Role of sampling in evaluating classical time autocorrelation functions

Tomáš Zimmermann and Jiří Vaníček

Additional information on J. Chem. Phys.

Role of sampling in evaluating classical time autocorrelation functions

Tomáš Zimmermann[a)] and Jiří Vaníček[b)]

Laboratory of Theoretical Physical Chemistry, Institut des Sciences et Ingénierie Chimiques,
Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

We analyze how the choice of the sampling weight affects efficiency of the Monte Carlo evaluation of classical time autocorrelation functions. Assuming uncorrelated sampling or sampling with constant correlation length, we propose a sampling weight for which the number of trajectories needed for convergence is independent of the correlated quantity, dimensionality, dynamics, and phase-space density. By contrast, it is shown that the computational cost of the "standard" algorithm sampling from the phase-space density may scale exponentially with the number of degrees of freedom. Yet, for the stationary Gaussian distribution of harmonic systems and for the autocorrelation function of a linear function of phase-space coordinates, the computational cost of this standard algorithm is also independent of dimensionality. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4820420]

INTRODUCTION: TIME CORRELATION FUNCTIONS

Many dynamical properties of stationary systems as well as the response of such systems to weak perturbations can be inferred from time autocorrelation functions.[1,2] Examples include the optical absorption line shapes computed from the dipole time autocorrelation function, the diffusion coefficient computed from the velocity time autocorrelation function, and various relaxation properties.[3] More general time correlation functions are in fact the principal ingredients of semiclassical[4,5] and path-integral[6–11] calculations of quantum dynamical properties. Trajectory-based methods for computing time correlation functions, however, may become too expensive in many-dimensional systems. Yet, dimensionality-independent algorithms have been found for special correlation functions, such as classical[12] and semiclassical[13] fidelity.[14] Motivated by the importance of correlation functions in many areas of physics, here we explore how these functions can be computed more efficiently in general. In particular, we propose a sampling weight for which the number of trajectories needed for convergence of any classical normalized time autocorrelation function is independent of dimensionality both of the phase space and of the studied observable.

Quantum mechanically, the unnormalized time autocorrelation function $C_u^{QM}(t)$ of a vector operator $\hat{\mathbf{A}}$ is defined as

$$C_u^{QM}(t) = \text{Tr}(\hat{\rho}^0 \hat{\mathbf{A}}^0 \cdot \hat{\mathbf{A}}^t), \tag{1}$$

where $\hat{\rho}^0$ is the density operator, $\hat{\mathbf{A}}^0$ is the operator evaluated at time $t = 0$, $\hat{\mathbf{A}}^t = e^{i\hat{H}t/\hbar}\hat{\mathbf{A}}e^{-i\hat{H}t/\hbar}$ is $\hat{\mathbf{A}}$ evolved with Hamiltonian $\hat{H}$ for time t, and subscript "u" emphasizes that the correlation function is not normalized. In the classical limit, correlation function (1) becomes

$$C_u(t) = h^{-D} \int dx \rho^0(x) \mathbf{A}^0(x) \cdot \mathbf{A}^t(x), \tag{2}$$

where $x := (q, p)$ is the $2D$-dimensional phase-space coordinate, $\rho^0(x)$ is the initial phase-space density, $\mathbf{A}^0(x)$ is the classical observable $\mathbf{A}$ evaluated at time $t = 0$, and $\mathbf{A}^t(x)$ $= e^{-\hat{L}t}\mathbf{A}^0(x)$ is this function $\mathbf{A}$ evolved classically for time t with the Liouville operator $\hat{L} = \{H, \cdot\}$. Note that besides a three-dimensional vector (such as the molecular dipole $\boldsymbol{\mu}$), $\mathbf{A}$ can also be a scalar (A) or a higher-dimensional phase-space vector. Also, $\mathbf{A}$ may be a function of only D' phase-space coordinates, where $D' < 2D$ or even $D' \ll 2D$. This occurs if one examines a property of a single molecule surrounded by an environment, an example being the velocity autocorrelation function of a molecule embedded in a solvent. Since the shape of the autocorrelation function is typically more interesting than its overall magnitude,[15] one often computes a normalized time autocorrelation function

