

# PCCP

Accepted Manuscript



This article can be cited before page numbers have been issued, to do this please use: C. Scherer and D. Andrienko, *Phys. Chem. Chem. Phys.*, 2018, DOI: 10.1039/C8CP00746B.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [author guidelines](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the ethical guidelines, outlined in our [author and reviewer resource centre](#), still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

# Understanding three-body contributions to coarse-grained force fields

Christoph Scherer<sup>1, a)</sup> and Denis Andrienko<sup>1, b)</sup>

*Max Planck Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany*

(Dated: August 9, 2018)

Coarse-graining is a systematic reduction of the number of degrees of freedom used to describe a system of interest. Coarse-graining can be thought of as a projection on the coarse-grained degrees of freedom and is therefore dependent on the number and type of basis functions used to represent the coarse-grained force field. We show that many-body extensions of the coarse-grained force field can result in substantial changes of the two-body interactions, making them much more attractive at short distances. This interplay can be alleviated by first parametrizing the two-body potential and then fitting the additional three-body contribution to the residual forces. The approach is illustrated on liquid water where three-body interactions are essential to reproduce the structural properties, and liquid methanol where two-body interactions are sufficient to reproduce the main structural features of the atomistic system. Furthermore, we demonstrate that the structural and thermodynamic accuracy of the coarse-grained models can be controlled by varying the magnitude of the three-body interactions. Our findings motivate basis set extensions which separate the many-body contributions of different order.

## I. INTRODUCTION

Coarse-graining (CG) is a systematic way to reduce the number of degrees of freedom describing a specific physical system. By combining coherently moving atoms of a molecule into a single interaction site one can increase the simulation times by ten to hundred times compared to the atomistic simulation. The computational cost is reduced in several ways: First, the coarse-grained system has less degrees of freedom. Second, smoother interaction potentials result in a smaller friction and, hence, faster dynamics. Third, one can use a bigger simulation timestep in the integration algorithm.

To coarse-grain a specific system, we need to choose coarse-grained degrees of freedom, relate them to the fine-grained description, identify a merit function which quantifies the difference between the fine- and coarse-grained representations, and optimize the coarse-grained potential energy surface (PES). Consistency between the coarse-grained and the fine-grained models can be defined in terms of consistency of the equilibrium probability densities resulting in unique expressions for the CG masses and interaction potential, i.e. the many-body potential of mean force (PMF).<sup>1</sup> Evaluating this potential is as computationally demanding as propagating the fine-grained system and, in practice, it is approximated using a limited set of basis functions chosen to represent the interactions at the CG level of resolution.

Several methods have been developed that approximate the many-body PMF, targeting specific properties of the underlying atomistic system. For example, iterative Boltzmann inversion (IBI)<sup>2</sup> and inverse Monte Carlo (IMC)<sup>3</sup> schemes target structural pair distribution functions. It can be shown that these methods minimize the

relative entropy between the CG and atomistic ensembles.<sup>4</sup> An alternative route is to match the forces of the CG system to those of the atomistic description, employing force-matching (FM) or multiscale coarse-graining (MS-CG).<sup>1,5,6</sup> This approach corresponds to projecting the many-body mean force, i.e., the negative gradient of the many-body PMF, into the space of force fields defined by the CG basis set. It allows to systematically increase the accuracy of the approximation of the many-body PMF by expanding the basis set. FM and structural coarse-graining can be connected via Yvon-Born-Green theory.<sup>7,8</sup>

The structural accuracy of the force-matched CG model depends on the CG mapping scheme<sup>9</sup> and on the ability of the CG basis set to capture the relevant many-body correlations of the mapped ensemble.<sup>8</sup> In some cases, force-matched pair potentials reproduce the atomistic pair correlation functions, as for one-site models of liquid methanol.<sup>10,11</sup> However, for liquids with a strong local orientational order, with water being a prominent example, CG pair potentials fail to fully capture the structural correlations of the underlying atomistic system.<sup>12–17</sup> Even when the pair potential is able to reproduce the atomistic pair correlation functions, thermodynamic properties are usually not reproduced correctly, for example the cohesive energy can be significantly underestimated.<sup>18</sup> This implies that an extension of the basis set is necessary to describe the many-body correlations more accurately.<sup>19–23</sup>

Indeed, coarse-grained models can be improved by including a local density dependent potential<sup>24–26</sup> or a short-ranged nonbonded three-body potential,<sup>12–14,27–30</sup> such as the Stillinger-Weber<sup>31</sup> (SW) potential. These basis set extensions lead to a significant change of the two-body interactions.<sup>12,13,24,25</sup> For example, an one-site CG SPC/E water model with the SW potential has a deep attractive well at 2.5 nm, as shown in Fig. 1. In this paper we demonstrate that the change of the pair

