Check for updates

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

Denis S. Tikhonov [ID] *[a] and Yury V. Vishnevskiy [ID] *[b]

In this work we discuss the generally applicable Wigner sampling and introduce a new, simplified Wigner sampling method, for computationally effective modeling of molecular properties containing nuclear quantum effects and vibrational anharmonicity. For various molecular systems test calculations of (a) vibrationally averaged rotational constants, (b) vibrational IR spectra and (c) photoelectron spectra have been performed. The performance of Wigner sampling has been assessed by comparing with experimental data and with results of other theoretical models, including harmonic and VPT2 approximations. The developed simplified Wigner sampling method shows advantages in application to large and flexible molecules.

1. Introduction

Molecular dynamics (MD) is a very powerful tool for modeling 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 $T_D = h\nu/k_B$ 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 $N_{\min} = \frac{h\nu}{k_B T}$. Therefore at low temperatures, this number gets very large thus diminishing the applicability of PIMD.

In contrast, 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

$$V(r) = \overbrace{V(r_e) + \frac{V_2}{2}\xi^2}^{V_h(r)} + \overbrace{\frac{V_3}{6}\xi^3 + \frac{V_4}{24}\xi^4 + \ldots}^{W_a(r)} \approx V_h(r). \quad (1)$$

Here, V denotes the potential energy, r_e is the equilibrium distance between atoms, $\xi = r - r_e$ is the displacement of the structure from the equilibrium geometry, $V_n = \frac{\partial^n V}{\partial r^n}(r_e)$, $V_h(r)$ is the harmonic potential, and $W_a(r)$ is the anharmonic correction to the harmonic potential. The quantum harmonic oscillator is one of the simplest quantum mechanical problems with a known analytical solution.[6] 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 symmetrical with respect to inversion of the displacement, $V_h(+\xi) = V_h(-\xi)$. On the other hand, the vibrational amplitude l, defined as

$$l^2 = \langle r^2\rangle - \langle r\rangle^2 \quad (2)$$

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

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

The harmonic approximation combined with the frequency scaling is a very robust, computationally affordable, and simple way for calculation of the vibrational properties.[7] For instance, the root-mean-square deviation (RMSD) of the scaled harmonic frequencies from the experimentally obtained fundamental

[a] Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany. E-mail: denis.tikhonov@desy.de

[b] 2Lehrstuhl für Anorganische Chemie und Strukturchemie, Fakultät für Chemie, Universität Bielefeld, Universitätsstr. 25, 33615 Bielefeld, Germany. E-mail: yury.vishnevskiy@uni-bielefeld.de

transitions is only around 20 cm^{-1} at the B3LYP/def2-TZVP level of theory.[8] However, the direct comparison of theoretical values computed in this way with experimental data may be of limited use, because real observables in the rotational or vibrational spectroscopy are generally anharmonic.[8,9] For example, in classical mechanics the angular vibrational frequency can be approximately expressed as[10]

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

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 (eqn (1)). This dependence indicates larger deviations from the harmonic frequency at higher vibrational excitations. A similar situation is observed in rotational spectroscopy, where rotational constants B can be considered as observables. For a diatomic molecule with the interatomic distance r and moment of inertia $I = \mu r^2$, taking into account the relationship

$$\frac{1}{I} \propto \frac{1}{r^2} = \frac{1}{(r_e + \xi)^2} \approx \frac{1}{r_e^2}\cdot\left(1 - \frac{2\xi}{r_e}\right),$$

the rotational constant may be given as[6]

$$B = \frac{\hbar}{4\pi c}\left\langle\frac{1}{I}\right\rangle \approx B_e - \frac{2B_e}{r_e}\langle\xi\rangle, \qquad (5)$$

where c is the speed of light, and $B_e = \hbar(4\pi cI_e)^{-1} = \hbar(4\pi c\mu r_e^2)^{-1}$ is the rotational constant corresponding to the structure at 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[11]

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

For bond stretching vibrations typically $V_3 < 0$ and $V_2 > 0$. Therefore, equilibrium rotational constants are usually larger than the 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.[12,13] 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 difficulties associated with these procedures limit the widespread usage of such methods.

The simplest quantum approach that offers anharmonic treatment of the molecular vibrations is the second-order perturbation theory, usually denoted as VPT2.[14,15] 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 a procedure can be automated and is already implemented in different software packages for quantum chemistry. 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.[13] Thus it is important to develop procedures for computation of molecular vibrational properties at low temperatures using classical MD simulations.

Wigner sampling[16,17] is one of possibilities for producing initial conditions in MD simulations, that emulate the quantum harmonic ground vibrational state.[18–20] In this work we combine Wigner sampling with MD simulations to provide anharmonic vibrational properties of molecules. We also propose a new simplified sampling procedure, which is based on Wigner sampling but does not require preliminary calculation of molecular Hessian.

The paper is structured as follows. First, we introduce and compare possible sampling procedures, including our simplified version of 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. 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). Next, we show the application of the sampling routines for the calculation of vibrational spectra for difluoromethane. Finally, different sampling methods are applied to several systems with known experimental data. All the molecules used throughout this work are shown in Fig. 1.

2. Theory

2.1 Maxwell–Boltzmann sampling

The classic procedure for the generation of initial conditions in MD simulations is Maxwell–Boltzmann sampling (MBS).[22] 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

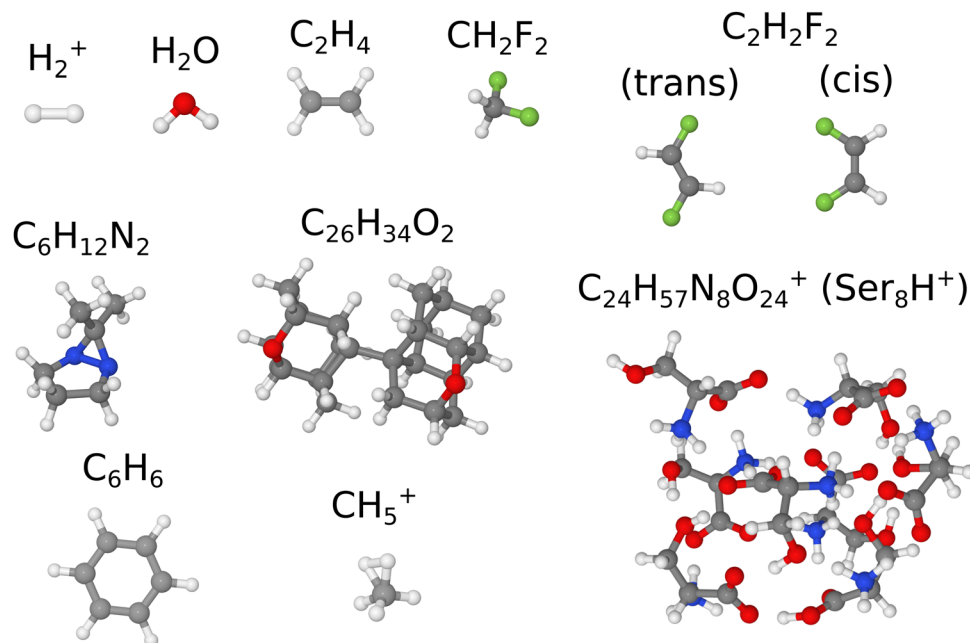$$v_\alpha \sim \exp\left(-\frac{mv_\alpha^2}{2k_BT}\right),$$

This journal is © the Owner Societies 2023

Phys. Chem. Chem. Phys., 2023, **25**, 18406–18423 | **18407**

Fig. 1 Test molecules. The standard Jmol[21] coloring scheme is used for the atoms: white (H), gray (C), blue (N), red (O), and green (F).

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 (eqn (2)) for the interatomic distance r in the classical harmonic approximation is:[11]

$$l_{class}^2 = \frac{k_B T}{\mu \omega^2},$$

whereas in the quantum case it is[11]

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

where l_0 is the vibrational amplitude for the ground vibrational state (see eqn (3)), which can also be defined as $l_0 = l_{quant}(T = 0)$. For each given temperature, the classical vibrational amplitude does not exceed the quantum one ($l_{class} \leq l_{quant}$). 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 $l_{class} = l_0$, we obtain

$$T = \frac{\hbar\omega}{2k_B} = \frac{T_D}{2},$$

i.e., at half of the Debye temperature $T_D = h\nu/k_B$ 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.[17,18,20,22] Let us consider the Hamiltonian

$\hat{H} = \frac{1}{2\mu}\hat{p}^2 + \frac{\mu\omega^2}{2}\xi^2$ for a single vibrational 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 eqn (6). The corresponding Wigner quasiprobability function for the momentum p can be obtained from any given wavefunction ψ via transformation[16]

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

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

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

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 **r** from their equilibrium positions ($\mathbf{r}_e$) can be related to the vector of displacements of vibrational modes ξ through the matrix of the mode shapes: $\mathcal{L}$: $\mathbf{R} = (\mathbf{r} - \mathbf{r}_e) = \mathcal{L}\xi$. A similar relationship connects the Cartesian momentums of the

atoms $\mathbf{P} = -i\hbar\partial_{\mathbf{R}}$ to the momentums of the independent modes $\mathbf{p} = -i\hbar\partial_{\xi}$: $\mathbf{P} = \mathcal{L}^{-1}\mathbf{p}$. The Wigner function for ξ and $\mathbf{p}$ is just a product of the one-dimensional quasi-probability distributions (eqn (7)), and it can be written as

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

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

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

Here, the effective variance matrix $\tilde{\Sigma}_{\mathbf{R}}$ is non-diagonal, unlike the initial Σ_{ξ}, which indicates the correlation between atomic displacements due to the collective nature of the vibrational modes. The replacement of the variables $\{\mathbf{p},\xi\} \rightarrow \{\mathbf{P}, \mathbf{R}\}$ is unitary, since $\mathrm{d}\mathbf{p}\mathrm{d}\xi = \mathcal{L}^{-1}\mathrm{d}\mathbf{R}\mathcal{L}\mathrm{d}\mathbf{P} = \mathrm{d}\mathbf{R}\mathrm{d}\mathbf{P}$. 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.[17] 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 eqn (6), we can also sample temperature-averaged vibrational states.

