

## Tests for, origins of, and corrections to non-Gaussian statistics. The dipole-flip model

Addison J. Schile, and Ward H. Thompson

Citation: *J. Chem. Phys.* **146**, 154109 (2017); doi: 10.1063/1.4981009

View online: <https://doi.org/10.1063/1.4981009>

View Table of Contents: <http://aip.scitation.org/toc/jcp/146/15>

Published by the [American Institute of Physics](#)

---

### Articles you may be interested in

[Perspective: Dissipative particle dynamics](#)

*The Journal of Chemical Physics* **146**, 150901 (2017); 10.1063/1.4979514

[Path integral molecular dynamics with surface hopping for thermal equilibrium sampling of nonadiabatic systems](#)

*The Journal of Chemical Physics* **146**, 154110 (2017); 10.1063/1.4981021

[IR spectral assignments for the hydrated excess proton in liquid water](#)

*The Journal of Chemical Physics* **146**, 154507 (2017); 10.1063/1.4980121

[Introducing the mean field approximation to CDFT/MMpol method: Statistically converged equilibrium and nonequilibrium free energy calculation for electron transfer reactions in condensed phases](#)

*The Journal of Chemical Physics* **146**, 154101 (2017); 10.1063/1.4979895

[Variational Koopman models: Slow collective variables and molecular kinetics from short off-equilibrium simulations](#)

*The Journal of Chemical Physics* **146**, 154104 (2017); 10.1063/1.4979344

[Multiscale simulations of anisotropic particles combining molecular dynamics and Green's function reaction dynamics](#)

*The Journal of Chemical Physics* **146**, 114106 (2017); 10.1063/1.4977515

---

**PHYSICS TODAY**

WHITEPAPERS

#### ADVANCED LIGHT CURE ADHESIVES

Take a closer look at what these environmentally friendly adhesive systems can do

READ NOW

PRESENTED BY  
 **MASTERBOND**  
ADHESIVES | SEALANTS | COATINGS

# Tests for, origins of, and corrections to non-Gaussian statistics. The dipole-flip model

Addison J. Schile<sup>a)</sup> and Ward H. Thompson<sup>b)</sup>

*Department of Chemistry, University of Kansas, Lawrence, Kansas 66045, USA*

(Received 15 November 2016; accepted 3 April 2017; published online 21 April 2017)

Linear response approximations are central to our understanding and simulations of nonequilibrium statistical mechanics. Despite the success of these approaches in predicting nonequilibrium dynamics, open questions remain. Laird and Thompson [J. Chem. Phys. **126**, 211104 (2007)] previously formalized, in the context of solvation dynamics, the connection between the static linear-response approximation and the assumption of Gaussian statistics. The Gaussian statistics perspective is useful in understanding why linear response approximations are still accurate for perturbations much larger than thermal energies. In this paper, we use this approach to address three outstanding issues in the context of the “dipole-flip” model, which is known to exhibit nonlinear response. First, we demonstrate how non-Gaussian statistics can be predicted from purely equilibrium molecular dynamics (MD) simulations (i.e., without resort to a full nonequilibrium MD as is the current practice). Second, we show that the Gaussian statistics approximation may also be used to identify the physical origins of nonlinear response residing in a small number of coordinates. Third, we explore an approach for correcting the Gaussian statistics approximation for nonlinear response effects using the same equilibrium simulation. The results are discussed in the context of several other examples of nonlinear responses throughout the literature. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4981009>]

## I. INTRODUCTION

Understanding the chemical dynamics of liquids is critical to the study of chemical reactions, spectroscopy, energy transfer, and catalysis. Linear response theories,<sup>1–5</sup> in which perturbations are described by fluctuations in the equilibrium state, are the dominant method for describing nonequilibrium dynamics. In many linear-response approaches,<sup>1–4</sup> the perturbation is assumed to be weak—on the order of thermal energies ( $k_B T$ )—so that higher-order correlations in the equilibrium fluctuations are neglected. Such linear response approximations have been remarkably successful in predicting nonequilibrium dynamical properties for not only weak perturbations, but also for perturbations much larger than those on the order of the thermal energy. It has been appreciated for some time,<sup>6,7</sup> and even shown directly,<sup>5,8</sup> that the latter is a consequence of systems exhibiting Gaussian statistics—which predicts not just local but *global*, linear response. Thus, Gaussian statistics approximations underlie the linear response approaches that assume small perturbations and represent a more broadly applicable framework for thinking about linear response.

A key application of linear-response approximations has been in describing time-dependent fluorescence (TDF) experiments that are used to study solvation dynamics. In these experiments, a dye molecule with a charge-transfer transition is photoexcited, and the subsequent return to the ground state via fluorescence is followed as a function of time after

excitation using femtosecond spectroscopy. Specifically, the fluorescence energy,  $\Delta E(t)$ , or Stokes shift at time  $t$  after excitation reports on the rearrangements of the solvent induced by the change in the dye-molecule charge distribution upon excitation. The result is often presented as the normalized time-dependent Stokes shift (TDSS)

$$S(t) = \frac{\langle \Delta E(t) \rangle_{ne} - \langle \Delta E(\infty) \rangle_{ne}}{\langle \Delta E(0) \rangle_{ne} - \langle \Delta E(\infty) \rangle_{ne}}, \quad (1)$$

where the  $\langle \dots \rangle_{ne}$  indicates a nonequilibrium average and  $\langle \Delta E(\infty) \rangle_{ne}$  is the relaxed fluorescence energy. This experiment can be modeled with nonequilibrium molecular dynamics (MDs) simulations in which the equilibrium ground-state dye molecule is excited, i.e., the charge distribution is changed at  $t = 0$ , and the resulting time-dependent fluorescence energy,  $\Delta E(t) = \Delta E_e(t) - \Delta E_g(t)$ , is monitored. Here, the subscripts indicate the excited ( $e$ ) and ground ( $g$ ) states and  $\Delta E(t)$  thus represents the difference in the dye-solvent interactions between the two electronic states (i.e., solute charge distributions).

Linear response approximations relate the nonequilibrium response to the fluctuations observed at equilibrium. There are two approaches to obtain linear response approximations which give  $S(t)$  in terms of only equilibrium simulations, i.e., in terms of a correlation function of the equilibrium fluctuations of the energy gap  $\Delta E$ . In the first approach, the dynamic linear-response approximation<sup>9</sup> gives the TDSS in terms of a ground-state correlation function,  $S(t) \approx C_g(t) = \langle \delta \Delta E(t) \delta \Delta E(0) \rangle_g / \langle \delta \Delta E^2 \rangle_g$ , where the angled brackets denote an equilibrium average in the ground ( $g$ ) state and  $\delta \Delta E(t) = \Delta E(t) - \langle \Delta E \rangle_g$  is the fluctuation in the energy gap at time  $t$  from its equilibrium value. This approximation has

<sup>a)</sup>Current address: Department of Chemistry, University of California, Berkeley, CA 94720, USA.

<sup>b)</sup>Electronic mail: [wthompson@ku.edu](mailto:wthompson@ku.edu)

been used to successfully describe the solvation dynamics of many systems and, indeed, most simulations of TDF measurements reported in the literature have been carried out using this method. The second approach, the static linear-response approximation, gives the TDSS in terms of the excited-state fluctuations,  $S(t) \approx C_e(t) = \langle \delta \Delta E(t) \delta \Delta E(0) \rangle_e / \langle \delta \Delta E^2 \rangle_e$ . This is obtained (*vide infra*) by assuming that fluctuations of the energy gap at equilibrium are small. Much like the dynamic linear-response approximation, the static linear-response approximation can accurately describe solvation dynamics in many systems. Interestingly, Maroncelli and Fleming found that it was generally superior to the dynamic linear-response approximation in tests against nonequilibrium simulations,<sup>10</sup> a result that was subsequently explained more directly by Laird and Thompson.<sup>5,8</sup> However, in a number of notable systems, neither linear-response approximation is accurate.<sup>11–15</sup> The reasons for these breakdowns are not yet fully understood in terms of the linear-response approximation, making it difficult to know if an approximation is justified absent a full nonequilibrium treatment. It is these issues that are the subject of this paper.

Laird and Thompson have previously shown that the static linear-response approximation, i.e.,  $S(t) \approx C_e(t)$ , actually represents the assumption that the energy gap,  $\Delta E$ , obeys Gaussian statistics.<sup>5,8</sup> Where the static linear response approximation ignores higher-order correlation functions (truncating the perturbation after the linear term), the Gaussian statistics assumption factorizes them into lower-order correlation functions. The Gaussian statistics assumption therefore does not rely upon the (usually unjustified) approximation that the energy gap fluctuations are small<sup>1–4,14</sup> and thus it represents a global approximation that is still applicable for systems driven far from equilibrium. We note that the connection between linear response and Gaussian statistics was widely recognized before this explicit relationship was derived (see, e.g., Refs. 6 and 7). For example, Marcus theory invokes linear response to illustrate the mechanism by which electron transfer reactions occur, that is, that solvent fluctuations on a *harmonic* free energy surface can adequately describe nonequilibrium dynamics.<sup>16</sup> Throughout the rest of this paper, we will refer to the approximation  $S(t) \approx C_e(t)$  as the Gaussian statistics, rather than static linear-response, approximation except when noting the differences between the two approaches.

In this paper, we address some of the outstanding fundamental questions about nonequilibrium chemical dynamics and the Gaussian statistics approximation. First, *can non-Gaussian statistics be predicted by relying only on equilibrium simulations?* Second, *what is the origin of the non-Gaussian statistics?* Third, *can the Gaussian statistics approximation be corrected to describe nonlinear response?*

We are motivated, in part, by the work of Ladanyi and co-workers, who studied solvation dynamics of model diatomic dye molecules in water and methanol to investigate linear response. They studied two models for the electronic transition. In the first model, excitations induced a dipole moment in the excited state from a nonpolar, ground-state diatomic solute.<sup>11,12</sup> In the second model, excitation induced a “dipole flip” in which the ground-state dipolar molecule charges switch in the excited state, as depicted in Fig. 1.<sup>12,13</sup> In both models,

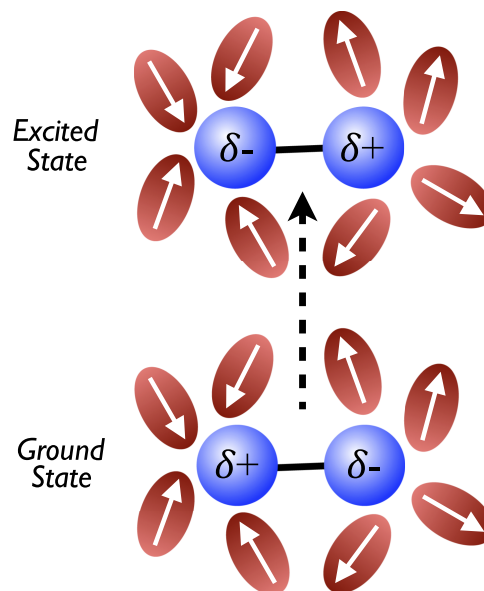


FIG. 1. Schematic illustration of the excitation of the dipole-flip solute model in which promotion from the ground to the excited state exchanges the charges on the two atoms; the solvent dipole moments are unchanged in the Franck-Condon transition.

linear-response approximations failed to predict the nonequilibrium dynamics: the nonequilibrium simulations showed faster relaxation times than were computed from equilibrium correlation functions. These linear-response breakdowns were attributed to changes in hydrogen-bonding between the solvent and solute after the excitation. The linear-response approximation accurately described the inertial relaxation at short times but showed significant deviations at longer times associated with molecular reorientations. The latter, dipole-flip, model is particularly interesting in that the ground- and excited-state equilibrium dynamics are the same so that the dynamic linear-response and Gaussian statistics approximations are equivalent. Not only does this model provide a convenient test for the approximations, but because of the breakdown despite the symmetry, it also hints at the way in which linear response fails, as will be discussed throughout this paper.

