Describing nuclear quantum effects in vibrational properties using molecular dynamics with Wigner sampling

Denis S. Tikhonov[1,3] and Yury V. Vishnevskiy[2]

[1]Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany
[2]Lehrstuhl für Anorganische Chemie und Strukturchemie, Fakultät für Chemie,Universität Bielefeld, Universitätsstr. 25, 33615 Bielefeld, Germany
[3]denis.tikhonov@desy.de

March 6, 2023

Abstract

In this work, we develop a molecular dynamics method, combining the simple Maxwell-Boltzmann statistics with the generally applicable standard Wigner sampling. The resulting implemented single-parameter procedure, termed simplified Wigner sampling, has been extensively benchmarked by calculating vibrationally averaged rotational constants and vibrational spectra of several molecules and by comparing with results from other methods. The procedure has been also combined with the Andersen thermostat for simulation of NVT-ensembles.

1 Introduction

Molecular dynamics (MD) is a very powerful tool for modeling of molecular systems at normal and elevated temperatures $T \geq 300$ K.[1–3] However, the simulation of nuclear quantum effects (NQEs) is still a challenging topic in MD.[1] At normal temperatures, the path-integral molecular dynamics (PIMD) approach[1, 4, 5] is very effective for modelling equilibrium properties of molecular systems. This technique takes advantage of the exact mapping between the quantum-mechanical system and its representation as a set of N interconnected replicas (also termed as beads) with the N-times elevated temperature. Taking enough beads of this so-called ring polymer, when their respective temperature ($T \times N$) overcomes the Debye temperature $T_\mathrm{D} = h\nu/k_\mathrm{B}$ of the highest possible vibrational mode with frequency ν, the representation approaches the exact limit. The minimally required number of beads is thus estimated as $N_\mathrm{min} = \frac{h\nu}{k_\mathrm{B}T}$. Therefore at low temperatures, this number is getting very large thus diminishing the applicability of PIMD.

Rotational or vibrational spectroscopy techniques in the gas phase often applied to molecules in the ground vibrational state,[6] especially if cooling techniques, such as the supersonic expansion, are used.[7] The most computationally inexpensive and simple vibrational model is the harmonic approximation. This approach assumes that the potential energy surface (PES) in the vicinity of the equilibrium geometry (local minimum) can be expressed with the second-order Taylor expansion. For a diatomic molecule, in which the only vibrational coordinate is the interatomic distance r, it reads as

$$V(r) = \overbrace{V(r_\mathrm{e}) + \frac{V_2}{2}\xi^2}^{V_\mathrm{h}(r)} + \overbrace{\frac{V_3}{6}\xi^3 + \frac{V_4}{24}\xi^4 + \ldots}^{W_\mathrm{a}(r)} \approx V_\mathrm{h}(r)\ . \tag{1}$$

Here, V denotes the potential energy, r_e is the equilibrium distance between atoms, $\xi = r - r_\mathrm{e}$ is the displacement of the structure from the equilibrium geometry, $V_n = \frac{\partial^n V}{\partial r^n}(r_\mathrm{e})$, $V_\mathrm{h}(r)$ is the harmonic potential, and $W_\mathrm{a}(r)$ is the anharmonic correction to the harmonic potential. The quantum harmonic oscillator is one of the simplest quantum mechanical problems with known analytical solution.[8] For a molecule with reduced mass μ, the stationary vibrational states $|v\rangle$ ($v = 0, 1, 2, \ldots$) have energies of $E_v = \hbar\omega(v + 1/2)$, where $\omega = \sqrt{V_2/\mu}$ is the angular vibrational harmonic frequency. The vibrationally averaged displacement $\langle\xi\rangle = \langle v|\xi|v\rangle = 0$, since the harmonic potential is symmetric with respect to inversion of the displacement, $V_\mathrm{h}(+\xi) = V_\mathrm{h}(-\xi)$. On the other hand, the vibrational amplitude l, defined as

$$l^2 = \langle r^2\rangle - \langle r\rangle^2 \tag{2}$$

for the harmonic oscillator state $|v\rangle$ is greater than zero:

$$l_v^2 = \langle v|r^2|v\rangle - \langle v|r|v\rangle^2 = \langle v|\xi^2|v\rangle = \frac{\hbar}{\mu\omega}\left(v + \frac{1}{2}\right)\ . \tag{3}$$

The harmonic approximation combined with the frequency scaling is a very robust, computationally affordable, and simple way for calculation of vibrational properties.[9] For instance, the root-mean-squared deviation (RMSD) for the scaled harmonic frequencies from the experimentally-obtained fundamental transitions is only around 20 cm^{-1} at B3LYP/def2-TZVP level of theory.[10] However, the direct comparison of computed in this way theoretical values with experimental data may be of limited use, because real observables in the rotational or vibrational spectroscopy are generally anharmonic.[10, 11] For instance, in the classical mechanics the angular vibrational frequency can be approximately expressed as[12]

$$\omega_{\text{anh}} \approx \omega + \overbrace{\frac{\omega l^2}{8}\left(\frac{V_4}{V_2} - \frac{5}{3}\frac{V_3^2}{V_2^2}\right)}^{\Delta\omega} . \tag{4}$$

As one can see, the anharmonic shift $\Delta\omega$ is proportional to the square of the amplitude (l^2) and is also induced by the anharmonic terms of the PES Taylor expansion (Equation (1)). This dependence indicates larger deviations from the harmonic frequency at higher vibrational excitations. A similar situation is observed in the rotational spectroscopy, where rotational constants B can be considered as observables. For a diatomic molecule with the moment of inertia $I = \mu r^2$, dependent on the internuclear distance r, the rotational constant is given as[8][1]

$$B = \frac{\hbar}{4\pi c}\left\langle\frac{1}{I}\right\rangle \approx B_{\text{e}} - \frac{2B_{\text{e}}}{r_{\text{e}}}\langle\xi\rangle , \tag{5}$$

where c is the speed of light, and $B_{\text{e}} = \hbar(4\pi c I_{\text{e}})^{-1} = \hbar(4\pi c\mu r_{\text{e}}^2)^{-1}$ is the rotational constant corresponding to the structure at the equilibrium with distance r_{e}. The average displacement in the harmonic approximation is zero. Therefore, the vibrationally-averaged rotational constant at the harmonic approximation is equal to the equilibrium rotational constant. However, the asymmetric terms of the PES Taylor expansion may introduce a vibrationally-averaged displacement. For example, due to the term $V_3\xi^3/6$ we have[13]

$$\langle\xi\rangle = -\frac{V_3}{2V_2}l^2 .$$

For bond stretch vibrations typically $V_3 < 0$ and $V_2 > 0$. Therefore, equilibrium rotational constants are usually larger than respective vibrationally-averaged rotational constants.

The full-dimensional quantum-mechanical treatment is the most accurate way to describe anharmonic motions of nuclei. The full solution of the vibrational Schrödinger equation is a computationally and theoretically challenging task.[14, 15] First, it requires extensive quantum-chemical calculations to produce the numerical potential energy surface (PES). Often these data are approximated with a suitable analytical form. Finally, it requires advanced techniques for solving multidimensional Schrödinger equations. The associated with these procedures difficulties limit the widespread usage of such methods.

The simplest quantum approach that offers the anharmonic treatment of the molecular vibrations is the second-order perturbation theory, usually denoted as VPT2.[16, 17] In this approach, harmonic, cubic ($\propto V_3\xi^3$) and semi-diagonal quartic force fields ($\propto V_4\xi^4$) are taken into consideration. Since the harmonic potential is symmetric and the cubic part is antisymmetric with respect to the displacements from the equilibrium values, the perturbative first order (PT1) energy correction with the cubic force field is zero. However, the quartic semi-diagonal part of the field with the terms of the form $V_{4,ij}\xi_i^2\xi_j^2$, where i and j enumerate displacements along different modes, gives nonzero corrections at the PT1 level. The second order corrections (PT2) are thus being performed only for the cubic part of the potential. Such procedure can be automated and is already implemented in multiple popular quantum-chemistry software, such as Gaussian,[18] ORCA,[19] GAMESS US,[20] and Cfour.[21] However, with the increasing number of atoms N in a molecule the computational cost of VPT2 increases dramatically due to the growing amount of cubic force field elements, which scales as N^3. For large molecules MD simulations may become computationally more effective.[15] Therefore, it is important to develop procedures for computation of molecular vibrational properties at low temperatures using classical MD simulations.

Wigner sampling[22, 23] is one of possibilities for producing initial conditions in MD simulations, that emulate the quantum harmonic ground vibrational state.[24–26] In this work, we combine this sampling routine with constant energy MD (NVE MD) simulations to provide anharmonic vibrational properties of molecules. In addition, we also propose another simplified routine, based on the Wigner sampling, that does not require preliminary calculations of molecular Hessian.

In this paper we first introduce and compare possible sampling routines, including our simplified version of the Wigner sampling. Then we perform a benchmarking of vibrationally-averaged rotational constants. We demonstrate the proof-of-principle for the approach using the diatomic system of the molecular hydrogen cation. Then, a benchmark set of five small (no more than six atoms) molecules is used to evaluate a method for calculation of rotational constants from MD trajectories. The resulting procedure is applied to another two moderately-sized molecules (20 and 62 atoms). Then we show the application of the sampling routines for the calculation of vibrational spectra for difluoromethane. In particular, we demonstrate the application of the novel sampling routine in conjunction with the Andersen thermostat.[28] In the end, we show the application of the developed methods to different systems with known experimental data. All molecular systems used throughout this work are provided in Figure 1.

[1] Since $\frac{1}{I} \propto \frac{1}{r^2} = \frac{1}{(r_{\text{e}}+\xi)^2} \approx \frac{1}{r_{\text{e}}^2}\cdot\left(1 - \frac{2\xi}{r_{\text{e}}}\right)$.

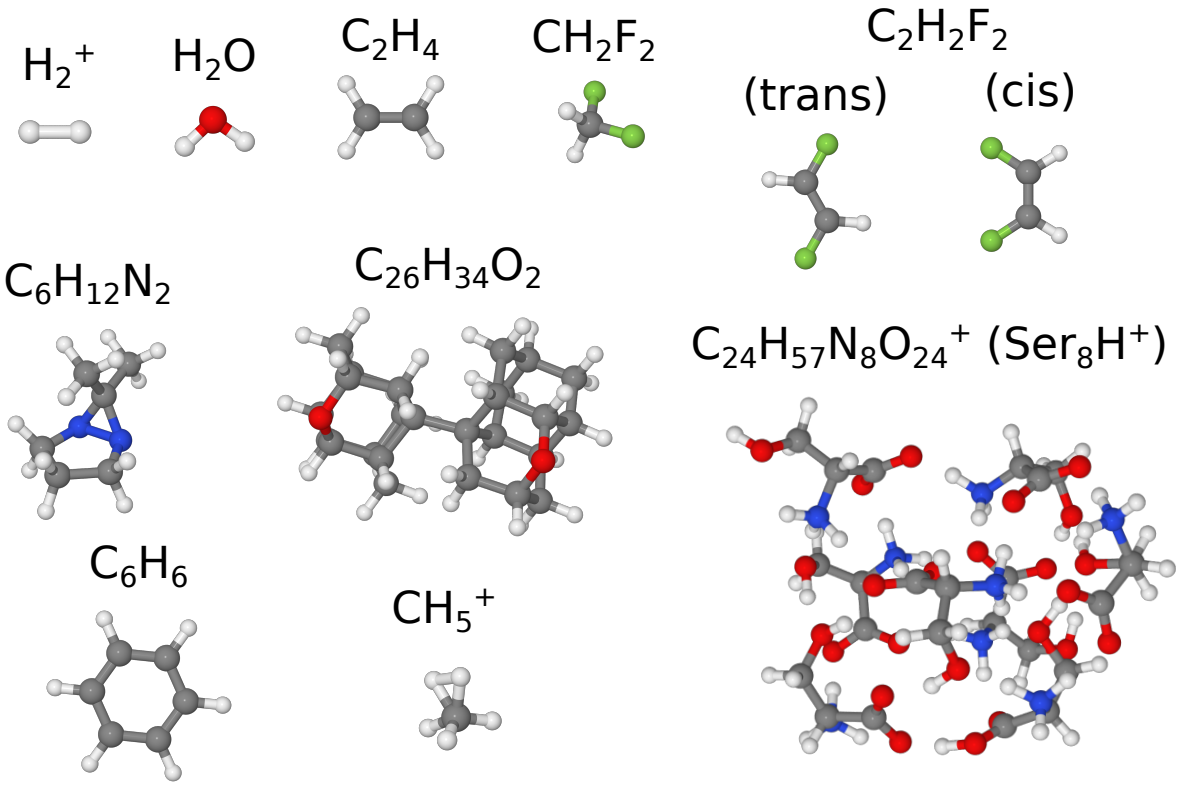

Figure 1: Testing molecules. Standard Jmol[27] coloring scheme is used for the atoms: white (H), gray (C), blue (N), red (O) and green (F).

2 Initial conditions

2.1 Maxwell-Boltzmann sampling

The classical procedure for the generation of initial conditions in MD simulations is the Maxwell-Boltzmann sampling (MBS).[29] Starting nuclear coordinates are defined externally, for example, they may correspond to the equilibrium molecular structure. Initial velocities of nuclei are generated from the Maxwell-Boltzmann distribution, e.g. for the temperature T the velocity v_α along a certain axis ($\alpha = x, y, z$) of a nucleus with the mass m is distributed as

$$v_\alpha \sim \exp\left(-\frac{m v_\alpha^2}{2k_\mathrm{B}T}\right) ,$$

where k_B is the Boltzmann constant. This generation of the initial conditions is very effective for classical ensembles. However, we can also use MBS for crude simulation of NQEs. In simulations at elevated temperatures, pair distribution functions (PDFs) are similar to those in the classical case. The squared vibrational amplitude (Equation 2) for the interatomic distance r in the classical harmonic approximation is:[13]

$$l^2_\mathrm{class} = \frac{k_\mathrm{B}T}{\mu\omega^2} ,$$

whereas in the quantum case it is[13]

$$l^2_\mathrm{quant} = \underbrace{\frac{\hbar}{2\mu\omega}}_{l_0^2} \coth\left(\frac{\hbar\omega}{2k_\mathrm{B}T}\right) , \tag{6}$$

where l_0 is the vibrational amplitude for the ground vibrational state (see Equation 3), which can also be defined as $l_0 = l_\mathrm{quant}(T = 0)$. For each given temperature, the classical vibrational amplitude does not exceed the quantum one ($l_\mathrm{class} \leq l_\mathrm{quant}$). However, it is possible to find the specific temperature at which the classical ensemble shows the same PDF as in the quantum case at $T = 0$. From the equation $l_\mathrm{class} = l_0$, we obtain

$$T = \frac{\hbar\omega}{2k_\mathrm{B}} = \frac{T_\mathrm{D}}{2} ,$$

i.e., at the half of the Debye temperature $T_\mathrm{D} = h\nu/k_\mathrm{B}$ the classical PDF of the harmonic oscillator is equivalent to the PDF of the ground vibrational quantum state.