2.3 New sampling procedure

WS requires initial optimization of the molecular structure and calculation of harmonic potential. However, we can try to construct a simplified sampling procedure similar to WS. Here we denote it as simplified Wigner sampling (SWS).

Let us consider the motion of a nucleus with the mass m along the coordinate x. We will assume that at each point in time, this motion is described with a Gaussian wavepacket with width σ_x in the coordinate representation and the corresponding width σ_p in the momentum representation. In this case, 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 eqn (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 and $p = mv$ 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. Alternatively, 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 larger the uncertainty for the momentum, the smaller it is for the coordinate, and *vice versa*. We can think of the uncertainty as the amount of information we have about x and p. As a guiding principle, we require the same amount of information for 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[23] defined as

$$D_{\mathrm{KL}}(\rho_{\mathrm{a}}||\rho_{\mathrm{b}}) = \int_{-\infty}^{+\infty} \rho_{\mathrm{a}}(q) \ln\left(\frac{\rho_{\mathrm{a}}(q)}{\rho_{\mathrm{b}}(q)}\right)\mathrm{d}q.$$

This function is not symmetric with respect to swapping ρ_{a} and ρ_{b}. Hence, here we will use the symmetrized KL distance:

$$J_{\mathrm{KL}}(\rho_{\mathrm{a}},\rho_{\mathrm{b}}) = J_{\mathrm{KL}}(\rho_{\mathrm{b}},\rho_{\mathrm{a}}) = D_{\mathrm{KL}}(\rho_{\mathrm{a}}||\rho_{\mathrm{b}}) + D_{\mathrm{KL}}(\rho_{\mathrm{b}}||\rho_{\mathrm{a}}).$$

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[24]

$$J_{\mathrm{KL}}(\rho_{\mathrm{a}}, \rho_{\mathrm{b}}) = \frac{1}{2}\left((\mu_{\mathrm{a}} - \mu_{\mathrm{b}})^2 \cdot \left(\frac{1}{\sigma_{\mathrm{a}}^2} + \frac{1}{\sigma_{\mathrm{b}}^2}\right) + \frac{\sigma_{\mathrm{a}}^2}{\sigma_{\mathrm{b}}^2} + \frac{\sigma_{\mathrm{b}}^2}{\sigma_{\mathrm{a}}^2} - 2\right).$$

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 another distribution for a 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 relationship $\sigma_x = \hbar/(2\sigma_p)$, we can find a condition of $J_{\mathrm{KL}}(\rho_X, \rho_p) \rightarrow \min$ as $\frac{\mathrm{d}J_{\mathrm{KL}}}{\mathrm{d}\sigma_p} = 0$, 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 a single control parameter τ and due to the independence on equilibrium molecular geometry.

The sampling procedure is simple. First, we choose a reasonable τ. Next, for each coordinate (x, y, z) of all nuclei we generate a displacement and velocity according to the distribution given in eqn (9). The procedure can also be defined in terms of a multivariate quasi-distribution of all Cartesian coordinates and momentums, similar to that in eqn (8). In this case, the sampling is done on the basis of the distribution

$$W_0(\mathbf{P}, \mathbf{R}) \propto \exp\left(-\frac{1}{2}\mathbf{R}^\dagger \Sigma_{\mathbf{R},\text{SWS}}^{-1} \mathbf{R} - \frac{2}{\hbar^2}\mathbf{P}^\dagger \Sigma_{\mathbf{R},\text{SWS}} \mathbf{P}\right), \quad (11)$$

where $\Sigma_{\mathbf{R},\text{SWS}} = \frac{\hbar\tau}{2}\text{diag}\left(\ldots, \frac{1}{m_k}, \ldots\right)$ and k enumerates all the Cartesian coordinates. This representation has two main differences compared to eqn (8). The first one is that the diagonal elements of matrix $\Sigma_{\mathbf{R},\text{SWS}}$ in eqn (11) are given by the parametrization from eqn (10), rather than computed from the harmonic potential. The second difference is that all the non-diagonal elements in $\Sigma_{\mathbf{R},\text{SWS}}$, which carry information on the correlations between different Cartesian coordinates of atoms, are zeros. Both differences appear due to the absence of information about the shape of the local potential of nuclei in the SWS procedure, which makes it physically less justified in comparison to WS. Accordingly, SWS is expected to be less accurate, although it is not yet exactly clear how large the accuracy loss will be in different use scenarios. At the same time SWS gains an advantage, that it does not require a preliminary optimization of molecular structure and calculation of Hessian.

A reasonable value for τ can be determined 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 of in terms of temperature, since for MBS $\sigma_{p,\text{MB}}^2 = mk_\text{B}T$, i.e. the effective temperature is defined as $T_\text{eff} = \frac{\hbar}{2k_\text{B}\tau}$. For velocities of nuclei $T_\text{eff} \leq 5000$ K can be achieved at $\tau > 0.8$ fs. This provides 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 ($\delta p \sim \exp(-\delta p^2/(2mk_\text{B}T))$). By replacing p with $p + \delta p$ we need to average $W(x,p)$ over δp, which leads to

$$W_T(\xi, p) = \frac{1}{\sqrt{2\pi mk_\text{B}T}}\int_{-\infty}^{+\infty}\exp\left(-\frac{\delta p^2}{2mk_\text{B}T}\right) \cdot W(p + \delta p, \xi)\text{d}\delta p$$

$$= \frac{1}{2\pi\sigma_x\sigma_{p,T}}\underbrace{\exp\left(-\frac{\xi^2}{2\sigma_x^2}\right)}_{\propto\rho_x} \cdot \underbrace{\exp\left(-\frac{p^2}{2\sigma_{p,T}^2}\right)}_{\propto\rho_{p,T}},$$

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

3. Computational methods

Most of the quantum-chemical calculations have been performed using the density functional theory approximation PBE/def2-SVP[25,26] (in some cases paired with D3BJ[27] empirical dispersion corrections) implemented in the ORCA 5 package.[28] In cases where a better agreement with experimental data was required, the composite method PBEh-3c[29] has been utilized. Vibrational corrections to rotational constants within VPT2[30] theory were computed using the Gaussian 16 package.[31] Details on methods and settings are provided in the ESI.†

4. Results and discussion

4.1 A simple case of H_2^+

To illustrate the general idea of WS and SWS in application to molecular vibrations, we have performed 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 (eqn (5)).

A comparison of the numerical anharmonic vibrational wavefunction with that from the harmonic approximation is shown in Fig. 2. The asymmetry of the anharmonic potential with respect to the equilibrium geometry r_e leads to a shift of the probability distribution into the range of larger interatomic distances r. 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.[32,33] 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.[11,13,34] This substitution is justified by 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 (eqn (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 performed a series of simulations (for details see the ESI†). Fig. 3 shows the results for the first two moments of the interatomic distribution, the averaged distance $\langle r\rangle$ and the amplitude l (eqn (2)).
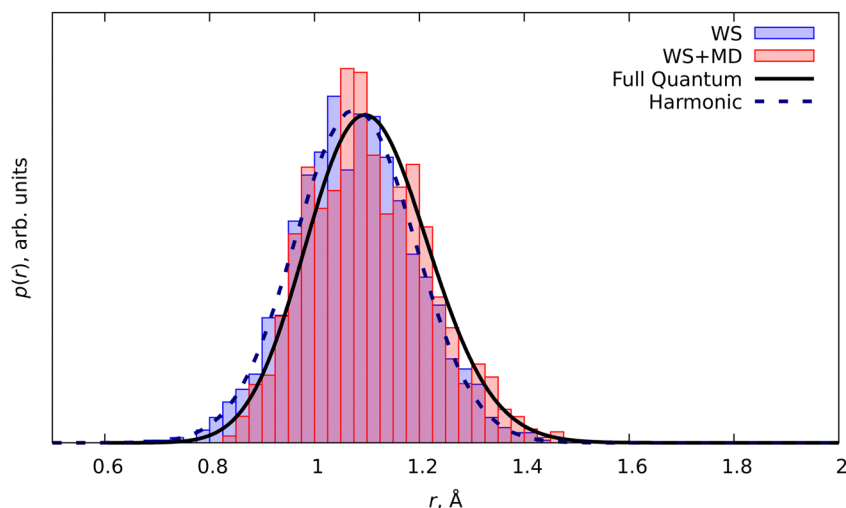
18410 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 18406–18423

This journal is © the Owner Societies 2023

Fig. 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).