The remainder of this paper is organized as follows. The linear-response and Gaussian statistics approximations are briefly outlined in Sec. II. The details of the solvation dynamics of the dipole-flip model in various solvents are given in Sec. III. The results of the simulations are then presented in Secs. IV–VI in the context of testing diagnostics for non-Gaussian statistics, uncovering the origins of the non-Gaussian behavior, and evaluating corrections to non-Gaussian dynamics. The present results are discussed in the context of other examples of nonlinear response in the literature in Sec. VII. Finally, some concluding remarks are offered in Sec. VIII.

## II. THEORY

In this section, we briefly outline the static linear response and Gaussian statistics approximations for the case of a change in solute state, e.g., solvation dynamics. Detailed derivations, including that for dynamic linear response, have been provided elsewhere.<sup>5,8</sup> We consider the change in some property  $A$  as a function of time after excitation from the ground,  $g$ , state to

the excited,  $e$ , state. This can be expressed as

$$\begin{aligned}\langle A(t) \rangle_{ne} &= \frac{\int d\mathbf{p} \int d\mathbf{q} A(t) e^{-\beta H_g}}{\int d\mathbf{p} \int d\mathbf{q} e^{-\beta H_g}} \\ &= \frac{\int d\mathbf{p} \int d\mathbf{q} [e^{-\hat{L}_e t} A(0)] e^{-\beta H_g}}{\int d\mathbf{p} \int d\mathbf{q} e^{-\beta H_g}},\end{aligned}\quad (2)$$

where  $H_g$  ( $H_e$ ) is the ground-state (excited-state) Hamiltonian and  $\hat{L}_e$  is the excited-state Liouville operator,  $\hat{L}_e = \{H_e, \cdot\}$ , with  $\{F, G\} = \sum_j [(\partial F / \partial p_j)(\partial G / \partial q_j) - (\partial F / \partial q_j)(\partial G / \partial p_j)]$  the Poisson bracket. Note that the time-dependence of the property  $A$  is determined by excited-state dynamics initiated from equilibrium ground-state configurations. The exact (classical) result for  $\langle A(t) \rangle_{ne}$  can thus be obtained from non-equilibrium molecular dynamics in which ground-state equilibrium simulations are used to generate configurations from which excited-state dynamics are initiated. Then,  $\langle A(t) \rangle_{ne}$  is calculated as an average of  $A$  over the non-equilibrium, excited-state trajectories as a function of time,  $t$ , after excitation.

Linear-response approximations, however, yield expressions for  $\langle A(t) \rangle_{ne}$  in terms of only *equilibrium* dynamics, resulting in significantly less computational effort. In these approximations, the ground-state Hamiltonian is expressed in terms of the excited-state Hamiltonian and the fluorescence energy,  $H_g = H_e - \Delta E$ . Then, Eq. (2) can be rearranged as

$$\langle A \rangle_{ne}(t) = \frac{\int d\mathbf{p} \int d\mathbf{q} A(t) e^{\beta \Delta E} e^{-\beta H_e}}{\int d\mathbf{p} \int d\mathbf{q} e^{\beta \Delta E} e^{-\beta H_e}},\quad (3)$$

where  $A(t)$  is propagated under the excited-state Hamiltonian (i.e.,  $A(t) = e^{-\hat{L}_e t} A(0)$ ). By dividing the numerator and the denominator by the excited-state partition function the dynamic nonequilibrium average of  $A$  is

$$\langle A \rangle_{ne}(t) = \frac{\langle A(t) e^{\beta \Delta E} \rangle_e}{\langle e^{\beta \Delta E} \rangle_e},\quad (4)$$

where the angled brackets indicate an equilibrium average and the subscript  $e$  signifies that the average is in the excited state. Defining the average energy gap  $\overline{\Delta E} \equiv \langle \Delta E \rangle_e$  and property  $\overline{A} \equiv \langle A \rangle_e$ , it is straightforward to show that

$$\begin{aligned}\langle A \rangle_{ne}(t) - \langle A \rangle_e &= \frac{\langle A(t) e^{\beta \delta \Delta E} \rangle_e}{\langle e^{\beta \delta \Delta E} \rangle_e} - \frac{\overline{A} \langle e^{\beta \delta \Delta E} \rangle_e}{\langle e^{\beta \delta \Delta E} \rangle_e} \\ &= \frac{\langle \delta A(t) e^{\beta \delta \Delta E} \rangle_e}{\langle e^{\beta \delta \Delta E} \rangle_e},\end{aligned}\quad (5)$$

where  $\delta \Delta E(t) = \Delta E(t) - \overline{\Delta E}$  and  $\delta A(t) = A(t) - \overline{A}$ .

So far no approximation has been invoked, but it is at this point that the derivations of the static linear-response and Gaussian statistics approximations diverge.

### A. Static linear response

The static linear-response approximation<sup>1-4</sup> is obtained by expanding the exponential in Eq. (5) in a Taylor series and truncating at the linear term

$$e^{\beta \delta \Delta E} \approx 1 + \beta \delta \Delta E.\quad (6)$$

This approximation can be introduced in both the numerator and denominator and is valid only when the equilibrium fluctuations in  $\Delta E$  are small relative to  $k_B T$  ( $\beta \delta \Delta E \ll 1$ ). This

gives the dynamic Stokes shift as

$$\begin{aligned}\langle A \rangle_{ne}(t) - \langle A \rangle_e &\approx \frac{\langle \delta A(t) \rangle_e + \beta \langle \delta A(t) \delta \Delta E(0) \rangle_e}{1 + \beta \langle \delta \Delta E \rangle_e} \\ &= \beta \langle \delta A(t) \delta \Delta E(0) \rangle_e,\end{aligned}\quad (7)$$

where the last equality results because  $\langle \delta A(t) \rangle_e = 0$  and  $\langle \delta \Delta E \rangle_e = 0$ .

### B. Gaussian statistics

The mathematical expression in Eq. (7) can, however, be obtained without invoking any approximations regarding the size of fluctuations. If, instead of truncating at the linear term the full Taylor series is included for the exponential, one has

$$\begin{aligned}\langle \delta A(t) e^{\beta \delta \Delta E(0)} \rangle_e &= \sum_{n=0}^{\infty} \frac{1}{n!} \beta^n \langle \delta A(t) \delta \Delta E(0)^n \rangle_e \\ &= \sum_{n=0}^{\infty} \frac{1}{(2n)!} \beta^{2n} \langle \delta A(t) \delta \Delta E(0)^{2n} \rangle_e \\ &\quad + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \beta^{2n+1} \langle \delta A(t) \delta \Delta E(0)^{2n+1} \rangle_e.\end{aligned}\quad (8)$$

Then, if  $\delta A(t)$  and  $\delta \Delta E(t)$  are Gaussian random variables, all of the higher-order correlation functions can be factorized into only two-point correlation functions, e.g.,  $\langle \delta A(t_i) \delta \Delta E(t_j) \rangle_e$ , according to Wick's theorem.<sup>17,18</sup> Thus, the correlation function  $\langle \delta A(t) \delta \Delta E(0)^{2n} \rangle_e$  can be factorized into terms of the form  $\langle \delta A(t) \delta \Delta E(0) \rangle_e \langle \delta \Delta E(0) \rangle_e$  and  $\langle \delta \Delta E(0)^2 \rangle_e \langle \delta A(t) \rangle_e$ . However, because  $\delta A$  and  $\delta \Delta E$  both have zero mean, then  $\langle \delta A(t) \delta \Delta E(0)^{2n} \rangle_e = 0$  for all  $n$  when the two variables obey Gaussian statistics. Applying similar factorizations to  $\langle \delta A(t) \delta \Delta E(0)^{2n+1} \rangle_e$  yields<sup>5,8</sup>

$$\begin{aligned}\langle \delta A(t) e^{\beta \delta \Delta E(0)} \rangle_e^G &= \sum_{n=0}^{\infty} \frac{1}{n! 2^n} \beta^{2n+1} \langle \delta A(t) \delta \Delta E(0) \rangle_e \langle \delta \Delta E^2 \rangle_e^n \\ &= \beta \langle \delta A(t) \delta \Delta E(0) \rangle_e e^{\beta^2 \langle \delta \Delta E^2 \rangle_e / 2}.\end{aligned}\quad (9)$$

If  $\delta \Delta E$  is a Gaussian variable then the denominator in Eq. (4) (which is the characteristic function evaluated at  $-i\beta$ ) can be similarly evaluated to give

$$\langle e^{\beta \delta \Delta E} \rangle_e \approx e^{\beta^2 \langle \delta \Delta E^2 \rangle_e / 2},\quad (10)$$

which gives the result for the nonequilibrium change in the property  $A$  within the Gaussian statistics approximation as

$$\langle A \rangle_{ne}(t) - \langle A \rangle_e \approx \beta \langle \delta A(t) \delta \Delta E(0) \rangle_e.\quad (11)$$

This is the same expression as the static linear-response approximation to the Stokes shift in Eq. (7). It is important, however, to note that the physical assumptions used here are quite different from those used in static linear response. In particular, there is no requirement that  $\beta \delta \Delta E \ll 1$ . As has been shown before<sup>8,19</sup> and will be demonstrated below, this assumption of the static linear-response approach is not satisfied and thus it is more accurate to consider the expression in Eqs. (7) and (11) as the Gaussian statistics approximation.



