PAPER

Check for updates

Fast Nosé–Hoover thermostat: molecular dynamics in quasi-thermodynamic equilibrium†

Dominik Sidler and Sereina Riniker [ID] *

An extension of the Nosé–Hoover thermostat equation for molecular dynamics (MD) simulation is introduced, which perturbs fast degrees of freedom out of canonical equilibrium, while preserving the average temperature of the system. Based on the generalised Liouville equation, it is shown that the newly introduced fast Nosé–Hoover (FNH) equations give rise to a dynamical system with a well-defined non-equilibrium probability distribution. Corresponding thermostat parameters are identified, which in principle allow to sample arbitrarily close to canonical equilibrium. Results show that the dynamic system properties (*e.g.* self-diffusion, shear viscosity, dielectric permittivity and rotational relaxation times) are only moderately perturbed for typical FNH setups. However, the non-equilibrium FNH equations modify the occurrence of rare events substantially and thus offer a novel approach for enhanced sampling in MD. In particular, it is shown that the FNH thermostat increased significantly the frequency of the folding and unfolding process of short β-peptides. The efficiency of the phase-space exploration solely depends on the additionally imposed quasi-equilibrium conditions, *i.e.* it does not rely on any modification of the potential-energy surface.

1 Introduction

Molecular dynamics (MD) simulations are typically performed under canonical (*NVT*) or isothermal–isobaric (*NPT*) conditions in order to compare simulated and experimental properties and to study temperature-dependent processes explicitly.[1] For this purpose, many different thermostat methods have been developed over the past decades to simulate canonical equilibrium[2–6] or closely related conditions.[7,8] A thorough overview of different thermostat approaches can for example be found in ref. 1 and 9.

Significant research effort has been aimed in the past years to develop thermostats that do not only generate a canonical ensemble but also damp numerical artefacts. This offers new ways to accelerate simulations. In particular, performance-wise it is tempting to partition a dynamical systems into multiple subproblems with different time scales in order to treat them using different time steps. However, such multi time-step integration (MTS) setups can lead to significant numerical artefacts.[10–12] A possible remedy for this issue is the use of specially designed thermostats, which stabilise large time-step integration. Typically, they reduce energy drifts and they disrupt high frequency modes from resonating with lower frequencies.[13] The most common choice is to couple (damp) each atom to a

white noise Langevin bath,[13–15] or by imposing isokinetic constraints on each degree of freedom.[16–18] However, the gain in performance obtained by using large time steps may be substantially reduced due to the disruption of the slow modes, thus greatly hindering diffusive and orientational motion.[13,19] As recently shown, this undesired side-effect can be mitigated with coloured-noise thermostats,[19–21] which couple fast modes strongly and slow modes weakly to the bath.

In this study, a contrary thermostat approach is developed, which has similarities to an idea originally proposed by Barth *et al.*[22] Instead of using MTS integration, they obtained a substantial sampling enhancement by modifying the Nosé–Hoover (NH) equation[3,4,23] to generate a non-canonical equilibrium distribution with increased tail probabilities.[22] In principle, this setup allows to accelerate the phase-space exploration without increasing the integration time steps or modifying the potential-energy surfaces. Hence, it would avoid the occurrence of undesired MTS resonance artefacts by construction. However, *a priori* it remains unclear if such a non-canonical equilibrium approach would not lead to significantly different dynamic properties, even for small perturbations compared to the canonical equilibrium.

Here, a novel perturbation of the fast degrees of freedom is introduced for the canonical NH equations. The resulting equations of motion are termed fast Nosé–Hoover (FNH) equations. In the Theory section, it is shown that the FNH equations give rise to a well-defined non-equilibrium probability distribution, which can be tuned arbitrarily close to the canonical equilibrium. In the Method section, simulation details and analysis tools are described,

Laboratory of Physical Chemistry, ETH Zürich, Vladimir-Prelog-Weg 2, 8093 Zürich, Switzerland. E-mail: sriniker@ethz.ch

This journal is © the Owner Societies 2019

Phys. Chem. Chem. Phys., 2019, **21**, 6059–6070 | 6059

which were used to validate the theoretical derivations. The Results and discussion section is split into two parts. First, relevant thermostat parameters and dynamic properties are discussed for a homogeneous water system using classical MD simulations. Second, the efficiency of the phase space exploration for non-equilibrium perturbations is investigated with respect to the occurrence of rare events. For this purpose, the folding/unfolding equilibrium of two β-peptides in methanol is analysed. The folding process of these β-peptides has been extensively studied in the past,[24–32] since it occurs on a relatively fast time scale. In the Conclusion section, relevant scientific findings are summarised and future research directions are discussed.

2 Theory

2.1 Nosé–Hoover equations of motion

Let a physical system be described by a classical Hamiltonian $H = E_{\text{kin}}(\boldsymbol{p}) + V(\boldsymbol{q})$ on a $2N$-dimensional phase space $\Gamma = (\boldsymbol{p},\boldsymbol{q})$. This implies that the dynamic system is conservative, *i.e.* the total energy is conserved, but not the temperature. Experimental measurements are, however, often performed under constant temperature (and constant pressure) conditions. In order to emulate these additional constraints, Nosé[3] and Hoover[4] developed the following non-Hamiltonian equations of motions,

$$\dot{q}_i = \frac{p_i}{m_i}, \tag{1}$$

$$\dot{p}_i = -\frac{\partial V(q)}{\partial q_i} - \xi p_i, \tag{2}$$

$$\dot{\xi} = \frac{1}{Q_1}\left[\sum_{i=1}^{N}\frac{p_i^2}{m_i} - N_{\text{df}}k_{\text{B}}T_0\right], \tag{3}$$

which are defined on an extended phase space $\Gamma_{\text{NH}}{}' = (\boldsymbol{p},\boldsymbol{q},q_\xi,\xi)$. The thermostat mass is labelled by Q_1 and particle masses of phase-space coordinates (p_i,q_i) are termed m_i. ξ is the Nosé–Hoover (NH) control variable and q_ξ is defined according to eqn (6). T_0 corresponds to the desired reference temperature of the system and N_{df} to the number of degrees of freedom. The Boltzmann constant is indicated by k_{B}.

The key property of the NH eqn (1)–(3) is related to their ability to sample from a canonical equilibrium distribution,

$$\mathscr{P}(\Gamma)_{\text{NH}} \propto e^{-\frac{H}{k_{\text{B}}T_0}}, \tag{4}$$

provided that some requirements of H on Γ are fulfilled. For example, there should be no additional conserved quantities in the system apart from a generalised energy,[3,4,23,33]

$$H_{\text{NH}} := H(\boldsymbol{p},\boldsymbol{q}) + Q_1\frac{\xi^2}{2} + N_{\text{df}}k_{\text{B}}T_0 q_\xi, \tag{5}$$

with

$$q_\xi = \int \xi \, \mathrm{d}t. \tag{6}$$

Typically this can be ensured by the removal of the centre of mass (COM) motion in order to avoid any impact of Newton's

third law $\sum_i \vec{F}_i = 0$ on the distribution $\mathscr{P}(\Gamma)_{\text{NH}}$ in the absence of other external forces. Additionally, the system should be "simple" enough such that the different time scales present in the system are well-behaved with regard to the chosen thermostat parameters. In practise, this often requires different NH thermostats for solvents and solutes.[1]

2.2 Fast Nosé–Hoover equations of motion

In the following, we introduce an extension of the NH eqn (1)–(3), which imposes an additional constraint on the fast degrees of freedom. Let the dynamical system be described by the following deterministic, non-Hamiltonian equations of motions,

$$\dot{q}_i = \frac{p_i}{m_i}, \tag{7}$$

$$\dot{p}_i = -\frac{\partial V(q)}{\partial q_i} - \xi p_i - \zeta p_i \mathscr{H}_{\mathscr{T}}(p_i), \tag{8}$$

$$\dot{\xi} = \frac{1}{Q_1}\left[\sum_{i=1}^{N}\frac{p_i^2}{m_i} - N_{\text{df}}k_{\text{B}}T_0\right], \tag{9}$$

$$\dot{\zeta} = \frac{1}{Q_2}\left[\frac{1}{3}\frac{\sum_{i=1}^{N}\frac{p_i^4}{m_i^3}}{\sum_{i=1}^{N}\frac{p_i^2}{m_i^2}} - \alpha k_{\text{B}}T_0\right], \tag{10}$$

defined on the extended phase space $\Gamma' = (\boldsymbol{p},\boldsymbol{q},q_\xi,q_\zeta,\xi,\zeta)$.

The mathematical form of the fast Nosé–Hoover (FNH) eqn (7)–(10) resembles earlier work from Barth *et al.*[22] as well as from Bulgac and Kusnezov,[34–36] who introduced a generalised form of the NH thermostat equations. T_0 corresponds to the desired reference temperature of the system, controlled by the ordinary NH variable ξ in eqn (8) and (9).[3,4,23] In addition, the auxiliary thermostat variable ζ is introduced, which perturbs fast motions according to eqn (10). For this purpose, a Heaviside step function is defined as,

$$\mathscr{H}_{\mathscr{T}}(p_i) := \begin{cases} 1 & \text{if } \frac{p_i^2}{m_i} \geq k_{\text{B}}\mathscr{T} = \frac{p_{\text{c}}^2}{m_i}, \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

where $\mathscr{T} \geq 0$ corresponds to the temperature beyond which p_i is additionally controlled by ζ, *i.e.* for $p_i \geq p_{\text{c}} := \sqrt{k_{\text{B}}\mathscr{T}m_i}$.

