Accelerated direct semiclassical molecular dynamics using a compact finite difference Hessian scheme

Michele Ceotto, Yu Zhuang, and William L. Hase

Citation: The Journal of Chemical Physics **138**, 054116 (2013); doi: 10.1063/1.4789759
View online: http://dx.doi.org/10.1063/1.4789759
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/138/5?ver=pdfcov
Published by the AIP Publishing

Accelerated direct semiclassical molecular dynamics using a compact finite difference Hessian scheme

Michele Ceotto,[1,a)] Yu Zhuang,[2] and William L. Hase[3]
[1]*Dipartimento di Chimica, Università degli Studi di Milano, via Golgi 19, 20133 Milano, Italy*
[2]*Department of Computer Science, Texas Tech University, Lubbock, Texas 79409-3104, USA*
[3]*Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409, USA*

(Received 19 December 2012; accepted 15 January 2013; published online 6 February 2013)

This paper shows how a compact finite difference Hessian approximation scheme can be proficiently implemented into semiclassical initial value representation molecular dynamics. Effects of the approximation on the monodromy matrix calculation are tested by propagating initial sampling distributions to determine power spectra for analytic potential energy surfaces and for "on the fly" carbon dioxide direct dynamics. With the approximation scheme the computational cost is significantly reduced, making *ab initio* direct semiclassical dynamics computationally more feasible and, at the same time, properly reproducing important quantum effects inherent in the monodromy matrix and the pre-exponential factor of the semiclassical propagator. © *2013 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4789759]

I. INTRODUCTION

Semiclassical (SC) molecular dynamics provides a general and well-defined tool for including *all* quantum effects in classical mechanics.[1–3] In particular, semiclassical initial value representation (SC-IVR) theories have proven to be quite accurate, as demonstrated by numerous applications.[4–20] The fundamental advantage of the SC-IVR propagator over basis set evaluations of the quantum propagator, such as the multiconfigurational time-dependent Hartree[21] or the coupled coherent states methodology,[22] is that it exclusively relies on classical trajectories and, therefore, it can be straightforwardly implemented with "on the fly" direct molecular dynamics calculations. In addition, basis sets and other[23,24] approximate evaluations of the quantum propagator need a global potential energy surface, while SC-IVR depends only on the local potential.[25] This aspect is of fundamental relevance, as a major challenge in quantum dynamics is the simulation of complex systems, i.e., ones with multiple degrees of freedom, for which the development of an analytic potential energy surface may be a formidable task.[26–33]

The version of the SC theory that has proven most useful for quantum dynamics is the coherent states implementation of the SC-IVR propagator,[24,34,35] which goes under the name of Herman-Kluk or Heller-Herman-Kluk-Kay SC-IVR approximation

$$e^{-i\hat{H}t/\hbar} = \frac{1}{(2\pi\hbar)^F}\int d\mathbf{p}(0)\int d\mathbf{q}(0)\, C_t\left(\mathbf{p}(0),\mathbf{q}(0)\right) \times e^{iS_t(\mathbf{p}(0),\mathbf{q}(0))/\hbar}\left|\mathbf{p}(t),\mathbf{q}(t)\right\rangle\left\langle\mathbf{p}(0),\mathbf{q}(0)\right|, \quad (1)$$

where $(\mathbf{p}(t),\mathbf{q}(t))$ is the set of $2F$-dimensional classically evolved phase space coordinates, S_t is the classical action and C_t is a pre-exponential factor that arises from local harmonic fluctuations about the classical paths.[1,8,13–15,17,19,36,37] The pre-exponential factor $C_t(\mathbf{p}(0),\mathbf{q}(0))$ represents the biggest stumbling block in ambitious applications of this theory and the computational effort required for its calculation is a major concern for SC-IVR simulations. In particular, it represents most of the SC-IVR computational cost when "on the fly" direct dynamics simulations are performed, since it requires the calculation of the Hessian matrix directly from the electronic wavefunction. Therefore, important implementations of the coherent states SC-IVR theories are those which provide an approximation that conserves the properties of the original propagator and, at the same time, reduces the computational costs so as to make "on the fly" simulations a viable tool.

To overcome the pre-exponential factor issue, in this paper we implement a compact finite difference (CFD) approximation for the Hessian calculation recently developed by Zhuang *et al.*[38] This is a numerical approximation of the pre-exponential factor, with the goal of reducing computational costs and, thus, making *ab initio* SC-IVR direct dynamics viable for an extensive range of simulations which accurately and properly include the pre-exponential factor. The accuracy of this approximation has been tested only on single trajectory simulations[39] and it is not obvious how it will perform for the phase space integration of Eq. (1). This paper deals with this issue and shows how the CFD approximation may be successfully implemented into the SC-IVR propagator for power spectra calculations and how it performs for an ensemble of trajectories instead of a single one.

Another issue regarding the pre-exponential factor is its numerical stability, especially for chaotic systems. Besides removing trajectories,[40] several paths have been taken in the past to obviate this issue in SC dynamics. Linearization of the SC-IVR propagator (LSC-IVR)[15,41–45] resulted in the most drastic approximation in terms of accuracy. Then, the Forward-Backward (FB) SC-IVR method was proposed for correlation function calculations.[46(a),13,47] Another route is one of filtering the classical trajectories

a)michele.ceotto@unimi.it.

to avoid the numerical instabilities associated with the pre-exponential factor. These filters were inspired by Filinov[48] and Wang *et al.*[46(e)] based on physical considerations.[49] One can also employ the log-derivative representation[50] for a more convenient numerical integration. Another approach involves introducing approximations of the pre-exponential factor, such as Johnson's multichannel WKB (Wenzel-Kramers-Brillouin) approximation[51] or the adiabatic approximation.[52] Also, the prefactor has been partially suppressed in Pollak's series propagator expansion[11] and totally suppressed in Takatsuka's amplitude-free quasicorrelation function.[37] More recently, a "poor person's" Frozen Gaussian propagator has been proposed,[25] where the pre-exponential factor is calculated from only a single trajectory by artificially extracting it from the phase space integral of Eq. (1).

In this paper, we employ the time-averaging filter[53] (Secs. II and III) which controls the numerical issues of $C_t(\mathbf{p}(0), \mathbf{q}(0))$. It also reduces the extent of the phase space integration but at the cost of a longer simulation time. The latter is alleviated by the CFD approximation described in Sec. IV, which reduces the computation cost of the Hessian calculation. The results are presented in Sec. V, followed by a discussion in Sec. VI. Section VII concludes the paper.

II. TIME-AVERAGING SC-IVR FOR POWER SPECTRA CALCULATIONS

We calculate the power spectra in the time-dependent representation

$$I(E) = \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} \langle\chi|e^{-i\hat{H}t/\hbar}|\chi\rangle e^{iEt/\hbar}dt, \tag{2}$$

and approximate the propagator $e^{-i\hat{H}t/\hbar}$ using the semiclassical initial value representation (SC-IVR) of Eq. (1). In the coherent-state[24,34] version of SC-IVR, the pre-exponential factor

$$C_t(\mathbf{p}(0), \mathbf{q}(0)) = \sqrt{\frac{1}{2}\left|\frac{\partial \mathbf{q}(t)}{\partial \mathbf{q}(0)} + \frac{\partial \mathbf{p}(t)}{\partial \mathbf{p}(0)} - i\hbar\gamma\frac{\partial \mathbf{q}(t)}{\partial \mathbf{p}(0)} + \frac{i}{\gamma\hbar}\frac{\partial \mathbf{p}(t)}{\partial \mathbf{q}(0)}\right|} \tag{3}$$

is given by the determinant of the combination of the four $F \times F$ size blocks of the $2F \times 2F$ symplectic (monodromy or stability) matrix $\mathbf{M}(t) \equiv (\partial(\mathbf{p}(t), \mathbf{q}(t))/\partial(\mathbf{p}(0), \mathbf{q}(0)))$. By employing Hamilton's equations, the time-evolution of $\mathbf{M}(t)$ is obtained as

$$\frac{d}{dt}\mathbf{M}(t) = \mathbf{K}\cdot\mathbf{M}(t), \tag{4}$$

where

$$\mathbf{K} = \begin{pmatrix} 0 & -\frac{\partial^2 V(\mathbf{q})}{\partial \mathbf{q}^2} \\ \frac{1}{\mathbf{m}} & 0 \end{pmatrix}. \tag{5}$$

We monitored the accuracy of the classical propagation by checking the deviation of the determinant of the monodromy matrix (or better of the positive-definite matrix $\mathbf{M}^T\mathbf{M}$) from unity. The coherent states in Eq. (1) are given by the direct product of one-dimensional coherent states

$$\langle\mathbf{q}|\mathbf{p}(t), \mathbf{q}(t)\rangle = \prod_i (\gamma_i/\pi)^{F/4} \times \exp\left[-\frac{\gamma_i}{2}\cdot(q_i - q_i(t))^2 + \frac{i}{\hbar}p_i(t)\cdot(q_i - q_i(t))\right], \tag{6}$$

where γ_i is fixed and equal to the width of the harmonic oscillator approximation to the vibrational wave function for the ith normal mode. Finally, the SC-IVR approximation for the survival probability in Eq. (2) is represented in the semiclassical approximation by the following Monte Carlo phase space integration,

$$\langle\chi|e^{-i\hat{H}t/\hbar}|\chi\rangle = \frac{1}{(2\pi\hbar)^F}\int d\mathbf{p}(0)\int d\mathbf{q}(0)\, C_t(\mathbf{p}(0), \mathbf{q}(0)) \times e^{iS_t(\mathbf{p}(0),\mathbf{q}(0))/\hbar}\langle\chi|\mathbf{p}(t), \mathbf{q}(t)\rangle\langle\mathbf{p}(0), \mathbf{q}(0)|\chi\rangle \tag{7}$$

for any given reference state $|\chi\rangle = |\mathbf{p}_{eq}, \mathbf{q}_{eq}\rangle$.