2.2 Wigner sampling

One of the currently popular sampling techniques is based on Wigner quasiprobability distributions obtained at the harmonic approximation.[23, 24, 26, 29] Let us consider the Hamiltonian $\hat{H} = \frac{1}{2\mu}\hat{p}^2 + \frac{\mu\omega^2}{2}\xi^2$ for a single vibraional mode with ξ being the displacement of the coordinate from the equilibrium position and $\hat{p} = -i\hbar\partial_\xi$. The ground vibrational state is then given by the wavefunction

$$\psi_0(\xi) = \frac{1}{\sqrt{\sqrt{2\pi}\cdot l_0}} \cdot \exp\left(-\frac{\xi^2}{4l_0^2}\right) ,$$

where l_0 is defined in Equation 6. The corresponding Wigner quasi-probability function for the momentum p can be obtained from any given wavefunction ψ via transformation [22]

$$W(p,\xi) = \frac{1}{\pi\hbar}\int_{-\infty}^{+\infty} \psi^*(\xi - s)\exp\left(\frac{2i\cdot p\cdot s}{\hbar}\right)\psi(\xi + s)ds .$$

For the ground state of the harmonic oscillator $\psi = \psi_0$ we get

$$W_0(p,\xi) = \frac{1}{\pi\hbar}\exp\left(-\frac{\xi^2}{2l_0^2} - \frac{2l_0^2p^2}{\hbar}\right) = \frac{1}{\pi\hbar}\cdot\exp\left(-\frac{\xi^2}{2\sigma_\xi^2}\right)\cdot\exp\left(-\frac{p^2}{2\sigma_p^2}\right) . \tag{7}$$

Thus, the resulting Wigner function is a product of two Gaussian distributions: for the coordinate ξ and the momentum p with widths $\sigma_\xi = l_0$ and $\sigma_p = \hbar/(2l_0)$, respectively. The product $\sigma_\xi \cdot \sigma_p = \hbar/2$ fulfills the Heisenberg inequality $\sigma_\xi \cdot \sigma_p \geq \hbar/2$.

In the case of a polyatomic molecule, there is a mapping between independent one-dimensional vibrational modes and the Cartesian coordinates and their corresponding momentums of atoms. The linear displacement of the atoms in the Cartesian coordinates $\mathbf{r}$ from their equilibrium positions ($\mathbf{r}_e$) can be related to the vector of displacements of vibrational modes $\boldsymbol{\xi}$ through the matrix of the mode shapes $\mathcal{L}$: $\mathbf{R} = (\mathbf{r} - \mathbf{r}_e) = \mathcal{L}\boldsymbol{\xi}$. A similar relation connects the Cartesian momentums of the atoms $\mathbf{P} = -i\hbar\partial_{\mathbf{R}}$ to the momentums of the independent modes $\mathbf{p} = -i\hbar\partial_{\boldsymbol{\xi}}$: $\mathbf{P} = \mathcal{L}^{-1}\mathbf{p}$. The Wigner function for $\boldsymbol{\xi}$ and $\mathbf{p}$ is just a product of the one-dimensional quasi-probability distributions (Equation 7), and it can be written as

$$W_0(\mathbf{p},\boldsymbol{\xi}) \propto \exp\left(-\frac{1}{2}\boldsymbol{\xi}^\dagger\Sigma_{\boldsymbol{\xi}}^{-1}\boldsymbol{\xi} - \frac{2}{\hbar^2}\mathbf{p}^\dagger\Sigma_{\boldsymbol{\xi}}\mathbf{p}\right) ,$$

where matrix $\Sigma_{\boldsymbol{\xi}} = \mathrm{diag}(\sigma^2_{\xi,1}, \sigma^2_{\xi,2}, \ldots)$ is the variance matrix, and the momentum variance is obtained from the Heisenberg equality relation. This distribution can be rewritten in terms of displacements ($\mathbf{R} = \mathcal{L}\boldsymbol{\xi}$) and momentums $\mathbf{P} = \mathcal{L}^{-1}\mathbf{p}$ in Cartiesian coordinates:

$$W_0(\mathbf{P},\mathbf{R}) \propto \exp\left(-\frac{1}{2}\mathbf{R}^\dagger\overbrace{\mathcal{L}^\dagger\Sigma_{\boldsymbol{\xi}}^{-1}\mathcal{L}}^{\tilde{\Sigma}_{\mathbf{R}}^{-1}}\mathbf{R} - \frac{2}{\hbar^2}\mathbf{P}^\dagger\overbrace{(\mathcal{L}^{-1})^\dagger\Sigma_{\boldsymbol{\xi}}\mathcal{L}^{-1}}^{\tilde{\Sigma}_{\mathbf{R}}}\mathbf{P}\right) . \tag{8}$$

Here, the effective variance matrix $\tilde{\Sigma}_{\mathbf{R}}$ is non-diagonal, unlike the initial $\Sigma_{\boldsymbol{\xi}}$, which indicates the correlation between atomic displacements due to the collective nature of the vibrational modes. The replacement of the variables $\{\mathbf{p},\boldsymbol{\xi}\} \rightarrow \{\mathbf{P},\mathbf{R}\}$ is unitary, since $d\mathbf{p}d\boldsymbol{\xi} = \mathcal{L}^{-1}d\mathbf{R}\mathcal{L}d\mathbf{P} = d\mathbf{R}d\mathbf{P}$. Using this distribution, it is possible to simultaneously sample both the linear displacements of the nuclei from the equilibrium geometry and nuclear velocities in polyatomic molecules.[23] This and similar sampling procedures for initial conditions are called Wigner sampling (WS). If we replace l_0 for each vibrational mode with the temperature-dependent l_{quant} from Equation 6, we can also sample temperature-averaged vibrational states.

2.3 Simplified Wigner sampling

WS requires initial optimization of molecular structure and calculation of its harmonic potential. However, we can try to construct a simplified sampling procedure similar to WS.

Let us consider the motion of a nucleus with the mass m along the coordinate x. The momentum p and the coordinate x should fulfill the uncertainty principle, which can be ensured by using the Wigner function similar to that in Equation 7:

$$W(p,\xi) = \frac{1}{\pi\hbar}\cdot\underbrace{\exp\left(-\frac{\xi^2}{2\sigma_x^2}\right)}_{\propto\rho_x}\cdot\underbrace{\exp\left(-\frac{p^2}{2\sigma_p^2}\right)}_{\propto\rho_p} , \tag{9}$$

where ξ is the displacement of the coordinate x, $p = mv$ is the corresponding momentum of the nucleus along the coordinate. By choosing a fixed σ_x we can parameterize the sampling procedure that fulfills the uncertainty principle along each given degree of freedom in the molecule.

The construction of the Wigner distribution requires solving the quantum-mechanical problem. However, we can try to apply heuristic considerations to make a self-sustained procedure for generating such distributions. As a starting point, we need to consider that the uncertainties of the coordinate and momentum should at least fulfill the Heisenberg equality principle ($\sigma_x \cdot \sigma_p = \hbar/2$). This principle makes both distributions for the coordinate ρ_x and the momentum ρ_p dependent on each other: the more uncertainty is for the momentum (σ_p is large), the less it is for the coordinate (σ_x is small), and *vice versa*. We can think of this uncertainty as the amount of information we have about x and p. As a guiding principle, we require the same information about x and p, i.e. the uncertainty in the distributions ρ_x and ρ_p should be approximately the same.

A simple way to define a metric between distributions ρ_a and ρ_b is the Kullback–Leibler (KL) divergence[30] defined as

$$D_{\mathrm{KL}}(\rho_a||\rho_b) = \int_{-\infty}^{+\infty} \rho_a(q) \ln\left(\frac{\rho_a(q)}{\rho_b(q)}\right) dq \ .$$

This function is not symmetric with respect to swapping ρ_a and ρ_b. Hence, here we will use the symmetrized KL distance:

$$J_{\mathrm{KL}}(\rho_a, \rho_b) = J_{\mathrm{KL}}(\rho_b, \rho_a) = D_{\mathrm{KL}}(\rho_a||\rho_b) + D_{\mathrm{KL}}(\rho_b||\rho_a) \ .$$

In the case of the Gaussian distribution $\rho_n(q) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp(-(q-\mu_n)^2/(2\sigma_n))$ we get[31]

$$J_{\mathrm{KL}}(\rho_a, \rho_b) = \frac{1}{2}\left((\mu_a - \mu_b)^2 \cdot \left(\frac{1}{\sigma_a^2} + \frac{1}{\sigma_b^2}\right) + \frac{\sigma_a^2}{\sigma_b^2} + \frac{\sigma_b^2}{\sigma_a^2} - 2\right) \ .$$

For the case of ρ_x and ρ_p the expectation value for the displacement and the momentum are $\mu_x = \mu_p = 0$. However, we cannot compare position and momentum distributions since they have different units. To correct for that, instead of the $\rho_x(\xi)$, we will consider an another distribution of the new parameter

$$X = \frac{m \cdot \xi}{\tau} \ ,$$

which shows the momentum originating from the displacement after some time τ. The distance between distributions ρ_X (with $\sigma_X = \frac{m\sigma_x}{\tau}$) and ρ_p is

$$J_{\mathrm{KL}}(\rho_X, \rho_p) = \frac{1}{2}\left(\frac{m^2\sigma_x^2}{\tau^2\sigma_p^2} + \frac{\tau^2\sigma_p^2}{m^2\sigma_x^2} - 2\right) \ .$$

Due to the relation $\sigma_x = \hbar/(2\sigma_p)$, we can find a condition of $J_{\mathrm{KL}}(\rho_X, \rho_p) \to \min$ as $\frac{dJ_{\mathrm{KL}}}{d\sigma_p} = 0$, which leads to

$$\sigma_p^2 = \frac{\hbar m}{2\tau} \ \Rightarrow \sigma_x^2 = \frac{\hbar\tau}{2m} \ . \tag{10}$$

The resulting parameterization of the displacement (ρ_x) and momentum (ρ_p) distributions depends on the single parameter τ in units of time. Although this approximation is physically less sound than the standard WS, it has the advantage of simplicity due to single control parameter τ and due to the independence on equilibrium molecular geometry.

The sampling procedure is simple. After choosing a reasonable τ, for each coordinate (x, y, z) of all nuclei we generate a displacement and velocity according to the distribution given in Equation 9. We denote this as simplified Wigner sampling (SWS). This procedure can also be rewritten in terms of sampling the multivariate quasi-distribution for all Cartesian coordinates and momentums, similar to the Equation 8. In this case, the sampling is done on the basis of the distribution

$$W_0(\mathbf{P}, \mathbf{R}) \propto \exp\left(-\frac{1}{2}\mathbf{R}^\dagger \Sigma_{\mathbf{R},\mathrm{SWS}}^{-1}\mathbf{R} - \frac{2}{\hbar^2}\mathbf{P}^\dagger \Sigma_{\mathbf{R},\mathrm{SWS}}\mathbf{P}\right) \ ,$$

where $\Sigma_{\mathbf{R},\mathrm{SWS}} = \frac{\hbar\tau}{2}\,\mathrm{diag}(\ldots, \frac{1}{m_k}, \ldots)$ and k enumerates all the Cartesian coordinates.

A reasonable value for τ can be chosen from general physical considerations. The largest displacement will be observed for the lightest nuclei, most commonly for hydrogen with the mass of 1 a.m.u. A displacement of more than 20% of the corresponding bond length will probably introduce too much energy into the system, that can even lead to a spurious fragmentation. Therefore, we require σ_x to be smaller than about 0.2 Å. For $m = 1$ a.m.u. this is achieved with $\tau < 13$ fs. From the other side, the momentum distribution $\sigma_p^2 = \frac{\hbar m}{2\tau}$ can be thought in terms of temperature, since for MBS $\sigma_{p,\mathrm{MB}}^2 = mk_{\mathrm{B}}T$, i.e. the effective temperature is defined as $T_{\mathrm{eff}} = \frac{\hbar}{2k_{\mathrm{B}}\tau}$. For velocities of nuclei $T_{\mathrm{eff}} \leq 5000$ K can be achieved at $\tau > 0.8$ fs. This provide a small range of reasonable values for τ, $1 \leq \tau \leq 10$ fs.