Note that the definition of an instantaneous temperature $T(t) := \frac{2E_{\text{kin}}}{N_{\text{df}}k_{\text{B}}}$ in eqn (9) follows from the equipartition theorem in the canonical equilibrium. Likewise, the auxiliary constraint set in eqn (10) can be related to the second order equipartition theorem, *i.e.* by integration by parts (I.P.) one can show that,

$$\left\langle \sum_{i=1}^{N}\frac{p_i^4}{m_i^3} \right\rangle \overset{\text{I.P.}}{=} 3k_{\text{B}}T_0\left\langle \sum_{i=1}^{N}\frac{p_i^2}{m_i^3} \right\rangle. \tag{12}$$

Consequently, the FNH equations defined by eqn (7)–(10) do not only allow to control the desired average temperature T of a

6060 | *Phys. Chem. Chem. Phys.*, 2019, **21**, 6059–6070

This journal is © the Owner Societies 2019

system H on Γ, but also the "variance"-like expression of the velocities according to eqn (10) *i.e.* based on (12). The parameter α defines the degree of perturbation acting on the fast motions according to eqn (8) (a more thorough discussion is given below). By increasing the probabilities in the tails of the velocity distributions using ζ, rare events (*e.g.* crossing of high-energy barriers) should become more probable within short simulation time,[22] while statistically averaged properties should not be affected substantially for small perturbations. Q_1, Q_2 correspond to the usual thermostat masses, which determine how fast the thermostat variables $(\dot\xi, \dot\zeta)$ respond to a change in the system. Note that the proposed definition for $\dot\zeta$ given in eqn (10) has the advantage that $\dot\zeta = \mathcal{O}(p_i^2)$, as it is the case for $\dot\xi$. Hence, typical numerical integration time-steps of MD simulations (*e.g.* 2 fs) should still be applicable with the FNH equations. This may not be the case anymore for smaller system sizes when the auxiliary thermostatting control variable is related directly to the instantaneous variance of the momenta, *e.g.* by imposing

$$\mathrm{Var}(E_{\mathrm{kin}} - \langle E_{\mathrm{kin}}\rangle) = \left\langle \left(\sum_i^N E_{\mathrm{kin},i}\right)^2\right\rangle \overset{!}{=} \left\langle\sum_i^N E_{\mathrm{kin},i}\right\rangle^2 = \left(\frac{N_{\mathrm{df}} k_{\mathrm{B}} T_0}{2}\right)^2,$$

which would result in $\dot\zeta = \mathcal{O}(p_i^4)$.

A priori there is no reason to assume that the proposed FNH eqn (7)–(10) sample a probability distribution close to canonical equilibrium. Therefore, in order to study the behaviour of the extended system dynamics on Γ', one has to rely on the generalised non-Hamiltonian Liouville equation,

$$0 = \frac{\partial \mathscr{P}}{\partial t} + \frac{\partial}{\partial \Gamma'}\cdot(\dot\Gamma'\mathscr{P}) \tag{13}$$

$$\Rightarrow$$

$$\frac{\mathrm{d}\mathscr{P}}{\mathrm{d}t} = -\left(\frac{\partial}{\partial\Gamma'}\cdot\dot\Gamma'\right)\mathscr{P}, \tag{14}$$

which describes the time evolution of the probability density $\mathscr{P}(\Gamma',t)$ of the dynamic system Γ'. Note that if an equilibrium solution $\mathscr{P}_0$ exists, it could be constructed by interpreting $\mathscr{P}_0 = f\sqrt{g}$ as a product of a distribution function f with non-trivial metric g associated to the curved phase space Γ'.[33,37,38] The equilibrium $\mathscr{P}_0$ then follows immediately from all conservation laws of the extended system.[33,37,38] However, for the following derivation of $\mathscr{P}(\Gamma',t)$ for the FNH equations, we follow an alternative pathway described by Nosé in ref. 23.

Let the generalised FNH energy be defined as,

$$H_{\mathrm{FNH}} := H(\boldsymbol{p},\boldsymbol{q}) + Q_1\frac{\xi^2}{2} + N_{\mathrm{df}}k_{\mathrm{B}}T_0 q_\xi + \int\sum_{i=1}^N \zeta\left\{\frac{p_i^2}{m_i}\mathscr{H}_{\mathscr{T},k}(p_i)\right\}\mathrm{d}t \tag{15}$$

$$= H_{\mathrm{NH}} + \int\sum_{i=1}^N \zeta\left\{\frac{p_i^2}{m_i}\mathscr{H}_{\mathscr{T},k}(p_i)\right\}\mathrm{d}t, \tag{16}$$

where the continuous logistic function $\mathscr{H}_{\mathscr{T},k}(p_i) = 1/(1+\exp[-k(p_i - p_{\mathrm{c}})/p_{\mathrm{c}}]) + 1/(1+\exp[-k(-p_i-p_{\mathrm{c}})/p_{\mathrm{c}}])$ is used as an approximation of the Heaviside step function

with $\lim_{k\to\infty}\mathscr{H}_{\mathscr{T},k}(p_i) = \mathscr{H}_{\mathscr{T}}(p_i)$. It is straightforward to show that H_{FNH} is conserved as,

$$\frac{\mathrm{d}H_{\mathrm{FNH}}}{\mathrm{d}t} = \sum_{i=1}^N\left(\dot p_i\frac{\partial H_{\mathrm{NH}}}{\partial p_i} + \dot q_i\frac{\partial H_{\mathrm{NH}}}{\partial q_i}\right) + \dot\xi\frac{\partial H_{\mathrm{NH}}}{\partial \xi} + \dot\xi\frac{\partial H_{\mathrm{NH}}}{\partial q_\xi}$$

$$+ \sum_{i=1}^N \zeta\left(\frac{p_i^2}{m_i}\mathscr{H}_{\mathscr{T},k}(p_i)\right)$$

$$= \sum_{i=1}^N\left\{\left(-\frac{\partial V}{\partial q_i} - \xi p_i - \zeta p_i\mathscr{H}_{\mathscr{T},k}(p_i)\right)\frac{p_i}{m_i} + \frac{p_i}{m_i}\frac{\partial V}{\partial q_i}\right\}$$

$$+ \frac{1}{Q_1}\left\{\sum_{i=1}^N\frac{p_i^2}{m_i} - N_{\mathrm{df}}k_{\mathrm{B}}T_0\right\}Q_1\xi$$

$$+ N_{\mathrm{df}}k_{\mathrm{B}}T_0\xi + \sum_{i=1}^N\zeta\frac{p_i^2}{m_i}\mathscr{H}_{\mathscr{T},k}(p_i) = 0. \tag{17}$$

From this conservation law, a solution for the generalised Liouville eqn (14) of the non-Hamiltonian system dynamics described by the FNH eqn (7)–(10) is obtained using,

$$\frac{\partial}{\partial\Gamma'}\cdot\dot\Gamma' = \sum_{i=1}^N\left(\frac{\partial}{\partial p_i}\dot p_i + \frac{\partial}{\partial q_i}\dot q_i\right) + \frac{\partial}{\partial\xi}\dot\xi + \frac{\partial}{\partial q_\xi}\dot q_\xi + \frac{\partial}{\partial\zeta}\dot\zeta + \frac{\partial}{\partial q_\zeta}\dot q_\zeta \tag{18}$$

$$= \sum_{i=1}^N\frac{\partial}{\partial p_i}\big(-\xi p_i - \zeta p_i\mathscr{H}_{\mathscr{T},k}(p_i)\big) \tag{19}$$

$$= -N_{\mathrm{df}}\xi - \zeta\left[\sum_{i=1}^N\big(\mathscr{H}_{\mathscr{T},k}(p_i) + |p_i|\delta_{\mathscr{T},k}(p_i)\big)\right], \tag{20}$$

where a reduced N_{df} is considered in the last step due to additionally imposed constraints (*e.g.* rigid molecular bonds) and $\delta_{\mathscr{T},k}(p_i) := \frac{\mathrm{d}}{\mathrm{d}p_i}\mathscr{H}_{\mathscr{T},k}(p_i) = \mathscr{H}_{\mathscr{T},k}(p_i)\big(1 - \mathscr{H}_{\mathscr{T},k}(p_i)\big)\frac{k}{p_{\mathrm{c}}}$. It can be shown that the resulting non-equilibrium probability density obeying eqn (14) can be written as,

$$\mathscr{P}(\Gamma',t) :=$$

$$C\mathrm{e}^{-\frac{1}{k_{\mathrm{B}}T_0}\left(H(\boldsymbol{p},\boldsymbol{q}) + Q_1\frac{\xi^2}{2} + \int\zeta\left\{\sum_{i=1}^N\frac{p_i^2}{m_i}\mathscr{H}_{\mathscr{T},k}(p_i) - k_{\mathrm{B}}T_0[\mathscr{H}_{\mathscr{T},k}(p_i) + |p_i|\delta_{\mathscr{T},k}(p_i)]\right\}\mathrm{d}t\right)} \tag{21}$$

$$\overset{(k\to\infty)}{=} C\mathrm{e}^{-\frac{1}{k_{\mathrm{B}}T_0}\left(H(\boldsymbol{p},\boldsymbol{q}) + Q_1\frac{\xi^2}{2} + \int\zeta\left\{\sum_{i=1}^N\frac{p_i^2}{m_i}\mathscr{H}_{\mathscr{T}}(p_i) - k_{\mathrm{B}}T_0[\mathscr{H}_{\mathscr{T}}(p_i) + |p_i|\delta_{\mathscr{T}}(p_i)]\right\}\mathrm{d}t\right)} \tag{22}$$

$$= C\mathrm{e}^{-\frac{1}{k_{\mathrm{B}}T_0}\left(H(\boldsymbol{p},\boldsymbol{q}) + Q_1\frac{\xi^2}{2} + \int\zeta\sum_{i=1}^n\phi(p_i)\mathrm{d}t\right)}, \tag{23}$$

where in the last step,

$$\phi(p_i) := \frac{p_i^2}{m_i}\mathscr{H}_{\mathscr{T},k}(p_i) - k_{\mathrm{B}}T_0\big[\mathscr{H}_{\mathscr{T},k}(p_i) + |p_i|\delta_{\mathscr{T}}(p_i)\big], \tag{24}$$

was introduced for the ease of notation.

