is in general difficult to determine *a priori* whether the inclusion of a third order cumulant correction will improve the spectral lineshape or cause unphysical spectral contributions, the benefit of extending the cumulant method beyond second order in realistic condensed phase systems is questionable.

In this work, we formulate and rigorously test an alternative approach to go beyond the second order cumulant expansion. The method relies on applying a dampening factor to the third order cumulant correction that approximately accounts for the effect of neglected higher order cumulants in cancelling divergences in the third order lineshape. We demonstrate that the dampening factor Φ is a function of the non-Gaussian nature of the energy gap fluctuations quantified via the skewness and kurtosis of the distribution. This function Φ can be directly parameterized using the exactly solvable model system of the GBOM Hamiltonian. We show that the approach rigorously removes unphysical features from linear absorption spectra in both harmonic and anharmonic model systems, as well as realistic condensed phase systems sampled from MD. Additionally, the approach yields an excellent agreement with the FC method for the GBOM Hamiltonian including Duschinsky mode mixing and changes in PES curvature, thus opening up the possibility of accounting for these effects from first principles in simulations of molecules embedded in condensed phase environments.

II. THEORY

In this work, we focus on the linear response of a two-state electronic system coupled to nuclear motion. Within the Born-Oppenheimer approximation[62] we can write:

$$|\Psi\rangle \in \{|e\rangle \otimes |\nu_e\rangle, |g\rangle \otimes |\nu_g\rangle\} \quad (1)$$

$$\hat{\mathrm{H}} = |e\rangle\hat{\mathrm{H}}_e\langle e| + |g\rangle\hat{\mathrm{H}}_g\langle g|. \quad (2)$$

Here, $\hat{\mathrm{H}}_g$ and $\hat{\mathrm{H}}_e$ denote nuclear Hamiltonians on the ground- and excited state potential energy surface (PES) respectively, $|e\rangle$ and $|g\rangle$ are pure electronic states, and ν_e and ν_g are nuclear wavefunctions. Under this form, the linear response function can be expressed as a trace over nuclear degrees of freedom in the condensed-phase system[52]

$$\chi^{(1)}(\mathrm{t}) = \mathrm{Tr}[\rho_g \hat{\mathrm{V}}_{ge}(\mathrm{t})\hat{\mathrm{V}}_{eg}(0)], \quad (3)$$

where ρ_g is the ground-state equilibrium density matrix of nuclear degrees of freedom and $\hat{V}_{eg} = \hat{V}_{ge}^{\dagger}$ is the transition dipole operator between electronic states $|e\rangle$ and $|g\rangle$. If an exact expression for Eqn. 3 can be obtained for a system of interest, the linear absorption spectrum can be constructed through its Fourier transform:[52]

$$\sigma(\omega) = \alpha\,\omega\mathbf{Re}[\mathcal{F}(\chi^{(1)}(\mathrm{t}))]. \quad (4)$$

Here, α is a constant factor necessary when comparing directly to experimental results,[36] but that will be set to 1 for the remainder of this work without loss of generality. Proposing approximate forms for Eqn. 3 that retain accuracy while remaining tractable in complex condensed phase systems is of key importance to the simulation of linear absorption spectra.[39]

A. Spectral Decomposition and Franck-Condon Response

Operating under the Condon approximation,[63,64] such that the electronic transition dipole operator is taken to be independent of nuclear degrees of freedom, and assuming the vibrational eigenstates of the ground and excited state nuclear Hamiltonians are known, Eqn. 3 can be recast into a spectral decomposition:[28,35]

$$\chi^{(1)}_{\mathrm{spectral}}(\mathrm{t}) = |\mathrm{V}_{eg}|^2 \sum_{i,j} \mathrm{P}(i)|\langle\nu_j^e|\nu_i^g\rangle|^2 \exp[-\mathrm{i}(\omega_{ji})\mathrm{t}], \quad (5)$$

where,

$$\mathrm{P}(i) = e^{-\mathrm{E}_i^g\beta} / \sum_j e^{-\mathrm{E}_j^g\beta}, \quad (6)$$

$$\omega_{ji} = \mathrm{E}_j^e - \mathrm{E}_i^g. \quad (7)$$

Here, E_i^e and E_i^g denote the eigenvalues corresponding to the ith eigenstate of the nuclear Hamiltonians $\hat{H}_e$ and $\hat{H}_g$ respectively, $\mathrm{P}(i)$ is the Boltzmann population of the ith energy level on the ground state PES in thermal equilibrium and ω_{ji} is the energy required to transition from state i on the ground state PES to state j on the excited state PES. For model systems where the eigenstates of the nuclear Hamiltonians can be computed exactly, the spectral decomposition yields the exact linear response function under the Condon approximation. This form provides a crucial benchmark within the scope of this work as a means to quantify the performance of approximate cumulant-based schemes (see Sec. II B). However, the sum-over-states formalism in Eqn. 5 has a number of drawbacks, especially when modeling realistic molecular systems in the condensed phase.

First, Eqn. 5 requires the exact nuclear wavefunctions $\{\nu_j^e\}, \{\nu_i^g\}$ to be known and the wavefunction overlap $\langle \nu_j^e | \nu_i^g \rangle$ to be easily evaluated for all i,j. In practice, this means that the full PESs of the molecule in the condensed phase have to be approximated through a model Hamiltonian. The most popular choice, invoked in the Franck-Condon (FC) method,[28–37] is to approximate the PESs as harmonic around their respective minima. Second, evaluating Eqn. 5 directly, for example by parameterizing a GBOM (see Appendix A 1), means that the condensed phase environment can often only be accounted for approximately in the resulting spectrum, either through collective representations of the solvent environment through a PCM, or by invoking a timescale separation between solute and solvent degrees of freedom.[5,49,65,66]

These shortcomings render a direct evaluation of the spectral decomposition expression potentially impractical in condensed phase systems, especially for systems with strong solute-solvent interactions and chromophores with low frequency anharmonic modes ill described by a harmonic model PES.[10,40,67]

B. Cumulant Response

An alternative formulation that avoids some of the issues of a direct evaluation of Eqn. 5 is the cumulant expansion. Beginning with Eqn. 3 and working under the Condon approximation, we define the energy gap fluctuation operator

$$\delta\hat{\mathrm{U}}(\hat{\mathbf{q}};\mathrm{t}) = (\hat{\mathrm{H}}_e - \hat{\mathrm{H}}_g) - \omega_{eg}^{\mathrm{av}}, \tag{8}$$

with $\omega_{eg}^{\mathrm{av}}$ denoting the thermal average of the energy gap between the two PESs in thermal equilibrium

$$\omega_{eg}^{\mathrm{av}} = \mathrm{Tr}[\rho_g \hat{\mathrm{U}}], \tag{9}$$

and $\hat{\mathrm{U}} = \hat{\mathrm{H}}_e - \hat{\mathrm{H}}_g$ is the energy gap operator. The exact response (Eqn. 3) can then be recast in form of a time-ordered exponential of $\delta\mathrm{U}(\hat{\mathbf{q}},\mathrm{t})$:

$$\chi^{(1)}(\mathrm{t}) = |\mathrm{V}_{eg}|^2 e^{-\mathrm{i}\omega_{eg}^{\mathrm{av}}\mathrm{t}} \mathrm{Tr}\left[\rho_g \exp_+\left(-\mathrm{i}\int_0^{\mathrm{t}} \mathrm{d}\tau\, \delta\mathrm{U}(\hat{\mathbf{q}};\tau)\right)\right]. \tag{10}$$

One can then leverage the ansatz that the infinite series expansion of time-ordered integrals may be treated as a moment-generating function and expressed in respective cumulants.[52,53] The response function of the system can then be expressed as:

$$\chi^{(1)}(\mathrm{t}) = |\mathrm{V}_{eg}|^2 e^{-\mathrm{i}\omega_{eg}^{\mathrm{av}}\mathrm{t}} \exp\left[-\sum_{i=2}^{\infty} \mathrm{g}_i(\mathrm{t})\right] \tag{11}$$

such that $\mathrm{g}_i(\mathrm{t})$ denotes the ith order cumulant of the energy gap fluctuation operator and $\sum_{i=2}^{\infty} \mathrm{g}_i(\mathrm{t}) \equiv \mathrm{g}(\mathrm{t})$ is known as the lineshape function. Being of substantial importance of this work, the general form of the second and third order cumulant contributions to the lineshape function are presented here:

$$\mathrm{g}_2(\mathrm{t}) = \int_0^{\mathrm{t}} \mathrm{dt}_1 \int_0^{\mathrm{t}_1} \mathrm{dt}_2\, \mathrm{C}_{\mathrm{qm}}^{(2)}(\mathrm{t}_1 - \mathrm{t}_2), \tag{12}$$

where $\mathrm{C}_{\mathrm{qm}}^{(2)}(\tau) = \mathrm{Tr}[\rho_g \delta\hat{\mathrm{U}}(\tau)\delta\hat{\mathrm{U}}(0)]$ is the quantum autocorrelation function of the energy gap fluctuation operator. The third order cumulant contribution to the lineshape takes the general form:

$$\mathrm{g}_3(\mathrm{t}) = -\mathrm{i}\int_0^{\mathrm{t}} \mathrm{dt}_1 \int_0^{\mathrm{t}_1} \mathrm{dt}_2 \int_0^{\mathrm{t}_2} \mathrm{dt}_3\, \mathrm{C}_{\mathrm{qm}}^{(3)}(\mathrm{t}_2 - \mathrm{t}_3, \mathrm{t}_1 - \mathrm{t}_3) \tag{13}$$

and $\mathrm{C}_{\mathrm{qm}}^{(3)}(\tau',\tau) = \mathrm{Tr}[\rho_g \delta\hat{\mathrm{U}}(\tau)\delta\hat{\mathrm{U}}(\tau')\delta\hat{\mathrm{U}}(0)]$ represents the two-time quantum correlation function. Analogous expression can be obtained for higher-order cumulants.[61]

1. Construction of Cumulants

A key distinction must be drawn between how the cumulants may be obtained in the context of realistic condensed-phase systems versus model Hamiltonians. For chromophores embedded in complex environments, the quantum correlation functions required to construct the cumulant expansion are generally inaccessible. Instead, cumulants are constructed through classical MD simulations in conjunction with calculations of the vertical excitation energy along the generated trajectory.[3,56,68,69] From these vertical energies, *classical* correlation functions can be computed, and the approximate *quantum* correlation functions can be reconstructed with the help of quantum correction factors (QCFs).[57,70–75] For the one- and two-time correlation functions necessary to construct the second and third order cumulant, the QCFs used in this work take the

following form in the frequency domain:[39,57,72,75]

$$\mathrm{C}^{(2)}_{\mathrm{qm}}(\omega) \approx \frac{\beta\omega}{1-e^{-\beta\omega}}\mathrm{C}^{(2)}_{\mathrm{cl}}(\omega) \quad (14)$$

$$\mathrm{C}^{(3)}_{\mathrm{qm}}(\omega,\omega') \approx \frac{\bar{\omega}\omega'\omega\beta^2\mathrm{C}^{(3)}_{\mathrm{cl}}(\omega,\omega')}{2(\omega' e^{-\beta\bar{\omega}} - \bar{\omega}e^{-\beta\omega'} + \omega)}, \quad (15)$$

with $\bar{\omega} = \omega + \omega'$. Substituting the Fourier representation of Eqns. 14 and 15 into Eqn. 12 and Eqn. 13, approximate expressions for $\mathrm{g}_2(\mathrm{t})$ and $\mathrm{g}_3(\mathrm{t})$ may be obtained that can be computed directly from energy gap fluctuations sampled along MD trajectories. To the best of our knowledge, there are no generally applicable correction factor analogs for higher order correlation functions. Thus the MD-type construction of the cumulant response is limited to the third order, and the implication of such truncation of the cumulant expansion will be explored in detail in this work.