<sup>a)</sup>Electronic mail: scherer@mpip-mainz.mpg.de

<sup>b)</sup>Electronic mail: denis.andrienko@mpip-mainz.mpg.de



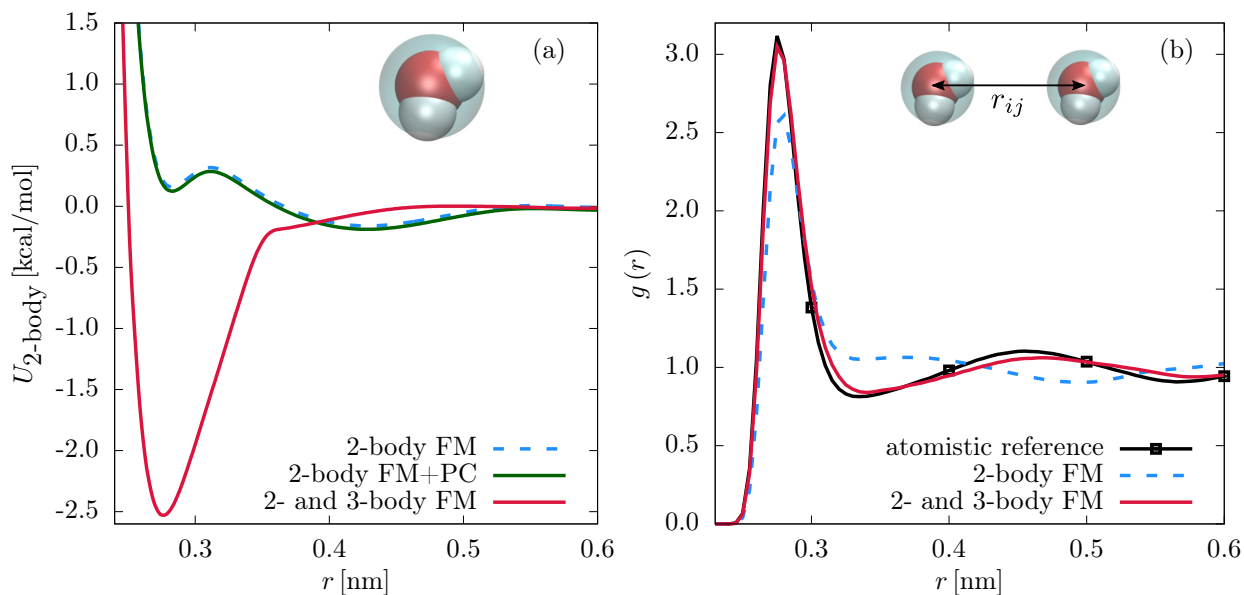


Figure 1. The pair potentials (a) and the radial distribution functions (b) of coarse-grained SPC/E water. The following parametrizations are shown: FM with a (tabulated) pair potential only (2-body FM), and FM with a tabulated pair potential and the short-range three-body SW potential (2- and 3-body FM). For the pair potential, we also show the curve with the additional pressure correction (2-body FM+PC).

potential can be tracked back to the non-orthogonality of the three-body and the two-body interactions. In addition, we propose a parametrization scheme which reduces the three-body contribution to the two-body PMF or, in other words, decouples the two- and three-body interaction terms.

## II. METHODS

### A. Atomistic simulations

For atomistic simulations we employ the GROMACS<sup>32</sup> package, version 5.1. We use the SPC/E water model<sup>33</sup> and the OPLS-AA force field<sup>34,35</sup> for methanol. Boxes of 1000 molecules are first equilibrated for 10 ns in the NPT ensemble at  $T = 300$  K and  $p = 1.0$  bar, followed by 10 ns NVT production runs at the average density of the preceding NPT simulations,  $\rho = 0.998$  g/cm<sup>3</sup> (water) and  $\rho = 0.776$  g/cm<sup>3</sup> (methanol). These values are very close to the experimental densities at normal conditions,  $\rho = 0.9971$  g/cm<sup>3</sup> (water)<sup>36</sup> and  $\rho = 0.7872$  g/cm<sup>3</sup> (methanol).<sup>37</sup> To integrate the equations of motion, we use a stochastic dynamics algorithm<sup>38</sup> in combination with a Berendsen barostat<sup>39</sup> with a time constant of 1 ps, the compressibility parameter of water ( $4.5 \cdot 10^{-5}$  1/bar), and a time step of 1 fs. Electrostatic interactions are treated with a smooth particle mesh Ewald method<sup>40</sup> with cubic interpolation, a grid spacing of 0.12 nm and an Ewald accuracy parameter of  $10^{-5}$ . We use a short-range cutoff of 1.2 nm and long-range dispersion corrections for energy and pressure. The OPLS-AA force field is used

without constraints.

### B. Coarse-graining procedure

We replace one water or methanol molecule with a CG bead positioned at its center of mass,

$$\mathbf{R}_i = \sum_{\alpha \in i} w_{\alpha} \mathbf{r}_{\alpha}, \quad (1)$$

where  $w_{\alpha} = m_{\alpha} / (\sum_{\alpha \in i} m_{\alpha})$  are the weights of each atom of molecule  $i$ . The mass of the CG bead is the sum of the atomistic masses:  $M_i = 18.0154$  amu (water) and  $M_i = 32.0374$  amu (methanol).

To parametrize the FM potentials, the reference force on each CG bead  $i$  is calculated as the sum of the atomistic forces,

$$\mathbf{f}_i^{\text{ref}} = \sum_{\alpha \in i} \mathbf{f}_{\alpha}. \quad (2)$$

The CG representation of the force is then determined by solving:

$$\mathbf{f}_{il}^{\text{CG}}(g_1, \dots, g_M) = \mathbf{f}_{il}^{\text{ref}}, \quad i = 1 \dots N, \quad l = 1 \dots L. \quad (3)$$

Here,  $g_1, \dots, g_M$  are the coefficients of the CG force field basis functions,  $N$  is the number of CG beads, and  $L$  is the number of simulation snapshots per block. We use  $N = 1000$  and  $L = 20$ . Given that the CG force field basis functions  $\mathbf{f}_{il}^{\text{CG}}$  depend linearly on the parameters  $g_1, \dots, g_M$ , eq. 3 is a set of  $N \times L$  overdetermined linear



equations ( $M < N \times L$ ) which we solve with a constrained least-squares solver.

All coarse-graining algorithms are implemented into the VOTCA-CSG package.<sup>18</sup>

### C. Coarse-grained potentials

Nonbonded two-body potentials are represented by cubic splines which depend linearly on the coefficients  $g_1, \dots, g_M$ . A set of  $K$  grid points leads to  $M = 2K$  spline coefficients where  $K$  of the coefficients are fixed due to constraints guaranteeing the continuity of the first derivatives.

The SW three-body potential,

$$U = \sum_{i,j \neq i, k > j} f^{(3b)}(\theta_{ijk}) \exp\left(\frac{\gamma_{ij}\sigma_{ij}}{r_{ij} - a_{ij}\sigma_{ij}} + \frac{\gamma_{ik}\sigma_{ik}}{r_{ik} - a_{ik}\sigma_{ik}}\right), \quad (4)$$

is used as a short-range three-body potential, where  $i$  is the index of the central atom, and  $j$  and  $k$  are the other two atom indices of a triplet of atoms with an angular interaction term  $f^{(3b)}(\theta_{ijk})$ . We do not limit ourselves to an analytic expression of  $f^{(3b)}(\theta_{ijk})$  as in the original SW potential, but allow for a flexible angular dependence of  $f^{(3b)}(\theta_{ijk})$ .

