Assessing the accuracy of improved force-matched water models derived from *ab-initio* molecular dynamics simulations

Andreas Köster[*] Thomas Spura[†] Gábor Rutkai[*] Jan Kessler[†]
Hendrik Wiebeler[†] Jadran Vrabec[*] Thomas D. Kühne[†,‡]

April 7, 2016

Abstract

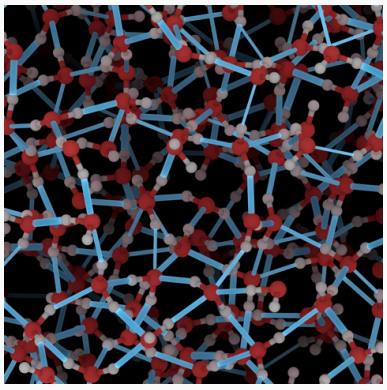
The accuracy of water models derived from *ab-initio* molecular dynamics simulations by means on an improved force-matching scheme is assessed for various thermodynamic, transport and structural properties. It is found that although the resulting force-matched water models are typically less accurate than fully empirical force fields in predicting thermodynamic properties, they are nevertheless much more accurate than generally appreciated in reproducing the structure of liquid water and in fact superseding most of the commonly used empirical water models. This development demonstrates the feasibility to routinely parametrize computationally efficient yet predictive potential energy functions based on accurate *ab-initio* MD simulations for a large variety of different systems.

Keywords: Liquid water, Force Matching, Ab-Initio, Molecular Dynamics, Monte Carlo ■

[*]Thermodynamics and Energy Technology, University of Paderborn, Warburger Str. 100, D-33098 Paderborn, Germany

[†]Dynamics of Condensed Matter, Department of Chemistry, University of Paderborn, Warburger Str. 100, D-33098 Paderborn, Germany

[‡]Corresponding author. E-mail address: tdkuehne@mail.upb.de.



The accuracy of water models derived from *ab-initio* molecular dynamics simulations by means on an improved force-matching scheme is assessed for a large variety of different thermodynamic, transport and structural properties. It is found that although the resulting force-matched water models are typically less accurate than fully empirical force fields in predicting thermodynamic properties, they are nevertheless much more accurate than generally appreciated in reproducing the structure of liquid water.

1 Introduction

Since the very first application of molecular dynamics (MD) to realistic systems[1], liquid water has been one of the most thoroughly studied systems and is arguably the most important liquid because of its role in chemistry, physics and biology[2,3]. In fact, the chemistry of living things crucially depends on the interplay of its unusual properties and anomalous behaviour and it is hard hard to imagine life without it[4]. It is also a desirable "green" solvent that offers great economic, safety and environmental benefits, since it is inexpensive, non-flammable and nontoxic at the same time[5]. Yet, in synthetic organic chemistry, water is generally not a popular choice of solvent. On the one hand because the oxygen atom of water molecules are rather reactive and may themselves react with the organic compounds, and, on the other hand, since most organic molecules are nonpolar and thus hydrophobic, which causes that the reactant are insoluble in liquid water. However, rather recently, the use of water as a solvent for organic synthesis was popularized by Sharpless and coworkers, who recognized that many important organic reactions exhibit greatly enhanced selectivity, improved yields and dramatically reduced reaction time when floating "on-water" [6–8]. Therefore, the behavior and properties of liquid water have been the subject of extensive scientific investigation[9–11]. Nevertheless, a detailed understanding of liquid water is still lacking[2,12,13]. This is a consequence that studying liquid water *in silico* is rather challenging, which is due to the difficulties in accurately modeling the numerous physical phenomena that conspire to make water unique, such as the cooperativity of the hydrogen bond (HB) network[14,15], strong permanent dipole moment, large polarizability effects and sizeable nuclear quantum effects[9].

Ab-Initio molecular dynamics (AIMD)[16–18], where the interatomic forces are calculated on-the-fly by accurate electronic structure calculations, has become widespread in computational studies of liquid water[19–42]. Despite constant advances in high-performance computing, the computational cost of AIMD has limited the attainable length- and timescales in defiance of substantial progress[17,43]. Therefore, great effort has been spent into developing empirical water potentials, which are fitted so as to reproduce experimental data, such as the structure factor, radial distribution function (RDF), heat of vaporization, vapor-liquid

equilibria (VLE), as well as the density maximum of liquid water[44–50]. Despite the fact that these force fields are typically remarkably successful in reproducing the underlying experiments, the transferability to regions of the phase diagram that are different from those in which they have been fitted is restricted.

Nevertheless, it is possible to circumvent the latter by deriving water potentials from parameter-free electronic structure calculations[51–57]. Apart from the trivial finite difference scheme[50], there a various different techniques to parameterize water models to match *ab-initio* data such as inverse Monte Carlo[58,59], iterative Boltzmann inversion[60] or force-matching[61] approaches. In this work, we assess the general accuracy, with a specific emphasis on thermodynamic quantities, of force-matched water potentials derived from first-principles electronic structure calculations as obtained by the recently devised improved force-matching method of Spura *et al.*[62].

The remainder of this paper in organized as follows. In section 2, we revisit the force-matching scheme to derive the parameters of a rigid, non-polarizable TIP4P-like water model. The computational details are described in Section 3, whereas in Section 4, the results as obtained by MD and Monte Carlo (MC) simulations are discussed. Section 5 is devoted to conclusion.

2 Force Matching

The force-matching technique of Ercolessi and Adams[61], where the interaction potential is derived so as to mimic the forces of accurate reference calculations, not only includes many-body environmental effects, but also allows to employ a rather high level of theory since in general relatively few electronic structure calculations are required. At variance to parametrization schemes that rely on Henderson's theorem[63], no computationally expensive generation of reference RDF from first-principle by means of *ab-initio* MD is required[19–42]. Specifically, to determine the parameters of an arbitrary interaction potential based on first-principles force calculations for a set of given configurations, we minimize the normalized L_1

force distance

$$\|\delta \boldsymbol{F}\|_1 = \frac{1}{3} \left\langle \sum_{i=1}^{N} \sum_{\alpha \in (x,y,z)} \left[\frac{|\boldsymbol{F}_{i,\alpha}^{\mathrm{QM}} - \boldsymbol{F}_{i,\alpha}^{\mathrm{FF}}|}{\sigma_i} \right] \right\rangle, \tag{1}$$

where N is the number of atoms and σ_i for the standard deviation of the force distribution $\boldsymbol{F}_{i,\alpha}$ of atom i in directions $\alpha \in (x, y, z)$, while $\langle \cdots \rangle$ implies the ensemble average of the selected configurations. The quantum mechanical reference forces are denoted as $\boldsymbol{F}_{i,\alpha}^{\mathrm{QM}}$, while $\boldsymbol{F}_{i,\alpha}^{\mathrm{FF}}$ are the nuclear forces of the classical interaction potential, respectively.

However, the minimization of Eq. 1 with respect to the variational parameters of $\boldsymbol{F}_{i,\alpha}^{\mathrm{FF}}$ represents an ill-posed problem, in particular when including atomic partial charges in the optimization procedure. As a consequence, the optimization may not be stable for small variations of the corresponding parameters. This is reflected in an error landscape with many saddle points and flat areas, where the Hessian matrix is nearly singular, which leads to numerical inaccuracies because of the limited precision of floating point arithmetic. Thus, an important problem of gradient-based minimization methods is the particular form of the objective function, whose derivative with respect to partial charges is often found to be ill-conditioned.

Although it is possible to ameliorate this difficulty by augmenting the penalty function with additional properties, such as the total force or torque with its respective weights[64–66], this is circumvented here by means of the sequential least-squares quadratic programming method (SLSQP) together with physically-sensible bound constraints[67]. The SLSQP scheme treats the original problem as a sequence of constrained least-squares problems, which is equivalent to a quadratic programming algorithm for nonlinearly-constrained gradient-based optimization, hence its name. More precisely, each SLSQP step involves solving a quadratic approximation of the original objective function, where the linear term is the gradient and the quadratic term is an approximate Hessian, with first-order affine approximations of the nonlinear constraints. The approximate Hessian, which is initialized to the identity matrix, is continuously updated, while maintaining it positive definite, based on the gradients and function values at subsequent steps similar to the BFGS quasi-Newton scheme[68]. Therefore, like any quasi-Newton method, the true Hessian is only approached in the limit of many iterations close to the minimum. Because of the ill-posed nature of the problem, we search

for the minimum along the direction of the modified quasi-Newton scheme by first bracketing the minimum and then using Brent's method[69]. Contrary to more elaborate techniques that exploit gradient information, the availability of the function's derivative is not required here. However, it should be noted that this procedure offers no guarantee about whether the global minimum of the optimization function is indeed found.

3 Computational Details

For the purpose to generate first-principles-based water models from *ab-initio* MD simulations that are as transferable as possible, we have extracted 1500 decorrelated snapshots from path-integral MD simulations[70] consisting of 125 water molecules in the isobaric-isothermal (NpT) ensemble using the q-TIP4P/F water potential of Habershon *et al.*[71]. More precisely, we have selected 125 different configurations at 1 bar pressure for each temperature over the whole liquid temperature range between 248 K to 358 K. Therefore, the resulting water model was not just parametrized to a single state point at ambient conditions, but over a wide range of state points from undercooled to nearly saturated liquid water.

