

# Extrapolation and interpolation strategies for efficiently estimating structural observables as a function of temperature and density

Jacob I. Monroe,<sup>1, a)</sup> Harold W. Hatch,<sup>1</sup> Nathan A. Mahynski,<sup>1</sup> M. Scott Shell,<sup>2</sup> and Vincent K. Shen<sup>1, b)</sup>

<sup>1)</sup>*National Institute of Standards and Technology, Gaithersburg, MD*

<sup>2)</sup>*University of California - Santa Barbara, Santa Barbara, CA*

(Dated: 20 October 2020)

Thermodynamic extrapolation has previously been used to predict arbitrary structural observables in molecular simulations at temperatures (or relative chemical potentials in open-system mixtures) different from those at which the simulation was performed. This greatly reduces the computational cost in mapping out phase and structural transitions. In this work we explore the limitations and accuracy of thermodynamic extrapolation applied to water, where qualitative shifts from anomalous to simple-fluid-like behavior are manifested through shifts in the liquid structure that occur as a function of both temperature and density. We present formulae for extrapolating in volume for canonical ensembles and demonstrate that linear extrapolations of water’s structural properties are only accurate over a limited density range. On the other hand, linear extrapolation in temperature can be accurate across the entire liquid state. We contrast these extrapolations with classical perturbation theory techniques, which are more conservative and slowly converging. Indeed, we show that such behavior is expected by demonstrating exact relationships between extrapolation of free energies and well-known techniques to predict free energy differences. An ideal gas in an external field is also studied to more clearly explain these results for a toy system with fully analytical solutions. We also present a recursive interpolation strategy for predicting arbitrary structural properties of molecular fluids over a predefined range of state conditions, demonstrating its success in mapping qualitative shifts in water structure with density.

## I. INTRODUCTION

With modern advances in computational power and algorithms, data generated via molecular simulation has become ubiquitous across numerous fields.<sup>1,2</sup> Despite the availability of past data and the ease with which new results may be generated, it is crucial that data be generated and utilized efficiently. Rather than improve computational efficiency directly through enhanced sampling algorithms<sup>3,4</sup> or refining hardware usage,<sup>5–11</sup> we focus here on general developments of statistical mechanical theory. It is often the case that simulation data is desired to demonstrate a trend with a specific state condition or other variable. For instance, it might be of interest to know the unfolding behavior of a protein as a function of temperature, or how the adsorption of fouling species at an interface changes over a wide range of temperatures, pressures, or even some adjustable parameter of the solute or interface. Many such quantities may be considered averages over configurations generated through Monte Carlo, molecular dynamics, or *ab initio* simulations. Recent work has provided “mapped averaging” formulas for calculating structural properties, greatly reducing the simulation time needed to achieve highly precise estimates of physical observables.<sup>12–14</sup> However, this technique does not address the need to run multiple simulations at different state points, which can consume time.

Complementary to mapped averaging, thermodynamic extrapolation<sup>15–18</sup> provides an efficient route to estimate the variation of observables with changes in simulation conditions, such as temperature or pressure. Thermodynamic extrapolation is based on statistical mechanical relationships

that, via a single simulation, directly provide estimates of the derivatives of an observable with respect to a specific variable. This technique has met with great success in predicting grand canonical number density distributions across temperatures<sup>17,18</sup> and relative chemical potentials.<sup>15</sup> It has also proven effective in predicting changes in virial coefficients upon alchemical variation of simulation parameters.<sup>19</sup> A number of structural properties as a function of temperature have also been successfully predicted for relatively simple model systems.<sup>20</sup> Use of extrapolation expressions to provide temperature derivatives of dynamical quantities has been developed independently<sup>21,22</sup> and used to more efficiently determine activation energies<sup>21</sup> and fit kinetic expressions associated with structural reorganization of water.<sup>23</sup>

An open question remains, however, as to the general accuracy and precision of these techniques and their relevance to more realistic fluids. For instance, liquid water exhibits a number of anomalies as a function of state conditions.<sup>24,25</sup> With changes to temperature and density, water’s fundamental structure may flip between what is expected for tetrahedral and simple fluids.<sup>26–28</sup> Such changes are expected to be difficult for thermodynamic extrapolation to capture — indeed, we clearly demonstrate that this technique is subject to the same pitfalls as any extrapolation strategy. Interestingly, a recent study utilizing first-order extrapolation techniques successfully captured the temperature dependence of oxygen-oxygen radial distribution functions of liquid water.<sup>29</sup> We find, however, that this is not necessarily general behavior. Namely, the success of the method relies on accurately capturing the local curvature in the observable of interest over a desired range of the extrapolation variable (e.g., temperature or density).

To better understand the errors inherent in extrapolation, we examine a toy ideal gas system for which analytic solutions exist. We develop intuition for the scaling of uncertainties in this toy system and propose practical guidelines for the

<sup>a)</sup>Electronic mail: jacob.monroe@nist.gov

<sup>b)</sup>Electronic mail: vincent.shen@nist.gov

use of extrapolation techniques. We also connect these results to rigorous theories underlying well-known free energy calculation techniques, which allows us to establish strategies and heuristics for consistently obtaining high accuracy from derivative estimation methods. Comparison of our methods to standard re-weighting of observables reveals significant gains in accuracy and efficiency, which we show is expected from previous literature concerning free energy calculations. Using these insights, we propose a polynomial interpolation approach based on thermodynamic extrapolation for efficiently using simulation data at multiple conditions, with popular free energy calculation techniques appearing as special cases of its implementation. Our presented algorithms are packaged into an easily extensible Python library that we use here to analyze both the ideal gas model and simulations of water across a wide range of liquid temperatures and densities. While we find linear extrapolation is capable of capturing the variation of many of water’s structural properties with temperature, more advanced strategies are necessary to describe its behavior with density.

We organize the paper as follows: Section II presents thermodynamic extrapolation and formalizes connections between estimation of general observables and free energies. Section II A covers various classes of observables and how extrapolation must be adapted to each, while II B discusses the differences between extrapolation techniques and re-weighting via standard perturbation theories.<sup>30</sup> Section III thoroughly investigates extrapolation, including extensive uncertainty analysis, for an analytic ideal gas system. With the lessons learned from this toy model, we propose a recursive interpolation algorithm in Section IV, providing fundamental connections to free energy calculation techniques, as well as a simple visual check for self-consistency. Section V applies all of these results to simulations of liquid water, testing their applicability in a more realistic system with non-trivial shifts in qualitative structural behavior.

## II. THERMODYNAMIC EXTRAPOLATION AND ITS RELATIONSHIP TO RE-WEIGHTING

### A. Classifications of thermodynamic extrapolation

Thermodynamic extrapolation is related to re-weighting techniques. To illustrate this, we first introduce common use cases that differ in the functional form of the quantity being extrapolated and its dependence on the extrapolation variable, with free energies appearing as our last example. This will be important later when we demonstrate that use of derivative information to perform polynomial interpolation with multiple data points is equivalent to both thermodynamic integration<sup>31</sup> and optimal combinations of cumulant expansions.<sup>32</sup> Though not completely exhaustive, the following classifications cover most cases of interest in molecular simulation. We define the function to be extrapolated as  $Y$ , the extrapolation variable as  $\alpha$ , and some arbitrary simulation observable to be  $X$ . For instance, if  $Y$  is the average end-to-end distance of a polymer chain and we want to know how this varies with temperature

( $\alpha$ ), then  $X$  would be the end-to-end distance at each simulation snapshot. In the simplest case, we want to average an observable of interest that does not explicitly depend on  $\alpha$

$$Y = \langle X \rangle \quad (1)$$

This applies to most structural observables, such as radial distribution functions (RDFs), the radius of gyration of a polymer, or particle density, when extrapolating over temperature. The second case involves a structural observable that explicitly depends on the extrapolation variable.

$$Y = \langle X(\alpha) \rangle \quad (2)$$

Though the above dependence is not common for most average observables extrapolated in temperature ( $\alpha = T$ ), it is highly relevant to extrapolation over volume or simulation parameters (such as those appearing in a potential energy function). An example is the variation of a canonical ensemble RDF ( $Y$ ) with respect to volume ( $\alpha = V$ ) where  $X$  is then the number of bin counts at a specific separation distance for a single system configuration. Though the first case is treated exhaustively in previous literature,<sup>16,17</sup> we reproduce a simplified version of its derivation in the next section. Thorough investigation of the second case has also been performed<sup>17</sup> and will not be repeated here, though we note that the released library of code is capable of handling such scenarios.

The next two cases are simple extensions of the first two. For quantities like excess chemical potentials, we are not interested in the average observable itself, but instead the negative of its natural logarithm.

$$Y = -\ln \langle X \rangle \quad (3)$$

$$Y = -\ln \langle X(\alpha) \rangle \quad (4)$$

Again,  $X$  may implicitly or explicitly depend on the extrapolation variable. The latter is true, for example, in the case of excess chemical potentials in the canonical ensemble given by  $\beta\mu_{\text{ex}} = -\ln \langle e^{-\beta\Delta U} \rangle$  with  $\beta$  as the inverse temperature and  $\Delta U$  as the potential energy difference with and without the added molecule. For the case of hard-sphere excess chemical potentials, however, only Eq. 3 is necessary as the explicit temperature derivatives go to zero. The latter two cases in Eqs. 3 and 4 may be simply handled in terms of the first two by noting that the derivative of the negative natural logarithm of a function can be expressed in closed form at arbitrary order in terms of the function itself and its derivatives up to the same order.

$$\frac{\partial^n Y}{\partial \alpha^n} = \sum_{k=1}^n (k-1)! \left( \frac{-1}{\langle X \rangle} \right)^k \times B_{n,k} \left( \frac{\partial \langle X \rangle}{\partial \alpha}, \frac{\partial^2 \langle X \rangle}{\partial \alpha^2}, \dots, \frac{\partial^{n-k+1} \langle X \rangle}{\partial \alpha^{n-k+1}} \right) \quad (5)$$

The above expression may be derived from Faà di Bruno’s formula<sup>33</sup> where  $B_{n,k}$  represents the partial Bell polynomials.<sup>34</sup>

Finally, the function of interest may be a free energy, in which case we wish to extrapolate the natural logarithm of a

partition function that will in most cases of interest have an explicit dependence on the extrapolation variable.

$$Y = -\ln X(\alpha) \quad (6)$$

Extrapolation of this quantity in terms of temperature for the canonical ensemble is almost identical to the cumulant expansions of Zwanzig in his original perturbation theory.<sup>30</sup> In contrast to our work, Zwanzig focused on the free energy difference with respect to a specific change in the potential energy function. More importantly, Zwanzig always uses the infinite temperature state (inverse temperature  $\beta = 1/k_B T$  equal to zero) as a reference when developing his expansions (hence the “high temperature equation of state”). If we wish to directly represent the free energy at a particular  $\beta$  in terms of a free energy and its derivatives at a reference  $\beta_0$ , we may expand to obtain the following.