### III. SIMULATION DETAILS

Equilibrium and nonequilibrium simulations were carried out for a dipolar solute in bulk solutions of methanol, water, and acetonitrile. Each system consisted of 256 total molecules: 1 solute molecule and 255 water, methanol, or acetonitrile molecules. The solute was modeled as a diatomic molecule with each atom having equal and opposite charges (Fig. 1). The molecules interact through site-site Lennard-Jones and Coulombic potentials; the force-field parameters are given in Table I. The solvent molecules are flexible with the water, methanol, and acetonitrile potentials similar to the similar to the extended simple point charge (SPC/E),<sup>20</sup> all-atom optimized potentials for liquid simulations (OPLS-AA),<sup>21</sup> and Nikitin and Lyubartsev<sup>22</sup> models, respectively; the intramolecular potential uses harmonic bond and angle bending potentials using the parameters given in Table II. Long-range interactions were computed using an Ewald sum with a tolerance of  $10^{-4}$  and a potential cutoff radius of 12.0 Å. The solute bond length was held fixed using the SHAKE algorithm.<sup>23</sup>

Two types of MD simulations were carried out: (1) equilibrium trajectories with the solute in the excited state and (2) nonequilibrium trajectories where the solute is promoted to the excited state from initial configurations obtained from an equilibrium ground state simulation and the dynamics are subsequently followed. Note that the dipole-flip solute has the same ground- and excited-state dynamics due to the symmetry of the charges. For each system, five equilibrium trajectories were sampled. Each simulation was propagated for a time between 400 and 600 ps in the NVE ensemble in which velocity rescaling was used to maintain the temperature. Following the NVE period, the system was propagated using the NVT ensemble at a temperature of 298.15 K with a Verlet integrator and a time step of 1 fs. The Nosé-Hoover thermostat was used to maintain temperature with a time constant of 100 fs. An equilibration time of 40–60 ps was followed by a production

TABLE I. Lennard-Jones and Coulombic site parameters for each atom used in the simulations.

Atom	$\epsilon$ (kcal/mol)	$\sigma$ (Å)	$q$	$m$ (g/mol)
H <sub>2</sub> O				
O	0.155 35	3.166	-0.8476	16.00
H	0	0	0.4238	1.0079
CH <sub>3</sub> OH				
O	0.17	3.12	-0.683	16.00
H	0	0	0.418	1.0079
C	0.066	3.5	0.145	12.01
H (CH <sub>3</sub> )	0.03	2.5	0.040	1.0079
CH <sub>3</sub> CN				
N	0.17	3.20	-0.56	14.019
C	0.066	3.5	0.46	12.01
C (CH <sub>3</sub> )	0.066	3.5	-0.08	12.01
H (CH <sub>3</sub> )	0.015	2.5	0.06	1.0079
Solute (ground)				
A	0.175	3.083	0.5	30.00
B	0.175	3.083	-0.5	30.00
Solute (excited)				
A	0.175	3.083	-0.5	30.00
B	0.175	3.083	0.5	30.00

TABLE II. Intramolecular potential parameters<sup>a</sup> for the H<sub>2</sub>O, CH<sub>3</sub>OH, and CH<sub>3</sub>CN solvents.

Bond	$r_0$ (Å)	$k_r$	Angle	$\theta_0$ (°)	$k_\theta$
H <sub>2</sub> O					
O-H	1.00	553.0	H-O-H	109.47	55.0
CH <sub>3</sub> OH					
C-O	1.41	320.0	O-C-H	109.5	35.0
O-H	0.96	553.0	C-O-H	108.5	55.0
C-H	1.09	331.0	H-C-H	109.5	35.0
CH <sub>3</sub> CN					
C-H	1.09	340.0	H-C-H	109.5	35.0
C-C	1.458	385.0	H-C-C	108.0	35.0
C-N	1.157	650.0	C-C-N	180.0	150.0

<sup>a</sup>The stretching potential is of the form  $V(r)=k_r(r-r_0)^2$  with  $k_r$  in units of (kcal/mol)/Å and the bending potential  $V(r)=k_\theta(\theta-\theta_0)^2$  with  $k_\theta$  in units of (kcal/mol)/radian<sup>2</sup>.

run of 5 ns for the solute in the ground state. The potential energy was then computed for the solute in the excited state in the configurations of the ground-state trajectory. Data for the simulations were recorded every 100 fs.

The nonequilibrium MD simulations were performed from the configurations produced by the ground state equilibrium trajectory. The solute at a particular configuration was promoted to the excited state and the dynamics in the excited state were followed for 50 ps. The ground-state dynamics were then computed along this excited-state trajectory to give the fluorescence energy,  $\Delta E(t)$ . This process was repeated for a total of 1000 nonequilibrium trajectories, taken from the ground-state equilibrium configurations every 5 ps. Data for the nonequilibrium simulations were recorded every 10 fs. Error bars were computed using block averaging. For the equilibrium simulations, five blocks were used where each block was taken to be an individual equilibrium simulation. Ten blocks were used for the nonequilibrium MD simulations, where each block was composed of 100 individual trajectories.

### IV. CAN NON-GAUSSIAN STATISTICS BE PREDICTED FROM EQUILIBRIUM SIMULATIONS?

We start by addressing an important practical issue involved in applying the Gaussian statistics approximation. Namely, how can one be confident that the results are accurate, i.e., the system exhibits Gaussian statistics, without carrying out a full nonequilibrium simulation?

We first show, as has been done before,<sup>6,12,13,24</sup> that the dipole-flip model exhibits a breakdown of the Gaussian statistics approximation for water and methanol solvents. It is first worthwhile, however, to note some general properties of the normalized time-dependent Stokes shift,  $S(t)$  for these two systems. The TDSS results obtained from nonequilibrium MD simulations are shown in Fig. 2. For the solute immersed in either solvent,  $S(t)$  is well fit by a tri-exponential function. In water, the shortest time scale is 20 fs and represents the inertial motion of the nearby solvent molecules. The intermediate time scale is  $\sim 0.35$  ps and represents the oscillating librational motion of water due to its hydrogen bonds. The longest time scale, which is associated with molecular rearrangements, is  $\sim 1.7$  ps. Methanol has an inertial time scale

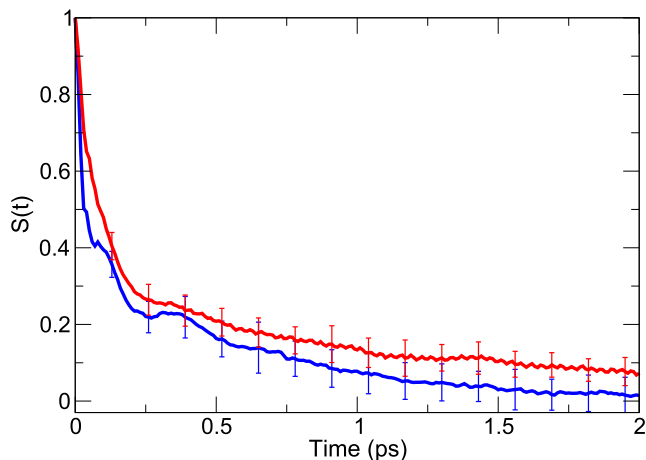


FIG. 2. Normalized time-dependent Stokes shift,  $S(t)$ , obtained from nonequilibrium MD simulations for water (blue) and methanol (red).

of  $\sim 65$  fs, a librational time scale of  $\sim 0.64$  ps, and a longest time scale of  $\sim 5$  ps. These relaxation times resulting from the simulations conducted in this study are consistent with those found previously.<sup>13</sup>

The difference in relaxation at short times can be attributed to key differences in the solvation dynamics between the two solvents. First, the three moments of inertia for water are all small in comparison to methanol which has one small and two large moments of inertia.<sup>13,24</sup> The larger moments of inertia describe the slower response time of methanol in the inertial regime. Second, and more relevant to the issue of the Gaussian behavior, is the number of solvent-solute hydrogen bonds. Water, because of its smaller size, donates more hydrogen bonds to the negatively charged solute site and thus exhibits a larger amplitude for the librational component of the solvation dynamics compared to methanol. The differences in the longer time scales, associated with molecular rearrangements, are consistent with the faster OH reorientation for water compared to methanol.<sup>25</sup>

For the dipole-flip model considered in this paper, the Gaussian statistics approximation has been implemented for the fluorescence energy, i.e.,  $A = \Delta E$ , to approximate the normalized TDSS response function in Eq. (1). This gives

$$S(t) \approx C(t) = \frac{\langle \delta \Delta E(t) \delta \Delta E(0) \rangle}{\langle \delta \Delta E^2 \rangle}, \quad (12)$$

where the subscript “e” has been suppressed on the thermal average due to the symmetry of the model. The full nonequilibrium  $S(t)$  and the Gaussian statistics approximation,  $C(t)$ , for water and methanol are shown in Fig. 3. The data show

that Gaussian statistics is an imperfect approximation for the nonequilibrium solvation dynamics of the dipole-flip model in water and a very poor one in methanol. For each system, the relaxation predicted by Gaussian statistics is slower and does not agree (within error bars) with the nonequilibrium dynamics until  $t \approx 0.5$  ps for water and  $t > 8$  ps for methanol.

Clearly, the error of the Gaussian statistics approximation is greater for methanol than it is for water. However, the larger question is how one could evaluate the approximation in either case without the direct comparison to the exact, nonequilibrium result, as presented in Fig. 3. As we now discuss, the Gaussian statistics approximation outlined in Sec. II provides an answer. Namely, Wick’s theorem states that if  $\Delta E$  is a Gaussian random variable then

$$C_m(t) \equiv \beta^{m-1} \langle \delta \Delta E(t) \delta \Delta E(0)^{m-1} \rangle = 0 \quad (13)$$

if  $m$  is odd, and, if  $m$  is even

$$C_m(t) = \frac{(2n+1)!}{n! 2^n} \beta^{2n+1} \langle \delta \Delta E(t) \delta \Delta E(0) \rangle \langle \delta \Delta E^2 \rangle^n \equiv C_m^{fac}(t), \quad (14)$$

where  $n = m/2 - 1$  (or  $m = 2n + 2$ ). Thus, the quality of the Gaussian statistics approximations depends on the accuracy of these factorizations and can be directly evaluated by the comparison of  $C_m(t)$  and  $C_m^{fac}(t)$  for the higher-order ( $m > 2$ ) correlation functions.