The harmonic oscillator model and the pure WS result in $\langle r \rangle_h = r_e$ because the harmonic potential is symmetrical. However, MD simulations with initial conditions from WS (WS + MD) produce an average value $\langle r \rangle_{WS + MD}$ very close to that computed with the numerical ground state anharmonic wavefunction $\langle r \rangle_{anh}$. In the case of the diatomic molecule, the SWS distribution can be exactly mapped to WS. Therefore we can get the same solution $\langle r \rangle_{SWS+MD} = \langle r \rangle_{WS+MD}$ by scanning the parameter τ. The resulting curve has a single minimum at $\tau = 2$ fs where $\langle r \rangle_{SWS+MD} = \langle r \rangle_{WS+MD} = \langle r \rangle_{anh}$. Fig. 3 also shows the results from MBS + MD simulations at various temperatures, where the temperature T is mapped to τ through the Debye temperature relationship $\tau = \nu^{-1} = h/(k_B T)$. 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 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.[11,34,35]

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 the optimal parameter τ for larger molecules. The obvious choice is an energy-related parameter, for example, the average total energy (electronic and nuclear) $\langle E_{tot} \rangle$. The other option is the kinetic energy of nuclei ($\langle E_{kin} \rangle$), which is often expressed in terms of temperature through the relationship $\langle E_{kin} \rangle = N_f k_B T/2$, where N_f is the number of degrees of freedom. For example, $N_f = 1$ in the case of the diatomic molecule H_2^+. The results in Fig. 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 Fig. 3 shows. The total vibrational energy $E_{vib} = \langle E_{tot} \rangle - E_e$ 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,[36] which 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 Fig. 4.

4.2 Rotational constants

The accuracy of calculated vibrationally-averaged (usually B_0) rotational constants is mostly determined by two factors, (a) equilibrium structure and (b) vibrational corrections $B_e - B_0$. The latter is relevant for the context of the present investigation. Thus, in the following we do not aim at reproducing the absolute values of rotational constants but benchmark only corrections $B_e - B_0$. For this, five small molecules have been chosen (see Fig. 1): water (H_2O), ethylene (C_2H_4), difluoromethane (CH_2F_2), and trans and cis isomers of 1,2-difluoroethylene ($C_2H_2F_2$). For each molecule the corrections $B_e - B_0$ were calculated by using Wigner sampling (in WS + MD and SWS + MD variants) and by VPT2. In our case with small semi-rigid molecules the values from VPT2 provide accurate reference values.[37,38]

Three of the test molecules (H_2O, C_2H_4, and CH_2F_2) were used to verify the position of optimal τ. Fig. 5 illustrates the effect of τ in the SWS procedure applied to these molecules. The minimal vibrational energies were in the range of 2–4 fs, very close to the optimal $\tau = 2$ fs determined for H_2^+. Thus, we used this value further in our simulations. However, in general we recommend to tune τ for obtaining the most accurate results.

Computing averaged rotational constants from MD trajectories is not straightforward. In this paper we tested the following three different procedures.

• The simplest approach is the calculation of rotational constants B_α ($\alpha = a, b, c$) at every time step with the following

This journal is © the Owner Societies 2023

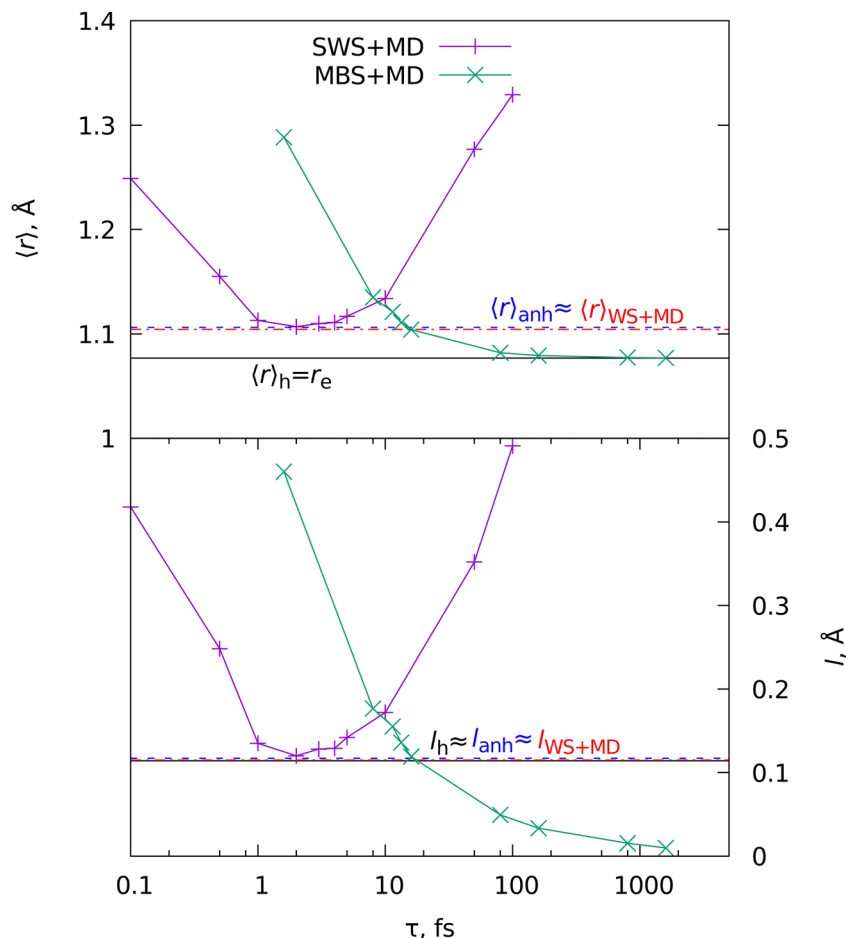*Phys. Chem. Chem. Phys.*, 2023, **25**, 18406–18423 | **18411**

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

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 is calculated at each MD step. The inverted tensor $^{-1}$ 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 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.[39] This

way of computing the mean rotational constants is denoted as mean inverse inertia tensor oriented (MIITO).

We can consider each MD trajectory 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 such MD trajectories started from different initial conditions. Therefore, the final values are the 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 the 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 $\langle\langle B_a \rangle^2\rangle = \frac{1}{N_{\mathrm{trj}}}\sum_{n=1}^{N_{\mathrm{trj}}} \langle B_a \rangle_n^2$.

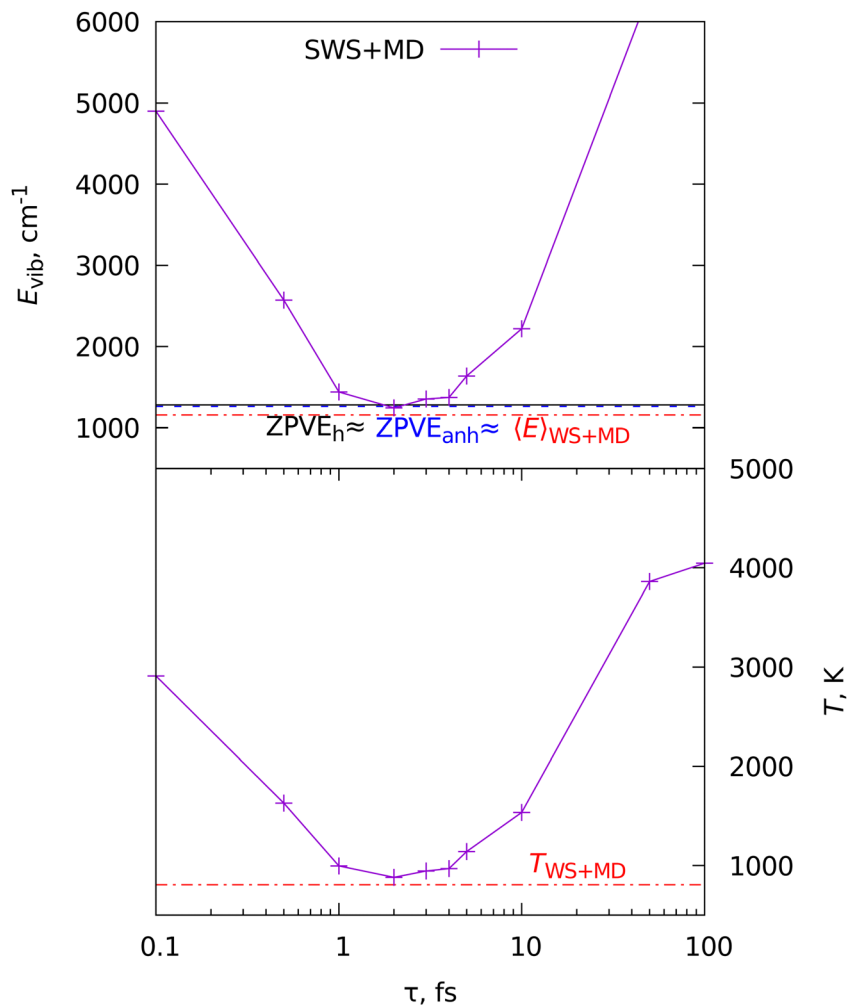**Fig. 4** Vibrational energies E_{vib} (above) and vibrational nuclear temperatures $T = 2 \langle E_{kin} \rangle / (N_f k_B)$ (below). Horizontal lines show the energy levels calculated in the harmonic (index "h", black line), anharmonic ("anh", blue line), and WS + MD (red line) approximations. The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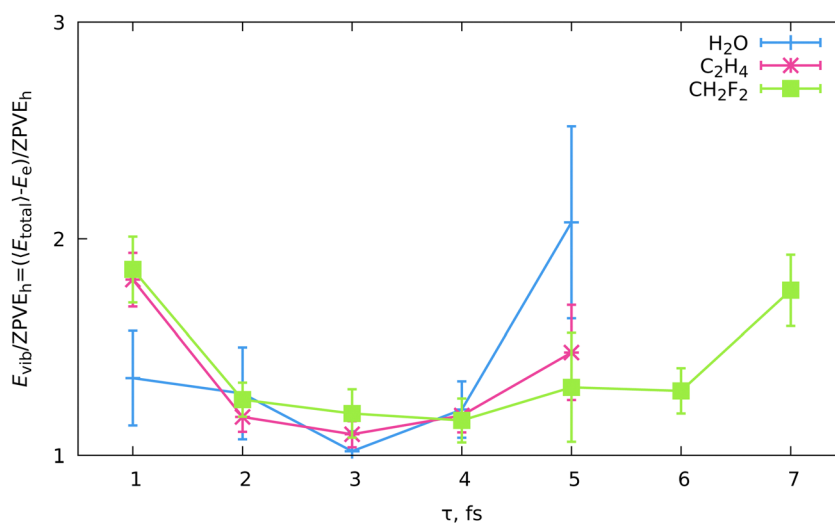