In contrast, for simplified model Hamiltonians, it is often possible to construct exact closed-form expressions of the second, third, and higher order cumulants.[39,60,61] Additionally, higher order cumulants for model systems can be constructed numerically.[61] These model systems can provide powerful insights into the types of errors made in evaluating low-order cumulant expansions using QCFs. In this work, we focus on the GBOM Hamiltonian[39] (see Sec. A 1), for which analytical expressions for $\mathrm{g}_2(\mathrm{t})$ and $\mathrm{g}_3(\mathrm{t})$ based on *quantum* and *classical* correlation functions and QCFs, as well as the exact result corresponding to the infinite order cumulant expansion[36,37] can be computed analytically (see SI Sec. I and Ref. 39).

2. Truncation of the Cumulant Expansion

For general molecular Hamiltonians, all orders of cumulants contribute to the spectral lineshape and a truncation at any finite order introduces errors. However, in the limit that the energy gap fluctuations obey Gaussian statistics, all cumulants beyond $\mathrm{g}_2(\mathrm{t})$ vanish,[39,60] such that exact response is recovered at the second order cumulant approximation. This behavior is observed for a system where all modes coupled to the electronic transition are harmonic and equal in ground and excited state curvature, resulting in an energy gap operator that is linear with respect to nuclear coordinates (See SI Sec I). Such a simplified model system is known as the Brownian oscillator model (BOM).

A departure from Gaussian statistics is quickly met in the realistic case of mismatched ground- and excited state frequencies in harmonic vibrational modes (as present in the GBOM Hamiltonian), as well as in the presence of any anharmonicity.[61] It can be shown analytically (see SI Secs. I and II) or numerically (SI Sec. III), that the introduction of higher order cumulants leads to an improved agreement with the exact response function in the short timescale limit, but that higher order terms become increasingly more volatile and divergent in the long timescale limit. In principle, a complete summation of the cumulant expansion to infinite order would lead to a cancellation of divergent terms and well-behaved response functions for all timescales. However, a truncation at finite order can yield unphysical lineshapes, as can be demonstrated even in the case of the harmonic GBOM Hamiltonian (SI Sec. II and Ref. 39), and only the second order cumulant approximation is guaranteed to yield linear response functions that do not diverge as $\mathrm{t} \to \infty$ for arbitrary Hamiltonians.[39]

This divergent behavior makes the benefit of evaluating higher order cumulants in realistic systems, such as by introducing a third order correction that can be computed in condensed phase systems from MD, questionable. In previous work, we have demonstrated numerically that monitoring the deviations of the energy gap fluctuations from Gaussian statistics, by calculating the skewness $\mu_{(3)}$ of the distribution, can serve as an indicator of the importance of higher order cumulant contributions.[39] It was found that for skewness values beyond 0.3, the third order cumulant correction becomes numerically unstable and unreliable due to the neglecting of higher order terms;[39] In these cases the numerically stable, albeit inaccurate, second order cumulant approximation should be used instead.

In this work, we instead introduce and justify the novel hypothesis that while the exact nature of neglected higher order cumulants in the expansion cannot be known in arbitrary systems, some information of their net effect on the lineshape can be inferred. By quantifying the degree of non-Gaussian behavior in the energy gap fluctuations, through metrics such as skewness $\mu_{(3)}$ and excess kurtosis $\mu_{(4)}$, one may approximately reintegrate the effects of the neglected cumulants in cancelling divergences of the lineshape. Specifically, we propose that an improved, and most importantly more stable, approximation to the exact lineshape compared to a pure third order cumulant expansion can be obtained through

$$\chi^{(1)}_{\mathrm{cumul},\Phi}(\mathrm{t}) = |\mathrm{V}_{eg}|^2 e^{-\mathrm{i}\omega^{\mathrm{av}}_{eg}\mathrm{t}} \times \exp\left[-\mathrm{g}_2(\mathrm{t}) - \Phi\mathrm{g}_3(\mathrm{t})\right]. \quad (16)$$

Here, Φ is a dampening factor approximately describing the effect of all neglected higher order cumulants such that $0 \leq \Phi \leq 1$. $\Phi \equiv \Phi\left(\mu_{(3)}, \mu_{(4)}\right)$ is taken to be an explicit function of the skewness $\mu_{(3)}$ and excess kurtosis $\mu_{(4)}$ quantifying the non-Gaussian nature of the energy gap fluctuations. The functional form of Eqn. 16 is further justified in Sec. II C.

The skewness and excess kurtosis of the energy gap

fluctuations can be defined in the following way:

$$\mu_{(1)} = \int \mathrm{d}\omega' \omega' \overline{\sigma}_{\mathrm{ens}}(\omega') \tag{17}$$

$$\mu_{(2)} = \int \mathrm{d}\omega' (\omega' - \mu_{(1)})^2 \overline{\sigma}_{\mathrm{ens}}(\omega') \tag{18}$$

$$\mu_{(3)} = \mu_{(2)}^{-\frac{3}{2}} \int \mathrm{d}\omega' (\omega' - \mu_{(1)})^3 \overline{\sigma}_{\mathrm{ens}}(\omega') \tag{19}$$

$$\mu_{(4)} = \mu_{(2)}^{-2} \int \mathrm{d}\omega' (\omega' - \mu_{(1)})^4 \overline{\sigma}_{\mathrm{ens}}(\omega') - 3 \tag{20}$$

where $\mu_{(1)}$ and $\mu_{(2)}$ are the mean and the variance, and

$$\overline{\sigma}_{\mathrm{ens}}(\omega) = \sigma_{\mathrm{ens}}(\omega) / \int \mathrm{d}\omega' \sigma_{\mathrm{ens}}(\omega') \tag{21}$$

is the normalized classical ensemble spectrum[39] of the absorption lineshape. The classical ensemble spectrum is obtained by discarding the quantum nature of the nuclei, as well as any vibronic effects, and directly encodes the statistics of the energy gap fluctuations. It can be expressed within a classical phase-space representation:[39]

$$\chi_{\mathrm{ens}}(\mathrm{t}) = |\mathrm{V}_{eg}|^2 \int \mathrm{d}\mathbf{q} \int \mathrm{d}\mathbf{p} \frac{e^{-\mathrm{H}_g(\mathbf{p},\mathbf{q})\beta}}{Z} e^{-\mathrm{iU}(\mathbf{q})\mathrm{t}}. \tag{22}$$

Eqn. 22 can be directly evaluated by sampling vertical excitation energies along a classical MD trajectory,[7,76–79] and can therefore be straightforwardly applied to any condensed phase system.

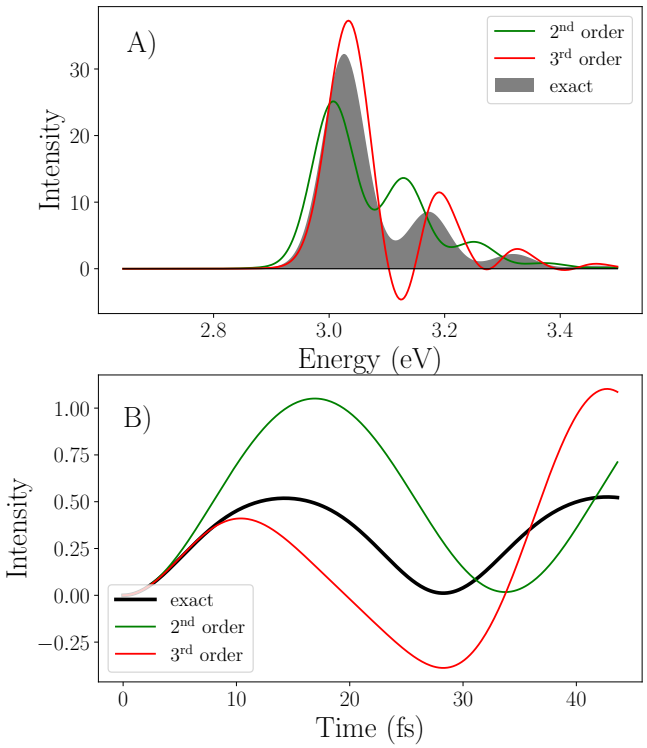

FIG. 1. A) Exact, second-order and third-order cumulant spectra for a 2-mode GBOM. A Gaussian broadening representing solvent effects has been added to all lineshapes B) Real part of the second and third order cumulant approximation to the lineshape function g(t) in comparison with the exact lineshape function (shown in black).

C. The dampening factor Φ in a harmonic model system

To illustrate the points made in Sec. II B 2, we demonstrate the effect of truncating the cumulant expansion at low order on a simple 2-mode harmonic model system described by the GBOM Hamiltonian (see Appendix A 1 and Appendix A 5 a for model parameters). For this Hamiltonian, the second order cumulant expansion is no longer exact, as the ground- and excited state PESs can have different curvatures and individual modes can be coupled. The resulting spectra for a specific model parameterization can be found in Fig. 1.

In the chosen parameterization, energy gap fluctuations are moderately non-Gaussian (with $\mu_{(3)} = 0.33$ and $\mu_{(4)} = 0.22$). The second order cumulant approximation fails to reproduce the correct vibronic progression, both underestimating the separation between vibronic peaks and predicting wrong intensities. When adding the third order correction to the lineshape, both peak intensities and positions are much improved, but at the cost of spurious unphysical spectral regions of negative absorbance. This spurious behavior of the truncated third order cumulant expansion is commonly observed for sufficiently non-Gaussian fluctuations.[39,61] We note that, while deviating significantly from the exact lineshape, the second order cumulant approximation is guaranteed to yield a positive-definite and thus physically meaningful absorption spectrum for arbitrary non-Gaussian energy gap fluctuations (see SI Sec. II).

The performance of the truncated cumulant expansion can be rationalized when considering the lineshape function g(t) determining the spectrum (see Fig. 1 B)). The third order cumulant approximation yields a short timescale correction towards the exact lineshape function, causing an improvement in the placement and intensity of vibronic progressions in third-order spectra of mildly non-Gaussian systems (See SI Sec. II A). However, this short time correction causes a divergence in the real part of the lineshape function proportional to t as $\mathrm{t} \to \infty$, with additional divergent terms present in the imaginary part (See SI Sec. II and Ref. 39). By inspection of Eqn. 5, it may be noted that an exact lineshape function should not exhibit this divergence. Characteristics of the divergent terms, such as the fact that the real part of the lineshape function can become negative in the short timescale (see Fig. 1 B)) can be seen as the origin of spurious negative spectral features in the resulting absorption spectrum. Thus, this spurious asymptotic behavior is a clear target for seeking methodological improvements to the computation of cumulant-based spectra. If the goal is to produce the most accurate, physically applicable spec-

trum possible, the inclusion of a correction that causes unphysical negative absorbances is undesirable, even if it leads to observed improvements in position and intensity of vibronic features. Therefore, it becomes of high interest to develop a methodology that retains the corrections afforded by the third order cumulant approach while reliably safeguarding against any unphysical spectral features. Ultimately, any technique developed must also be freely applicable to MD-type simulations of molecules in the condensed phase in order to be useful in realistic systems.

We propose that a low order cumulant expansion can be significantly improved by introducing a dampening factor Φ to the third order contribution $g_3(t)$, as outlined in Eqn. 16. To demonstrate this concept, we again turn to the harmonic model system in Fig. 1. We first define a goodness-of-fit metric for the cumulant spectrum with respect to the exact spectrum:

$$\kappa = \frac{1}{n}\sum_{i}^{n} |\sigma_{\text{cumul}}(\omega_i) - \sigma_{\text{exact}}(\omega_i)| \tag{23}$$

The metric κ is simply taken as the unsigned difference between the cumulant and the exact spectrum, averaged over n numerical grid points along the frequency axis. We evaluate σ_{cumul} using Eqn. 16, where the parameter Φ is systematically varied from 0 (corresponding to a pure second order cumulant spectrum) to 1 (corresponding to an undampened cumulant expansion truncated at third order). A plot of the resulting $\kappa(\Phi)$ can be found in Fig. 2 A).