$$C(t) = C_u(t)/C_u(0). \tag{3}$$

ALGORITHMS

Most methods for evaluating Eqs. (2) and (3) in many dimensions employ classical trajectories. Two general approaches are currently used:[16] (i) the direct approach in which initial conditions for many trajectories are sampled from the stationary distribution ρ, and (ii) the single-trajectory approach in which the desired autocorrelation function is computed as an average of many correlation functions computed using a single trajectory either as an average of correlation functions initiated at different times or using the Fourier transform and Wiener-Khinchin theorem. The direct approach is more general and does not require the ergodicity of the time evolution, whereas the single trajectory approach is often simpler, as it avoids explicit sampling of ρ.

Here we explore modifications of the direct approach using generalized sampling weights and start by expressing the correlation function (2) in terms of trajectories,

$$C_u(t) = h^{-D} \int dx^0 \rho(x^0) \mathbf{A}(x^0) \cdot \mathbf{A}(x^{-t}), \tag{4}$$

a)Electronic mail: tomas.zimmermann@epfl.ch
b)Electronic mail: jiri.vanicek@epfl.ch

139, 104105-1

where $x^t := \Phi^t(x^0)$ is the phase-space coordinate at time t of a trajectory of the Hamiltonian flow Φ^t with initial condition x^0. We further rewrite Eq. (4) in a form suitable for Monte Carlo (MC) evaluation, i.e., as an average

$$\langle E(x^0, t)\rangle_W := \frac{\int dx^0 E(x^0, t) W(x^0)}{\int dx^0 W(x^0)}, \tag{5}$$

where the positive definite function W is the sampling weight and E is the estimator. In the Monte Carlo method, average (5) is evaluated numerically as an average

$$E_W(N, t) = N^{-1} \sum_{j=1}^{N} E(x_j^0, t) \tag{6}$$

over N trajectories whose initial conditions x_j^0 are sampled from the weight W.

The convergence rate of the sum (6) usually depends on W. Among many possible weights W, the form of Eq. (4) immediately suggests the following three: $W(x) = \rho(x)$, $\rho(x)|\mathbf{A}(x)|$, and $\rho(x)\mathbf{A}(x)^2$. These three weights lead to three different algorithms, which may be written as

$$C_{u,W}(t) = I_W \langle E_W(x^0, t)\rangle_W, \tag{7}$$

where $I_W := h^{-D} \int W(x) dx$ denotes the norm of W and the estimators are

$$E_\rho(x^0, t) = \mathbf{A}(x^0) \cdot \mathbf{A}(x^{-t}), \tag{8}$$

$$E_{\rho|\mathbf{A}|}(x^0, t) = \frac{\mathbf{A}(x^0) \cdot \mathbf{A}(x^{-t})}{|\mathbf{A}(x^0)|}, \tag{9}$$

$$E_{\rho\mathbf{A}^2}(x^0, t) = \frac{\mathbf{A}(x^0) \cdot \mathbf{A}(x^{-t})}{|\mathbf{A}(x^0)|^2}. \tag{10}$$

Substitution of Eq. (7) into the definition (3) yields a Monte Carlo prescription for the normalized correlation function:

$$C_W(t) = \frac{\langle E_W(x^0, t)\rangle_W}{\langle E_W(x^0, 0)\rangle_W}. \tag{11}$$

Since $E_{\rho\mathbf{A}^2}(x^0, 0) = 1$, no normalization is needed for the $\rho\mathbf{A}^2$ algorithm. The two averages in Eq. (11) may be evaluated with two independent Monte Carlo simulations, or—as we do—in a single Monte Carlo simulation, which is faster and normalizes both $C_\rho(0)$ and $C_{\rho|\mathbf{A}|}(0)$ exactly.

STATISTICAL ERRORS

The three algorithms differ by the sampling weight W used, and hence also by the estimator E_W. The cost of all three algorithms is $O(c\frac{t}{\Delta t}N)$, where N is the number of trajectories, Δt the time step used, and c the combined cost of a single evaluation of the force (needed for the dynamics) and of the estimator E_W. Usually, the cost of evaluating the estimator is or can be made negligible to that of evaluating the force. Therefore, the costs of the algorithms differ mainly in the number N of trajectories needed to achieve a desired precision (i.e., discretization error) σ_{discr}.