The FM parametrization is implemented in the VOTCA-CSG package<sup>18</sup> using a cubic spline representation of  $f^{(3b)}(\theta)$  with  $K$  grid points and a linear dependence on the  $M = 2K$  spline coefficients. Treating the two exponential terms,  $\exp\left(\frac{\gamma\sigma}{r-a\sigma}\right)$ , as prefactors yields a linear set of equations which can be solved by a constrained least-squares solver as in the case of the pair forces. To do so, the remaining parameters  $a_{ij}$ ,  $a_{ik}$ ,  $\sigma_{ij}$ ,  $\sigma_{ik}$ ,  $\gamma_{ij}$ , and  $\gamma_{ik}$ , have to be set beforehand. For one CG bead type, as in our case, the number of parameters reduces to  $a_{ij} = a_{ik} = a$ ,  $\sigma_{ij} = \sigma_{ik} = \sigma$ , and  $\gamma_{ij} = \gamma_{ik} = \gamma$ . Here,  $a$  is the three-body short-range cutoff which we choose in a way that the three-body potential is fully switched on in the first coordination shell. Setting  $\sigma = 1$ , the parameter  $\gamma$  controls the steepness of this switching on. In the limit of  $\gamma \rightarrow 0$ , the SW potential is instantaneously switched on. For water, we refer to the cutoff radius of Ref. 13, namely  $a = 0.37$  nm. In the case of liquid methanol, we choose  $a = 0.45$  nm. We systematically vary  $\gamma$  until the structural fit does not improve anymore. The optimal value, in both cases, is  $\gamma = 0.08$  nm. This means that  $U^{\text{SW}}$  is fully switched on in the first and switched off within the second coordination shell. For more technical details on the FM implementation we refer to the supporting information.

### D. CG simulations

Coarse-grained MD simulations are performed at the atomistic density using the LAMMPS package.<sup>41</sup> Two-

body CG interactions are treated as linear interpolation tables with a short-range cutoff of 1.2 nm. We extend the LAMMPS implementation of the Stillinger-Weber interaction potential to read in tabulated angular dependent potentials. We employ a time step of 1 fs and a chain of three Nose-Hoover thermostats<sup>42,43</sup> with a damping parameter of 200 fs when integrating the equations of motion<sup>44</sup> for 10 ns in the NVT ensemble. When simulating in the NPT ensemble, we apply an additional chain of three Nose-Hoover barostats with a damping parameter of 1000 fs.

### E. Calculation of the pair PMF

To separate the two-body and three-body contributions of the CG interaction potential, we evaluate the pair PMF between two coarse-grained sites:

$$U_{\text{PMF}}(r) = - \int_0^r F_{\text{rad}}(r') dr'. \quad (5)$$

Here,  $F_{\text{rad}}(r) = \langle \mathbf{F}_j \cdot \hat{\mathbf{r}}_{ij} \rangle \equiv \langle -\mathbf{F}_i \cdot \hat{\mathbf{r}}_{ij} \rangle$  is the projection of the total CG force acting on a CG bead  $i$  or  $j$  onto the unit distance vector  $\hat{\mathbf{r}}_{ij}$  connecting this pair of beads, averaged over all CG bead pairs with distance  $r$ . Note that the total pair  $U_{\text{PMF}}(r)$  is equivalent to the radial distribution function, since  $g(r) \sim \exp(-U_{\text{PMF}}(r)/k_{\text{B}}T)$ . When calculating  $U_{\text{PMF}}(r)$  for the atomistic simulations,  $\mathbf{F}_i$  and  $\mathbf{F}_j$  are evaluated as a weighted sum of all atomistic forces acting on each CG bead  $i$  and  $j$ , according to eq. (2).

$F_{\text{rad}}(r)$  and, thus,  $U_{\text{PMF}}(r)$  can be evaluated separately for the two-body and three-body interactions by a rerun of the CG trajectory and switching on and off the individual force contributions. That gives the contributions of the two-body and three-body interactions to the total pair PMF, including the effects of correlated forces from the environment. This is comparable to calculating the decomposition of the pair PMF in terms of the MS-CG G-matrix.<sup>45</sup>

### F. Inner product of force field functions

To quantify the degree of correlation between two-body and three-body force field functions, we define an inner product as<sup>7</sup>

$$\mathbf{F}^{(2)} \cdot \mathbf{F}^{(3)} = \int d\mathbf{R} p_{\mathbf{R}}(\mathbf{R}) \sum_{i=1}^N \mathbf{F}_i^{(2)}(\mathbf{R}) \cdot \mathbf{F}_i^{(3)}(\mathbf{R}). \quad (6)$$

Here,  $\mathbf{F}_i^{(\xi)}$  is the total force on bead  $i$ ,  $\xi = 2, 3$  corresponds to the two- and three-body contributions, and  $p_{\mathbf{R}}(\mathbf{R})$  is the CG configuration distribution. We evaluate a normalized inner product of the two-body and three-body contributions to the total force on each CG bead



as:

$$p = \frac{\mathbf{F}^{(2)} \cdot \mathbf{F}^{(3)}}{|\mathbf{F}^{(2)}| |\mathbf{F}^{(3)}|}. \quad (7)$$

The normalized inner product corresponds to the ensemble average of the inner product of the two-body and three-body contribution on each CG bead and can be converted into an angle:

$$\Phi = \arccos(p) = \arccos\left(\frac{\mathbf{F}^{(2)} \cdot \mathbf{F}^{(3)}}{|\mathbf{F}^{(2)}| |\mathbf{F}^{(3)}|}\right). \quad (8)$$

### III. RESULTS AND DISCUSSION

We choose one-site models of liquid water and methanol as test systems where two-body basis functions fail (water) and succeed (methanol) to accurately reproduce the atomistic pair correlation functions. We employ force-matching (FM) to parametrize the tabulated two-body in combination with the three-body Stillinger-Weber (SW) interactions with a flexible angular dependence.

#### A. Structural Properties

First, we compare the structural correlation functions of the different models. As mentioned above and shown in Fig. 1(b), the CG model of SPC/E water with the three-body SW potential (2- and 3-body FM) reproduces the radial distribution function (RDF) of the atomistic reference better than the model with only pairwise interactions (2-body FM). This also holds for the angular distribution function, as shown in Fig. 5(b). For liquid methanol, the two-body FM model already reproduces the RDF nearly perfectly. Therefore, there is no visible change upon including the three-body SW term, as shown in Fig. 6 of the supporting information. At the same time, the angular distribution function significantly improves, as shown in Fig. 2(b).

Next, we focus on the interplay of the two-body and three-body interactions. In Fig. 3, we show the total pair PMF (red solid lines), the two-body (grey dashed lines) and three-body (grey dotted lines) contributions to this potential of mean force for water and methanol. In addition, we show the pair PMF generated by the two-body FM parametrizations as blue crosses. The evaluation of the pair PMF and the decomposition are described in subsection II E. To some extent, the attractive forces of the two-body potential within the first coordination shell are compensated by correlation effects due to two-body interactions with surrounding molecules. This can be seen in the reduced magnitude of the two-body contribution to the pair PMF compared to the direct pair potential (compare Figs. 1(a) and 3(a) for water and 2(a)

and 3(b) for methanol). However, most of the compensation of the two-body attractions comes from short-range repulsive forces of the three-body interactions.