Fig. 5 The ratio of the vibrational energy E_{vib} to the harmonic zero-point vibrational energy (ZPVE$_h$) in SWS + MD simulations of H_2O, C_2H_4 and CH_2F_2. Error bars demonstrate standard errors for 30 independent simulation runs.

This journal is © the Owner Societies 2023

Phys. Chem. Chem. Phys., 2023, **25**, 18406–18423 | **18413**

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

	MRC	MIITU	MIITO
ρ(WS + MD,VPT2)	0.77	0.76	0.80
ρ(SWS + MD,VPT2)	0.83	0.89	0.90

Using these three procedures, we have calculated the vibrational shifts of rotational constants $B_{e,\alpha} - \langle B_\alpha \rangle$, which were compared with those from the standard VPT2 computations. Since the shifts are sensitive to the absolute values of rotational constants, in statistical assessments (Table 1 and Fig. 6) we used normalized quantities $(B_{e,\alpha} - \langle B \rangle)/B_{e,\alpha} = 1 - \langle B \rangle/B_{e,\alpha}$. 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 $\hat{H}_{rot} = \langle 0|\frac{1}{2}\hat{\mathbf{L}}^\dagger \mathcal{I}^{-1}\hat{\mathbf{L}}|0\rangle = \frac{1}{2}\hat{\mathbf{L}}^\dagger \langle 0|\mathcal{I}^{-1}|0\rangle\hat{\mathbf{L}}$, where $|0\rangle$ is the ground vibrational state, and $\hat{\mathbf{L}}$ is 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 the 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 is negligible in comparison to the 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.[14]

Two additional tests of the proposed computational methods were done using the moderate-size molecular systems (see Fig. 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 previously investigated experimentally[40,41] by gas electron diffraction (GED). For the latter molecule GED data were supplemented by rotational constants from microwave measurements. Results of our calculations for these two molecules are given in Table 2. The $(B_e - B_0)$ values obtained by WS + MD simulations are in reasonable agreement with those from VPT2. The relative deviations are about of the same order of magnitude as for the smaller systems discussed above. Thus, WS + MD can probably be applied to larger and more complicated molecules.

4.3 Vibrational spectra

Another possible application of MD is the prediction of vibrational spectra. Here we discuss only infrared (IR) one-photon spectra. Their intensity can be computed as[42]

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

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.[42,43] The duration (t_{tot}) and the time step (Δt) of ACF are the same as for the original MD trajectory. In discrete Fourier transform,
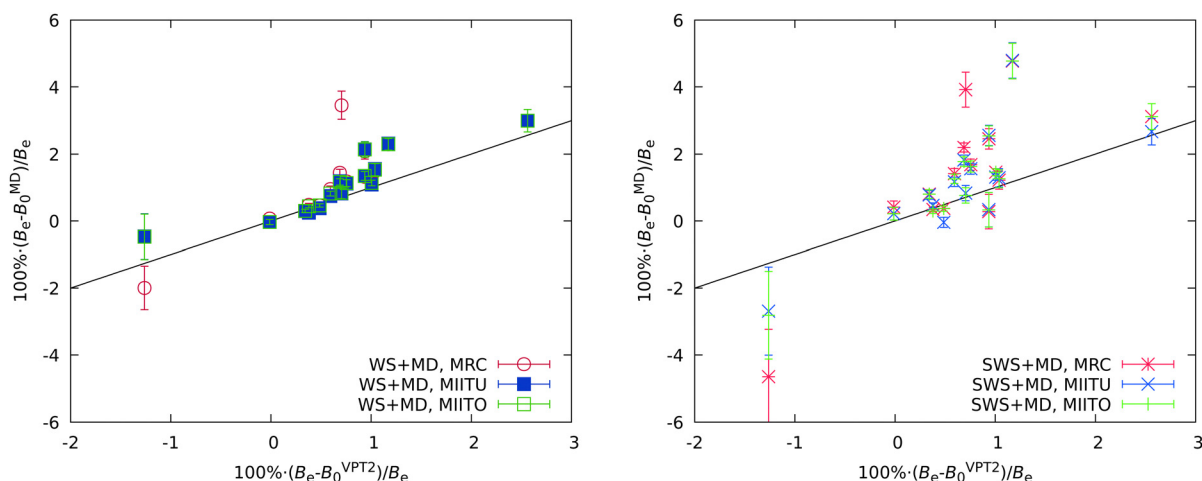
Fig. 6 Corrections to rotational constants $B_e - B_0$ calculated using different methods in comparison to reference values computed from cubic force fields. A solid line indicates the diagonal.

18414 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 18406–18423

This journal is © the Owner Societies 2023

Table 2 Equilibrium rotational constants B_e and their anharmonic shifts in the ground vibrational state $B_e - B_0$ for the seven model asymmetric top molecules[a]

| Molecule	$A/B/C$	B_e	$B_e - B_0$			
			VPT2	VPT2 + cent.	WS + MD	SWS + MD(τ = 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
$C_6H_{12}N_2$	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 quantum-chemical approximation used for the 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. The uncertainties of the MD values are the ± standard errors.

the frequency resolution is determined by the trajectory duration as $\Delta\nu \propto t_{tot}^{-1}$, while the time step determines the maximal frequency as $\nu_{max} \propto \Delta t^{-1}$. 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.[44] However, such shifts can be corrected using a simple replacement of Fourier-transform frequencies ν_{FT} with[45]

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

We also applied this correction throughout the work.
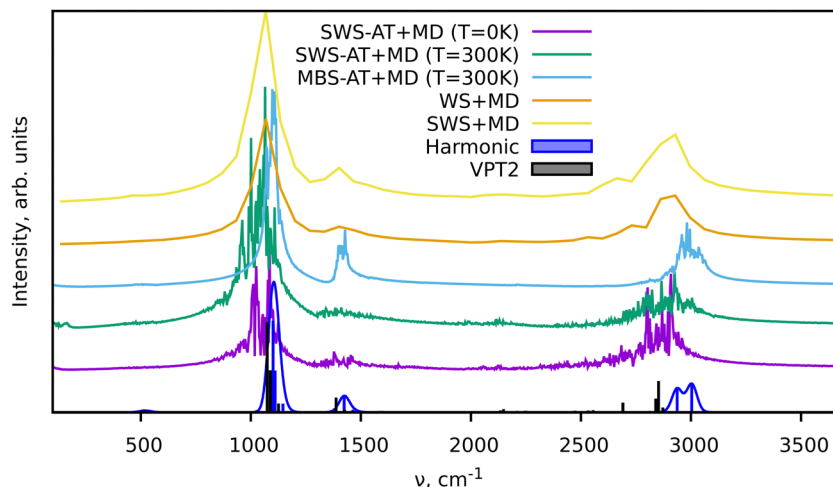
Fig. 7 Vibrational infrared (IR) spectra of CH_2F_2. The bars indicate the harmonic and VPT2 anharmonic spectra. The line around the harmonic spectra is the Gaussian-convoluted spectra with each fundamental peak broadened with a FWHM of 10 cm^{-1}. The lines above show the spectra obtained from the MD simulations.

This journal is © the Owner Societies 2023

Phys. Chem. Chem. Phys., 2023, **25**, 18406–18423 | **18415**

It is worth mentioning the 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 $N = N_t \cdot N_s$. For the predefined constant N we may balance N_t and N_s but cannot increase both parameters. Smaller N_s causes lowering of the spectral resolution, because of the duration of trajectories $t = N_s \cdot \Delta t$, 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.[13]

The Andersen thermostat is the most straightforward to implement.[46] The WS method relies on equilibrium molecular geometry and cannot be used for anything except for sampling of the initial conditions. SWS, on the other hand, has the same principle as MBS and thus can be used with the Andersen thermostat. Test calculations have been performed for difluoromethane CH_2F_2 at the PBE/def2-SVP level of theory using a combination of SWS and Andersen thermostat (abbreviated here as SWS-AT + MD). We also carried out classical constant temperature MD simulations with the Andersen thermostat applying the MBS scheme (MBS-AT + MD). Fig. 7 shows the obtained vibrational spectra in comparison to harmonic and VPT2 fundamentals. Also, for demonstrational purposes, we calculated and plotted spectra from the WS + MD and SWS + MD simulations, which were used in computation of rotational constants described above. 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 duration of trajectories the spectral resolution in SWS-AT + MD exceeds by far that 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.