Alternatively, the algorithms can be compared by evaluating the discretization errors $\sigma_{discr,W}$ resulting from a given

number N of trajectories. For an unbiased estimator, the discretization error σ_{discr} is equal to the statistical error σ_W, where $\sigma_W(N, t)^2 = \overline{C_W(N, t)^2} - \overline{C_W(N, t)}^2$ and the overline denotes an average over an infinite number of simulations with different sets of N trajectories. Assuming for now that the N trajectories are uncorrelated, one can show that the error of the unnormalized $C_u(t)$ satisfies

$$\sigma_{u,W}(N, t)^2 = \frac{I_W^2}{N}\big[\langle E_W(x^0, t)^2\rangle_W - \langle E_W(x^0, t)\rangle_W^2\big]. \tag{12}$$

For $W = \rho\mathbf{A}^2$, the error of normalized $C(t)$ satisfies a relation obtained by removing factors of I_W from Eq. (12). Statistical errors of algorithms with weights ρ and $\rho|\mathbf{A}|$, which must be normalized according to Eq. (11), are found from the formula for the statistical error of a ratio of random variables:

$$\left(\frac{\sigma_{S/T}}{\bar{S}/\bar{T}}\right)^2 = \left(\frac{\sigma_S}{\bar{S}}\right)^2 + \left(\frac{\sigma_T}{\bar{T}}\right)^2 - 2\frac{\overline{ST} - \bar{S}\bar{T}}{\bar{S}\bar{T}}. \tag{13}$$

In our case, $S = C_{u,W}(N, t)$ and $T = C_{u,W}(N, 0)$. Realizing that $\overline{C_{u,W}(N, t)} = C_u(t)$, we obtain the following general expression for the statistical errors of the three algorithms:

$$\sigma_W(N, t)^2 = \frac{1}{N d_W}[a_W C(t)^2 - 2b_W C(t) + c_W], \tag{14}$$

where $a_W = \langle|\mathbf{A}^0|^4\rho/W\rangle_\rho$, $b_W = \langle|\mathbf{A}^0|^2(\mathbf{A}^0 \cdot \mathbf{A}^t)\rho/W\rangle_\rho$, $c_W = \langle(\mathbf{A}^0 \cdot \mathbf{A}^t)^2\rho/W\rangle_\rho$, $d_W = \langle|\mathbf{A}^0|^2\rho/W\rangle_\rho^2$, and an abbreviated notation $\mathbf{A}^t := \mathbf{A}(x^{-t})$ was used. The special cases are obtained by replacing W with ρ, $\rho|\mathbf{A}|$, or $\rho\mathbf{A}^2$.

For $W = \rho\mathbf{A}^2$, the coefficients can be rearranged as $a_{\rho\mathbf{A}^2} = -d_{\rho\mathbf{A}^2}$, $b_{\rho\mathbf{A}^2} = 0$, $c_{\rho\mathbf{A}^2} = \langle(\mathbf{A}^0 \cdot \mathbf{A}^t)^2/|\mathbf{A}^0|^2\rangle_\rho$, and $d_{\rho\mathbf{A}^2} = \langle|\mathbf{A}^0|^2\rangle_\rho$. Using the Cauchy-Schwarz inequality $(\mathbf{A}^0 \cdot \mathbf{A}^t)^2 \leq |\mathbf{A}^0|^2|\mathbf{A}^t|^2$ in the expression for $c_{\rho\mathbf{A}^2}$ and the fact that for stationary distributions $\langle|\mathbf{A}^0|^2\rangle_W = \langle|\mathbf{A}^t|^2\rangle_W$, we find that $c_{\rho\mathbf{A}^2} \leq \langle\mathbf{A}(x^{-t})^2\rangle_\rho = d_{\rho\mathbf{A}^2}$. Hence, for the weight $\rho\mathbf{A}^2$ the upper bound for the statistical error depends only on N and the value $C(t)$ of the autocorrelation function:

$$\sigma_{\rho\mathbf{A}^2}^2(N, t) \leq N^{-1}[1 - C(t)^2]. \tag{15}$$

In particular, the error does not explicitly depend on the dimensionality D of the system, dimensionality D' and nature of the observable $\mathbf{A}$, chaoticity of the dynamics, or time t. This remarkable fact is the main thesis of this paper.