The strong coupling of the two-body and three-body interactions motivates a procedure similar to the subtraction of the Coulomb force from the atomistic reference.<sup>46</sup> This can be viewed as a Gram-Schmidt orthogonalization scheme which reduces the coupling between different terms in the CG force field. First, we obtain the CG pair potential by two-body FM. In a second step, we determine the residual force  $\Delta \mathbf{f}_{\text{FM}}$  acting on each CG bead by subtracting the two-body force from the total reference force,  $\mathbf{f}^{\text{ref}}$ , given by eq. 2:

$$\Delta \mathbf{f}_{\text{FM}} = \mathbf{f}^{\text{ref}} - \mathbf{f}^{\text{FM}}. \quad (9)$$

We now parametrize the three-body SW potential using the residual instead of the reference force. The main effect of the residual force parametrization is a vertical shift of the angular part of the SW potential  $f^{(3b)}(\theta)$  to smaller values (see Figs. 2 and 4 of the supporting information). The decompositions of the pair PMFs for the new parametrizations (see Fig. 4, orange solid and dashed lines with dots) clearly show significantly reduced three-body contributions and less attractive two-body contributions to the pair PMFs compared to the unconstrained parametrizations (red solid and dashed lines).

It is essential to check how the residual force parametrizations perform in terms of the liquid structure. In Fig. 5, we compare the RDFs and angular distribution functions of CG beads within the first coordination shell for liquid water. The angular distribution functions of the two different parametrization schemes are very similar. The constrained parametrization (2-b FM and 3-b  $\Delta \mathbf{f}_{\text{FM}}$ ) also improves the RDF compared to the two-body FM result. For liquid methanol, the same is observed for the angular distribution functions (see Fig. 2(b)). Regarding the RDF, the first peak is slightly suppressed in case of the residual force parametrization (2-body FM and 3-body  $\Delta \mathbf{f}_{\text{FM}}$ , Fig. 6 of the supporting information). A similar effect has been observed when adding a many-body local density-dependent potential to a two-body pair potential for liquid methanol.<sup>26</sup>

As a complementary information to the decomposition of the pair PMF into the two- and three-body interaction terms, we evaluate the inner product  $p$  and average angle  $\Phi$  between the two-body and three-body forces on each CG bead (see eqns. (7) and (8)). We find that the constrained parametrization scheme indeed reduces the inner products from  $p \approx -0.5$  ( $\Phi \approx 120^\circ$ ) to  $p \approx -0.03$  ( $\Phi \approx 92^\circ$ ) (water) and  $p \approx -0.6$  ( $\Phi \approx 125^\circ$ ) to  $p \approx 0.05$  ( $\Phi \approx 87^\circ$ ) (methanol).

#### B. Thermodynamics

We now assess the performance of the CG models in terms of thermodynamic quantities. In particular, we evaluate the virial pressure  $P$  which should be equal to



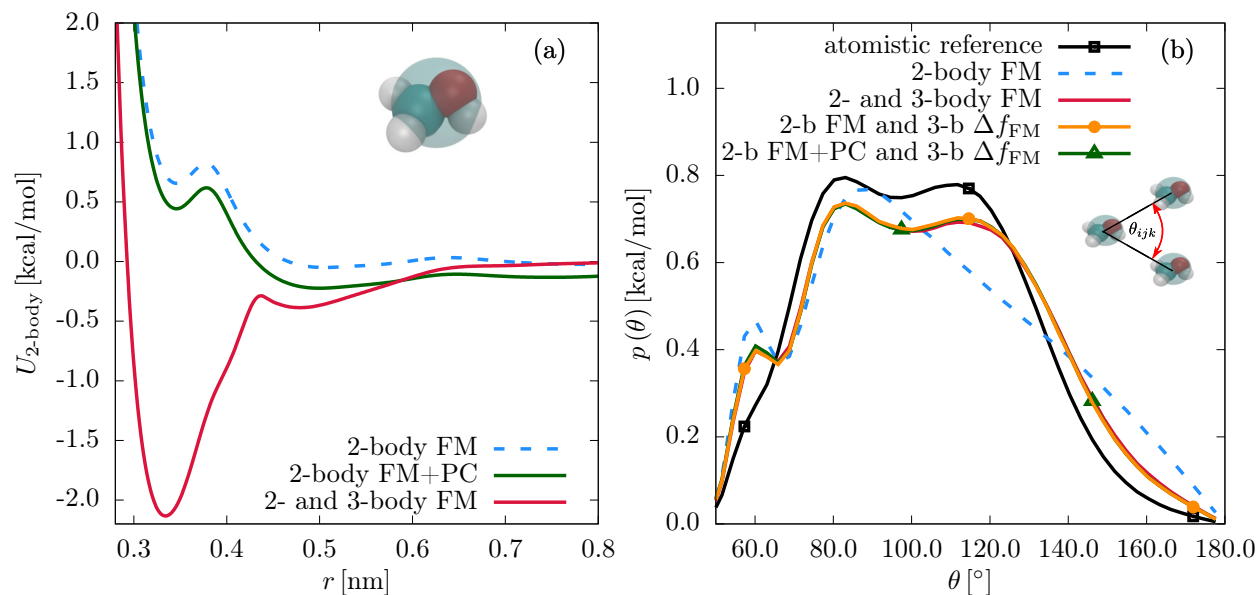


Figure 2. The pair potentials (a) and the angular distribution functions (b) of coarse-grained liquid methanol. Four parametrizations are shown: FM with a (tabulated) pair potential only (2-body FM), FM with a tabulated pair potential and the short-range three-body SW potential (2- and 3-body FM), three-body FM using the residual force of the two-body FM potential (2-b FM and 3-b  $\Delta f_{\text{FM}}$ ), as well as, three-body FM using the residual force of the two-body FM potential with additional pressure correction (2-b FM+PC and 3-b  $\Delta f_{\text{FM}}$ ). The second pair potential (2-body FM+PC) refers to the pair potential with the additional pressure correction. The angular distribution function is calculated using a cutoff of 0.38 nm.