We find that $\kappa(\Phi)$ is a convex function, and thus by selecting the prefactor value corresponding to the minimum of $\kappa(\Phi)$, one may optimize the performance of the third order cumulant approximation with respect to the exact spectrum. Selecting the optimal dampening factor Φ when constructing the third order cumulant lineshape following Eqn. 16 results in a spectral lineshape in very close agreement to the exact spectrum, both in the intensity and positioning of vibronic peaks (see Fig. 2 B and SI Sec. V). Additionally, the unphysical negative absorbance is removed from the spectrum. While Fig. 2 shows results for a specific parameterization of the GBOM Hamiltonian, we find that similar results can be obtained over a wide parameter range, suggesting that the functional form of a dampened cumulant response proposed in Eqn. 16 is widely applicable.

D. The dampening factor Φ in general systems

The results in Fig. 2 are obtained for a harmonic model system. In this work, we argue that 1) the general functional form of Eqn. 16 provides a pathway of improving cumulant lineshapes even in realistic systems with more complex (generally anharmonic) potential energy surfaces as encountered in condensed phase systems sampled with MD; and that 2) the optimal dampening factor

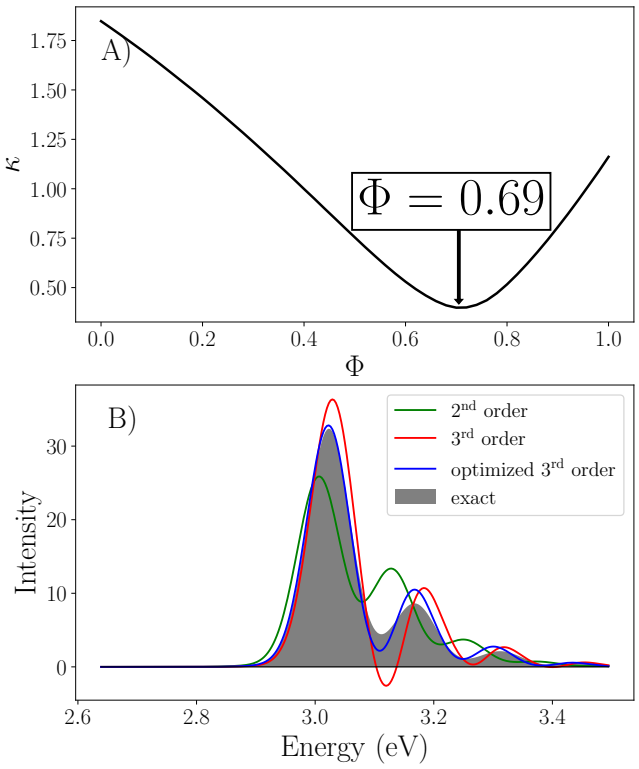

FIG. 2. A) $\kappa(\Phi)$ for an arbitrary GBOM. The optimal prefactor located at the minimum of $\kappa(\Phi)$ ($\Phi = 0.68$) is used to generate an optimized spectrum in B).

Φ can be expressed as a function of the non-Gaussian features of the underlying energy gap fluctuations, such that $\Phi \equiv \Phi\left(\mu_{(3)}, \mu_{(4)}\right)$.

The above points can be justified by the following observations. First, as demonstrated in Fig. 1, the third order cumulant correction does improve the lineshape function in the short timescale limit, but introduces oscillations with amplitudes growing linearly in time, leading to an overestimation of oscillatory terms at longer timescales (see SI Secs. II, III and V). Therefore, a net effect of higher order cumulant terms must be a dampening of the linear divergence observed in the third order cumulant correction. This finding can be confirmed by evaluating higher order cumulants for model systems numerically (see SI Sec. III). Second, the timescale of divergences in the third order cumulant contribution, and thus the importance of higher order cumulant terms, is directly related to the degree of non-Gaussian behavior in the underlying energy gap fluctuations. Thus, we expect Φ to be a function of $\mu_{(3)}$ and $\mu_{(4)}$.

To obtain a functional form of $\Phi(\mu_{(3)}, \mu_{(4)})$ we apply an iterative approach as outlined in the schematic of Fig. 3. A large set of GBOM Hamiltonians is constructed through the random sampling of parameters across a broad, yet realistic, domain relevant to molecular systems (see SI Sec. VI). For a given GBOM, the ideal value for the dampening factor Φ is determined by varying Φ be-

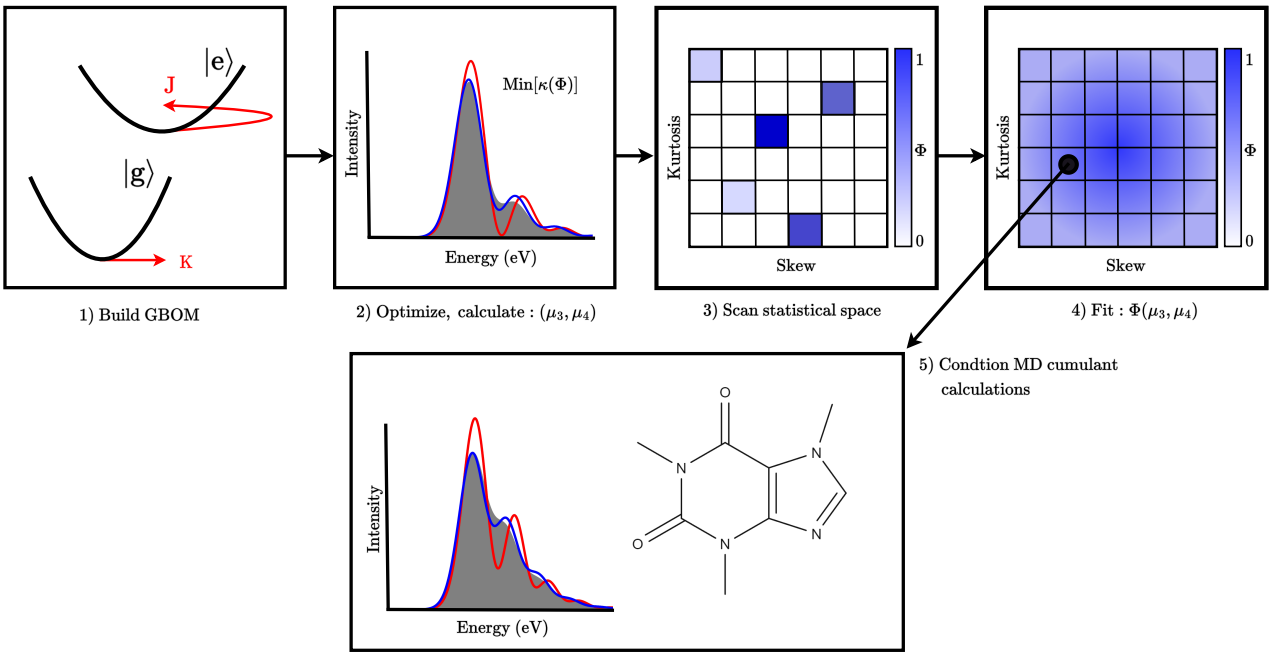

FIG. 3. Schematic for generating predictive prefactor plot. A large number of GBOMs is sampled to produce a predictive grid of Φ as a function of $\mu_{(3)}$ and $\mu_{(4)}$ that may then be applied to molecular dynamics (MD) simulations in realistic condensed phase systems.

tween 0 and 1, and finding the minimum of the resulting convex function $\kappa(\Phi)$, as outlined in Fig. 2. Additionally, the degree of non-Gaussian behavior in the energy gap fluctuations is evaluated by determining skew and excess kurtosis via Eqn. 22, thus obtaining a single point of the function $\Phi(\mu_{(3)}, \mu_{(4)})$ in statistical space. Repeating this process across a large number of unique GBOMs with randomly sampled parameters, we then construct a continuous function $\Phi(\mu_{(3)}, \mu_{(4)})$ by fitting an analytic function to the sampled data. This allows us to predict the ideal dampening factor Φ for arbitrary systems from the measures $\mu_{(3)}$ and $\mu_{(4)}$ of the underlying non-Gaussian energy gap fluctuations alone, with the aim of improving the lineshape in realistic systems sampled from MD. Thus regardless of the model system or realistic condensed phase molecule that generated an energy gap fluctuation distribution, the degree of dampening which would occur from the inclusion of higher order cumulants is now being inferred from quantifiable properties of the underlying distribution. In Sec. IV, we rigorously test the performance of the $\Phi(\mu_{(3)}, \mu_{(4)})$ derived from the GBOM sampling on a range of model systems and real molecules in the condensed phase.

III. COMPUTATIONAL DETAILS

A. Stochastic sampling of GBOMs

To construct a plot of $\Phi\left(\mu_{(3)}, \mu_{(4)}\right)$ as outlined in Sec. II D, $\approx 125,000$ individual GBOM parameterizations (see Appendix A 1) were sampled. The number of modes in the GBOMs was varied systematically from 2 to 50. The ground and excited state frequencies and shift vectors were sampled from uniform distributions over a physically realistic range of molecular vibrations. The n-mode Duschinsky rotation matrix relating ground- to excited state normal modes was constructed by filling the off-diagonal elements of each row with randomly generated values such that their sum equals a predetermined value; this value being sampled from a uniform distribution as well. The coupling matrix was subsequently made unitary through a Gram-Schmidt orthonormalization of column vectors. Full details of the distributions of model parameters and how they relate to realistic molecular systems can be found in SI Sec. VI A.

For all systems sampled, a fixed solvent response was coupled to each GBOM to obtain realistically broadened spectral lineshapes (see Appendix A 3). It is emphasized that the solvent parameters need not be varied to explore the statistical nature of the prefactor, as the solvent modes are assumed to follow Gaussian statistics and are thus described exactly by the second order cumulant approximation. Since in the cumulant approach, no distinction has to be made between whether fluc-

tuations arise from solvent or chromophore degrees of freedom, a stronger or weaker coupling to solvent environment can be interpreted as adding or removing non-interacting Gaussian modes from the underlying "chromophore" GBOM. Therefore, the effect of stronger solvent coupling through the addition of more Gaussian fluctuations is directly incorporated in the GBOM sampling scheme.

After sampling 125,000 GBOMs, evaluating $\mu_{(3)}$ and $\mu_{(4)}$ and computing the ideal prefactor Φ for each GBOM using the metric $\kappa(\Phi)$, resulting data of $\Phi(\mu_{(3)}, \mu_{(4)})$ was then averaged across cells of $\Delta\mu_{(3)} \times \Delta\mu_{(4)}$, with $\Delta\mu_{(3)} = 0.026$ and $\Delta\mu_{(4)} = 0.030$. A cubic bivariate spline was fit through the resulting data to create a continuous function for Φ (see SI Sec. VI B), that can then be used to predict the appropriate correction factor for arbitrary systems. Only cells of $\Delta\mu_{(3)} \times \Delta\mu_{(4)}$ sampled with at least three distinct GBOMs were considered in the spline fit. We emphasize that predicted prefactor values outside of the contour of Fig. 4 are extrapolated from the collected model system data using the spline fitting. While prefactors obtained in this fashion are likely reliable in regions close to the sampled bounds, the third order cumulant correction should be discarded entirely for systems exhibiting extreme values of $|\mu_{(3)}|$ or $|\mu_{(4)}|$, as even the corrected third order cumulant approach likely becomes unreliable due to the missing higher order cumulant contributions. We will explore how well the parameter space explicitly sampled by the 125,000 GBOMs corresponds to skewness and kurtosis values in realistic condensed phase systems in Sec. IV C.