4.4 Case studies

To further demonstrate the applicability of WS + MD and SWS + MD for the calculation of vibrational properties 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. We have chosen benzene (C_6H_6), protonated methane cation (CH_5^+) and protonated serine octamer (Ser_8H^+, $(C_3H_7NO_3)_8H^+$), see Fig. 1. For all of them there have been

published high quality experimental and theoretical data.[47–54] Benzene is an example of a simple and semi-rigid system, for which are available experimental rotational constants with respective theoretical anharmonic corrections,[47,49] vibrational[48] and photoelectron (PE) spectra.[50] 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.

4.4.1 Benzene. The semi-empirical equilibrium molecular structure of benzene has been determined[49] previously from experimental rotational constants and calculated vibrational corrections at the coupled cluster level of theory. In this work we have tested several types of simulations (see the ESI† for details) at the PBEh-3c level of theory.[29] From WS + MD simulations we have computed the vibrational corrections to rotational constants to be $B_e − B_0 = 0.192 \pm 0.01$ m^{-1} and
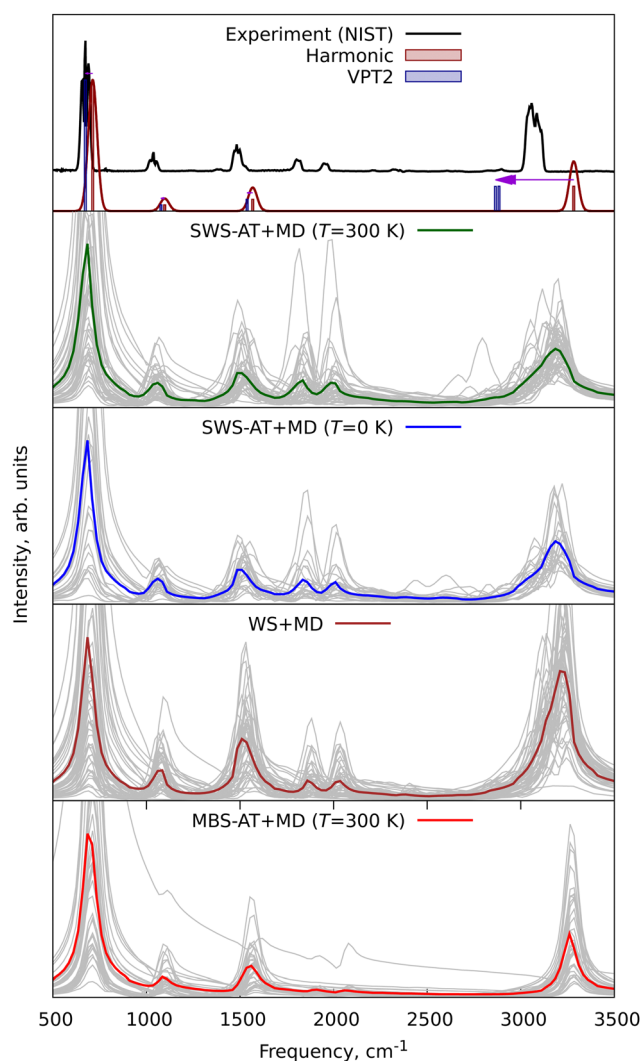

Fig. 8 Experimental and theoretical infrared (IR) vibrational spectra of benzene. Grey lines are individual spectra from each single trajectory, and colored lines are averaged spectra. The horizontal arrow indicates the anharmonic VPT2 shift. The line around the harmonic spectra has been obtained by Gaussian convolution with a FWHM of 10 cm^{-1}.

18416 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 18406–18423

This journal is © the Owner Societies 2023

$C_e - C_0 = 0.083 \pm 0.004$ 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 experiments. For the C-constant our vibrational correction is in reasonable agreement with the reference theoretical value of 0.14 m^{-1} obtained at the very high level.[49]

Furthermore, the experimental vibrational IR spectra of benzene were taken from NIST Chemistry WebBook[48] and compared with those calculated in this work, see Fig. 8. The averaged IR spectra from SWS-AT + MD and WS + MD reproduce all experimental bands. MBS-AT + MD underestimates the intensity for the overtones within the 1700–2000 cm^{-1} range (not shown in the harmonic and VPT2 plots). It is interesting to discuss the C–H stretching region, corresponding to vibrations with the highest Debye temperature in benzene. In MBS-AT + MD the position of the corresponding peak at 3260 cm^{-1} is essentially the same as in the harmonic approximation (3262 cm^{-1}). This is explained by underestimation of vibrational amplitudes in the classical MD treatment (see the discussion of eqn (4)). WS + MD results in a larger anharmonic shift with the peak maximum at 3223 cm^{-1}. In SWS-AT + MD at both temperatures (0 and 300 K) this band is at 3185 cm^{-1}, the closest to the experimental value at about 3058 cm^{-1}. Notably, the MD models qualitatively correctly predict frequencies in the C–H stretching region around 3000 cm^{-1}, whereas in the VPT2 calculation we obtained an unrealistically large anharmonic shift of 409 cm^{-1} for one of the modes. This VPT2 artifact probably arises from resonances, which have been discussed previously.[55,56] Overall, WS and SWS demonstrate a better agreement with experimental data in comparison to other tested methods.

We also computed photoelectron spectra (PES) of benzene and compared them with the available experimental data.[50] Several 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. As a preliminary step, structure optimizations and harmonic

frequency calculations were carried out for C_6H_6 and $C_6H_6{}^+$. However, $C_6H_6{}^+$ is a well-known example[57,58] 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 obtained using this procedure. The WS procedure with harmonic modes has been used to account for vibrational motions. With this simplistic treatment of vibrational degrees of freedom in the harmonic approximation we computed an approximate PE spectrum, see Fig. 9. For the anharmonic treatment, trajectories from SWS-AT + MD modeling were used (see the ESI† for details). A conventional Franck-Condon computation would allow direct evaluation of 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 and SWS can be used as the fast and simple methods for estimation of line shapes, as we demonstrate in Fig. 9. 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 due to anharmonicity.

4.4.2 Protonated methane cation. Protonated methane ($CH_5{}^+$) is a strongly nonrigid molecular system. Its vibrational spectra[51,52] and molecular structure and dynamics[53,59–61] were extensively investigated. The availability of experimental data makes this highly flexible system a good test target for assessing theoretical vibrational spectra. We have performed a series of simulations in different variants (WS, SWS and MBS) similar to those for benzene, as described above. Technical details are provided in the ESI.† Experimental vibrational spectra of $CH_5{}^+$ were measured earlier using action spectroscopy,[51] monitoring the yield of the OCOH$^+$ cation according to the reaction

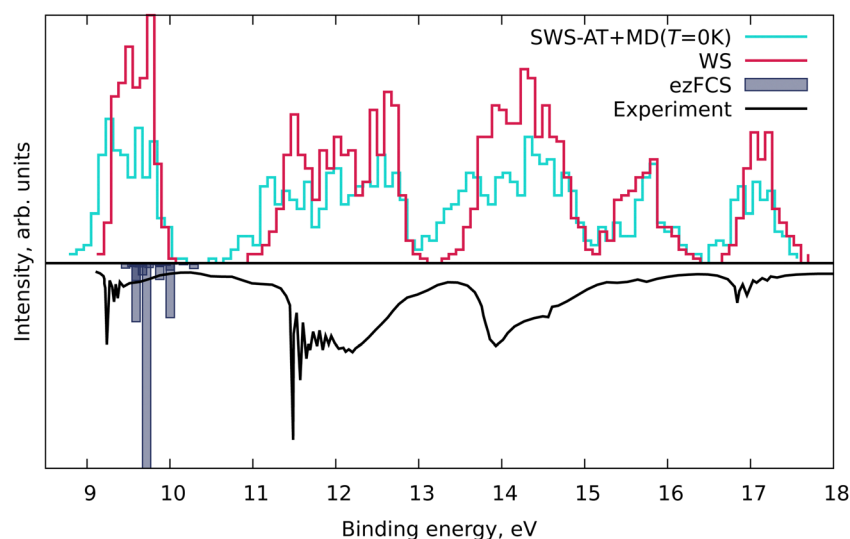$$CH_5{}^+ + CO_2 + h\nu \rightarrow CH_4 + OCOH^+, \qquad (12)$$