$$\begin{aligned} \beta A(\beta) &= \beta_0 A(\beta_0) - \left. \frac{\partial \ln Q}{\partial \beta} \right|_{\beta_0} (\beta - \beta_0) \\ &\quad - \frac{1}{2!} \left. \frac{\partial^2 \ln Q}{\partial \beta^2} \right|_{\beta_0} (\beta - \beta_0)^2 + \dots \\ \beta A(\beta) &= \beta_0 A(\beta_0) + \langle H \rangle_{\beta_0} (\beta - \beta_0) \\ &\quad - \frac{1}{2!} \left( \langle H^2 \rangle_{\beta_0} - \langle H \rangle_{\beta_0}^2 \right) (\beta - \beta_0)^2 + \dots \end{aligned} \quad (7)$$

The Helmholtz free energy  $A$  is related to the canonical ensemble partition function  $Q = \int e^{-\beta H(\mathbf{r}^N, \mathbf{p}^N)} d\mathbf{r}^N d\mathbf{p}^N$  as  $\beta A = -\ln Q$  with  $\mathbf{r}^N$  and  $\mathbf{p}^N$  representing the positions and momenta of all particles in the system. Averages and derivatives evaluated at specific values of the extrapolation variable are indicated with subscripts. Eq. 7 is a cumulant expansion in terms of the difference in  $\beta$  rather than  $\beta$  itself, with  $H$  being the Hamiltonian of the system, including both the potential and kinetic energies. A particularly useful form for many types of free energy calculations results from expanding in terms of an arbitrary variable in the Hamiltonian  $\lambda$

$$\begin{aligned} \beta A(\lambda) &= \beta A(\lambda_0) + \left\langle \beta \frac{\partial H}{\partial \lambda} \right\rangle_{\lambda_0} (\lambda - \lambda_0) \\ &\quad - \frac{1}{2!} \left( \left\langle \left( \beta \frac{\partial H}{\partial \lambda} \right)^2 \right\rangle_{\lambda_0} - \left\langle \beta \frac{\partial H}{\partial \lambda} \right\rangle_{\lambda_0}^2 \right) (\lambda - \lambda_0)^2 \\ &\quad + \dots \end{aligned} \quad (8)$$

### B. Extrapolation versus perturbation in the canonical ensemble

Before contrasting extrapolation with perturbation theory, we derive temperature extrapolation in the canonical ensemble, which corresponds to the simplest case described above (i.e.,  $Y = \langle X \rangle$ ). More thorough treatments, including extension to other variables, may be found elsewhere,<sup>15,17</sup> while more involved derivations for extrapolation of volume in the canonical ensemble are provided in the Appendix. For any observable  $X(\mathbf{r}^N)$  that only depends on coordinates  $\mathbf{r}^N$  of the

$N$  atoms in the system and does not depend on momenta or explicitly on temperature, we can write the extrapolation from  $\beta_0$  by an amount  $\delta\beta$  for the ensemble average of  $X$  as

$$\langle X \rangle_{\beta} = \langle X \rangle_{\beta_0} + \left. \frac{\partial \langle X \rangle}{\partial \beta} \right|_{\beta_0} (\delta\beta) + \left. \frac{\partial^2 \langle X \rangle}{\partial \beta^2} \right|_{\beta_0} \frac{(\delta\beta)^2}{2!} + \dots \quad (9)$$

For structural averages, the relevant normalization factor is the configurational part  $Z$  of the canonical ensemble partition function  $Q$  and the appropriate configuration weight is then  $\frac{e^{-\beta U(\mathbf{r}^N)}}{Z}$ , with  $U$  being the potential energy. The derivatives in Eq. 9 (up to second order in  $\beta$ ) become

$$\left. \frac{\partial \langle X \rangle}{\partial \beta} \right|_{\beta_0} = \langle X \rangle_{\beta_0} \langle U \rangle_{\beta_0} - \langle XU \rangle_{\beta_0} \quad (10)$$

$$\begin{aligned} \left. \frac{\partial^2 \langle X \rangle}{\partial \beta^2} \right|_{\beta_0} &= 2 \langle X \rangle_{\beta_0} \langle U \rangle_{\beta_0}^2 - \langle X \rangle_{\beta_0} \langle U^2 \rangle_{\beta_0} \\ &\quad - 2 \langle XU \rangle_{\beta_0} \langle U \rangle_{\beta_0} + \langle XU^2 \rangle_{\beta_0} \end{aligned} \quad (11)$$

As noted by Mahynski et al.,<sup>20</sup> configurational quantities of interest can almost always be expressed as ensemble averages. This includes each point in any structural distribution function such as a RDF.

To relate thermodynamic extrapolation to more well known re-weighting methods, we briefly revisit perturbation theory. In the original theory developed by Zwanzig,<sup>30</sup> we may re-weight averages in one ensemble in order to obtain the same average in another ensemble. For two canonical ensembles at different temperatures, with the observable having no explicit dependence on temperature or momenta, this becomes

$$\langle X \rangle_{\beta} = \frac{\langle X e^{-(\beta - \beta_0)U} \rangle_{\beta_0}}{\langle e^{-(\beta - \beta_0)U} \rangle_{\beta_0}} \quad (12)$$

Or more explicitly in terms of configuration weights in a given ensemble,  $w = \frac{e^{-\beta U}}{Z}$

$$\langle X \rangle_{\beta} = \frac{\langle X \frac{w}{w_0} \rangle_{\beta_0}}{\langle \frac{w}{w_0} \rangle_{\beta_0}} \quad (13)$$

As each is mathematically rigorous, perturbation and extrapolation utilizing an infinite number of terms should yield identical results. In practice, the results will differ due to finite sampling and series truncation, despite both techniques utilizing the exact same data, namely the potential energies and observable values for each sampled configuration. In each case, we manipulate the data by taking various averages, then combine them to produce the result. To illustrate differences between the two approaches, we Taylor expand the numerator and denominator in Eq. 12 around an inverse temperature difference  $\delta\beta$  of zero

$$\langle X \rangle_{\beta} = \frac{\langle X \rangle_{\beta_0} - \langle XU \rangle_{\beta_0} \delta\beta + \langle XU^2 \rangle_{\beta_0} \frac{\delta\beta^2}{2} - \dots}{1 - \langle U \rangle_{\beta_0} \delta\beta + \langle U^2 \rangle_{\beta_0} \frac{\delta\beta^2}{2} - \dots} \quad (14)$$

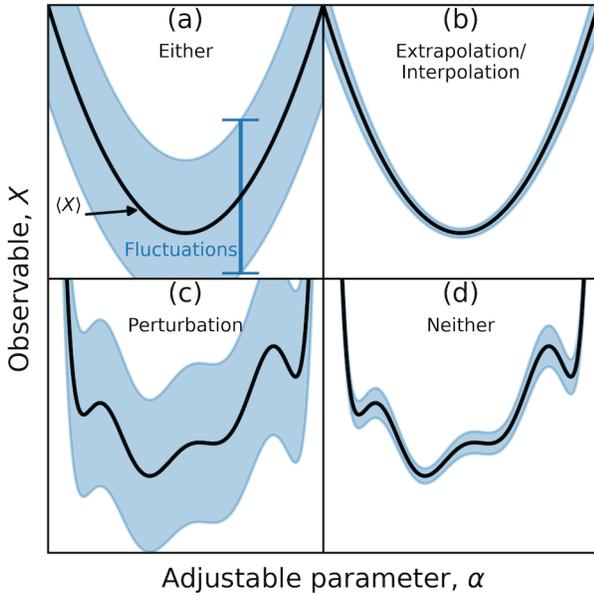


FIG. 1. Schematic representations of scenarios in which extrapolation/interpolation or perturbation (re-weighting) are expected to be efficient. When the behavior of  $X$  with respect to  $\alpha$  is well-approximated by a low-order polynomial (a and b), only derivatives of low order (here second for extrapolation, first for polynomial interpolation) are needed to accurately capture a wide range of  $\alpha$  values and extrapolation/interpolation is appropriate. Perturbation is suitable when fluctuations in  $X$  are large compared to its variations with  $\alpha$  (a and c). Neither method is efficient for systems with many extrema and relatively small fluctuations (d). The adjustable parameter  $\alpha$  may be a thermodynamic variable (e.g., temperature or density) or model parameter (e.g., appearing in a potential energy function).

We emphasize that there is no practical reason for expanding Eq. 12 in this way as it can be directly applied to simulation data. However, Eq. 14 demonstrates that we are effectively taking the ratio of two series expansions, with both extending to infinite order. Actually taking this ratio should yield Eq. 9 substituted with Eqs. 10 and 11. Indeed, we may obtain the numerator of Eq. 14 by multiplying the denominator by the combined Eqs. 9–11 and collecting terms of the same order of  $\delta\beta$ . In practice, error is introduced into extrapolation predominantly through truncation of Eq. 9, while perturbation via Eq. 12 evaluates the series in Eq. 14 at infinite order but suffers from finite sampling errors stemming from a lack of sufficient phase-space overlap.

Finite sampling limits both perturbation theory and extrapolation, with the latter also limited by the order at which we truncate the expansion. As demonstrated later, however, extrapolation may significantly outperform perturbation for the special cases of observables that vary approximately linearly or quadratically as a function of the extrapolation variable (i.e., first- or second-order extrapolation, respectively) over large ranges of temperatures or densities. This is because perturbation does not utilize information regarding the local curvature in terms of the dependence of the observable as a func-

tion of the extrapolation variable (Fig. 1). Instead, perturbation relies on selecting configurations that have high weights in the new ensemble but were sampled in the reference ensemble, as is the case for large observable fluctuations shown in Fig. 1. If no such configurations (or phase-space overlaps) exist, perturbation theory cannot work, as can be seen from the ratio of weights appearing in the numerator and denominator of Eq. 13. Moving further from the reference ensemble tends to greatly reduce overlap, rendering perturbation estimates unreliable. Lack of overlap becomes even more pronounced for large systems due to reductions in relative fluctuations with system size. On the other hand, extrapolation at a given order is only accurate while the observable of interest varies in a similar fashion to the highest-order term of the polynomial expansion (Fig. 1). Similar to perturbation theory, the highest extrapolation order we can use, and hence the distance we can perturb, is limited by the amount of sampling that we have performed. This is because extrapolation requires converged estimates of the moments of an observable.