Introducing temperature effects in the distribution 9 is rather straightforward. This can be done in a classical sense as the addition of the momentum δp, which is distributed according to the Maxwell-Boltzmann statistics ($\delta p \sim \exp(-\delta p^2/(2mk_{\mathrm{B}}T))$). By replacing p with $p + \delta p$ we need to average $W(x, p)$ over δp, which leads to

$$W_T(\xi, p) = \frac{1}{\sqrt{2\pi m k_{\mathrm{B}}T}} \int_{-\infty}^{+\infty} \exp\left(-\frac{\delta p^2}{2mk_{\mathrm{B}}T}\right) \cdot W(p + \delta p, \xi) d\delta p = \frac{1}{2\pi\sigma_x\sigma_{p,T}} \underbrace{\exp\left(-\frac{\xi^2}{2\sigma_x^2}\right)}_{\propto\rho_x} \cdot \underbrace{\exp\left(-\frac{p^2}{2\sigma_{p,T}^2}\right)}_{\propto\rho_{p,T}} \ ,$$

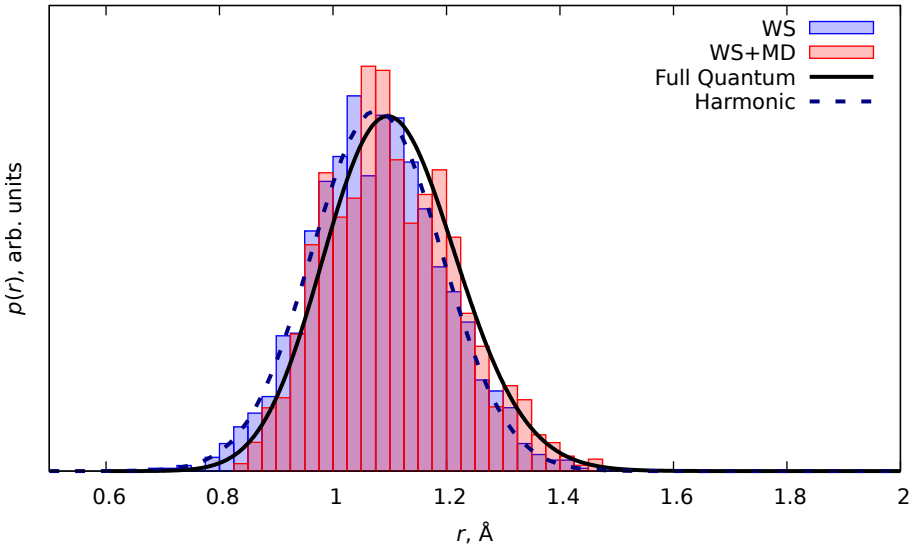

Figure 2: PDF of the molecular cation $H_2{}^+$. The solid line shows the anharmonic ground vibrational state density $|\psi_0|^2$, the dashed line is the density distribution from the harmonic approximation. The blue bar plot shows the PDF obtained from the WS routine. The red bar plot shows the PDF from MD simulations with initial conditions from the WS routine (WS+MD).

where $\sigma^2_{p,T} = \sigma^2_p + mk_\mathrm{B}T$. In the case of $\sigma_x = 0$ and $\sigma_p = 0$, this sampling procedure is equivalent to the standard MBS.

3 Results and discussion

3.1 A simple case of H_2^+

To illustrate the general idea of WS and SWS in application to molecular vibrations, we have done calculations for a simple diatomic model, the molecular cation H_2^+. Here we focus on the interatomic pair distribution function, as this provides a general representation of the molecular ensemble, from which other observables can be computed. For example, the rotational constant of the diatomic molecule can be explicitly expressed via the average interatomic distance (Equation 5).

To simplify simulations, we calculated potential energy at the HF/aug-cc-pV6Z level using ORCA software[19] and then approximated it with Morse potential $E(r) = E_\mathrm{e} + D_\mathrm{e} \cdot (1 - \exp(-\beta \cdot (r - r_\mathrm{e})))^2$, where $E_\mathrm{e} = -0.606674$ Ha is the equilibrium electronic energy, $D_\mathrm{e} = 0.103666$ Ha is the equilibrium dissociation energy, $r_\mathrm{e} = 1.07691$ Å is the equilibrium distance, and $\beta = 1.46583$ Å^{-1} is the parameter, that controls the steepness of the potential. The one-dimensional stationary Schrödinger equation for the vibrational motion was solved with help of the MOLINC package,[32] employing sinc-DVR kinetic energy representation.[33] The comparison of the numerical anharmonic vibrational wavefunction with that from the harmonic approximation is shown in Figure 2. The asymmetry of the anharmonic potential with respect to the equilibrium geometry r_e leads to the shift of the probability distribution into the range of larger interatomic distances r.

The Wigner distribution at the harmonic approximation gives symmetric probability distribution, as it is obtained from the model symmetric potential. Therefore no anharmonic shifts of the interatomic distance can be observed from the ensemble of structures sampled in this case. Although MD can be used with anharmonic PES, being a classical simulation technique, it cannot recover quantum effects (NQEs). Nevertheless, some successful examples of modified MD describing quantum properties exist in the literature.[34, 35] It is possible to replace the classical vibrational amplitude $l = \sqrt{\langle r^2 \rangle - \langle r \rangle^2}$ with its quantum analogue in harmonic approximation.[13, 15, 36] This substitution is justified by the domination of the harmonic potential in the second moment of the interatomic distribution.

The same logic can be transferred to WS. By default, the initial ensemble of the structures produced from the harmonic wavefunction via WS (Equation 7) has no anharmonicity. However, the system evolves at the anharmonic PES for some time, thus allowing the classical dynamics capturing the anharmonicity to some extent. In other words, the WS procedure ensures the quantum behavior, while the anharmonicity of the PES is considered through MD. To test this, we ran 150 trajectories with three sampling routines, MBS at various initial temperatures, WS, and SWS with different τ parameters ranging from 0.1 to 100 fs. The total time of each simulation was 0.5 ps with a time step of 0.5 fs.

The simplest characterization of the one-dimensional interatomic distribution is through its first two moments, the averaged distance $\langle r \rangle$ and the amplitude l (Equation 2). The corresponding values calculated for H_2^+ are given in Figure 3. The harmonic oscillator model and the pure WS result in $\langle r \rangle_\mathrm{h} = r_\mathrm{e}$ because the harmonic potential is

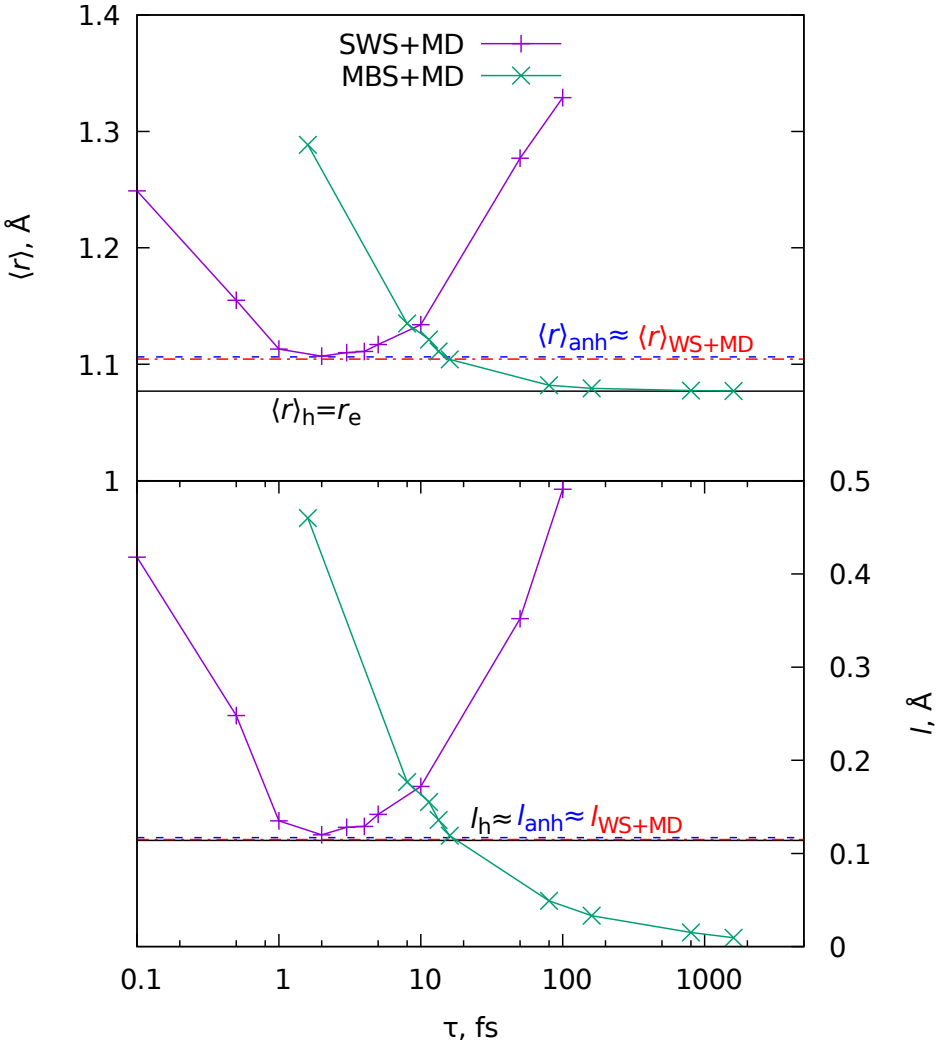

Figure 3: Mean internuclear distances $\langle r \rangle$ and vibrational amplitudes l (Equation 2) for the ground vibrational state of the molecular cation H_2^+. The horizontal lines show the results of the harmonic (index "h", black line), anharmonic ("anh", blue line), and MD with WS ("WS+MD", red line). The curves show the results of the MBS+MD and SWS+MD simulations at different $\tau = h/(k_B T)$.

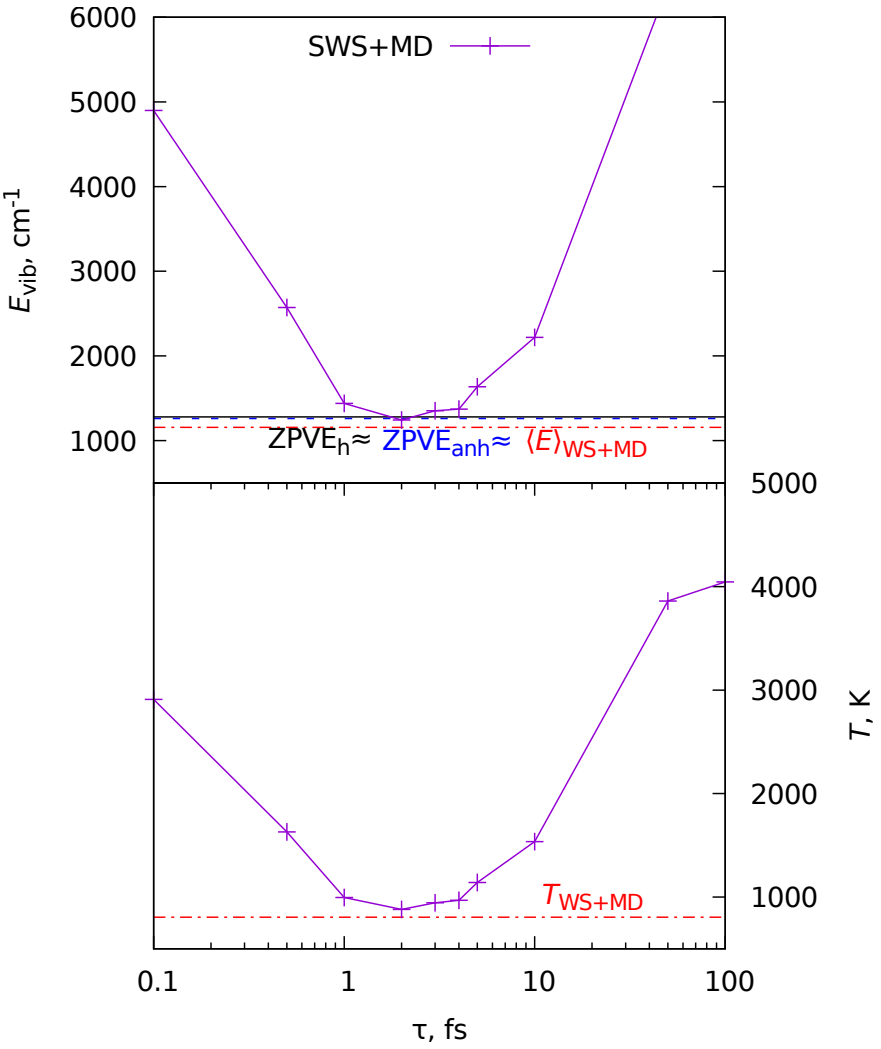

Figure 4: Vibrational energies $E_{\rm vib}$ (above) and vibrational nuclear temperatures $T = 2\langle E_{\rm kin}\rangle/(N_{\rm f}k_{\rm B})$ (below). Horizontal lines show the energy levels calculated in the harmonic (index "h", black line), anharmonic ("anh", blue line), and WS+MD (red line) approximations. The curves show the results of the SWS+MD simulations at different τ.

symmetric. However, MD simulations with initial conditions from WS (WS+MD) produce an average value $\langle r\rangle_{\rm WS+MD}$ very close to that computed with the numerical ground state anharmonic wavefunction $\langle r\rangle_{\rm anh}$. In the case of the diatomic molecule, the SWS distribution can be exactly mapped to WS. Therefore we can get the same solution $\langle r\rangle_{\rm SWS+MD} = \langle r\rangle_{\rm WS+MD}$ by scanning the parameter τ. The resulting curve has a single minimum at $\tau = 2$ fs where $\langle r\rangle_{\rm SWS+MD} = \langle r\rangle_{\rm WS+MD} = \langle r\rangle_{\rm anh}$. Figure 3 also shows the results from MBS+MD simulations at various temperatures, where the temperature T is mapped to τ through the Debye temperature relation $\tau = \nu^{-1} = h/(k_{\rm B}T)$. The resulting value coincides with the expected one at T around 3000 K ($\tau = 16$ fs). The trend for MBS shows only the growth of $\langle r\rangle$ with the increase in temperature. Thus MBS requires an *a priori* knowledge of the appropriate temperature, in contrast to SWS+MD, where the best τ can be found in the variational procedure by minimizing the $\langle r\rangle$. The results for the amplitude l (Figure 3) show that the harmonic value is close to those from anharmonic and WS+MD calculations. This is expected since the second moment of the internuclear distribution is dominated by the harmonic potential.[13, 36, 37]

The single minimum of $\langle r\rangle$ and l in the SWS+MD simulations suggest a similar behavior for other observables, which can be used for finding optimal parameter τ for larger molecules. The obvious choice is an energy-related parameter, for example, the average total energy (electronic and nuclear) $\langle E_{\rm tot}\rangle$. The other option is the kinetic energy of nuclei ($\langle E_{\rm kin}\rangle$), which is often expressed in terms of temperature through the relation $\langle E_{\rm kin}\rangle = N_{\rm f}k_{\rm B}T/2$, where $N_{\rm f}$ is the number of degrees of freedom. For example, $N_{\rm f} = 1$ in the case of the diatomic molecule $H_2{}^+$. The results in Figure 4 show that the minimum energy is reached at $\tau = 2$ fs. Note, at this value the average distance $\langle r\rangle$ and the amplitude l are minimal, as Figure 3 shows. The total vibrational energy $E_{\rm vib} = \langle E_{\rm tot}\rangle - E_{\rm e}$ was compared to the quantum zero-point vibrational energy (ZPVE) of the system. Similar to the amplitude, ZPVE is dominated by the harmonic potential. Therefore, the harmonic and anharmonic ZPVE values are very close. There is also another consideration regarding the τ parameter. According to the equipartition theorem,[38] which also assumes the dominating role of the harmonic potential, the nuclear temperature can be used to guide the choice of the τ parameter in SWS. This is also confirmed by the numerical test demonstrated in Figure 4.