For all GBOMs sampled, the second and third order cumulant lineshapes were evaluated using the analytical expressions for the exact quantum correlation functions, that are generally inaccessible in realistic condensed phase systems. This was done so that the prefactor $\Phi(\mu_{(3)}, \mu_{(4)})$ was independent of errors introduced in the lineshape through constructing cumulants from classical correlation functions using QCFs. We have performed tests (see SI Sec. V A) to confirm that the errors introduced through QCFs are generally small in comparison to the errors introduced by a low-order truncation of the cumulant expansion. Thus the $\Phi(\mu_{(3)}, \mu_{(4)})$ derived for exact quantum correlation functions in the GBOM Hamiltonian is expected to perform well for condensed phase systems sampled in MD, where lineshape functions have to be evaluated from classical correlation functions using QCFs.

B. MD sampling of condensed phase systems

To assess the performance of the proposed approach on realistic systems, we consider a number of molecules where the classical energy gap fluctuations are sampled directly from MD (see SI Sec. VII). A specific focus is the coumarin-153 dye in toluene.

To sample the fluctuations of the coumarin dye, mixed quantum mechanical/molecular mechanical (QM/MM) dynamics[80] of the molecule in a 30 Å solvent sphere in open boundary conditions were performed, with the QM region confined to the chromophore. QM/MM dynamics were run using the interface between Amber[81] and TeraChem,[7,82] and the force field parameters for the Toluene solvent was generated using AmberTools. The QM region was treated using density-funtional theory (DFT) at the CAM-B3LYP[83]/6-31+G*[84] level of theory. A timestep of 0.5 fs was used throughout and the system was kept at 300 K using a Langevin thermostat with a collision frequency of 1 ps^{-1}. A 50 ps pure MM equilibration was carried out, before switching to QM/MM dynamics for 22 ps. The first 2 ps of the QM/MM trajectory were discarded to allow for additional equilibration upon switching the chromophore Hamiltonian from an MM to a QM representation.

Along the QM/MM trajectory, vertical excitation energies were computed in 2 fs intervals using time-dependent density-functional theory (TDDFT) as implemented in the TeraChem code,[85] with the same basis set and functional as for the ground state dynamics. A total of 10,000 individual vertical excitation energies were then used to compute classical correlation functions $C_{cl}^{(2)}$, $C_{cl}^{(3)}$, as well as measures of the non-Gaussian nature of energy gap fluctuations $\mu_{(3)}$ and $\mu_{(4)}$.

IV. RESULTS AND DISCUSSION

A. The dampening factor $\Phi\left(\mu_{(3)}, \mu_{(4)}\right)$ constructed from randomly sampled GBOMs

The spline fit of $\Phi\left(\mu_{(3)}, \mu_{(4)}\right)$ resulting from sampling ≈125,000 GBOMs can be found in Fig. 4, whereas the raw data prior to fitting the spline can be found in SI Sec. VI. Close to the Gaussian limit ($\mu_{(3)} = 0, \mu_{(4)} = 0$), the third order contribution should only provide a small correction to the second order cumulant response function, as all higher order cumulants strictly vanish for Gaussian fluctuations. As such, we expect that the largest prefactor values of $\Phi\left(\mu_{(3)}, \mu_{(4)}\right) \approx 1$ should reside close to the Gaussian limit, as is indeed observed. In this region, the undampened third order cumulant approximation provides the best approximation to the exact spectrum.

Additionally, we observe that neither the statistical region sampled by the GBOMs, nor the spline fit to $\Phi(\mu_{(3)}, \mu_{(4)})$ are symmetric around the origin of $\mu_{(3)} = 0, \mu_{(4)} = 0$. Thus, the volatility of the third order cumulant approximation does not depend solely on the magnitude of non-Gaussian behavior ($|\mu_{(3)}|, |\mu_{(4)}|$), and specifically the sign of skewness of the energy gap fluctuations imparts a substantial effect on the value of the dampening factor Φ. This finding can be rationalized by examining the relative performance of the undampened (Φ=0) third order cumulant approximation against the second order cumulant approximation. We define the relative

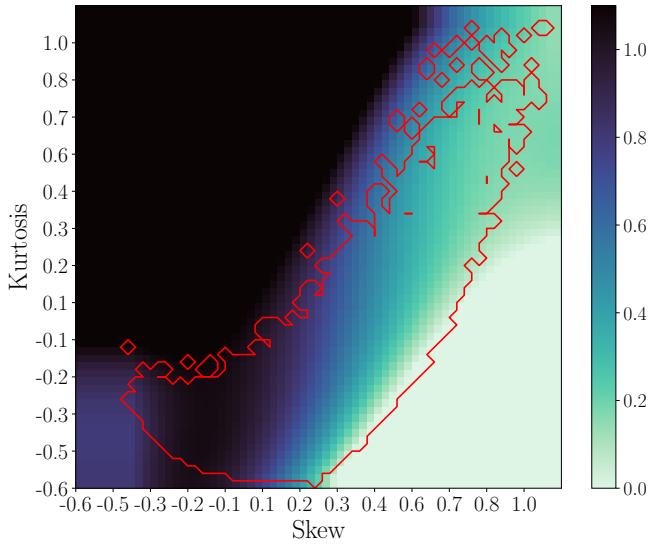

FIG. 4. Spline-fit mapping of $\Phi(\mu_{(3)}, \mu_{(4)})$. Regions within the contour are being interpolated with a cubic bivariate spline fit to the GBOM sampling. Outside of the contour, prefactor values are extrapolated.

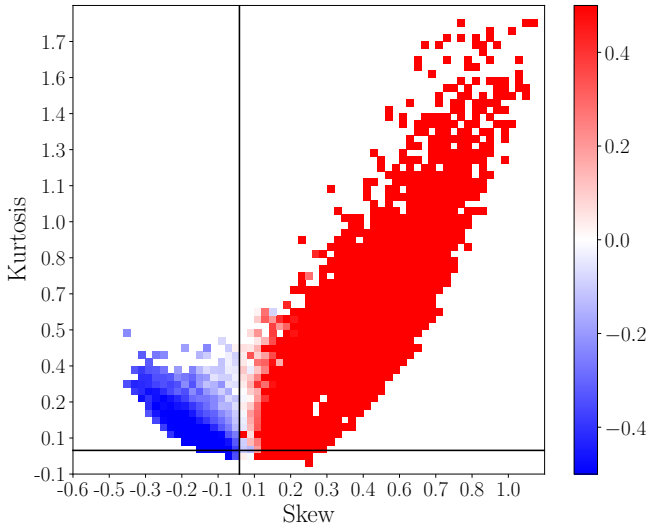

FIG. 5. κ_{rel} of the third order cumulant spectra versus the second order spectra across the range of GBOMs sampled.

metric $\kappa_{\text{rel}} = (\kappa_3 - \kappa_2)/\kappa_2$, where κ_2 and κ_3 correspond to the metric of Eqn. 23 evaluated for the second order and undampened third order cumulant approximation respectively. A plot of κ_{rel} across the range of sampled GBOMs can be found in Fig. 5. As can be seen, systems with energy gap fluctuation statistics presenting with a negative skew always yield a third order cumulant spectrum that outperforms the second order cumulant approximation. It may be demonstrated analytically (see SI Sec. II. A) for a system with uncoupled GBOM modes (corresponding to the Duschinsky rotation matrix being equal to the identity matrix, [J] = [I]), that a negative skew is indicative of a relaxation in vibrational frequency upon excitation ($\omega_e < \omega_g$). This causes a divergence proportional to t^2 in the real part of the third order cumulant lineshape function, leading to a dampening of the overall response function and a physical linear absorption spectrum. Conversely, as $\omega_e > \omega_g$ a positive skew occurs in the simple model system. As ω_e/ω_g becomes increasingly large, the lineshape function diverges proportionally to $-t^2$, causing an unphysical divergent response function. Additionally, oscillatory divergences with linearly growing amplitudes appear more dominantly in the short timescale relevant to linear optical response and these terms can cause unphysical features in the resulting spectral lineshape. Interestingly, our calculations indicate that an analogous statement holds for GBOMs with couplings between individual modes as described by the Duschinsky rotation: A negative skew is indicative of a stable third order cumulant approximation that improves over the second order cumulant lineshape. For positive skewness values, truncating a cumulant expansion at third order almost universally leads to a deterioration of the spectrum in comparison to the second order approximation.

Evaluation of the performance of the prefactor-optimized third order cumulant approach versus the second order cumulant approach with an analogous relative metric κ_{rel} (see Fig. 6 A) reveals a substantial expansion in statistical space where the third order cumulant approximation may be applied to good effect. Specifically, introducing the dampening factor Φ significantly improves the performance for systems with a positive skewness value in the energy gap fluctuations. As is demonstrated in Sec. IV C, this expansion of the third order cumulant method to positive skew, low kurtosis energy gap fluctuations is key for obtaining reliable corrections to the lineshape in real molecular systems in the condensed phase sampled with MD.

We also note that there is a low skewness, high kurtosis region where the optimized third order cumulant method under-preforms the second order cumulant approximation. This behavior may be rationalized in the following way: Within this region, one observes both numerically unstable GBOM parameterizations and GBOMs for which the unaltered third order cumulant and second order cumulant are in close agreement. With respect to the first effect, we observe nonphysical artifacts in this region not only for the third order cumulant approximation, but also with FC and ensemble computational methods which are generally assumed to be stable. This suggests that this region likely corresponds to model parameterizations that lie outside the domain of realistic molecular systems. This observation is reinforced by the fact that none of the real molecular systems sampled with MD investigated in this study (see SI Sec. VII and Sec. IV C) reside within this domain of statistical space. In fact, with all molecules studied for this work we find that their energy gap fluctuations reside in regions where the optimized third order cumulant approximation is predicted to out-

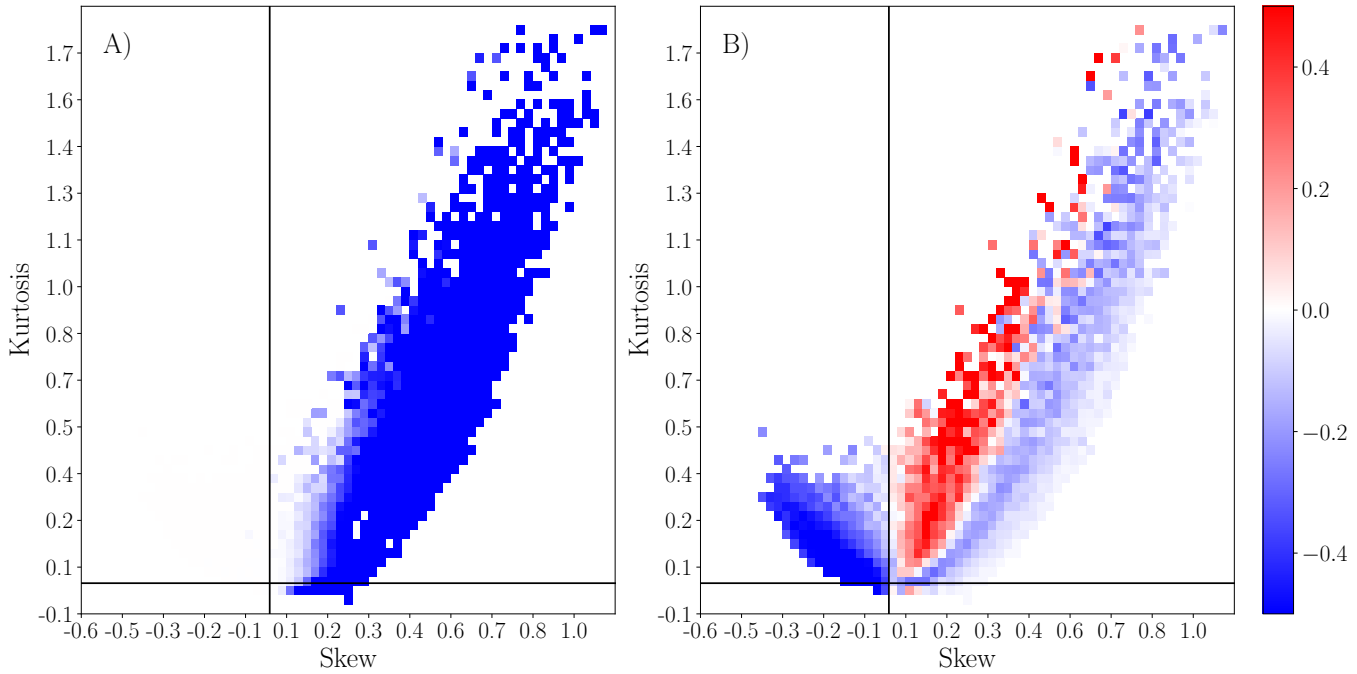

FIG. 6. A) κ_{rel} of the optimized third order cumulant against the unaltered third order cumulant. B) κ_{rel} of the optimized third order cumulant versus the second order cumulant.

preform the second order cumulant approximation.