### III. ANALYTICAL EXAMPLE: IDEAL GAS IN AN EXTERNAL FIELD

To concretely compare perturbation and extrapolation and understand why each fails or succeeds, we examine the simple case of a one-dimensional ideal gas in an external potential depending linearly on the particle positions. This model is analytically tractable, but still exhibits non-trivial behavior in the temperature dependence of its structural properties. We treat a canonical ensemble with temperature  $T$ , number of particles  $N$ , and length of  $L$  in the single dimension with the coordinate of the  $i^{\text{th}}$  particle being  $x_i$ . The particles do not interact with each other, but do interact with an external field linear in the particle positions,  $\Phi(x) = ax$ . The total  $N$ -particle partition function  $Q$  is given by the product of single-particle partition functions  $q$  as  $Q = \frac{q^N}{N!}$  with

$$q = \frac{z}{\Lambda} \quad (15)$$

where  $\Lambda$  accounts for kinetic degrees of freedom and will not be of interest here, and  $z$  is the configurational partition function given by

$$\begin{aligned} z &= \int_0^L e^{-\beta ax} dx \\ &= \frac{1 - e^{-\beta aL}}{\beta a} \end{aligned} \quad (16)$$

The total potential energy for a system with  $N$  particles is simply  $U = a \sum_i x_i$ . For a single particle, the probability of a given  $x$  is

$$P(x) = \frac{e^{-\beta ax}}{z} \quad (17)$$

and the total probability for  $N$  particles with different positions is the product of the probability in Eq. 17 due to lack of

particle-particle interactions. For a single particle the potential energy distribution is also given by  $P(x)$  because a particle coordinate uniquely maps to a potential energy. For a sufficiently large number of particles, however, the distribution of potential energies becomes Gaussian and can be completely characterized by only its mean and variance. Throughout, we use 1000 particles, which satisfies this condition and is on an order of magnitude with the degrees of freedom that might be present in a molecular simulation. The average  $x$  position (for a single particle or over all particles) is given by

$$\langle x \rangle = \frac{1}{\beta a} - \frac{L}{e^{\beta a L} - 1} \quad (18)$$

The average potential energy is then given by  $\langle U \rangle = Na \langle x \rangle$  and, similarly, the variance of the potential energy is  $\sigma_U^2 = Na^2 \sigma_x^2$ , where

$$\sigma_x^2 = \frac{1}{\beta^2 a^2} - \frac{L^2 e^{\beta a L}}{(e^{\beta a L} - 1)^2} \quad (19)$$

We will treat the average position  $\langle x \rangle$  as our structural property of interest that we perturb and extrapolate in temperature space. It is important that  $\langle x \rangle$  is not a simple function of  $\beta$  and in fact exhibits significant non-linearity (Fig. 2). This makes the extrapolation non-trivial even though analytical results for extrapolation at each order (i.e., the exact  $n^{\text{th}}$  derivative with respect to  $\beta$ ) are trivial to compute. Without loss of generality, we henceforth set  $a$  and  $L$  equal to 1 for simplicity.

Fig. 2 shows finite sampling results for perturbation and extrapolation along with analytical results in the limit of infinite sampling. Finite sampling consists of generating  $N_s$  configurations by sampling from  $P(x)$  for each particle and calculating  $\langle x \rangle$  and  $U$  for each configuration. It is clear that finite sampling limits both perturbation and extrapolation, but in different ways. Perturbation works well close to the reference temperature but reaches a saturation value further away. The accuracy range is extended slightly with more samples, but the increases in accuracy (i.e., distance from the infinite sampling, infinite-order result) are marginal. The reason for this can be inferred from  $P(x)$  at various temperatures (Fig. S1). To obtain the correct average at one temperature, samples are required from all relevant configurations with high  $P(x)$  at that temperature. In other words, the distributions of potential energies at the two temperatures must overlap. Thus, sampling only a subset of these configurations amounts to implicit truncation of the average.<sup>35</sup> The result is a highly conservative (i.e., predicting no variation with temperature) estimate due to overly concentrated sampling of phase space relevant only to the reference temperature. This behavior is well-known in perturbation theory — most modern techniques involving free energy calculations and enhanced sampling seek to overcome exactly the difficulties described above.<sup>3,4</sup>

For the ideal gas system investigated in Fig. 2, both linear and quadratic extrapolation are superior to perturbation theory, even though extrapolation suffers from not only sampling error, but also truncation error that is not present in perturbation theory. Depending on the system, however, both errors may be much lower than those obtained from re-weighting

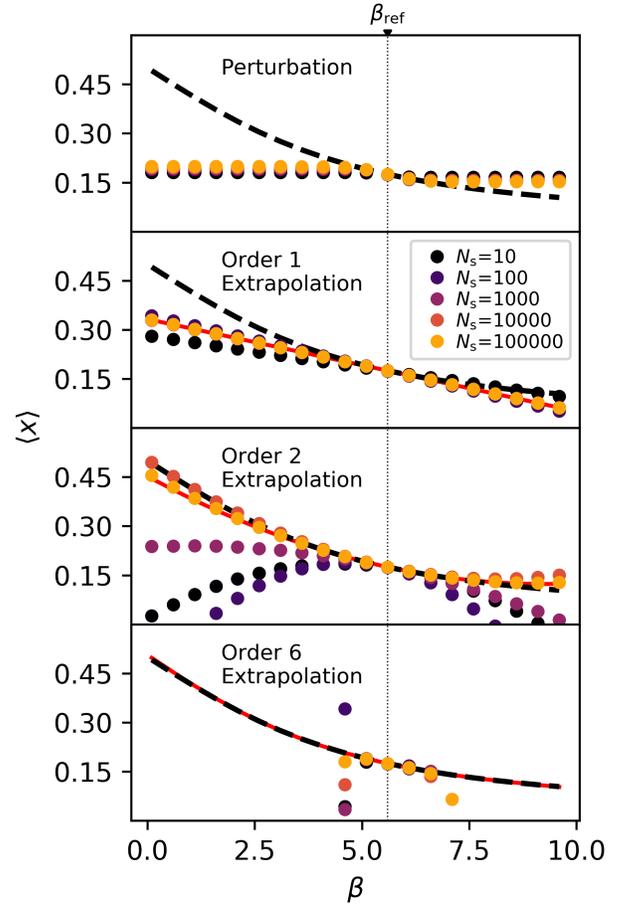


FIG. 2. Perturbation theory and extrapolation at various orders are compared for predicting the average position  $\langle x \rangle$  of an ideal gas particle in a linear external field. Analytical behavior at infinite order and sampling is shown as a black dashed line, the analytical result with infinite sampling for each order of extrapolation is shown as a red solid line, and perturbation or extrapolation results with varying numbers of randomly sampled configurations from the reference state of  $\beta = 5.6$  are represented by circles.

techniques. Fig. 2 also provides some sense of the behavior of both errors as a function of extrapolation order and sample size. While first-order extrapolation quickly converges to the infinite sampling limit, second order takes many more samples to attain similar accuracy. At higher orders, finite sampling errors result in inaccurate estimates of local derivatives, which in turn lead to highly inaccurate estimates far from the reference temperature. This is in contrast to the very low truncation error for sixth-order extrapolation with infinite sampling, which nearly exactly tracks the analytical result of Eq. 18.

There is clearly a trade-off between sampling and truncation errors when using extrapolation. As shown in Figs. 2 and 3a error and uncertainty (or imprecision) also grow with extrapolation distance  $\delta\beta$ . This is expected since increasing extrapolation distance magnifies errors in derivative estimates according to Eq. 9. The dependence of the uncertainty on

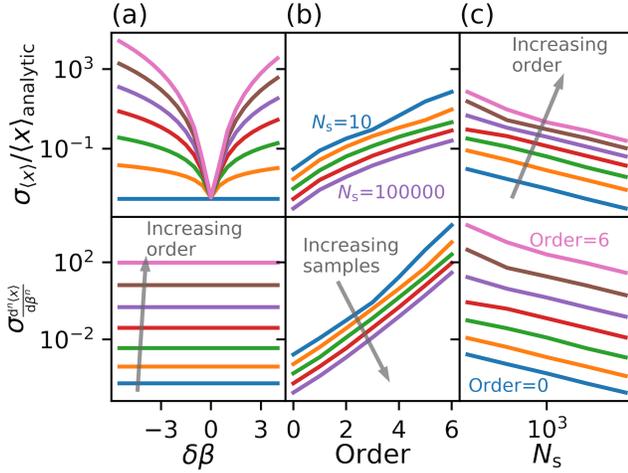


FIG. 3. Uncertainty (standard deviations over 1000 independent data draws at the reference condition) in the average position  $\langle x \rangle$  relative to the infinite sampling extrapolation result  $\langle x \rangle_{\text{analytic}}$  (top row) and in the derivative estimates (bottom) are shown as functions of extrapolation distance  $\delta\beta$  (a), extrapolation order (b), and number of samples  $N_s$  (c). In (b) and (c) the extrapolation distance is fixed at  $\delta\beta = 1.0$ . Uncertainty increases with greater distance from the reference point, higher order, and fewer numbers of samples. In the bottom panel of (a), uncertainties in derivatives do not grow with  $\delta\beta$  as these are locally estimated at the reference state.

sample size and order is summarized in Fig. 3b-c for a fixed temperature difference. Estimation of higher-order derivatives involves higher-order moments of the potential energy, leading to exponentially increasing uncertainty (Fig. 3b). At any order, the uncertainty in extrapolation estimates or derivative estimates vary as  $1/\sqrt{N_s}$  or  $1/N_s$ , respectively, with  $N_s$  being the number of samples (Fig. 3c). Relative unsigned errors from infinite-sampling results also decrease significantly with sampling and increase quickly as the order is increased (Fig. S2), which is also obvious from Fig. 2. In most practical settings, it is in fact the sampling error that will dominate, limiting the order and determining the truncation error. Even for the simple system investigated here, more than 10000 uncorrelated samples are required to achieve less than 10% uncertainty at third order. In molecular systems involving biomolecules, obtaining on the order of 100000 uncorrelated samples can require hundreds of nanoseconds or more of simulation time. As we demonstrate below, it is likely more efficient to combine the results of short simulations at multiple conditions, using only lower-order information.

#### IV. AN INTERPOLATION STRATEGY FOR OBSERVABLE ESTIMATION

Instead of extrapolating from a single point, we can simultaneously extrapolate from multiple points to improve our estimate. Simulations at additional state points are already necessary to check the accuracy of the extrapolation, even if the uncertainty may be estimated through bootstrapping. Previ-

ous work has suggested weighting each extrapolation estimate by the distance to the new state point,<sup>15</sup> which we refer to as “weighted extrapolation”, where subscripts 1 and 2 below indicate the two reference points

$$\langle X \rangle_\beta = \frac{\langle X \rangle_{\beta_1} |\beta - \beta_2|^m + \langle X \rangle_{\beta_2} |\beta - \beta_1|^m}{|\beta - \beta_1|^m + |\beta - \beta_2|^m} \quad (20)$$

This formulation affords a great deal of flexibility in that  $m$  may be optimized as a free parameter in order to satisfy criteria such as the Gibbs-Duhem relation.<sup>15</sup> For simplicity, however, we set  $m = 20$  throughout this work. This approach is most useful in the case of interpolation, when we desire the behavior of the observable over a range between two points at which data is available. From the perspective of free energy calculations, this is similar to techniques that utilize information from two states, such as Bennett’s Acceptance Ratio (BAR)<sup>36</sup>. Determining the behavior of an observable over an exactly or approximately known region between two states is in fact a common task. We may then use derivative information up to the maximum order possible to interpolate within the interval of interest, greatly reducing our uncertainty and increasing accuracy.