Force matching, as alluded to above, was conducted using reference forces from density functional theory (DFT) calculations[72,73]. Specifically, the mixed Gaussian and plane wave approach[74], as implemented in the CP2K/Quickstep code, was employed[75]. In this approach the Kohn-Sham orbitals are represented by a TZV2P Gaussian basis set[76], while the charge density is expanded by plane waves using a density cutoff of 320 Ry. The exchange and correlation (XC) energy was described by the Tao-Perdew-Staroverov-Scuseria (TPSS) meta-generalized gradient approximation[77], and norm-conserving Goedecker-Teter-Hutter pseudopotentials were used to describe the interactions between the valence electrons and the ionic cores[78,79]. Van der Waals interactions, which are typically neglected by common local and semi-local XC functionals, were approximated by an additional, attractive pair-potential[80,81]. In its most recent and elaborate form, the latter is referred to as D3 correction[82].

The parameters of the q-TIP4P/F-like water potentials were obtained by minimizing Eq. 1 using the SLSQP algorithm of Kraft with a convergence tolerance of 10^{-6} between the

individual iterations[67]. Gradients with respect to the various variational parameters were computed using finite differences with a displacement of 10^{-8}. The initial parameters were taken from the original q-TIP4P/F water model[71]. The resulting DFT-based water model is denoted as TIP4P/TPSS-D3 in the following[62]. The usage of the TPSS XC functional is due to the observation that, when combined with Grimme's D3 correction[82], the resulting TPSS-D3 level of theory is remarkable accurate in reproducing the structure of liquid water[62]. In fact, as will demonstrated in the present paper, when comparing with the recently revised experimental x-ray and neutron measurements[83,84], the present TIP4P/TPSS-D3 water model is even more accurate than most of the commonly used empirical water potentials. Nevertheless, we find it important to note that the present fixed point-charge water model is neither polarizable nor able to mimic cooperativity effects and chemical reactions that may take place in liquid water. Even though these many-body effects are explicitly taken into account by DFT, in this way it is generally only possible to simulate liquid water in qualitative, but not in quantitative agreement with experiment[40,42,85,86].

For comparison, two empirical water potentials, i.e. TIP4P/2005 by Abascal and Vega[48] and another one of Huang *et al.*[87], were employed in this work. Being one of the most important "general purpose" water potentials, TIP4P/2005 is an obvious choice for this task and performs rather well for a large variety of different thermodynamic properties[48,50]. In contrast, the potential of Huang *et al.* was mainly optimized to reproduce the vapor-liquid equilibrium[87]. To that extent, the intramolecular OH distance is chosen to be 40 % larger to yield a more localized hydrogen bond network. However, the magnitude of the point charges is significantly smaller than that of other TIP4P-like water models, while the attractive force is compensated by a relatively high LJ energy parameter.

All water models used in the present work consist of two positive charge sites of magnitude $q/2$ on the hydrogen atoms and a negative charge of magnitude q positioned at $\boldsymbol{r}_{\mathrm{M}} = \gamma \boldsymbol{r}_{\mathrm{O}} + (1-\gamma)(\boldsymbol{r}_{\mathrm{H1}} - \boldsymbol{r}_{\mathrm{H2}})/2$ to ensure local charge neutrality of each water molecule. These so-called M-sites and the hydrogen atoms on different water molecules interact with each other through a simple Coulomb potential. Together with a Lennard-Jones potential between the

oxygen atoms, this constitutes the following pairwise-additive intermolecular potential

$$V_{\text{inter}} = \sum_i \sum_{j>i} 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] + \sum_{m \in i} \sum_{n \in j} \frac{q_m q_n}{r_{mn}},$$

where r_{ij} is the distance between the oxygen atoms and r_{mn} the distance between the partial charges in molecules i and j. All molecular interaction parameters are listed in table 1.

Table 1: Parameters of the molecular interaction models for water.

Interaction model	r_{OH}	r_{OM}	θ_{HOH}	σ	ϵ/k_{B}	q_{M}	q_{H}
	Å	Å	deg	Å	K	e	e
TIP4P-TPSS[62]	0.9662	0.1654	107.21	3.2081	64.95	-1.0552	0.5276
TIP4P-TPSS-D3[62]	0.9666	0.1547	107.38	3.1625	79.37	-1.0318	0.5159
TIP4P/2005[48]	0.9572	0.1546	104.52	3.1589	93.20	-1.1128	0.5564
Huang *et al.*[87]	1.3338	0.2048	104.52	3.1183	208.08	-0.8391	0.4196

All simulations in the present work were carried out with the molecular simulation tool *ms*2[88]. Except for the transport property calculations with the Green-Kubo formalism[89,90], where MD is mandatory[91,92], all results were generated by MC sampling[93,94]. In the case of the latter, a cubic simulation volume consisting of 864 water molecules was considered. The Lennard-Jones interactions beyond a cut-off radius of 12.25 Å were corrected as proposed by Lustig[95], while the long-range electrostatic interactions were taken into account by means of the reaction field method[96]. Each MC move in the NpT ensemble consisted 864 displacement, 864 rotation and one volume move. For the MD simulations to compute the transport properties, a simulation volume of 3000 water molecules together with a Lennard-Jones cut-off of 17.5 Å was considered. The equations of motion were discretized by a timestep of 0.87 fs and solved numerically by employing a fifth-order Gear predictor-corrector integrator[96]. The temperature was kept constant by means of an isokinetic rescaling of the velocities, while the pressure was controlled with the piston method of Andersen[97].

4 Results and discussion

In order to validate the present results, they were compared to a highly accurate fundamental equation of state (EOS), to experimental data and a correlation thereof. In the case of time independent thermodynamic properties, the reference class[98] EOS of Wagner and Pruß[99] was used as recommended by the International Association for the Properties of Water and Steam (IAPWS). This equation was parameterized to around 6000 carefully selected experimental data points. The uncertainty for most properties is below 0.1 % over a large temperature and pressure range. Uncertainties for the vapor pressure and the saturated liquid density are even below 0.0025 %. The shear viscosity was compared to the correlation of Huber *et al.*[100] which has an uncertainty of 1 % in the temperature range considered here. Relative deviations between the simulation results and the reference EOS were quantified by

$$\Delta X = 100 \left(\frac{X_{\mathrm{SIM}} - X_{\mathrm{EOS}}}{X_{\mathrm{EOS}}} \right). \tag{2}$$

Based on this definition, the average absolute (unsigned) deviation (AAD) was defined as

$$AAD = \frac{1}{N} \sum_{i=1}^{N} |\Delta X_i| \,. \tag{3}$$

The corresponding results of the considered water models are listed in Table 2.

4.1 Second virial coefficient

Over a temperature range from 263 K to 1200 K, the second virial coefficient B was calculated by evaluating Mayer's f-function as described e.g. by Eckl *et al.*[101]. Results for the four water potentials are presented in Fig. 1 and compared to the reference EOS[99], as well as to a correlation of experimental data from the DIPPR database[102]. Above 600 K, the agreement between the reference data and the various potentials is very good, while for lower temperatures the second virial coefficient is generally underestimated. The latter is particularly true for the TIP4P/2005 force field, whereas the deviation is systematically smaller for the potential of Huang *et al.*, as well as for the DFT-derived TIP4P-TPSS and TIP4P-TPSS-D3 water models, which are of comparable quality.

Table 2: Average absolute (unsigned) deviations (AAD) of four water potentials in comparison to the reference fundamental equation of state by Wagner and Pruß[99]. Vapor-liquid equilibrium properties were evaluated between 300 K and 525 K.

Thermodynamic property	Molecular interaction model			
	TIP4P-TPSS[62]	TIP4P-TPSS-D3[62]	TIP4P/2005[48]	Huang *et al.*[87]
	AAD / %			
Vapor pressure p_v	234.8	162.6	55.4	12.3
Saturated liquid density ρ'	14.0	10.0	0.5	1.5
Saturated vapor density ρ"	309.0	205.8	54.2	11.9
Enthalpy of vaporization Δh_v	23.2	18.8	11.7	2.4
Res. isochoric heat capacity c_v^{res}	7.1	7.2	22.1	11.0
Res. isobaric heat capacity c_p^{res}	16.8	15.1	18.8	6.8
Pressure p	50.9	38.5	5.2	30.2
Isothermal compressibility β_T	16.4	15.6	7.3	26.0
Residual enthalpy h^{res}	19.7	15.5	24.0	11.4
Speed of sound w	7.0	6.1	2.6	19.7

4.2 Vapor-liquid equilibrium

VLE were determined with the grand equilibrium method[103]. In this two-step procedure, the coexisting phases are simulated independently. First, at a specified temperature one molecular simulation run in the liquid phase is carried out in the NpT ensemble to obtain the chemical potential as a function of pressure. This was done here using the gradual insertion method[104,105]. The second step includes one pseudo-grand canonical (μVT) ensemble simulation for the vapor phase and yields the saturated vapor state point. A detailed description of this method can be found elsewhere[103].