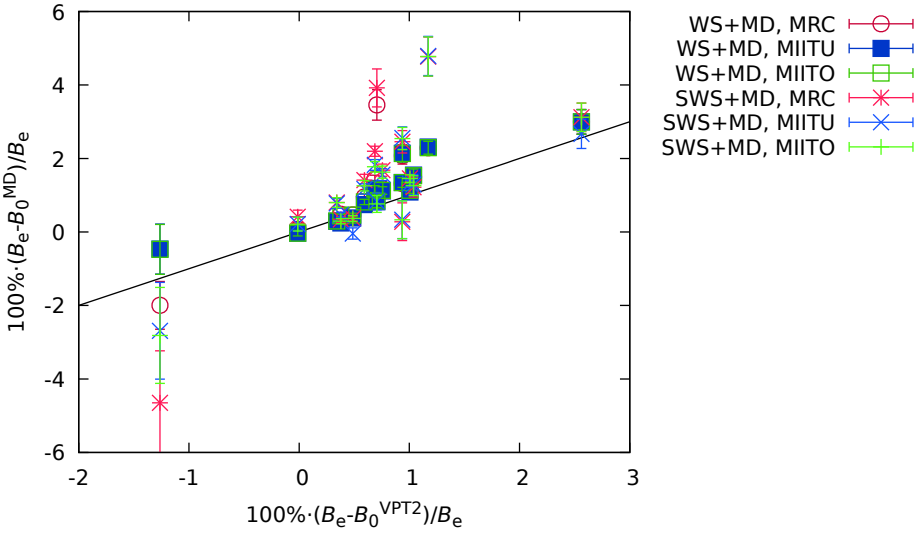

Figure 5: Corrections to rotational constants $B_0 - B_e$ calculated using different methods in comparison to reference values computed from cubic force fields.

3.2 Rotational constants

To test the applicability of WS+MD for the calculation of anharmonic corrections to rotational constants we have chosen five small molecules (see Figure 1): water (H_2O), ethylene (C_2H_4), difluoromethane (CH_2F_2), trans and cis isomers of 1,2-difluoroethylene ($C_2H_2F_2$). For each of these molecules, we have computed vibrational corrections to rotational constants using numeric cubic force fields as implemented in the VPT2[39] procedure in the Gaussian 16 [18] program package. For small semi-rigid molecules, this level of approximation combined with a sufficient treatment of electronic problem can provide accurate reference values.[40, 41] As an alternative method, MD simulations for these molecules were done for 0.5 ps. The slowest harmonic vibration among the test molecules was in cis – $C_2H_2F_2$ with the frequency of 219 cm^{-1}, which corresponds to the vibrational period of 0.15 ps. Thus during 0.5 ps this motion was likely sampled three times. The time step was 0.5 fs. For each of the molecules, 50 MD trajectories with WS and SWS initial conditions were started. The gradients for the molecules were obtained using ORCA 5.[19] The used quantum-chemical approximation for the electronic problem in all cases was PBE/def2-SVP.[42, 43] The WS procedure was performed using scripts from the SHARC-MD package.[44–46] Rotational and translational motions in the molecules at the beginning of the simulations were frozen. In SWS+MD simulations, the τ parameter was not scanned but fixed at the optimal value found for the H_2^+ ion ($\tau = 2$ fs). MD simulations were run with the BOMoND script from the PyRAMD repository.[47, 48] Obtained MD trajectories were processed with scripts from the MOLINC repository.[32]

First, a proper way of computing the averaged rotational constants from the MD trajectories should be chosen. For this, there are at least three ways.

- The simplest approach is the calculation of rotational constants B_α ($\alpha = a, b, c$) at every time step with the following averaging of the B_α over the complete MD trajectory. This procedure is denoted as mean rotational constants (MRC).

- A more sophisticated procedure, in which the complete tensor of inertia $\mathcal{I}$ is calculated at each MD step. The inverted tensor $\mathcal{I}^{-1}$ is averaged along the trajectory, producing $\langle \mathcal{I}^{-1} \rangle$. The eigenvalues of $\langle \mathcal{I}^{-1} \rangle$ are $\mathcal{I}_a^{-1} \geq \mathcal{I}_b^{-1} \geq \mathcal{I}_c^{-1}$, which are related to the rotational constants as $B_\alpha = \frac{\hbar}{4\pi c}\mathcal{I}_\alpha^{-1}$ ($\alpha = a, b, c$). Since the rotation is frozen during the simulation, we may assume, that the principal axes of the molecule do not deviate significantly from the initial positions. Therefore there is no need for a reorientation of the molecule with respect to the reference structure. This procedure is denoted as mean inverse inertia tensor unoriented (MIITU).

- A method similar to the previous but with the orientation of all molecular frames with respect to the reference equilibrium structure. The orientation is done by minimizing the mass-weighted functional $\sum_i m_i(\mathbf{r}_i - \mathbf{r}_{e,i})^2$, where i enumerates all the nuclei in the molecule, m are the masses and $\mathbf{r}$ are the coordinates of the nuclei, the index "e" denotes the reference equilibrium configuration of the nuclei. The minimization can be efficiently performed using the Kabsch algorithm.[49] This way of computing the mean rotational constants is denoted as mean inverse inertia tensor oriented (MIITO).

Each MD trajectory we can consider as a single measurement of the system with the resulting constants $\langle A \rangle = \langle B_a \rangle$, $\langle B \rangle = \langle B_b \rangle$, $\langle C \rangle = \langle B_c \rangle$ obtained by either one of the MRC/MIITU/MIITO procedures. The actual result, however, requires averaging over several of such MD trajectories started from different initial conditions. Therefore, final values

from the MD are results of double averaging from multiple trajectories, defined as

$$B_{\alpha,\mathrm{MD}} = \langle\langle B_a\rangle\rangle = \frac{1}{N_{\mathrm{trj}}}\sum_{n=1}^{N_{\mathrm{trj}}}\langle B_a\rangle_n \ , \ \alpha = a,b,c$$

where N_{trj} is the total number of MD trajectories, and $\langle B_a\rangle_n$ are the mean rotational constants obtained from n-th trajectory. To quantify the statistical significance and convergence of the simulation, one can use the standard errors, calculated as

$$\mathrm{SE}_\alpha = \sqrt{\frac{\langle\langle B_a\rangle^2\rangle - \langle\langle B_a\rangle\rangle^2}{N_{\mathrm{trj}}}} \ ,$$

where $\langle\langle B_a\rangle^2\rangle = \frac{1}{N_{\mathrm{trj}}}\sum_{n=1}^{N_{\mathrm{trj}}}\langle B_a\rangle_n^2$.

Table 1: Pearson correlation coefficients ρ between the normalized rotational constant shifts $(B_e - B_0)/B_e$, computed from the VPT2 model and the MD simulations with different procedures. See text for further details.

	MRC	MIITU	MIITO
$\rho(\mathrm{WS+MD, VPT2})$	0.77	0.76	0.80
$\rho(\mathrm{SWS+MD, VPT2})$	0.83	0.89	0.90

Using these three procedures, we have calculated the vibrational shifts of rotational constants $B_{\mathrm{e},\alpha} - \langle B_\alpha\rangle$, which were compared with those from the standard VPT2 computations. Since the values of the shifts are very sensitive to the values of rotational constants, in statistical assessments (Table 1, Figure 5) we have used the normalized quantities $(B_{\mathrm{e},\alpha} - \langle B\rangle)/B_{\mathrm{e},\alpha} = 1 - \langle B\rangle/B_{\mathrm{e},\alpha}$. In general, results from all tested combinations of methods for trajectory simulations (WS+MD, SWS+MD) and for calculation of rotational constants (MRC, MIITU, MIITO) correlated reasonably with reference VPT2 values. The MIITU and MIITO routines are the most physically justified, since the vibrationally-averaged rotational parameter appears from the averaging of the rotational Hamiltonian as $\hat{H}_{\mathrm{rot}} = \langle 0|\frac{1}{2}\hat{\mathbf{L}}^\dagger\mathcal{I}^{-1}\hat{\mathbf{L}}|0\rangle = \frac{1}{2}\hat{\mathbf{L}}^\dagger\langle 0|\mathcal{I}^{-1}|0\rangle\hat{\mathbf{L}}$, where $|0\rangle$ is the ground vibrational state, and $\hat{\mathbf{L}}$ are the angular momentum operator. To quantify the performance of the methods, we have computed Pearson correlation coefficients ρ for the observed normalized rotational constant shifts from the MD trajectories and VPT2. The results are summarized in Table 1. MIITO showed the best performance among all of the routines applied, and therefore it was further used as a default procedure for calculation of vibrationally-averaged rotational constants from MD trajectories. The resulting vibrational corrections to rotational constants are listed in Table 2. As one can see, the MD values correlate quite well with the reference values from the VPT2 model.

A very important aspect is that MD simulations with frozen out rotations cannot recover centrifugal effects. Centrifugal expansion is a few orders of magnitude smaller than the anharmonic shift (see Table 2) and negligible in comparison to standard errors of our MD values. Thus, even with allowed rotations the uncertainty of MD prevents its usage for calculation of this particular effect. On the other side, it is mostly dominated by the harmonic potential and therefore is well predicted in the harmonic approximation with a much smaller computational cost.[16]

Table 2: Equilibrium rotational constants B_{e} and their anharmonic shifts in the ground vibrational state $B_{\mathrm{e}} - B_0$ for the five model asymmetric top molecules.[a] The uncertainties of the MD-based values are the ± standard errors.

Molecule	$A/B/C$	B_{e}	$B_{\mathrm{e}} - B_0$			
			VPT2	VPT2+cent.	WS+MD	SWS+MD($\tau = 2$ fs)
H_2O	A	2505.328	−31.632	−32.781	-12 ± 17	-71 ± 33
	B	1456.758	10.273	10.567	12 ± 3	11 ± 3
	C	921.145	23.594	24.057	28 ± 3	29 ± 4
C_2H_4	A	475.622	4.457	4.457	10 ± 1	12 ± 1
	B	98.657	0.678	0.679	1.2 ± 0.1	1.8 ± 0.1
	C	81.708	0.821	0.820	0.9 ± 0.1	1.2 ± 0.1
CH_2F_2	A	164.710	1.710	1.710	2.5 ± 0.3	2.1 ± 0.4
	B	34.706	0.206	0.206	0.28 ± 0.03	0.4 ± 0.1
	C	30.442	0.230	0.230	0.34 ± 0.03	0.5 ± 0.1
trans − $C_2H_2F_2$	A	190.383	2.230	2.230	4.4 ± 0.3	9 ± 1
	B	13.199	0.046	0.046	0.06 ± 0.01	0.04 ± 0.01
	C	12.343	0.058	0.058	0.06 ± 0.01	0.05 ± 0.01
cis − $C_2H_2F_2$	A	71.606	0.670	0.670	1.0 ± 0.1	0.2 ± 0.4
	B	18.895	−0.002	−0.003	-0.01 ± 0.01	0.04 ± 0.04
	C	14.950	0.051	0.051	0.04 ± 0.01	0.12 ± 0.02

[a] The quantum-chemical approximation used for simulation was PBE/def2-SVP. All values given are in m^{-1} (for conversion to cm^{-1}, the values must be divided by 100). All the MD values were obtained from the MIITO routine.

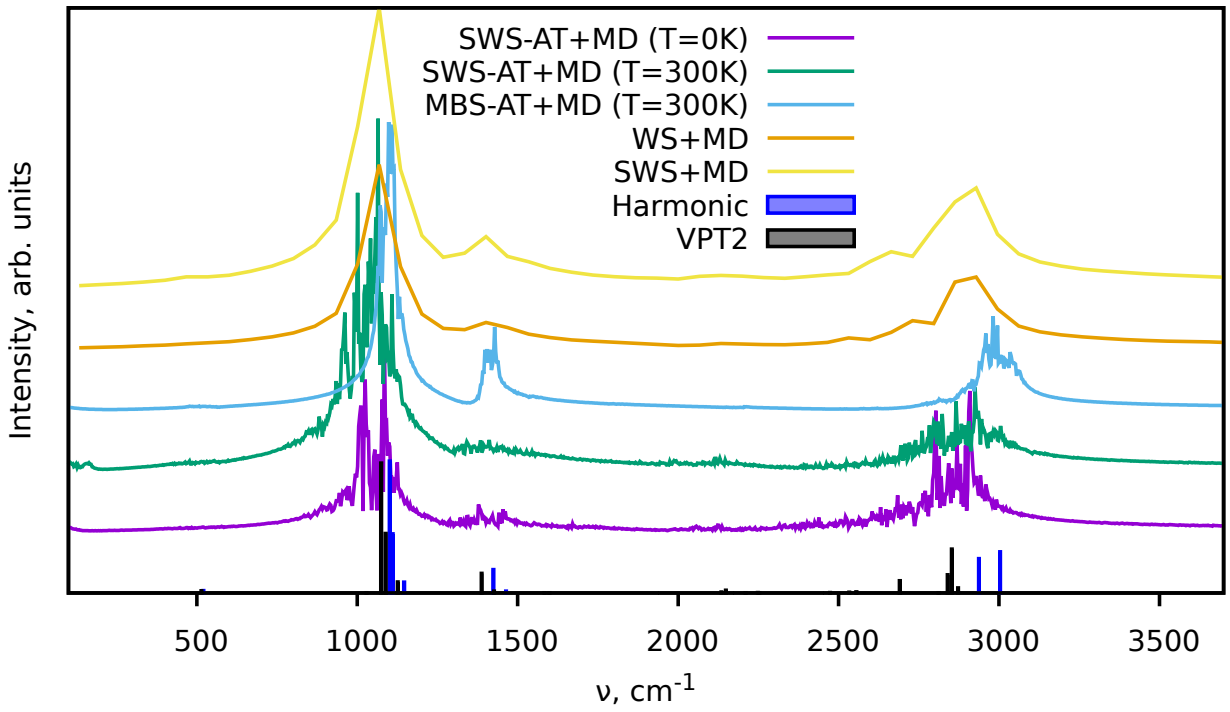

Figure 6: Vibrational infrared (IR) spectra of CH_2F_2. The bars indicate the harmonic and VPT2 anharmonic spectra. The lines show the spectra obtained from the MD simulations.