This journal is © the Owner Societies 2019

Phys. Chem. Chem. Phys., 2019, **21**, 6059–6070 | **6061**

From eqn (23), it follows that the non-equilibrium FNH probabilities are modified with respect to the canonical NH probabilities in eqn (4) based on the following mapping,

$$\frac{p_i^2}{2m_i} \mapsto \begin{cases} \dfrac{p_i^2}{2m_i} + \displaystyle\int \zeta\phi(p_i)\mathrm{d}t & \text{if } \dfrac{p_i^2}{m_i} \geq k_{\mathrm{B}}\mathscr{T} \\[2ex] \dfrac{p_i^2}{2m_i} & \text{otherwise} \end{cases}. \qquad (25)$$

Note that the non-equilibrium perturbations are discontinuous in the chosen setup with $(k \to \infty)$, which is expected to disrupt the smooth motion compared to the canonical equilibrium. In addition, if $\mathscr{T} > T_0$ the following property holds,

$$\phi(p_i) = \begin{cases} -\infty & \text{if } \dfrac{p_i^2}{m_i} = k_{\mathrm{B}}\mathscr{T} \\[2ex] > 0 & \text{otherwise} \end{cases}. \qquad (26)$$

Based on aforementioned theoretical considerations it is possible to discuss several practical aspects of the proposed FNH equations:

(1) A canonical equilibrium probability distribution for Γ can only be obtained if $\int \zeta\phi(p_i)\mathrm{d}t = 0 \, \forall i, t$. For arbitrary initial conditions, this is only possible for $\lim_{Q_2 \to \infty}\zeta = 0$ or $\lim_{\mathscr{T} \to \infty}\phi(p_i) = 0$ due to eqn (26), which means the normal NH eqn (1)–(3) are recovered.

(2) Condition (1) automatically implies that $\nexists\alpha_0$ with a canonical equilibrium solution of the Liouville equation. This result is in agreement with the properties obtained in ref. 39 for multi-temperature dynamics. Therefore, the proposed FNH equations can at best approximate a (canonical) equilibrium distribution for finite $\mathscr{T}$, Q_2.

(3) For practical purposes, one needs to ensure that the tail perturbations of the FNH equations remain small with respect to the canonical equilibrium, i.e. $\mathscr{P}(\Gamma, t)_{\mathrm{FNH}} \approx \mathscr{P}(\Gamma)_{\mathrm{NH}}$. The requirement of such a quasi equilibrium distribution leads to the following conditions for the thermostat parameters and the simulated systems:

(a) $\mathscr{T} \gg T_0$ restricts perturbations to the tails of the kinetic-energy distribution.

(b) $N \gg 1$ ensures that $\exists p_i$ with $\dfrac{p_i^2}{m_i} \geq k_{\mathrm{B}}\mathscr{T}, \forall t$, which can be efficiently controlled by ζ at every time-step.

(c) To limit the non-equilibrium perturbations to a small value, one can define a "quasi" canonical equilibrium at $\alpha_0^{\mathrm{exact}}$ as follows,

$$\left\langle \dot{\zeta}\left(\alpha_0^{\mathrm{exact}}\right) \right\rangle \overset{!}{=} 0, \qquad (27)$$

which is non-trivial to solve. In practice, $\alpha_0^{\mathrm{exact}}$ can be estimated based on eqn (10),

$$\alpha_0^{\mathrm{exact}} \approx \frac{1}{3k_{\mathrm{B}}T_0} \left\langle \frac{\sum\limits_{i=1}^{N}\dfrac{p_i^4}{m_i^3}}{\sum\limits_{i=1}^{N}\dfrac{p_i^2}{m_i^2}} \right\rangle_{\mathrm{NH}} \qquad (28)$$

$$\approx \frac{Q_2\langle\dot{\zeta}\rangle}{3k_{\mathrm{B}}T_0} = \alpha_0,$$

by setting $Q_2 \gg Q_1$ and $\alpha = 0$. This approximation relies on condition (1), for which the FNH equations approximate a canonical

equilibrium distribution. Any coupling effects between ξ and ζ, which may occur for finite Q_2, are neglected. Typically these effects should be small for systems with a high number of degrees of freedom. Thus, the accuracy of the estimated α_0 depends on N_{df} as well as on the particular system studied. For well-behaved systems, a small perturbation $\delta\alpha$ of α_0 is expected to drive the system gently out of quasi-canonical equilibrium.

(4) In order to obtain a rough estimate for reasonable choices of the thermostat mass Q_i, one can investigate the dynamics of the extended system variables ξ, ζ for linearised FNH equations assuming $\dfrac{\partial V(q)}{\partial q_i} = 0$ and $\alpha = 1$. For simplicity, we additionally assume that the dynamics of ξ and ζ are completely decoupled by setting $\dot{p}_i^{(\xi)} := -\xi p_i$ and $\dot{p}_i^{(\zeta)} := -\zeta p_i$. For a rough estimate, this is a reasonable assumption if one considers that ζ only controls the tails, whereas ξ mainly controls the mean of $\mathscr{P}$. These assumptions lead to the following two decoupled ordinary differential equations (see Appendix A),

$$\ddot{\xi} \approx -\frac{2N_{\mathrm{df}}k_{\mathrm{B}}T_0}{Q_1}\xi \qquad (29)$$

$$\ddot{\zeta} \approx -\frac{2\alpha k_{\mathrm{B}}T_0}{Q_2}\zeta, \qquad (30)$$

with an implied thermostat times scale, which can be defined as,

$$\tau_1 := \sqrt{\frac{Q_1}{2N_{\mathrm{df}}k_{\mathrm{B}}T_0}} \qquad (31)$$

$$\tau_2 := \sqrt{\frac{Q_2}{2k_{\mathrm{B}}T_0}}. \qquad (32)$$

In classical MD, typically $\tau_1 \approx 0.1$ ps is chosen with NH. For the FNH extension, we recommend $\tau_1 < \tau_2 \approx 1$ ps as a default, in order to reduce instabilities in case condition (3b) is violated.

(5) To anticipate the preservation of hydrodynamic properties by the FNH equations, one can apply a similar argument for $N_{\mathrm{df}} \to \infty$ as given in ref. 40. In order to ensure efficient equilibration, above-mentioned implied thermostat time scales τ_1 and τ_2 should remain fixed independently of N_{df}. Using eqn (29), it follows $Q_1 \to \infty$ if $N_{\mathrm{df}} \to \infty$, which automatically implies $\dot{\xi} \to 0$ according to eqn (9). Therefore, one can show that in canonical equilibrium $\langle\xi^2\rangle = \dfrac{k_{\mathrm{B}}T_0}{Q_1} \to 0$ for NH equations, which in turn leads to $\xi \to 0$.[40] Consequently, the NH equations reduce to Newtonian equations if the system size is increased. Therefore, hydrodynamic properties are preserved for large systems. In contrast, since τ_2 is independent of N_{df} for a FNH thermostat, $\dot{\zeta}$ is not necessarily 0, $\forall t$. For this reason, the FNH equations do not strictly converge to Newtonian equations of motion for $N_{\mathrm{df}} \to \infty$, which means that hydrodynamic properties remain affected also for larger system sizes. The underlying reason for Q_2 not depending on N_{df} is related to the fractional definition of $\dot{\zeta} \propto \dfrac{\sum\limits_{i}^{N}\dfrac{p_i^4}{m_i^3}}{\sum\limits_{i}^{N}\dfrac{p_i^2}{m_i^2}}$ given in eqn (10), which ensures numerical stability for typical integration

6062 | *Phys. Chem. Chem. Phys.*, 2019, **21**, 6059–6070

This journal is © the Owner Societies 2019

time steps and intermediate N_{df} values based on the $\mathcal{O}(p_i^2)$-scaling. However, hydrodynamic properties of the FNH thermostat remain mostly preserved in practice for typical simulation setups, as will be shown in the Results & discussion section.

(6) Note that the computational complexity of the FNH equations remains unaffected by the additional control variable and constraint in eqn (8) and (10). The number of operations for the force calculation in eqn (10) is, however, slightly increased for every degree of freedom due to the additional control term. Practically more relevant is the evaluation of a global expression for $\dot{\xi}$ and $\dot{\zeta}$, which involves global communication across processors. Similar to NH, this is a disadvantage with respect to an efficient parallel implementation.[40] However, the additional control term ζ, which acts on the forces, should not affect the speed of the computation significantly compared to a normal NH thermostat. As mentioned earlier, the auxiliary constraint is also designed such that the integration time step does not have to be adjusted, which ensures an efficient numerical integration.