Fig. 9 Experimental and modeled (see the text and ESI† for details) photoelectron spectra of benzene. ezFCS bars indicate results of the computation of Franck–Condon factors. The WS and SWS-AT + MD lines are manually shifted by 0.5 eV for a better match with the experimental peaks.
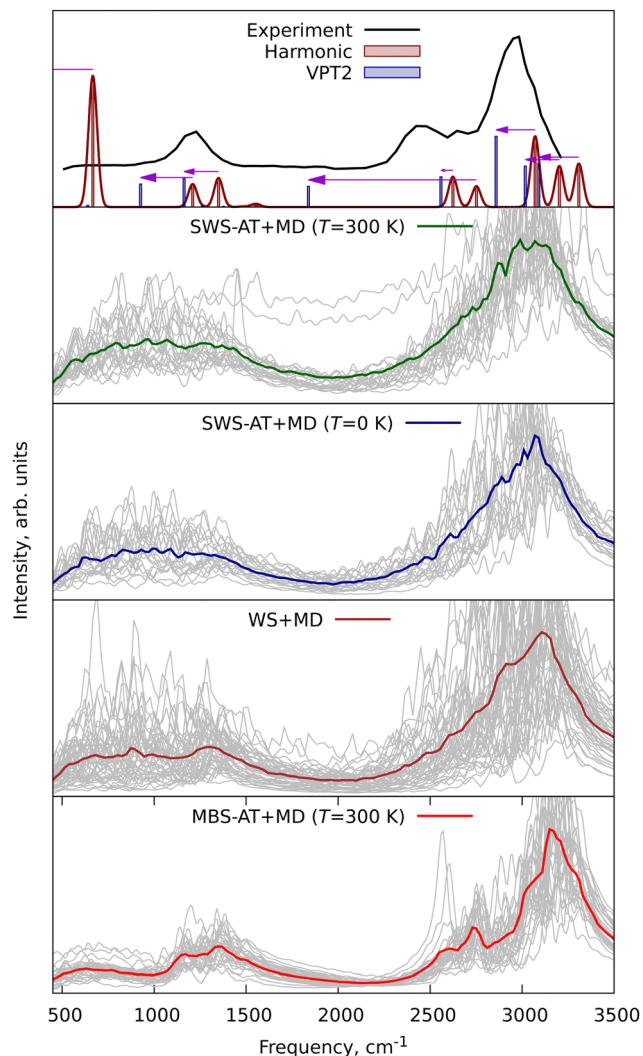
This journal is © the Owner Societies 2023

Phys. Chem. Chem. Phys., 2023, **25**, 18406–18423 | **18417**

Fig. 10 Experimental[51] and modeled IR spectra of CH_5^+. Arrows indicate anharmonic VPT2 shifts. Broadening of the theoretical harmonic lines has been done by Gaussian convolution with a FWHM of 10 cm^{-1}.

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

$$f(\nu) = \begin{cases} 0, & h\nu \leq D, \\ 1 - \dfrac{D}{h\nu}, & h\nu > D, \end{cases} \qquad (13)$$

where ν is the photon frequency, and the D is the energy threshold for the observed reaction. This procedure allows us to take into account processes occurring in the measurements and practically damps the low-frequency wide band below 1000 cm^{-1}, which is predicted by theory but could not be observed experimentally (see ref. 51 for a discussion). In our work we assumed $D = 383$ cm^{-1}, which corresponds to the Gibbs free energy of the probing reaction (eqn (12)). A comparison of experimental and all spectra modeled in this work is

shown in Fig. 10. Expectedly, VPT2 fails by producing a negative value for the lowest fundamental frequency, due to large-amplitude motions. MBS-AT + MD shows the fine band structure, whereas the WS-based methods demonstrate the broadening of the peaks without significant anharmonic shifts. For comparison, a path-integral MD simulation of CH_5^{+62} did not show significant broadening. This may indicate a too aggressive sampling in WS and SWS for such a flexible molecule. Thus, the classical MD appears to be more effective for reproducing the vibrational band structure in this case.

4.4.3 Protonated serine octamer. The protonated serine octamer (Ser$_8$H$^+$, $C_{24}H_{57}N_8O_{24}^+$, see Fig. 1) is a relatively large molecular system comprising 113 atoms. Its structure was investigated using helium-tagging IR spectroscopy by comparing computed harmonic frequencies with experimental spectra.[54] An application of the VPT2 method for this system is computationally not feasible. We have done a simple benchmark for computational timings (details are in the ESI†) and determined that in this case a VPT2 calculation is equivalent to a MD simulation with about $6-12 \times 10^3$ frames. In addition, 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$_8$H$^+$ molecular ensemble can be reached even with fewer than 6,000 frames.

To demonstrate that, we have performed two simulations, MBS-AT + MD at $T = 300$ K and SWS-AT + MD at $T = 0$ K. Since experimental spectra were obtained using a helium tag,[54] we applied the intensity correction factor given in eqn (13) with the calculated helium detachment energy $D = 379$ cm^{-1}. Fig. 11 shows comparison of the experimental and calculated spectra.

It is generally difficult to achieve a very good agreement between theoretical and experimental spectra for a large and complicated system such as Ser$_8$H$^+$. The spectra obtained here qualitatively describe intensity distribution within the ranges 1000–1800 cm^{-1} and 2900–3500 cm^{-1}. There are, however, some particular deviations worth mentioning. In the experiments there were no significant signals around 2000 cm^{-1}, whereas harmonic and SWS calculations predict some bands in this region. Only the SWS spectrum can reproduce large intensities (although with somewhat shifted frequencies) of the sharp lines at about 3350 cm^{-1} in the experiment, which were assigned to N–H and O–H stretchings. However, neither of our simulations showed an intense signal for another O–H stretching vibration at 3700 cm^{-1}. Note, the same problem of very low intensity was observed in calculations performed in the original work.[54] Overall, we demonstrate here the principal applicability of SWS for simulation of the vibrational spectra of large systems. At that, the accuracy in some cases may be higher than in harmonic approximation.

5. Conclusions

Wigner sampling (WS) combined with classical molecular dynamics is a powerful approach for describing nuclear quantum effects and vibrational anharmonicity at a reasonable
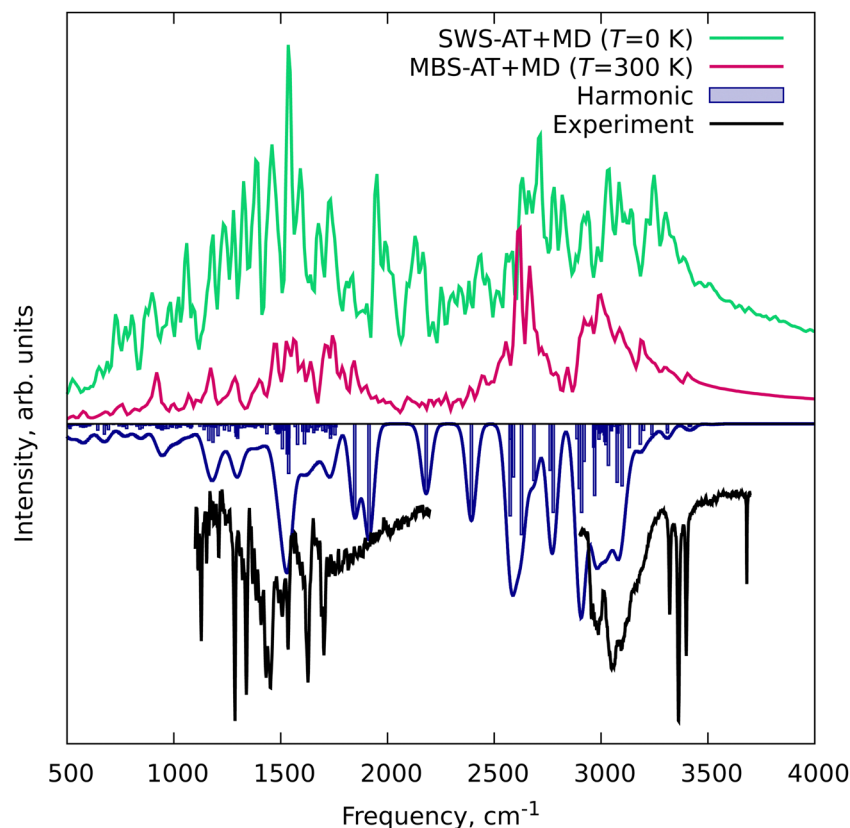
Fig. 11 Experimental (black lines below) and theoretical IR vibrational spectra of Ser_8H^+. The two independent sections of the experimental data, 1100–2200 cm^{-1} and 2900–3700 cm^{-1}, were independently scaled for a better view. Vertical bars (below) show the harmonic fundamentals. The calculated harmonic spectrum has been obtained by Gaussian convolution of harmonic frequencies with FWHM of 10 cm^{-1}.

computational cost. In this work we briefly describe the underlying theory and introduce a new procedure, denoted as simplified Wigner sampling (SWS). In contrast to WS, this procedure does not require initial calculation of Hessian. Instead, it can be controlled by a single parameter τ, whose optimal values can be found in the interval $1 \leq \tau \leq 10$ fs.

To test the sampling methods a series of calculations were performed for several molecular systems. Among the modeled properties were (a) vibrational IR spectra, (b) photoelectron spectra and (c) vibrationally averaged rotational constants. For the latter, three computational procedures have been introduced and tested. The accuracy of WS and SWS has been assessed by comparing with experimental data (where available) and with results of harmonic and VPT2 anharmonic calculations. The new SWS procedure showed a promising performance taking into account its computational simplicity. Its particular advantage over the VPT2 method has been observed for the large system of the protonated serine octamer and in the cases with large amplitude motions.