In order to smooth the oscillatory integrand in Eq. (7), the time averaging filter was introduced,[53] where the number of trajectories required for the Monte Carlo integration is reduced at the cost of longer simulation times.[54] The TA (time-averaging) SC-IVR expression for the spectral density is

$$I(E) = \frac{1}{(2\pi\hbar)^F}\int d\mathbf{p}(0)\int d\mathbf{q}(0)\frac{\mathrm{Re}}{\pi\hbar T}\int_0^T dt_1 \times \int_{t_1}^T dt_2\, C_{t_2}(\mathbf{p}(t_1), \mathbf{q}(t_1)) \times \langle\chi \mid \mathbf{p}(t_2), \mathbf{q}(t_2)\rangle\, e^{i\left(S_{t_2}(\mathbf{p}(0),\mathbf{q}(0))+Et_2\right)/\hbar} \times [\langle\chi \mid \mathbf{p}(t_1), \mathbf{q}(t_1)\rangle e^{i(S_{t_1}(\mathbf{p}(0),\mathbf{q}(0))+Et_1)/\hbar}]^*, \tag{8}$$

where the sets $(\mathbf{p}(t_1), \mathbf{q}(t_1))$ and $(\mathbf{p}(t_2), \mathbf{q}(t_2))$ of the position and momentum variables are the evolution of the initial phase space point $(\mathbf{p}(0), \mathbf{q}(0))$ at times t_1 and t_2, respectively, and T is the total simulation time. The time integration in t_1 acts as the Fourier transform of Eq. (2), while the one in t_2 does the filtering job. It is important to note that the two time-integrals imply two nested do-cycles. In other words, for a trajectory run from time 0 to time T all possible time-intervals from t_1 and t_2 are considered. In this way, this integration sums over all possible trajectories that are generated as segments of the time length $T - t_1$ for a single long trajectory from time 0 to time T.

The most computational intense part of Eq. (8) is the calculation of the prefactor $C_{t_2}(\mathbf{p}(t_1), \mathbf{q}(t_1))$, for two reasons. First, because it requires the Hessian calculation; second, because it depends on two time variables, since its calculation is done by taking all possible trajectories for the $T - t_1$ time interval. This last requirement can be controlled by adopting the separable approximation, where $C_{t_2}(\mathbf{p}(t_1), \mathbf{q}(t_1)) \approx \exp[i(\phi(t_2) - \phi(t_1))/\hbar]$ and $\phi(t) = \mathrm{phase}[C_t(\mathbf{p}(0), \mathbf{q}(0))]$.

Then, Eq. (8) becomes

$$I(E) = \frac{1}{(2\pi\hbar)^F}\frac{1}{2\pi\hbar T}\int d\mathbf{p}(0)\int d\mathbf{q}(0) \times \left|\int_0^T dt\,\langle\chi \mid \mathbf{p}(t),\mathbf{q}(t)\rangle \times e^{i\left(S_t(\mathbf{p}(0),\mathbf{q}(0))+Et+\phi_t(\mathbf{p}(0),\mathbf{q}(0))/\hbar\right)}\right|^2, \quad (9)$$

where now the double time integral is simplified to a single and positive-definite time integral. In this paper we will employ this approximation for Eq. (2), since we have found it to be accurate[53,55] and an order of magnitude less computational demanding than Eq. (8).

III. TIME AVERAGING MC-SC-IVR IMPLEMENTATION

In order to perform *ab initio* direct semiclassical dynamics simulations, the number of trajectories employed for integrating Eq. (7) should be greatly reduced. This is done only in part by the time-averaging filter, since about a thousands of trajectories are needed *per* each degree of freedom[53] to converge the integration of Eqs. (8) and (9). For these reasons, recently the MC-SC-IVR method was developed by Ceotto *et al.*[55–57] This SC-IVR implementation enhances as much as possible the overlap between the reference state $|\chi\rangle$ and the exact quantum eigenfunctions. As a result, representation of spectral peaks for excited vibrational states are less noisy, more intense and accurate. To reach such a goal, the extra coherent states of the MC-SC-IVR method are placed either nearby the classical turning points of excited vibrational states or, equivalently, at higher momentum (i.e., kinetic energy) values comparable with the eigenstate energy (see Figure 1 in Ref. 55). Then, one can further reduce the number of trajectories to a few "eigen-trajectories," each one crossing a coherent state location, or sample at different coherent states locations per time. Importantly, the coherent states momenta do not need to be chosen at an energy close to the peak location because the Gaussian spreading of each coherent state is generally wide enough to include the peak energy shell, as tested numerically.[55,59]

Thus, in the MC-SC-IVR approximation, the reference states $|\chi\rangle$ are chosen to be a combination of coherent states,

$$|\chi\rangle = \sum_{i=1}^{N_{states}} \left|\mathbf{p}^i_{eq},\mathbf{q}^i_{eq}\right\rangle, \quad (10)$$

which have the equilibrium molecular position $\mathbf{q}_{eq}$ and initial momenta such that $p^2_{eq,j}/2m = \hbar\omega_j(n+1/2)$ for each normal mode frequency ω_j. Quantum mechanical delocalization is reproduced here because the coherent states are placed in a non-local fashion, even if their centers are kept fixed during the simulation time. Now, by inserting Eq. (10) into Eq. (8) the final expression for the multiple coherent states spectra calculation is

$$I(E) = \frac{1}{(2\pi\hbar)^F}\frac{\mathrm{Re}}{\pi\hbar T}\int d\mathbf{p}(0)\int d\mathbf{q}(0)\int_0^T dt_1 \times \int_{t_1}^T dt_2\, C_{t_2}(\mathbf{p}(t_1),\mathbf{q}(t_1)) \times \sum_{i=1}^{N_{states}}\left\langle \mathbf{p}^i_{eq},\mathbf{q}^i_{eq} \mid \mathbf{p}(t_2),\mathbf{q}(t_2)\right\rangle e^{i(S_{t_2}(\mathbf{p}(0),\mathbf{q}(0))+Et_2)/\hbar} \times \left[\sum_{i=1}^{N_{states}}\left\langle \mathbf{p}^i_{eq},\mathbf{q}^i_{eq}\middle|\mathbf{p}(t_1),\mathbf{q}(t_1)\right\rangle e^{i(S_{t_1}(\mathbf{p}(0),\mathbf{q}(0))+Et_1)/\hbar}\right]^*, \quad (11)$$

where the convenient combination of N_{states} coherent state centers of Eq. (10) has been introduced. An analogous expression holds after the separable approximation is invoked. Due to the approximation in the coherent state combination of Eq. (10), the MC-SC-IVR method provides information about the peak locations and not about their relative intensities with respect to a ground state transition. In order to further reduce the computational effort for a direct dynamics calculation, the phase space integral may be approximated by a sum of trajectories starting from each set of coherent state coordinates

$$I(E) = \frac{1}{(2\pi\hbar)^F}\frac{\mathrm{Re}}{\pi\hbar T}\sum_{j=1}^{N_{states}}\int_0^T dt_1\int_{t_1}^T dt_2\, C_{t_2}(\mathbf{p}^j(t_1),\mathbf{q}^j(t_1)) \times \sum_{i=1}^{N_{states}}\left\langle \mathbf{p}^i_{eq},\mathbf{q}^i_{eq} \mid \mathbf{p}^j(t_2),\mathbf{q}^j(t_2)\right\rangle e^{i(S_{t_2}(\mathbf{p}(0),\mathbf{q}(0))+Et_2)/\hbar} \times \left[\sum_{i=1}^{N_{states}}\left\langle \mathbf{p}^i_{eq},\mathbf{q}^i_{eq}\middle|\mathbf{p}^j(t_1),\mathbf{q}^j(t_1)\right\rangle e^{i(S_{t_1}(\mathbf{p}(0),\mathbf{q}(0))+Et_1)/\hbar}\right]^*. \quad (12)$$

Applications of Eq. (12) include accurate power spectra for the gas phase H_2O molecule[55] and CO molecules chemisorbed on Cu(100) using analytical potentials,[58] while an "on the fly" approach has been employed for CO_2[55,59] and H_2CO[57] vibrational energy level and vibrational eigenfunction calculations.[56]

IV. THE CFD METHOD FOR THE MONODROMY MATRIX CALCULATIONS

To calculate the monodromy matrix M(t) elements, the Hessian is needed at every time step to solve Eq. (4). An *ab initio* Hessian calculation is computationally very expensive, and hence an accurate Hessian approximation that can preserve properties of the *ab initio* Hessian is highly desirable. Hessian updating is a way to approximate the Hessian in a stepwise fashion with one direction of the Hessian approximated using the latest available information, while the remaining directions of the Hessian remain almost unchanged.[60]

Hessian update schemes have evolved since 1950s and were originally developed for optimization (see Refs. 60–65 and references therein), and some have been used in direct dynamics simulations.[66–68] Most of the schemes are based on a first-order Taylor expansion which, for optimization, is the equation that quasi-Newton methods are based upon. The accuracy of the first-order Taylor expansion is sufficient for optimization, as evidenced by fast convergence of quasi-Newton methods.[62] For direct dynamics, a higher accuracy is desirable and a recent study[38] has shown the effectiveness of a highly accurate Hessian approximation in attaining high simulation quality. The Hessian approximation schemes reported in Ref. 38 were developed using a CFD method,[69–73] which has an approximation accuracy one order higher than that of the first-order Taylor expansion.

CFD is a high-order finite difference approximation for differentiation of functions without incurring a larger stencil, the set of sampling points in the approximation formula. The high accuracy is attained by including differentiated terms at more locations within the stencil. The "compact" refers to the compactness of the stencil for attaining a high-accuracy approximation. The CFD-based Hessian approximations start with the equation

$$\frac{1}{2}[H(q_1) + H(q_2)](q_2 - q_1) = G(q_2) - G(q_1), \tag{13}$$

where $G(q_1)$ and $G(q_2)$ are the gradients of the potential energy at q_1 and q_2, respectively, and $H(q_1)$ and $H(q_2)$ are the Hessians at q_1 and q_2. Compared with the first-order Taylor expansion

$$H(q_2)(q_2 - q_1) = G(q_2) - G(q_1) \tag{14}$$

of $G(q_1)$ about the point q_2, the CFD-based Eq. (13) has an error of $O(\|q_1 - q_2\|^3)$ while the Taylor expansion in Eq. (14) has an error of $O(\|q_1 - q_2\|^2)$. Note that Eq. (13) has the same stencil as Eq. (14), but with the Hessian included at one more location of the variable q. When $G(q_1)$, $G(q_2)$, and $H(q_1)$ are given, a Hessian update scheme is to approximate $H(q_2)$ using the given information of $G(q_1)$, $G(q_2)$, $H(q_1)$, and q_1 and q_2. Equation (14) is the equation for which most existing Hessian update schemes are based upon. When the coordinate q has n scalar entries, that is, $q = (x_1, x_2, \cdots, x_n)$, both Eqs. (13) and (14) have an $n(n-1)/2$-dimensional solution space. A dimension-restriction technique pioneered by Bofill,[65] originally for the first-order Eq. (14), can be applied to the CFD Eq. (13), with which one can obtain a family of Hessian update schemes

$$\Delta H = (1-\lambda)\frac{R\,R^T}{R^T\Delta q} + \lambda\left(\frac{\Delta q\,R^T + R\,\Delta q^T}{\|\Delta q\|^2} - \frac{R^T\Delta q}{\|\Delta q\|^4}\Delta X q\,\Delta q^T\right), \tag{15}$$

from which $H(q_2)$ can be obtained using $H(q_2) = H(q_1) + \Delta H$, where λ is a parameter allowed to vary, $\Delta q = q_2 - q_1$, and

$$R = 2[G(q_2) - G(q_1) - H(q_1)(q_2 - q_1)]. \tag{16}$$

Equation (15) is called the CFD-Bofill family[38] of Hessian update schemes, where the CFD contribution is effected through Eq. (16) for R, and Eq. (15) reduces to the CFD-Powell Symmetric Broyden (CFD-PSB) scheme when $\lambda = 1$. It is worth noting that when R is half of Eq. (16), that is, $R = G(q_2) - G(q_1) - H(q_1)(q_2 - q_1)$, Eq. (15) gives the family of Bofill update schemes that are solutions of the first-order Eq. (14). A practical value for λ is

$$\lambda = 1 - \frac{(R^T\,\Delta q)^2}{\|R\|^2 \cdot \|\Delta q\|^2}. \tag{17}$$

Equation (17) for λ is due to Bofill,[65] and Eq. (15) with λ given by Eq. (17) is called the CFD-Bofill Hessian update scheme, which was found to be a simple and accurate Hessian approximation.[38]

With the Hessian update schemes, the monodromy matrix M(t) can be calculated by solving the governing Eq. (4) with the velocity Verlet algorithm or higher order symplectic algorithms. Hessian updating is used in a manner that for every K time steps, an *ab initio* Hessian is calculated in the first step followed by K—1 steps of Hessian updates.