(7) Note that it should be straightforward to replace the NH thermostat within the FNH equations by a NH chain thermostat[5] in order to the control the dynamics of ξ.

3 Methods

3.1 Simulation details

Two different types of systems were used for the investigation of the FNH properties: homogeneous systems (liquids) and short β-peptides. First, a cubic box of 1000 SPC water molecules[41] was simulated using different thermostat parameters. In addition, auxiliary systems consisting of a cubic box with 1000 chloroform molecules[42] and two boxes with 1000 and 5000 argon atoms were used to investigate the impact of different system properties onto the estimated quasi-equilibrium parameter α_0.

Second, two β-peptides (Fig. 1) solvated in methanol were studied (1096 and 1123 molecules, respectively).[31,32] Their topology and initial coordinates were taken from ref. 32. For peptide 1, a left-handed 3_{14} helical fold was used as starting structure, and for peptide 2 a right-handed $2.7_{10/12}$ helical fold.[32] The β-peptides offer an ideal test system as they fold

(a) Peptide 1

(b) Peptide 2

Fig. 1 Amino acid sequence of two β-peptides: (a) peptide 1: H$_2^+$-β^3-HVal-β^3-HAla-β^3-HLeu-(S,S)-β^3-HAla(αMe)-β^3-HVal-β^3-HAla-β^3-HLeu-OH; (b) peptide 2: H$_2^+$-β^2-HVal-β^3-HAla-β^2-HLeu-β^3-HVal-β^2-HAla-β^3-HLeu-OH.
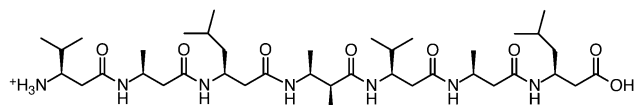
on a time scale of nanoseconds in methanol, which is easily accessible with classical MD.[24] However, the folding events are rare with respect to other implied time scales of the dynamical system (e.g. rotational relaxation times τ or thermostat times τ_1 and τ_2). Therefore, the effect of the perturbed FNH tail distributions onto rare events can be investigated using these β-peptides.

For all simulations, a modified version of the GROMOS software package[43,44] was used in combination with the GROMOS 54A7β force field.[45,46] Periodic boundary conditions were imposed for all systems and all simulations were performed under NVT conditions. The water and chloroform systems were simulated at 300 K, the argon system at 80 K, and the β-peptides at 340 K. As a canonical reference system, the NH thermostat was used with a relaxation time of 0.1 ps.[3,4] Newton's equations of motion were integrated using the leapfrog scheme[47] with a time step of 2 fs using a single cutoff at 1.4 nm with a pairlist update every 10 fs.[12] The centre of mass (COM) motion was stopped every 1 ps. Bond lengths were constrained using the SHAKE algorithm.[48] Long-range electrostatic interactions were described using a homogeneous RF approach,[49] with a charge-group based cutoff $R = 1.4$ nm.[50] For all simulations involving SPC, a dielectric permittivity of 78.4 (experimental value of water[51]) was used. For methanol $\varepsilon_{rf}^{CH_3OH} = 19.8$ (calculated value from ref. 52) and for chloroform $\varepsilon_{rf}^{CHCl_3} = 5.0$ (experimental value from ref. 53) were chosen. For argon no long-range electrostatics have to be considered. The water systems were simulated for 50 ns, the auxiliary systems for 1.5 ns, and the β-peptides for 350 ns. Coordinates, velocities and energies were written out every 1 ps, unless stated otherwise. The quasi canonical equilibrium parameter α_0 was obtained from eqn (28) by setting $\tau_2 = 1000$ μs and by taking the ensemble average from 1 ns trajectories after 0.5 ns equilibration. In all FNH simulations, $\mathcal{T} = 7T_0$ was used, which was considered a reasonable value to restrict the non-canonical perturbations of $\mathcal{P}$ to the tails.

3.2 Analysis

Different statistical quantities were extracted from the simulation trajectories and averaged using the GROMOS++ analysis tools.[54] Self-diffusion coefficients D were obtained according to the Einstein relation[55,56] using the diffus program. For this purpose, 12 short simulations of 1.5 ns length (after 2 ns equilibration) were performed with different initial setups. The uncertainty of the self-diffusion was quantified by calculating the standard deviation of the 12 different values. Similarly, shear viscosities η_{shear} were extracted from the off-diagonal elements of the pressure tensor using a modified version of the visco program.[56,57] For this purpose, the slope of a linear least-square fit of the squared pressure tensor integrals was evaluated between 0 and 10 ps. The data for the pressure tensor integrals was gathered from 1 ns of simulation with a time resolution of 2 fs. The uncertainty of the shear viscosity was quantified based on the standard deviation of the fitted slope. The rotational relaxation time τ was calculated from fitting an exponential decaying function $Ce^{-t/\tau}$ to the first-order autocorrelation functions, which were calculated for vectors
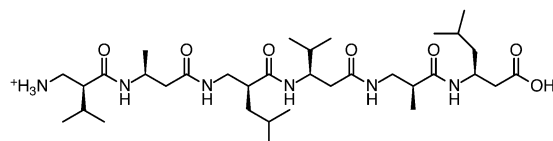
This journal is © the Owner Societies 2019

Phys. Chem. Chem. Phys., 2019, **21**, 6059–6070 | **6063**

Table 1 Comparison of estimated α_0 with corresponding standard deviations $\Delta\alpha_0$ for different systems. For the homogeneous systems, α_0 values were estimated from 1 ns simulation trajectories after equilibration for 0.5 ns. For the β-peptides, 1 ns equilibration and 2 ns production run were performed. In all cases, the following parameters were set during the estimation run: τ_{Q_1} = 0.1 ps, τ_{Q_2} = 1 ms (!) and α = 0. Note that the removal of the COM is considered in the calculation of N_{df}

System	Argon	Argon	Water	Chloroform	Peptide 1	Peptide 2
N_{atoms}	1000	5000	3000	5000	3352	3425
N_{df}	2997	14 997	5997	5997	6702	6848
$\alpha_0 \pm \Delta\alpha_0$	0.998 ± 0.050	0.999 ± 0.036	0.658 ± 0.044	0.450 ± 0.031	0.637 ± 0.034	0.637 ± 0.036

pointing along different directions (*e.g.* H–H-bond) of the molecular reference frame. The autocorrelation function was obtained using the rot_rel program. For this analysis, the coordinates were written out every 2 fs over 10 ps of simulation time, after an equilibration time of 25 ns. The dielectric permittivity ε was calculated from 50 ns trajectories (discarding the first 1 ns as equilibration) by solving a Kirkwood–Fröhlich-type equation,[58] implemented in the epsilon program. Note that this approach involves the dielectric permittivity of the reaction field as an input parameter. In order to quantify the uncertainty of the dielectric permittivity estimate, the standard deviation was calculated from 10 simulations with identical starting coordinates but different initial velocities. In order to measure the folding and unfolding of the β-peptides, the atom-positional root-mean-square deviation (RMSD) of the backbone atoms between residues 2–6 (peptide 1) and residues 2–5 (peptide 2) were calculated using the rmsd program.

4 Results & discussion

4.1 Homogeneous systems

In a first step, the quasi-equilibrium parameter α_0 was determined according to eqn (28) for all simulated systems. As can be seen in Table 1, α_0 = 1 for the argon simulation. This is an expected result in the absence of any constraints (apart from the COM removal) due to the generalised equipartition theorem given in eqn (12). However, in the presence of constraints (*i.e.* bonds with fixed length) the estimated α_0 parameters depend on the system studied and cannot be anticipated without preparatory simulations. This becomes apparent if one compares $\alpha_0^{CHCl_3} = 0.450$ and $\alpha_0^{H_2O} = 0.658$, which were both obtained for systems containing 1000 molecules with an equally number of degrees of freedom $N_{df}^{H_2O} = N_{df}^{CHCl_3} = 6000$ (both solvent models are fully constrained, *i.e.* without internal degrees of freedom).

The effect of the quasi equilibrium arising from ζ onto the probability distribution $\mathscr{P}$ with respect to different perturbations $\delta\alpha$ of α_0 was investigated in detail for the water box (Fig. 2 and Table 2). For this system, $\alpha_0 = 0.658$ was estimated using eqn (28).