The first of these higher-order correlation functions and their factorizations based on Wick’s theorem are displayed in Fig. 4, where the non-normalized Gaussian statistics approximation,  $C_2(t) = \beta \langle \delta \Delta E(t) \delta \Delta E(0) \rangle_e$ , and the first term that is neglected according to Eq. (13),  $C_3(t) = \beta^2 \langle \delta \Delta E(t) \delta \Delta E(0)^2 \rangle_e / 2$ , are shown. For both water and methanol solvents, the latter, higher-order correlation function is nearly, but not fully, zero within error bars, which is expected for Gaussian statistics. For water it is nonzero up to  $t \approx 0.4$  ps, while for methanol the results suggest that  $C_3(t)$  may be non-zero for much longer times, but the (substantial) error bars overlap zero for all times. The next highest, odd-order correlation functions,  $C_5(t) = \beta^2 \langle \delta \Delta E(t) \delta \Delta E(0)^4 \rangle_e / 24$  (not shown), are essentially zero for all times for water but non-zero and positive for methanol with a rapid ( $< 30$  fs) initial drop followed by an  $\sim 1.5$  ps decay.

The even-order time correlation functions are all non-zero and it is the factorization in Eq. (14) that can be checked for deviations from Gaussian statistics. The first example is shown in Fig. 5 where  $C_4(t)$  and its factorized form,  $C_4^{fac}(t)$ , are compared for both water and methanol solvents.

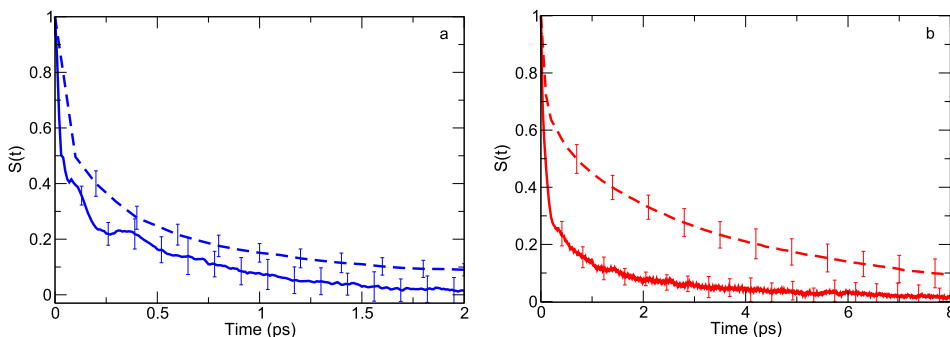


FIG. 3. The normalized time-dependent Stokes shift,  $S(t)$ , obtained from nonequilibrium MD (solid lines) is compared to the Gaussian statistics approximation,  $C(t)$ , (dashed lines) for (a) water (blue) and (b) methanol (red).

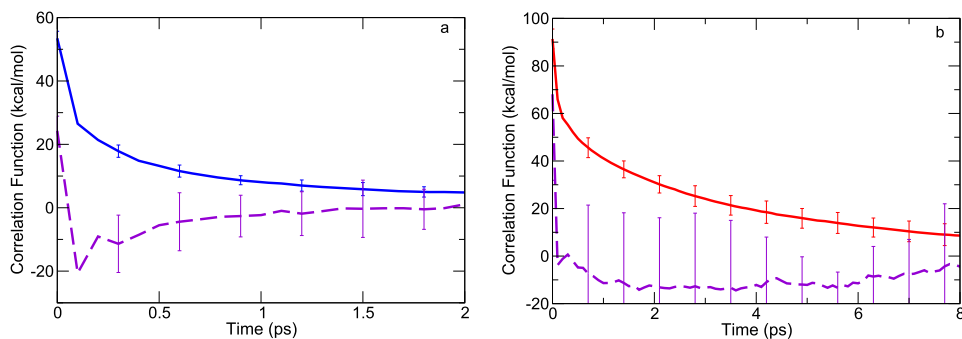


FIG. 4. Shown are the first two terms in the Taylor series,  $C_2(t)$ , (solid lines) and  $C_3(t)$  (dashed violet lines) for (a) water (blue) and (b) methanol (red).

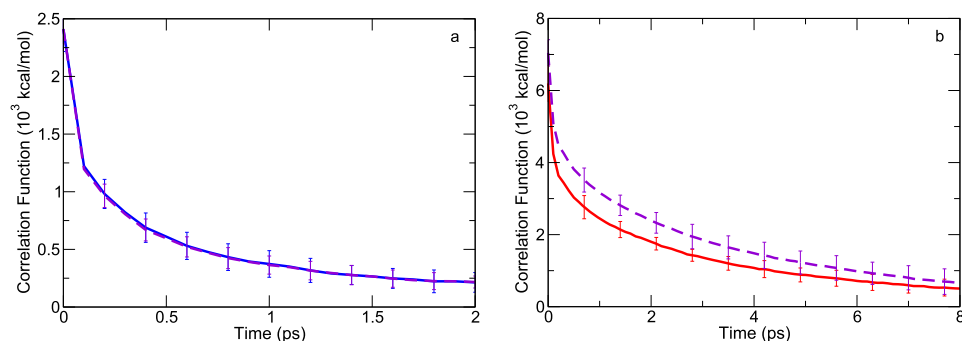


FIG. 5. Shown are the fourth-order correlation functions and Gaussian statistics factorizations,  $C_4(t)$  (solid lines) and  $C_4^{fac}(t)$  (dashed violet lines), respectively, for (a) water (blue) and (b) methanol (red).

Similarly, the sixth-order correlation functions,  $C_6(t)$ , are compared to the factorized forms,  $C_6^{fac}(t)$ , in Fig. 6. Note that the magnitude of these correlation functions increases with the order in direct contradiction of the approximation invoked in static linear response. Interestingly, when water is the solvent, the factorization of these even-order time correlation functions is an excellent approximation, with no discernable difference with the full time correlation functions. In contrast, the results for the methanol case show significant disagreement between the full and factorized forms. For both  $C_4(t)$  and  $C_6(t)$ , the corresponding factorized correlation function is larger than the full version. The difference is larger for  $C_6^{fac}(t)$  than  $C_4^{fac}(t)$ , but in both cases the factorized form decays slightly faster than the full correlation function such that they are in better agreement at longer times. These results are consistent with the comparison of the nonequilibrium response,  $S(t)$ , to the Gaussian statistics approximation,  $C(t)$ , in Fig. 2, where the latter is significantly larger than the former.

Taken together, these comparisons of the full, higher-order time correlation functions to the factorized forms assumed in the Gaussian statistics approximation indicate a strong breakdown of the approximation for methanol and a quite weak one for water. This is generally consistent with the

comparison of the Gaussian statistics result with the nonequilibrium response in Fig. 2, where the methanol case shows significant disagreement while for water the differences are smaller and more quantitative.

It is interesting to compare these results to an analysis of the equilibrium energy gap distribution,  $P(\delta\Delta E)$ . This is done in Fig. 7, where the distributions for water and methanol are shown. For both solvents, the results are interesting. In water, the energy gap distribution is well described by a Gaussian, with deviations only observed in the very wings of the distribution. In methanol, however, the statistics of the energy gap are non-Gaussian exhibiting a shoulder on the low energy side of the distribution. These results make prescribing a cause of non-Gaussian statistics difficult: in some cases, like methanol, the breakdown is expressly found in the energy gap distribution, but in others, like water, the breakdown is subtle, or even “hidden,” and not clearly evident in the distribution.<sup>26</sup> This was previously observed for the dipole-flip model in water by Geissler and Chandler who observed a similar Gaussian behavior but showed, by analysis of the nonequilibrium energy gap distributions after excitation, that it was nonstationary, i.e., the distribution changed with time.<sup>6</sup>

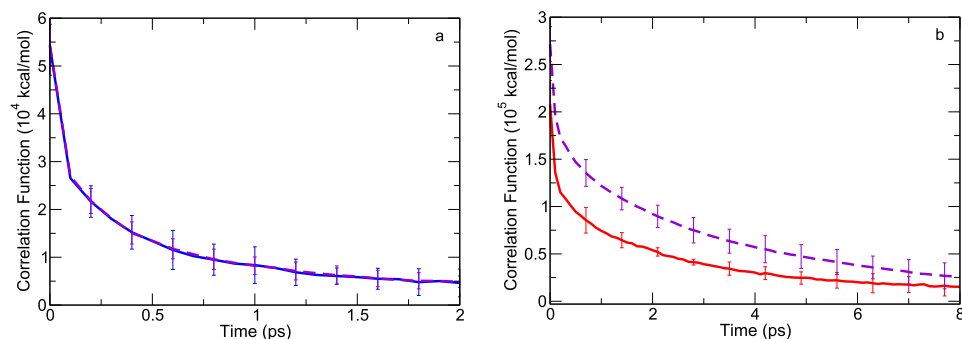


FIG. 6. Shown are the sixth-order correlation functions and Wick's factorizations,  $C_6(t)$  (solid lines) and  $C_6^{fac}(t)$  (dashed violet lines), respectively, for (a) water (blue) and (b) methanol (red).

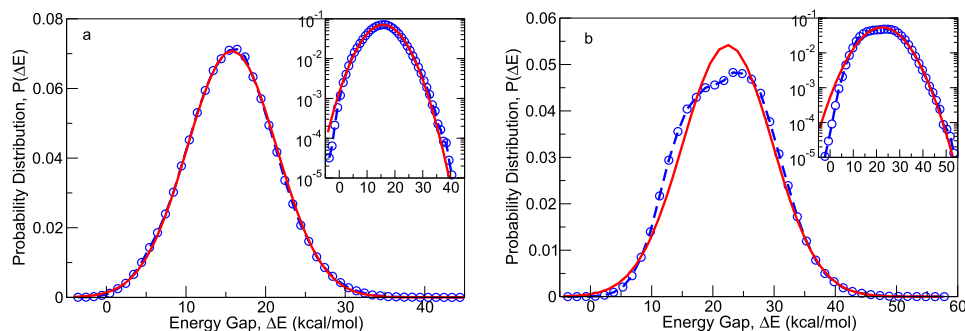


FIG. 7. Energy gap probability distributions for (a) water and (b) methanol computed directly from a histogram of the simulation data (blue circles and dashed line) are compared with a Gaussian having the same mean and standard deviation as the simulation data (solid red line); insets show the data on a semi-log plot.

## V. WHAT IS THE PHYSICAL ORIGIN OF THE NON-GAUSSIAN BEHAVIOR?

The previous results show how breakdowns of the Gaussian statistics approximation can be predicted without resorting to full nonequilibrium simulations. However, it is also important to understand the origin of the breakdown. For the dipole-flip model, it has been proposed that the breakdown of the Gaussian statistics approximation is caused by changes in the solute-solvent hydrogen-bonding (H-bonding) network upon excitation.<sup>11,13</sup> Here, we investigate this hypothesis in two ways. First, we consider an aprotic solvent that cannot hydrogen bond with the dipole-flip solute. Second, we examine how the Gaussian statistics approximation allows a partitioning between different contributions to the solvation energy and thus an isolation of the non-Gaussian motions responsible for the breakdown.