As an alternative to weighted extrapolation, we suggest directly fitting observable values and their derivatives, which are related to fluctuations (e.g., Eqs. 10 and 11), with an interpolating polynomial. Given two sets of observable and derivative values up to order  $k$ , it is possible to exactly fit a polynomial of order  $2k + 1$ . For example, with derivative information up to first order we obtain the third-order polynomial that best approximates the observable over the interval of interest. By design, this polynomial will exactly match the provided observables and their derivatives at the values of the interpolation variable (e.g., temperature or density) for which data was collected. Using two points up to arbitrary order of derivatives, this is obtained by solving for the polynomial coefficients  $c_i$  of order  $i$  in the set of Eqs.

$$\begin{bmatrix} 1 & \beta_1 & \beta_1^2 & \dots & \beta_1^{2k+1} \\ 0 & 1 & 2\beta_1 & \dots & (2k+1)\beta_1^{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \frac{(2k+1)!}{(k+1)!} \beta_1^{k+1} \\ 1 & \beta_2 & \beta_2^2 & \dots & \beta_2^{2k+1} \\ 0 & 1 & 2\beta_2 & \dots & (2k+1)\beta_2^{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \frac{(2k+1)!}{(k+1)!} \beta_2^{k+1} \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{2k+1} \end{bmatrix} = \begin{bmatrix} \langle X \rangle_{\beta_1} \\ \left. \frac{\partial \langle X \rangle}{\partial \beta} \right|_{\beta_1} \\ \vdots \\ \left. \frac{\partial^k \langle X \rangle}{\partial \beta^k} \right|_{\beta_1} \\ \langle X \rangle_{\beta_2} \\ \left. \frac{\partial \langle X \rangle}{\partial \beta} \right|_{\beta_2} \\ \vdots \\ \left. \frac{\partial^k \langle X \rangle}{\partial \beta^k} \right|_{\beta_2} \end{bmatrix} \quad (21)$$

It is also possible to incorporate more than two points, further increasing the polynomial order. However, this can lead to overfitting with large numbers of data points. As described below, we found a piecewise interpolation procedure to be more reliable.

We directly compare weighted extrapolation, polynomial interpolation, and re-weighting using BAR in Fig. 4a for the average  $x$  position in the ideal gas model. In the last technique, we use BAR to estimate the free energy difference, and hence weight of each state, for the lowest and highest temperatures. To re-weight observations from each state to the state of interest, weights from BAR are combined with observed Boltzmann weights. This is very similar to using standard perturbation theory to re-weight, the only difference being that we are also incorporating relative weights between the two states from which samples are drawn. Polynomial interpolation provides the smoothest curve, which is expected in comparison to the weighted extrapolation procedure of stitching together two first-order extrapolations rather than cleanly fitting a higher-order polynomial. Notably, polynomial interpolation also has the highest uncertainty predicted by bootstrap resampling. This indicates that the results for weighted extrapolation and BAR re-weighting exhibit less variance, but in this case also exhibit greater bias. Lower accuracy with BAR re-weighting is due to the same deficiencies as for standard perturbation theory extrapolations. Namely, the lack of phase-space overlap results in a highly conservative estimate heavily weighted towards the state with the lower free energy.

This exact malady has been known since the inception of BAR.<sup>36</sup> In his original paper, Bennett recommends against using BAR in cases where there is very little overlap of the distributions of potential energy differences. In such cases, Bennett instead suggests to fit the logarithm of each distribution to a polynomial and graphically determine the shift necessary to make the polynomials align, which is the free energy difference. Recent work has also used this same polynomial interpolation of potential energies to efficiently predict free energy differences between crystal polymorphs.<sup>38</sup> This is closely related to work that optimally combines cumulant expansions built from two states.<sup>32</sup> We may recover identical expressions in the current context by performing polynomial interpolation to determine the free energy as a function of the variable of interest. In this case, we assume that the first state has a free energy of zero and lose an order of the polynomial as we must also estimate the unknown free energy difference.

$$\begin{bmatrix} 1 & \beta_1 & \dots & \lambda_1^{2k} & 0 \\ 0 & 1 & \dots & (2k)\lambda_1^{2k-1} & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & \frac{(2k)!}{(k)!}\lambda_1^k & 0 \\ 1 & \beta_2 & \dots & \lambda_2^{2k} & -1 \\ 0 & 1 & \dots & (2k)\lambda_2^{2k-1} & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & \frac{(2k)!}{(k)!}\lambda_2^k & 0 \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{2k} \\ \Delta A \end{bmatrix} = \begin{bmatrix} 0 \\ \beta \frac{\partial H}{\partial \lambda} \Big|_{\lambda_1} \\ \vdots \\ \beta \frac{\partial^k H}{\partial \lambda^k} \Big|_{\lambda_1} \\ 0 \\ \beta \frac{\partial H}{\partial \lambda} \Big|_{\lambda_2} \\ \vdots \\ \frac{\partial^k H}{\partial \lambda^k} \Big|_{\lambda_2} \end{bmatrix} \quad (22)$$

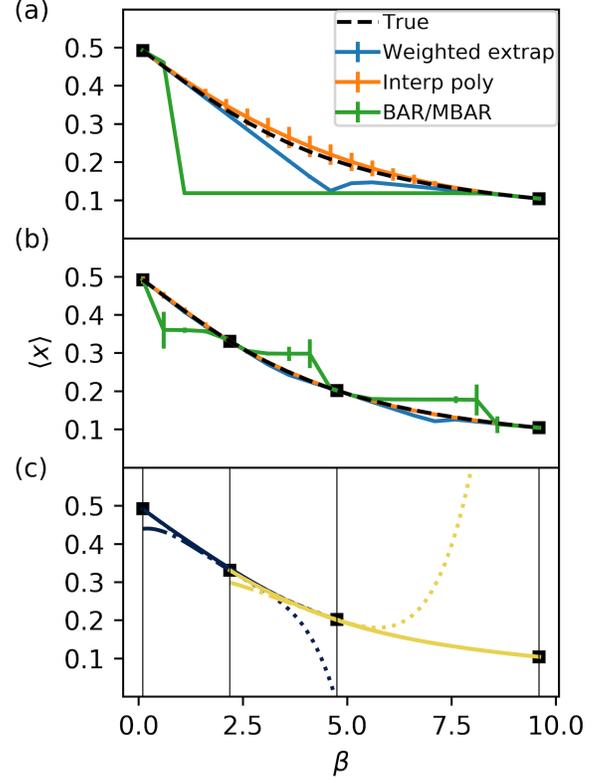


FIG. 4. Interpolation procedures are shown for weighted extrapolation and interpolating polynomials (both using only first-order derivatives), as well as BAR or MBAR depending on the number of data points (using the pymbar package<sup>37</sup>). Interpolation from two points is shown in (a), while interpolation using all of the points selected during a recursive interpolation run using polynomial interpolation and an error tolerance of 2% are shown in (b). Eq. 18 is shown as a black dashed line. The same 10000 samples at each temperature are used for each method, with error bars representing one standard deviation in 100 bootstrap resamples. (c) shows polynomial fits of three points selected during recursive interpolation, which are indicated by black squares and vertical lines. For each set, a polynomial in the same color is plotted based on each pair of points (solid lines indicate the outermost points of the set, dotted the left two points, and dash-dot the right two points). In many cases, polynomials diverge significantly outside the interval over which they were fit. In the shared interval, however, good overlap indicates that the local curvature of the polynomial fits is consistent, meaning additional points are not necessary.

With first-order derivatives, solving the above system yields

$$\beta \Delta A = \frac{1}{2} (\lambda_2 - \lambda_1) \left( \left\langle \beta \frac{\partial H}{\partial \lambda} \right\rangle_{\lambda_2} + \left\langle \beta \frac{\partial H}{\partial \lambda} \right\rangle_{\lambda_1} \right) \quad (23)$$

If we let  $\lambda_1 = 0$  and  $\lambda_2 = 1$  and  $H(\lambda) = H_0 + \lambda \Delta H$ , as is common in free energy calculation literature and equivalent to the choices in Ref. 32, we find that

$$\beta \Delta A = \frac{\beta}{2} \left( \langle \Delta H \rangle_{\lambda_2} + \langle \Delta H \rangle_{\lambda_1} \right) \quad (24)$$

This is identical to the result found by Hummer<sup>32</sup> if we note that non-equilibrium work in one direction is the negative of the work in the opposite direction. Going to fourth order in the derivatives (yielding an exact eighth-order polynomial and free energy difference) results in

$$\beta\Delta A = \frac{\beta}{2} (C_{\lambda_2}^{(1)} + C_{\lambda_1}^{(1)}) - \frac{3\beta}{28} (C_{\lambda_2}^{(2)} - C_{\lambda_1}^{(2)}) + \frac{\beta}{84} (C_{\lambda_2}^{(3)} + C_{\lambda_1}^{(3)}) - \frac{\beta}{1680} (C_{\lambda_2}^{(4)} - C_{\lambda_1}^{(4)}) \quad (25)$$

In the above,  $C^{(i)}$  represents the  $i^{\text{th}}$  cumulant of the free energy evaluated at a specific state indicated by  $\lambda$ . This is again identical to the cumulant expansion derived by Hummer, but with the explicit interpretation as an interpolating polynomial. If we also consider thermodynamic integration via the trapezoid rule, we are assuming that the derivative of the free energy as a function of the variable  $\lambda$  is linear. This is equivalent to assuming that the free energy itself is a second-order polynomial and using the estimated derivatives at the end points to simultaneously fit this polynomial and determine the free energy difference, as in Eq. 22. In the context of estimating free energy differences, we have thus shown that our polynomial interpolation strategy using up to first derivatives is equivalent to thermodynamic integration or optimal combination of cumulant expansions. We also obtain a closed-form polynomial expression of the free energy estimate over the entire interpolation range, whereas other techniques solely provide the free energy difference.