SPECIAL CASES

One cannot make a similar general statement about either of the algorithms using weight ρ or $\rho|\mathbf{A}|$. We therefore turn to three special cases permitting analytical evaluation of the statistical errors.

Observable A with constant dimensionality $D' < 2D$

When $\mathbf{A}$ depends only on a fixed subset D' of all $2D$ phase-space coordinates, we may trivially integrate out remaining $2D - D'$ coordinates in all averages in Eq. (14) under the assumption that these are dynamically uncoupled from D' coordinates on which $\mathbf{A}$ depends. In that case, σ is trivially independent of D for all studied sampling weights. When

the coupling is present, the partial integration is not possible because $\mathbf{A}^t$ becomes function of all $2D$ initial phase-space coordinates for $t > 0$. As demonstrated below, in this case, σ generally depends on D for algorithms using weights ρ and $\rho|\mathbf{A}|$. However, as the numerical results and experience with molecular dynamics simulations show, in many cases the dependence on D is not very strong. By contrast, for weight $\rho\mathbf{A}^2$, as our proof shows, the error is *always* independent of both D and D'.

Remaining analytical examples concern observables $\mathbf{A}$ depending explicitly on $D' = D$ phase-space coordinates. They both involve a many-dimensional harmonic oscillator (HO) $H = (1/2)(p^2/m + kq^2)$ and its stationary Gaussian distribution

$$\rho(x) = [2\tanh(u/2)]^D \exp[-\tanh(u/2)(q^2/a^2 + p^2 a^2/\hbar^2)], \tag{16}$$

given by the Wigner transform of the Boltzmann density operator. Above, $u := \beta\hbar\omega$, $\omega^2 = k/m$, $a^2 = \hbar/(m\omega)$. [Note that the ground state density and the classical Boltzmann distribution can be obtained as the limits of Eq. (16) for $\beta \to \infty$ and $\beta \to 0$, respectively.]

Exponential growth of σ with D

First consider A to be the product of coordinates: $A = q_1 q_2 \cdots q_D$. The statistical error for $W = \rho A^2$ is described by Eq. (15) in full generality and thus is independent of D. On the other hand, straightforward but somewhat tedious calculations using Eq. (14) show that statistical errors for both weights ρ and $\rho|A|$ grow exponentially with the number of dimensions D:

$$\sigma_\rho(N, t)^2 = \frac{1}{N}\left\{\left[1 + 2\sqrt[D]{C(t)^2}\right]^D - 3^D C(t)^2\right\}, \tag{17}$$

$$\sigma_{\rho|A|}(N, t)^2 = \frac{1}{N}\left(\frac{2}{\pi}\right)^D \left\{\left[1 + \sqrt[D]{C(t)^2}\right]^D - 2^D C(t)^2\right\}. \tag{18}$$

The fact that for $W = \rho$ and $\rho|A|$ there exist observables for which the error grows exponentially with D is our second main result. Similar behavior of σ is expected for any multiplicatively separable function A of phase-space coordinates, such as the Gaussian $A = \exp(-q^2/a^2)$.

Independence of D

Yet, the situation is not always so bleak. Consider the correlated function $A = \mu' \cdot q$ to be a linear function of coordinates q (μ' is a D-dimensional vector). In this important special case, all three sampling methods have statistical errors independent of dimensionality:

$$\sigma_{\rho \text{ or } \rho A^2}(N, t)^2 = N^{-1}[1 - C(t)^2], \tag{19}$$

$$\sigma_{\rho|A|}(N, t)^2 = (2/\pi)N^{-1}[1 - C(t)^2]. \tag{20}$$

The proof of Eq. (20) for weight $\rho|A|$ is somewhat involved and was done only for the case $\mu_1 = \cdots = \mu_D$. On the other hand, Eq. (19) remains valid even for HOs with different frequencies in different dimensions. Note that the statistical error is slightly lower for $W = \rho|A|$ than for $W = \rho$ or ρA^2.