In Fig. 6 B), we focus on the relative performance between the dampened and the undampened third order cumulant approximation. We note that in the negative skew region, the dampened third order cumulant approximation is indistinguishable from the pure third order cumulant approximation in average performance, again indicating that the pure third order cumulant approximation systematically improves the lineshape in this region. In the positive yet small skew limit, it is generally observed that physically reasonable GBOMs only have a very small third order correction, resulting in very little effect in the overall response function by conditioning the cumulant expansion through the ideal dampening factor Φ. In the larger valued positive skew region, we observe a strong and consistent improvement of the dampened third order cumulant approximation over the pure third order cumulant approximation, with Φ becoming small for large skewness values. We find these results consistent with the principles that led us to propose the functional form of Eqn. 16: As we reach regions with a more strong departure from Gaussian behavior, the size of the third order cumulant contribution must increase. In parallel, the (unaccounted for) higher order cumulant contributions must increase in size as well, requiring a larger degree of dampening.

B. Application to Model Systems

1. The GBOM

Once the function $\Phi(\mu_{(3)}, \mu_{(4)})$ has been parameterized by fitting a bivariate spline to the sampled data points of ≈125,000 GBOMs, the prefactor conditioning of the third order cumulant approximation can now be applied in a predictive manner. To do so, we apply the prefactor method to a set of GBOMs where, rather than selecting the ideal dampening factor from minimizing the metric $\kappa(\Phi)$, we evaluate Φ directly from the statistical properties of the energy gap fluctuations, namely the skew $\mu_{(3)}$ and excess kurtosis $\mu_{(4)}$, without the need to refer to the exact analytical solution of the spectral lineshape. Three example GBOMs can be found in Fig. 7 (information on model parameters is provided in Appendix A 5 b). As can be seen, the fitted spline function for $\Phi(\mu_{(3)}, \mu_{(4)})$ yields an appropriate dampening factor for all GBOM parameterizations in different areas of statistical space. In all cases, the dampened third order cumulant lineshape is much improved over the pure third order cumulant lineshape, removing strong divergences and unphysical negative absorbances in the case of Fig. 7 A) and C). The dampened lineshape also provides a significant improvement over the second order cumulant approximation, both in terms of the position and intensity of vibronic peaks. While Fig. 7 shows three representative GBOM parameterizations, we observe a similar performance over a wide range of model parameterizations studied.

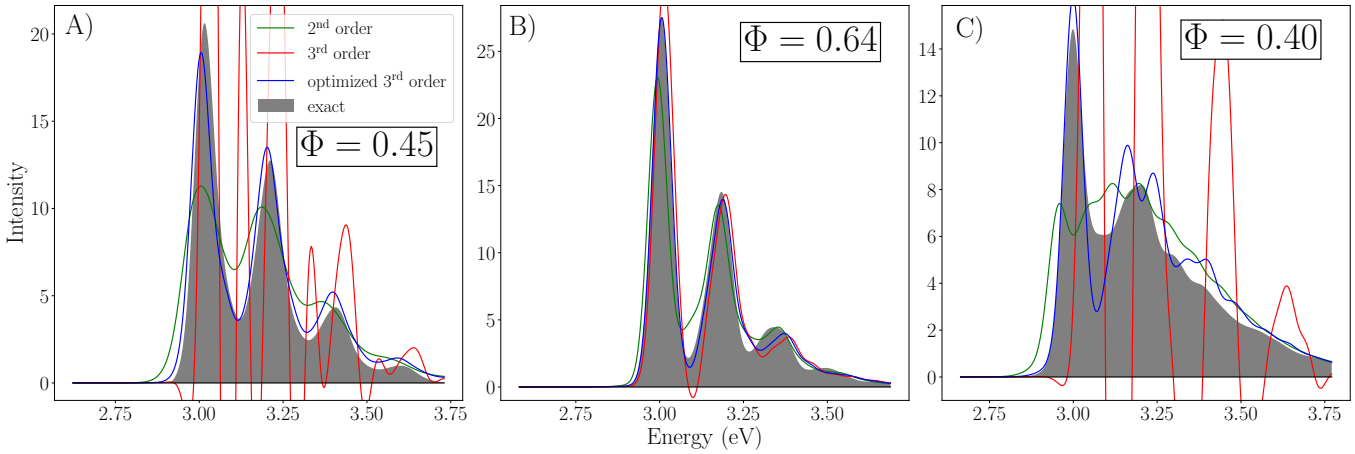

FIG. 7. Three different GBOM Hamiltonian parameterizations sampled across statistical space. A) A 15-mode GBOM ($\mu_{(3)} = 0.26, \mu_{(4)} = 0.15$). B) An 8-mode GBOM ($\mu_{(3)} = 0.11, \mu_{(4)} = 0.04$). C) A 16-mode GBOM ($\mu_{(3)} = 0.25, \mu_{(4)} = 0.4$).

The fact that the correction factor can be applied with good accuracy to the same class of models it was parameterized for is not entirely surprising, but serves at a good empirical demonstration that Φ can be truly represented as a function of the non-Gaussian features of the energy gap fluctuations, rather than explicit parameters of the underlying Hamiltonian. In the context of GBOMs, this implies that the optimal prefactor within a region of non-Gaussian fluctuations obtained from a set of randomly sampled GBOMs can still be effectively applied to a GBOM that may strongly vary in underlying parameters, as long as it shares similar energy gap fluctuation statistics. What remains to be demonstrated is that a dampening factor $\Phi(\mu_{(3)}, \mu_{(4)})$ derived for a simple set of harmonic model can be applied to more general (anharmonic) systems. However, the good performance of the dampening factor shown in Fig. 7 is highly promising, as it suggests that failures of the cumulant approach in correctly capturing Duschinsky mode mixing effects and changes in PES curvature upon excitation can be effectively cured even in complex condensed phase systems, where the FC method cannot be applied.

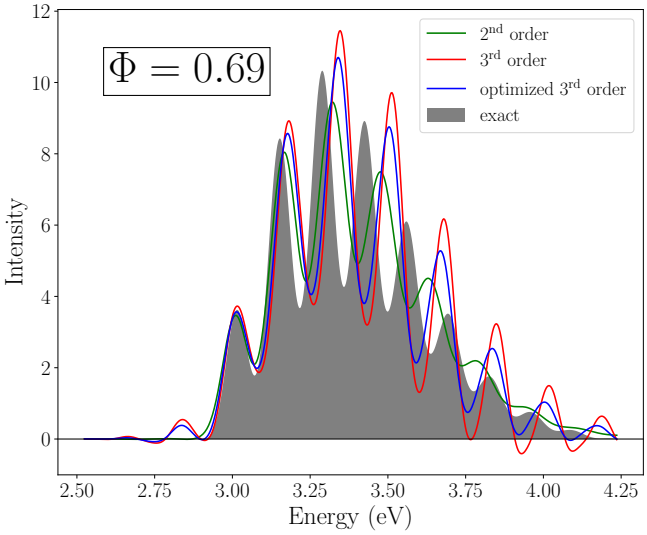

FIG. 8. Exact, second, third and prefactor optimized spectra for a model system containing a single Morse oscillator and a 9 mode GBOM ($\mu_{(3)} = 0.08, \mu_{(4)} = 0.034$).

2. The Morse oscillator

To test the validity of the prefactor mapping on molecular systems which contain strongly anharmonic modes, we construct model systems under the scheme outlined in Appendix A 2, where the chromophore is approximated through a set of harmonic modes described as a GBOM and a single anharmonic mode described as a Morse oscillator. The results of pure and dampened cumulant approximations are once again assessed against the exact spectrum. The results for a specific model parameterization can be found in Fig. 8, with the exact model parameters specified in Appendix A 5 c.

As found in our previous work, both second and third order cumulant expansions struggle to replicate higher order vibronic progressions in the Morse spectrum.[39] However, we find that the predicted prefactor Φ removes nonphysical characteristics of the third order spectrum entirely. While the degree of improvement for this anharmonic model is not appreciable in comparison to the harmonic systems studied in Fig. 7, conditioning through the dampening factor Φ remains an effective way to safeguard against unphysical spectral lineshapes.

Furthermore, the ability to apply a prefactor parameterized for harmonic systems to a system which now contains anharmonic PESs supports the hypothesis that the dampening factor is indeed independent of the underlying Hamiltonian of the physical system, and instead depends on the non-Gaussian features of the energy gap fluctu-

ations alone. This suggests that the fitted dampening factor $\Phi(\mu_{(3)}, \mu_{(4)})$ can be applied to complex condensed phase systems with energy gap fluctuations sampled directly from MD simulations.

C. Molecular Systems sampled from MD

The ultimate objective of the optimization approach developed within this work is to apply the inference gained in model systems to the simulated absorption spectra of realistic molecular systems in the condensed phase, where the exact spectrum can no longer be computed. For these complex systems, the aim of evaluating the parameterized dampening factor $\Phi(\mu_{(3)}, \mu_{(4)})$ is to both improve the computed lineshapes and reduce computational cost.

MD-type calculations required to sample energy gap fluctuations in the condensed phase to construct cumulant spectra are computationally expensive[4,55,58] compared to calculations on simple model systems (generally requiring the computation of tens of thousands vertical excitation energies along the trajectories). Additionally, computing reliable third order cumulant corrections requires significantly more data than second order cumulant spectra, as two-time correlation functions of the energy gap have to be converged.[57] Thus, there is merit in being able to predict the overall stability of the third order cumulant approximation based on statistics of energy gap fluctuations alone, which can be computed cheaply from preliminary data sets. If the skew and excess kurtosis of energy gap fluctuations of a molecule fall into a region of low confidence in Fig. 4, it can be concluded that expending the extra computational cost to construct third order cumulant corrections is unlikely to yield improved spectra. In other systems, specifically those with moderate positive skew in the energy gap fluctuations, we expect a prefactor-conditioned MD-based third order cumulant approach to yield improved lineshapes, both by removing unphysical negative spectral features and by improving the underlying vibronic fine-structure of the transition.