From the above considerations, it is not surprising that an interpolating polynomial for predicting structural observables outperforms re-weighting with BAR (Fig. 4a). Being equivalent to thermodynamic integration in the context of free energy calculations, polynomial interpolation only relies on accurately capturing the local curvature of the observable as a function of the interpolation variable, not on overlapping phase-space probability distributions as in BAR.<sup>39</sup> While statistically optimal re-weighting methods such as BAR or the Multistate Bennett Acceptance Ratio (MBAR)<sup>40</sup> can reliably provide highly precise free energy estimates, they are not always optimal compared to thermodynamic integration.<sup>39</sup> When curvature along the chosen thermodynamic path is high and there is significant phase-space overlap, as is typical in computing solvation free energies of small molecules, re-weighting techniques significantly outperform thermodynamic integration.<sup>41,42</sup> Such behavior is not observed in our present work, but is represented by panel c in Fig. 1. Our purpose here is not, then, to advocate for the general use of polynomial interpolation (or equivalently thermodynamic integration) to compute free energy differences or potentials of mean force. We make the connection only to explain differences in performance observed in Fig. 4 and Section V, where there is very little configurational overlap over large changes in temperature or density. This is analogous to the situation shown in Fig. 1(b), where extrapolation or interpolation is efficient while perturbation is not.

If we are not interested in computing the free energy but only an observable over a specific range of conditions, methods that rely on free energy calculations to perform re-

weighting will consistently provide estimates heavily biased towards the reference state ensemble. Re-weighting will only be preferred to interpolation if the fluctuations in the observable at a fixed state point rival its variations with changing simulation conditions (Fig. 1(a,c)). For extensive observables, this criterion will be met with vanishing frequency as system size increases. In any scenario, it is expected that if free energies are computed accurately, observables will be as well. However, a reduced computational cost is expected for computing observables alone, as performing the extensive sampling necessary for accurate free energy calculations is typically a computationally expensive endeavor. In such a case, direct polynomial interpolation of the observables further reduces the simulation cost by providing a prediction of the observable over the specified range. Interpolation relies on the local curvature of the observable (as a function of interpolation variable) being of equal or lower order to the estimated polynomial over the interval of interest. In the worst-case scenario shown schematically in Fig. 1(c,d), this is not true and the observable function is rugged. For small fluctuations in the observable (Fig. 1(d)), this will require collecting a similar number of data points as if a re-weighting procedure were implemented.

To address this issue, we require an algorithm to guide selection of additional data points, as well as a simple method to gauge convergence of the interpolation over the range of interest. Heuristics along these lines already exist in the free energy calculation literature, namely that efficiency and accuracy of re-weighting methods may be enhanced by selecting intermediate states that maximize phase-space overlap and enforce constant entropy differences at each step.<sup>35,43,44</sup> As mentioned previously, we have found that it is advantageous to fit interpolating polynomials in a piecewise fashion rather than use all data points to estimate coefficients for a higher-order polynomial. This is accomplished through a recursive algorithm. We start with the two most extreme points over the range of interest and collect data at each. Estimates of the observable of interest along with bootstrapped uncertainty estimates are obtained over the entire interval on a fine grid. The relative uncertainties given by  $|\frac{\sigma_Y}{Y}|$  are computed and compared to a specified error tolerance. If the tolerance is not met for all points on the interval, the point with maximum uncertainty is selected for new data collection. This algorithm recurses, dividing all high-uncertainty intervals into two parts, until the error tolerance is satisfied on all subintervals. Fig. 4b shows the result of executing this algorithm with a tolerance of 2% on the ideal gas model. A total of four points are necessary to achieve the desired accuracy.

This interpolation method crucially requires accurate estimates of local curvature on each interval. We can check this self-consistently with a simple visualization, similar to checking for histogram overlap when performing free energy calculations.<sup>45</sup> This visual check is shown in Fig. 4c. If we consider all sets of three adjacent points used to calculate polynomials, we can compare the local curvature of interpolations using the outer points of the interval to those using the points on the two sub-intervals. If the polynomials on the sub-intervals match closely the polynomial on the larger interval

then the curvature of that interval is well-approximated to the given polynomial order and no additional points are needed. In the ideal gas case, overlap occurs for only a small number of points when using third-order polynomials (Fig. 4c). More generally, local curvature of a third-order polynomial will capture scenarios where the sign of the derivative changes, covering a wide variety of scenarios. However, a third-order polynomial only requires observable values and first derivative information, which is exponentially easier to obtain to a given uncertainty compared to higher-order derivative information (Fig. 3). Additionally, the interpolations on different intervals agree in their value and first derivative at the selected data points. Thus the resulting piecewise function is continuous in both its value and first derivative by construction.

## V. PREDICTING STRUCTURAL PROPERTIES OF LIQUID WATER

To demonstrate our above theory and algorithms on a more realistic system, we examine the TIP4P/2005 model of water<sup>46</sup> in its liquid state. Simulations of 1400 water molecules spanning a wide range of temperatures (250-350 K) and densities (0.87-1.40 g/cm<sup>3</sup>) were performed with GROMACS 2016.1,<sup>47</sup> with additional details published previously.<sup>28</sup> The range of temperatures and densities was selected to span the structural anomaly envelop first described by Errington and Debenedetti,<sup>27</sup> which encompasses much of the liquid state including metastable regions (for instance, the state point discussed later at 0.87 g/cm<sup>3</sup> and 300 K is very near the liquid-vapor spinodal). Each simulation consists of snapshots saved every 1 ps over 5 ns trajectories, with simulations at 300 K and 1.00 g/cm<sup>3</sup> extended to 50 ns to examine the effect of increasing sampling by an order of magnitude. We consider both translational and orientational properties that classify water structure in addition to hard-sphere insertions (0.33 nm radius), which characterize local density fluctuations.<sup>48,49</sup> In all results below, error bars are the standard deviation as determined by bootstrap resampling (the calculation is repeated 100 times with random samples of the same size as the original data set drawn with replacement).

### A. Variation in temperature

Fig. 5a demonstrates that oxygen-oxygen RDFs at 1.00 g/cm<sup>3</sup> may be extrapolated from ambient conditions over nearly the entire liquid-state temperature range using only first-order derivative information. We accomplish this by individually extrapolating the count in each RDF bin before normalization.<sup>20</sup> In other words, the count in each bin of the RDF corresponds to the observable  $X$  in Eq. 9. The accuracy in Fig. 5 (and Fig. S10) is remarkable considering that a structural anomaly boundary is crossed over this temperature range (see Fig. 4 in Ref. 28). These results hold true even at the highest and lowest densities (Figs. S4-S7). Extrapolations remain quantitatively accurate even when RDFs undergo qualitative changes, such as developing shoulders at

lower temperatures. This implies that each set of histogram bin counts varies approximately linearly in temperature according to its own temperature derivative. Since we allow these derivatives to vary for each bin, we capture qualitative shifts in RDF behavior. Using second-order derivative information greatly increases uncertainty in the extrapolation, which remains the case even with an order of magnitude more simulation time (Fig. S3). Given the success of first-order extrapolation, any improvements associated with adding second derivative information are likely not worth the increased cost of sampling to the point of low uncertainty (on the order of 100 ns). Extrapolating RDFs in temperature also allows us to capture variation of the translational order parameter of Errington and Debenedetti<sup>27</sup> (Fig. 6a). Based on the integration of RDFs, this metric characterizes radial translational order and is defined as  $t = \frac{1}{\xi_c} \int_0^{\xi_c} |g(\xi) - 1| d\xi$  where  $\xi = r\rho^{1/3}$  with  $\xi_c = 2.843$ .<sup>27</sup> In contrast, re-weighting via perturbation theory only performs well for temperatures close to the reference simulation.

Similar success is found in temperature extrapolations of orientational properties, while perturbation encounters the same difficulties that were previously discussed. Three-body angles are defined as the angle between two water oxygens within a predefined cutoff (here 0.34 nm) of a central water oxygen. This structural feature has proven critical to developing monatomic models of water<sup>50</sup> and has also been shown to qualitatively track water's response to changes in state conditions or the presence of solutes.<sup>28</sup> Again, only 5 ns are necessary to successfully extrapolate entire three-body angle distributions over the entire temperature range (Fig. 5b). Qualitative shifts in the distribution are even more pronounced compared to RDFs, but are nonetheless predicted accurately with only first derivative information. Second derivative information based on these short trajectories only serves to significantly increase uncertainty, as was the case with translational measures of structure. Fig. 6b also demonstrates that extrapolation is successful in predicting the tetrahedral order parameter, which is defined as  $q = \left\langle 1 - \sum_{i,j} \left( \cos \theta_{i,j} + \frac{1}{3} \right)^2 \right\rangle$ .<sup>27,51</sup> The summation runs over all unique pairs of the four nearest neighbors of a central water oxygen with  $\theta_{i,j}$  representing the three-body angle of a pair. Averaging is performed over all water oxygens and system configurations.  $q$  is related to but can be calculated independently of three-body angle distributions.<sup>28,52</sup> In this case, it is more obvious why first order is appropriate — the order parameter only varies linearly in temperature. This is not only the case at ambient conditions, but also at lower and higher densities (Figs. S4-S7).

Water's density fluctuations as a function of temperature are more difficult to capture with extrapolation. Such fluctuations are critical to understanding solvation phenomena,<sup>53</sup> hydrophobic association,<sup>48,54</sup> and characterizing the nature of interfaces in contact with water.<sup>49,55</sup> Density fluctuations are commonly characterized by the excess chemical potential of a hard-sphere solute, which is given by the probability of having zero water oxygens within the hard-sphere volume