All setups have in common that the average system temperature corresponds to the desired system temperature ($T \approx T_0$). However, depending on the chosen FNH perturbation, the standard deviation of the system's temperature is decreased for $\alpha < \alpha_0$ or increased for $\alpha > \alpha_0$. This is an expected behaviour since the fluctuations in kinetic energy ΔE_{kin} (*i.e.* temperature) can be used to probe the canonicity of a thermostat.[59] For finite systems of size N without constraints (*e.g.* without rigid bonds) in canonical equilibrium,
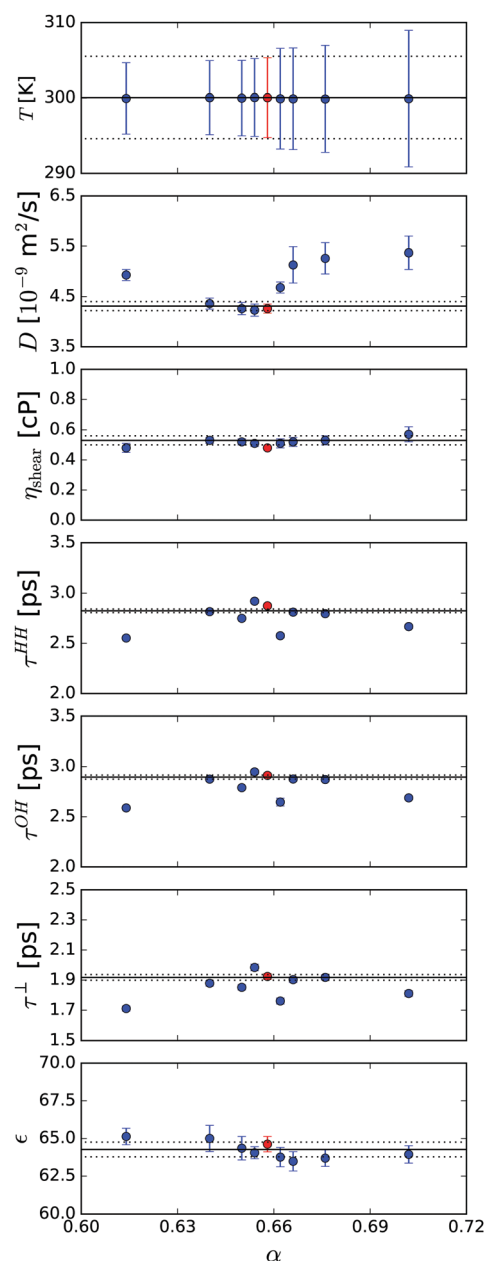

Fig. 2 Impact of different α-parameters on the time-averaged temperature T and dynamic system properties of water (self-diffusion D, shear viscosity η_{shear}, rotational relaxation time τ and dielectric permittivity ε). The data with $\alpha = \alpha_0$ is shown in red. Canonical equilibrium values from NH simulation are indicated as a horizontal black line. If displayed, dotted black lines correspond to error bars of the NH reference. For all simulations, the following thermostat parameters were chosen: τ_{Q_1} = 0.1 ps, τ_{Q_2} = 1.0 ps, $\mathscr{T}$ = 2100 K, T_0 = 300 K.

6064 | *Phys. Chem. Chem. Phys.*, 2019, **21**, 6059–6070

This journal is © the Owner Societies 2019

Table 2 Impact of the choice of the α-parameter on the time-averaged static temperature T and dynamic system properties (self-diffusion D, shear viscosity η_{shear}, rotational relaxation time τ, dielectric permittivity ε). Canonical equilibrium values from NH simulation are given in the last column as reference. For all simulations, the following thermostating parameters were chosen: $\tau_{Q_1} = 0.1$ ps, $\tau_{Q_2} = 1.0$ ps, $\mathscr{T} = 2100$ K, $T_0 = 300$ K

α	0.614	0.640	0.650	0.654	0.658	0.662	0.666	0.676	0.702	Nosé–Hoover
$\langle T \rangle$ [K]	299.94 ± 4.73	300.04 ± 4.90	299.98 ± 4.99	300.06 ± 5.15	300.04 ± 5.29	299.99 ± 6.67	299.89 ± 6.73	299.87 ± 7.09	299.91 ± 9.06	300.06 ± 5.46
$\langle D \rangle$ [10^{-9} m^2 s^{-1}]	4.94 ± 0.11	4.36 ± 0.11	4.26 ± 0.12	4.23 ± 0.12	4.26 ± 0.09	4.68 ± 0.12	5.13 ± 0.36	5.26 ± 0.31	5.36 ± 0.33	4.31 ± 0.09
η_{shear} [cP]	0.48 ± 0.03	0.53 ± 0.02	0.52 ± 0.02	0.51 ± 0.02	0.48 ± 0.01	0.51 ± 0.03	0.52 ± 0.03	0.53 ± 0.03	0.57 ± 0.05	0.53 ± 0.03
τ_{HH} [ps]	2.553 ± 0.016	2.816 ± 0.017	2.749 ± 0.016	2.920 ± 0.021	2.874 ± 0.019	2.576 ± 0.021	2.811 ± 0.018	2.796 ± 0.016	2.667 ± 0.019	2.824 ± 0.018
τ^{OH} [ps]	2.588 ± 0.017	2.875 ± 0.020	2.790 ± 0.017	2.947 ± 0.022	2.912 ± 0.022	2.646 ± 0.039	2.876 ± 0.018	2.871 ± 0.018	2.688 ± 0.018	2.895 ± 0.020
$\tau^{\perp}$ [ps]	1.711 ± 0.016	1.879 ± 0.018	1.852 ± 0.017	1.984 ± 0.022	1.924 ± 0.019	1.761 ± 0.020	1.811 ± 0.018	1.903 ± 0.018	1.918 ± 0.018	1.918 ± 0.019
ε	65.7 ± 0.55	65.00 ± 0.87	64.35 ± 0.78	64.05 ± 0.40	64.62 ± 0.51	63.76 ± 0.64	63.48 ± 0.64	63.69 ± 0.55	63.94 ± 0.58	64.27 ± 0.49

the deviation of the average kinetic energy can be written as

$$\frac{\Delta E_{kin}}{\langle E_{kin} \rangle} = \frac{\sqrt{\langle E_{kin}^2 \rangle - \langle E_{kin} \rangle^2}}{\langle E_{kin} \rangle} = \sqrt{\frac{2}{3N}}.$$ [59] Any non-canonical equilibrium distribution of the velocities will result in deviations regarding the observed temperature fluctuations for finite systems. Therefore, the non-equilibrium property of the FNH equation leads to α-dependent temperature fluctuations as observed in Fig. 2.

The analysis of the self-diffusion coefficient D shows a relatively small sigmoid-like increase for $\alpha > \alpha_0$. Thus, significant deviations from canonical equilibrium values were already observed for very small perturbations $\delta\alpha = \pm 0.004$, which are well within the standard deviation $\Delta\alpha_0 = 0.044$ of the estimated α_0. For example, a reduction of $\alpha = \alpha_0 - 0.004$ leads to a slightly slower diffusion and *vice versa* for an increased α (Fig. 2). However, overall, the self-diffusion remains relatively well preserved over the entire α-parameter range. In particular, all values are well within the range of computed values, which were obtained from different simulation studies of SPC water.[57,60–67]

Interestingly, the second calculated transport property, *i.e.* the shear viscosity η_{shear} seems to be almost unaffected by the non-equilibrium perturbation of the fast motions, except for the most extrem α-values. Similarly to the diffusivity, all obtained shear viscosities agree well with computational values from literature.[57,60,61,67,68] The FNH thermostat seems to be capable of calculating shear viscosity relatively accurately for reasonable thermostatting parameter choices, despite its inability to preserve momentum exactly for $N_{df} \to \infty$ as it is the case for the normal NH equations. For the accurate calculation of transport properties, often *NVE* simulations that start from a well initialised temperature and pressure level are performed in practice, which preserve hydrodynamic properties.[56,69] If the *NVE* simulations are short enough, the temperature and pressure remains approximately constant. However, in the vicinity of a critical point (*i.e.* close to a phase transition) long simulation runs are required due to critical slowing down, causing a problem with such *NVE* simulations.[69,70] Therefore, often simulations with NH or dissipative particle dynamics (DPD) thermostats are employed instead.[40,69,71] For NH, it is known that self-diffusion and shear viscosity are adequately derived. With DPD, also accurate bulk viscosities are accessible, despite the reduced capability of DPD thermostats to control temperature.[69] Note that transport properties can alternatively also be derived from non-equilibrium MD simulations (see *e.g.* ref. 71). Based on these considerations and the results displayed in Fig. 2, one can conclude that the FNH equations preserve the known NH capabilities with respect to the calculation of transport properties (diffusivity and shear viscosity) for α close to α_0, *i.e.* for small non-equilibrium perturbations. Results for rotational correlation times and dielectric permittivities are shown in Fig. 2 and Table 2. All values are mostly preserved over the entire range of α-values. Especially the dielectric permittivity shows nearly no dependence on α.