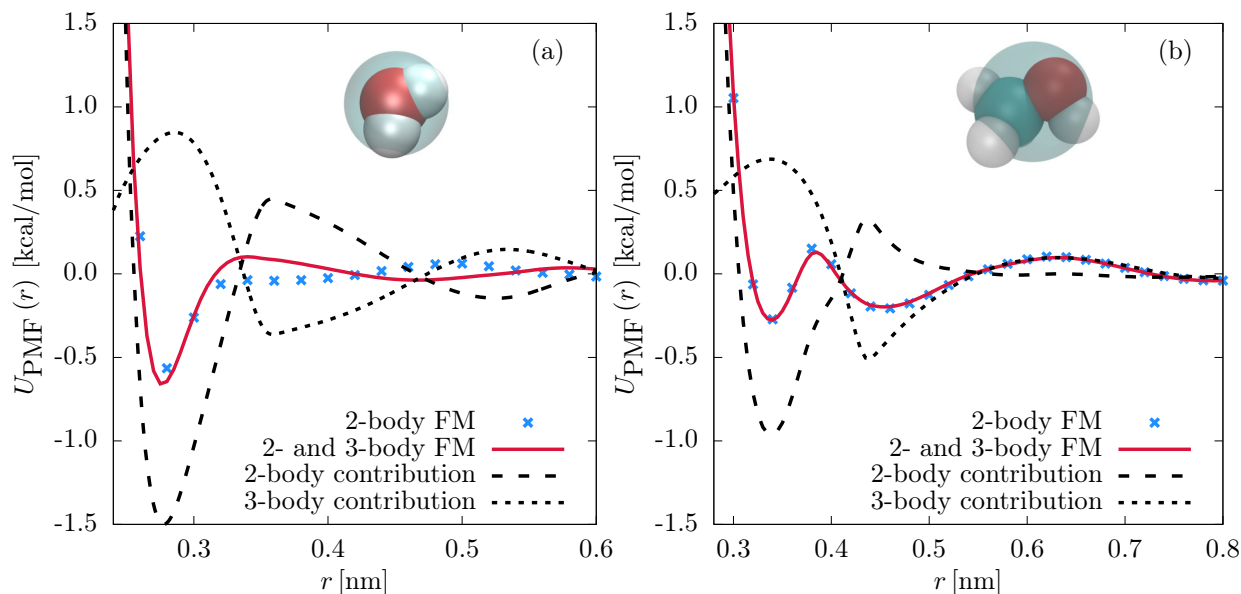


Figure 3. Decomposition of the CG two-body PMF for (a) SPC/E water and (b) liquid methanol. The labels 2-body FM, 2- and 3-body FM refer to the two different CG parametrizations: FM with a (tabulated) pair potential only, and concurrent FM with tabulated pair and short-range three-body SW potential. The 2-body and 3-body contributions refer to the decomposition of the concurrent two- and three-body FM parametrization.

the ambient pressure when simulated at the density of the atomistic model, as well as the enthalpy of vaporization  $\Delta H = H_{\text{gas}} - H_{\text{liq}}$ . Note that we do not explicitly conduct CG gas phase simulations and assume an ideal CG gas.<sup>28,37</sup> In addition, we take into account the ac-

tual average pressure of the liquid phase CG simulations. Calculation details can be found in the supporting information.

So far, all structural properties referred to a SW interaction parameter of  $\gamma = 0.08$  nm (see eqn. (4)), which



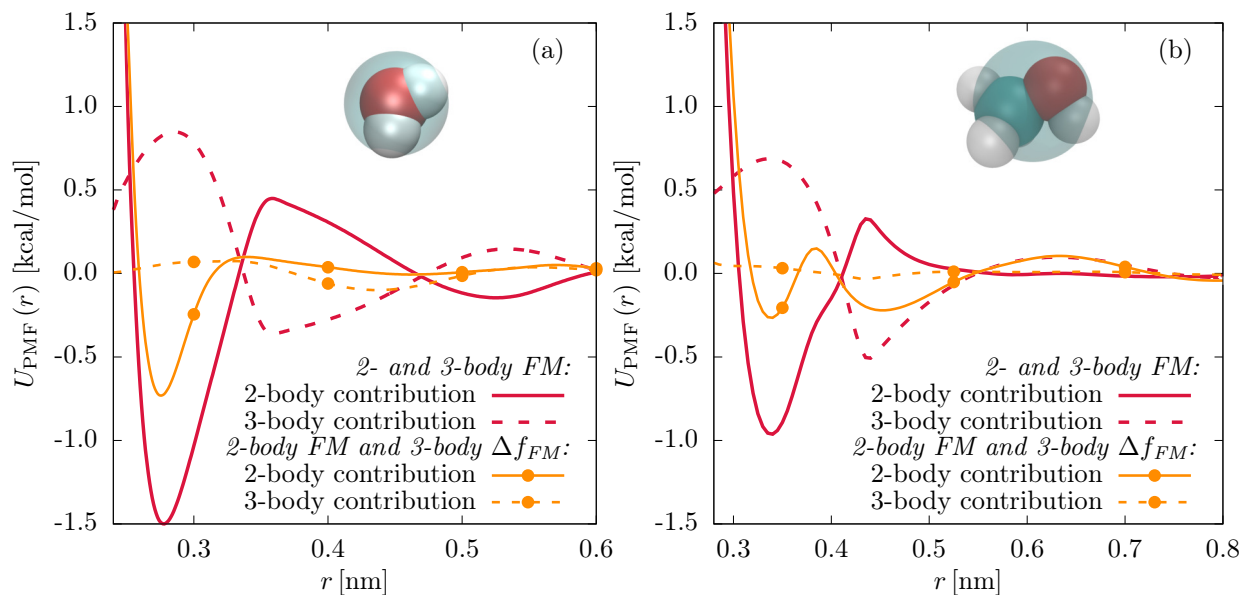


Figure 4. Decomposition of the CG two-body PMF for different CG parametrization schemes for (a) SPC/E water and (b) methanol. The results shown are for the concurrent two-body and three-body FM parametrization (2- and 3-body FM), and the three-body FM using the residual force of the two-body FM potential (2-body FM and 3-body  $\Delta f_{\text{FM}}$ ).