In Figs. 2 and 3, the representation of the vapor pressure as determined by the four considered water potentials is shown and compared to the reference EOS[99]. The TIP4P/2005 force field yields a vapor pressure that is throughout too low with decreasing deviations at higher temperatures resulting in an AAD of 55 %, cf. Table 2. In contrast, the two DFT-derived water models TIP4P-TPSS and TIP4P-TPSS-D3 throughout overestimate the experimental data by at least 100 %. The potential of Huang *et al.*, however, overestimates

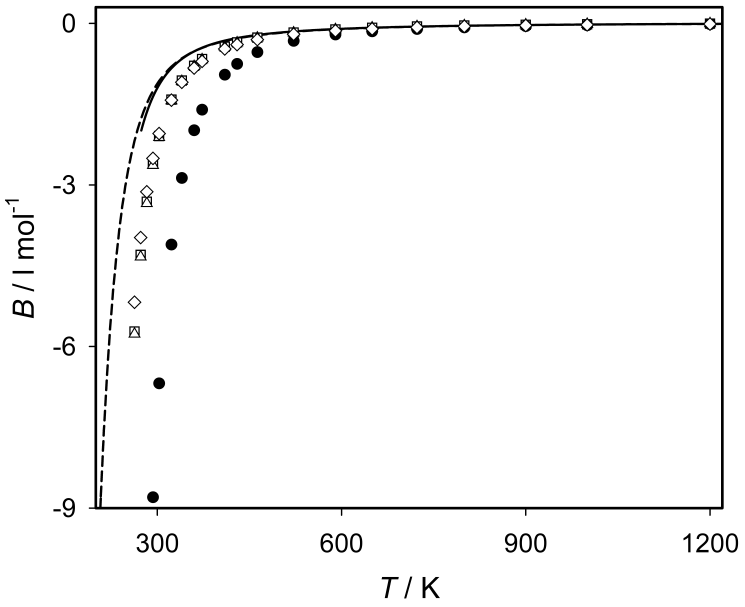


Figure 1: Second virial coefficient of water: (—) Reference fundamental equation of state by Wagner and Pruß[99]. (- -) DIPPR correlation of experimental data[102]. Computational data: (•) TIP4P/2005; (⋄) Huang *et al.*; (□) TIP4P-TPSS; (△) TIP4P-TPSS-D3.

the vapor pressure at low temperatures by about 30 %, while underestimating it at 625 K by around 10 %, which results in an AAD of only 12 %.

The coexisting liquid and vapor densities are presented in Fig. 4. With an AAD of 0.5 %, the TIP4P/2005 force field shows the best agreement with the reference EOS[99] on the liquid side, whereas the potential of Huang *et al.* yields a slope that is slightly too steep. Nonetheless, the $AAD = 1.5$ % of the potential of Huang *et al.* is not much larger than that of the TIP4P/2005 force field. Both exhibit deviations of around 5 % at 625 K, which is close to the critical point. However, the DFT-derived TIP4P-TPSS and TIP4P-TPSS-D3 water models, systematically underestimate the saturated liquid density, yielding AAD values of 14 % and 10 %, respectively. As a consequence, the critical temperature of both potentials is around 100 K below the experimental value.

Regarding the saturated vapor density, which is closely related to the vapor pressure, the TIP4P/2005 force field underestimates the reference value by approximately 54 %. The potential of Huang *et al.* exhibits the best agreement with the reference EOS[99], i.e. $AAD =$

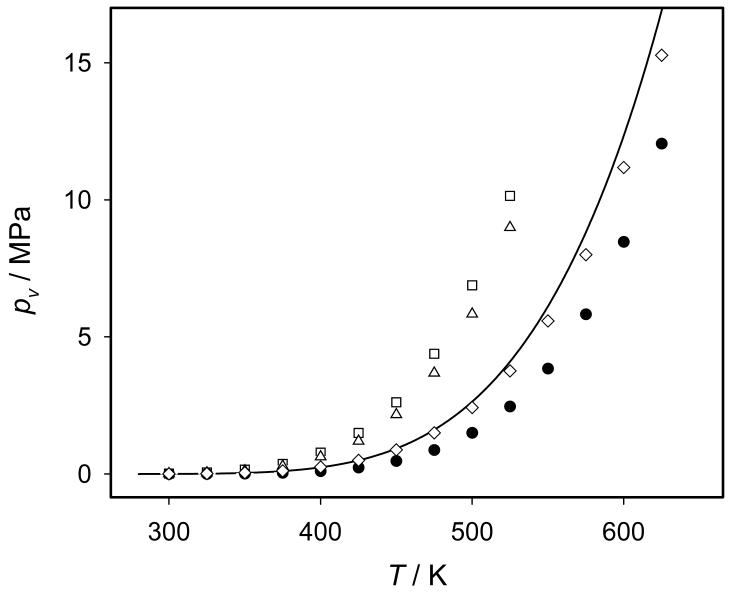


Figure 2: Vapor pressure of water: (—) Reference fundamental equation of state by Wagner and Pruß[99]. Computational data: (•) TIP4P/2005; (◇) Huang *et al.*; (□) TIP4P-TPSS; (△) TIP4P-TPSS-D3. The statistical simulation uncertainties are within symbol size.

12 %, while the TIP4P-TPSS and TIP4P-TPSS-D3 water models overestimate this property on average by more than 200 %.

Similar to the representation of the vapor pressure and saturated vapor density, the potential of Huang *et al.* also performs rather well in accurately in terms of the enthalpy of vaporization and yields an $AAD = 2.4$ %, cf. Fig. 5. At temperatures above 575 K, in the extended critical region, the TIP4P/2005 force field also shows a very good agreement with the reference EOS[99]. At lower temperatures, however, the deviations increase up to 14 %, which results in an $AAD = 12$ %. The two DFT-derived water models underestimate the enthalpy of vaporization throughout with deviations of around 6 % for lower temperatures. Over the whole sampled temperature range the DFT-derived TIP4P-TPSS and TIP4P-TPSS-D3 water models yield an AAD of 23 and 19 %, respectively.

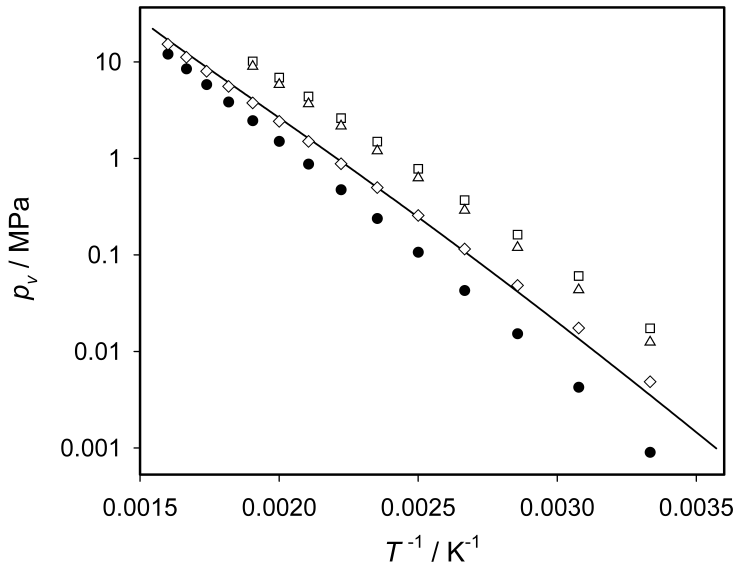


Figure 3: Vapor pressure of water: (—) Reference fundamental equation of state by Wagner and Pruß[99]. Computational data: (•) TIP4P/2005; (◇) Huang *et al.*; (□) TIP4P-TPSS; (△) TIP4P-TPSS-D3. The statistical simulation uncertainties are within symbol size.

4.3 Homogeneous properties

A wide range of different static properties in the homogeneous fluid region were generated using the formalism proposed by Lustig[106,107]: Residual isobaric heat capacity c_p^{res}, residual isochoric heat capacity c_v^{res}, residual enthalpy h^{res}, pressure p, speed of sound w, isothermal compressibility β_T and Joule-Thomson coefficient μ_{JT}. For a given state point, this formalism allows to calculate an arbitrary number of time independent thermodynamic properties from a single simulation run within the canonical (NVT) ensemble. More precisely, the method was designed to directly yield the temperature and density derivatives

$$A_{mn} = (1/T)^m \rho^n \frac{\partial^{m+n}(a^{\mathrm{o}} + a^{\mathrm{r}})/(RT)}{\partial(1/T)^m \partial\rho^n}, \tag{4}$$

of the residual Helmholtz free energy $a(T,\rho) = a^{\mathrm{o}}(T,\rho) + a^{\mathrm{r}}(T,\rho)$ for any $m > 0$ or $n > 0$, where ρ is the molar density, R is the molar gas constant, while the factor $1/(RT)$ renders A_{mn} dimensionless. The molar Helmholtz free energy a, and thus the derivatives A_{mn}, can be additively separated into an ideal and residual contribution. The ideal contribution

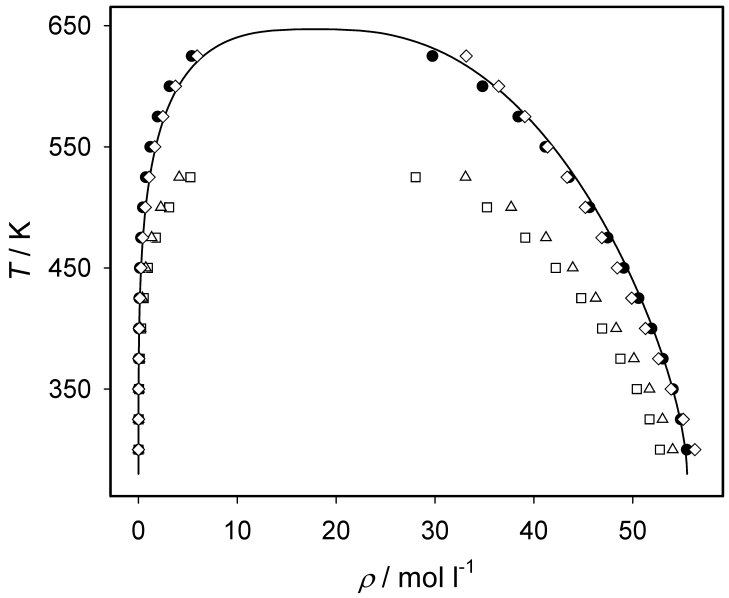


Figure 4: Saturated densities of water: (—) Reference fundamental equation of state by Wagner and Pruß[99]. Computational data: (•) TIP4P/2005; (⋄) Huang *et al.*; (□) TIP4P-TPSS; (△) TIP4P-TPSS-D3. The statistical simulation uncertainties are within symbol size.