Overall, these results show that the dynamic properties of the system studied remain close to the respective canonical equilibrium values for FNH setups with small perturbation. The deviations are on the same order of magnitude as observed for
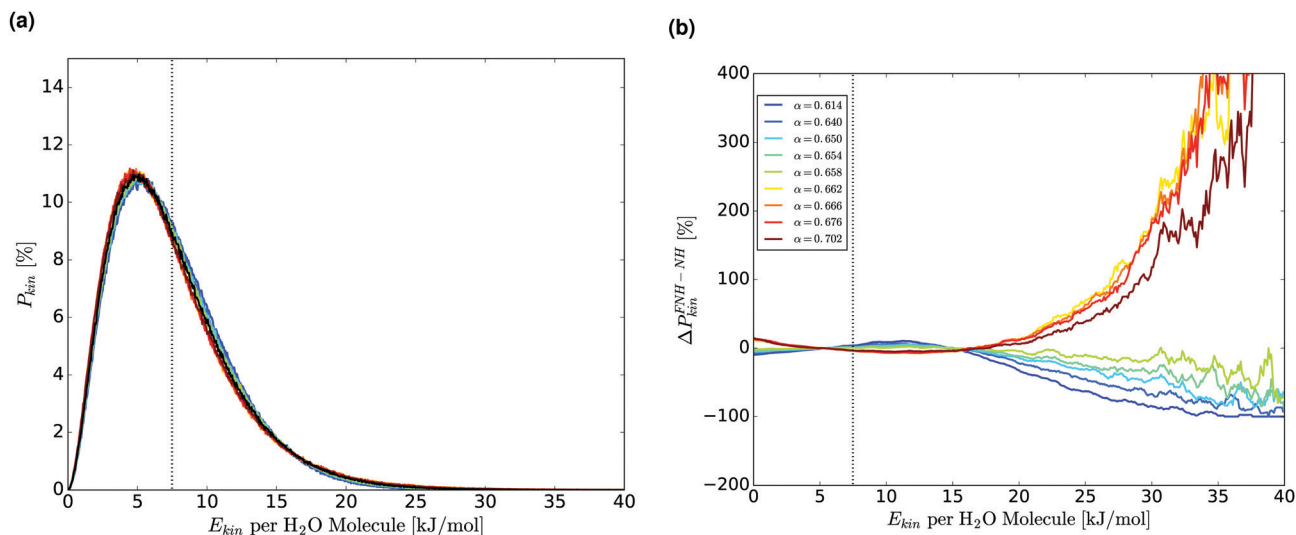
(a)

(b)

Fig. 3 (a) Probability density of the molecular kinetic energy P_{kin} of water for different FNH thermostat parameters α at 300 K. The reference Nosé–Hoover (NH) distribution is shown in black. The dotted vertical line denotes $\langle E_{kin} \rangle = 3k_BT$. (b) Corresponding relative differences of the tail probabilities with respect to the canonical equilibrium values, by calculating $\Delta P_{kin}^{FNH-NH} = \dfrac{P_{kin}^{FNH} - P_{kin}^{NH}}{P_{kin}^{NH}}$.

a coloured-noise thermostat applied to suppress MTS integration errors,[13] and substantially lower than for other MTS-optimised thermostats.[13,19] In particular, the quasi-canonical equilibrium solution α_0 corresponds closely to the canonical NH solution for every observed quantity as expected from theory. This observation indicates that eqn (28) provides indeed a reasonable estimator for α_0.

In order to quantify the effect of the perturbed probability density function $\mathscr{P}$ further, the corresponding probability distribution functions of the molecular kinetic energies P_{kin} were compared for the water box (Fig. 3a). A plot of the relative probability differences $\Delta P_{kin}^{FNH-NH} = \dfrac{P_{kin}^{FNH} - P_{kin}^{NH}}{P_{kin}^{NH}}$ with respect to canonical equilibrium is given in Fig. 3b. Both plots nicely illustrate that the FNH equations indeed generate kinetic energy distributions close to canonical equilibrium, but with substantially perturbed tails of the distribution. This result was theoretically predicted from $\mathscr{P}(\Gamma',t)$ given in eqn (23), and for small perturbations in eqn (25). Surprisingly, the highest positive perturbation $\Delta\alpha_0 = 0.044$ has a weaker impact on the tail probabilities (dark red line) than smaller positive perturbations $\alpha > \alpha_0$. The reason for that becomes evident if one investigates the time evolution of the thermostat control variables ξ and ζ with respect to α (Fig. 4). For $\alpha = \alpha_0 + 0.044 = 0.702$, these do not converge to a quasi-equilibrium with constant sign as it is the case for smaller perturbations. Therefore, according to eqn (25), the perturbations of the fast motions are not always negative for this case. This results in a reduced increase of the time-averaged tail probabilities (Fig. 3b) and larger temperature fluctuations (Fig. 2).

In addition, Fig. 4 reveals four more interesting properties of the FNH equations. First, the FNH thermostat is very sensitive regarding the choice of α, since substantial deviations of $\langle\zeta\rangle$ from its quasi canonical equilibrium value $\langle\zeta(\alpha_0^{exact})\rangle = 0$ can already occur for very small deviations of α. In particular, the simulations showed that $\alpha_0^{exact} \in (0.658, 0.662)$ for the water
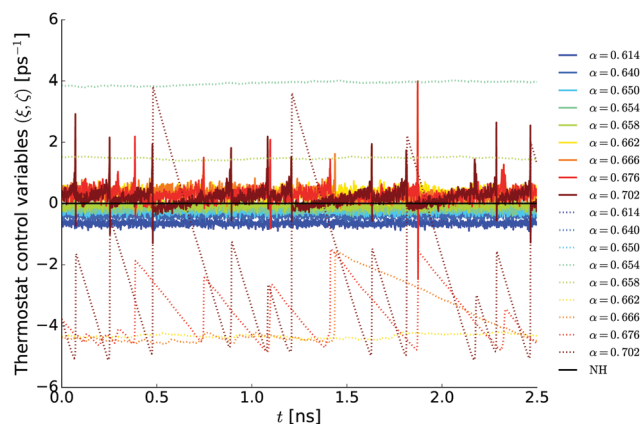

Fig. 4 Time evolution of the thermostat control variables ξ (solid lines) and ζ (dotted lines) with respect to different fast Nosé–Hoover (FNH) perturbations α for a test system of 1000 water molecules. The estimated $\alpha_0 = 0.658$ (green lines). The Nosé–Hoover (NH) reference value for canonical equilibrium is shown in black. Note that some doted lines are outside the chosen scale for the y-axis (e.g. $\langle\zeta_{FNH}(\alpha = 0.614)\rangle = 31.99$).

box. Second, the control variables ξ and ζ are substantially more noisy for all FNH setups compared to ξ of the NH thermostat (black line in Fig. 4), as one would expect from the discontinuities introduced by the Heaviside function. Third, imposing the additional constraint for the fast degrees of freedom in the FNH eqn (10) is an asymmetric and nonlinear problem. This becomes evident for symmetric perturbations $\pm\Delta\alpha_0$, which lead to a relatively stable solution for $\langle\zeta(\alpha_0 - 0.044)\rangle \approx 32$ ps^{-1}, but cause oscillations for $\langle\zeta(\alpha_0 + 0.044)\rangle \in (-5, 0)$ ps^{-1}. Fourth, the FNH equations remain relatively close to canonical equilibrium for the estimated α_0 values (green lines in Fig. 4), i.e. $\langle\xi(\alpha_0)\rangle \approx \langle\xi^{NH}\rangle$ but with slightly reduced tail probabilities since $\zeta(\alpha_0) > 0 \forall t$.

6066 | Phys. Chem. Chem. Phys., 2019, **21**, 6059–6070

This journal is © the Owner Societies 2019

(a)

(b)