Two additional tests of the proposed computational methods were done using the moderate-size molecular systems (see Figure 1), 6,6-dimethyl-1,5-diazabicyclo[3.1.0]hexane ($C_6H_{12}N_2$) and *syn*-conformer of 6,6'-bis(3-oxadiamantane) ($C_{26}H_{34}O_2$). Both structures were investigated experimentally [50, 51] using gas electron diffraction (GED). For the latter molecule GED data were supplemented by rotational constants from microwave measurements. Due to the larger role of dispersion interactions in the structures of these systems, the applied in our work quantum-chemical approximation included D3BJ empirical corrections.[52] Optimizations of the molecular structures and calculations of VPT2 corrections were done using Gaussian 16[18]. The lowest and the highest harmonic frequencies among vibrations in both molecules were 78 and 3110 cm^{-1}, which correspond to the vibrational periods of 430 and 11 fs, respectively. To provide an approximate coverage of at least three vibrational periods for each of the modes, the total duration of MD trajectories was set to 1.5 ps. The time step for the simulation was chosen to be 1 fs, to be at least ten times smaller than the period of the fastest vibration. For $C_6H_{12}N_2$ and $C_{26}H_{34}O_2$ 35 and 46 WS+MD trajectories were obtained, respectively. The comparison of the resulting anharmonic corrections to the rotational constants ($B_e - B_0$) is given in Table 3. The relative deviations of the WS+MD corrections from the VPT2 values are about the same order of magnitude as for the smaller systems. Thus, WS+MD can probably be applied to larger and more complicated molecules.

Table 3: Equilibrium rotational constants B_e (in m^{-1}) and their anharmonic shifts in the ground vibrational state $B_e - B_0$ for $C_6H_{12}N_2$ and $C_{26}H_{34}O_2$ molecules. The uncertainties of the MD-based values are the ± standard errors.

	$A/B/C$	B_e	$B_e - B_0$		
			VPT2	VPT2+cent.	WS+MD[a]
$C_6H_{12}N_2$	A	12.02	0.12	0.12	0.16 ± 0.01
	B	5.85	0.06	0.06	0.10 ± 0.01
	C	5.79	0.07	0.07	0.11 ± 0.01
$C_{26}H_{34}O_2$	A	1.3151	0.0126	0.0126	0.0164 ± 0.0002
	B	0.5286	0.0060	0.0060	0.0078 ± 0.0001
	C	0.5281	0.0061	0.0061	0.0078 ± 0.0001

[a]The values were obtained using the MIITO routine.

3.3 Vibrational spectra

Another possible application of MD is the prediction of vibrational spectra. Here we discuss only infrared (IR) one-photon spectra. The intensity of IR spectra can be computed as[53]

$$I(\nu) = \left| \int \langle \dot{\mathbf{d}}(\tau)\dot{\mathbf{d}}(\tau + t)\rangle_\tau \exp(-i2\pi\nu t)dt \right| ,$$

where $\langle \dot{\mathbf{d}}(\tau)\dot{\mathbf{d}}(\tau + t)\rangle_\tau$ is the autocorrelation function (ACF) of the dipole moment time-derivative ($\dot{\mathbf{d}}$), t the is time, and ν is the frequency. Other types of vibrational spectra, such as Raman or vibrational circular dichroism, can be computed similarly.[53, 54] The duration (t_{tot}) and the time step (Δt) of ACF are the same as for the original MD trajectory. In discrete Fourier transform, the frequency resolution is determined by the trajectory duration as $\Delta\nu \propto t_{\mathrm{tot}}^{-1}$, while the time step determines the maximal frequency as $\nu_{\mathrm{max}} \propto \Delta t^{-1}$. Generally, it is recommended to use small time steps in MD simulations, otherwise large integration steps can lead to artificial blue shifts in the high-frequency region.[55] However, such shifts can be corrected using a simple replacement of Fourier-transform frequencies ν_{FT} with[56]

$$\nu = \frac{\sqrt{2 \cdot (1 - \cos(2\pi \cdot \Delta t \cdot \nu_{\mathrm{FT}}))}}{2\pi \cdot \Delta t} .$$

We also applied this correction throughout the work.

It is worth to mention practical limitations of NVE MD simulations due to the limited total number of collected frames N, determined by available computational and time resources. We calculate N_{t} trajectories each with N_{s} steps, so that the total number of frames is $N = N_{\mathrm{t}} \cdot N_{\mathrm{s}}$. For the predefined constant N we may balance N_{t} and N_{s} but cannot increase both parameters. Smaller N_{s} cause the lowering of the spectral resolution, because of the duration of trajectories $t = N_{\mathrm{s}} \cdot \Delta t$, where Δt is the integration time step. On the other side, smaller N_{t} will result in poorer phase-space sampling. To overcome this problem of NVE MD, we can try switching to NVT MD with a thermostat, in which the phase space sampling is done on-the-fly. In this case, with a similar number of computational operations it is possible to obtain fewer trajectories (smaller N_{t}) but with larger lengths (larger N_{s}) and optimal phase space sampling. However, the thermostat must not interfere too much with the dynamics of the system.[15]

Andersen thermostat is the most straightforward to implement.[28] The WS method relies on equilibrium molecular geometry and cannot be used for anything except for sampling of initial conditions. SWS, on the other hand, has the same principle as MBS and thus can be used in combination with the Andersen thermostat. We have performed two simulations for difluoromethane with SWS and Andersen thermostat (abbreviated as SWS-AT+MD) with the collision period of 300 fs, time step of 1 fs and total trajectory duration of 10 ps at two different temperatures, $T = 0$ K and $T = 300$ K. Three and four trajectories were collected in simulations for $T = 0$ K and $T = 300$ K, respectively. For comparison, we also performed classical constant temperature MD simulations with the Andersen thermostat applying the MBS scheme (MBS-AT+MD).

Figure 6 shows the computed vibrational spectra for the CH_2F_2 molecule. Classical MBS introduces relatively small vibrational amplitudes. Therefore, high-frequency vibrations from MBS-AT+MD are essentially harmonic. Taking into account nuclear quantum effects in WS+MD, SWS+MD, and SWS-AT+MD simulations leads to a better agreement of the MD-based spectra with the VPT2 calculations. Short lengths of MD trajectories in the NVE MD simulations limit the frequency resolution in the WS+MD and SWS+MD simulations. The much longer SWS-AT+MD simulations show results very similar to WS+MD and SWS+MD, which is seen in the similar shapes and positions of the peaks in the spectra. However, due to the much longer trajectories duration the spectral resolution of SWS-AT+MD exceeds by far the spectral resolution in WS+MD and SWS+MD. On the basis of the obtained results we can recommend the SWS-AT+MD technique for calculation of vibrational spectra.

3.4 Examples

To further demonstrate the applicability of WS+MD and SWS+MD for the calculation of vibrational properties of molecules and to compare the performance of MD with harmonic and VPT2 vibrational models, we have performed three additional sets of calculations for different molecular systems (see Figure 1). We have chosen benzene (C_6H_6), protonated methane cation (CH_5^+) and protonated serine octamer (Ser_8H^+, $(C_3H_7NO_3)_8H^+$). For all of them there have been published high quality experimental and theoretical data.[57–64] Benzene is an example of a simple and semi-rigid system, for which are available experimental rotational constants with respective theoretical anharmonic corrections [57, 59], vibrational [58] and photoelectron (PE) spectra [60]. CH_5^+ was taken as a system with large amplitude motions (LAMs) requiring treatment beyond the localized vibrational harmonic and VPT2 approximations. Ser_8H^+ is used to demonstrate the increasing computational expenses of the VPT2 approach compared to MD.

3.4.1 Benzene

To provide the comparison with experimental and high-level theoretical data, we have performed various types of MD simulations for benzene at the PBEh-3c level of theory[65] using BOMoND and ORCA 5. The following numbers of trajectories were obtained: 50 for WS+MD, 24 for SWS-AT+MD at $T = 0$ K, 47 for SWS-AT+MD at $T = 300$ K and 35 for MBS-AT+MD at $T = 300$ K. The time step in all cases was 1 fs, the total lengths of each trajectory were

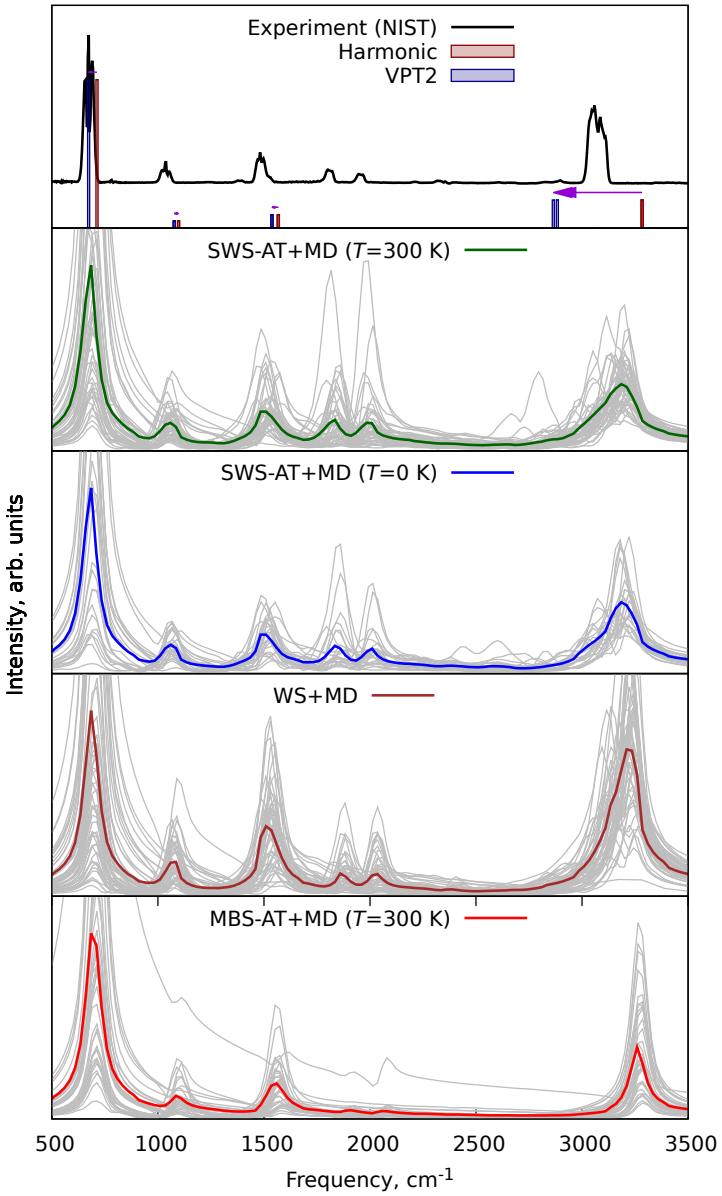

Figure 7: Experimental and theoretical infrared (IR) vibrational spectra of benzene. The horizontal arrow indicate the anharmonic VPT2 shift.

2 ps, collision period for AT was 300 fs. From the WS+MD trajectories we have computed vibrational corrections to rotational constants, $B_e - B_0 = 0.192 \pm 0.01$ m^{-1} and $C_e - C_0 = 0.083 \pm 0.004$ m^{-1}. $^{12}C_6{}^1H_6$ benzene is an oblate symmetric top molecule, thus the $A = B$ rotational constants cannot be obtained directly from experiment. For the C-constant our value of the vibrational correction is in reasonable agreement with the reference theoretical value of 0.14 m^{-1}.[59]

Experimental vibrational IR spectra were taken from NIST Chemistry WebBook [58] and compared with calculated in this work, see Figure 7. Since the PBEh-3c model is not implemented in Gaussian, we also performed harmonic and VPT2 frequency calculations using ORCA 5. As the VPT2 module of ORCA does not compute anharmonic intensities, we used respective values from the harmonic approximation. The IR spectra from all the MD simulations were calculated for each trajectory separately (gray lines in Fig. 7) with subsequent their averaging for each simulation type set (colored lines in Fig. 7). First 0.5 ps in each trajectory were considered as equilibration phase and therefore ignored. The resulting spectra were convoluted using Gaussian function with the full width at half maximum (FWHM) parameter of 50 cm^{-1}. The averaged IR spectra from SWS-AT+MD and WS+MD reproduce all experimental bands. MBS-AT+MD underestimates intensity for the overtones in the 1700-2000 cm^{-1} range (not shown in harmonic and VPT2 plots). It is interesting to observe variations of the C-H stretching peak, corresponding to vibrations with highest Debye temperature in benzene. Its position in MBS-AT+MD (3260 cm^{-1}) is essentially the position of the harmonic vibration (3262 cm^{-1}). This is explained by underestimation of vibrational amplitudes in the classical MD

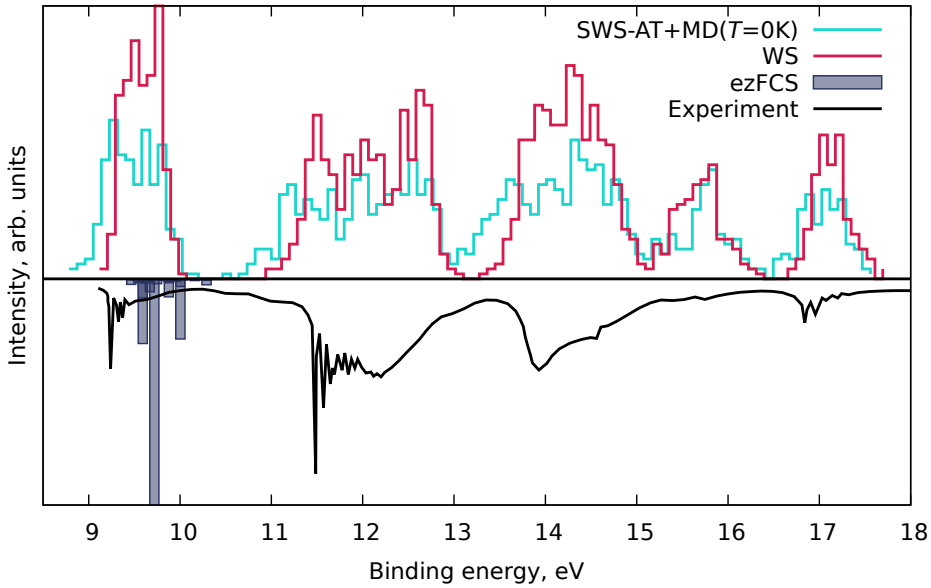

Figure 8: Experimental and theoretical photoelectron spectra (PES) of benzene. The ezFCS results were computed at the PBEh-3c approximation (no shifts applied). The WS and SWS-AT+MD results were obtained at the ACD(2)/cc-pVTZ//PBEh-3c approximation and manually shifted by 0.5 eV for better match with the experimental peaks.

treatment (see discussion of Equation 4). The WS+MD results show larger anharmonic shift with the peak maximum at 3223 cm^{-1}. In SWS-AT+MD at both temperatures (0 and 300 K) this band has the same position at 3185 cm^{-1}. Notably, all MD models underestimate the anharmonic shift and cannot reproduce accurately the experimental band at 2874 cm^{-1}. However, we see acceptable general agreement of the experimental IR spectrum with theoretical from MD methods accounting for NQEs within the WS paradigm.