$a^{\mathrm{o}}(T,\rho) = a^{\mathrm{o}}(T) + RT\ln(\rho/\rho_{\mathrm{ref}})$ by definition corresponds to the value of $a(T,\rho)$ when no intermolecular interactions are at work. The exclusively temperature dependent part of the ideal contribution $a^{\mathrm{o}}(T)$ has a non-trivial temperature dependence and is directly related to the isochoric heat capacity of the ideal gas $A^{\mathrm{o}}_{20}(T) = -c^{\mathrm{o}}_v/R$. However, not every derivative A_{mn} requires the knowledge of $a^{\mathrm{o}}(T) = RTA^{\mathrm{o}}_{00}(T)$, namely

- $A^{\mathrm{o}}_{mn}(T,\rho) = 0$, for $m > 0$ and $n > 0$,
- $A^{\mathrm{o}}_{mn}(T,\rho) = A^{\mathrm{o}}_{mn}(T) + 0$, for $m > 0$ and $n = 0$,
- $A^{\mathrm{o}}_{mn}(T,\rho) = 0 + (-1)^{1+n}$, for $m = 0$ and $n > 0$.

The value of $A^{\mathrm{o}}_{00}(T) = a^{\mathrm{o}}(T)/(RT)$ is often determined by spectroscopy or *ab-initio* calculations, whereas the residual contribution is typically the target of molecular simulations.

Since $a/(RT)$ is a thermodynamic potential, any other time independent thermodynamic property is a combination of its derivatives with respect to its independent variables[106,107]:

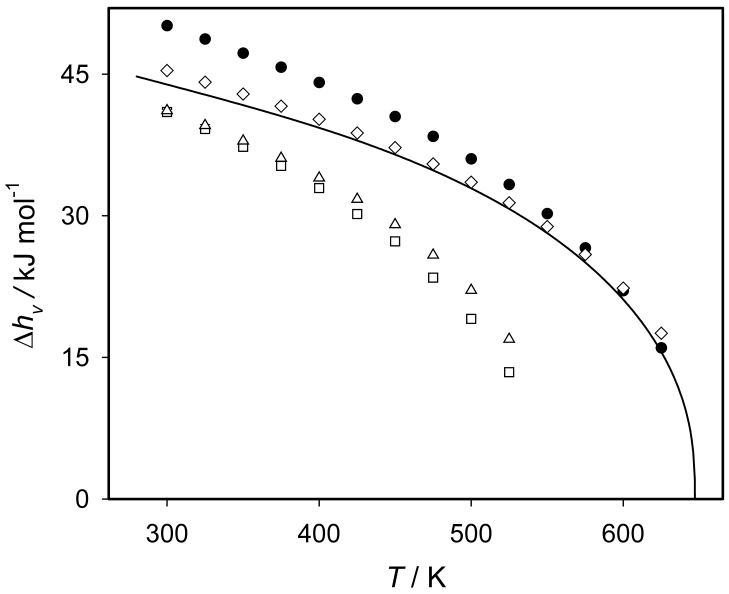


Figure 5: Enthalpy of vaporization of water: (—) Reference fundamental equation of state by Wagner and Pruß[99]. Computational data: (•) TIP4P/2005; (⋄) Huang *et al.*; (□) TIP4P-TPSS; (△) TIP4P-TPSS-D3. The statistical simulation uncertainties are within symbol size.

Numerous properties, such as the internal energy u, pressure p, enthalpy h, Gibbs energy g, and isochoric heat capacity c_v, are linear functions

$$\frac{u}{RT} = A^{\mathrm{o}}_{10}(T) + A^{\mathrm{r}}_{10}, \tag{5}$$

$$\frac{p}{\rho RT} = 1 + A^{\mathrm{r}}_{01}, \tag{6}$$

$$\frac{1}{RT}\left(\frac{\partial p}{\partial \rho}\right)_T = 1 + 2A^{\mathrm{r}}_{01} + A^{\mathrm{r}}_{02}, \tag{7}$$

$$\frac{1}{\rho R}\left(\frac{\partial p}{\partial T}\right)_\rho = 1 + A^{\mathrm{r}}_{01} - A^{\mathrm{r}}_{11}, \tag{8}$$

$$\frac{h}{RT} = 1 + A^{\mathrm{r}}_{01} + A^{\mathrm{o}}_{10}(T) + A^{\mathrm{r}}_{10}, \tag{9}$$

$$\frac{g}{RT} = 1 + A^{\mathrm{r}}_{01} + A^{\mathrm{o}}_{00} + A^{\mathrm{r}}_{00}, \tag{10}$$

$$\frac{c_v}{R} = -A^{\mathrm{o}}_{20}(T) - A^{\mathrm{r}}_{20}, \tag{11}$$

whereas the molar isobaric heat capacity c_p, speed of sound w, and Joule-Thomson coefficient μ_{JT} are non-linear combinations of derivatives

$$\frac{c_p}{R} = -A^{\mathrm{o}}_{20}(T) - A^{\mathrm{r}}_{20} + \frac{(1 + A^{\mathrm{r}}_{01} - A^{\mathrm{r}}_{11})^2}{1 + 2A^{\mathrm{r}}_{01} + A^{\mathrm{r}}_{02}}, \tag{12}$$

$$\frac{Mw^2}{RT} = 1 + 2A^{\mathrm{r}}_{01} + A^{\mathrm{r}}_{02} - \frac{(1 + A^{\mathrm{r}}_{01} - A^{\mathrm{r}}_{11})^2}{A^{\mathrm{o}}_{20}(T) + A^{\mathrm{r}}_{20}}, \tag{13}$$

$$\frac{\rho R}{\mu_{\mathrm{JT}}} = \frac{(A^{\mathrm{o}}_{20}(T) + A^{\mathrm{r}}_{20})(1 + 2A^{\mathrm{r}}_{01} + A^{\mathrm{r}}_{02})}{A^{\mathrm{r}}_{01} + A^{\mathrm{r}}_{02} + A^{\mathrm{r}}_{11}} - \frac{(1 + A^{\mathrm{r}}_{01} - A^{\mathrm{r}}_{11})^2}{A^{\mathrm{r}}_{01} + A^{\mathrm{r}}_{02} + A^{\mathrm{r}}_{11}}, \tag{14}$$

where M is the molar mass. For properties that cannot be additively separated into ideal and residual part, such as w and μ_{JT}, the ideal contribution $A^{\mathrm{o}}_{20}(T) = -c^{\mathrm{o}}_v/R$ was taken from the reference EOS[99] because it does not depend on the intermolecular interactions.

Properties in the homogeneous fluid region were compared to the reference EOS[99] along four isochores (i.e. 6, 24, 42 and 58 mol/l) up to temperatures of 1000 K. The corresponding state points are shown in Fig. 6.

The results of the four investigated water potentials for the residual isochoric heat capacity (top) and the residual isobaric heat capacity (bottom) are shown in Fig. 7 and compared to the reference EOS[99]. The residual isochoric heat capacity is throughout overestimated by the TIP4P/2005 force field, which performs particularly poorly at low densities, leading to an AAD of 22 %. On the contrary, the potential of Huang *et al.* underestimates the isochoric heat capacity for densities above 6 mol/l by around 11 %. With an AAD of around 7 % over the entire temperature and density range, both the TIP4P-TPSS and TIP4P-TPSS-D3 water models perform quite well in comparison with the reference. However, regarding the residual isobaric heat capacity, the agreement between the TIP4P/2005 force field and the reference EOS[99] is better than for the residual isochoric heat capacity. Again, the largest deviations can be observed for lower densities resulting in an $AAD = 19$ %. With an $AAD = 17$ % and 15 %, respectively, the DFT-derived TIP4P-TPSS and TIP4P-TPSS-D3 water models

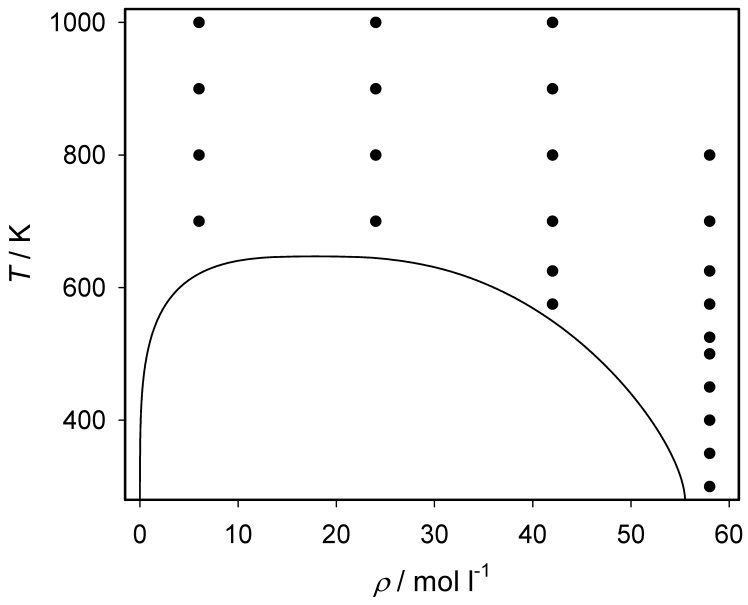


Figure 6: Studied state points in the homogeneous fluid region of water along four isochores. The solid line indicates saturated states according to the reference equation of state for water[99].

are slightly superior. Nevertheless, both water models underestimate the residual isobaric heat capacity along the 24 mol/l isochore by more than 25 %. The best agreement with the reference was observed for the potential of Huang *et al.* with an $AAD = 6.8$ %. But it has to be noted that the state point at 700 K and 24 mol/l is within the extended critical region, which leads to relatively large statistical uncertainties of the simulations.