Choosing a coordinate to partition, as will be shown in Secs. V A and V B, for the dipole-flip model is fairly intuitive, and previous studies have made the identification simple. For other systems, choosing an adequate partition may not be as straightforward and may require further analysis of equilibrium simulations. These coordinates must be sufficiently small in number such that the central limit theorem does not apply but their contributions must be sufficiently large that their effect is important in the overall nonequilibrium response.<sup>14,27,28</sup>

### A. Dipole-flip solvation in an aprotic solvent

If H-bonding rearrangements are the origin of the failure of the Gaussian statistics approximation, there should be no such breakdown for solvation in an aprotic solvent. We have tested this idea by comparing the nonequilibrium TDSS result,  $S(t)$ , for the dipole-flip model in acetonitrile to the Gaussian statistics result,  $C(t)$ , as shown in Fig. 8. The data show that the Gaussian statistics approximation is in excellent agreement with the full, nonequilibrium MD result. This supports the implication of H-bonding dynamics in the non-Gaussian behavior.

This can be examined further by comparison of the higher-order correlation functions and their Wick's factorizations. These are shown for the dipole-flip solute in acetonitrile in Fig. 9. The results clearly show that for this non-H-bonding solvent, the Gaussian statistics approximation is an excellent description of the nonequilibrium solvation dynamics, with only a small deviation from zero of the lowest odd-order correlation function,  $C_3(t)$ , at short times. The data also reinforce the

diagnostic utility of the comparison of the full and factorized higher-order correlation functions for evaluating the Gaussian statistics approximation from purely equilibrium simulations. Finally, they illustrate why the static linear-response approximation is not the appropriate interpretation of Eq. (12). Even in this case where  $C(t) \approx S(t)$ , each higher-order correlation function is larger than all the lower order ones and is thus not remotely negligible, while, in contrast, the factorization assumed in the Gaussian statistics approximation is highly accurate.

### B. Partitioning Gaussian and non-Gaussian solvent coordinates

These results implicate the H-bonds between the solute and solvent molecules in the breakdown of the Gaussian statistics approximation in water and methanol. To explicitly demonstrate this we take a different tack. Specifically, we have partitioned the trajectories into a hydrogen-bonded partition—which is presumably non-Gaussian—in which only those solvent molecules that are H-bonded to the solute and the solute itself are included and a non-H-bonded partition—which is presumably Gaussian—in which only those solvent molecules that are not H-bonded to the solute and the solute itself are included. Solute-solvent H-bonds were identified using geometric criteria for both methanol and water. Specifically the criteria were based on the oxygen-site distance  $r_{O\alpha} \leq 3.5 \text{ \AA}$ , the hydrogen-site distance  $r_{H\alpha} \leq 2.45 \text{ \AA}$ , and the

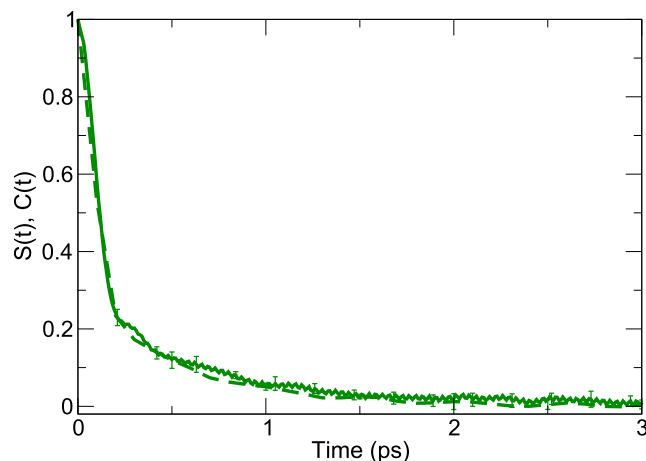


FIG. 8. Time-dependent fluorescence correlation functions using the full nonequilibrium MD,  $S(t)$ , (solid green line) and the Gaussian statistics approximation,  $C(t)$ , (dashed green line) for acetonitrile.



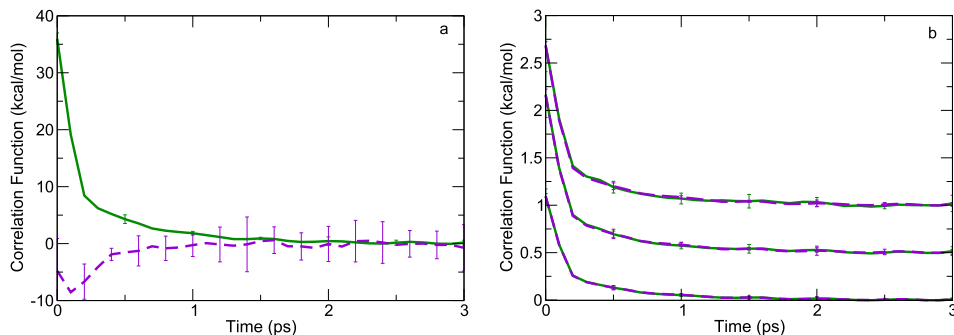


FIG. 9. Higher-order correlation functions (solid green lines) and corresponding Wick's factorizations (dashed violet lines) are shown for the dipole-flip model in acetonitrile, (a)  $C_2(t)$  and  $C_3(t)$  are compared, and (b) from bottom to top  $C_4(t)$  and  $C_4^{fac}(t)$  (scaled by  $10^3$ ),  $C_6(t)$  and  $C_6^{fac}(t)$  (scaled by  $10^4$  and shifted up by 0.5 for clarity), and  $C_8(t)$  and  $C_8^{fac}(t)$  (scaled by  $10^5$  and shifted up by 1.0 for clarity).

hydrogen-oxygen-site angle  $\angle_{HO\alpha} \leq 30^\circ$ , where  $\alpha = A$  or  $B$  of the solute. Transiently broken H-bonds were also included in the H-bonding partition.

One can easily show that the Gaussian statistics approximation for the nonequilibrium response of a property  $A$  can be decomposed into various contributions to  $A$  by an additive partitioning. Specifically, we can write the fluorescence energy as  $\Delta E(t) = \Delta E_{HB}(t) + \Delta E_{NHB}(t)$ , where  $\Delta E_{HB}$  is the contribution to the energy gap due to solvent molecules H-bonded to the solute and  $\Delta E_{NHB}$  the contribution from all other, non-H-bonded, molecules. The Gaussian statistics approximation is distributive such that

$$\begin{aligned} C(t) \langle \delta \Delta E^2 \rangle_e &= \langle \delta \Delta E(t) \delta \Delta E(0) \rangle_e \\ &= \langle \delta \Delta E_{HB}(t) \delta \Delta E(0) \rangle_e + \langle \delta \Delta E_{NHB}(t) \delta \Delta E(0) \rangle_e \\ &\equiv C_{HB}(t) + C_{NHB}(t), \end{aligned} \quad (15)$$

where the Gaussian statistics approximation gives

$$\langle \Delta E_{HB}(t) \rangle_{ne} - \langle \Delta E_{HB}(\infty) \rangle_{ne} \simeq C_{HB}(t), \quad (16)$$

and similarly  $\langle \Delta E_{NHB}(t) \rangle_{ne} - \langle \Delta E_{NHB}(\infty) \rangle_{ne} \simeq C_{NHB}(t)$ ; note that  $C_{HB}(t)$  and  $C_{NHB}(t)$ , unlike  $C(t)$ , are not normalized. The correlation functions  $C_{HB}(t)$  and  $C_{NHB}(t)$  can be further divided into components using  $\Delta E(0) = \Delta E_{HB}(0) + \Delta E_{NHB}(0)$  in Eq. (15). However, as is clear from the above derivation, this is not necessary for the present purposes because  $C_{HB}(t)$  ( $C_{NHB}(t)$ ) itself represents the contribution of the H-bonded (non-H-bonded) molecules to the unnormalized TDSS. The same diagnostics applied in Sec. IV can be independently applied to  $C_{HB}(t)$  and  $C_{NHB}(t)$  by comparison of higher-order correlation functions to the corresponding factorized forms. In this way, the Gaussian statistics approximation can be independently tested in different partitions of the relevant coordinates in the same way as was done for the total fluorescence energy.

Note that the energy gaps,  $\Delta E_{HB}$  and  $\Delta E_{NHB}$ , involve only the Coulombic contributions to the energy difference of the solute-solvent interactions between the excited and ground states. The Lennard-Jones interactions are the same in the ground and excited states and thus cancel in the energy gap. Although there is a contribution to the energy gap from the long-range electrostatics due to the solute interaction with itself in different periodic images in the Ewald sum, we have not included it here.

The even-order time correlation functions for the H-bonded and non-H-bonded partitions are compared with their factorized forms in Figs. 10 and 11. In Sec. IV we found that for water, the factorized forms were in excellent agreement with the full correlation functions (see Figs. 5(a) and 6(a)). Thus it is not surprising that the factorization is accurate for both the fourth-order and sixth-order partitioned correlation functions shown in Figs. 10(a) and 11(a), respectively. On the other hand, the same is not true for the odd-order correlation function which is zero in the Gaussian statistics approximation but was shown to be non-zero at short times for the dipole-flip model in water (see Fig. 4(a)). The results for the corresponding partitioned correlation functions are presented in Fig. 12 and show that  $C_{3,NHB}(t) \simeq 0$  for all times and the deviations from Gaussian statistics are isolated in the H-bonding contributions represented in  $C_{3,HB}(t)$ . This indicates that the non-Gaussian character of the dipole-flip model in water is present primarily in an asymmetry of the energy gap distribution as shown in Fig. 7(a) and is associated solely with the waters H-bonded to the solute.