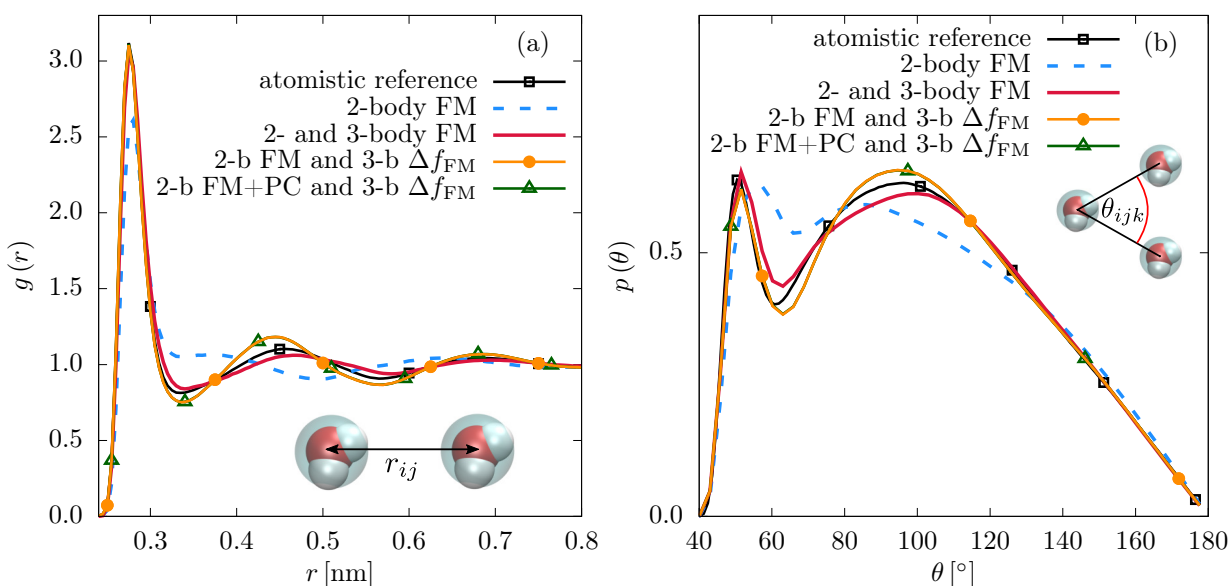


Figure 5. (a) Radial and (b) angular atomistic reference distribution functions of SPC/E water compared to the ones of different CG parametrizations: FM with a (tabulated) pair potential only (2-body FM), concurrent two-body and three-body FM parametrization (2- and 3-body FM), the three-body FM parametrization using the residual force of the two-body FM potential (2-b FM and 3-b  $\Delta f_{\text{FM}}$ ), and the three-body FM parametrization using the residual force of the two-body FM potential with additional pressure correction (2-b FM+PC and 3-b  $\Delta f_{\text{FM}}$ ). The angular distribution function is calculated for the first coordination shell using a cutoff of 0.37 nm.

is the optimal value regarding the structural accuracy of the models. Fig. 6 shows how  $P$  changes with the SW interaction parameter  $\gamma$  for all different parametrization schemes. For the unconstrained models,  $P$  decreases with increasing  $\gamma$  until reaching its minimum value at  $\gamma = 0.08$  nm, namely  $P = 0.12$  kbar for water and  $P = 0.83$  kbar for methanol. This corresponds to system

densities of  $\rho = 0.990$  g/cm<sup>3</sup> (water) and  $\rho = 0.613$  g/cm<sup>3</sup> (methanol) for NPT simulations at 1 atm. These values are close to the atomistic values of  $\rho = 0.998$  g/cm<sup>3</sup> (water) and  $\rho = 0.776$  g/cm<sup>3</sup> (methanol), and the experimental reference values at normal conditions:  $\rho = 0.9971$  g/cm<sup>3</sup> (water<sup>36</sup>) and  $\rho = 0.7872$  g/cm<sup>3</sup> (methanol<sup>37</sup>), even though these thermodynamic quantities have not



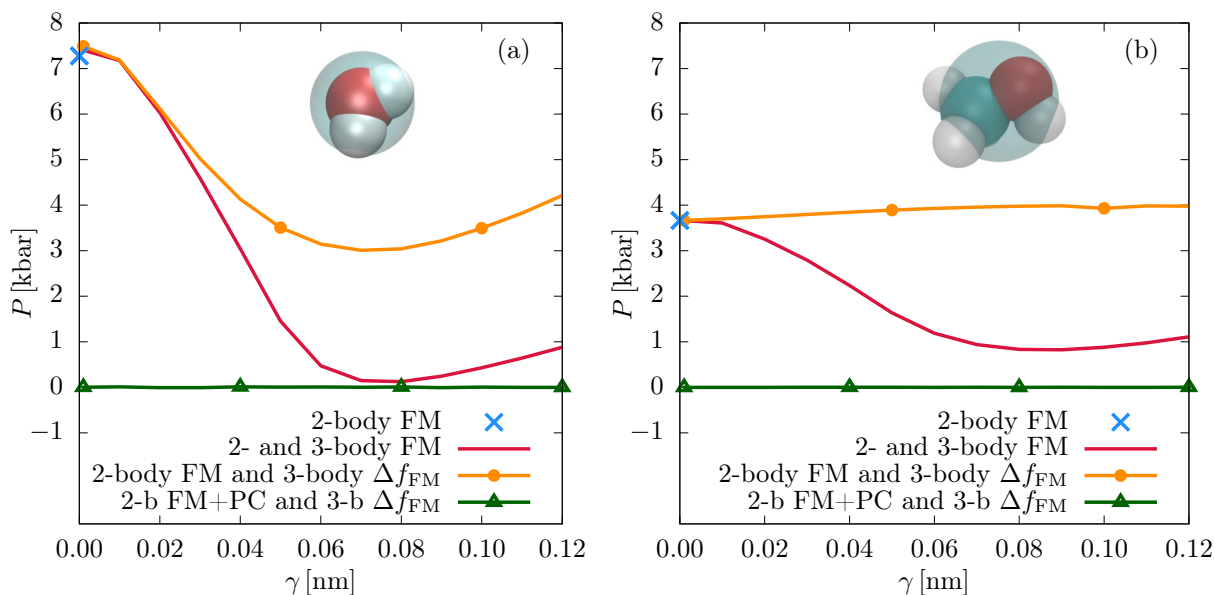


Figure 6. Average pressures  $P$  for different CG parametrization schemes for (a) SPC/E water and (b) methanol. The results shown are for the two-body FM (2-body FM), the concurrent two-body and three-body FM parametrization (2- and 3-body FM), the three-body FM using the residual force of the two-body FM potential (2-body FM and 3-body  $\Delta f_{FM}$ ), and the three-body FM parametrization using the residual force of the two-body FM potential with additional pressure correction (2-b FM+PC and 3-b  $\Delta f_{FM}$ ).