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)
AT	Andersen thermostat‡
FWHM	Full width at half maximum
IR	Infrared (spectrum)
PE	Photoelectron (spectrum)
IP	Ionization potential

Author contributions

Denis S. Tikhonov: project administration, conceptualization, data curation, investigation, software, visualization, writing –

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

This journal is © the Owner Societies 2023

Phys. Chem. Chem. Phys., 2023, **25**, 18406–18423 | **18419**

original draft. Yury V. Vishnevskiy: conceptualization, data curation, writing – review & editing.

Conflicts of interest

There are no conflicts to declare.

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 bound 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

Fig. 12 Scheme of the heuristic derivation of the damping factor in the action-type spectroscopy. See the text for details.

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 the 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 Fig. 12) 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 $N_{diss} = 0$. Otherwise, when $h\nu > D$, the number of the dissociated molecules is calculated as $N_{diss} = (h\nu - D) \cdot \cos(45°) = \frac{\sqrt{2}(h\nu - D)}{2}$ (see Fig. 12). Thus we can find the quantum yield of the dissociation per absorbed photon with frequency ν as

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

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.

The general form of the damping function $f(\nu)$ from eqn (13) is given in Fig. 13. Application of this function effectively zeroes signals with energies below the threshold ($h\nu < D$). Signals above the threshold are scaled by factors, increasing with the increase of the photon energy $h\nu$.
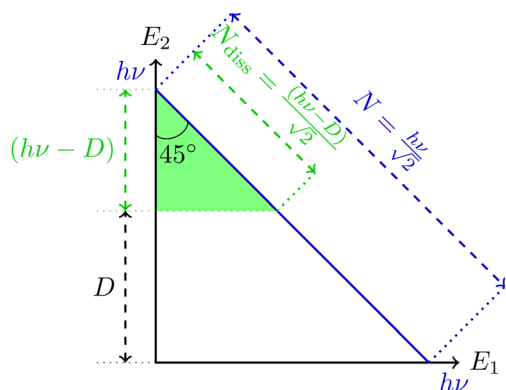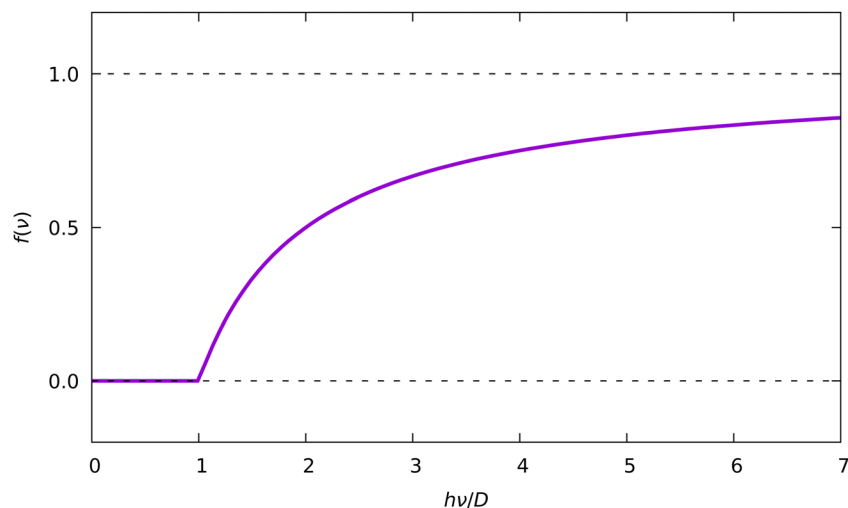

Fig. 13 Damping function $f(\nu)$ (eqn (13)).

18420 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 18406–18423

This journal is © the Owner Societies 2023

Acknowledgements

References

1 T. E. Markland and M. Ceriotti, Nuclear quantum effects enter the mainstream, *Nat. Rev. Chem.*, 2018, **20**(3), 0109, DOI: **10.1038/s41570-017-0109**.

2 S. A. Hollingsworth and R. O. Dror, Molecular dynamics simulation for all, *Neuron*, 2018, **990**(6), 1129–1143, DOI: **10.1016/j.neuron.2018.08.011**.

3 R. Iftimie, P. Minary and M. E. Tuckerman, *Ab initio* molecular dynamics: Concepts, recent developments, and future trends, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **1020**(19), 6654–6659, DOI: **10.1073/pnas.0500193102**.

4 D. Marx and M. Parrinello, Ab initio path integral molecular dynamics: Basic ideas, *J. Chem. Phys.*, 1996, **1040**(11), 4077–4082, DOI: **10.1063/1.471221**.

5 S. C. Althorpe, Path-integral approximations to quantum dynamics, *Eur. Phys. J. B*, 2021, **940**(7), 155, DOI: **10.1140/epjb/s10051-021-00155-2**.

6 P. W. Atkins and R. S. Friedman, *Molecular Quantum Mechanics*, OUP, Oxford, 2011.

7 P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs and A. 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, *J. Am. Chem. Soc.*, 1983, **1050**(24), 7037–7047, DOI: **10.1021/ja00362a005**.

8 M. K. Kesharwani, B. Brauer and J. 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?, *J. Phys. Chem. A*, 2015, **1190**(9), 1701–1714, DOI: **10.1021/jp508422u**.

9 R. A. Mata and M. A. Suhm, Benchmarking quantum chemical methods: Are we heading in the right direction?, *Angew. Chem., Int. Ed.*, 2017, **560**(37), 11011–11018, DOI: **10.1002/anie.201611308**.

10 L. D. Landau and E. M. Lifshitz, *Mechanics: Volume 1 (Course of Theoretical Physics)*, Butterworth-Heinemann, 3rd edn, 1976.

11 Y. V. Vishnevskiy and D. Tikhonov, Quantum corrections to parameters of interatomic distance distributions in molecular dynamics simulations, *Theor. Chem. Acc.*, 2016, **1350**(4), 88, DOI: **10.1007/s00214-016-1848-2**.

12 C. Qu and J. M. Bowman, Quantum approaches to vibrational dynamics and spectroscopy: is ease of interpretation sacrificed as rigor increases?, *Phys. Chem. Chem. Phys.*, 2019, **21**, 3397–3413, DOI: **10.1039/C8CP04990D**.

13 D. S. Tikhonov, D. I. Sharapa, J. Schwabedissen and V. V. Rybkin, Application of classical simulations for the computation of vibrational properties of free molecules, *Phys. Chem. Chem. Phys.*, 2016, **18**, 28325–28338, DOI: **10.1039/C6CP05849C**.

14 V. Barone, Anharmonic vibrational properties by a fully automated second-order perturbative approach, *J. Chem. Phys.*, 2005, **1220**(1), 014108, DOI: **10.1063/1.1824881**.

15 Q. Yang, M. Mendolicchio, V. Barone and J. 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*, 2021, **8**, DOI: **10.3389/fspas.2021.665232**.

16 E. Wigner, On the quantum correction for thermodynamic equilibrium, *Phys. Rev.*, 1932, **40**, 749–759, DOI: **10.1103/PhysRev.40.749**.

17 L. Sun and W. L. Hase, Comparisons of classical and wigner sampling of transition state energy levels for quasiclassical trajectory chemical dynamics simulations, *J. Chem. Phys.*, 2010, **1330**(4), 044313, DOI: **10.1063/1.3463717**.

18 J. P. Zobel, J. J. Nogueira and L. González, Finite-temperature wigner phase-space sampling and temperature effects on the excited-state dynamics of 2-nitronaphthalene, *Phys. Chem. Chem. Phys.*, 2019, **21**, 13906–13915, DOI: **10.1039/C8CP03273D**.

19 J. P. Zobel, M. Heindl, J. J. Nogueira and L. González, Vibrational sampling and solvent effects on the electronic structure of the absorption spectrum of 2-nitronaphthalene, *J. Chem. Theory Comput.*, 2018, **140**(6), 3205–3217, DOI: **10.1021/acs.jctc.8b00198**.

20 D. Avagliano, E. Lorini and L. González, Sampling effects in quantum mechanical/molecular mechanics trajectory surface hopping non-adiabatic dynamics, *Philos. Trans. R. Soc., A*, 2022, **3800**(2223), 20200381, DOI: **10.1098/rsta.2020.0381**.

21 Jmol development team, Jmol: an open-source java viewer for chemical structures in 3d, **https://www.jmol.org/. URL http://jmol.sourceforge.net/**.

22 M. Barbatti and K. Sen, Effects of different initial condition samplings on photodynamics and spectrum of pyrrole, *Int. J. Quantum Chem.*, 2016, **1160**(10), 762–771, DOI: **10.1002/qua.25049**.

23 S. Kullback and R. A. Leibler, On Information and Sufficiency, *Ann. Math. Stat.*, 1951, **220**(1), 79–86, DOI: **10.1214/aoms/1177729694**.

24 D. S. Tikhonov, Y. V. Vishnevskiy, A. N. Rykov, O. E. Grikina and L. S. Khaikin, Semi-experimental equilibrium structure of pyrazinamide from gas-phase electron diffraction. how much experimental is it?, *J. Mol. Struct.*, 2017, **1132**, 20–27, DOI: **10.1016/j.molstruc.2016.05.090**.

25 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple [phys. rev. lett. 77, 3865 (1996)], *Phys. Rev. Lett.*, 1997, **78**, 1396, DOI: **10.1103/PhysRevLett.78.1396**.