We also computed photoelectron spectra (PES) of benzene and compared them with the experimental data [60] extracted using the WebPlotDigitizer software.[66] The three types of calculations were performed. First, conventional Franck-Condon factors were computed for the ground electronic states of the neutral benzene and for its cation using the ezFCS software.[67] As a preliminary step, structure optimizations and harmonic frequency calculations were done for C_6H_6 and $C_6H_6{}^+$. However, $C_6H_6{}^+$ is a well known example [68, 69] of the Jahn-Teller effect. Therefore, the same procedures cannot be applied to the excited electronic states of $C_6H_6{}^+$ and only the first band of the PE spectrum could be calculated using this procedure.

Next we tried to account for the vibrational motions by sampling 150 structures of benzene using the WS procedure and harmonic vibrational modes at the PBEh-3c level. For these structures ionization potentials (IPs) were calculated using ADC(2)/cc-pVTZ method. Binning the obtained IPs allowed us to computed an approximate PE spectrum, see Figure 8. This is a simplistic treatment of vibrational degrees of freedom in the harmonic approximation.

For anharmonic treatment of vibrational motions, 5 structures were randomly taken from each of 24 SWS-AT+MD ($T = 0$ K) trajectories, providing 120 structures in total. For them single-point IPs were computed at the ADC(2)/cc-pVTZ level. Then, doing a similar binning, we obtained the corresponding PE spectrum (Fig. 8). The conventional Franck-Condon modeling directly evaluates the vibrational structure. However, this requires optimization of cationic excited states, which can be considerably complicated due to their strong multireference character, as in the case of benzene. Alternatively WS can be used as a fast and simple method for estimation of line shapes. The spectra computed on the basis of MD are slightly red-shifted in the first three bands with respect to the WS-based values. This is probably an effect of anharmonicity.

3.4.2 Protonated methane cation

Protonated methane (CH_5^+) is a strongly nonrigid molecular system. There have been measured vibrational spectra [61, 62] and investigated its structure and dynamics.[63, 70–72] Due to the significant flexibility of this system, we use it as a test case for the calculation of vibrational spectra. Four sets of MD simulations were performed, MBS-AT+MD at $T = 300$ K (29 trajectories), WS+MD (45 trajectories), SWS-AT+MD at $T = 0$ K (21 trajectories) and $T = 300$ K (23 trajectories). The gradients were calculated at the PBEh-3c level of theory. The time step was 1 fs for each trajectory. The total time span for each trajectory was set to 4 ps, except for WS+MD with 2 ps. First 0.5 ps were removed from each trajectory due to equilibration. To compare with standard models, harmonic and VPT2 calculations were performed in ORCA 5.

Vibrational spectra of CH_5^+ were measured using action spectroscopy [61], monitoring the yield of the $OCOH^+$ cation according to the reaction

$$CH_5^+ + CO_2 + h\nu \rightarrow CH_4 + OCOH^+ \ , \tag{11}$$

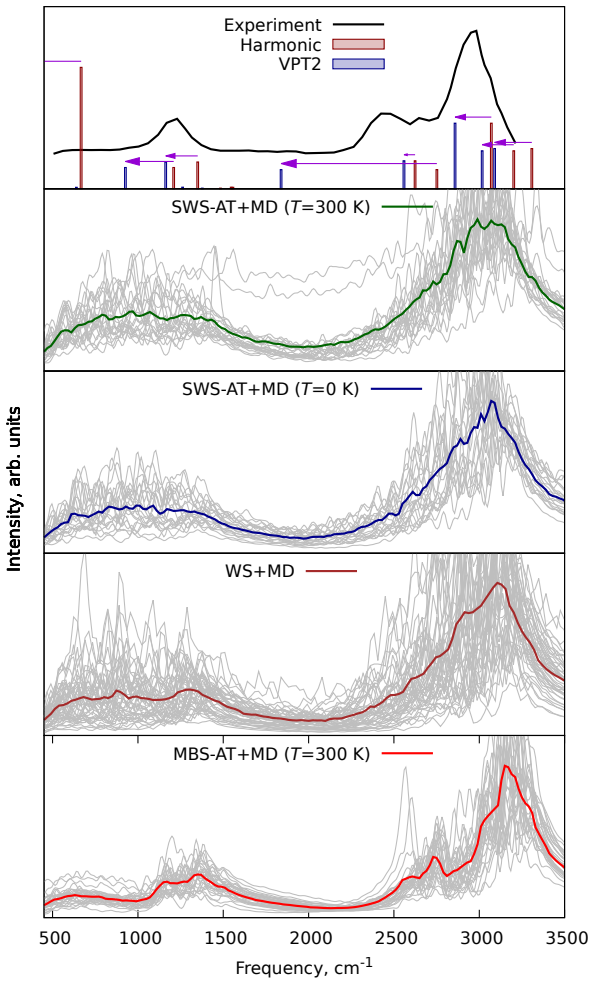

Figure 9: Experimental [61] (retrieved with WebPlotDigitizer [66]) and theoretical IR spectra of CH_5^+. Arrows indicate anharmonic VPT2 shifts.

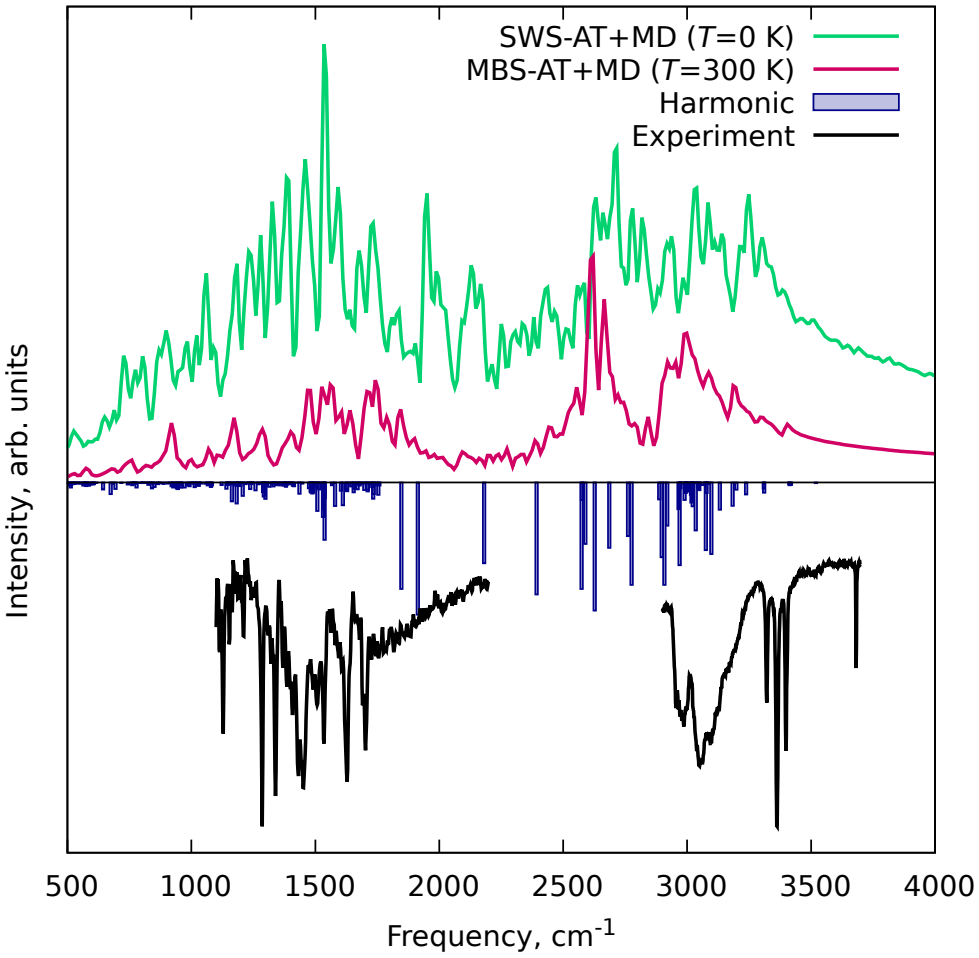

Figure 10: Experimental (black lines below) and theoretical IR vibrational spectra of the protonated serine octamer (Ser_8H^+). The two independent sections of the experimental data (1100–2200 cm^{-1} and 2900-3700 cm^{-1}) were independently scaled for a better view. Vertical bars (below) show the harmonic fundamentals.

where $h\nu$ is the IR photon. To account for this, we multiplied our IR spectra with a factor (see Appendix for details)

$$f(\nu) = \begin{cases} 0 \ , \ h\nu \leq D, \\ 1 - \frac{D}{h\nu} \ , \ h\nu > D, \end{cases} \tag{12}$$

where ν is the photon frequency, and the D is the energy threshold for the observed reaction. In our work we assumed $D = 383$ cm^{-1}, which corresponds to the Gibbs free energy of the probing reaction (Equation 11), computed in the following hybrid scheme. Geometry optimizations and frequency calculations were done at the PBEh-3c level of theory for 298 K, whereas single point energies were computed using the ae-CCSD(T) theory applying complete basis set (CBS) extrapolation technique based on the cc-pVTZ and cc-pVQZ sets.[73–75] The resulting spectra are shown in Figure 9. VPT2 fails by producing a negative value for the lowest fundamental frequency. MBS-AT+MD shows the fine band structure of the spectra, whereas the WS-based methods demonstrate broadening of the peaks without significant shifts.

3.4.3 Protonated serine octamer

The protonated serine octamer (Ser_8H^+, $C_{24}H_{57}N_8O_{24}{}^+$, see Figure 1) is a relatively large molecular system comprising 113 atoms. Its structure was investigated using helium-tagging IR spectroscopy by comparing harmonic frequencies with experimental spectra.[64] The VPT2 method for this system with $3 \times 113 - 6 = 333$ vibrational degrees of freedom requires calculation in total 667 Hessians – for two displaced structures along each of the 333 modes plus for the equilibrium structure. Our simple timing benchmark in Orca 5 has shown that a parallel VPT2 calculation for this system using a GGA functional with the 6-31G basis set[76] is equivalent to a MD simulation of about 13,000 frames. As it has been shown the VPT2 method is not reliable in application to vibrations with large amplitudes. On the other side, using molecular dynamics a decent representation of the Ser_8H^+ molecular ensemble can be reached even with fewer than 13,000 frames.

To demonstrate that, we have performed two MD simulations of Ser_8H^+ at the BLYP-D3BJ/6-31G level of theory, MD-AT+MD at $T = 300$ K and SWS-AT+MD at $T = 0$ K. The lengths of trajectories in both simulation sets were 3 ps. First 100 fs of the trajectories were discarded due to equilibration. Since experimental spectra were obtained using helium tag,[64] we applied the intensity correction factor given in Equation 12. The helium detachment energy D in the $He - Ser_8H^+$ complex was estimated to be 379 cm^{-1}, on the basis of GFN2-xTB[77] semi-empiric calculations. Figure 10 shows comparison of experimental and calculated spectra. The MD results demonstrate a much better agreement with experimental data in comparison to harmonic calculations.

4 Conclusions

Wigner sampling (WS) technique combined with classical molecular dynamics (MD) is a powerful combination describing nuclear quantum effects (NQEs) and vibrational anharmonicity at a reasonable computational cost. We have performed a benchmark evaluation of respective procedures for computing vibrationally-averaged ground state rotational constants for a set of molecules. The most robust way is the calculation of the mean inverse tensor of inertia along the trajectory. Molecular structures at each snapshot of trajectory should be oriented with respect to the reference structure.

The physically-justified WS procedure can also be reduced to a one-parameter heuristic procedure denoted as simplified Wigner sampling (SWS). The single parameter τ with the dimension of time controls the SWS method. The determined interval of optimal τ is $1 \leq \tau \leq 10$ fs and the value of $\tau = 2$ fs is optimal starting approximation for MD simulations. The adjustment of τ can be done using a variational procedure by minimizing the mean energy or temperature of the system.

WS and SWS also allow for effective treatment of NQEs in calculations of vibrational spectra. A better frequency resolution can be gained by combining the phase-space sampling with long trajectory duration. Constant temperature MD with the Andersen thermostat combined with the SWS routine partially allows to recover NQEs at a low computational cost.

5 Acknowledgements

This work has been supported by Deutsches Elektronen-Synchtrotron DESY, a member of the Helmholtz Association (HGF). Molecular dynamics calculations were enabled through the Maxwell computational resources operated at Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany. YuVV is grateful to Deutsche Forschungsgemeinschaft (grant VI 713/1-3) for support. A part of computations has been done at HPC facilities of the Universität zu Köln.

6 Conflict of interest

There are no conflicts to declare.