The results are different for the methanol solvent, which can be seen from the comparison of the fourth- and sixth-order correlation functions to their factorized forms in Figs. 10(b) and 11(b) for the H-bonding and the non-H-bonding partitions. As in the case of water, the non-H-bonding partitions obey Gaussian statistics based on the close agreement between

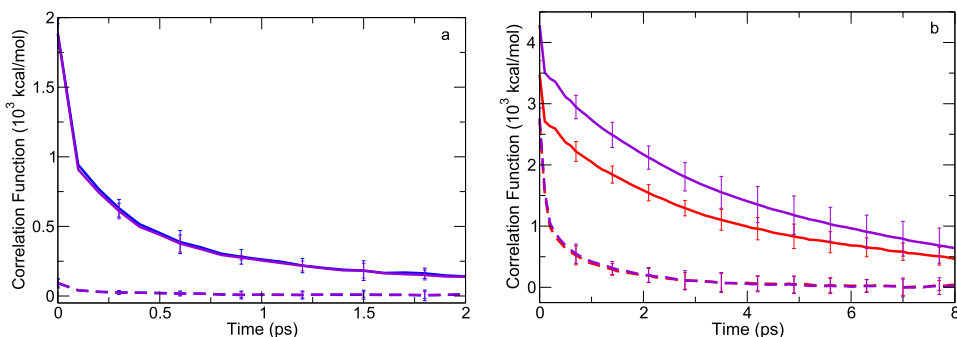


FIG. 10. Fourth-order correlation functions (blue and red lines) and Wick's factorizations (violet lines) are presented. Results are shown for the contributions from H-bonded solvent molecules,  $C_{4,HB}(t)$  and  $C_{4,HB}^{fac}(t)$ , (solid lines) and non-H-bonded molecules,  $C_{4,NHB}(t)$  and  $C_{4,NHB}^{fac}(t)$ , (dashed lines) for (a) water and (b) methanol.

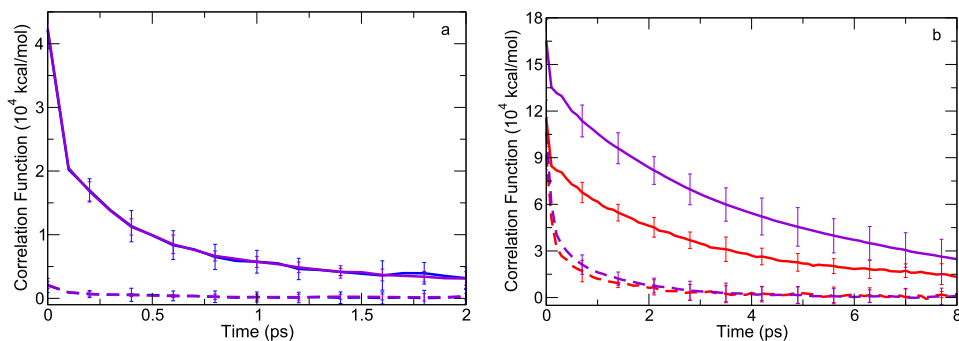
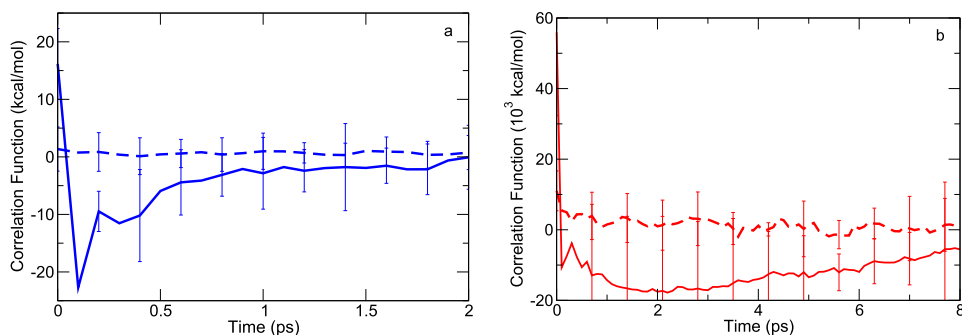


FIG. 11. Same as Fig. 10 but for the sixth-order correlation functions.

FIG. 12. The third-order correlation functions (that are zero in the Gaussian statistics approximation) are presented. Results are shown for the contributions from H-bonded solvent molecules,  $C_{3,HB}(t)$ , (solid lines) and non-H-bonded molecules,  $C_{3,NHB}(t)$ , (dashed lines) for (a) water (blue) and (b) methanol (red).

the correlation functions and their factorizations. However, in methanol, the H-bonding partition shows a significant deviation from the Gaussian statistics approximation in both the fourth- and sixth-order correlation functions. A similar behavior is observed for the partitioned third-order correlation functions shown in Fig. 12(b), for which the non-H-bonding contribution is effectively zero at all times (and thus consistent with Gaussian statistics) while the H-bonding correlation function is non-zero (albeit with significant error bars). Thus, it is clear that the non-Gaussian behavior is associated with the H-bonding of solvent molecules to the dipole-flip solute.

This analysis represents a general approach to identifying the origins of non-Gaussian statistics from only equilibrium simulations. It is interesting to note that in the case of water, the non-Gaussian behavior only arises from the odd-order correlation functions that are related to the asymmetry of the energy gap distribution. In contrast, for methanol, the non-Gaussian behavior is observed at all orders (even and odd), indicative of the stronger failure of the Gaussian statistics approximation that is also reflected in the energy gap distribution. This is presumably related to the smaller number of methanol molecules that contribute to the dipole-flip solute solvation through H-bonding, compared to the case of water, that is, the central limit theorem does not apply for the H-bonding solvent molecules in methanol but nearly does in water. It is important to note that the partitioning used here was relatively easy to find and motivated by the previous work on the dipole-flip model.<sup>6,13</sup> Indeed, we cannot be sure that the partitioning has completely identified the origin of the non-Gaussian statistics in that the H-bonding partition we have defined may still contain degrees-of-freedom that are Gaussian. The approach outlined here does not provide a prescription for how to determine the non-Gaussian coordinates but rather gives a straightforward method for evaluating hypotheses for such coordinates.

## VI. CAN NON-GAUSSIAN STATISTICS BE CORRECTED?

A key aim of investigating the origins of failures of the Gaussian statistics approximation is to ultimately develop approaches for correcting for non-Gaussian effects and thereby obviate the need for performing nonequilibrium MD simulations. A possible approach for this is offered by the deviations from Gaussian statistics manifested in the difference between the higher-order time correlation functions and their factorized forms.

As shown in Sec. II B, the Gaussian approximation is obtained by replacing the higher-order correlation functions with their factorized forms. Corrections to this approximation are most straightforwardly examined by considering the even- and odd-order correlation functions separately. The latter have no contribution in the Gaussian approximation and thus the corrections for odd-order terms can be included simply by addition of any non-zero higher-order correlation functions. Thus, the lowest-order correction to the nonequilibrium energy gap would be

$$\begin{aligned} \langle \Delta E \rangle_{ne}(t) - \langle \Delta E \rangle_e &\simeq \beta \langle \delta \Delta E(t) \delta \Delta E(0) \rangle \\ &+ \frac{\beta^2}{2} \langle \delta \Delta E(t) \delta \Delta E(0)^2 \rangle \\ &\equiv C_{3,corr}(t), \end{aligned} \quad (17)$$

where the subscript “3,corr” indicates that this correlation function represents a correction up to the third-order. This corrected result has been applied to both the water and methanol solvent cases and the results are shown (dashed violet lines) in Fig. 13, plotted in normalized form. In the case of water, the effect of the correction is to predict a significantly more rapid decay in the energy gap compared to the Gaussian statistics approximation alone followed by a nearly constant but non-zero value. Despite the fact that the correction is in better

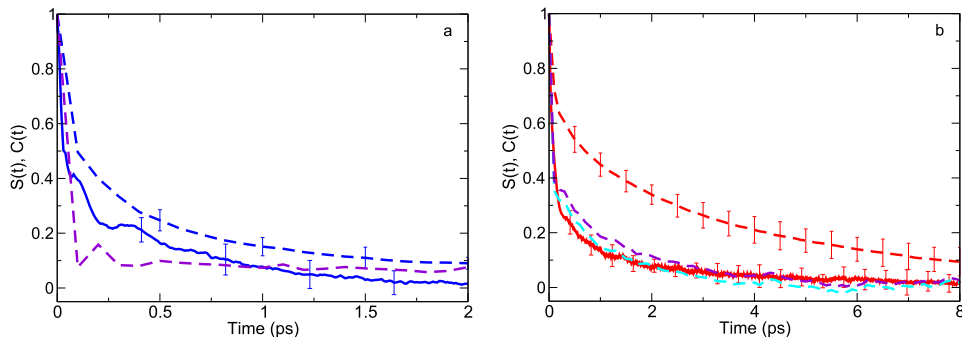


FIG. 13. The normalized TDSS,  $S(t)$ , (solid line) is compared to the Gaussian statistics approximation,  $C(t)$ , (dashed line) and is shown for (a) water (blue lines) and (b) methanol (red lines). They are compared to the corrections to the Gaussian statistics approximation,  $C_{3,corr}(t)$  (dashed violet lines) and the averaged correction,  $C_{3-6,corr}(t)$  (dashed cyan line).

quantitative agreement with the full, nonequilibrium TDSS at most times compared to the Gaussian statistics result, it more poorly predicts the shape of  $S(t)$ . In contrast, for the methanol solvent, the  $C_{3,corr}(t)$  corrected result is in significantly better qualitative and quantitative agreement with the nonequilibrium TDSS compared to the Gaussian statistics approximation. In particular, it exhibits a larger amplitude for the initial, inertial component as well as a faster decay at longer times. We note that the higher-order correlation functions are obtained with essentially no additional effort above that required for the Gaussian statistics approximation. However, a key issue with this correction is the large statistical errors associated with the  $C_3(t)$  correlation function it is based on as illustrated in Fig. 4.

It is less clear how to correct for deviations from Gaussian statistics in the even-order correlation functions. One approach is to note that the Gaussian approximation assumes

$$\langle \delta A(t) \delta \Delta E(0)^{2n+1} \rangle_e = \frac{(2n+1)!}{n! 2^n} \langle \delta \Delta A(t) \delta \Delta E(0) \rangle \langle \delta \Delta E^2 \rangle^n. \quad (18)$$

Then an approximation to the Gaussian statistics result for a given  $n$  can be obtained as

$$\langle \delta \Delta A(t) \delta \Delta E(0) \rangle_{2n+2} = \frac{\langle \delta A(t) \delta \Delta E(0)^{2n+1} \rangle_e}{(2n+1)! \langle \delta \Delta E^2 \rangle^n / (n! 2^n)}. \quad (19)$$

Clearly, for  $n = 0$  this gives the Gaussian statistics approximation,  $C(t)$ , but for  $n > 0$  it yields a correction to this  $C(t)$  based on the time-dependence of the higher-order correlation function. Then a reasonable scheme for correcting the Gaussian statistics approximation is to average the estimates of this correlation function from the different orders, e.g.,

$$\begin{aligned} \langle \Delta A \rangle_{ne}(t) - \langle \Delta A \rangle_e &\simeq \frac{1}{3} \sum_{n=0}^2 \beta \langle \delta \Delta A(t) \delta \Delta E(0) \rangle_{2n+2} \\ &+ \frac{\beta^2}{2} \langle \delta \Delta A(t) \delta \Delta E(0)^2 \rangle \equiv C_{3-6,corr}(t). \end{aligned} \quad (20)$$

This result is shown for methanol solvent in Fig. 13(b) and is in agreement with the full TDSS result for all times (within error bars). It represents a slight improvement over the  $C_{3,corr}(t)$  correction, particularly for times less than  $\sim 3$  ps. However, it is clear that the corrections for the even-order terms are less significant than the first odd-order correction represented in  $C_{3,corr}(t)$ . For water, the higher-order, even correlation functions are in good agreement with the Gaussian approximations so that  $C_{3-6,corr}(t) = C_{3,corr}(t)$  within our calculations.