V. RESULTS

A. Preliminary pre-exponential factor monitoring

The accuracy of the pre-exponential factor $C_{t_2}(\mathbf{p}(t_1), \mathbf{q}(t_1))$ is preliminarily monitored by looking at the phase $\phi(t) = \text{phase}[C_t(\mathbf{p}(0), \mathbf{q}(0))]$ and at the modulus for calculations with different K values. Here K is the number of time-steps within a fixed interval for which the Hessian is estimated according to the CFD approximation. In Figure 1, the phase differences between the phase of the exact pre-exponential factor and the CFD approximated ones are plotted. The left-panel of Figure 1 is for a CO_2 ground state energy trajectory simulation (10 a.u. step-size for 5000 time-steps) and it shows that accuracy is preserved up to K = 32. Though there are a few points with a π off-phase pre-exponential factor, the imaginary part is still the same. The right-panel of the same Figure is for a H_2O ground state energy trajectory with 4000 time-steps of 5 a.u. In this case, the approximation is much less accurate. This is probably due to the presence of the light hydrogen atoms and chaotic dynamics, where the CFD approximation performs more poorly.[39] However, the phase differences in Figure 1 represent those for a single trajectory and in the following calculations are presented for thousands of trajectories.

Thus, it is not possible to extend *a priori* the accuracy shown in Figure 1 to each trajectory of the Monte Carlo phase space integration for the TA-SC-IVR with the separable approximation. Lower kinetic energy trajectories better preserve the accuracy of the CFD approximation (see Sec. V B), since for high kinetic energies it becomes more probable to visit chaotic regions of the potential energy surface where at least one of the local Hessian eigenvalues is negative.[39]

If one does not adopt the separable approximation of Eq. (9), it is important to establish the resulting error in the absolute value of the prefactor when the CFD approximation is used. The relative error of the modulus of the pre-exponential

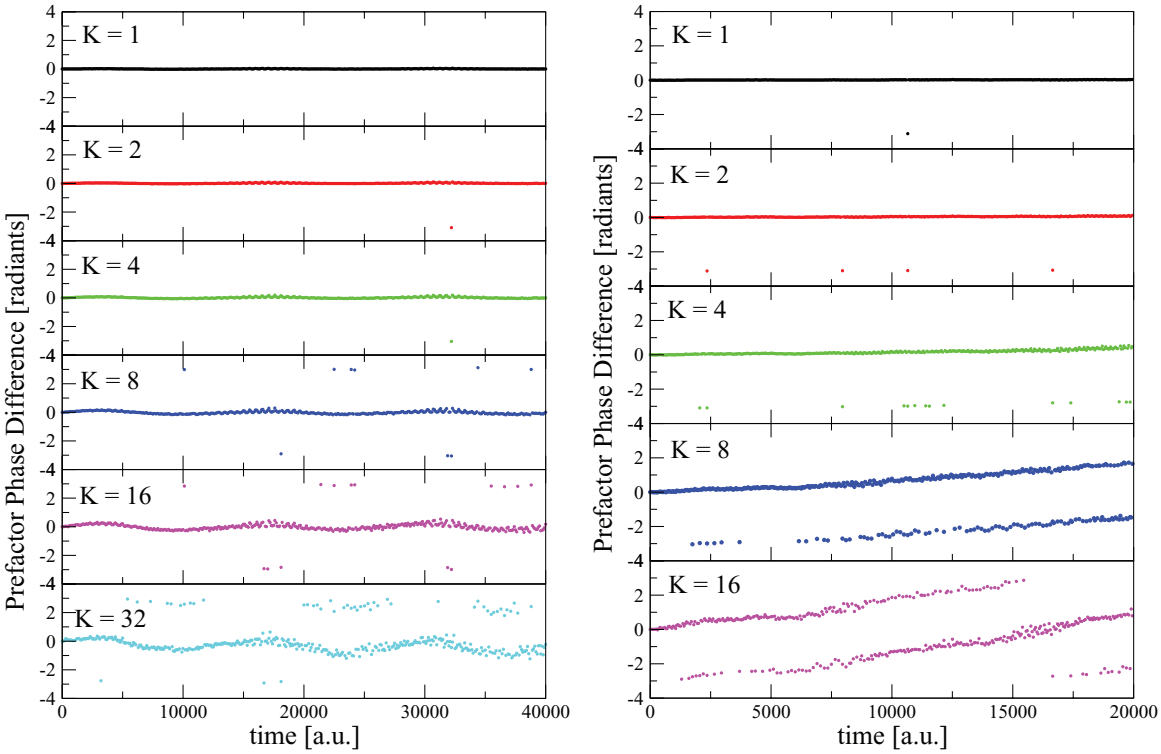

FIG. 1. Prefactor phase differences $\phi(t) = \text{phase}[C_t(\mathbf{p}(0), \mathbf{q}(0))]$ at different levels K of the CFD-PSB approximation for the ground state energy. The left panel is for a CO_2 trajectory and the right for a H_2O trajectory.

factor $C_{t_2}(\mathbf{p}(t_1), \mathbf{q}(t_1))$, at each time step, is reported in Figure 2. The notation is the same as in Figure 1. As far as the CO_2 set of simulations are concerned, the accuracy limit of K = 16 is observed, showing that the approximation is less accurate for the modulus than the phase calculation. The same considerations hold for the H_2O simulations, where the accuracy limit is about K = 8. These results are encouraging and show that the CFD approximation will be helpful even for SC-IVR simulations where the separable approximation is not employed. In order to have a comprehensive evaluation of the CFD Hessian approximation for the SC-IVR integrator, a better approach is to directly monitor the power spectra peaks for increasing K values, as done in Sec. V B.

B. Accelerated SC-IVR power spectra with the CFD approximation

From Sec. IV and Figure 1, it is clear that the accuracy of the CFD approximation deteriorates when the number of Hessian updates between two consecutive *ab initio* Hessian calculations increases. Also, the CFD Hessian calculations for high kinetic energy trajectories are less accurate than for lower kinetic energy trajectories as stated above. Thus, when performing the Monte Carlo integration of Eqs. (8) and (9) with the CFD approximation, high energy trajectories are more probable to be rejected, since their monodromy matrix determinants deviate substantially from unity as a result of round-off error propagation discussed above. This bias increases when the CFD Hessian approximation is invoked for more time-steps and for higher kinetic energy trajectories. To have a clear picture of this sampling issue, we performed a preliminary statistical study of time averaging SC-IVR calculations with different sampling choices in order to understand how discarding high momentum trajectories approximates the power spectrum. In particular, the goal of this preliminary set of exact (i.e., without CFD approximation) semiclassical calculations is to find the largest sampling standard deviation which gives accurate results for a given power spectrum energy range. In a second stage, the CFD approximation is applied to determine a K-threshold for an accurate calculation.

The water molecule analytical potential of Thiel[78] was chosen for a realistic simulation of a molecular spectrum. This molecule has stiff stretching modes strongly coupled to a floppy bending one. To perform the Monte Carlo integration of Eqs. (8) and (9), the Husimi sampling distribution of Eq. (A1) is usually employed[2,3,10] and this approach was used here. This distribution guarantees that a wide enough phase space region is sampled. The sensitivity of the Monte Carlo integration was studied for different sampling widths of the Husimi distribution and different numbers of trajectories calculated for the distribution. The left panel of Figure 3 shows the distribution of trajectories versus the total (kinetic plus potential) trajectory energy for the different sampling widths and different numbers of trajectories. The sampling width was narrowed by increasing the **a**, **b** parameters introduced in the Appendix in Eq. (A1) to 5**a**,5**b**; 10**a**,10**b**; and 15**a**,15**b**. Two ensemble sizes were considered for sampling the Husimi distributions; i.e., one with 32 000 trajectories and the other with 3200. The former were integrated for 3000 time-steps with a 5 a.u. time interval, while the latter was integrated for 4000 time-steps with the same time interval. The trajectory integration was

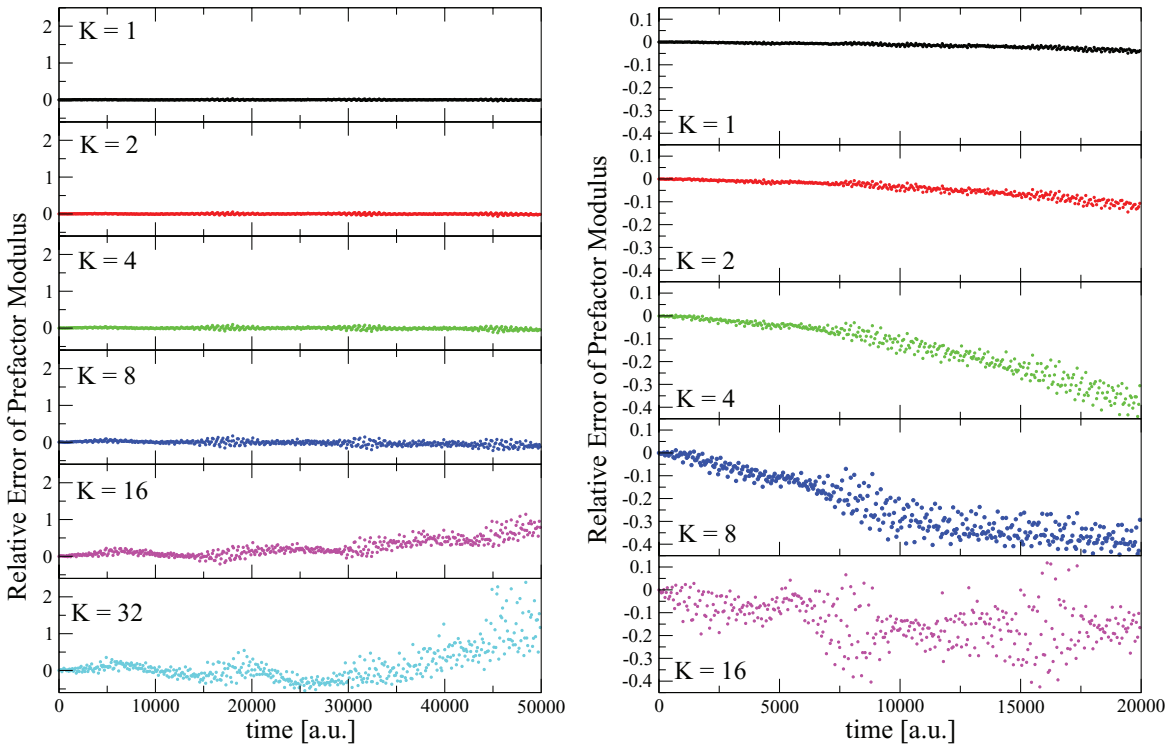

FIG. 2. Prefactor modulus relative error at different levels K of the CFD-PSB approximation for the ground state energy. The left panel is for a CO_2 trajectory and the right for a H_2O trajectory.