SAMPLING METHODS AND CORRELATION LENGTH

Before presenting numerical examples, let us briefly discuss the sampling methods. In many dimensions, sampling from a general weight W is often performed with the Metropolis method.[17–19] Two variants are used here: The original Metropolis method proposes the new point x_{new} using a random walk step from the last accepted point x_{old}; x_{new} is accepted with probability $p_{\text{acc}} = \min[W(x_{\text{new}})/W(x_{\text{old}}), 1]$. If x_{new} is rejected, the last accepted point x_{old} is duplicated. In the "product" Metropolis method, W is factorized as $W = YZ$, where Y can be sampled directly to propose a new point x_{new}, which is subsequently accepted with probability $p_{\text{acc}} = \min[Z(x_{\text{new}})/Z(x_{\text{old}}), 1]$.

Unfortunately, except for a few distributions W (such as the uniform or normal distributions, which may be sampled directly), points generated by Metropolis methods are correlated, leading to a correlation length $N_{\text{corr}} > 1$ between samples. This increases the statistical error for a given number of samples N. As a consequence, in all of our analytical expressions, N should be replaced by N/N_{corr}, which can affect (slightly) the dependence of σ on D. An important factor increasing N_{corr} is the rejection of proposed moves, which results in exactly identical samples. In a properly designed code, however, these repeated samples do not increase the computational cost; they are accounted for by increasing the statistical weight of the original (not yet duplicated) sample. Thus, the efficiency of a sampling algorithm depends on the number N_{uniq} of unique trajectories needed for convergence rather than on the total number N of trajectories. In situations like ours, where the cost of evaluating the "estimator" is much larger than the cost of a MC move, the correlation length N_{corr} can be reduced by considering only every nth point generated. While we took N_{corr} into account in the numerical calculations, a detailed analysis of N_{corr}, which can both increase (slowly) or decrease (slowly) with D, is beyond the scope of this paper.

NUMERICAL RESULTS

We first confirmed our analytical results for HOs numerically using $k = m = \hbar = \beta = 1$. Numerical statistical errors were estimated by averaging these errors over 100 independent simulations, each with the same number of unique trajectories $N_{\text{uniq}} = 5 \times 10^5$. In order to compare with the analytical results, the effect of correlation was removed by converting the numerical statistical error σ to an error per trajectory $\sigma_1 := (N/N_{\text{corr}})^{1/2}\sigma$. The correlation lengths N_{corr} were estimated using the method of block averages.[20]

Figure 1(a) shows that for $A = q_1 q_2 \cdots q_D$, the error σ_1 grows exponentially with D for both weights ρ and $\rho|A|$, while it is independent of D for $W = \rho A^2$. Moreover, numerical results agree with the analytical predictions (15) and (17), and (18). The original Metropolis method was used, since the acceptance rate of the product Metropolis method was
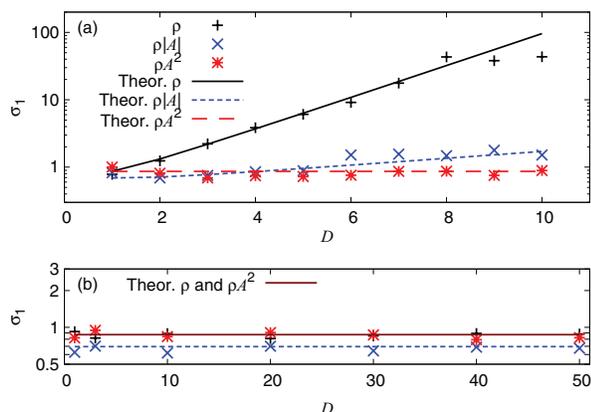
FIG. 1. Expected statistical error per trajectory of the autocorrelation function $C(t)$ of a function A in a D-dimensional harmonic oscillator. Time t was chosen separately for each D so that $C(t) \approx 0.5$. (a) $A = q_1 q_2 \cdots q_D$. Statistical error is independent of dimensionality for the algorithm with weight $W = \rho A^2$ and grows exponentially with D for the other two weights. (b) $A = \mu' \cdot q$. Statistical error is independent of dimensionality for all three sampling weights studied.
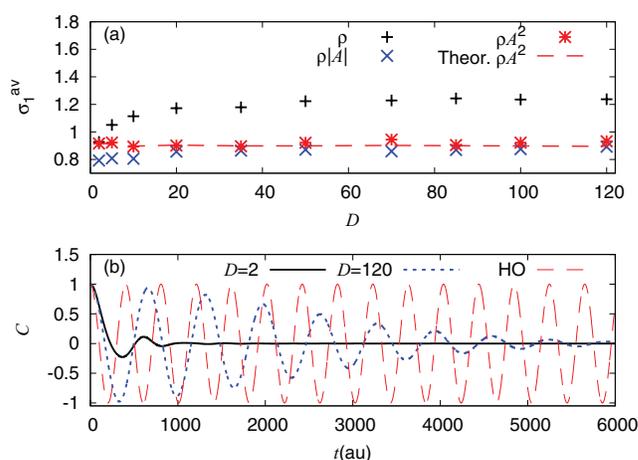