The pressure and the isothermal compressibility in the homogeneous fluid region are shown in Fig. 8. In comparison to the reference EOS, the two DFT-derived water models TIP4P-TPSS and TIP4P-TPSS-D3 yield a too high pressure. With an $AAD = 39$ %, the TIP4P-TPSS-D3 water model is somewhat better than the TIP4P-TPSS potential ($AAD = 51$ %). The potential of Huang *et al.* exhibits good agreement at lower temperatures with increasing deviations for higher temperatures, leading to an $AAD = 30$ %. Yet, the best agreement with the reference EOS can be observed for the TIP4P/2005 force field. This also applies to the isothermal compressibility β_T with an $AAD = 7.3$ %. The DFT-derived TIP4P-TPSS and TIP4P-TPSS-D3 water models tend to underestimate the

isothermal compressibility, especially at lower densities. However, the largest AAD of 26 % is observed for the potential of Huang *et al.*, which is mainly due to larger deviations at higher densities. The rather large statistical uncertainties for the potential of Huang *et al.* and the TIP4P/2005 force field at 700 K and 24 mol/l indicate the proximity to the critical point.

The results for the residual enthalpy, speed of sound and Joule-Thomson coefficient can be found in Fig. 9. Regarding the former, the TIP4P/2005 force field yields the largest deviations from the reference EOS[99] and underestimates it throughout. By contrast, both the TIP4P-TPSS and TIP4P-TPSS-D3 water models overestimate the residual enthalpy. The AAD of the DFT-derived TIP4P-TPSS and TIP4P-TPSS-D3 water models are 20 % and 16 %, respectively. Good agreement with the reference EOS, especially for higher densities, can be found for the potential of Huang *et al.* with an $AAD = 11$ %. Regarding the speed of sound, the TIP4P/2005 force field yields the best results, i.e. $AAD = 2.6$ %. The results of the TIP4P-TPSS and TIP4P-TPSS-D3 water models are again very similar to each other and lead to an $AAD = 6$ % to 7 %. It has to be noted that the agreement to the reference EOS[99] of those potentials at 58 mol/l is even better than the agreement of the TIP4P/2005 force field. The worst agreement with the reference for the speed of sound was found for the potential of Huang *et al.* The deviations are between 20 and 30 % for the two highest isochores.

In case of the Joule-Thomson coefficient, a relative evaluation of the simulation data in comparison to the reference EOS is not applicable because this property changes its sign. Therefore, absolute values were considered. Except for the 6 mol/l isochore, the TIP4P/2005 force field performs quite well, yielding an average deviation of 0.1 K/MPa. It can be seen that the potential of Huang *et al.*[87] shows a good agreement throughout. Here, average deviations of 0.07 K/MPa can be found. The DFT-derived TIP4P-TPSS and TIP4P-TPSS-D3 water models underestimate the Joule-Thomson coefficient throughout, yielding higher average deviations, i.e. 0.2 K/MPa.

4.4 Transport properties

Transport properties were calculated by MD in the NpT ensemble with the Green-Kubo formalism[89,90]. Simulation results for the translational self-diffusion coefficient and the shear viscosity of the four discussed potentials are presented in Fig. 10 at 0.1 MPa for temperatures between 280 K and 340 K, respectively. The results for the translational self-diffusion coefficient were compared to experimental measurements by Harris and Woolf[108], Dullien[109], Holz *et al.*[110], and Easteal *et al.*[111], whereas the shear viscosity data were compared to a correlation of experimental data by Huber *et al.*[100]. The TIP4P/2005 force field results are in good agreement with the references for both considered properties. The shear viscosity is slightly overestimated, but nonetheless is, within the statistical uncertainties of the MD simulations, in agreement with the correlation of Huber *et al.*[100]. However, it can be seen that DFT-derived TIP4P-TPSS and TIP4P-TPSS-D3 water models, as well as the potential of Huang *et al.* significantly overestimate the translational self-diffusion coefficient. As a consequence, a significant underestimation of the shear viscosity is found for the latter three potentials.

4.5 Radial distribution function

For the purpose to determine the liquid water structure, the RDF in the liquid state at 298 K and 0.1 MPa were sampled in the NpT ensemble. The RDF $g(r)$ relates the local particle density $\rho^L(r)$ around a position within a molecule to the overall density $\rho = N/V$

$$g(r) = \frac{\rho^L(r)}{\rho} = \frac{1}{\rho}\frac{dN(r)}{dV} = \frac{1}{4\pi r^2 \rho}\frac{dN(r)}{dr}, \tag{15}$$

where $dN(r)$ is the differential number of atoms in a spherical shell volume element dV with the width dr, which is located at a distance r from the regarded position within a molecule[112]. For comparison, very recent experimental measurements from Skinner *et al.*[83] and Soper[84] were selected. In Fig. 11, the three partial RDF for O-H, O-O and H-H are shown, which were calculated using the four considered water potentials. It is apparent that the TIP4P/2005 force field as well as the two DFT-derived water models qualitatively reproduce the structure of liquid water. The potential of Huang *et al.*, however, is not able to correctly yield the first maximum of the O-O RDF, failing one of the most crucial

experimental tests of water models[113]. In addition, it also incapable to reproduce all further solvation shells of liquid water. This is most likely a result of the intramolecular O-H bond length r_{OH} of this potential, which is roughly 0.4 Å longer than the O-H bond length of the other potentials. In the light of the recent revision of the experimental RDF[83,84], it is particularly impressive that the DFT-derived water models exhibit the best agreement with the experimental reference, despite the fact that they have not been considered during their parametrization. In fact, they are even able to reproduce the correct ordering of the relative O-H peak heights. As such, they even outperform the well-established TIP4P/2005 force field, which is arguably the current benchmark among rigid, non-polarizable water potentials[50].

5 Conclusion

To summarize, in this paper we have assessed the accuracy of DFT-derived water models, which were recently developed based on an improved force-matching method[62]. To that extent the resulting water models are compared with the potential of Huang, which was optimized to reproduce the vapor-liquid equilibrium[87], and the well-appreciated TIP4P/2005 force field[48] as well as experimental measurements and a reference class EOS[99,100]. Specifically, it is found that the TIP4P/2005 force field exhibits a rather well-balanced accuracy for a large variety of different observables, in particular various thermodynamic properties. The potential of Huang *et al.* is indeed also an accurate interaction potential for multiple thermodynamic properties, but falls short of qualitatively reproducing the structure of liquid water. As opposed to this, both DFT-derived water models are only able to qualitatively reproduce thermodynamic quantities, but more importantly are in best agreement with the most recent experimental RDF, even superseding the TIP4P/2005 force field. Nevertheless, there are many improvements that could be made to built on this study. In particular, the TIP4P/TPSS-D3 water model developed here is not polarizable, an effect which has the potential to further improve the agreement with experiment and that in the future could be easily incorporated into the current force-matching scheme.

We conclude by noting that the present force-matching scheme will facilitate to routinely

parametrize computationally efficient yet predictive potential energy functions based on accurate *ab-initio* MD simulations for a large variety of different systems. The development of transferable interaction potentials for more complex liquids, which are in good qualitative agreement with the experiment, is work in progress and will be reported elsewhere.

Acknowledgements

We gratefully acknowledge the Paderborn Center for Parallel Computing (PC2) for the generous allocation of computer time on the OCuLUS cluster.