Finally, we note that an alternative approach to correct the Gaussian statistics approximation is to add the cumulants, i.e., the differences between the higher-order correlation functions and their factorized forms, to the Gaussian statistics approximation result.<sup>29</sup> If, however, the resulting series of higher-order terms does not converge (as has been the case in previous studies<sup>8,19</sup>), then this construction will fail at least theoretically if not also practically. Indeed, this is what we have observed for both the water and methanol solvents for the dipole-flip solute.

## VII. RELATION TO OTHER EXAMPLES OF NONLINEAR RESPONSE

The dynamic linear response and Gaussian statistics approximations are used frequently, particularly in applications to solvation dynamics. Because of the general success of these approaches, the cases where they fail have drawn significant theoretical and experimental interest.

The dipole-flip model, in particular, has been studied by others following the early work by Ladanyi and co-workers establishing it as a key example of nonlinear response. Geissler and Chandler computed the nonequilibrium distribution of the excited- and ground-state energy gaps for the dipole-flip model in water and found that the statistics of the energy gap are Gaussian but are nonstationary.<sup>6</sup> These nonstationary Gaussian statistics were attributed to the change in the H-bond network due to the sensitivity of dielectric response on boundary conditions. As the solvent molecules rearrange themselves upon solute excitation, boundary conditions of the solvent region change, influencing the solvent response. These results also indicate that the non-Gaussian statistics result from some alternative, non-Gaussian degree(s) of freedom, which cause the overall energy gap to be non-Gaussian and result in nonlinear response. The present work is consistent with these results, particularly in explicitly isolating the role of the H-bonding network, through the projection of the energy gap fluctuations (Sec. V), as the few degrees of freedom responsible for the non-Gaussian behavior.

Breakdowns in Gaussian statistics for solvation dynamics have also been reported in a number of other systems. Schwartz and co-workers used MD to study a single atom capable of carrying a charge dissolved in water.<sup>27,28</sup> They found that solute size changes (e.g., induced by ionization) cause a breakdown in linear response. In particular, both the dynamic linear-response and Gaussian statistics approximations give incorrect results in the inertial regime of relaxation. This is associated with the size change of the solute, which results in a restructuring of

the H-bond network due to inertial motion of proximal OH groups.

Systems of the solvated electron serve as another class of non-Gaussian solvation dynamics. Rossky and co-workers have shown that the Gaussian statistics approximation fails to describe the solvation dynamics of an excess electron in methanol.<sup>15,30</sup> They conducted mixed quantum-classical MD (MQC-MD) simulations and computed the normalized TDSS from nonequilibrium simulations as well as linear response correlation functions from equilibrium simulations in two states: the initial state (the neat solvent) and the final state (the solvated electron system). Both linear response correlation functions differed from the TDSS, which was, like the previous example, attributed to large structural changes in the H-bond network.<sup>15</sup> Upon excitation, mechanical forces from localization of the electron push H-bonds away, so at short times (<500 fs) the number of H-bonds around the electron decreases.

Bedard-Hearn *et al.* used MQC-MD to study the hydrated electron to show how their previous findings on Gaussian statistics breakdowns from solute size changes apply to the system.<sup>31</sup> Using Steele theory,<sup>32</sup> they found that the Gaussian statistics breakdown for size changes in the hydrated electron from excitation occurs from significant deviations in the solvent librations that interfere with the electron *p*-like wavefunction. These studies on solute size changes are possible explanations for the breakdown of Marcus theory (i.e., linear response) for reactions of cation transition metals with the hydrated electron.<sup>33</sup> Specifically, Kanjana *et al.* studied the kinetics of hydrated electron reactions with a number of dication transition metals using UV-Vis absorption spectroscopy, which showed deviations from Marcus theory.<sup>33</sup> That Marcus theory fails for these reactions is perplexing; the reactions are Arrhenius, are not diffusion limited, and are thermally activated, all of which point to the validity of Marcus theory for this system.<sup>34</sup>

Laird and Thompson studied TDF and dye molecule diffusion in nanoconfined solvents.<sup>8</sup> The results showed that the Gaussian statistics approximation agrees reasonably with nonequilibrium MD results for the TDSS of a nanoconfined dye molecule (whereas dynamic linear response does not) but more poorly for the dye molecule position. Additionally, in the case of nanoconfined solvation dynamics, the assumption invoked by static linear response (neglect of higher-order terms) is not justified. In fact, the higher-order correlation functions not only do not decrease but also increase by orders of magnitude with each term in the series. The energy gap was found to obey Gaussian statistics for a fixed dye molecule position but with position-dependent properties leading to the non-Gaussian behavior. This is consistent with the nonstationary Gaussian statistics observed by Geissler and Chandler.<sup>6</sup>

Li and co-workers have studied the validity of linear response for TDF of *Staphylococcus* Nuclease (SNase).<sup>19,35</sup> They performed equilibrium and nonequilibrium simulations and found that the Gaussian statistics approximation proved more favorable than dynamic linear response based on comparison with nonequilibrium simulations.<sup>35</sup> Furthermore, they showed that the static linear-response approximation was not valid for the system they studied by computing higher-order

correlation functions, which also increased with increasing order.<sup>19</sup>

These examples should not be used to infer that breakdowns of the Gaussian statistics approximation are limited to systems where H-bonding is present. There are, indeed, many systems displaying nonlinear response in aprotic solvents. One example is the solvation of neutral sodium atoms in tetrahydrofuran (THF). Schwartz and co-workers have investigated the solvation of a neutral Na atom in liquid THF using two approaches: one in which a sodium atom was produced by removing an electron from Na<sup>+</sup> and the other in which an electron was added to Na<sup>+</sup>.<sup>36</sup> Were the system to obey Gaussian statistics, the dynamics of both experiments should exhibit the same time scales because they occur on different sides of the same *harmonic* free energy surface. Bragg *et al.* however, observed that the solvation dynamics of the sodium atom prepared from Na<sup>+</sup> were much slower than that of the sodium atom prepared from Na<sup>+</sup>.<sup>37</sup> Bedard-Hearn *et al.* used MD to study sodium deionizing in THF; however, instead of an explicit breakdown in linear response resulting from comparison of *C(t)* and *S(t)*, they found that the breakdowns were hidden in the mechanism of relaxation for each property, that is, *S(t)* and *C(t)* were in good agreement but the components for each differed.<sup>26</sup> Using Steele theory, they analyzed the contributions to *C(t)* and *S(t)* from the Coulomb and Lennard-Jones partitions of the energy and found that the relaxation from each contribution differed between the nonequilibrium value and linear-response approximation.

Another example is the relaxation of rotational energy. Moskun *et al.* studied the energy dissipation in highly rotationally excited CN dissolved in liquid argon.<sup>38</sup> The rotational excitation in the gas phase would be observed for an extended period compared to the liquid phase in which, as linear response would predict, the rotational energy would rapidly dissipate. They observed, however, that the dynamics went against the intuition of both phases: initially a linear-response-like dissipation was observed followed by a sudden switch to gas phase-like dynamics. This switch corresponds to a molecular event that cannot be predicted by linear response theory. In fact, the rotations of CN produce a distinct cavity in the liquid at short times, which extends the coherence of the rotations. This was confirmed by Tao and Stratt in which they derived a linear-response approximation to the normalized response function and computed both linear-response correlation functions and nonequilibrium MD simulations.<sup>7</sup> Similar findings were reported for simulations of OH rotors.<sup>39</sup> An analogous result was observed in a combined experimental and computational study of vibrational relaxation of HCN formed as the product of a chemical reaction.<sup>40</sup>

In each of the aforementioned examples of a breakdown in linear response, i.e., non-Gaussian statistics, specific molecular rearrangements of the solvent are reported that are strongly coupled to the solvation dynamics. It is from these molecular events—cavitation of a solvated rotor, H-bond restructuring, solute cavity expansion or contraction, etc.—that a breakdown of the Gaussian statistics approximation can be attributed. A key commonality is that each involves only a small number of degrees-of-freedom preventing Gaussian behavior via the central limit theorem,<sup>1</sup> as has been previously suggested by



Schwartz and co-workers.<sup>27,28</sup> This is supported in the case of the dipole-flip model by the more nonlinear response observed in methanol compared to water (Fig. 3) and observed by Fonseca and Ladanyi for a smaller solute compared to a larger one in methanol.<sup>12</sup> That is, the response is less linear when fewer solvent molecules are involved in the molecular rearrangements, which for the dipole-flip model are associated with H-bonding to the solute and influenced by both the solvent and solute sizes. Similarly, the nonlinear response observed for the solvated electron<sup>15,30</sup> and solute size changes<sup>27,28</sup> is associated with rearrangements of the liquid structure in the first solvation shell, i.e., involving a small number of molecules.

Importantly, Tao and Stratt noted that changes in liquid structure were a common feature in many of the studies cited above and further suggested “. . .that the relative slowness of structural relaxation could be an important prerequisite for linear-response breakdowns.”<sup>7</sup> The observation of non-Gaussian structural rearrangements occurring on slower time scales has been observed in the case of rotational relaxation, where it is associated with the formation of a solute cavity<sup>7,38,39</sup> as well as for vibrational relaxation which exhibits an analogous behavior.<sup>40</sup> In the case of solvation dynamics, a similar picture was posited by Turi *et al.* for the solvated electron in methanol, again associated with a change in solvent structure around the electron.<sup>15</sup> For the dipole-flip model, the picture is more complicated. The linear response breakdown tends to appear at longer times, after the inertial dynamics, as was pointed out by Fonseca and Ladanyi,<sup>12</sup> who noted “. . .both fast and slow solvation dynamics have substantial contributions from structural rearrangements within the first solvation shell. After the perturbation is applied, the destruction of an existing local structure is rapid, but a build-up of a new one is much slower.” The present analysis shows deviations from Gaussian statistics at all time scales, which can be primarily observed in the odd-order correlation functions for both water and methanol (Fig. 4) but also in the even-order correlation functions for methanol (Figs. 5 and 6). However, it is also the case that the H-bonding rearrangements responsible for the nonlinear response in this model do occur on longer time scales. Thus, while the current results appear to suggest that structural relaxation need not be slow on the time scale of the response to lead to nonlinear response, this topic deserves additional examination for other systems and from other perspectives.