$$\beta\mu_{\text{ex}}^{\text{HS}} = -\ln P(N=0) \quad (26)$$

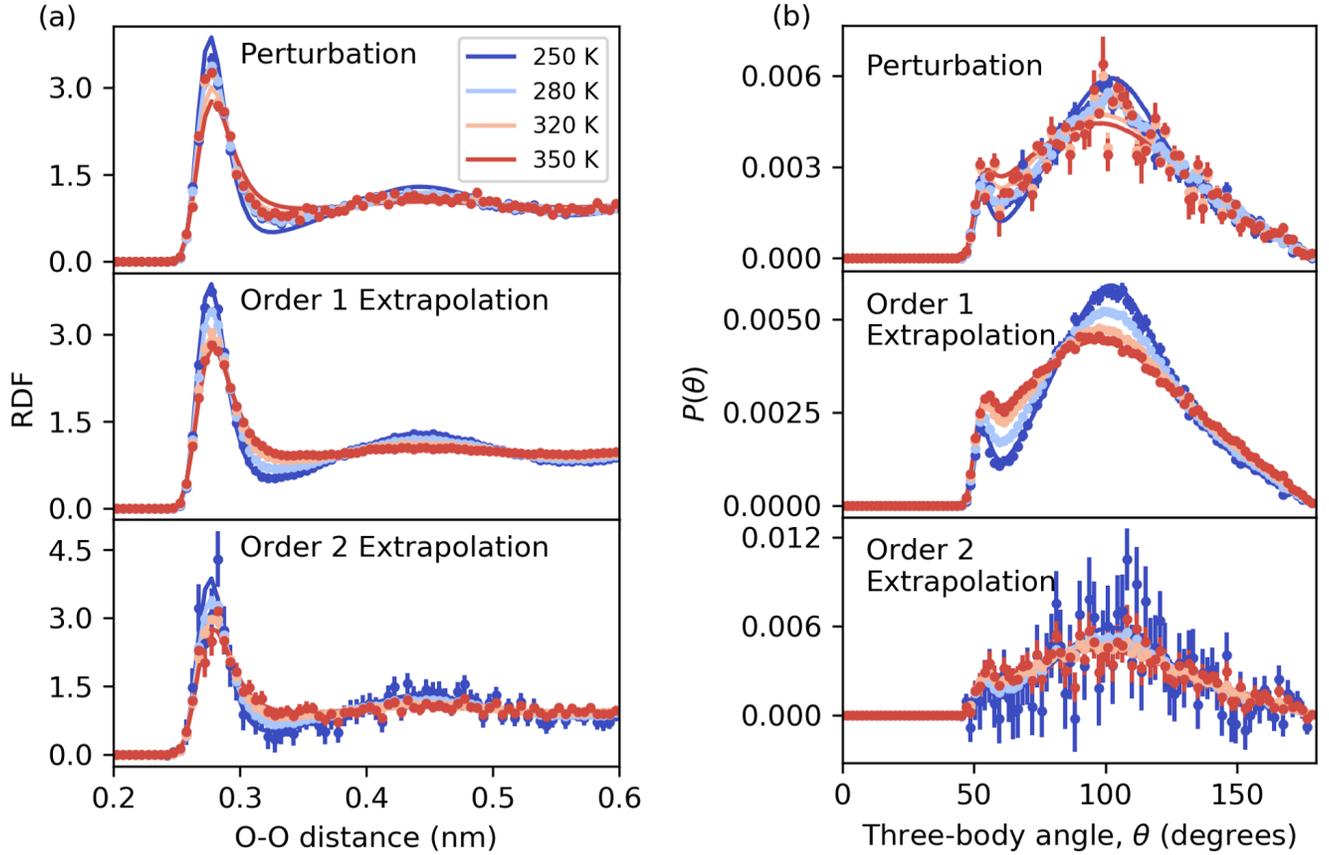


FIG. 5. In all panels, lines represent direct simulation results at each temperature for RDFs (a) or three-body angle distributions (b) while points are the estimates via extrapolation from 300 K (constant density of  $1.00 \text{ g/cm}^3$ ) with 5000 snapshots obtained at regular intervals over a 5 ns simulation. Error bars represent one standard deviation and are determined through bootstrap resampling. The neighbor cutoff was set to 0.34 nm in computing three-body angles.

As Eq. 26 depends on histogram counts of a quantity that only depends on sampled configurations, we categorize hard-sphere chemical potential as a structural observable akin to others discussed in this paper. Fig. 6c-d shows extrapolations in temperature from 300 K for hard-sphere excess chemical potentials, also revealing non-monotonic behavior with temperature (extrapolations of  $P(N=0)$  are also shown in Fig. S8). With non-monotonicity in  $P(N=0)$  and  $\beta\mu_{\text{ex}}^{\text{HS}}$ , it is no surprise that first-order extrapolation fails for some portions of the temperature range. It is interesting to note that, at this density, the temperature dependence of  $\beta\mu_{\text{ex}}^{\text{HS}}$  is weak, with the changes in  $\mu_{\text{ex}}^{\text{HS}}$  dominated by its linear dependence on  $\beta$  (Fig. 6c-d). This weak temperature dependence means that extremely precise estimates of the  $P(N=0)$  are needed to correctly capture the variation. As such, we utilize 50 ns at the reference temperature to perform extrapolation in Fig. 6c-d, as it is clear in Fig. S9 that the first derivative estimate with 5 ns of simulation time is highly inaccurate. In this case, the inherent noise in the method used to calculate the observable is similar to its variations with temperature. To reduce the overall simulation time in this scenario, faster converging methods of observable estimation should be employed. Other methods

to estimate  $\beta\mu_{\text{ex}}^{\text{HS}}$  could then be extrapolated in the same way but at higher certainty and accuracy.

Polynomial interpolations over temperature are presented in Fig. S11, converging to very low uncertainty (and high accuracy) using simulations at only the two most extreme temperatures. Since the translational and tetrahedral order parameters vary approximately linearly over the studied temperature range, extrapolation already performs well and interpolation displays only marginally better accuracy. In the case of hard-sphere chemical potentials, which vary non-monotonically with temperature, interpolation leads to starker improvement over extrapolation (Fig. S11c). Variations with density are non-monotonic for all observables and are explored in detail in the next section.

## B. Variation in volume

First-order derivatives in the canonical ensemble of arbitrary observables with respect to volume may be computed

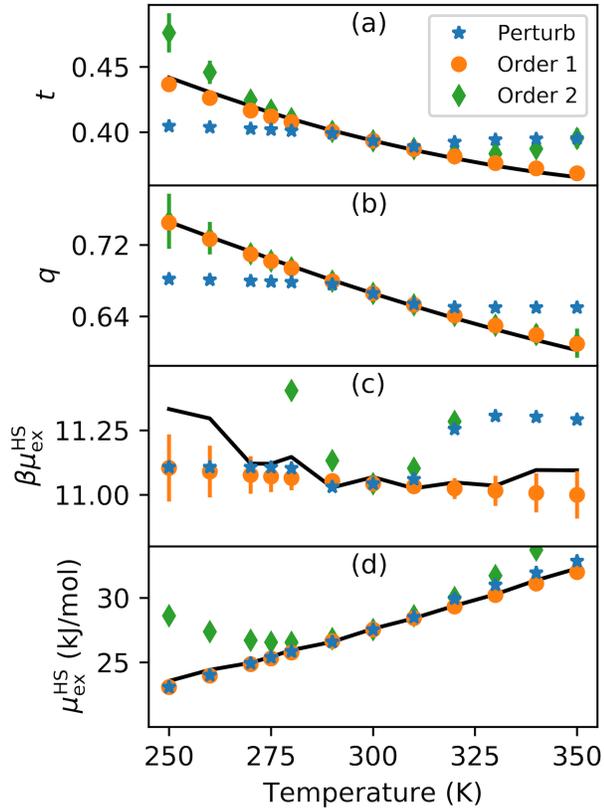


FIG. 6. Direct simulation results (solid lines), perturbation theory estimates (stars), and first- (circles) and second-order (diamonds) extrapolations are shown versus temperature for (a) translational order parameters, (b) tetrahedral order parameters, (c)/(d) excess chemical potentials of 0.33 nm radii hard-spheres. Translational order parameters are computed from RDFs extrapolated in temperature from 5 ns simulations at 300 K and  $1.00 \text{ g/cm}^3$ , while all other quantities are extrapolated directly from the same reference state. Due to the inherently high ratio of noise to temperature variation for excess chemical potential estimates at this density, 50 ns of simulation are used for extrapolations in (c) and (d). For clarity, error bars are not shown for perturbation or second-order extrapolation in (c) or (d), as they are on the same scale as the entire ordinate axis.

according to

$$\frac{d\langle X \rangle}{dV} = \frac{1}{3V} \left[ -\beta \langle X \rangle \langle W \rangle + \beta \langle XW \rangle + \left\langle \sum_i \frac{\partial X}{\partial x_i} x_i \right\rangle \right] \quad (27)$$

where  $X$  is the observable of interest,  $W$  is the virial, and  $x_i$  represents the  $i^{\text{th}}$  degree of freedom in the system. The last term appearing on the right hand side involves the response of the observable itself to changes in volume, and as such will be unique to each observable. A derivation of Eq. 27 along with expressions necessary to compute the last term for a variety of observables are provided in the Appendix. We do not pursue more than first-order derivative information in what follows, as higher-order derivatives in volume require derivatives of forces, which are not readily available in most MD simulation packages. Given the difference in observed uncertainties of

first and second-order extrapolations with respect to temperature, it is likely that high uncertainty would prevent the use of such extrapolations. In contrast, perturbation is not feasible at all due to negligible phase-space overlap between simulations at different densities.

On the whole, extrapolation in density is much less successful than in temperature, as shown in Figs. 7-8 and Fig. S12. Contributing to this is the fact that canonical ensemble simulations at different densities are spaced further apart in terms of their phase-space overlap, resulting in higher uncertainties in extrapolation estimates. Liquid structure is typically highly sensitive to density, with pressure changing by multiple orders of magnitude over the range studied here. Additionally, many properties of water naturally exhibit non-monotonic behavior as a function of density. With changing density, water structure changes significantly, crossing the structural anomaly boundary twice<sup>27,28</sup> and leading the translational and tetrahedral order parameters to behave non-monotonically. As such, first-order extrapolations should not in fact be expected to produce accurate results over large density differences. On the other hand, perturbation is highly impractical (and as such is not presented) as there is effectively no configurational overlap between even neighboring densities.