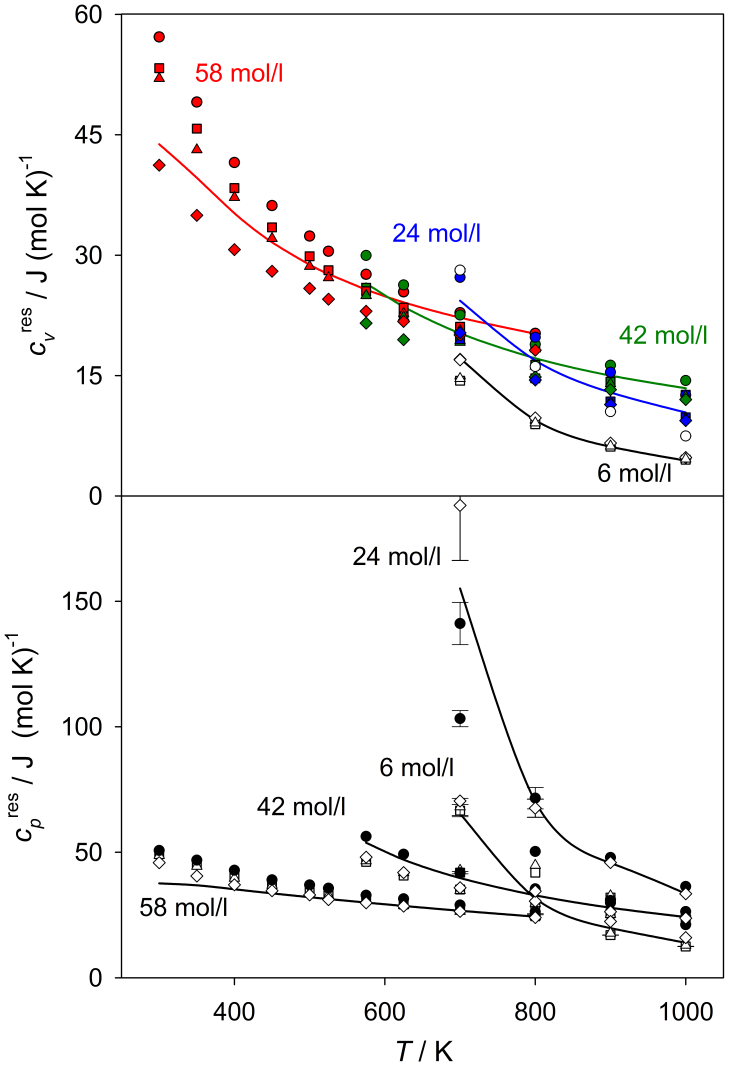


Figure 7: Residual isochoric (top) and residual isobaric heat capacity (bottom) along four isochores in the homogeneous fluid region of water: (—) Reference fundamental equation of state by Wagner and Pruß[99]. Computational data: (•) TIP4P/2005; (◇) Huang *et al.*; (□) TIP4P-TPSS; (△) TIP4P-TPSS-D3. The statistical uncertainties of the simulations are shown if they exceed symbol size.

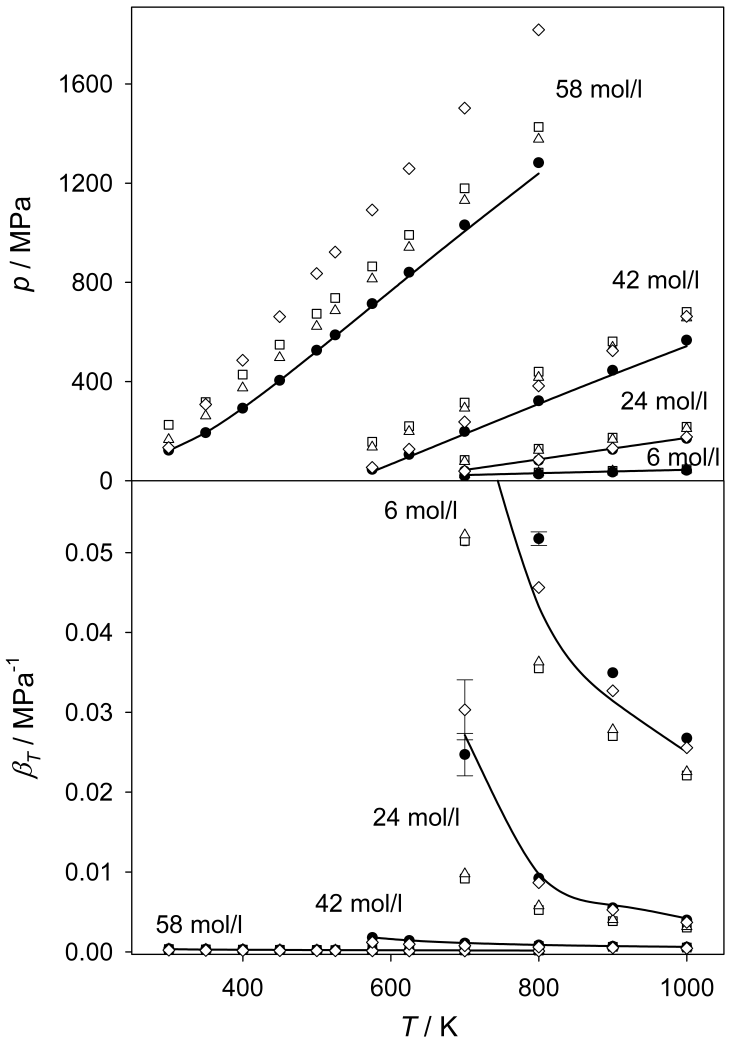


Figure 8: Pressure (top) and isothermal compressibility (bottom) along four isochores in the homogeneous fluid region of water: (—) Reference fundamental equation of state by Wagner and Pruß[99]. Computational data: (•) TIP4P/2005; (◇) Huang *et al.*; (□) TIP4P-TPSS; (△) TIP4P-TPSS-D3. The statistical simulation uncertainties are shown if they exceed symbol size.

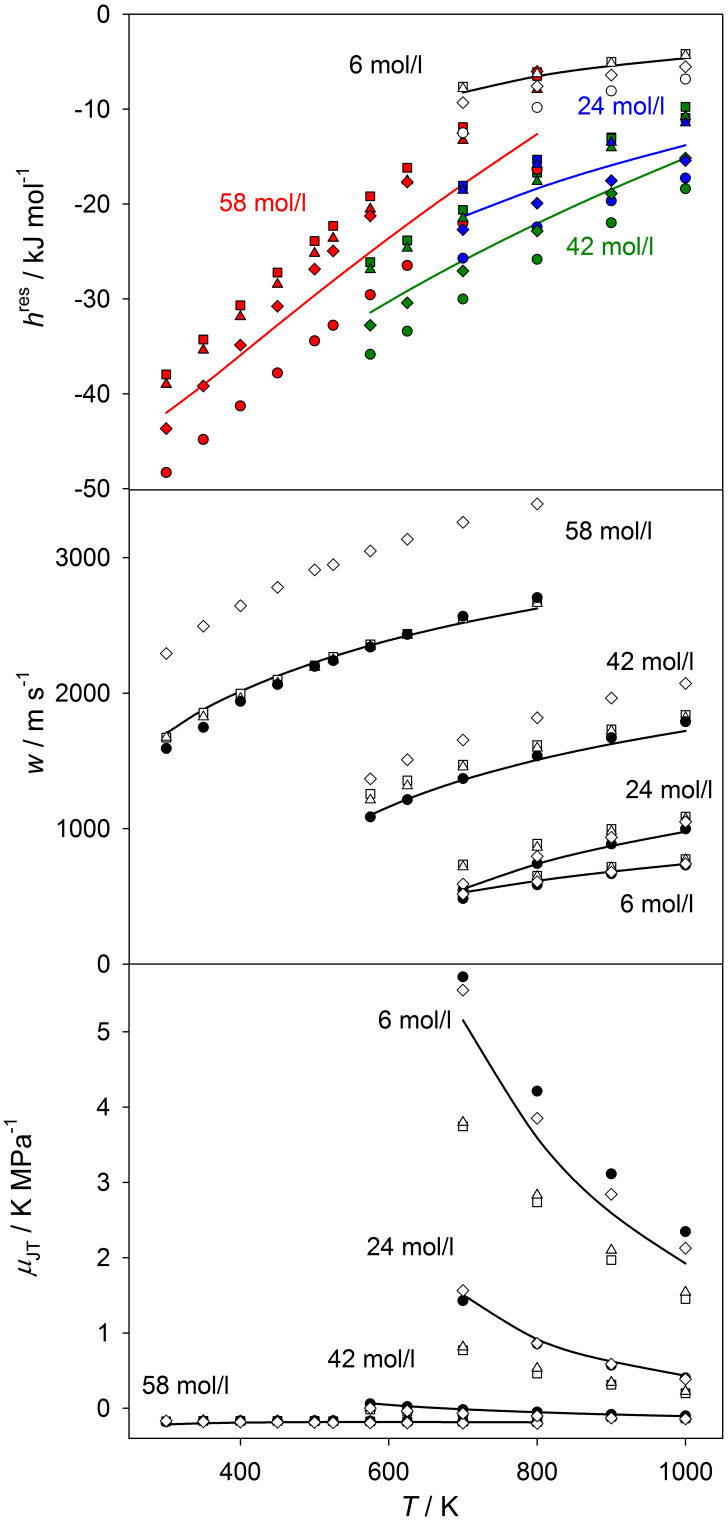


Figure 9: Residual enthalpy (top), speed of sound (center) and Joule-Thomson coefficient (bottom) along four isochores in the homogeneous fluid region of water: (—) Reference fundamental equation of state by Wagner and Pruß[99]. Computational data: (•) TIP4P/2005; (◇) Huang *et al.*; (□) TIP4P-TPSS; (△) TIP4P-TPSS-D3. The statistical simulation uncertainties are within symbol size.