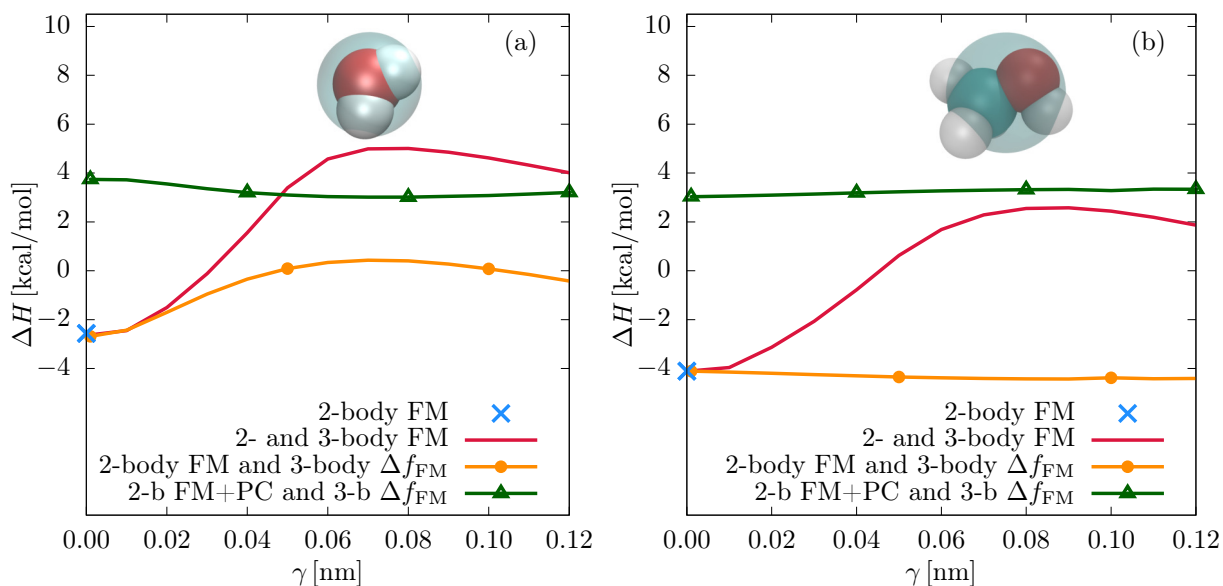


Figure 7. Average enthalpies of vaporization  $\Delta H$  for different CG parametrization schemes for (a) SPC/E water and (b) methanol. The results shown are for the two-body FM (2-body FM), the concurrent two-body and three-body FM parametrization (2- and 3-body FM), the three-body FM using the residual force of the two-body FM potential (2-body FM and 3-body  $\Delta f_{FM}$ ), and the three-body FM parametrization using the residual force of the two-body FM potential with additional pressure correction (2-b FM+PC and 3-b  $\Delta f_{FM}$ ).

been explicitly targeted in our parametrizations as, for example, with an explicit pressure matching scheme.<sup>47</sup>

The constrained parametrizations (2-body FM and 3-body  $\Delta f_{FM}$ ) yield significantly larger values of  $P$ . This is a known issue of the FM method which can be alleviated by adding a small linear perturbation to the two-body

FM potential (pressure correction):<sup>2,18</sup>

$$\Delta U_{PC} = -A \left( 1 - \frac{r}{r_{cut}} \right), \quad r_{cut} = 1.2 \text{ nm}, \quad (10)$$

where  $A$  is a positive constant. The pressure correction adds a small constant attractive force to each particle





pair within the cutoff radius  $r_{\text{cut}}$ . The value of  $A$  can be adjusted for each  $\gamma$  (see Tab. I of the supporting information) to shift the pressure to zero (see Fig. 6). The pressure corrected two-body FM potentials are shown in Figs. 1(a) (water) and 2(a) (methanol). For the full potential range until  $r_{\text{cut}} = 1.2$  nm we refer to Figs. 2(a) and 5(a) of the supporting information. NPT simulations at 1 atm yield system densities of  $\rho = 0.996$  g/cm<sup>3</sup> (water) and  $\rho = 0.777$  g/cm<sup>3</sup> (methanol) which are identical to the atomistic values.

The pressure correction only slightly perturbs the structure of the system. Indeed, the pair and the angular distribution functions of the corrected and uncorrected constrained parametrizations are practically identical (see Figs. 2(b) and 5, and Fig. 6 of the supporting information). The same holds for the decompositions of the two-body PMFs, implying that the pressure correction does not change the inner product of the two- and three-body interactions (see Figs. 3, 4, 7, and 8 of the supporting information).

Fig. 7 shows the enthalpies of vaporization  $\Delta H$  for all different models. Again, the two-body FM value is indicated with a blue cross. The unconstrained parametrizations (2- and 3-body FM) result in a maximum of  $\Delta H$  at  $\gamma = 0.08$  nm, where  $\Delta H = 5.01$  kcal/mol (water) and  $\Delta H = 2.55$  kcal/mol (methanol). The corresponding atomistic (experimental) values are 11.76 kcal/mol (10.52 kcal/mol) for water<sup>28</sup> and 8.94 kcal/mol (8.95 kcal/mol) for methanol.<sup>37</sup>

The pressure correction significantly improves the vaporization enthalpies of the constrained parametrizations by increasing the cohesive energy. For example, for  $\gamma = 0.08$  nm,  $\Delta H$  of water changes from 0.40 kcal/mol to 3.01 kcal/mol and the vaporization enthalpy of methanol changes from  $-4.42$  kcal/mol to 3.32 kcal/mol.

#### IV. CONCLUSIONS

We have shown that adding a Stillinger-Weber three-body potential to the coarse-grained force field of liquid water and methanol leads to significantly more attractive pair potentials. This short-range attraction is then compensated by the three-body term. Parametrization of the three-body term using residual forces substantially reduces the three-body contribution to the two-body potential of mean force. At the same time it worsens the thermodynamic properties of the coarse-grained model, quantified by pressure and enthalpy of vaporization. The thermodynamic properties can be corrected by adding a weak long-range attractive potential (pressure correction) without changing the structure of the system and the orthogonality of the three- and two-body contributions. Our findings motivate basis set extensions which separate the many-body contributions of different order.