performed with a 4th order symplectic algorithm[14,79] and trajectories with a monodromy matrix determinant deviation from unity greater than 10^{-4} were discarded. The latter becomes important for trajectories with a large kinetic energy.

The power spectra, calculated using Eq. (9), are shown in the right panel of Figure 3 for the different Husimi ensembles of trajectories in the left panel. Overall, the spectra are insensitive to the samplings of the Husimi distribution. For the Husimi distribution with parameters 1**a**, 1**b** the energy range extends to 14 000 cm^{-1} beyond the ZPE level. The number of rejected trajectories is large for this broad sampling width, as a result of the large kinetic energies of some of the trajectories, and it is significant that this sampling gives an accurate spectrum. It is also important that the same Husimi distribu-

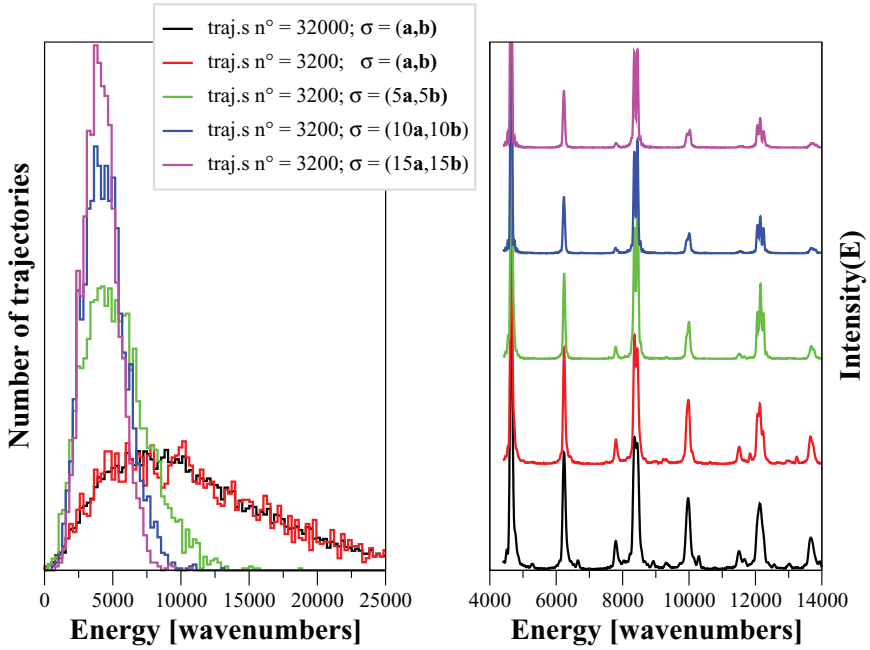

FIG. 3. Water molecule vibrational dynamics. (Left panel) Normalized trajectory distribution versus the total initial trajectory energy for different Gaussian **a**, **b** sampling standard deviations; the bin width for the histogram distributions is 200 cm^{-1}. (Right panel) The resulting power spectrum.

TABLE I. Water vibrational energy levels for the sampling distributions reported in the left panel of Figure 3.[a]

$\nu_1\nu_2\nu_3$		σ	σ	5σ	10σ	15σ
traj.s	Harm.	32 000	3200	3200	3200	3200
000 (A_1)	4711	4644	4646	4648	4644	4640
100 (A_1)	6361	6236	6238	6239	6231	6231
200 (A_1)	8011	7793	7796	7791	7787	7796
010 (A_1)	8541	8361	8360	8356	8352	8351
001 (B_2)	8652	8441	8441	8450	8446	8446
300 (A_1)	9661	9314	9313	9307	9348*	9368*
110 (A_1)	10 191	9933*	9923*	9928	9932	9945
101 (B_2)	10 302	9974	9980	9997	10 008	10 025
210 (A_1)	11 841	11 505	11 512	11 516	11 558	11 567
201 (B_2)	11 952	11 666	11 662	11 662	11 693*	
020 (A_1)	12 372	12 053*	12 065	12 061	12 058	12 058
011 (B_2)	12 483	12 134	12 134	12 153	12 146	12 145
002 (A_1)	12 593	12 226*	12 227	12 242	12 240	12 245
120 (A_1)	14 022	13 513*	13 482*	13 513		
111 (B_2)	14 133	13 667	13 657	13 681	13 683	13 700
102 (A_1)	14 244	13 750*	13 724*	13 762	13 807	13 833

[a]The 3200 Husimi sampling power spectrum (4th column) is at convergence by comparison with the 32 000 simulation (3rd column). Husimi distribution sampling is for $\sigma = (\mathbf{a}; \mathbf{b})$. Uncertain peaks are marked with (*). For multiples of σ the spectra become gradually less accurate.

tion, but with ensembles of trajectories which vary in size by an order of magnitude give the same spectrum (see black versus red lines). This is a result of the time-averaging filtering. In addition, it is seen that the intensity of the higher energy peaks gradually decrease as the sampling distribution is narrowed with larger **a**,**b** parameters.

To have a better comparison between the different sampling strategies adopted in Figure 3, the vibrational eigenvalues from the spectra are reported in Table I. The second column (labeled "Harm.") gives the harmonic frequencies for the analytic potential and comparison with the following one shows the degree of anharmonicity. Comparison with the third and fourth column shows that a 3200 trajectory simulation is already at convergence thanks to the time-averaging filter. Then, by comparing the frequencies in the following columns, it is apparent that as the sampling parameters **a**, **b** are increased and the distribution narrowed, the accuracy of the semiclassical calculation for the highest vibrational states decreases. However, discrepancies are always contained within $10-20\,\mathrm{cm}^{-1}$ and with respect to the vibrational eigenvalues this is an uncertainty of about a 1–2 per thousand. These results show that only drastic changes in the Husimi sampling will significantly bias the power spectra results and reasonable sampling variations (up to a multiple of 5 in the sampling standard deviation) will leave the spectra invariant.

A comparison between the current semiclassical levels and the quantum mechanical ones is not given in Table I because this was already done in a previous paper[55] and because, here, we are interested in comparing the approximation for the SC-IVR method with the original SC-IVR formulation. Any agreement between our approximated SC-IVR levels and the quantum mechanical, better than that found with the original SC-IVR, would be completely fortuitous.

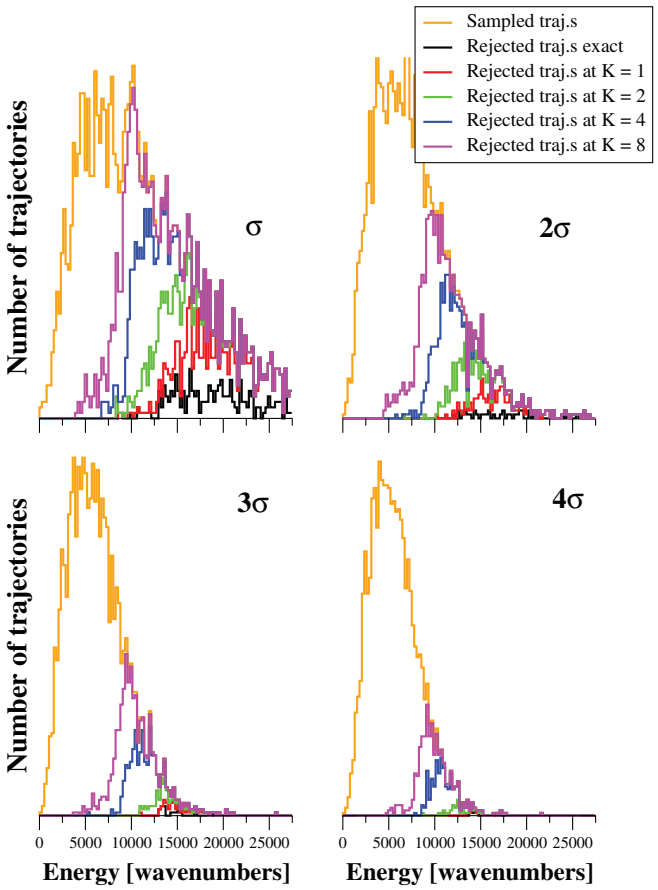

FIG. 4. Dependence of the distribution of rejected sampled trajectories with respect to the value of K for the CFD interpolated Hessian. Orange line is for the total number of sampled trajectories and other color lines are for the rejected trajectories. The sampling width is $n\sigma = (na, nb)$ where n is an integer.