FIG. 2. (a) Expected average statistical error per trajectory σ_1^{av} of the autocorrelation function $C(t)$ of a function $A = \mu' q_1$ in a D-dimensional harmonic oscillator perturbed by the Gaussian bump. Note that, for all D, A depends explicitly only on a single coordinate q_1. For $W = \rho A^2$, the error σ_1^{av} follows theoretical expectations. For $W = \rho$, σ_1^{av} is higher than for $W = \rho A^2$ and grows slowly with D. (b) Autocorrelation functions $C(t)$. Initial decay of $C(t)$ in the perturbed system is comparable for all D and is slower than in the unperturbed harmonic oscillator (HO). Parameters of $V(q)$ in atomic units were $k = 0.883$, $m = 1822.9$, $a_1 = 0.1$, and $b_1 = 0.3$. Temperature was set to 3.167×10^{-2} a.u. and $\mu' = 1$.

prohibitively low for high D. The step size of the random walk was the same for all three weights but varied weakly with D for the sake of a reasonable acceptance rate.

Figure 1(b) compares the analytical predictions with numerically computed errors for $A = \mu' \cdot q$, where μ' is a D-dimensional vector with all entries equal to 1. Such A can be interpreted as a linear approximation to the electric dipole of a nonpolar molecule. Figure 1(b) confirms that the statistical error σ_1 is independent of D for all three algorithms. Initial conditions were sampled using the product Metropolis algorithm with $W = YZ$ and $Y = \rho$ in all cases. Function Z used in the acceptance criterion was equal to 1, $|A|$, and A^2, for $W = \rho$, $\rho|A|$, and ρA^2, respectively. Therefore, for $W = \rho$, $N_{corr} = 1$, and $N = N_{uniq}$, while for $W = \rho|A|$ and ρA^2, $N_{corr} > 1$ and $N > N_{uniq}$.

The independence of D, however, is not robust, and breaks down for the standard algorithm ($W = \rho$) even for small perturbations of the harmonic system and for one-dimensional linear observables. In Fig. 2, we numerically demonstrate that for $W = \rho$ or $\rho|A|$ in anharmonic systems, σ_1 in general depends on D even for a linear function $A = \mu' q_1$ depending on a single coordinate q_1. The anharmonic system was obtained by perturbing the HO with a Gaussian bump to obtain a system described by the potential

$$V(q) = \frac{k}{2} \sum_{i=1}^{N} q_i^2 + a \exp\left(\frac{-\sum_{i=1}^{N} q_i^2}{b^2}\right), \quad (21)$$

where parameters $a = D a_1$ and $b^2 = D b_1^2$ grow linearly with D in order to keep the initial decay of the autocorrelation function $C(t)$ similar for all D. As can be seen in Fig. 2(b), the envelope of $C(t)$ decays more slowly for higher D, signifying that in a certain sense the perturbation in Eq. (21) becomes less important in higher dimensions.