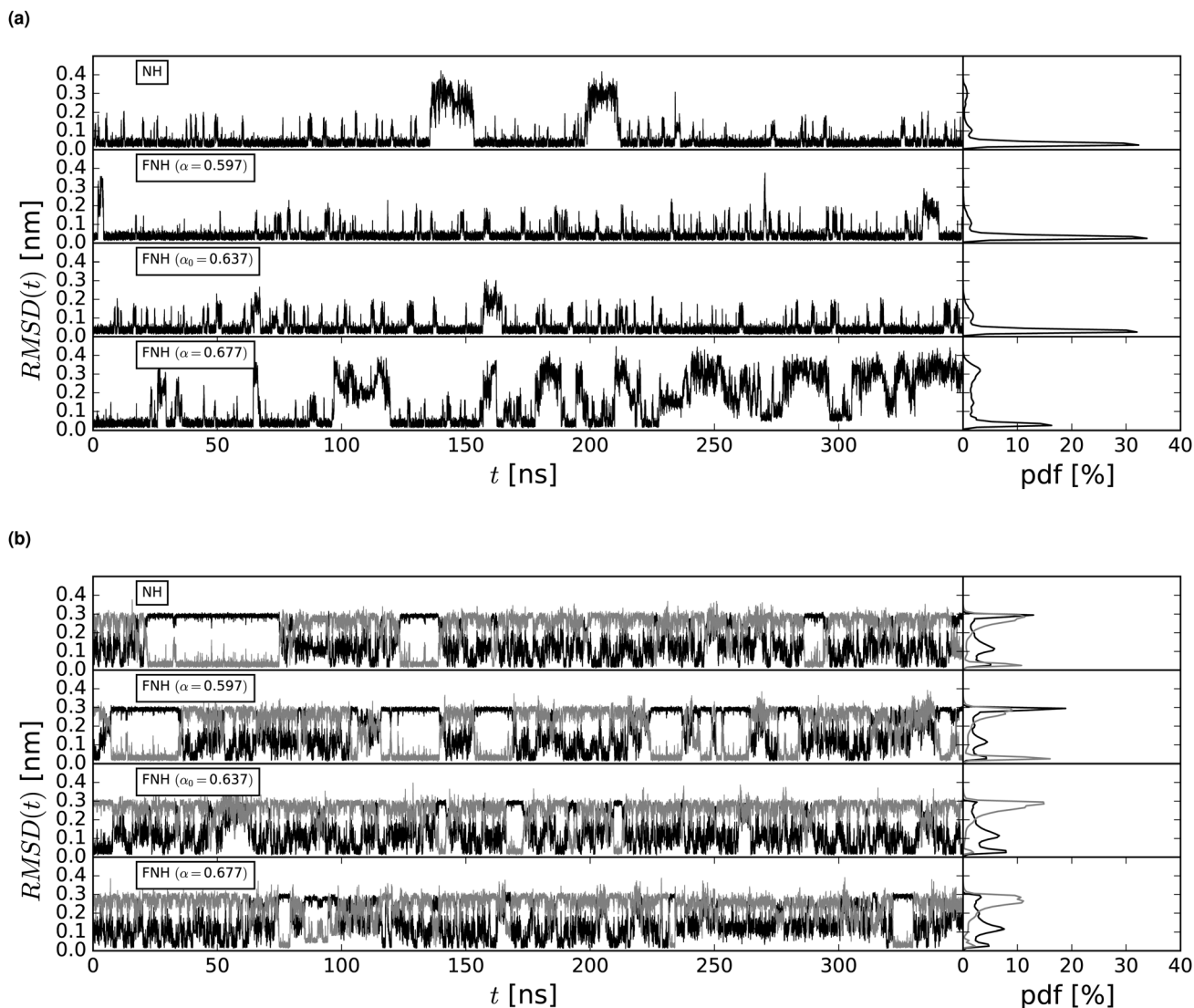

Fig. 5 Results of the simulations of β-peptides 1 and 2 in methanol at 340 K using the canonical Nosé–Hoover (NH) or the fast Nosé–Hoover (FNH) thermostat with different perturbations. (a) Time series (left) and distribution (right) of the atom-positional backbone RMSD with respect to the left-handed 3_{14}-helical fold of peptide 1. Residues 2–6 were used for the RMSD calculation. (b) Time evolution (left) and distribution (right) of the atom-positional backbone RMSD with respect to a right-handed $2.7_{10/12}$-helical fold (black line) and left-handed 3_{14}-helical fold (grey line) of peptide 2. Residues 2–5 were used for the RMSD calculation.

4.2 β-Peptides in methanol

Having analysed the effect of the FNH equations on dynamic properties of homogeneous systems, we investigated the effect of a perturbed tail distribution on the efficiency of phase-space exploration in a next step. The focus was thereby on the variation in the probability distribution of rare events. A good example of such rare events are transitions between folded and unfolded states of peptides, which occur relatively infrequently during typical simulation times. For peptide 1 (Fig. 1), the time evolution of the backbone atom-positional RMSD is shown in Fig. 5a with corresponding probability distributions for the canonical equilibrium and for estimated α_0, $\alpha_0 \pm \Delta\alpha_0$.

The increase of the tail probability of the velocity distributions was found to have a substantial impact on the number of transitions between the folded and unfolded states of peptide 1, which indicates

a faster exploration of the phase space. A similar observation was made for peptide 2 (Fig. 5b). In addition, when suppressing the fast motions using $\alpha = \alpha_0 - \Delta\alpha_0$, peptide 2 remained often trapped in either the left-handed 3_{14}-helical fold (grey line) or the initial right-handed $2.7_{10/12}$ helical fold state (black line). By increasing α, a more balanced distribution of the RMSD values was obtained. These results are in line with ref. 72, where faster folding was observed for the folding of a 16-residue helical peptide when non-canonical modifications of the potential-energy surface according to Tsallis[73,74] were introduced. Time-averages of the thermostat control variables are given in Table 3, with the related time series in Fig. S1 and S2 of the ESI.†

All findings are in agreement with the results of the water system. In particular, for a positive perturbation $\alpha = \alpha_0 + \Delta\alpha_0$,

This journal is © the Owner Societies 2019

Phys. Chem. Chem. Phys., 2019, 21, 6059–6070 | 6067

Table 3 Time averages of the thermostat control variables ξ, ζ with respect to different perturbations α for the two β-peptides in methanol. The sign ($\pm$) of $\langle\zeta\rangle$ indicates whether the probability of fast motions is increased ($-$) or decreased ($+$) with respect to the canonical equilibrium

		$\alpha = \alpha_0 - \Delta\alpha_0 = 0.597$	$\alpha = \alpha_0 = 0.637$	$\alpha = \alpha_0 + \Delta\alpha_0 = 0.677$	Nosé–Hoover
Peptide 1	$\langle\xi\rangle\ [\mathrm{ps}^{-1}]$	-0.168 ± 0.075	-0.012 ± 0.084	0.028 ± 0.113	0.002 ± 0.121
	$\langle\zeta\rangle\ [\mathrm{ps}^{-1}]$	10.402 ± 0.073	0.381 ± 0.131	-0.384 ± 1.267	—
Peptide 2	$\langle\xi\rangle\ [\mathrm{ps}^{-1}]$	-0.161 ± 0.075	-0.008 ± 0.084	0.021 ± 0.104	0.0014 ± 0.1202
	$\langle\zeta\rangle\ [\mathrm{ps}^{-1}]$	10.047 ± 0.072	0.267 ± 0.108	-0.301 ± 1.053	—

$\mathrm{sgn}(\langle\zeta\rangle) = -1$ which indicates an increased probability of the fast degrees of freedom, leading to faster transitions. However, for the estimated quasi-equilibrium α_0, the results are ambiguous. For peptide 1, it appears that the tail probabilities are reduced on the chosen simulation timescale, whereas the opposite was observed for peptide 2, *i.e.* significantly more transitions occurred than in canonical equilibrium. This highlights that it is difficult to determine the quasi-equilibrium $\alpha_0^{\mathrm{exact}}$, which solves eqn (27) exactly, and thus small deviations may already have a significant impact on the dynamics of the system.

5 Conclusion

In this work, a modification of the Nosé–Hoover thermostat equations was introduced, which allows to gently force the system out of canonical equilibrium. Rigorous mathematical properties of the FNH equations were derived, showing that the dynamics of the non-equilibrium system remains close to the canonical equilibrium for reasonable thermostat parameters and a sufficiently high number of degrees of freedom in the system. In more detail, a quasi-equilibrium parameter α was identified, which determines whether tail probabilities are increased or decreased with respect to canonical equilibrium, while the time-averaged temperature is preserved. Simulations showed that it is easily possible to determine an approximative value α_0, which is close to quasi-canonical equilibrium. However, it was found to be non-trivial to control the exact strength of the perturbation, since the tail distributions are highly sensitive to perturbations of α. The effect of the FNH equations for relevant test systems showed that slightly increased tail probabilities substantially effect the occurrence of rare events (*i.e.* folding and unfolding of helical β-peptides). Interestingly, averaged transport properties (*e.g.* self-diffusion and shear-viscosity of water) were mostly preserved for the tested FNH setups. This observation in combination with the well-defined perturbation of the canonical equilibrium points to interesting future applications of the FNH thermostat. In particular, the well preserved shear viscosity potentially offers the opportunity to reduce critical slowing down in the vicinity of critical points by perturbations of the tail probabilities. Moreover, it should also be possible to obtain canonical ensemble averages from FNH simulations by using appropriate ensemble reweighting techniques. Hence, equilibrium simulations could be accelerated by using the non-equilibrium FNH thermostat. In the future, we plan to investigate corresponding reweighting techniques and applications on interesting chemical systems.

Conflicts of interest

There are no conflicts to declare.

Appendix: Thermostat masses Q_1, Q_2

By using the assumptions for the thermostat masses stated in Section 2, one can decouple and simplify the FNH equations as follows:

$$\ddot{\xi} = \frac{1}{Q_1}\sum_i^N \frac{2p_i\dot{p}_i}{m_i} \approx \frac{1}{Q_1}\sum_i^N \frac{2p_i^{(\xi)}\dot{p}_i^{(\xi)}}{m_i}$$

$$= -\frac{1}{Q_1}\sum_i^N \left(\xi\frac{2\left(p_i^{(\xi)}\right)^2}{m_i}\right) \overset{\langle\cdot\rangle}{\approx} -\frac{2N_{\mathrm{df}}k_{\mathrm{B}}T_0}{Q_1}\xi \qquad (33)$$

and

$$\ddot{\zeta} = \frac{1}{3Q_2}\left[\frac{\sum_{i=1}^N \frac{4p_i^3\dot{p}_i}{m_i^3}}{\sum_{i=1}^N v_i^2} - \frac{\sum_{i=1}^N \frac{p_i^4}{m_i^3}\left(\sum_{i=1}^N \frac{2p_i\dot{p}_i}{m_i^2}\right)}{\left(\sum_{i=1}^N v_i^2\right)^2}\right]$$

$$\approx \frac{1}{3Q_2}\left[\frac{\sum_{i=1}^N \frac{4\left(p_i^{(\zeta)}\right)^3\dot{p}_i^{(\zeta)}}{m_i^3}}{\sum_{i=1}^N \left(v_i^{(\zeta)}\right)^2} - \frac{\sum_{i=1}^N \frac{\left(p_i^{(\zeta)}\right)^4}{m_i^3}\left(\sum_{i=1}^N \frac{2p_i^{(\zeta)}\dot{p}_i^{(\zeta)}}{m_i^2}\right)}{\left(\sum_{i=1}^N \left(v_i^{(\zeta)}\right)^2\right)^2}\right] \qquad (34)$$

$$\overset{\mathrm{eqn\ (10)}}{\approx} -\frac{2\alpha k_{\mathrm{B}}T_0}{Q_2}\zeta,$$

which leads to the simple expressions given in eqn (29) and (30).

Acknowledgements

References

1 P. H. Hünenberger, *Adv. Comp. Sim.*, Springer, 2005, pp. 105–149.
2 H. C. Andersen, *J. Chem. Phys.*, 1980, **72**, 2384–2393.

6068 | *Phys. Chem. Chem. Phys.*, 2019, **21**, 6059–6070

This journal is © the Owner Societies 2019

3 S. Nosé, *J. Chem. Phys.*, 1984, **81**, 511–519.

4 W. G. Hoover, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1985, **31**, 1695.

5 G. J. Martyna, M. L. Klein and M. E. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635–2643.

6 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.