26 F. Weigend and R. 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.*, 2005, **7**, 3297–3305, DOI: **10.1039/B508541A**.

27 S. Grimme, S. Ehrlich and L. Goerigk, Effect of the damping function in dispersion corrected density functional theory, *J. Comput. Chem.*, 2011, **320**(7), 1456–1465, DOI: **10.1002/jcc.21759**.

28 F. Neese, F. Wennmohs, U. Becker and C. Riplinger, The orca quantum chemistry program package, *J. Chem. Phys.*, 2020, **1520**(22), 224108, DOI: **10.1063/5.0004608**.

29 S. Grimme, J. G. Brandenburg, C. Bannwarth and A. Hansen, Consistent structures and interactions by density functional theory with small atomic orbital basis sets, *J. Chem. Phys.*, 2015, **1430**(5), 054107, DOI: **10.1063/1.4927476**.

30 V. Barone, Anharmonic vibrational properties by a fully automated second-order perturbative approach, *J. Chem. Phys.*, 2005, **1220**(1), 014108, DOI: **10.1063/1.1824881**.

31 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*, Gaussian Inc., Wallingford CT, 2016.

32 M. Ceriotti, G. Bussi and M. Parrinello, Nuclear quantum effects in solids using a colored-noise thermostat, *Phys. Rev. Lett.*, 2009, **103**, 030603, DOI: **10.1103/PhysRevLett.103.030603**.

33 H. Dammak, Y. Chalopin, M. Laroche, M. Hayoun and J.-J. Greffet, Quantum thermal bath for molecular dynamics simulation, *Phys. Rev. Lett.*, 2009, **103**, 190601, DOI: **10.1103/PhysRevLett.103.190601**.

34 V. A. Levashov, S. J. L. Billinge and M. F. Thorpe, Quantum correction to the pair distribution function, *J. Comput. Chem.*, 2007, **280**(11), 1865–1882, DOI: **10.1002/jcc.20713**.

35 V. A. Sipachev., Diffraction measurements and equilibrium parameters, *Adv. Phys. Chem.*, 2011, 1–14.

36 P. Atkins and J. Paula, *Atkins' physical chemistry*, Oxford University press, 2008.

37 N. Vogt, J. Vogt and J. Demaison, Accuracy of the rotational constants, *J. Mol. Struct.*, 2011, **9880**(1), 119–127, DOI: **10.1016/j.molstruc.2010.12.040**.

38 S. Alessandrini, J. Gauss and C. Puzzarini, Accuracy of rotational parameters predicted by high-level quantum-chemical calculations: Case study of sulfur-containing molecules of astrochemical interest, *J. Chem. Theory Comput.*, 2018, **140**(10), 5360–5371, DOI: **10.1021/acs.jctc.8b00695**.

39 W. Kabsch, A solution for the best rotation to relate two sets of vectors, *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.*, 1976, **320**(5), 922–923, DOI: **10.1107/S0567739476001873**.

40 Y. V. Vishnevskiy, J. Schwabedissen, A. N. Rykov, V. V. Kuznetsov and N. 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, *J. Phys. Chem. A*, 2015, **1190**(44), 10871–10881, DOI: **10.1021/acs.jpca.5b08228**.

41 A. A. Fokin, T. S. Zhuk, S. Blomeyer, C. Pérez, L. V. Chernish, A. E. Pashenko, J. Antony, Y. V. Vishnevskiy, R. J. F. Berger, S. Grimme, C. Logemann, M. Schnell, N. W. Mitzel and P. R. Schreiner, Intramolecular london dispersion interaction effects on gas-phase and solid-state structures of diamondoid dimers, *J. Am. Chem. Soc.*, 2017, **1390**(46), 16696–16707, DOI: **10.1021/jacs.7b07884**.

42 M. Thomas, M. Brehm, R. Fligg, P. Vöhringer and B. Kirchner, Computing vibrational spectra from ab initio molecular dynamics, *Phys. Chem. Chem. Phys.*, 2013, **15**, 6608–6622, DOI: **10.1039/C3CP44302G**.

43 E. Ditler and S. Luber, Vibrational spectroscopy by means of first-principles molecular dynamics simulations, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**(5), e1605, DOI: **10.1002/wcms.1605**.

44 J. Horníček, P. Kaprálová and P. Bouř, Simulations of vibrational spectra from classical trajectories: Calibration with ab initio force fields, *J. Chem. Phys.*, 2007, **1270**(8), 084502, DOI: **10.1063/1.2756837**.

45 D. S. Tikhonov, Simple posterior frequency correction for vibrational spectra from molecular dynamics, *J. Chem. Phys.*, 2016, **1440**(17), 174108, DOI: **10.1063/1.4948320**.

46 H. C. Andersen, Molecular dynamics simulations at constant pressure and/or temperature, *J. Chem. Phys.*, 1980, **720**(4), 2384–2393, DOI: **10.1063/1.439486**.

47 I. Heo, J. C. Lee, B. R. Özer and T. Schultz, Structure of benzene from mass-correlated rotational Raman spectroscopy, *RSC Adv.*, 2022, **12**, 21406–21416, DOI: **10.1039/D2RA03431J**.

48 "Infrared Spectra" by NIST Mass Spectrometry Data Center, William E. Wallace, director, in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, ed. 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).

49 J. Gauss and J. F. Stanton, The equilibrium structure of benzene, *J. Phys. Chem. A*, 2000, **1040**(13), 2865–2868, DOI: **10.1021/jp994408y**.

50 A. J. Yencha, R. I. Hall, L. Avaldi, G. Dawber, A. G. McConkey, M. A. MacDonald and G. C. King, Threshold photoelectron spectroscopy of benzene up to 26.5 ev, *Can. J. Chem.*, 2004, **820**(6), 1061–1066, DOI: **10.1139/v04-057**.

51 O. Asvany, P. Kumar, B. Redlich, I. Hegemann, S. Schlemmer and D. Marx, Understanding the infrared

spectrum of bare ch5+, *Science*, 2005, **3090**(5738), 1219–1222, DOI: **10.1126/science.1113729**.

52 O. Asvany, K. M. T. Yamada, S. Brünken, A. Potapov and S. Schlemmer, Experimental ground-state combination differences of ch5+, *Science*, 2015, **3470**(6228), 1346–1349, DOI: **10.1126/science.aaa3304**.

53 X. Huang, A. B. McCoy, J. M. Bowman, L. M. Johnson, C. Savage, F. Dong and D. J. Nesbitt, Quantum deconstruction of the infrared spectrum of ch5, *Science*, 2006, **3110**(5757), 60–63, DOI: **10.1126/science.1121166**.

54 V. Scutelnic, M. A. S. Perez, M. Marianski, S. Warnke, A. Gregor, U. Rothlisberger, M. T. Bowers, C. Baldauf, G. von Helden, T. R. Rizzo and J. Seo, The structure of the protonated serine octamer, *J. Am. Chem. Soc.*, 2018, **1400**(24), 7554–7560, DOI: **10.1021/jacs.8b02118**.

55 P. E. Maslen, N. C. Handy, R. D. Amos and D. Jayatilaka, Higher analytic derivatives. IV. Anharmonic effects in the benzene spectrum, *J. Chem. Phys.*, 1992, **970**(6), 4233–4254, DOI: **10.1063/1.463926**.

56 A. Miani, E. Cané, P. Palmieri, A. Trombetti and N. C. Handy, Experimental and theoretical anharmonicity for benzene using density functional theory, *J. Chem. Phys.*, 2000, **1120**(1), 248–259, DOI: **10.1063/1.480577**.

57 M. L. Vidal, M. Epshtein, V. Scutelnic, Z. Yang, T. Xue, S. R. Leone, A. I. Krylov and S. Coriani, Interplay of open-shell spin-coupling and Jahn–Teller distortion in benzene radical cation probed by x-ray spectroscopy, *J. Phys. Chem. A*, 2020, **1240**(46), 9532–9541, DOI: **10.1021/acs.jpca.0c08732**.

58 M. Epshtein, V. Scutelnic, Z. Yang, T. Xue, M. L. Vidal, A. I. Krylov, S. Coriani and S. R. Leone, Table-top x-ray spectroscopy of benzene radical cation, *J. Phys. Chem. A*, 2020, **1240**(46), 9524–9531, DOI: **10.1021/acs.jpca.0c08736**.

59 D. Marx and M. Parrinello, Ch5+: The cheshire cat smiles, *Science*, 1999, **2840**(5411), 59–61, DOI: **10.1126/science.284.5411.59**.

60 P. Padma Kumar and M. Dominik, Understanding hydrogen scrambling and infrared spectrum of bare ch5+ based on ab initio simulations, *Phys. Chem. Chem. Phys.*, 2006, **8**, 573–586, DOI: **10.1039/B513089C**.

61 K. C. Thompson, D. L. Crittenden and M. J. T. Jordan, Ch5 +: Chemistryss chameleon unmasked, *J. Am. Chem. Soc.*, 2005, **1270**(13), 4954–4958, DOI: **10.1021/ja0482280**.

62 S. D. Ivanov, O. Asvany, A. Witt, E. Hugo, G. Mathias, B. Redlich, D. Marx and S. Schlemmer, Quantum-induced symmetry breaking explains infrared spectra of ch5+ isotopologues, *Nat. Chem.*, 2010, **20**(4), 298–302, DOI: **10.1038/nchem.574**.