To compare statistical errors for different D in a way which takes into account all important parts of $C(t)$, we computed statistical errors per trajectory σ_1^{av} which are averaged

over $C(t)$ from $t = 0$ to the last time where $|C(t)| > 0.02$ using the time step $\Delta t = 1$ a.u. Only every 2000th trajectory generated by the Metropolis algorithm was used in order to reduce N_{corr}. This allowed to reduce N_{uniq} to 5000 and to increase the number of independent simulations to 1000. Correlation lengths were computed automatically by block averaging with the block size equal to $N/50$. As can be seen in Fig. 2(a), for $W = \rho A^2$, numerical value of $\sigma_{1,\rho A^2}^{av}$ follows the theoretical prediction which is essentially a constant function of D. For $W = \rho$, $\sigma_{1,\rho}^{av}$ grows slowly and is approximately 1.4 times higher than $\sigma_{1,\rho A^2}^{av}$ for $D > 50$. For $W = \rho|A|$, only a very slight growth of $\sigma_{1,\rho|A|}^{av}$ can be observed. Despite the modest growth of $\sigma_{1,\rho}^{av}$, this example shows that even for a weakly anharmonic system and the simplest correlated function $A = \mu' q_1$, the dimensionality independent sampling weight $W = \rho A^2$ is advantageous to $W = \rho$. Interestingly, in both the harmonic system and, for $D \leqq 120$, even in the weakly anharmonic system, the sampling weight $W = \rho|A|$ gives the lowest σ_1, suggesting that in some systems this weight may be the best choice for linear observables $A = \mu' \cdot q$. Nonetheless, since $\sigma_{\rho|A|}$ is difficult to predict in general, in other systems this error may be substantially larger than the always well-behaved $\sigma_{\rho A^2}$.

Finally, to demonstrate that all three algorithms indeed converge to the same result, we calculated the vibrational spectrum of a 48-dimensional harmonic model of the ground electronic state of azulene computed at the CASSCF(4,6)/6-31G* level of theory. Observable $\mathbf{A}$ was a linear approximation of the dipole moment of azulene, $\mathbf{A} = \boldsymbol{\mu} = \boldsymbol{\mu}_0 + \boldsymbol{\mu}' \cdot q$, where $\boldsymbol{\mu}_0 := \boldsymbol{\mu}(0)$ is the equilibrium dipole moment (a three-dimensional vector) and $\boldsymbol{\mu}'$ the $3 \times D$ matrix of derivatives of the dipole moment at $q = 0$. Sampling was performed as in the previous example. The dipole autocorrelation function $C(t)$ was computed up to time $t_{tot} = 1.45$ ps, the
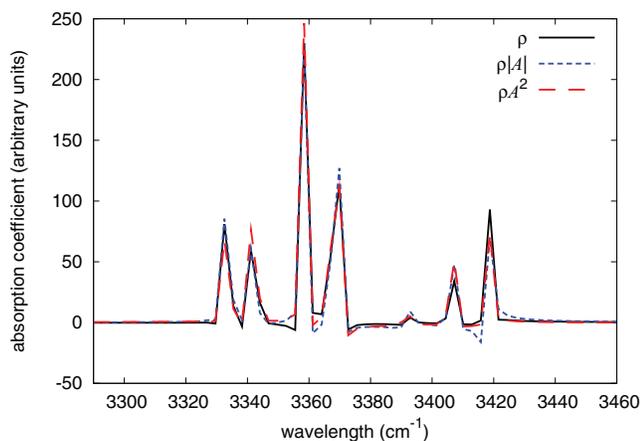
FIG. 3. High frequency part of the vibrational spectrum of a harmonic model of azulene computed via the Fourier transform of the dipole time autocorrelation function.