7 H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola and J. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

8 P. Minary, G. J. Martyna and M. E. Tuckerman, *J. Chem. Phys.*, 2003, **118**, 2510–2526.

9 B. Leimkuhler and C. Matthews, *Molecular Dynamics*, Springer, 2016.

10 J. J. Biesiadecki and R. D. Skeel, *J. Comput. Phys.*, 1993, **109**, 318–328.

11 Q. Ma, J. A. Izaguirre and R. D. Skeel, *SIAM J. Sci. Comput.*, 2003, **24**, 1951–1973.

12 D. Sidler, M. Lehner, S. Frasch, M. Cristófol-Clough and S. Riniker, *F1000Research*, 2019, **7**, 1745.

13 J. A. Morrone, T. E. Markland, M. Ceriotti and B. J. Berne, *J. Chem. Phys.*, 2011, **134**, 014103.

14 E. Barth and T. Schlick, *J. Chem. Phys.*, 1998, **109**, 1617–1632.

15 X. Qian and T. Schlick, *J. Chem. Phys.*, 2002, **116**, 5971–5983.

16 P. Minary, M. E. Tuckerman and G. J. Martyna, *Phys. Rev. Lett.*, 2004, **93**, 150201.

17 B. Leimkuhler, D. T. Margul and M. E. Tuckerman, *Mol. Phys.*, 2013, **111**, 3579–3594.

18 D. T. Margul and M. E. Tuckerman, *J. Chem. Theory Comput.*, 2016, **12**, 2170–2180.

19 J. A. Izaguirre, D. P. Catarello, J. M. Wozniak and R. D. Skeel, *J. Chem. Phys.*, 2001, **114**, 2090–2098.

20 M. Ceriotti, G. Bussi and M. Parrinello, *Phys. Rev. Lett.*, 2009, **102**, 020601.

21 M. Ceriotti, G. Bussi and M. Parrinello, *J. Chem. Theory Comput.*, 2010, **6**, 1170–1180.

22 E. J. Barth, B. B. Laird and B. J. Leimkuhler, *J. Chem. Phys.*, 2003, **118**, 5759–5768.

23 N. Shuichi, *Prog. Theor. Phys. Suppl.*, 1991, **103**, 1–46.

24 X. Daura, W. F. van Gunsteren and A. E. Mark, *Proteins*, 1999, **34**, 269–280.

25 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angew. Chem., Int. Ed.*, 1999, **38**, 236–240.

26 X. Daura, K. Gademann, H. Schäfer, B. Jaun, D. Seebach and W. F. van Gunsteren, *J. Am. Chem. Soc.*, 2001, **123**, 2393–2404.

27 W. F. van Gunsteren, R. Bürgi, C. Peter and X. Daura, *Angew. Chem., Int. Ed.*, 2001, **40**, 351–355.

28 X. Daura, A. Glättli, P. Gee, C. Peter and W. F. Van Gunsteren, *Adv. Protein Chem.*, 2002, **62**, 341–360.

29 W. F. van Gunsteren and Z. Gattin, in *Foldamers: Structure, Properties and Application* ed. S. Hecht and I. Huc, 2007, pp. 173–192.

30 Z. Lin, N. Schmid and W. F. van Gunsteren, *Mol. Phys.*, 2011, **109**, 493–506.

31 W. Huang, Z. Lin and W. F. van Gunsteren, *J. Chem. Theory Comput.*, 2011, **7**, 1237–1243.

32 W. Huang, S. Riniker and W. F. van Gunsteren, *J. Chem. Theory Comput.*, 2014, **10**, 2213–2223.

33 M. Tuckerman, *Statistical mechanics: Theory and molecular simulation*, Oxford University Press, 2010.

34 A. Bulgac and D. Kusnezov, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1990, **42**, 5045.

35 D. Kusnezov, A. Bulgac and W. Bauer, *Ann. Phys.*, 1990, **204**, 155–185.

36 B. Leimkuhler, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2010, **81**, 026703.

37 M. E. Tuckerman, C. J. Mundy and G. J. Martyna, *EPL*, 1999, **45**, 149.

38 M. E. Tuckerman, Y. Liu, G. Ciccotti and G. J. Martyna, *J. Chem. Phys.*, 2001, **115**, 1678–1702.

39 Z. Jia and B. Leimkuhler, *ESAIM: Math. Modell. Numer. Anal.*, 2007, **41**, 333–350.

40 T. Soddemann, B. Dünweg and K. Kremer, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2003, **68**, 046702.

41 H. J. Berendsen, J. P. Postma, W. F. van Gunsteren and J. Hermans, *Intermolecular forces*, Springer, 1981, pp. 331–342.

42 I. G. Tironi and W. F. van Gunsteren, *Mol. Phys.*, 1994, **83**, 381–403.

43 N. Schmid, C. D. Christ, M. Christen, A. P. Eichenberger and W. F. van Gunsteren, *Comput. Phys. Commun.*, 2012, **183**, 890–903.

44 S. Riniker, C. D. Christ, H. S. Hansen, P. H. Hünenberger, C. Oostenbrink, D. Steiner and W. F. van Gunsteren, *J. Phys. Chem. B*, 2011, **115**, 13570–13577.

45 N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark and W. F. van Gunsteren, *Eur. Biophys. J.*, 2011, **40**, 843.

46 Z. Lin and W. F. van Gunsteren, *J. Comput. Chem.*, 2013, **34**, 2796–2805.

47 R. W. Hockney, *Potential calculation and some applications*, Langley research center, hampton, va. technical report, 1970.

48 J.-P. Ryckaert, G. Ciccotti and H. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.

49 I. G. Tironi, R. Sperb, P. E. Smith and W. F. van Gunsteren, *J. Chem. Phys.*, 1995, **102**, 5451–5459.

50 W. R. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger and W. F. van Gunsteren, *J. Phys. Chem. A*, 1999, **103**, 3596–3607.

51 W. M. Haynes, *CRC Handbook of Chemistry and Physics*, 2016.

52 R. Walser, A. E. Mark, W. F. van Gunsteren, M. Lauterbach and G. Wipff, *J. Chem. Phys.*, 2000, **112**, 10450–10459.

53 D. R. Lide, *Handbook of Chemistry and Physics*, CRC Press/Taylor & Francis, Boca Raton, FL, 88th edn, 2007–2008.

54 A. P. Eichenberger, J. R. Allison, J. Dolenc, D. P. Geerke, B. A. Horta, K. Meier, C. Oostenbrink, N. Schmid, D. Steiner, D. Wang and W. F. van Gunsteren, *J. Chem. Theory Comput.*, 2011, **7**, 3379–3390.

55 M. P. Allen and D. J. Tildesley, *Computer simulation of liquids*, Oxford University Press, 2017.

56 E. J. Maginn, R. A. Messerly, D. J. Carlson, D. R. Roe and J. R. Elliott, *Living J. Comput. Mol. Sci.*, 2019, **1**, 6324.

This journal is © the Owner Societies 2019

Phys. Chem. Chem. Phys., 2019, **21**, 6059–6070 | **6069**

57 P. E. Smith and W. F. van Gunsteren, *Chem. Phys. Lett.*, 1993, **215**, 315–318.

58 M. Neumann, *Mol. Phys.*, 1983, **50**, 841–858.

59 R. Dimelow and A. Masters, *Mol. Simul.*, 2007, **33**, 1165–1166.

60 Y. Wu, H. L. Tepper and G. A. Voth, *J. Chem. Phys.*, 2006, **124**, 024503.

61 A. Glättli, X. Daura and W. F. van Gunsteren, *J. Chem. Phys.*, 2002, **116**, 9811–9828.

62 D. van der Spoel, P. J. van Maaren and H. J. C. Berendsen, *J. Chem. Phys.*, 1998, **108**, 10220–10230.

63 D. Van Der Spoel and P. J. van Maaren, *J. Chem. Theory Comput.*, 2006, **2**, 1–11.

64 K. Watanabe and M. L. Klein, *Chem. Phys.*, 1989, **131**, 157–167.

65 P. Mark and L. Nilsson, *J. Phys. Chem. A*, 2001, **105**, 9954–9960.

66 M. W. Mahoney and W. L. Jorgensen, *J. Chem. Phys.*, 2001, **114**, 363–366.

67 G. Guevara-Carrion, J. Vrabec and H. Hasse, *J. Chem. Phys.*, 2011, **134**, 074508.

68 B. Hess, *J. Chem. Phys.*, 2002, **116**, 209–217.

69 S. Roy and S. K. Das, *Eur. Phys. J. E: Soft Matter Biol. Phys.*, 2015, **38**, 132.

70 S. Roy and S. K. Das, *Europhys. Lett.*, 2011, **94**, 36001.

71 G. Jung and F. Schmid, *J. Chem. Phys.*, 2016, **144**, 204104.

72 Y. Pak and S. Wang, *J. Chem. Phys.*, 1999, **111**, 4359–4361.

73 C. Tsallis, *J. Stat. Phys.*, 1988, **52**, 479–487.

74 E. M. Curado and C. Tsallis, *J. Phys. A: Gen. Phys.*, 1991, **24**, L69.

6070 | *Phys. Chem. Chem. Phys.*, 2019, **21**, 6059–6070

This journal is © the Owner Societies 2019