As discussed above, trajectories with a monodromy matrix determinant deviation from unity greater than 10^{-4} were discarded. Reported in Figure 4 are the energy distributions of the discarded trajectories for different widths of the Husimi sampling distribution and different levels of the CFD approximation. The energy distribution for the total number of trajectories is given by the orange line in each panel, while the distributions for the discarded ones are given by the other colors, depending on the level of the CFD approximation. A higher energy trajectory is more probable to visit chaotic regions of the potential and, consequently, is more probable to be rejected. By comparing the panels in Figure 4, it is seen that fewer trajectories are rejected when the sampling is peaked at smaller energy values, i.e., the Husimi sampling coefficients **a** and **b** are increased from the Husimi value of σ. In addition, within each panel it is apparent there are more rejected trajectories when the CFD interpolation interval K is increased**.** This shows that the effect of the CFD approximation is either to make the already negative local Hessian eigenvalues more negative or to introduce additional negative eigenvalues.[39]

Here, we want to understand when these higher rejection rates are prejudicial for the power spectrum calculation. As an example, the **a**, **b** sampling parameters are fixed at the Husimi value of σ and the water power spectra (shown in Figure 5)

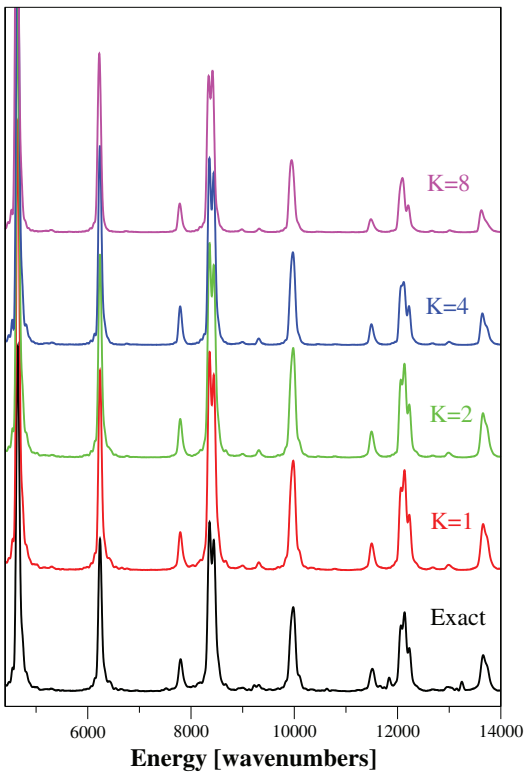

FIG. 5. Water power spectra, given by the colored lines, calculated with different levels of approximation of the CFD-Bofill Hessian update scheme. A total of 3200 trajectories are sampled according to the Husimi distribution with the **a**,**b** parameters set to σ. The black line is the spectrum without the Hessian approximation.

calculated for different K values. K is the number of consecutive integration time-steps for which the Hessian is estimated according to the CFD approximation, before calculating it accurately. The effect of increasing K is analogous to that of restricting the sampling width, i.e., higher energy trajectories are discarded and higher vibrational energy level peaks are less intense. Nevertheless, the spectra in Figure 5 appear quite similar.

To obtain a quantitative comparison of the above power spectra calculations, the vibrational eigenvalues obtained with different sampling parameters and different levels of the CFD approximation are listed in Table II, where they are compared with the "exact" semiclassical values. These results, together with the ones in Figure 5, show a quite accurate threshold of K = 4 for the CFD approximation. K = 4 shows very few peaks deviating by more than 20 cm^{-1} from the "exact" semiclassical simulation, while the deviation is about double for K = 8. Since we are employing a 4th order algorithm, with K = 4 the Hessian is evaluated 16 times less than for an "exact" simulation. For a semiclassical direct dynamics simulation, calculation of the Hessian dominates the computational time and, thus, with the CFD approximation the simulation is about 16 times faster than without. Another important observation from Table II is that, as the sampling parameters **a** and **b** are increased and the sampling distribution becomes more different from the Husimi one, the spectrum accuracy is poorer and some peaks are missing. Thus, direct application of the CFD approximation to the Husimi distribution is the preferred approach. Sampling distributions that are more narrow than the Husimi one may make the CFD approximation more accurate for some vibrational peaks, but the spectra precision for the higher vibrational peaks is substantially degraded.

C. MC-SC-IVR power spectra and the CFD approximation

As presented in Sec. III, a viable tool for an *ab initio* semiclassical calculation is the MC-SC-IVR approximation, where the number of trajectories is reduced to very few. In this section we study application of the CFD approximation to a MC-SC-IVR direct dynamics calculation. As described in the Appendix, for the MC-SC-IVR method the trajectory initial conditions are chosen by performing Box-Muller[76] samplings

TABLE II. Water vibrational energy levels.[a]

$\nu_1\nu_2\nu_3$ traj.s	σ 32 000	σ K = 4	σ K = 8	2σ K = 4	2σ K = 8	3σ K = 4	3σ K = 8	4σ K = 4	4σ K = 8	5σ K = 4	5σ K = 8
000 (A_1)	4644	4642	4627	4641	4627	4639	4626	4646	4623	4645	4635
100 (A_1)	6236	6235	6223	6230	6218	6233	6222	6234	6231	6234	6225
200 (A_1)	7793	7790	7782	7792	7777	7791	7791	7791	7797	7785	7769
010 (A_1)	8361	8356	8341	8350	8334	8350	8333	8359	8333	8352	8339
001 (B_2)	8441	8439	8416	8441	8427	8441	8424	8446	8426	8445	8429
300 (A_1)	9314	9312	9321	9300	9329	9313	9335	9297	9286*	9305	9298
110 (A_1)	9933*	9923*		9942	9925	9934*		9929*		9988	9915
101 (B_2)	9974	9973	9948	9987	9974*	9998	9959	9984	9961	9931	9974
210 (A_1)	11 505	11 492*	11 485*	11 510	11 506	11 494	11 500	11 508	11 483	11 510	11 509
201 (B_2)	11 666			11 688		11 573		11 617		11 646*	11 637
020 (A_1)	12 053*	12 072	12 072*	12 054	12 041*	12 055	12 038	12 061	12 039*	12 060	12 039*
011 (B_2)	12 134	12 118	12 094	12 128	12 109	12 129	12 109	12 146	12 118*	12 149	12 117
002 (A_1)	12 226*	12 221	12 206	12 220	12 205	12 217	12 197	12 221	12 213*	12 235	12 209
120 (A_1)	13 513*	13 607*		13 604*				13 595*		13 610*	13 602*
111 (B_2)	13 667	13 643	13 623	13 672	13 662	13 682	13 648	13 675	13 658	13 673	13 661
102 (A_1)	13 750*	13 716	13 725*	13 768*	13 776*	13 779	13 761*	13 744		13 753	13 743

[a]The 2nd column is for the 32 000 trajectory simulation from Table I. The remaining columns are for the 3200 trajectory simulations in this table. Uncertain peaks are marked with (*).

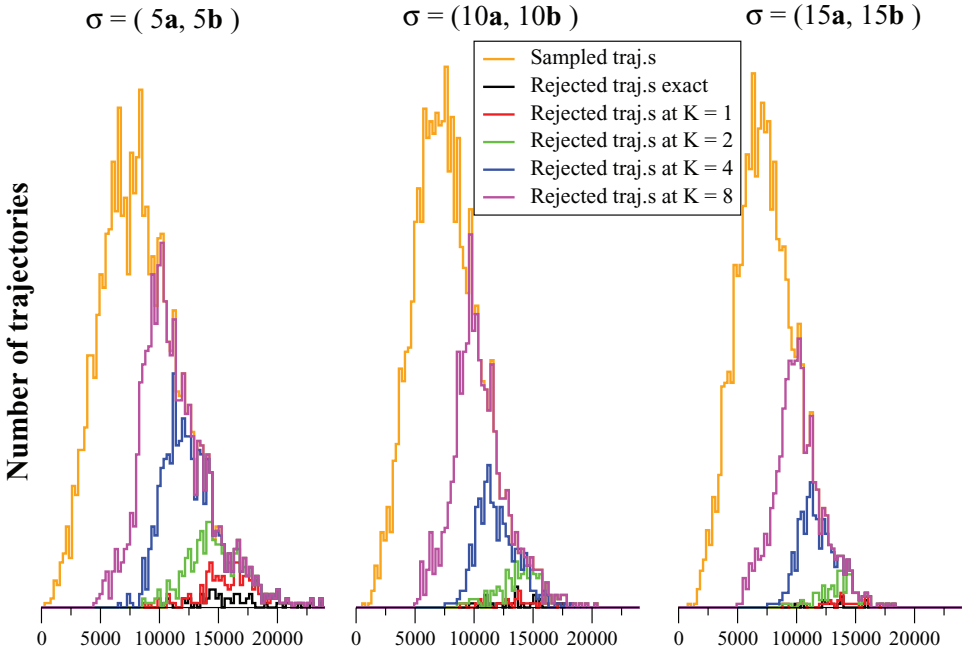

FIG. 6. Energy distributions for sampled and rejected trajectories for multiple coherent states sampling of the water molecule with eight Gaussian sampling centers (see text). The orange line is the distribution for the total trajectories and the other lines are for the rejected trajectories at different levels of the CFD approximation.

centered at each coherent state per time. Figure 6 shows the energy histogram distribution for the sampled trajectories using the same water potential as above. Also shown are the distributions of rejected trajectories without the CFD approximation and with different levels of the CFD approximation. Each panel gives the results for a fixed sampling standard deviation, equal for all centers. The original MC-SC-IVR formulation, where the number of trajectories is equal to a small number (few) coherent states as given by Eq. (10), corresponds to a sampling standard deviation equal to infinity, i.e., only the centers $(\mathbf{p}^i_{eq}, \mathbf{q}^i_{eq})$ are sampled. For this reason, in Figure 6 finite large values of the standard sampling deviation are considered to have a gradual approach to the few trajectory MC-SC-IVR sampling case. The results in Figure 6 are very similar to the above SC-IVR calculations; i.e., the wider the sampling distribution the higher the number of rejected trajectories when employing the CFD approximation. The main difference with respect to the sampling results of Figure 4 is that for the larger standard deviations the single-center Husimi sampling distribution is biased around the ZPE energy part of the spectrum, while the MC-SC-IVR sampling procedure keeps an even sampling distribution across the power spectrum. This last kind of sampling is clearly going to be more representative of the power spectrum as compared to an arbitrarily chosen single-center Husimi standard deviation.