minimum time needed to resolve all vibrational peaks, and with only $N_{\text{uniq}} = 10^4$ unique trajectories, for which $C(t)$ starts to converge. Prior to computing the spectrum via a Fourier transform, $C(t)$ was damped with a factor $\cos(\pi t/2t_{\text{tot}})^2$. After the transform, $\mathcal{F}[C(t)](\omega)$ was multiplied by $2\omega \tanh (\beta\hbar\omega/2)$, which includes the standard "quantum correction"[2] for the lack of detailed balance in the classical $C(t)$. While this correction is *not* exact even for HOs if ρ is the classical Boltzmann density, it becomes exact for harmonic systems if ρ is the Wigner Boltzmann density (16). Figure 3, showing the high-frequency region of the spectrum containing the C–H bond stretches, confirms that all three algorithms converge to the same result, agreeing, within the resolution, with the exact spectrum. Moreover, even in this slightly more general harmonic case [than that considered in Fig. 1(b)], the statistical errors associated with all three sampling weights stayed approximately independent of D. (Systems with $D < 48$ were generated by progressively cutting off the lowest frequency normal modes of azulene.)

CONCLUSIONS

We have demonstrated the existence of a sampling weight for which the number of trajectories needed for convergence of the normalized time autocorrelation function of any phase-space function **A** is independent of the system's dimensionality and underlying dynamics. This weight is $W = \rho\mathbf{A}^2$, which may not be surprising at time $t = 0$, when this W represents the ideal importance sampling weight with all trajectories contributing unity to the sum (6). Here we have

shown that this weight retains its favorable properties also for $t > 0$ by proving that $\sigma_{\rho\mathbf{A}^2}$ depends explicitly only on $C(t)$ itself, and not on other parameters of the system.

While best suited for normalized autocorrelation functions, weight $\rho\mathbf{A}^2$ can also be used in calculations of unnormalized autocorrelation functions $C_u(t)$ via the relation $C_u(t) = C_u(0)C(t)$. There, the weight $\rho\mathbf{A}^2$ is retained for the dynamical calculation of $C(t)$, typically the most time-consuming task by far. Although the initial norm $C_u(0)$ must be computed separately using a normalized sampling weight such as ρ, one can afford many more trajectories, since calculation of $C_u(0)$ does not require any dynamics.

To conclude, we hope that the dimensionality-independent sampling weight will find its use in other classical, semiclassical,[4,5] and even quantum mechanical trajectory-based applications, such as those using the centroid[6,9–11] or ring-polymer[7–11] molecular dynamics.

[1]A. Nitzan, *Chemical Dynamics in Condensed Phases* (Oxford University Press, New York, 2006).
[2]B. J. Berne and G. D. Harp, Adv. Chem. Phys. **17**, 63 (1970).
[3]G. Williams, Chem. Rev. **72**, 55 (1972).
[4]X. Sun and W. H. Miller, J. Chem. Phys. **110**, 6635 (1999).
[5]J. Liu, W. H. Miller, G. S. Fanourgakis, S. S. Xantheas, S. Imoto, and S. Saito, J. Chem. Phys. **135**, 244503 (2011).
[6]J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5106 (1994).
[7]I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **121**, 3368 (2004).
[8]T. F. Miller and I. D. E. Manolopoulos, J. Chem. Phys. **123**, 154504 (2005).
[9]S. Habershon, G. S. Fanourgakis, and D. E. Manolopoulos, J. Chem. Phys. **129**, 074501 (2008).
[10]A. Witt, S. D. Ivanov, M. Shiga, H. Forbert, and D. Marx, J. Chem. Phys. **130**, 194510 (2009).
[11]A. Perez, M. E. Tuckerman, and M. H. Muser, J. Chem. Phys. **130**, 184105 (2009).
[12]C. Mollica, T. Zimmermann, and J. Vaníček, Phys. Rev. E **84**, 066205 (2011).
[13]C. Mollica and J. Vaníček, Phys. Rev. Lett. **107**, 214101 (2011).
[14]T. Gorin, T. Prosen, T. H. Seligman, and M. Žnidarič, Phys. Rep. **435**, 33 (2006).
[15]D. A. McQuarrie, *Statistical Mechanics* (Harper and Row publishers, New York, 1976).
[16]M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford University Press, New York, 2010).
[17]N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[18]W. Hastings, Biometrika **57**, 97 (1970).
[19]S. Chib and E. Greenberg, Am. Stat. **49**, 327 (1995).
[20]H. Flyvbjerg and H. G. Petersen, J. Chem. Phys. **91**, 461 (1989).