## VIII. CONCLUSION

The static linear response and Gaussian statistics approximations have been outlined and investigated in the context of the dipole-flip model. Where the static linear response approximation assumes all higher-order correlation functions are negligible, the Gaussian statistics approximation factorizes them into lower-order correlation functions based on Wick's theorem and, upon summation, gives the same mathematical expression as static linear response.

For a dipolar solute immersed in water or methanol, the approximation fails, confirming prior knowledge on the subject,<sup>6,12,13</sup> but it has also been shown that the physical approximation made in static linear response is unjustified:

the higher-order correlation functions are not negligible and, in fact, grow by orders of magnitude. Thus, what is often referred to as the static linear response approximation is really the assumption of Gaussian statistics.

The Gaussian statistics approximation assumes that the relevant dynamical variable, e.g., the energy gap in solvation dynamics, is a Gaussian random variable. An accepted approach for evaluating this approximation is to compare it with full nonequilibrium MD simulations. This is rarely done, however, because the nonequilibrium simulations are significantly more computationally expensive than the equilibrium MD simulations required for the Gaussian statistics result. Thus, we have addressed the question *can non-Gaussian statistics be predicted by relying only on equilibrium simulations?* The answer is yes. Specifically, Gaussian statistics assumes that higher-order time correlation functions can be factorized into two-point time correlation functions; both the higher-order and factorized time correlation functions can be evaluated from the same equilibrium MD simulation. The dipole-flip model in water and methanol exhibits a breakdown of the Gaussian statistics approximation based on comparison with nonequilibrium MD results. This breakdown is also reflected in differences between the equilibrium higher-order and factorized time correlation functions, which thus offers a method for the evaluation of the Gaussian statistics approximation. This complements a direct examination of the distribution of the energy gap, which for simulations in methanol is clearly non-Gaussian, but for water is nearly Gaussian.

When there is a breakdown of the Gaussian statistics approximation it is natural to ask *what is the origin of the non-Gaussian statistics?* It has been proposed that the nonlinear response for the dipole-flip model in H-bonding solvents can be attributed to the solvent-solute H-bonds<sup>6,12,13</sup> which must rearrange upon the electronic excitation that changes the identity of the dipole-flip solute atom that can act as an H-bond acceptor. Here, we have shown that the Gaussian statistics approximation agrees with nonequilibrium dynamics for a non-H-bonding solvent, acetonitrile, indicating that the lack of H-bonding recovers agreement between the Gaussian statistics approximation and nonequilibrium dynamics.

An additional, more direct, test to the non-Gaussian nature of the solute-solvent hydrogen bonds was performed in which the equilibrium trajectories were partitioned into a solvent-solute H-bonding component and a bath (non-H-bonding) component. The same comparisons of higher-order correlation functions and their factorizations that, as discussed above, can be used to identify non-Gaussian statistics can then be applied to each partition separately. This analysis clearly shows that solvent-solute H-bonds are responsible for the breakdowns of the Gaussian statistics approximation for methanol and water. Moreover, it shows how the origins of the non-Gaussian behavior can be identified and isolated in general through such partitioning.

Ideally, one would like to find a way to improve upon the Gaussian statistics approximation when necessary. Thus, we asked *can the Gaussian statistics approximation be corrected to describe nonlinear response?* This is a more challenging problem than identifying the non-Gaussian behavior and its origins. Nevertheless, the results indicate that there is some

promise in using the information about the non-Gaussian character of the dynamics embodied in the higher-order time correlation functions to improve upon the Gaussian approximation. The results for methanol, where the non-Gaussian behavior is more prominent, are encouraging, while those for water are less so. This is an area that is ripe for future exploration.

We have placed the present study of the dipole-flip model in various solvents in the context of other examples of breakdowns of the Gaussian statistics approximation throughout the literature. In each case discussed, one can reasonably attribute the origin of this breakdown to some specific solvent mechanism, involving a small number of relevant coordinates, that occurs in the nonequilibrium dynamics that does not occur at equilibrium. Though there have been no tests of these systems based on the assumptions invoked in the static linear response and Gaussian statistics approximations such as that presented here, they represent both model and real, experimentally accessible systems to which the present analysis could be applied to advantage.

## ACKNOWLEDGMENTS

Support for this work was provided by the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy (Grant No. DE-FG02-05ER15708). The calculations presented were performed at the University of Kansas Advanced Compute Facility. W.H.T. thanks Professor Branka M. Ladanyi for helpful comments in the early stages of this work. The authors thank Professor Brian B. Laird for many useful discussions.

- <sup>1</sup>R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II* (Springer-Verlag, New York, 1978).
- <sup>2</sup>D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).
- <sup>3</sup>R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, New York, 2001).
- <sup>4</sup>A. Nitzan, *Chemical Dynamics in Condensed Phases: Relaxation, Transfer, and Reactions in Condensed Molecular Systems* (Oxford University Press, New York, 2006).
- <sup>5</sup>B. B. Laird and W. H. Thompson, *J. Chem. Phys.* **126**, 211104 (2007).
- <sup>6</sup>P. L. Geissler and D. Chandler, *J. Chem. Phys.* **113**, 9759 (2000).
- <sup>7</sup>G. H. Tao and R. M. Stratt, *J. Chem. Phys.* **125**, 114501 (2006).

- <sup>8</sup>B. B. Laird and W. H. Thompson, *J. Chem. Phys.* **135**, 084511 (2011).
- <sup>9</sup>Throughout the paper we use terminology that distinguishes between the different mathematical approaches for obtaining linear response approximations. In particular, we adopt the nomenclature of, e.g., Refs. 3 and 4, who use “dynamic” and “static” linear response to denote approximations focused on the time-evolution and the Boltzmann distribution, respectively. These two approaches assume that the perturbation leading to the response is weak. In contrast, the Gaussian statistics approximation applies even for large perturbations by only assuming that the relevant variables are Gaussian. While they are based on different physical and mathematical assumptions—which are a key focus of this work—they are all linear response theories.
- <sup>10</sup>M. Maroncelli and G. R. Fleming, *J. Chem. Phys.* **89**, 5044 (1988).
- <sup>11</sup>T. Fonseca and B. M. Ladanyi, *J. Phys. Chem.* **95**, 2116 (1991).
- <sup>12</sup>T. Fonseca and B. M. Ladanyi, *J. Mol. Liq.* **60**, 1 (1994).
- <sup>13</sup>M. S. Skaf and B. M. Ladanyi, *J. Phys. Chem.* **100**, 18258 (1996).
- <sup>14</sup>R. M. Stratt and M. Maroncelli, *J. Phys. Chem.* **100**, 12981 (1996).
- <sup>15</sup>L. Turi, P. Mináry, and P. J. Rossky, *Chem. Phys. Lett.* **316**, 465 (2000).
- <sup>16</sup>R. A. Marcus, *Annu. Rev. Phys. Chem.* **15**, 155 (1964).
- <sup>17</sup>L. Isserlis, *Biometrika* **12**, 134 (1918).
- <sup>18</sup>G. C. Wick, *Phys. Rev.* **80**, 268 (1950).
- <sup>19</sup>T. Li and R. Kumar, *J. Chem. Phys.* **143**, 174501 (2015).
- <sup>20</sup>H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *J. Phys. Chem.* **91**, 6269 (1987).
- <sup>21</sup>W. Jorgensen, D. Maxwell, and J. TiradoRives, *J. Am. Chem. Soc.* **118**, 11225 (1996).
- <sup>22</sup>A. M. Nikitin and A. P. Lyubartsev, *J. Comput. Chem.* **28**, 2020 (2007).
- <sup>23</sup>G. Ciccotti and J. P. Ryckaert, *Comput. Phys. Rep.* **4**, 346 (1986).
- <sup>24</sup>B. M. Ladanyi and M. S. Skaf, *Annu. Rev. Phys. Chem.* **44**, 335 (1993).
- <sup>25</sup>A. A. Vartia, K. R. Mitchell-Koch, G. Stirnemann, D. Laage, and W. H. Thompson, *J. Phys. Chem. B* **115**, 12173 (2011).
- <sup>26</sup>M. J. Bedard-Hearn, R. E. Larsen, and B. J. Schwartz, *J. Phys. Chem. A* **107**, 4773 (2003).
- <sup>27</sup>V. Tran and B. J. Schwartz, *J. Phys. Chem. B* **103**, 5570 (1999).
- <sup>28</sup>D. Aherne, V. Tran, and B. J. Schwartz, *J. Phys. Chem. B* **104**, 5382 (2000).
- <sup>29</sup>A. A. Mosyak, O. V. Prezhdo, and P. J. Rossky, *J. Chem. Phys.* **109**, 6390 (1998).
- <sup>30</sup>P. Mináry, L. Turi, and P. J. Rossky, *J. Chem. Phys.* **110**, 10953 (1999).
- <sup>31</sup>M. J. Bedard-Hearn, R. E. Larsen, and B. J. Schwartz, *Phys. Rev. Lett.* **97**, 130403 (2006).
- <sup>32</sup>W. A. Steele, *Mol. Phys.* **61**, 1031 (1987).
- <sup>33</sup>K. Kanjana, B. Courtin, A. MacConnell, and D. M. Bartels, *J. Phys. Chem. A* **119**, 11094 (2015).
- <sup>34</sup>M. Anbar, Z. B. Alfassi, and H. Bregman-Reisler, *J. Am. Chem. Soc.* **89**, 1263 (1967).
- <sup>35</sup>T. Li, *J. Phys. Chem. B* **118**, 12952 (2014).
- <sup>36</sup>M. C. Larsen and B. J. Schwartz, *J. Chem. Phys.* **131**, 154506 (2009).
- <sup>37</sup>A. E. Bragg, M. C. Cavanagh, and B. J. Schwarz, *Science* **321**, 1817 (2008).
- <sup>38</sup>A. C. Moskun, A. E. Jailaubekov, S. E. Bradforth, G. H. Tao, and R. M. Stratt, *Science* **311**, 1907 (2006).
- <sup>39</sup>B. J. Savitsky and R. M. Stratt, *J. Phys. Chem. B* **112**, 13326 (2008).
- <sup>40</sup>D. R. Glowacki, R. A. Rose, S. J. Greaves, A. J. Orr-Ewing, and J. N. Harvey, *Nat. Chem.* **3**, 850 (2011).