Details of the MC-SC-IVR calculations may be appreciated by inspecting the power spectra. In Figure 7 power spectra of the water molecule are given for MC-SC-IVR simulations with both 8 and 3200 trajectories. The coherent states in Eq. (10) are restricted to the lower eight vibrational levels, including the ZPE one, which are represented by the harmonic approximation as described in Ref. 55. If only 8 trajectories are used, we obtain the black line spectrum at the

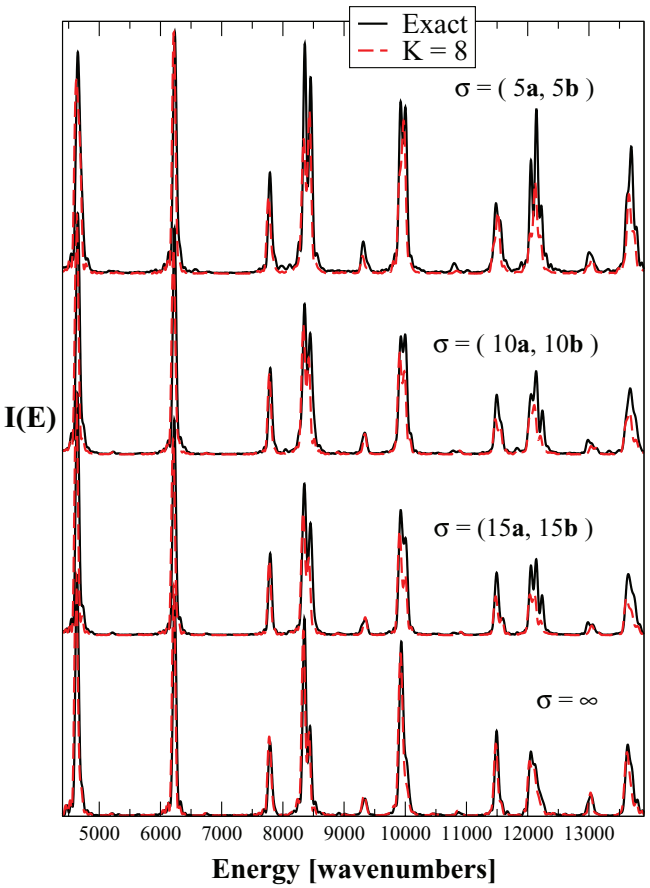

FIG. 7. Water power spectra for MC-SC-IVR simulations with different sampling parameters. For each set of parameters the K = 8 CFD approximated spectrum (red line) is compared with the MC-SC-IVR spectrum without the CFD approximation (black line). The $\sigma = \infty$ is the 8-trajectory simulation.

TABLE III. Water vibrational eigenvalues with the CFD approximation included with the MC-SC-IVR method.[a]

$\nu_1\nu_2\nu_3$			5σ	5σ	10σ	10σ	10σ	15σ	15σ	15σ	∞	∞	∞
CFD	σ	5σ	K = 4	K = 8		K = 4	K = 8		K = 4	K = 8		K = 4	K = 8
traj.s	32 000	3200	3200	3200	3200	3200	3200	3200	3200	3200	8	8	8
000 (A_1)	4644	4648	4648	4625	4651	4645	4631	4642	4639	4623	4639	4634	4620
100 (A_1)	6236	6234	6228	6218	6230	6225	6212	6226	6221	6207	6231	6228	6216
200 (A_1)	7793	7789	7784	7765	7794	7789	7777	7793	7789	7777	7790	7787	7777
010 (A_1)	8361	8358	8354	8340	8354	8350	8335	8351	8349	8331	8353	8347	8336
001 (B_2)	8441	8454	8447	8434	8447	8444	8421	8451	8448	8422	8444	8439	8426
300 (A_1)	9314	9314	9313	9301	9337	9341	9339	9350	9348	9344	9341	9338	9327
110 (A_1)	9933*	9925	9917	9906*	9928	9928	9908	9926	9923	9907	9935	9931	9922
101 (B_2)	9974	10 001	9997	9973	9998	9991	9985	10 003	9999	9998	10 008*	10 008*	10 020*
210 (A_1)	11 505	11 483	11 483	11 450*	11 494*	11 484	11 473	11 493	11 489	11 476	11 493	11 488	11 478
201 (B_2)	11 666	11 545	11 527	11 513	11 559	11 562	11 548	11 592	11 590	11 576	11 569*	11 569*	11 550
020 (A_1)	12 053*	12 053	12 051	12 040	12 054	12 057	12 043	12 054	12 052	12 036	12 052	12 043	12 030
011 (B_2)	12 134	12 142	12 136	12 120	12 139	12 128	12 108	12 138	12 132	12 112	12 121*	12 111*	12 093*
002 (A_1)	12 226*	12 222	12 221	12 201	12 234	12 229	12 206	12 234	12 230	12 210	12 226*	12 241	12 232
120 (A_1)	13 513*	13 629*	13 628*		13 610*	13 610*	13 609*				13 633	13 631	13 622
111 (B_2)	13 667	13 692	13 689	13 652	13 674	13 660	13 666	13 643	13 638	13 613	13 701		
102 (A_1)	13 750*	13 775	13 764	13 726	13 750	13 770*	13 668*	13 731*	13 727*	13 687*	13 791*	13 785	13 774

[a]First row gives the type of sampling, the second gives the level of the CFD approximation, and the third gives the number of trajectories. If the column is blank the CFD approximation is not employed. The last three columns refer to the 8-trajectory MC-SC-IVR calculations. Uncertain peaks are marked with (*).

bottom of Figure 7 and if the CFD approximation with K = 8 is additionally invoked we obtain the red line spectrum. From the bottom spectra of Figure 7 it is seen that the CFD approximation is not introducing any significant deviation to the 8-trajectory MC-SC-IVR approximation. The situation is different when 3200 trajectories are employed for the MC-SC-IVR, as represented by the other spectra in the figure.

To understand how the CFD approximation should be applied for a MC-SC-IVR direct dynamics simulation, water vibrational eigenvalues are reported in Table III for several sets of sampling parameters and K values. The first row reports the sampling choice, the second row reports the level of the CFD approximation (if blank, the approximation was not employed), and the third row gives the number of trajectories. To assist in the comparison, in the second column of Table III the eigenvalues from Table I are reported for the "exact" 32 000 trajectory SC-IVR simulation. The remaining columns give the eigenvalues for different sampling widths without the CFD approximation and with the K = 4 and K = 8 CFD approximation. If the CFD approximation is not applied, the 8-trajectory MC-SC-IVR simulation and the 5σ sampled one are of the same accuracy and comparable with the 32 000 SC-IVR trajectory simulation, while the others are of lower accuracy. The accuracy of the 5σ sampling 3200 trajectory MC-SC-IVR simulation in Table III can be explained by considering that the sampled phase space is quite similar to the 3200 trajectory Husimi sampling SC-IVR one. However, it is quite unexpected that the 8-trajectory MC-SC-IVR simulation performs better than those with 3200 trajectories when sampling is wider than 5σ. This may be due to the *ad hoc* MC-SC-IVR sampling procedure adopted here and described in the Appendix. As far as the accuracy of the CFD approximation is concerned, we find again that using K = 4 does not significantly change the spectrum for any of the sampling choices. This means that the threshold for the CFD approximation can be chosen irrespective of the SC-IVR version used and, by comparing with the results of Ref. 39, this is also the case for a single trajectory simulation.

D. First principles semiclassical direct dynamics results

The ultimate goal of approximating the Hessian in SC-IVR molecular dynamics is to enhance *ab initio* semiclassical molecular dynamics. In this section, an illustrative example is reported of the power spectrum calculation for carbon dioxide. This spectrum is quite a challenging one, given the strong anharmonic couplings between the symmetric and bending modes and the numerous Fermi resonances. First-principles classical trajectories and Hessian calculations were performed at the level of B3LYP/cc-pVDZ level of theory with the VENUS package[80,81] interfaced with NWCHEM.[82] A version of the integrated VENUS/NWCHEM was developed to determine all the classical quantities needed for the semiclassical calculations. The simulations were performed in normal mode coordinates, after an initial Hessian diagonalization. During the time-evolution the Hessian is calculated every K time-steps, as illustrated above.

Figure 8 shows "on-the-fly" CO_2 power spectra calculated for various values of K. Major difference is apparent at higher energies, where some peaks significantly lose intensity at the level of spectral noise. In order to better appreciate these differences, the vibrational energy values in cm^{-1} are reported in Table IV. In the first column of the Table, the "traditional" terminology is used to label the vibrational levels, where the first number refers to the symmetric stretch, the second to the bending modes and the last one to the asymmetric stretch. Fermi resonances occur between the symmetric stretch and

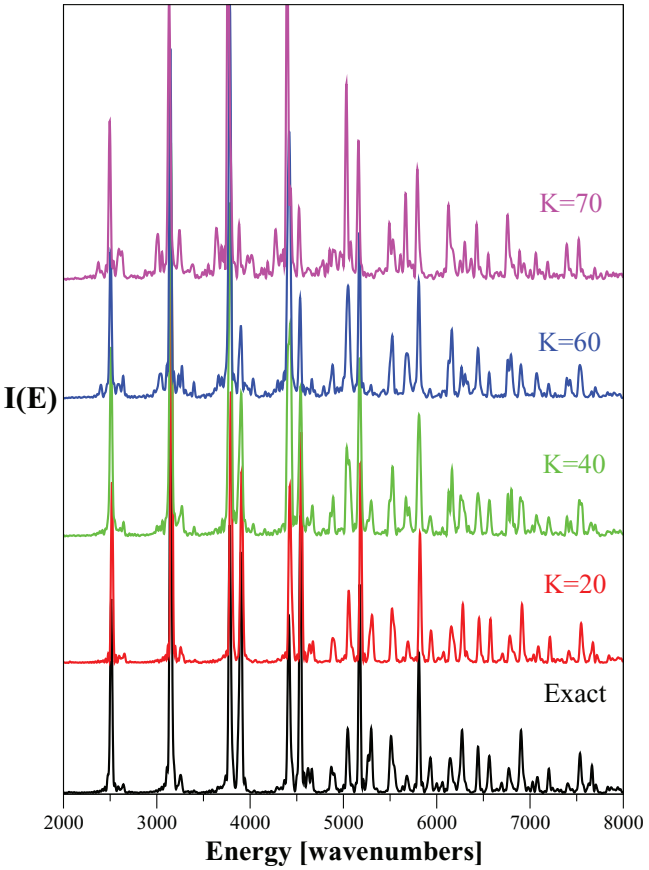

FIG. 8. Direct dynamics power spectra for carbon dioxide calculated at different levels of the CFD-Bofill Hessian update scheme. The simulations were performed with the MC-SC-IVR method using 8 trajectories. The black line is the spectrum without the Hessian CFD approximation and the colored lines are for different levels of the CFD Hessian approximation.

bending modes, as previously reported.[55,59] These Fermi coupled states are denoted by superscript symbols. In the second column, the vibrational levels in the harmonic approximation are reported, where several degeneracies are present which are removed by the Fermi couplings. The third column gives the MC-SC-IVR results obtained with 8 classical trajectories. For the trajectory initial conditions, the coordinates are not displaced from equilibrium and momenta is added to the modes so that the first eight vibrational levels are harmonically spaced (see discussion above in Sec. III). Table IV shows that, by using Bofill's algorithm for the CFD approximation, only at K = 70 does the power spectrum start to have missing peaks, including deviations of the peak positions on the order of 10 cm^{-1} or greater. In terms of *ab initio* molecular dynamics, the use of K = 70 corresponds to a two order of magnitude decrease in computational time for a velocity Verlet algorithm, since calculating the Hessian requires most of this time.