#### SUPPLEMENTARY MATERIAL

The supplementary material contains technical details of the CG implementations, the calculation of thermodynamic properties, the tabulated two-body and three-body potentials, and additional figures of CG potentials, distribution functions and PMFs.

#### ACKNOWLEDGMENTS

We thank Joseph Rudzinski, Tristan Bereau, Leanne Paterson, Anton Melnyk, and Kurt Kremer for fruitful discussions as well as Gregory A. Voth for providing reference data of the SW water potential. DFG is acknowledged for financial support through the collaborative research center TRR 146.

#### REFERENCES

- W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, *J. Chem. Phys.* **128**, 244114 (2008).
- D. Reith, M. Pütz, and F. Müller-Plathe, *Journal of computational chemistry* **24**, 1624 (2003).
- A. P. Lyubartsev and A. Laaksonen, *Phys. Rev. E* **52**, 3730 (1995).
- M. S. Shell, *J. Chem. Phys.* **129**, 144108 (2008).
- S. Izvekov, M. Parrinello, C. J. Burnham, and G. A. Voth, *J. Chem. Phys.* **120**, 10896 (2004).
- G. Tóth, *J. Phys-Condens. Mat.* **19**, 335222 (2007).
- J. W. Mullinax and W. G. Noid, *J. Phys. Chem. C* **114**, 5661 (2010).
- J. F. Rudzinski and W. G. Noid, *The European Physical Journal Special Topics* **224**, 2193 (2015).
- J. F. Rudzinski and W. G. Noid, *J. Phys. Chem. B* **118**, 8295 (2014).
- S. Izvekov and G. A. Voth, *The Journal of chemical physics* **123**, 134105 (2005).
- J. W. Mullinax and W. G. Noid, *J. Chem. Phys.* **131**, 104110 (2009).
- V. Molinero and E. B. Moore, *J. Phys. Chem. B* **113**, 4008 (2009).
- L. Larini, L. Lu, and G. A. Voth, *J. Chem. Phys.* **132**, 164107 (2010).
- A. Das and H. C. Andersen, *J. Chem. Phys.* **136**, 194114 (2012).
- A. Das and H. C. Andersen, *J. Chem. Phys.* **131**, 034102 (2009).
- M. Jochum, D. Andrienko, K. Kremer, and C. Peter, *J. Chem. Phys.* **137**, 064102 (2012).
- C. Scherer and D. Andrienko, *The European Physical Journal Special Topics* **225**, 1441 (2016).
- V. Rühle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko, *J. Chem. Theory. Comput.* **5**, 3211 (2009).
- A. P. Lyubartsev and A. Laaksonen, *Phys. Rev. E* **55**, 5689 (1997).
- A. P. Bartók and G. Csányi, *Int. J. Quantum Chem.* **115**, 1051 (2015).
- A. Tsourtis, V. Harmandaris, and D. Tsagkarogiannis, *Entropy* **19**, 395 (2017).
- M. Dijkstra, J. M. Brader, and R. Evans, *J. Phys-Condens. Mat.* **11**, 10079 (1999).
- P. G. Bolhuis, A. A. Louis, and J. P. Hansen, *Phys. Rev. E* **64** (2001).
- T. Sanyal and M. S. Shell, *J. Chem. Phys.* **145**, 034109 (2016).
- J. W. Wagner, T. Dannenhöffer-Lafage, J. Jin, and G. A. Voth, *J. Chem. Phys.* **147**, 044113 (2017).



- <sup>26</sup>M. R. DeLyser and W. G. Noid, *J. Chem. Phys.* **147**, 134111 (2017).
- <sup>27</sup>F. Zipoli, T. Laino, S. Stolz, E. Martin, C. Winkelmann, and A. Curioni, *J. Chem. Phys.* **139**, 094501 (2013).
- <sup>28</sup>J. Lu, Y. Qiu, R. Baron, and V. Molinero, *J. Chem. Theory. Comput.* **10**, 4104 (2014).
- <sup>29</sup>G. A. Cisneros, K. T. Wikfeldt, L. Ojamäe, J. Lu, Y. Xu, H. Torabifard, A. P. Bartók, G. Csányi, V. Molinero, and F. Paesani, *Chem. Rev.* **116**, 7501 (2016).
- <sup>30</sup>A. Das and H. C. Andersen, *J. Chem. Phys.* **136**, 194113 (2012).
- <sup>31</sup>F. H. Stillinger and T. A. Weber, *Physical review B* **31**, 5262 (1985).
- <sup>32</sup>M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, *SoftwareX* **1-2**, 19 (2015).
- <sup>33</sup>H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *Journal of Physical Chemistry* **91**, 6269 (1987).
- <sup>34</sup>W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* **110**, 1657 (1988).
- <sup>35</sup>W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, *J. Am. Chem. Soc.* **118**, 11225 (1996).
- <sup>36</sup>J. L. F. Abascal and C. Vega, *J. Chem. Phys.* **123**, 234505 (2005).
- <sup>37</sup>C. Caleman, P. J. van Maaren, M. Hong, J. S. Hub, L. T. Costa, and D. van der Spoel, *J. Chem. Theory. Comput.* **8**, 61 (2012).
- <sup>38</sup>W. F. Van Gunsteren and H. J. C. Berendsen, *Molecular Simulation* **1**, 173 (1988).
- <sup>39</sup>H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *The Journal of chemical physics* **81**, 3684 (1984).
- <sup>40</sup>U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, *J. Chem. Phys.* **103**, 8577 (1995).
- <sup>41</sup>S. Plimton, *J. Comput. Phys.* **117**, 1 (1995).
- <sup>42</sup>S. Nosé, *Molecular Physics* **52**, 255 (1984).
- <sup>43</sup>W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).
- <sup>44</sup>W. Shinoda, M. Shiga, and M. Mikami, *Phys. Rev. B* **69** (2004).
- <sup>45</sup>J. F. Rudzinski and W. G. Noid, *J. Phys. Chem. B* **116**, 8621 (2012).
- <sup>46</sup>W. G. Noid, P. Liu, Y. Wang, J.-W. Chu, G. S. Ayton, S. Izvekov, H. C. Andersen, and G. A. Voth, *J. Chem. Phys.* **128**, 244115 (2008).
- <sup>47</sup>N. J. H. Dunn and W. G. Noid, *J. Chem. Phys.* **143**, 243148 (2015).