7 List of abbreviations

- MD – molecular dynamics,
- NQEs – nuclear quantum effects,
- PIMD – path-integral molecular dynamics,
- RMSD – root-mean-square deviation,
- PES – potential energy surface,
- VPT2 – vibrational perturbation theory of the second order,
- PTn – perturbation theory of the n-th order,
- MBS – Maxwell-Boltzmann sampling,
- PDF – pair distribution function,
- WS – Wigner sampling,
- SWS – simplified Wigner sampling,
- MRC – mean rotational constants,
- MIITU – mean inverse inertia tensor (unoriented),
- MIITO – mean inverse inertia tensor (oriented),

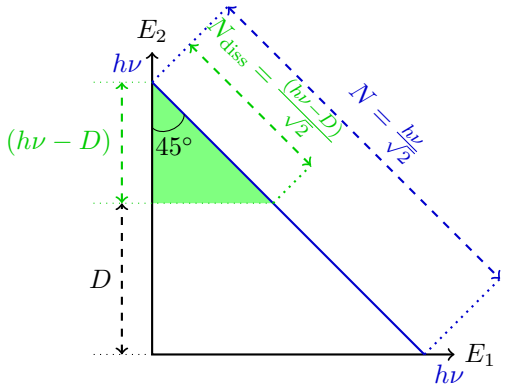

Figure 11: Scheme of the heuristic derivation of the damping factor in the action-type spectroscopy. See text for details.

- AT – Andersen thermostat,[2]
- FWHM – full width at half maximum,
- IR – infrared (spectrum),
- PE – photoelectron (spectrum),
- IP – ionization potential.

8 Appendix: derivation of the heuristic damping function for the action spectroscopy

Action spectroscopy detects absorption of photons by tracing down photon-induced changes in the molecular system. For example, in the case of tag spectroscopy the molecular system is tagged by a weakly bounded inert particle like inert gas atom or nitrogen molecule. If a vibrational mode absorbs a single photon with frequency ν, the resulting energy $h\nu$ is redistributed between modes, such that this leads to the detachment of the tag. Let us assume the existence of two vibrational modes in the molecule: #1, which is photoexcited, and #2 corresponding to the dissociation of the tag. We also assume that vibrations are classical and thus the total energy (kinetic plus potential) of each mode $E_n \geq 0$ ($n = 1, 2$). Upon the photon absorption, the total energy of the system $E = E_1 + E_2$ increases by the photon energy $h\nu$ and becomes $E = h\nu = E_1 + E_2$. The energy of the second mode (E_2) during redistribution process may overcome the dissociation energy (D). In this case a signal is produced and detected. We can imagine the total number of states N with the redistributed energy $h\nu$ as a straight line $h\nu = E_1 + E_2$ ($0 \leq E_1, E_2 \leq h\nu$) in coordinates $E_1 - E_2$ (see Figure 11) and $N = h\nu \cdot \cos(45°) = \frac{\sqrt{2}h\nu}{2}$. The number of tag-dissociating states is then a subsection of this line with $D < E_2 \leq h\nu$. If the photon energy is smaller than the dissociation energy ($h\nu \leq D$), the removal of the tag cannot happen, and thus the number of dissociated molecules is $N_{\text{diss}} = 0$. Otherwise, when $h\nu > D$, the number of the dissociated molecules is calculated as $N_{\text{diss}} = (h\nu - D) \cdot \cos(45°) = \frac{\sqrt{2}(h\nu - D)}{2}$ (see Figure 11). Thus we can find the quantum yield of the dissociation per absorbed photon with frequency ν as

$$f(\nu) = \frac{N_{\text{diss}}}{N} = \begin{cases} 0\ , & h\nu \leq D, \\ 1 - \frac{D}{h\nu}\ , & h\nu > D. \end{cases}$$

We can approximate the final observable action spectroscopy signal as $I(\nu) \cdot f(\nu)$, where $I(\nu)$ is the absorption spectrum of the molecule and $f(\nu)$ is the quantum yield of the dissociation.

References

[1] Thomas E. Markland and Michele Ceriotti. Nuclear quantum effects enter the mainstream. *Nature Reviews Chemistry*, 2(3):0109, Feb 2018. ISSN 2397-3358. doi: 10.1038/s41570-017-0109. URL https://doi.org/10.1038/s41570-017-0109.

[2] Scott A. Hollingsworth and Ron O. Dror. Molecular dynamics simulation for all. *Neuron*, 99(6):1129–1143, 2018. ISSN 0896-6273. doi: https://doi.org/10.1016/j.neuron.2018.08.011. URL https://www.sciencedirect.com/science/article/pii/S0896627318306846.

[2] X-AT with X=MBS, SWS means the same thermostat applied in the canonical fashion, and with modified SWS sampling procedure.

[3] Radu Iftimie, Peter Minary, and Mark E. Tuckerman. *Ab initio* molecular dynamics: Concepts, recent developments, and future trends. *Proceedings of the National Academy of Sciences*, 102(19):6654–6659, 2005. doi: 10.1073/pnas.0500193102. URL https://www.pnas.org/doi/abs/10.1073/pnas.0500193102.

[4] Dominik Marx and Michele Parrinello. Ab initio path integral molecular dynamics: Basic ideas. *The Journal of Chemical Physics*, 104(11):4077–4082, 1996. doi: 10.1063/1.471221. URL https://doi.org/10.1063/1.471221.

[5] Stuart C. Althorpe. Path-integral approximations to quantum dynamics. *The European Physical Journal B*, 94 (7):155, Jul 2021. ISSN 1434-6036. doi: 10.1140/epjb/s10051-021-00155-2. URL https://doi.org/10.1140/epjb/s10051-021-00155-2.

[6] Walther Caminati and Jens-Uwe Grabow. Chapter 15 - microwave spectroscopy: Molecular systems. In Jaan Laane, editor, *Frontiers of Molecular Spectroscopy*, pages 455–552. Elsevier, Amsterdam, 2009. ISBN 978-0-444-53175-9. doi: https://doi.org/10.1016/B978-0-444-53175-9.00015-5. URL https://www.sciencedirect.com/science/article/pii/B9780444531759000155.

[7] Richard E. Smalley, Lennard Wharton, and Donald H. Levy. Molecular optical spectroscopy with supersonic beams and jets. *Accounts of Chemical Research*, 10(4):139–145, 1977. doi: 10.1021/ar50112a006. URL https://doi.org/10.1021/ar50112a006.

[8] P.W. Atkins and R.S. Friedman. *Molecular Quantum Mechanics*. OUP Oxford, 2011. ISBN 9780199541423. URL https://books.google.de/books?id=9k-cAQAAQBAJ.

[9] Peter Pulay, Geza Fogarasi, Gabor Pongor, James E. Boggs, and Anna Vargha. Combination of theoretical ab initio and experimental information to obtain reliable harmonic force constants. scaled quantum mechanical (qm) force fields for glyoxal, acrolein, butadiene, formaldehyde, and ethylene. *Journal of the American Chemical Society*, 105(24):7037–7047, 1983. doi: 10.1021/ja00362a005. URL https://doi.org/10.1021/ja00362a005.

[10] Manoj K. Kesharwani, Brina Brauer, and Jan M. L. Martin. Frequency and zero-point vibrational energy scale factors for double-hybrid density functionals (and other selected methods): Can anharmonic force fields be avoided? *The Journal of Physical Chemistry A*, 119(9):1701–1714, 2015. doi: 10.1021/jp508422u. URL https://doi.org/10.1021/jp508422u. PMID: 25296165.

[11] Ricardo A. Mata and Martin A. Suhm. Benchmarking quantum chemical methods: Are we heading in the right direction? *Angewandte Chemie International Edition*, 56(37):11011–11018, 2017. doi: https://doi.org/10.1002/anie.201611308. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201611308.

[12] *Mechanics, Third Edition: Volume 1 (Course of Theoretical Physics)*. Butterworth-Heinemann, 3 edition, January 1976. ISBN 0750628960.

[13] Yury V. Vishnevskiy and Denis Tikhonov. Quantum corrections to parameters of interatomic distance distributions in molecular dynamics simulations. *Theoretical Chemistry Accounts*, 135(4):88, Mar 2016. ISSN 1432-2234. doi: 10.1007/s00214-016-1848-2. URL https://doi.org/10.1007/s00214-016-1848-2.

[14] Chen Qu and Joel M. Bowman. Quantum approaches to vibrational dynamics and spectroscopy: is ease of interpretation sacrificed as rigor increases? *Phys. Chem. Chem. Phys.*, 21:3397–3413, 2019. doi: 10.1039/C8CP04990D. URL http://dx.doi.org/10.1039/C8CP04990D.

[15] Denis S. Tikhonov, Dmitry I. Sharapa, Jan Schwabedissen, and Vladimir V. Rybkin. Application of classical simulations for the computation of vibrational properties of free molecules. *Phys. Chem. Chem. Phys.*, 18:28325–28338, 2016. doi: 10.1039/C6CP05849C. URL http://dx.doi.org/10.1039/C6CP05849C.

[16] Vincenzo Barone. Anharmonic vibrational properties by a fully automated second-order perturbative approach. *The Journal of Chemical Physics*, 122(1):014108, 2005. doi: 10.1063/1.1824881. URL https://doi.org/10.1063/1.1824881.

[17] Qin Yang, Marco Mendolicchio, Vincenzo Barone, and Julien Bloino. Accuracy and reliability in the simulation of vibrational spectra: A comprehensive benchmark of energies and intensities issuing from generalized vibrational perturbation theory to second order (gvpt2). *Frontiers in Astronomy and Space Sciences*, 8, 2021. ISSN 2296-987X. doi: 10.3389/fspas.2021.665232. URL https://www.frontiersin.org/article/10.3389/fspas.2021.665232.

[18] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand,

K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian 16 Revision C.01, 2016. Gaussian Inc. Wallingford CT.

[19] Frank Neese, Frank Wennmohs, Ute Becker, and Christoph Riplinger. The orca quantum chemistry program package. *The Journal of Chemical Physics*, 152(22):224108, 2020. doi: 10.1063/5.0004608. URL https://doi.org/10.1063/5.0004608.

[20] Giuseppe M. J. Barca, Colleen Bertoni, Laura Carrington, Dipayan Datta, Nuwan De Silva, J. Emiliano Deustua, Dmitri G. Fedorov, Jeffrey R. Gour, Anastasia O. Gunina, Emilie Guidez, Taylor Harville, Stephan Irle, Joe Ivanic, Karol Kowalski, Sarom S. Leang, Hui Li, Wei Li, Jesse J. Lutz, Ilias Magoulas, Joani Mato, Vladimir Mironov, Hiroya Nakata, Buu Q. Pham, Piotr Piecuch, David Poole, Spencer R. Pruitt, Alistair P. Rendell, Luke B. Roskop, Klaus Ruedenberg, Tosaporn Sattasathuchana, Michael W. Schmidt, Jun Shen, Lyudmila Slipchenko, Masha Sosonkina, Vaibhav Sundriyal, Ananta Tiwari, Jorge L. Galvez Vallejo, Bryce Westheimer, Marta Wloch, Peng Xu, Federico Zahariev, and Mark S. Gordon. Recent developments in the general atomic and molecular electronic structure system. *The Journal of Chemical Physics*, 152(15):154102, April 2020. ISSN 0021-9606, 1089-7690. doi: 10.1063/5.0005188. URL http://aip.scitation.org/doi/10.1063/5.0005188.

[21] Devin A. Matthews, Lan Cheng, Michael E. Harding, Filippo Lipparini, Stella Stopkowicz, Thomas-C. Jagau, Péter G. Szalay, Jürgen Gauss, and John F. Stanton. Coupled-cluster techniques for computational chemistry: The CFOUR program package. *J. Chem. Phys.*, 152(21):214108, 2020. doi: 10.1063/5.0004837.

[22] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:749–759, Jun 1932. doi: 10.1103/PhysRev.40.749. URL https://link.aps.org/doi/10.1103/PhysRev.40.749.

[23] Lipeng Sun and William L. Hase. Comparisons of classical and wigner sampling of transition state energy levels for quasiclassical trajectory chemical dynamics simulations. *The Journal of Chemical Physics*, 133(4):044313, 2010. doi: 10.1063/1.3463717. URL https://doi.org/10.1063/1.3463717.

[24] J. Patrick Zobel, Juan J. Nogueira, and Leticia González. Finite-temperature wigner phase-space sampling and temperature effects on the excited-state dynamics of 2-nitronaphthalene. *Phys. Chem. Chem. Phys.*, 21:13906–13915, 2019. doi: 10.1039/C8CP03273D. URL http://dx.doi.org/10.1039/C8CP03273D.

[25] J. Patrick Zobel, Moritz Heindl, Juan J. Nogueira, and Leticia González. Vibrational sampling and solvent effects on the electronic structure of the absorption spectrum of 2-nitronaphthalene. *Journal of Chemical Theory and Computation*, 14(6):3205–3217, 2018. doi: 10.1021/acs.jctc.8b00198. URL https://doi.org/10.1021/acs.jctc.8b00198. PMID: 29694042.

[26] Davide Avagliano, Emilio Lorini, and Leticia González. Sampling effects in quantum mechanical/molecular mechanics trajectory surface hopping non-adiabatic dynamics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 380(2223):20200381, 2022. doi: 10.1098/rsta.2020.0381. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2020.0381.

[27] Jmol development team. Jmol: an open-source java viewer for chemical structures in 3d. http://www.jmol.org/. URL http://jmol.sourceforge.net/.

[28] Hans C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72(4):2384–2393, 1980. doi: 10.1063/1.439486. URL https://doi.org/10.1063/1.439486.

[29] Mario Barbatti and Kakali Sen. Effects of different initial condition samplings on photodynamics and spectrum of pyrrole. *International Journal of Quantum Chemistry*, 116(10):762–771, 2016. doi: https://doi.org/10.1002/qua.25049. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.25049.

[30] S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79 – 86, 1951. doi: 10.1214/aoms/1177729694. URL https://doi.org/10.1214/aoms/1177729694.

[31] Denis S. Tikhonov, Yury V. Vishnevskiy, Anatolii N. Rykov, Olga E. Grikina, and Leonid S. Khaikin. Semi-experimental equilibrium structure of pyrazinamide from gas-phase electron diffraction. how much experimental is it? *Journal of Molecular Structure*, 1132:20–27, 2017. ISSN 0022-2860. doi: https://doi.org/10.1016/j.molstruc.2016.05.090. URL https://www.sciencedirect.com/science/article/pii/S0022286016305579. Gas electron diffraction and molecular structure.

[32] Denis S. Tikhonov. https://stash.desy.de/projects/MOLINC, 2020.

[33] Daniel T. Colbert and William H. Miller. A novel discrete variable representation for quantum mechanical reactive scattering via the s-matrix kohn method. *The Journal of Chemical Physics*, 96(3):1982–1991, 1992. doi: 10.1063/1.462100. URL https://doi.org/10.1063/1.462100.