TABLE IV. MC-SC-IVR vibrational spectra for CO_2 obtained with direct dynamics and the CFD Hessian approximation.[a]

States	Harmonic	Exact	K = 10	K = 20	K = 40	K = 60	K = 70
ZPE	2538	2512	2518	2520	2510	2504	2495
$0, 1^1, 0$	3190	3148	3156	3158	3150	3145	3132
$0, 2^0, 0^\star$	3843	3782	3792	3794	3790	3784	3763
$1, 0^0, 0^\star$	3895	3901	3910	3910	3905	3901	3882
$0, 3^1, 0^\dagger$	4495	4415	4425	4425	4430	4419	4398
$0, 3^3, 0$	4495	4537	4547	4548	4542	4536	4525
$1, 1^1, 0^\dagger$	4547	4622	4630	4637	4614		
$0, 0^0, 1$	4950	4870	4879	4886	4990	4884	4852*
$0, 4^0, 0^\ddagger$	5148	5046	5056	5058	5036	5051	5031
$0, 4^2, 0^\S$	5148	5174	5182	5185	5175	5173	5161
$0, 4^4, 0^\ddagger$	5148	5298	5307	5306	5298	5297	
$1, 2^2, 0^\S$	5200	5511	5519	5521	5526	5524	5492
$2, 0^0, 0^\ddagger$	5252	5679	5690	5691	5671	5679	5668
$1, 0^0, 1$	6307	5829	5819	5820	5810	5808	5791

[a]The first column gives the vibrational spectroscopic terms, with Fermi resonance states indicated by a distinct superscript symbol for each group of resonant states. The second column gives the harmonic results and the third gives the semiclassical energy levels without a Hessian approximation. The remaining columns give the energy levels obtained with different levels of the CFD Hessian approximation according to Bofill's algorithm. The energies are in cm^{-1}. Uncertain peaks are marked with (*).

VI. DISCUSSION

Of the many methods elaborated in the past to enhance the monodromy matrix calculation, we have found that only the one of Garashchuk and Light[83] does not invoke an additional approximation beyond a numerical one. For this reason, their approximation may be directly compared to the present CFD one. Garashchuk and Light proposed to calculate the second derivatives for the Hessian by generating 2N (where N is the number of degrees of freedom) auxiliary classical trajectories, whose initial conditions differ from the main reference trajectory by a displacement in one of the phase space variables.[83] Then, the monodromy matrix was calculated as the product of monodromy matrices for infinitesimal time-steps. The overall computational cost scales linearly instead of quadratically.

We find the present CFD approximation to offer several advantages with respect to the Garashchuk and Light method, which, however, has the desirable property to be numerical. First, the CFD method allows for "tuning" the approximation by varying the number of consecutive time intervals for which the Hessian is estimated, i.e., by changing the value of K. This feature allows any user to set her/his level of approximation, according to the accuracy desired for observables of interest. Second, the CFD approximation does not require any further calculation of auxiliary trajectories, except those used for the statistical convergence of the Monte Carlo integration. Third, Garashchuk and Light formulation of the second derivatives is accurate up to first order. In contrast, the CFD approximation is higher order, depending on the choice of the algorithm. For these reasons, our understanding is that the present numerical approximation of the SC-IVR propagator represents a significant step forward with respect to other monodromy matrix approximations.[11,19,50,52,83]

VII. CONCLUSIONS

A semiclassical (SC) IVR calculation requires numerical integration of classical trajectories for evaluations of the monodromy matrix versus time. Evaluating this matrix requires the Hessian, which is computationally quite expensive for di-

rect dynamics simulations. The work reported here shows that the CFD approximation for the Hessian is a very useful approach that can significantly reduce the computational time required for a SC direct dynamics calculation. The CFD approximation may decrease the computational time required for a SC-IVR direct dynamics calculation by orders of magnitude. Such an enhancement in the efficiency of the calculation makes SC-IVR direct dynamics computationally tractable for a broad range of problems.

ACKNOWLEDGMENTS

The University of Milan is thanked for funding ("5 per mille" grant) and CILEA (Consorzio Interuniversitario Lombardo per L'Elaborazione Automatica) for computational time allocation. The Texas Tech contribution to this research is based upon work supported by the National Science Foundation under Grant Nos. CHE-0957521 and OISE-0730114, and the Robert A. Welch Foundation under Grant No. D-0005. Support was also provided by the High Performance Computing Center (HPCC) at Texas Tech University, under the direction of Philip W. Smith.

APPENDIX: MONTE CARLO SAMPLING

The natural Monte Carlo sampling distribution of Eq. (8) is given by the value of the integrand at time zero. Given a reference state $|\chi\rangle = |\mathbf{p}_{eq}, \mathbf{q}_{eq}\rangle$, this is proportional to

$$\langle\chi|\mathbf{p}(0), \mathbf{q}(0)\rangle\langle\mathbf{p}(0), \mathbf{q}(0)|\chi\rangle = e^{-\mathbf{a}(\mathbf{q}_0-\mathbf{q}_{eq})^2-\mathbf{b}(\mathbf{p}_0-\mathbf{p}_{eq})^2}, \quad \text{(A1)}$$

where the vectors $\mathbf{a}$ and $\mathbf{b}$ are such that the ith element is $a(i) = \omega_i/2$ and $b(i) = 1/(2\omega_i)$, and ω_i is the harmonic vibrational frequency of the ith normal mode. This phase space probability distribution is called the Husimi distribution.[74] Recently, an efficient time-dependent importance sampling approach has been advanced for SC-IVR molecular dynamics of time correlation functions.[75] Here, we have shown how the power spectra depends on the sampling parameters $\mathbf{a}$ and $\mathbf{b}$. A correct and converged Monte Carlo sampling is such that the integration results do not depend on the sampling parameters. However, since the CFD Hessian approximation is invoked in the prefactor calculation, trajectories are run and the contribution of a trajectory is removed if the determinant of the matrices product $\mathbf{M}^T\mathbf{M}$ deviates from unity more than a given threshold. In principle, when using a symplectic propagator, the deviation from unity should be within a round-off error.[14(c)] However, we have previously proved that the effect of the CFD approximation may introduce negative Hessian eigenvalues that cause the monodromy matrix elements to become significantly large in magnitude, with an associated round-off error as well. For this reason, more trajectories are rejected as the CFD approximation becomes more severe, and the SC-IVR sampling is consequently biased. Here we have shown how the power spectra and the number of rejected trajectories change by changing the sampling parameters and the level of the CFD approximation.

For a MC-SC-IVR calculation of Eq. (11), the reference state is given by Eq. (10) and, consequently, there are multiple sampling centers. This is seen from the integrand expression at zero-time,

$$\langle\chi \mid \mathbf{p}(0), \mathbf{q}(0)\rangle\langle\mathbf{p}(0), \mathbf{q}(0)\mid\chi\rangle = \sum_i e^{-\mathbf{a}(\mathbf{q}_0-\mathbf{q}^i_{eq})^2-\mathbf{b}(\mathbf{p}_0-\mathbf{p}^i_{eq})^2}, \quad \text{(A2)}$$

where the sum runs over the coherent states' centers. As explained above, the MC-SC-IVR approximation was introduced to drastically reduce the number of trajectories for "on the fly" semiclassical calculations. The minimum number of trajectories is equal to the number of coherent states, as given by Eq. (12). In the examples quoted above, these have never been more than a few trajectories. The need for such a low number of trajectories opens up the possibility of *ab initio* semiclassical direct dynamics calculations.[55,57,59] This choice is equivalent to putting Dirac-delta distributions at the positions of the Gaussian distributions in Eq. (A2) and initiating trajectories one by one from each $(\mathbf{q}^i_{eq}, \mathbf{p}^i_{eq})$ phase space point. In contrast, to sample trajectories according to Eq. (A2), positions and momenta should be randomly chosen within given intervals. Equation (A2) may then be sampled, with the phase space point rejected or accepted by the von Neumann rejection technique.[77] This acceptance/rejection algorithm heavily relies on the initial sampling interval, whose unbiased determination is postponed for future work. Instead, in this paper, we look for a close comparison between the MC-SC-IVR calculations of Eq. (12) and those of Eq. (11). In order to accomplish this, we found it reasonable to have the trajectories sampled separately from each Gaussian distribution of Eq. (A2) by a Box-Muller method.[76]

[1] W. H. Miller, J. Chem. Phys. **53**, 3578 (1970); **53**, 1949 (1970); J. Phys. Chem. A **105**, 2942 (2001); Adv. Chem. Phys. **25**, 69 (1974).
[2] W. H. Miller, Proc. Natl. Acad. Sci. U.S.A. **102**, 6660 (2005).
[3] K. G. Kay, Annu. Rev. Phys. Chem. **56**, 255 (2005).
[4] W. H. Miller, J. Chem. Phys. **125**, 132305 (2006).
[5] K. G. Kay, Chem. Phys. **322**, 3 (2006).
[6] T. Sklarz and K. G. Kay, J. Chem. Phys. **120**, 2606 (2004).
[7] C. Harabati and K. G. Kay, J. Chem. Phys. **127**, 084104 (2007).
[8] G. Hochman and K. G. Kay, J. Chem. Phys. **130**, 061104 (2009).
[9] M. F. Herman, Annu. Rev. Phys. Chem. **45**, 83 (1994).
[10] N. Makri, Annu. Rev. Phys. Chem. **50**, 167 (1999).
[11] S. Zhang and E. Pollak, J. Chem. Phys. **121**, 3384 (2004); J. Chem. Theory Comput. **1**, 345 (2005).
[12] J. Shao and E. Pollak, J. Chem. Phys. **125**, 133502 (2006).
[13] E. Pollak and E. Martin-Fierro, J. Chem. Phys. **126**, 164107 (2007); E. Martin-Fierro and E. Pollak, *ibid.* **125**, 164104 (2006).
[14] (a) A. R. Walton and D. E. Manolopoulos, Mol. Phys. **87**, 961 (1996); (b) Chem. Phys. Lett. **244**, 448 (1995); (c) M. L. Brewer, J. S. Hulme, and D. E. Manolopoulos, J. Chem. Phys. **106**, 4832 (1997).
[15] S. Bonella, D. Montemayor, and D. F. Coker, Proc. Natl. Am. Soc. **102**, 6715 (2005).
[16] S. Bonella and D. F. Coker, J. Chem. Phys. **118**, 4370 (2003).
[17] C. Harabati, J. M. Rost, and F. Grossmann, J. Chem. Phys. **120**, 26 (2004); F. Grossmann, Comments At. Mol. Phys. **34**, 141 (1999).
[18] T. F. Viscondi and M. A. M. de Aguiar, J. Chem. Phys. **134**, 234105 (2011).
[19] B. B. Issack and P. N. Roy, J. Chem. Phys. **127**, 054105 (2007).
[20] C. Venkataraman, J. Chem. Phys. **135**, 204503 (2011).
[21] H.-D. Meyer, U. Manthe, and L. S. Cederbaum, Chem. Phys. Lett. **165**, 73 (1990); U. Manthe, H.-D. Meyer, and L. S. Cederbaum, J. Chem. Phys. **97**, 3199 (1992); M. H. Beck, A. Jaäckle, G. A. Worth, and H.-D. Meyer, Phys. Rep. **324**, 1 (2000); H. Wang and M. Thoss, J. Chem. Phys. **119**, 1289 (2003).
[22] D. V. Shalashilin and M. S. Child, J. Chem. Phys. **113**, 10028 (2000); Chem. Phys. **304**, 103 (2004); J. Chem. Phys. **128**, 054102 (2008).
[23] M. J. Davis and E. J. Heller, J. Chem. Phys. **71**, 3383 (1979).