To examine the usefulness of the dampening factor $\Phi(\mu_{(3)}, \mu_{(4)})$ in realistic systems, we first demonstrate that typical condensed phase systems sampled from MD fall into the statistical range of non-Gaussian energy gap fluctuations sampled by the randomly parameterized GBOMs used in this work. Such an analysis is provided for nine selected molecules in SI Sec. VII. Both isolated molecules in vacuum and chromophores in condensed phase environments are included in the data set. For all nine molecules, it is found that the non-Gaussian features of the energy fluctuations, $\mu_{(3)}$ and $\mu_{(4)}$, fall into the region sampled by the GBOMs. Additionally, they all fall into a region where the optimized third order cumulant approximation is predicted to outperform the second order cumulant approximation, thus indicating that the dampening factor Φ can likely be used to improve spectral lineshapes in these realistic systems. Here, we focus on a single selected system in more detail, namely coumarin-135 in toluene, due to the fact that vibronic peaks are well resolved in the lineshape, and an experimental spectrum is readily available.[86]

We emphasize that to produce what one may define as a sufficiently accurate spectrum for a molecule in the condensed phase using the MD-based cumulant method involves additional challenges when compared to the study of simple model systems where the exact spectrum can be readily computed. Vibronic lineshapes are often found to differ considerably depending on the level of theory used for modeling the ground- and excited state potential energy surfaces with TDDFT.[87,88] Thus, any discrepancies with respect to the experimental spectrum cannot be easily ascribed to errors introduced in the low order cumulant expansion alone. For this reason, we more broadly focus on observed changes in the vibronic fine structure under the different cumulant approximations, rather than direct a quantitative comparison to the experimental lineshape.

1. Coumarin-153

Experimental[86] and simulated MD-based cumulant spectra for coumarin-153 can be found in Fig. 9. We note that available experimental spectrum in Ref. 86 uses hexane as the non-polar solvent, rather than toluene as used in our calculations. Due to the fact that both are non-polar solvents with similar dielectric constants, we expect the spectra in the two solvents to match each other closely. Additionally, we note that the weakly interacting nature of the solvent means that this system is likely well-described by more commonly used and computationally affordable Franck-Condon type approaches. While strongly interacting solvents provide a more suitable application for the cumulant method, as direct solute-solvent interactions and slow collective chromophore-environment motions cannot be straightforwardly included in the Franck-Condon method, the well-defined vibronic finestructure of coumarin-153 in weakly interacting solvents allows for a more detailed analysis of the performance of the corrected cumulant approach and is therefore presented here as a test case.

The results in Fig. 9 indicate that the pure third order cumulant spectrum for this mildly non-Gaussian system already exhibits negative contributions to the spectral lineshape. For the underlying non-Gaussian statistics of the energy gap, a prefactor of $\Phi = 0.43$ is predicted for this system. When applying the dampening factor to the spectral lineshape we indeed successfully remove the unphysical negative absorbance from the onset of the spectrum at 2.8 eV.

Focusing on the spectral shape, we note that the pure third order cumulant approach predicts a much more pronounced vibronic fine structure than the second order cumulant approach. Additionally, while in the second order

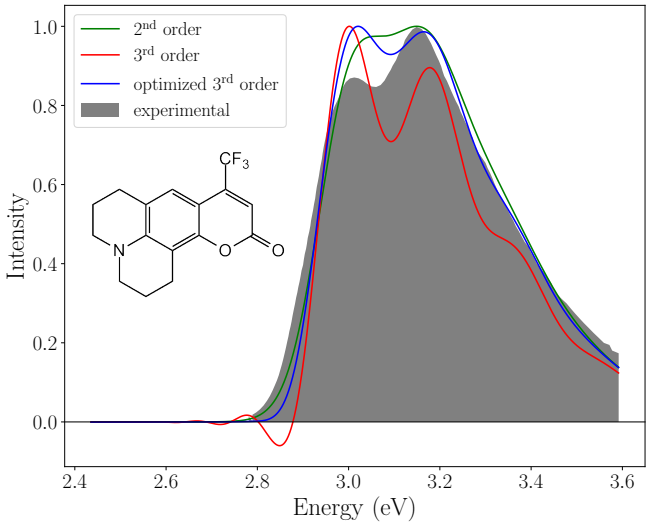

FIG. 9. Experimental,[86] second order, third order and optimized third order spectrum of Coumarin-153 in toluene (for simulated spectra) and hexane (experimental spectrum). The optimal prefactor was determined to be $\Phi = 0.43$ based on a skew of 0.22 and an excess kurtosis of 0.05.

cumulant approximation and in the experimental spectrum, the second vibronic peak has the highest intensity, the third order cumulant approach erroneously predicts the first vibronic peak to be more bright. The dampened third order cumulant method produces a spectrum with a more resolved vibronic fine-structure and improved peak separation compared to the second order cumulant approach, but still predicts a high intensity for the second vibronic peak. In general, the dampened spectrum is in good agreement with the experimental spectrum, with only a minor overestimation of the first vibronic peak. The origin of remaining discrepancy with the experimental spectrum is unclear, but can potentially be ascribed to inaccuracies in the (TD)DFT description of the system Hamiltonian, rather than the truncation of the cumulant expansion.

We take the fact that the prefactor $\Phi(\mu_{(3)}, \mu_{(4)})$ successfully removes non-physical features from the third order cumulant spectrum in a realistic condensed phase system as an additional proof of the wide applicability of the dampened cumulant response developed in this work (see Eqn. 16). No details of the underlying system Hamiltonian are needed to obtain the ideal dampening factor Φ, and only non-Gaussian features of the energy gap fluctuations determine this quantity. We have demonstrated that constructing $\Phi(\mu_{(3)}, \mu_{(4)})$ from a stochastic sampling of simple, exactly solvable model systems provides a recipe for improving cumulant spectra truncated at low order in complex condensed phase systems.

V. CONCLUSION

In this work, we have outlined an approach to account for third order corrections to the linear absorption spectrum computed in the widely-used cumulant framework. The approach takes into account the effect of moderately non-Gaussian energy gap fluctuations without exhibiting unphysical divergences and regions of negative absorbance in the resulting spectra. The method promises to yield more accurate and robust linear spectra in model systems and chromophores embedded in complex condensed phase environments alike.

The key insight in this work consists of introducing a dampening factor $\Phi(\mu_{(3)}, \mu_{(4)})$ that is applied to the bare third order cumulant correction, where Φ is taken to be an explicit function of the skew and excess kurtosis, measures of the non-Gaussian nature of the energy gap fluctuations. We have rationalized the functional form of this correction factor by noting that the third order cumulant term improves the spectral lineshape in the short timescale, but exhibits divergences at longer timescales. These divergences would be cancelled by higher order cumulant contributions that have to be neglected in any practical application to condensed phase systems. Thus the factor Φ approximately accounts for the collective dampening contribution to the lineshape of higher order cumulants.

We have shown that the functional form of $\Phi(\mu_{(3)}, \mu_{(4)})$ can be parameterized by stochastically sampling model parameters of the GBOM, a harmonic model Hamiltonian that is widely applied to the prediction optical spectra of semi-rigid molecules. By constructing Φ for a model system where the exact spectrum cam be computed analytically, we were able to parameterize an ideal dampening factor as a function of the non-Gaussian fluctuations only, rather than the parameters of the underlying system Hamiltonian. With this predictive plot in place, we were able to find optimal dampening factors Φ in MD-type simulations of molecular systems in the condensed phase.

We have demonstrated that the parameterized dampening factor rigorously removes unphysical and divergent lineshapes in the third order cumulant approximation, both for model systems and for condensed phase systems sampled with MD. In the GBOM Hamiltonian, we have also shown that the approach improves the agreement with the exact spectrum, both compared to the second order and the pure third order cumulant approach, for a wide range of model system parameterizations. For the realistic condensed phase system of coumarin-135 in toluene, our method yields a final spectrum in excellent agreement with experiment. The results presented indicate that the approach outlined in this work provides an efficient, computationally affordable pathway for correcting the main shortcomings of the second order cumulant approach in condensed phase systems, namely the inability to account for non-Gaussian fluctuations introduced by anharmonic effects, Duschinsky mode mixing effects

and changes in the PES curvature upon excitation. Since our parameterization of Φ only relies on the skewness and kurtosis of the underlying energy gap fluctuations, we expect the method to be widely applicable to condensed phase systems sampled with MD. The dampened third order cumulant approach developed in this work has been implemented in the open-source software package MolSpeckPy developed within our group.[89]

Appendix A: Model systems

In this work, we turn to two exactly solvable model systems. The Generalized Brownian Oscillator Model (GBOM) is a harmonic system that will be used to infer the functional form of the dampening factor $\Phi\left(\mu_{(3)}, \mu_{(4)}\right)$, whereas the 1D Morse oscillator is invoked to explore the effectiveness of the dampening factor in systems that go beyond the harmonic approximation.

1. Generalized Brownian oscillator model (GBOM)

The GBOM is a convenient, numerically robust model in which the second order cumulant approximation is no longer exact, but in which the exact second and third order cumulants can be evaluated analytically (see SI Sec. II and Refs. 39 and 60). Furthermore, the exact response function (Eqn. 5) can be computed analytically,[36,37] enabling us to examine in detail the errors introduced by truncating the cumulant expansion at some finite order.

For an n-mode GBOM, we define the following ground- and excited state nuclear Hamiltonians:

$$\hat{\mathrm{H}}_g(\hat{\mathbf{q}}, \hat{\mathbf{p}}) = \frac{1}{2}\sum_i^n \left[\hat{\mathrm{p}}_i^2 + \omega_{g,i}^2 \hat{\mathrm{q}}_i^2\right] \tag{A1}$$

$$\hat{\mathrm{H}}_e(\hat{\mathbf{r}}, \hat{\boldsymbol{\pi}}) = \frac{1}{2}\sum_i^n \left[\hat{\pi}_i^2 + \omega_{e,i}^2 \hat{\mathrm{r}}_i^2\right] + \Delta, \tag{A2}$$

where $\{\omega_{g,i}\}$ and $\{\omega_{e,i}\}$ are the set of ground- and excited state vibrational frequencies, Δ is the adiabatic energy gap between the two electronic surfaces and $\{\hat{\mathrm{q}}_i\}$ and $\{\hat{\mathrm{r}}_i\}$ are ground and excited state normal mode coordinates respectively. The coordinates are related to each other by a linear transformation in terms of a shift-vector ($\mathbf{k}$) between the ground and excited state minima, and the mode-mixing Duschinsky rotation matrix ($\mathbf{J}$):[38]

$$\hat{\mathrm{q}}_i = \sum_m \mathrm{J}_{im}\hat{\mathrm{r}}_m - \mathrm{k}_i. \tag{A3}$$

Both the Duschinsky mode mixing and the mismatch between the ground- and excited state normal mode frequencies cause non-Gaussian energy gap fluctuations in this model Hamiltonian.[39]

2. Morse oscillator model

A main shortcoming of the GBOM as applied to realistic systems is that it does not exhibit any anharmonicity in the potential energy surface. To probe the influence of anharmonic effects on computed lineshapes, we turn to the 1D Morse oscillator with the system Hamiltonian taking the following form:

$$\hat{\mathrm{H}}_g(\hat{\mathrm{q}}, \hat{\mathrm{p}}) = \frac{\hat{\mathrm{p}}^2}{2} + \mathrm{D}_g[1 - e^{-\alpha_g \hat{\mathrm{q}}}]^2 \tag{A4}$$

$$\hat{\mathrm{H}}_e(\hat{\mathrm{q}}, \hat{\mathrm{p}}) = \frac{\hat{\mathrm{p}}^2}{2} + \mathrm{D}_e[1 - e^{-\alpha_e(\hat{\mathrm{q}}-\mathrm{k})}]^2 + \Delta. \tag{A5}$$

Here, D is the well depth of the potential, k is the displacement between the ground and excited potential minima, and α_g and α_e are the anharmonicity parameters for the ground- and excited state. Analytical expressions for the Morse oscillator wavefunctions exist,[90] such that the exact response function for this Hamiltonian can be constructed directly through Eqn. 5. Additionally, cumulants based on the exact quantum correlation function can also be constructed numerically.[39]

The exact analytical expressions for the Morse oscillator wavefunctions are numerically unstable, meaning that the parameters this model have to be chosen somewhat carefully to guarantee well-defined results. As such, it less apt for the high-throughput screening of parameter space desired in this study. Instead, we construct a few selected Morse oscillator parameterizations inspired by realistic molecular systems to test whether the conclusions obtained for the GBOM Hamiltonian can be carried over to the anharmonic case. In these model calculations, we combine the Morse oscillator with a GBOM to model a chromophore with a few harmonic and a single anharmonic mode. The anharmonic mode is taken to be decoupled from the harmonic ones, such that the total response function can be written as:

$$\chi(\mathrm{t}) = \chi_{\mathrm{Morse}}(\mathrm{t})\chi_{\mathrm{GBOM}}(\mathrm{t}), \tag{A6}$$

where each response contribution may be computed through FC, ensemble and cumulant-type approximations.