[34] Michele Ceriotti, Giovanni Bussi, and Michele Parrinello. Nuclear quantum effects in solids using a colored-noise thermostat. *Phys. Rev. Lett.*, 103:030603, Jul 2009. doi: 10.1103/PhysRevLett.103.030603. URL https://link.aps.org/doi/10.1103/PhysRevLett.103.030603.

[35] Hichem Dammak, Yann Chalopin, Marine Laroche, Marc Hayoun, and Jean-Jacques Greffet. Quantum thermal bath for molecular dynamics simulation. *Phys. Rev. Lett.*, 103:190601, Nov 2009. doi: 10.1103/PhysRevLett.103.190601. URL https://link.aps.org/doi/10.1103/PhysRevLett.103.190601.

[36] V. A. Levashov, S. J. L. Billinge, and M. F. Thorpe. Quantum correction to the pair distribution function. *Journal of Computational Chemistry*, 28(11):1865–1882, 2007. doi: https://doi.org/10.1002/jcc.20713. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20713.

[37] Victor A. Sipachev. Diffraction measurements and equilibrium parameters. *Adv. Phys. Chem.*, 2011:1–14, 2011.

[38] Peter Atkins and Julio Paula. *Atkins' physical chemistry*. Oxford University press, 2008. ISBN 9780195685220.

[39] Vincenzo Barone. Anharmonic vibrational properties by a fully automated second-order perturbative approach. *The Journal of Chemical Physics*, 122(1):014108, 2005. doi: 10.1063/1.1824881. URL http://scitation.aip.org/content/aip/journal/jcp/122/1/10.1063/1.1824881.

[40] Natalaja Vogt, Jürgen Vogt, and Jean Demaison. Accuracy of the rotational constants. *Journal of Molecular Structure*, 988(1):119–127, 2011. ISSN 0022-2860. doi: https://doi.org/10.1016/j.molstruc.2010.12.040. URL https://www.sciencedirect.com/science/article/pii/S0022286010010070.

[41] Silvia Alessandrini, Jürgen Gauss, and Cristina Puzzarini. Accuracy of rotational parameters predicted by high-level quantum-chemical calculations: Case study of sulfur-containing molecules of astrochemical interest. *Journal of Chemical Theory and Computation*, 14(10):5360–5371, 2018. doi: 10.1021/acs.jctc.8b00695. URL https://doi.org/10.1021/acs.jctc.8b00695. PMID: 30141928.

[42] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple [phys. rev. lett. 77, 3865 (1996)]. *Phys. Rev. Lett.*, 78:1396–1396, Feb 1997. doi: 10.1103/PhysRevLett.78.1396. URL https://link.aps.org/doi/10.1103/PhysRevLett.78.1396.

[43] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.*, 7:3297–3305, 2005. doi: 10.1039/B508541A. URL http://dx.doi.org/10.1039/B508541A.

[44] Sebastian Mai, Martin Richter, Moritz Heindl, Maximilian F. S. J. Menger, Andrew Atkins, Matthias Ruckenbauer, Felix Plasser, Lea Maria Ibele, Simon Kropf, Markus Oppel, Philipp Marquetand, and Leticia González. Sharc2.1: Surface hopping including arbitrary couplings — program package for non-adiabatic dynamics. sharc-md.org, 2019.

[45] Martin Richter, Philipp Marquetand, Jesús González-Vázquez, Ignacio Sola, and Leticia González. SHARC: ab initio molecular dynamics with surface hopping in the adiabatic representation including arbitrary couplings. *J. Chem. Theory Comput.*, 7(5):1253–1258, 2011. doi: 10.1021/ct1007394. URL http://dx.doi.org/10.1021/ct1007394.

[46] Sebastian Mai, Philipp Marquetand, and Leticia González. Nonadiabatic dynamics: The sharc approach. *WIREs Comput. Mol. Sci.*, 8:e1370, 2018. doi: 10.1002/wcms.1370. URL http://dx.doi.org/10.1002/wcms.1370.

[47] Denis S. Tikhonov. Pyramd. https://stash.desy.de/projects/PYRAMD/repos/pyramd/browse, 2021.

[48] Denis S. Tikhonov. Pyramd manual. https://confluence.desy.de/display/CFA/PyRAMD, 2021.

[49] W. Kabsch. A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, 32(5):922–923, Sep 1976. doi: 10.1107/S0567739476001873. URL https://doi.org/10.1107/S0567739476001873.

[50] Yury V. Vishnevskiy, Jan Schwabedissen, Anatolii N. Rykov, Vladimir V. Kuznetsov, and Nina N. Makhova. Conformational and bonding properties of 3,3-dimethyl- and 6,6-dimethyl-1,5-diazabicyclo[3.1.0]hexane: A case study employing the monte carlo method in gas electron diffraction. *The Journal of Physical Chemistry A*, 119(44):10871–10881, 2015. doi: 10.1021/acs.jpca.5b08228. URL https://doi.org/10.1021/acs.jpca.5b08228. PMID: 26461037.

[51] Andrey A. Fokin, Tatyana S. Zhuk, Sebastian Blomeyer, Cristóbal Pérez, Lesya V. Chernish, Alexander E. Pashenko, Jens Antony, Yury V. Vishnevskiy, Raphael J. F. Berger, Stefan Grimme, Christian Logemann, Melanie Schnell, Norbert W. Mitzel, and Peter R. Schreiner. Intramolecular london dispersion interaction effects on gas-phase and solid-state structures of diamondoid dimers. *Journal of the American Chemical Society*, 139(46):16696–16707, 2017. doi: 10.1021/jacs.7b07884. URL https://doi.org/10.1021/jacs.7b07884. PMID: 29037036.

[52] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry*, 32(7):1456–1465, 2011. doi: 10.1002/jcc.21759. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21759.

[53] Martin Thomas, Martin Brehm, Reinhold Fligg, Peter Vöhringer, and Barbara Kirchner. Computing vibrational spectra from ab initio molecular dynamics. *Phys. Chem. Chem. Phys.*, 15:6608–6622, 2013. doi: 10.1039/C3CP44302G. URL http://dx.doi.org/10.1039/C3CP44302G.

[54] Edward Ditler and Sandra Luber. Vibrational spectroscopy by means of first-principles molecular dynamics simulations. *WIREs Computational Molecular Science*, n/a(n/a):e1605. doi: https://doi.org/10.1002/wcms.1605.

[55] Jan Horníček, Petra Kaprálová, and Petr Bouř. Simulations of vibrational spectra from classical trajectories: Calibration with ab initio force fields. *The Journal of Chemical Physics*, 127(8):084502, 2007. doi: 10.1063/1.2756837. URL http://scitation.aip.org/content/aip/journal/jcp/127/8/10.1063/1.2756837.

[56] Denis S. Tikhonov. Simple posterior frequency correction for vibrational spectra from molecular dynamics. *The Journal of Chemical Physics*, 144(17):174108, 2016. doi: 10.1063/1.4948320. URL https://doi.org/10.1063/1.4948320.

[57] In Heo, Jong Chan Lee, Begüm Rukiye Özer, and Thomas Schultz. Structure of benzene from mass-correlated rotational raman spectroscopy. *RSC Adv.*, 12:21406–21416, 2022. doi: 10.1039/D2RA03431J. URL http://dx.doi.org/10.1039/D2RA03431J.

[58] "Infrared Spectra" by NIST Mass Spectrometry Data Center, William E. Wallace, director, in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899, https://webbook.nist.gov/chemistry/, (retrieved January 11, 2023).

[59] Jürgen Gauss and John F. Stanton. The equilibrium structure of benzene. *The Journal of Physical Chemistry A*, 104(13):2865–2868, 2000. doi: 10.1021/jp994408y. URL https://doi.org/10.1021/jp994408y.

[60] Andrew J Yencha, Richard I Hall, Lorenzo Avaldi, Grant Dawber, Andrew G McConkey, Michael A MacDonald, and George C King. Threshold photoelectron spectroscopy of benzene up to 26.5 ev. *Canadian Journal of Chemistry*, 82(6):1061–1066, 2004. doi: 10.1139/v04-057. URL https://doi.org/10.1139/v04-057.

[61] Oskar Asvany, Padma Kumar P, Britta Redlich, Ilka Hegemann, Stephan Schlemmer, and Dominik Marx. Understanding the infrared spectrum of bare ch¡sub¿5¡/sub¿¡sup¿+¡/sup¿. *Science*, 309(5738):1219–1222, 2005. doi: 10.1126/science.1113729. URL https://www.science.org/doi/abs/10.1126/science.1113729.

[62] Oskar Asvany, Koichi M. T. Yamada, Sandra Brünken, Alexey Potapov, and Stephan Schlemmer. Experimental ground-state combination differences of ch¡sub¿5¡/sub¿¡sup¿+¡/sup¿. *Science*, 347(6228):1346–1349, 2015. doi: 10.1126/science.aaa3304. URL https://www.science.org/doi/abs/10.1126/science.aaa3304.

[63] Xinchuan Huang, Anne B. McCoy, Joel M. Bowman, Lindsay M. Johnson, Chandra Savage, Feng Dong, and David J. Nesbitt. Quantum deconstruction of the infrared spectrum of ch¡sub¿5¡/sub¿¡sup¿+¡/sup¿. *Science*, 311(5757):60–63, 2006. doi: 10.1126/science.1121166. URL https://www.science.org/doi/abs/10.1126/science.1121166.

[64] Valeriu Scutelnic, Marta A. S. Perez, Mateusz Marianski, Stephan Warnke, Aurelien Gregor, Ursula Rothlisberger, Michael T. Bowers, Carsten Baldauf, Gert von Helden, Thomas R. Rizzo, and Jongcheol Seo. The structure of the protonated serine octamer. *Journal of the American Chemical Society*, 140(24):7554–7560, 2018. doi: 10.1021/jacs.8b02118. URL https://doi.org/10.1021/jacs.8b02118. PMID: 29637771.

[65] Stefan Grimme, Jan Gerit Brandenburg, Christoph Bannwarth, and Andreas Hansen. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *The Journal of Chemical Physics*, 143(5):054107, 2015. doi: 10.1063/1.4927476. URL https://doi.org/10.1063/1.4927476.

[66] Ankit Rohatgi. Webplotdigitizer: Version 4.6, 2022. URL https://automeris.io/WebPlotDigitizer.

[67] Samer Gozem and Anna I. Krylov. The ezspectra suite: An easy-to-use toolkit for spectroscopy modeling. *WIREs Computational Molecular Science*, 12(2):e1546, 2022. doi: https://doi.org/10.1002/wcms.1546. URL https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1546.

[68] Marta L. Vidal, Michael Epshtein, Valeriu Scutelnic, Zheyue Yang, Tian Xue, Stephen R. Leone, Anna I. Krylov, and Sonia Coriani. Interplay of open-shell spin-coupling and jahn–teller distortion in benzene radical cation probed by x-ray spectroscopy. *The Journal of Physical Chemistry A*, 124(46):9532–9541, 2020. doi: 10.1021/acs.jpca.0c08732. URL https://doi.org/10.1021/acs.jpca.0c08732. PMID: 33103904.

[69] Michael Epshtein, Valeriu Scutelnic, Zheyue Yang, Tian Xue, Marta L. Vidal, Anna I. Krylov, Sonia Coriani, and Stephen R. Leone. Table-top x-ray spectroscopy of benzene radical cation. *The Journal of Physical Chemistry A*, 124(46):9524–9531, 2020. doi: 10.1021/acs.jpca.0c08736. URL https://doi.org/10.1021/acs.jpca.0c08736. PMID: 33107734.

[70] Dominik Marx and Michele Parrinello. Ch¡sub¿5¡/sub¿¡sup¿+¡/sup¿: The cheshire cat smiles. *Science*, 284(5411): 59–61, 1999. doi: 10.1126/science.284.5411.59. URL https://www.science.org/doi/abs/10.1126/science.284.5411.59.

[71] Padma Kumar P. and Dominik Marx. Understanding hydrogen scrambling and infrared spectrum of bare ch5+ based on ab initio simulations. *Phys. Chem. Chem. Phys.*, 8:573–586, 2006. doi: 10.1039/B513089C. URL http://dx.doi.org/10.1039/B513089C.

[72] Keiran C. Thompson, Deborah L. Crittenden, and Meredith J. T. Jordan. Ch5+: Chemistryss chameleon unmasked. *Journal of the American Chemical Society*, 127(13):4954–4958, 2005. doi: 10.1021/ja0482280. URL https://doi.org/10.1021/ja0482280. PMID: 15796561.

[73] Shijun Zhong, Ericka C. Barnes, and George A. Petersson. Uniformly convergent n-tuple-ζ augmented polarized (nzap) basis sets for complete basis set extrapolations. i. self-consistent field energies. *The Journal of Chemical Physics*, 129(18):184116, 2008. doi: 10.1063/1.3009651. URL https://doi.org/10.1063/1.3009651.

[74] Frank Neese and Edward F. Valeev. Revisiting the atomic natural orbital approach for basis sets: Robust systematic basis sets for explicitly correlated and conventional correlated ab initio methods? *Journal of Chemical Theory and Computation*, 7(1):33–43, 2011. doi: 10.1021/ct100396y. URL https://doi.org/10.1021/ct100396y. PMID: 26606216.

[75] Trygve Helgaker, Wim Klopper, Henrik Koch, and Jozef Noga. Basis-set convergence of correlated calculations on water. *The Journal of Chemical Physics*, 106(23):9639–9646, 1997. doi: 10.1063/1.473863. URL https://doi.org/10.1063/1.473863.

[76] W. J. Hehre, R. Ditchfield, and J. A. Pople. Self—consistent molecular orbital methods. xii. further extensions of gaussian—type basis sets for use in molecular orbital studies of organic molecules. *The Journal of Chemical Physics*, 56(5):2257–2261, 1972. doi: 10.1063/1.1677527. URL https://doi.org/10.1063/1.1677527.

[77] Christoph Bannwarth, Sebastian Ehlert, and Stefan Grimme. Gfn2-xtb—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *Journal of Chemical Theory and Computation*, 15(3):1652–1671, 2019. doi: 10.1021/acs.jctc.8b01176. URL https://doi.org/10.1021/acs.jctc.8b01176. PMID: 30741547.