[24]E. J. Heller, J. Chem. Phys. **62**, 1544 (1975); **75**, 2923 (1981).
[25]J. Tatchen, E. Pollak, G. Tao, and W. H. Miller, J. Chem. Phys. **134**, 134104 (2011).
[26]W. Chen, W. L. Hase, and H. B. Schlegel, Chem. Phys. Lett. **228**, 436 (1994).
[27]J. M. Millam, V. Bakken, W. Chen, W. L. Hase, and H. B. Schlegel, J. Chem. Phys. **111**, 3800 (1999).
[28]L. Sun and W. L. Hase, Rev. Comput. Chem. **19**, 79 (2003).
[29]L. Sun, K. Song, and W. L. Hase, Science **296**, 875 (2002).
[30]T. Zimmermann, J. Ruppen, B. Li, and J. Vaníček, Int. J. Quantum Chem. **110**, 2426 (2010).
[31]M. Ben-Nun and T. J. Martinez, Adv. Chem. Phys. **121**, 439 (2002); B. G. Levine, J. D. Coe, A. M. Virshup, and T. J. Martinez, Chem. Phys. **347**, 3 (2008); J. D. Coe, B. G. Levine, and T. J. Martinez, J. Phys. Chem. **111**, 11302 (2007).
[32]R. M. Wentzcovitch and J. L. Martins, Solid State Commun. **78**, 831 (1991).
[33]D. Marx and J. Hutter, in *Modern Methods and Algorithms of Quantum Chemistry*, 2nd ed., edited by J. Grotendorst (NIC, FZ Jülich, 2000), Vol. 3, pp. 301–449.
[34]M. F. Herman and E. Kluk, Chem. Phys. **91**, 27 (1984).
[35]K. G. Kay, J. Chem. Phys. **100**, 4377 (1994); **100**, 4432 (1994).
[36]H. Wang, X. Sun, and W. H. Miller, J. Chem. Phys. **108**, 9726 (1988); T. Yamamoto, H. Wang, and W. H. Miller, *ibid.* **116**, 7335 (2002); T. Yamamoto and W. H. Miller, *ibid.* **118**, 2135 (2003).
[37]H. Ushiyama and K. Takatsuka, J. Chem. Phys. **122**, 224112 (2005); S. Takahashi and K. Takatsuka, *ibid.* **127**, 084112 (2007).
[38]H. Wu, M. Rahman, J. Wang, U. Louderaj, W. L. Hase, and Y. Zhuang, J. Chem. Phys. **133**, 074101 (2010).
[39]Y. Zhuang, M. R. Siebert, W. L. Hase, K. G. Kay, and M. Ceotto, J. Chem. Theory Comput. **9**, 54 (2013).
[40]K. G. Kay, J. Chem. Phys. **101**, 2250 (1994).
[41]X. Sun, H. Wang, and W. H. Miller, J. Chem. Phys. **109**, 7064 (1998).
[42]W. H. Miller, Faraday Discuss. Chem. Soc. **110**, 1 (1998).
[43]W. H. Miller, J. Phys. Chem. A **103**, 9384 (1999).
[44]J. Liu and W. H. Miller, J. Chem. Phys. **125**, 224104 (2006); **126**, 234110 (2007); **127**, 114506 (2007); **128**, 144511 (2008).
[45]I. Navrotskaya and E. Geva, J. Phys. Chem. A **111**, 460 (2007); B. K. Ka, Q. Shi, and E. Geva, *ibid.* **109**, 5527 (2005); F. X. Vazquez, S. Talapatra, and E. Geva, *ibid.* **115**, 9775 (2011).
[46]X. Sun and W. H. Miller, J. Chem. Phys. **110**, 6635 (1999); H. Wang, M. Thoss, K. Sorge, R. Gelabert, X. Gimenez, and W. H. Miller, *ibid.* **114**, 2562 (2001); R. Gelabert, X. Gimenez, M. Thoss, H. Wang, and W. H. Miller, *ibid.* **114**, 2572 (2001); M. Thoss, H. Wang, and W. H. Miller, *ibid.* **114**, 9220 (2001); H. Wang, D. E. Manolopoulos, and W. H. Miller, *ibid.* **115**, 6317 (2001).
[47]K. Thompson and N. Makri, Phys. Rev. E **59**, R4729 (1999); J. Shao and N. Makri, J. Phys. Chem. A **103**, 7753 (1999); **103**, 9479 (1999).
[48]V. S. Filinov, Nucl. Phys. B **271**, 717 (1986); N. Makri and W. H. Miller, Chem. Phys. Lett. **139**, 10 (1987); J. D. Doll, D. L. Freeman, and T. L. Beck, Adv. Chem. Phys. **78**, 61 (1994); S. M. Anderson, D. Neuhauser, and R. Baer, J. Chem. Phys. **118**, 9103 (2003).
[49]N. T. Maitra, J. Chem. Phys. **112**, 531 (2000).
[50]R. Gelabert, X. Gimenez, M. Thoss, H. Wang, and W. H. Miller, J. Phys. Chem. A **104**, 10321–10327 (2000).
[51]B. Harland and P.-N. Roy, J. Chem. Phys. **118**, 4791 (2003).
[52]V. Guallar, V. S. Batista, and W. H. Miller, J. Chem. Phys. **110**, 9922 (1999).
[53]A. L. Kaledin and W. H. Miller, J. Chem. Phys. **118**, 7174 (2003); **119**, 3078 (2003).
[54]Y. Elran and K. G. Kay, J. Chem. Phys. **110**, 3653 (1999); **110**, 8912 (1999).
[55]M. Ceotto, S. Atahan, G. F. Tantardini, and A. Aspuru-Guzik, J. Chem. Phys. **130**, 234113 (2009).
[56]M. Ceotto, S. Valleau, G. F. Tantardini, and A. Aspuru-Guzik, J. Chem. Phys. **134**, 234103 (2011).
[57]M. Ceotto, G. F. Tantardini, and A. Aspuru-Guzik, J. Chem. Phys. **135**, 214108 (2011).
[58]M. Ceotto, D. dell'Angelo, and G. F. Tantardini, J. Chem. Phys. **133**, 054701 (2010).
[59]M. Ceotto, S. Atahan, S. Shim, G. F. Tantardini, and A. Aspuru-Guzik, Phys. Chem. Chem. Phys. **11**, 3861 (2009).
[60]C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," Math. Comput. **19**, 577 (1965).
[61]M. J. D. Powell, "Recent advances in unconstrained optimization," Math. Program. **1**, 26 (1971).
[62]J. E. Dennis, Jr. and J. J. Moré, "Quasi-Newton methods, motivation and theory," SIAM Rev. **19**, 46 (1977).
[63]J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, Englewood Cliffs, NJ, 1983).
[64]J. Nocedal, "Theory of algorithms for unconstrained optimization," Acta Numerica **1**, 199 (1992).
[65]J. M. Bofill, "Updated Hessian matrix and the restricted step method for locating transition structures," J. Comput. Chem. **15**, 1 (1994).
[66]V. Bakken, J. M. Millam, and H. B. Schlegel, "Ab initio classical trajectories on the Born-Oppenheimer surface: Updating methods for Hessian-based integrators," J. Chem. Phys. **111**, 8773 (1999).
[67]H. P. Hratchian and H. B. Schlegel, "Using Hessian updating to increase the efficiency of a Hessian based predictor-corrector reaction path following method," J. Chem. Theory Comput. **1**, 61 (2005).
[68]U. Lourderaj, K. Song, T. L. Windus, Y. Zhuang, and W. L. Hase, "Direct dynamics simulations using Hessian-based predictor-corrector Integration Algorithms," J. Chem. Phys. **126**, 044105 (2007).
[69]S. Lele, "Compact finite difference schemes with spectral-like resolution," J. Comput. Phys. **103**, 16 (1992).
[70]R. E. Lynch and J. R. Rice, "High accuracy finite difference approximation to solutions of elliptic partial differential equations," Proc. Natl. Acad. Sci. U.S.A. **75**, 2541 (1978).
[71]I. Singer and E. Turkel, "High-order finite difference methods for the Helmholtz equation," Comput. Methods Appl. Mech. Eng. **163**, 343 (1998).
[72]Yu Zhuang and X.-H. Sun, "A high order ADI method for separable generalized Helmholtz equations," Adv. Eng. Software **31**, 585 (2000).
[73]Yu Zhuang and X.-H. Sun, "A high-order fast direct solver for singular Poisson equations," J. Comput. Phys. **171**, 79 (2001).
[74]K. Husimi, Proc. Phys. Math. Soc. Jpn. **22**, 264 (1940); K. Takahashi and N. Saito, Phys. Rev. Lett. **55**, 645 (1985); R. C. Brown and E. J. Heller, J. Chem. Phys. **75**, 186 (1981); E. J. Heller, *ibid.* **65**, 1289 (1976); M. J. Davis and E. J. Heller, *ibid.* **80**, 5036 (1984).
[75]G. Tao and W. H. Miller, J. Chem. Phys. **135**, 024104 (2011); **137**, 124105 (2012).
[76]W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recepies in Fortran: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1992).
[77]J. von Neumann, Natl. Bur. Stand. Appl. Math. Ser. **12**, 36 (1951).
[78]S. Dressler and W. Thiel, Chem. Phys. Lett. **273**, 71 (1997).
[79]A. Odell, A. Delin, B. Johansson, N. Bock, M. Challacombe, and A. M. N. Niklasson, J. Chem. Phys. **131**, 244106 (2009).
[80]W. L. Hase, R. J. Duchovic, X. Hu *et al.*, Quantum Chem. Program Exch. Bull. **16**, 671 (1996).
[81]X. Hu, W. L. Hase, and T. Pirraglia, J. Comput. Chem. **12**, 1014 (1991).
[82]E. Apra, T. L. Windus, T. P. Straatsma *et al.*, NWCHEM, version 5.0, a computational chemistry package for parallel computers, Pacific Northwest National Laboratory, Richland, Washington, 2007.
[83]S. Garashchuk and J. C. Light, J. Chem. Phys. **113**, 9390 (2000).