3. Solvent model

In model systems with a finite number of modes, solvent effects must be reintroduced to produce realistically broadened spectra. Here, solvent effects are accounted for through a continuous set of low frequency Brownian oscillator modes described by a spectral density of the Debye form:

$$J_{\mathrm{env}}(\omega) = 2\lambda_{\mathrm{env}}\omega_c \frac{\omega}{\omega_c^2 + \omega^2}, \tag{A7}$$

where λ_{env} is the solvent reorganization energy quantifying the strength of the solvent coupling and ω_c is a

cutoff frequency determining the relaxation timescale of the bath.[39] Assuming the solvent modes are fully separable from the modes describing the chromophore degrees of freedom, solvent-driven broadening effects on the lineshape can subsequently be described through a separable second order cumulant contribution:

$$\chi_{\mathrm{env}}(\mathrm{t}) = e^{-\mathrm{g}_2^{\mathrm{env}}[J_{\mathrm{env}}(\omega)](\mathrm{t})} \tag{A8}$$
$$\chi_{\mathrm{tot}}(\mathrm{t}) = \chi_{\mathrm{chromophore}}(\mathrm{t})\chi_{\mathrm{env}}(\mathrm{t}) \tag{A9}$$

In this work, the solvent model parameters $\omega_c = 22\ \mathrm{cm}^{-1}$ and $\lambda_{\mathrm{env}} = 0.0009$ Ha are used throughout, representing weak solute-solvent interactions typically found in non-polar solvents in realistic systems.

4. The infinite order lineshape function for a GBOM Hamiltonian

In practice, the exact response function Eqn. 5 for a GBOM Hamiltonian is not resloved directly through a sum over states method. Instead, a path integral approach beginning with Fermi's golden rule for linear response is applied to analytically evaluate Eqn. 5 in the time-domain.[37] The newly obtained form of Eqn. 5 allows one to find $\mathrm{g}_\infty(\mathrm{t})$; the exact lineshape function for a GBOM Hamiltonian:

$$\begin{aligned}\mathrm{g}_\infty(\mathrm{t}) = -\ln\left(\sqrt{\det(\mathbf{R})}\right) - \mathrm{i}(\lambda_0+\xi)\mathrm{t} \\ -i\mathbf{k}^{\mathrm{T}}\mathbf{C}\mathbf{k} + \frac{\mathrm{i}}{2}\mathbf{E}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{E}.\end{aligned} \tag{A10}$$

Where $\mathbf{k}$ is the shift-vector in Eqn. A3. The definition of λ_0 and ξ for the GBOM can be found in SI Sec. I A. To compute the matrices $\mathbf{R}$, $\mathbf{C}$, $\mathbf{D}$ and $\mathbf{E}$, we follow the approach outlined by de Souza and coworkers[37] and define the following auxiliary matrices $\mathbf{a}, \overline{\mathbf{a}}, \mathbf{A}, \mathbf{b}, \overline{\mathbf{b}}, \mathbf{B}, \mathbf{P}$:

$$\mathrm{a}_{jk} = \delta_{jk}\omega_{g,j}(\sin(\omega_{g,j}(-\mathrm{t}-\mathrm{i}\beta))^{-1} \tag{A11}$$
$$\overline{\mathrm{a}}_{jk} = \delta_{jk}\omega_{e,j}(\sin(\omega_{e,j}(\mathrm{t}))^{-1} \tag{A12}$$
$$\mathbf{A} = \overline{\mathbf{a}} + \mathbf{J}^T\mathbf{a}\mathbf{J} \tag{A13}$$
$$\mathrm{b}_{jk} = \delta_{jk}\omega_{g,j}(\tan(\omega_{g,j}(-\mathrm{t}-\mathrm{i}\beta))^{-1} \tag{A14}$$
$$\overline{\mathrm{b}}_{jk} = \delta_{jk}\omega_{e,j}(\tan(\omega_{e,j}(\mathrm{t}))^{-1} \tag{A15}$$
$$\mathbf{B} = \overline{\mathbf{b}} + \mathbf{J}^T\mathbf{b}\mathbf{J} \tag{A16}$$
$$\mathrm{P}_{jk} = \delta_{jk}2\sinh\left(\frac{\omega_{g,j}\beta}{2}\right). \tag{A17}$$

The matrices of Eqn. A10 can then be defined as:

$$\mathbf{R} = \mathbf{a}\overline{\mathbf{a}}\left[\mathbf{B}(\mathbf{B}-\mathbf{A}\mathbf{B}^{-1}\mathbf{A})\right]^{-1}\mathbf{P}^2 \tag{A18}$$
$$\mathbf{C} = \mathbf{b}-\mathbf{a} \tag{A19}$$
$$\mathbf{E} = \begin{bmatrix}\mathbf{k}^{\mathrm{T}}\mathbf{C}\mathbf{J}\\ \mathbf{k}^{\mathrm{T}}\mathbf{C}\mathbf{J}\end{bmatrix} \tag{A20}$$
$$\mathbf{D} = \begin{bmatrix}\mathrm{B} & -\mathrm{A}\\ -\mathrm{A} & \mathrm{B}\end{bmatrix} \tag{A21}$$

Here, $\mathbf{J}$ is again the Duschinsky rotation matrix corresponding to the ground- and excited state normal mode coordinate transformation (Eqn. A3).

5. Parameters of Model Systems

a. Parameters of Figs. 1 and 2

The same two-mode GBOM was presented in Figs. 1 and 2. The following parameters were selected:

$$\begin{aligned}\boldsymbol{\omega}_g &= [0.0045, 0.0045]\\ \boldsymbol{\omega}_e &= [0.0053, 0.0055]\\ \mathbf{k} &= [7, 8]\\ \theta &= 0.25\\ \mathbf{J}(\theta) &= \begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix}.\end{aligned}$$

Here, frequencies are in Hartree atomic units and elements of the shift vector $\mathbf{k}$ are in units of $\sqrt{\omega_{g,i}}^{-1}$.

b. Parameters of Fig. 7

The GBOMs presented in Fig. 7 contain nontrivial rotation matrices which are large enough to be cumbersome to report here. The frequencies, shift vectors and rotation matrices of the examples presented within Fig. 7 may be found in the Zenodo repository listed in the Data Availability section. We note that these particular files need to be executed in a default version of the Spectroscopy Python Code developed in our group.[89]

c. Parameters of Fig. 8

For Fig. 8, the parameters of the GBOM coupled to the Morse potential mode are as follows:

$$\begin{aligned}\boldsymbol{\omega}_g &= [0.0045, 0.0045, 0.005, 0.006,\\ &\quad 0.0057, 0.005, 0.0042, 0.004, 0.003] \quad \text{(A22)}\\ \boldsymbol{\omega}_e &= [0.0045, 0.0045, 0.005, 0.0058,\\ &\quad 0.0057, 0.0051, 0.0042, 0.004, 0.003] \quad \text{(A23)}\\ \mathbf{k} &= [5, 4, 10, 6, 8, 0, 5, 6, 0]\\ \theta &= 0\\ \mathbf{J}(\theta) &= [\mathbf{I}].\end{aligned}$$

Here, all units follow the conventions of Appendix A 5 a. The 1D Morse oscillator added to this GBOM has the following parameters:

$$\begin{aligned}\mathrm{D}_g &= 0.48\\ \mathrm{D}_e &= 0.2\\ \alpha_g &= 0.7\\ \alpha_e &= 0.9\\ \mathrm{k} &= 0.2,\end{aligned}$$

where well depths ($D_{g,e}$) are in units of Hartree energy, anharmonic factors ($\alpha_{g,e}$) are in units of a_0^{-1} and the displacement (k) is in units of a_0, where a_0 is the Bohr radius.

II. SUPPLEMENTARY MATERIAL

See the supplementary material for the analytic expressions of the second and third order cumulants of a GBOM Hamiltonian, an analysis of their asymptotic behavior and an inspection of higher order cumulants for a number of model systems. Stochastic sampling techniques used to construct the prefactor plot and an analysis of the statistical properties of simulated molecular systems are provided as well.

ACKNOWLEDGMENTS

T. J. Z. acknowledges startup funding provided by Oregon State University. The authors thank Zephyr Solabella for providing the raw MD-based cumulant data for some of the molecules described in SI Sec. VII.

III. DATA AVAILABILITY

Raw data for the GBOM scan conducted in this work as well as the code used to do so may be obtained under the following persistent DOI: https://doi.org/10.5281/zenodo.8361624.[91]

[1]D. Jacquemin, B. Mennucci, and C. Adamo, "Excited-state calculations with td-dft: from benchmarks to simulations in complex environments," Phys. Chem. Chem. Phys. **13**, 16987–16998 (2011).
[2]F. Santoro and D. Jaquemin, "Going beyond the vertical approximation with time-dependent density functional theory," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **6**, 460–486 (2016).
[3]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).
[4]T. J. Zuehlsdorff and C. M. Isborn, "Modeling absorption spectra of molecules in solution," Int. J. Quantum Chem. **119** (2019).
[5]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).
[6]M. Bondanza, M. Nottoli, L. Cupellini, F. Lipparini, and B. Mennucci, "Polarizable embedding qm/mm: the future gold standard for complex (bio)systems?" Phys. Chem. Chem. Phys. **22**, 14433–14448 (2020).
[7]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).
[8]D. J. Cole, A. W. Chin, N. D. M. Hine, P. D. Haynes, and M. C. Payne, "Toward ab initio optical spectroscopy of the fenna-matthews-olson complex," J. Phys. Chem. Lett. **4**, 4206–4212 (2013).
[9]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).
[10]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).
[11]M. Nottoli, M. Bondanza, P. Mazzeo, L. Cupellini, C. Curutchet, D. Loco, L. Lagardère, J.-P. Piquemal, B. Mennucci, and F. Lipparini, "Qm/amoeba description of properties and dynamics of embedded molecules," WIREs Computational Molecular Science **n/a**, e1674.
[12]M. L. Chaillet, F. Lengauer, J. Adolphs, F. Müh, A. S. Fokas, D. J. Cole, A. W. Chin, and T. Renger, "Static disorder in excitation energies of the fenna–matthews–olson protein: Structure-based theory meets experiment," J. Phys. Chem. Lett. **11**, 10306–10314 (2020).
[13]E. Cignoni, V. Slama, L. Cupellini, and B. Mennucci, "The atomistic modeling of light-harvesting complexes from the physical models to the computational protocol," J. Chem. Phys **156**, 120901 (2022).
[14]M. Wendling, M. A. Przyjalgowski, D. Gülen, S. I. E. Vulto, T. J. Aartsma, and R. van Gondelle, "The quantitative relationship between structure and polarized spectroscopy in the FMO complex of *Prosthecochloris aestuarii*: refining experiments and simulations," Photosynth. Res. **71**, 99–123 (2002).
[15]J. Adolphs and T. Renger, "How Proteins Trigger Excitation Energy Transfer in the FMO Complex of Green Sulfur Bacteria," Biophys. J. **91**, 2778–2797 (2006).
[16]C. Olbrich, J. Strümpfer, K. Schulten, and U. Kleinekathöfer, "Theory and Simulation of the Environmental Effects on FMO Electronic Transitions," J. Phys. Chem. Lett. **14**, 1771–1776 (2011).
[17]C. Kreisbeck and T. Kramer, "Long-Lived Electronic Coherence in Dissipative Exciton Dynamics of Light-Harvesting Complexes," J. Phys. Chem. Lett. **3**, 2828–2833 (2012).
[18]S. Shim, P. Rebentrost, S. Valleau, and A. Aspuru-Guzik, "Atomistic Study of the Long-Lived Quantum Coherences in the Fenna-Matthews-Olson Complex," Biophys. J. **102**, 649–660 (2012).
[19]L. Viani, M. Corbella, C. Curutchet, E. J. O'Reilly, A. Olaya-Castro, and B. Mennucci, "Molecular basis of the exciton–phonon interactions in the PE545 light-harvesting complex," Phys. Chem. Chem. Phys. **16**, 16302 (2014).
[20]S. Chandrasekaran, M. Aghtar, S. Valleau, A. Aspuru-Guzik, and U. Kleinekathöfer, "Influence of Force Fields and Qauntum Chemistry Approach on Spectral Densities of BChl *a* in Solution and in FMO Proteins," J. Phys. Chem. B **119**, 9995–10004 (2015).
[21]M. K. Lee and D. F. Coker, "Modeling Electronic-Nuclear Interactions for Excitation Energy Transfer Processes in Light-Harvesting Complexes," J. Phys. Chem. Lett. **7**, 3171–3178 (2016).
[22]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).
[23]M. K. Lee, K. B. Bravaya, and D. F. Coker, "First-Principles Models for Biological Light-Harvesting: Phycobiliprotein Complexes from Cryptophyte Algae," J. Am. Chem. Soc. **139**, 7803–1814 (2017).
[24]S. M. Blau, I. G. Bennett, C. Kreisbeck, G. D. Scholes, and A. Aspuru-Guzik, "Local protein solvation drives direct down-conversion in phycobiliprotein PC645 via incoherent vibronic transport," Proc. Natl. Acad. Sci. U.S.A. **115**, E3342–E3350 (2018).
[25]M. I. Mallus, Y. Shakya, J. D. Prajapati, and U. Kleinekathöfer, "Environmental effects on the dynamics in the light-harvesting complexes LH2 and LH3 based on molecular simulations," Chem. Phys. **515**, 141–151 (2018).