It is clear from Figs. 7-8 that successful extrapolation in density would require higher-order derivative information. However, relative uncertainties will grow exponentially with derivative order, as established in Section III. As a result, orders of magnitude more simulation time will be required to estimate even second-order derivatives. More modest increases in simulation time will further lower uncertainty in first-order extrapolations, but will not significantly improve accuracy. We turn instead to our recursive algorithm for polynomial interpolation, using the tetrahedral order parameter as a test case. This metric exhibits non-monotonic behavior in density, which is easy to visualize, as it is a scalar, and can be directly extrapolated or interpolated (unlike the translational order parameter). Fig. 8b shows that interpolation using two state points correctly captures the basic qualitative behavior of the tetrahedral order parameter. However, the local curvature is not accurate across the range of interest. This is largely due to the steep slope estimated at the lowest density point, which is in very close proximity to the liquid-vapor spinodal.<sup>28</sup> As a result, the slope of  $q$  changes rapidly across low-density, metastable state points (see Fig. S14), representing a mild case of what is shown schematically in Fig. 1d. Though our statistical analysis indicates low uncertainty in presented  $q$  values and derivatives, we caution that, as stability limits are approached, all properties become more difficult to determine with high statistical certainty due to increasing wavelengths and magnitudes of density fluctuations.<sup>56</sup>

Complicating matters further, the bootstrapped uncertainty is already within 1% over the entire region using just the two most extreme densities. In this case, it is necessary to check the consistency of the local curvature in a piecewise fashion by adding an additional point within the interval (specifically where the bootstrapped uncertainty is greatest). The resulting visual consistency check is shown in Fig. 9a. It is clear that the interpolation over the full interval (solid line) does

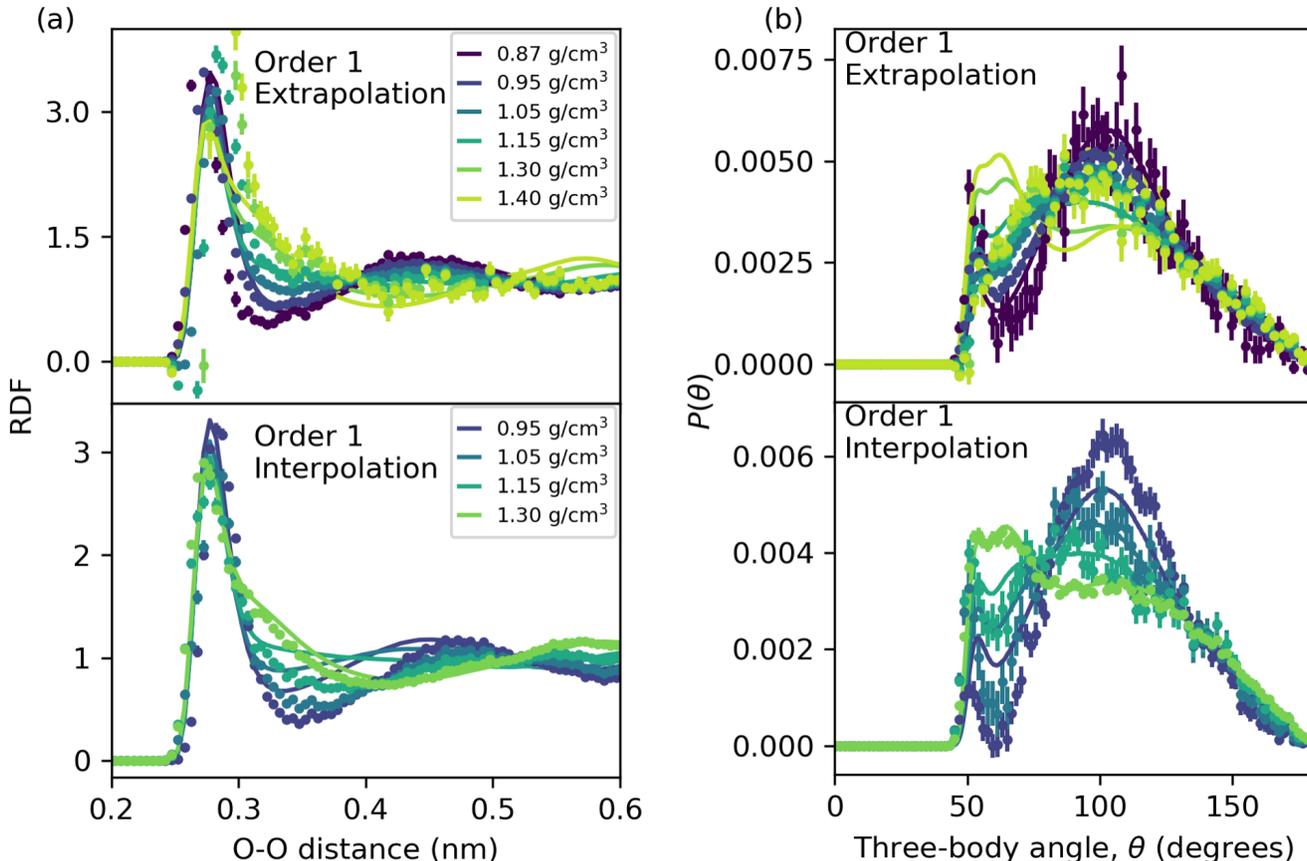


FIG. 7. Lines represent direct simulation results at each density for RDFs (a) or three-body angle distributions (b) while points are the estimates via extrapolation from a density of  $1.00 \text{ g/cm}^3$  (top) or interpolation from the two most extreme densities of  $0.87$  and  $1.40 \text{ g/cm}^3$  (bottom). Temperature is constant at  $300 \text{ K}$ . Error bars represent one standard deviation and are determined through bootstrap resampling.

not overlap very well with either of the interpolations over subintervals (dotted and dashed lines). This indicates a lack of consistency in the local curvature and suggests that at least one additional point should be added. The result of adding a single point on the high-density interval is shown in Fig. 9b, revealing that the local curvature is consistent on that interval, but not on the low-density interval where the curvature is in fact the largest. We can reduce the error tolerance for the recursive algorithm until the visual consistency check is passed, or (likely at lower computational cost) manually inspect each interval and add points where the local curvature is highest until self-consistent results are obtained. A complete implementation of this procedure is illustrated in Fig. S13, revealing that 50% or less of the original simulation data is required to achieve the same accuracy. We note, however, that this procedure is only necessary for quantitatively accurate results in the present case. With two 5 ns simulations providing 5000 snapshots each, we obtain correct qualitative behavior for the tetrahedral order parameter over the density range from  $0.87$  to  $1.40 \text{ g/cm}^3$ . Such information would be sufficient to determine the presence of an anomaly boundary and roughly approximate its location.

## VI. CONCLUSIONS

We have demonstrated, both for analytic and more realistic systems, that the use of thermodynamic extrapolation or interpolation performs no worse, and in many cases better, than re-weighting observables to other state conditions using standard perturbation theory. This result is not in fact surprising once connections between thermodynamic extrapolation and the extensive free energy calculation literature are made explicit. We find that the use of interpolating polynomials to calculate free energy differences and profiles along predefined coordinates is a special case of applying interpolating polynomials to general observables. This then illuminates why the same techniques outperform re-weighting for estimating structural observables in the case of very little overlap in phase space.<sup>36</sup> Successful interpolations of water structural properties in volume make this point particularly clear, as configurations satisfying other densities are exceptionally rare, rendering re-weighting impractical.

The connection with free energy calculations (or just rudimentary knowledge of numerical methods) also makes it clear that interpolation should be preferred over extrapolation. A particular range of interest for parameters or state variables is

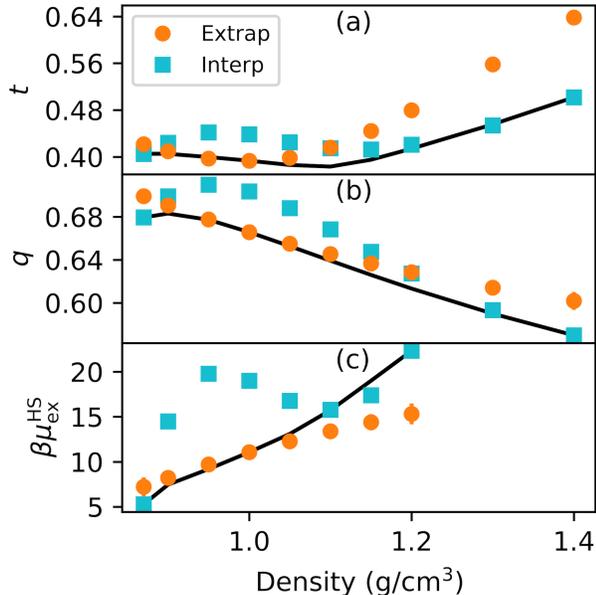


FIG. 8. Direct simulation results (solid lines), first-order extrapolations from 1.0 g/cm<sup>3</sup> (circles), and interpolations from the most extreme densities (squares) are shown versus density (at constant temperature of 300 K) for (a) translational order parameters, (b) tetrahedral order parameters, (c) excess chemical potentials of 0.33 nm radii hard-spheres. Translational order parameters are computed from RDFs extrapolated in density from 5 ns simulations at 300 K and 1.00 g/cm<sup>3</sup>, while all other quantities are extrapolated directly from the same reference state. Interpolations utilize two 5 ns simulations, one from each of the most extreme densities. Hard-sphere chemical potentials could not be computed at densities above 1.20 g/cm<sup>3</sup>.

not always known, but can usually be determined at little computational cost. Based on extensive uncertainty analysis of our analytic ideal gas system, we have also provided guidelines for allocation of computational resources: as higher-order derivatives are exponentially more difficult to converge to low uncertainty, it is recommended to use only low (first or second) order derivatives and run more short simulations at different conditions. This is corroborated by our studies of water, finding that only a few nanoseconds are required to converge first derivative information while hundreds of nanoseconds may be required for even second-order derivatives. Our subsequent results indicate that the behavior of a system can be determined straightforwardly over a continuous range of conditions by collecting low-order derivative information, which requires minimal additional computational cost. To facilitate the use of such methods, we have presented a recursive interpolation algorithm as well as a simple self-consistency check. These procedures are available as a Python library (available at <https://github.com/usnistgov/thermo-extrap>) that was used for both the ideal gas and water systems presented here. Such techniques may not only be used to efficiently generate new data, but also to extract additional information from previously generated trajectories, as we have done.

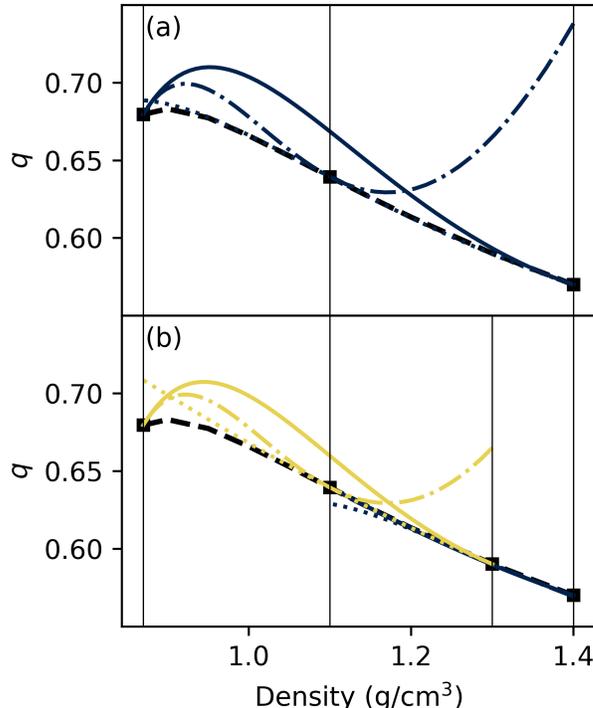


FIG. 9. Polynomial fits are shown on sliding sets of three points, which are indicated by solid black vertical lines. For each set, a polynomial in the same color is plotted based on each pair of points (solid lines indicate the outermost points of the set, dotted the two higher-density points, and dash-dot the two lower-density points). Piecewise interpolations using three (a) and four (b) data points are compared to demonstrate improvement of local curvature consistency with the number of points. Results from direct simulation at each density are shown as a black dashed line.