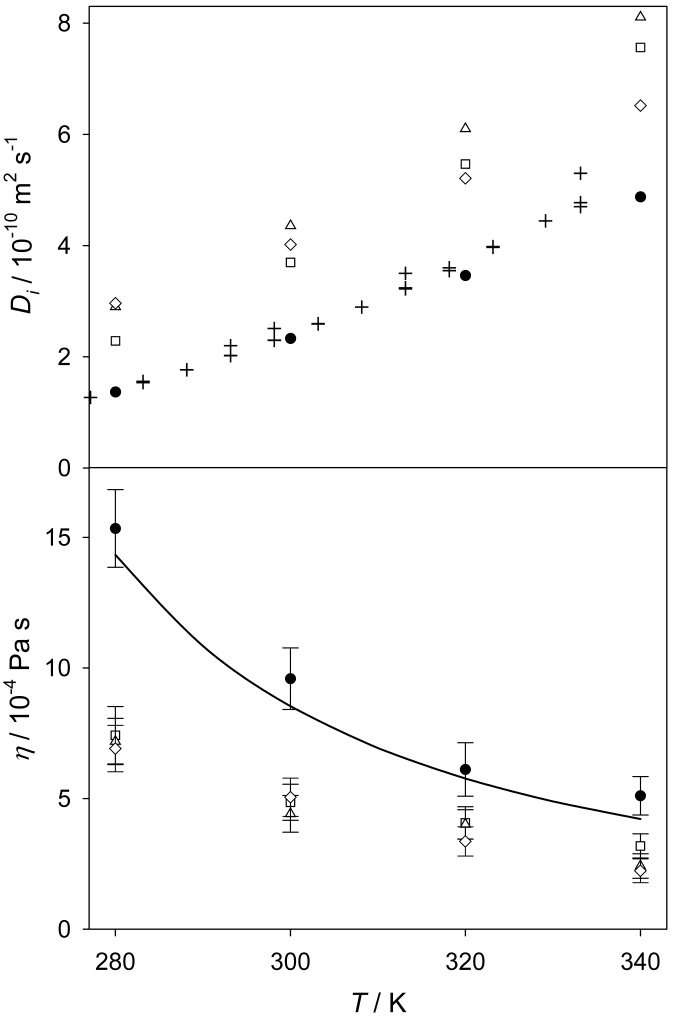


Figure 10: Self-diffusion coefficient (top) and shear viscosity (bottom) of liquid water at 0.1 MPa. (+) Experimental data[108–111]. (—) Shear viscosity correlation of experimental data by Huber *et al.*[100]. Computational data: (●) TIP4P/2005; (◇) Huang *et al.*; (□) TIP4P-TPSS; (△) TIP4P-TPSS-D3. The statistical uncertainties of the MD simulations are shown only if they exceed symbol size.

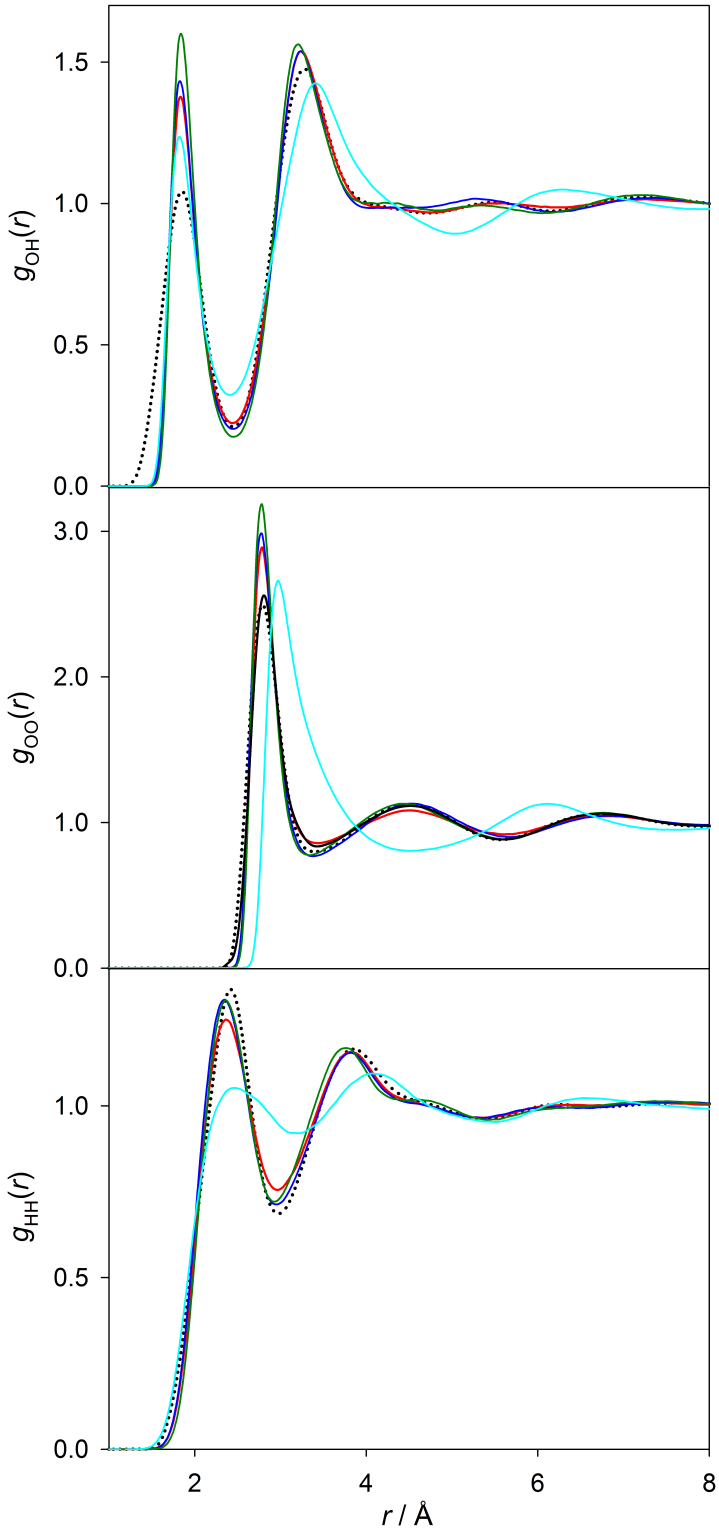


Figure 11: Radial distribution functions in the liquid state at 298 K and 0.1 MPa: (—) Skinner *et al.*[83]; (...) Soper[84]; (—) TIP4P-TPSS; (—) TIP4P-TPSS-D3; (—) TIP4P/2005; (—) Huang *et al.*

References

1. A. Rahman and F. H. Stillinger, J. Chem. Phys. **55**, 3336 (1971).

2. D. Eisenberg and W. Kauzmann, *The Structure and Properties of Water* (Oxford University Press, Oxford, 1969).

3. M. Chaplin, Nature Rev. Mol. Cell Biol. **7**, 861 (2006).

4. P. Ball, Chem. Rev. **108**, 74 (2008).

5. S. Narayan, V. V. Fokin, and K. B. Sharpless, in *Organic Reactions in Water: Principles, Strategies and Applications*, edited by U. M. Lindstrom (Blackwell Publishing, Oxford, 2007).

6. S. Narayan, J. Muldoon, M. G. Finn, V. V. Fokin, H. C. Kolb, and K. B. Sharpless, Angew. Chem. Int. Ed. **44**, 3275 (2005).

7. Y. Jung and R. A. Marcus, J. Am. Chem. Soc. **129**, 5492 (2007).

8. K. Karhan, R. Z. Khaliullin, and T. D. Kühne, J. Chem. Phys. **141**, 22D528 (2014).

9. F. H. Stillinger, Science **209**, 451 (1980).

10. R. Ludwig, Angew. Chem. Int. Ed. **40**, 1808 (2001).

11. M. Dömer, T. Spura, R. Z. Khaliullin, and T. D. Kühne, Nachr. Chemie **61**, 1203 (2013).

12. R. Ludwig and D. Paschek, Chem. Unserer Zeit **39**, 164 (2005).

13. P. Ball, Nature **452**, 291 (2008).

14. T. D. Kühne and R. Z. Khaliullin, Nature Commun. **4**, 1450 (2013).

15. R. Z. Khaliullin and T. D. Kühne, Phys. Chem. Chem. Phys. **15**, 15746 (2013).

16. D. Marx and J. Hutter, *Ab Initio Molecular Dynamics - Basic Theory and Advanced Methods* (Cambridge University Press, Cambridge, 2009).

17. T. D. Kühne, WIREs Comput. Mol. Sci. **4**, 391 (2014).

18. T. D. Kühne, P. Partovi-Azar, and H. Elgabarty, Nachr. Chemie **63**, 327 (2015).

19. K. Laasonen, M. Sprik, M. Parrinello, and R. Car, J. Chem. Phys. **99**, 9080 (1993).

20. M. Sprik, J. Hutter, and M. Parrinello, J. Chem. Phys. **105**, 1142 (1996).

21. J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, J. Chem. Phys. **120**, 300 (2004).

22. I.-F. W. Kuo, C. J. Mundy, M. J. McGrath, J. I. Siepmann, J. VandeVondele, M. Sprik, J. Hutter, B. Chen, M. L. Klein, F. Mohamed, et al., J. Phys. Chem. B **108**, 12990 (2004).

23. P. H.-L. Sit and N. Marzari, J. Chem. Phys. **122**, 204510 (2005).

24. T. Todorova, A. P. Seitsonen, J. Hutter, I.-F. W. Kuo, and C. J. Mundy, J. Phys. Chem. B **110**, 3685 (2006).

25. J. A. Morrone and R. Car, Phys. Rev. Lett. **101**, 17801 (2008).

26. J. Schmidt, J. VandeVondele, I.-F. W. Kuo, D. Sebastiani, J. I. Siepmann, J. Hutter, and C. J. Mundy, J. Phys. Chem. B **113**, 11959 (2009).