[26]S. J. Jang and B. Mennucci, "Delocalized excitons in natural light-harvesting complexes," Rev. Mod. Phys. **90**, 035003 (2018).
[27]F. Segatta, D. A. Ruiz, F. Aleotti, M. Yaghoubi, S. Mukamel, M. Garavelli, F. Santoro, and A. Nenov, "Nonlinear molecular electronic spectroscopy via mctdh quantum dynamics: From exact to approximate expressions," J. Chem. Theory Comput. **19**, 2075–2091 (2023).
[28]A. Warshel and M. Karplus, "Vibrational structure of electronic transitions in conjugated molecules," Chem. Phys. Lett. **17**, 7–14 (1972).
[29]L. S. Cederbaum and W. Domcke, "A many-body approach to the vibrational structure in molecular electronic spectra. I. Theory," J. Chem. Phys. **64**, 603 (1976).
[30]E. V. Doktorov, I. A. Malkin, and V. I. Manko, "Dynamical symmetry of vibronic transitions in polyatomic molecules and the Franck-Condon principle," J. Mol. Spectrosc. **64**, 302–326 (1977).
[31]K.-M. Chen and C.-C. Pei, "Exact evaluation of two-dimensional Franck–Condon integrals under the Duschinksy mixing effect," Chem. Phys. Lett. **165**, 523–527 (1990).
[32]M. Roche, "On the polyatomic franck-condon factors," Chem. Phys. Lett. **168**, 556–558 (1990).
[33]P. T. Ruhrhoff, "Recursion relations for multi-dimensional Franck-Condon overlap integrals," Chem. Phys. **186**, 335–374 (1994).
[34]M. Dierksen and S. Grimme, "An efficient approach for the calculation of franck–condon integrals of large molecules," The Journal of Chemical Physics **122** (2005).
[35]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).
[36]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).
[37]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).
[38]F. Duschinsky, "Zur Deutung der Elektronenspektren Mehratomiger Moleküle," Acta Physicochim. URSS. **7**, 551 (1937).
[39]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).
[40]F. J. Avila Ferrer, J. Cerezo, J. Soto, R. Improta, and F. Santoro, "First-principle computation of absorption and fluorescence spectra in solution accounting for vibronic structure, temperature effects and solvent inhomogenous broadening," Comput. Theor. Chem. **1040-1041**, 328–337 (2014).
[41]R. Improta, F. J. Avila Ferrer, E. Standardo, and F. Santoro, "Quantum-classical calculation of the absorption and emission spectral shapes of oligothiophenes at low and room temperature by first-principle calculations," Chem. Phys. Chem. **15**, 3320–3333 (2014).
[42]A. Baiardi, J. Bloino, and V. Barone, "General formulation of vibronic spectroscopy in internal coordinates," J. Chem. Phys. **144**, 084114 (2016).
[43]J. Cerezo and F. Santoro, "Revisiting vertical models to simulate the line shape of electronic spectra adopting cartesian and internal coordinates," J. Chem. Theory Comput. **12**, 4970–4985 (2016).
[44]J. Cerezo, D. Aranda, F. J. A. Ferrer, G. Prampolini, G. Mazzeo, G. Longhi, S. Abbate, and F. Santoro, "Toward a general mixed quantum/classical method for the calculation of the vibronic ECD of a flexible dye molecule with different stable conformers: Revisiting the case of 2,2,2-trifluoro-anthrylethanol," Chirality **30**, 730–743 (2018).
[45]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 (2019).
[46]R. Cammi, S. Corni, B. Mennucci, and J. Tomasi, "Electronic excitation energies of molecules in solution: State specific and linear response methods for nonequilibrium continuum solvation models," J. Chem. Phys. **122**, 104513 (2005).
[47]B. Mennucci, "Polarizable continuum model," Wiley Interdisciplinary Reviews: Computational Molecular Science **2**, 386–404 (2012).
[48]F. Lipparini and B. Mennucci, "Perspective: Polarizable continuum models for quantum-mechanical descriptions," J. Chem. Phys. **144**, 160901 (2016).
[49]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).
[50]E. F. Petrusevich, M. H. E. Bousquet, B. Ośmiałowski, D. Jacquemin, J. M. Luis, and R. Zaleśny, "Cost-effective simulations of vibrationally-resolved absorption spectra of fluorophores with machine-learning-based inhomogeneous broadening," J. Chem. Theory Comput. **19**, 2304–2315 (2023).
[51]J. Cerezo, C. García-Iriepa, F. Santoro, I. Navizet, and G. Prampolini, "Unraveling the contributions to the spectral shape of flexible dyes in solution: insights on the absorption spectrum of an oxyluciferin analogue," Phys. Chem. Chem. Phys. **25**, 5007–5020 (2023).
[52]S. Mukamel, *Principles of Nonlinear Optical Spectroscopy* (Oxford University Press, New York, 1995).
[53]S. Mukamel, "Fluorescence and absorption of large anharmonic molecules. Spectroscopy without eigenstates," J. Phys. Chem. **89**, 1077 (1985).
[54]S. Mukamel and D. Abramavicius, "Many-Body Approaches for Simulating Coherent Nonlinear Spectroscopies for Electronic and Vibrational Excitons," Chem. Rev. **104**, 2073–2098 (2004).
[55]M. C. Zwier, J. M. Shorb, and B. P. Krueger, "Hybrid Molecular Dynamics-Quantum Mechanics Simulations of Solute Spectral Properties in the Condensed Phase: Evaluation of Simulation Parameters," J. Comput. Chem. **28**, 1572–1581 (2007).
[56]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).
[57]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).
[58]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).
[59]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).
[60]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).
[61]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).
[62]M. Born and R. Oppenheimer, "Zur quantentheorie der molekeln," Ann. Phys. **389**, 457–484 (1927).
[63]E. Condon, "A Theory of Intensity Distribution in Band Systems," Phys. Rev. **28**, 1182–1201 (1926).

[64]E. Condon, "Nuclear Motion Associated with Electron Transitions in Diatomic molecles," Phys. Rev. **32**, 858–872 (1928).
[65]T. J. Zuehlsdorff and C. M. Isborn, "Combining the ensemble and Franck-Condon approaches for calculating spectral shapes of molecules in solution," J. Chem. Phys. **148**, 024110 (2018).
[66]A. Segalina, D. Aranda, J. A. Green, V. Cristino, S. Caramori, G. Prampolini, M. Pastore, and F. Santoro, "How the interplay among conformational disorder, solvation, local, and charge-transfer excitations affects the absorption spectrum and photoinduced dynamics of perylene diimide dimers: A molecular dynamics/quantum vibronic approach," J. Chem. Theory Comput **18**, 3718–3736 (2022).
[67]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).
[68]H. Kim and P. J. Rossky, "Evaluation of Quantum Correlation Functions from Classical Data," J. Phys. Chem. B **106**, 8240 (2002).
[69]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).
[70]R. Kubo, "Generalized cumulant expansion method," J. Phys. Soc. Jpn. **17**, 1100–1120 (1962).
[71]J. S. Bader and B. J. Berne, "Quantum and classical relaxation rates from classical simulations," J. Chem. Phys. **100**, 8359 (1994).
[72]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).
[73]C. P. Lawrence and J. L. Skinner, "Quantum corrections in vibrational and electronic condensed phase spectroscopy: Line shapes and echoes," Proc. Natl. Acad. Sci. U.S.A. **102**, 6720–6725 (2005).
[74]S. A. Egorov, K. F. Everitt, and J. L. Skinner, "Quantum Dynamics and Vibrational Relaxation," J. Phys. Chem. A **103**, 9494–9499 (1999).
[75]K. A. Jung, P. E. Videla, and V. S. Batista, "Inclusion of nuclear quantum effects for simulations of nonlinear spectroscopy," J. Chem. Phys. **148**, 244105 (2018).
[76]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).
[77]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).
[78]A. V. Marenich, C. J. Cramer, and D. G. Truhlar, "Electronic absorption spectra and solvatochromic shifts by the vertical excitation model: Solvated clusters and molecular dynamics sampling," J. Phys. Chem. B **119**, 958–967 (2015).
[79]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).
[80]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).
[81]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.; Steinbrecher, 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.
[82]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).
[83]T. Yanai, D. P. Tew, and N. C. Handy, "A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP)," Chem. Phys. Lett. **393**, 51–57 (2004).
[84]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).
[85]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).
[86]M. Sayed, D. K. Maity, and H. Pal, "A comparative photophysical study on the structurally related coumarin 102 and coumarin 153 dyes," J. Photochem. Photobiol. A **434**, 114265 (2023).
[87]A. D. Laurent and D. Jacquemin, "Td-dft benchmarks: A review," Int. J. Quantum Chem. **113**, 2019–2039 (2013).
[88]J. S. Sandoval and D. W. McCamant, "The best models of bodipy's electronic excited state: Comparing predictions from various dft functionals with measurements from femtosecond stimulated raman spectroscopy," J. Phys. Chem. A **127**, 8238–8251 (2023).
[89]T. J. Zuehlsdorff, "Molspeckpy: Spectroscopy python code, available on github: https://github.com/tjz21/spectroscopy_python_code," (2021).
[90]J. P. Dahl and M. Springborg, "The morse oscillator in position space, momentum space, and phase space," J. Chem. Phys. **88**, 4535–4547 (1988).
[91]L. Allan and T. J. Zuehlsdorff, "Publication data for "taming the 3rd order cumulant approximation to linear optical spectroscopy"',zenodo repository: https://zenodo.org/records/8361624,".