27. T. D. Kühne, M. Krack, and M. Parrinello, J. Chem. Theory Comput. **5**, 235 (2009).

28. J. Wang, G. Roman-Perez, J. M. Soler, E. Artacho, and M.-V. Fernandez-Serra, J. Chem. Phys. **134**, 024516 (2011).

29. C. Zhang, D. Donadio, F. Gygi, and G. Galli, J. Chem. Theory Comput. **7**, 1443 (2011).

30. I.-C. Lin, A. P. Seitsonen, I. Tavernelli, and U. Rothlisberger, J. Chem. Theory Comput. **8**, 3902 (2012).

31. T. A. Pascal, D. Schärf, Y. Jung, and T. D. Kühne, J. Chem. Phys. **137**, 244507 (2012).

32. Z. Ma, Y. Zhang, and M. E. Tuckerman, J. Chem. Phys. **137**, 044506 (2012).

33. C. Zhang, R. Z. Khaliullin, D. Bovi, L. Guidoni, and T. D. Kühne, J. Phys. Chem. Lett. **4**, 3245 (2013).

34. M. Ceriotti, J. Cuny, M. Parrinello, and D. E. Manolopoulos, Proc. Nat. Acad. Sci. USA **110**, 15591 (2013).

35. A. A. Hassanali, F. Giberti, J. Cuny, T. D. Kühne, and M. Parrinello, Proc. Nat. Acad. Sci. USA **110**, 13723 (2013).

36. R. A. DiStasio Jr., B. Santra, Z. Li, X. Wu, and R. Car, J. Chem. Phys. **141**, 084502 (2014).

37. T. D. Kühne and R. Z. Khaliullin, J. Am. Chem. Soc. **136**, 3395 (2014).

38. K. Forster-Tonigold and A. Groß, J. Chem. Phys. **141**, 064501 (2014).

39. H. Elgabarty, R. Z. Khaliullin, and T. D. Kühne, Nature Commun. **6**, 8318 (2015).

40. M. Del Ben, J. Hutter, and J. VandeVondele, J. Chem. Phys. **143**, 054506 (2015).

41. C. Zhang, L. Guidoni, and T. D. Kühne, J. Mol. Liq. **205**, 42 (2015).

42. S. Y. Willow, M. A. Salim, K. S. Kim, and S. Hirata, Sci. Rep. **5**, 14358 (2015).

43. T. D. Kühne, M. Krack, F. R. Mohamed, and M. Parrinello, Phys. Rev. Lett. **98**, 66401 (2007).

44. A. Rahman and F. H. Stillinger, J. Chem. Phys. **60**, 1545 (1974).

45. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

46. H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).

47. M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).

48. J. L. F. Abascal and C. Vega, J. Chem. Phys. **123**, 234505 (2005).

49. B. Guillot, J. Mol. Liq. **101**, 219 (2002).

50. C. Vega and J. L. F. Abascal, Phys. Chem. Chem. Phys. **13**, 19663 (2011).

51. G. C. Lie and E. Clementi, Phys. Rev. A **33**, 2679 (1986).

52. A. P. Lyubartsev and A. Laaksonen, Phys. Chem. Lett. **325**, 15 (2000).

53. S. Izvekov, M. Parrinello, C. J. Burnham, and G. A. Voth, J. Chem. Phys. **120**, 10896 (2004).

54. R. Bukowski, K. Szalewicz, G. C. Groenenboom, and A. van der Avoird, Science **315**, 1249 (2007).

55. G. S. Fanourgakis and S. S. Xantheas, J. Chem. Phys. **128**, 074506 (2008).

56. V. Babin, G. R. Medders, and F. Paesani, J. Phys. Chem. Lett. **3**, 3765 (2012).

57. G. R. Medders, A. W. Götz, M. A. Morales, P. Bajaj, and F. Paesani, J. Chem. Phys. **143**, 104102 (2015).

58. M. Ostheimer and H. Bertagnolli, Mol. Phys. **3**, 227 (1989).

59. A. P. Lyubartsev and A. Laaksonen, Phys. Rev. E **52**, 3730 (1995).

60. D. Reith, M. Pütz, and F. Müller-Plathe, J. Comp. Chem. **24**, 1624 (2003).

61. F. Ercolessi and J. B. Adams, Europhys. Lett. **26**, 583 (1994).

62. T. Spura, C. John, S. Habershon, and T. D. Kühne, Mol. Phys. **113**, 808 (2014).

63. R. L. Henderson, Phys. Lett. A **49**, 197 (1974).

64. O. Akin-Ojo, Y. Song, and F. Wang, J. Chem. Phys. **129**, 064108 (2008).

65. J. Sala, E. Guardia, and M. Masia, Comp. Phys. Commun. **182** (2011).

66. J. Sala, E. Guardia, J. Marti, D. Spangberg, and M. Masia, J. Chem. Phys. **136** (2012).

67. D. Kraft, ACM Trans. Math. Software **20**, 262 (1994).

68. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes* (Cambridge University Press, Cambridge, 1992).

69. R. P. Brent, *Algorithms for Minimization without Derivatives* (Prentice-Hall, Englewood Cliffs, 1973).

70. S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. Miller, Annu. Rev. Phys. Chem. **64**, 387 (2013).

71. S. Habershon, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **131**, 024501 (2009).

72. W. Kohn, Rev. Mod. Phys. **71**, 1253 (1999).

73. R. O. Jones, Rev. Mod. Phys. **87**, 897 (2015).

74. G. Lippert, J. Hutter, and M. Parrinello, Mol. Phys. **92** (1997).

75. J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, Comp. Phys. Commun. **167**, 103 (2005).

76. J. VandeVondele and J. Hutter, J. Chem. Phys. **127**, 114105 (2007).

77. J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).

78. S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).

79. M. Krack, Theor. Chem. Acc. **114**, 145 (2005).

80. J. Hepburn, G. Scoles, and R. Penco, Chem. Phys. Lett. **36**, 451 (1975).

81. R. Ahlrichs, R. Penco, and G. Scoles, Chem. Phys. **19**, 119 (1977).

82. S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).

83. L. B. Skinner, C. Huang, D. Schlesinger, L. G. M. Pettersson, A. Nilsson, and C. J. Benmore, J. Chem. Phys. **138**, 074506 (2013).

84. A. K. Soper, ISRN Physical Chemistry **2013**, 297463 (2013).

85. M. A. Morales, J. R. Gergely, J. McMinis, J. M. McMahon, J. Kim, and D. M. Ceperley, J. Chem. Theory Comput. **10**, 2355 (2014).

86. M. J. Gillan, D. Alfe, and A. Michaelides, J. Chem. Phys. **144**, 130901 (2016).

87. Y.-L. Huang, T. Merker, M. Heilig, H. Hasse, and J. Vrabec, Ind. Eng. Chem. Res. **51**, 7428 (2012).

88. C. W. Glass, S. Reiser, G. Rutkai, S. Deublein, A. Köster, G. Guevara-Carrion, A. Wafai, M. Horsch, M. Bernreuther, T. Windmann, et al., Comp. Phys. Commun, **185**, 3302 (2014).

89. M. S. Green, J. Chem. Phys. **22**, 398 (1954).

90. R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).

91. D. Frenkel and B. Smit, *Understanding Molecular Simulation : from algorithms to applications* (Academic Press, San Diego, 2002).

92. M. Griebel, S. Knapek, and G. Zumbusch, *A Guide to Monte Carlo Simulations in Statistical Physics* (Springer-Verlag, Berlin, 2007).

93. K. Binder, Rep. Prog. Phys. **60**, 487 (1997).

94. D. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, 2013).

95. R. Lustig, Mol. Phys. **65**, 175 (1988).

96. M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Oxford University Press, Oxford, 1989).

97. H. C. Andersen, J. Chem. Phys. **72**, 2384 (1980).

98. R. Span, *Multiparameter Equations of State: An Accurate Source of Thermodynamic Property Data* (Springer-Verlag, Berlin, 2000).

99. W. Wagner and A. Pruß, J. Phys. Chem. Ref. Data **31**, 387 (2002).

100. M. Huber, R. Perkins, A. Laesecke, D. Friend, J. Sengers, M. Assael, I. Metaxa, E. Vogel, R. Mareš, and K. Miyagawa, J. Phys. Chem. Ref. Data **38**, 101 (2009).

101. B. Eckl, J. Vrabec, and H. Hasse, Mol. Phys. **106**, 1039 (2008).

102. R. L. Rowley, W. V. Wilding, J. L. Oscarson, Y. Yang, , N. A. Zundel, T. E. Daubert, and R. P. Danner, *DIPPR Data Compilation of Pure Compound Properties: Design Institute for Physical Properties* (AIChE, New York, 2006).

103. J. Vrabec and H. Hasse, Mol. Phys. **100**, 3375 (2002).

104. I. Nezbeda and J. Kolafa, Mol. Sim. **5**, 391 (1991).

105. J. Vrabec, M. Kettler, and H. Hasse, Chem. Phys. Lett. **356**, 431 (2002).

106. R. Lustig, Mol. Sim. **37**, 457 (2011).

107. R. Lustig, Mol. Phys. **110**, 3041 (2012).

108. K. R. Harris and L. A. Woolf, J. Chem. Soc., Faraday Trans. 1 **76**, 377 (1980).

109. F. A. L. Dullien, AIChE J. **18**, 62 (1972).

110. M. Holz, S. R. Heil, and A. Sacco, Phys. Chem. Chem. Phys. **2**, 4740 (2000).

111. A. J. Easteal, W. E. Price, and L. A. Woolf, J. Chem. Soc., Faraday Trans. 1 **85**, 1091 (1989).

112. K. A. F. Röhrig and T. D. Kühne, Phys. Rev. E **87**, 045301 (2013).

113. W. L. Jorgensen and J. Tirado-Rives, Proc. Nat. Acad. Sci. USA **102**, 6665 (2005).