We expect that the presented theory and techniques will improve the efficiency of calculations based on molecular simulation. Derivative information for a very general set of observables is readily available, though typically underused. Our results suggest that the use of this information can greatly reduce the amount of required simulation, even when compared to other techniques for re-using simulation data such as standard perturbation theories. It has recently been pointed out that this derivative information also provides easy access to thermodynamic properties.<sup>29</sup> Indeed, entropies are immediately obtained by applying expressions presented here and elsewhere<sup>15–17</sup> to determine first temperature derivatives of potentials of mean force. This is possible because potentials of mean force represent differences in natural logarithms of probabilities and we have demonstrated in two cases how to extrapolate bin counts that form the basis of determining relative probabilities. It is not immediately clear that such information may be usefully interpreted in the case of RDFs or three-body angle distributions, but could, for example, quickly determine entropies associated with solvation free energies. Remarkably, this information can be directly extracted from the simulations used to determine solvation free ener-

gies without additional computational expense. To further enhance efficiency, it may be fruitful to consider classes of functions beyond polynomials when interpolating or extrapolating. For instance, Gaussian process regression<sup>57</sup> only assumes that functions are smooth on a specified length scale, positing the most likely functions satisfying both this criteria and the provided observable values and derivative information. Such improvements are currently under investigation.

## SUPPLEMENTARY MATERIAL

Additional supporting figures are provided in the Supplementary Material.

## ACKNOWLEDGMENTS

We thank Vikram Khanna for fruitful conversations and careful reading of the manuscript. This research was performed while J.I.M. held an NRC Research Associateship award at the National Institute of Standards and Technology.

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Appendix A: Volume extrapolation in the canonical ensemble

To compute the derivative  $\frac{d\langle X \rangle}{dV}$  at constant temperature and number of particles, we follow the derivations for  $\frac{dZ}{dV}$  found in McQuarrie<sup>58</sup> or Shell.<sup>59</sup> Both derivations rely on the scaling of the coordinates by the characteristic length of the system volume. We will here treat cubic volumes where  $V = L^3$  so that we can compute  $\frac{d\langle X \rangle}{dL}$  and in turn the desired derivative with respect to volume through the chain rule. From the above references and use of the chain rule, we have that

$$\frac{dZ}{dL} = \frac{3NZ}{L} - \beta Z \frac{\left\langle \sum_i \frac{\partial U}{\partial x_i} x_i \right\rangle}{L} \quad (\text{A1})$$

The summation runs over all degrees of freedom in a specific cartesian dimension  $x$  (i.e., atomic coordinates in this dimension). If the potential energy is pairwise, this may simplify the expression to only include pairwise distances in this dimension. As written, however, the above equation includes potentials where the system experiences an external field dependent on the atomic coordinates, or includes angle, dihedral or other terms to describe intramolecular interactions.

The derivative we want is

$$\frac{d\langle X \rangle}{dL} = \frac{d \int_0^L X \frac{e^{-\beta U}}{Z} dx_1 dy_1 dz_1 \dots dx_N dy_N dz_N}{dL} \quad (\text{A2})$$

We have only written the integration over translational coordinates in the simulation box. This is sufficient to fully describe monatomic particles, but for full molecules we should also integrate over rotations and intramolecular motions, not just center of mass translations. If we assume that such degrees of freedom do not depend on the size of the box, these terms do not affect the above derivative and simply contribute to the definition of the average. This treatment is exact for rigid molecules or for molecules that are small relative to the typical box size. For large flexible molecules, however, even intramolecular interactions may depend on the box size (following best practices in determining simulation box sizes for periodic systems this should in practice occur only rarely). In such cases we may instead simply assume that the integration in Eq. A2 does indeed run over all atomic positions and that the potential energy function includes bonds, angles, etc. in terms of these coordinates (in practice this is in fact the case). In real simulations, the potential energy will also depend on the box size through the application of minimum imaging in systems with periodic boundaries. This is not explicitly treated in the following derivation, but is assumed to be accurately captured if the virial (defined below) is properly computed.

First define a scaled coordinate as  $x'_i = \frac{x_i}{L}$  and transform the integrals in Eq. A2 so that we can take derivatives directly inside the integral.

$$\begin{aligned} \frac{d\langle X \rangle}{dL} &= \frac{d \int_0^1 X \frac{e^{-\beta U}}{Z} L^{3N} dx'_1 \dots dx'_N}{dL} \\ &= \frac{1}{Z^2} \int_0^1 Z e^{-\beta U} \left[ 3NL^{3N-1} X + L^{3N} \left( \frac{dX}{dL} - \beta X \frac{dU}{dL} \right) \right] \\ &\quad - L^{3N} X e^{-\beta U} \frac{dZ}{dL} dx'_1 \dots dx'_N \\ &= \int_0^1 \frac{e^{-\beta U}}{Z} \left[ 3NL^{3N-1} X + L^{3N} \left( \frac{dX}{dL} - \beta X \frac{dU}{dL} \right) \right] \\ &\quad - \frac{1}{Z} L^{3N} X \frac{dZ}{dL} dx'_1 \dots dx'_N \quad (\text{A3}) \end{aligned}$$

We can substitute in the derivative with respect to  $Z$  from Eq. A1, but must determine the derivatives with respect to  $U$  and  $X$ . If we have not transformed the coordinates inside these functions, we can write

$$\begin{aligned} \frac{dU}{dL} &= \sum_i \frac{\partial U}{\partial x_i} \frac{\partial x_i}{\partial L} \\ &= \sum_i \frac{\partial U}{\partial x_i} x'_i \\ &= \frac{1}{L} \sum_i \frac{\partial U}{\partial x_i} x_i \quad (\text{A4}) \end{aligned}$$

The same goes for  $X$ . Substituting in all of the derivatives

$$\frac{d\langle X \rangle}{dL} = \int_0^L \frac{e^{-\beta U}}{Z} \left[ 3NL^{3N-1}X + L^{3N} \left( \frac{1}{L} \sum_i \frac{\partial X}{\partial x_i} x_i - \beta X \frac{1}{L} \sum_i \frac{\partial U}{\partial x_i} x_i \right) - \frac{1}{Z} L^{3N} X \left( \frac{3NZ}{L} - \beta Z \frac{\langle \sum_i \frac{\partial U}{\partial x_i} x_i \rangle}{L} \right) \right] dx'_1 \dots dx'_N \quad (\text{A5})$$

We now transform the integral back to the original coordinates and clean everything up.

$$\begin{aligned} \frac{d\langle X \rangle}{dL} &= \int_0^L \frac{e^{-\beta U}}{ZL^{3N}} \left[ 3NL^{3N-1}X + L^{3N} \left( \frac{1}{L} \sum_i \frac{\partial X}{\partial x_i} x_i - \beta X \frac{1}{L} \sum_i \frac{\partial U}{\partial x_i} x_i \right) - \frac{1}{Z} L^{3N} X \left( \frac{3NZ}{L} - \beta Z \frac{\langle \sum_i \frac{\partial U}{\partial x_i} x_i \rangle}{L} \right) \right] dx_1 \dots dx_N \\ &= \int_0^L \frac{e^{-\beta U}}{ZL} \left[ 3NX + \left( \sum_i \frac{\partial X}{\partial x_i} x_i - \beta X \sum_i \frac{\partial U}{\partial x_i} x_i \right) - \frac{1}{Z} X \left( 3NZ - \beta Z \left\langle \sum_i \frac{\partial U}{\partial x_i} x_i \right\rangle \right) \right] dx_1 \dots dx_N \\ &= \int_0^L \frac{e^{-\beta U}}{ZL} \left[ 3NX + \left( \sum_i \frac{\partial X}{\partial x_i} x_i - \beta X \sum_i \frac{\partial U}{\partial x_i} x_i \right) - 3NX + \beta X \left\langle \sum_i \frac{\partial U}{\partial x_i} x_i \right\rangle \right] dx_1 \dots dx_N \\ &= \int_0^L \frac{e^{-\beta U}}{ZL} \left[ \sum_i \frac{\partial X}{\partial x_i} x_i - \beta X \sum_i \frac{\partial U}{\partial x_i} x_i + \beta X \left\langle \sum_i \frac{\partial U}{\partial x_i} x_i \right\rangle \right] dx_1 \dots dx_N \\ &= \frac{1}{L} \left[ \left\langle \sum_i \frac{\partial X}{\partial x_i} x_i \right\rangle - \beta \left\langle X \sum_i \frac{\partial U}{\partial x_i} x_i \right\rangle + \beta \langle X \rangle \left\langle \sum_i \frac{\partial U}{\partial x_i} x_i \right\rangle \right] \\ &= \frac{1}{L} \left[ \beta \langle X \rangle \left\langle \sum_i \frac{\partial U}{\partial x_i} x_i \right\rangle - \beta \left\langle X \sum_i \frac{\partial U}{\partial x_i} x_i \right\rangle + \left\langle \sum_i \frac{\partial X}{\partial x_i} x_i \right\rangle \right] \end{aligned} \quad (\text{A6})$$

Using the chain rule we can write

$$\begin{aligned} \frac{d\langle X \rangle}{dV} &= \frac{d\langle X \rangle}{dL} \left( \frac{dV}{dL} \right)^{-1} \\ &= \frac{1}{3L^2} \frac{d\langle X \rangle}{dL} \\ &= \frac{1}{3V} \left[ \beta \langle X \rangle \left\langle \sum_i \frac{\partial U}{\partial x_i} x_i \right\rangle - \beta \left\langle X \sum_i \frac{\partial U}{\partial x_i} x_i \right\rangle \right. \\ &\quad \left. + \left\langle \sum_i \frac{\partial X}{\partial x_i} x_i \right\rangle \right] \end{aligned} \quad (\text{A7})$$

It is typical to define the virial as  $W = -\sum_i \frac{\partial U}{\partial x_i} x_i$ . Using this definition we obtain Eq. 27 in the main text

$$\frac{d\langle X \rangle}{dV} = \frac{1}{3V} \left[ -\beta \langle X \rangle \langle W \rangle + \beta \langle XW \rangle + \left\langle \sum_i \frac{\partial X}{\partial x_i} x_i \right\rangle \right] \quad (\text{A8})$$

We can obtain the virial  $W$  from most simulation software packages, with rigid-body constraints and periodic boundary conditions with minimum imaging properly accounted for. Note that the above derivation only allows extrapolation up to first order. Continuing to second order will result in second derivatives of the potential energy with respect to the box length, which in turn results in second derivatives of the potential energy with respect to the coordinates. This means that we would need to keep track of the derivatives of the forces, which is not typically trivial in most simulation packages. With the value of the observable and the virial for each simulation configurational snapshot, we can easily compute the first two terms in Eq. 27. A remaining difficulty, however,

is the last term involving derivatives of the observable with respect to coordinates. There is no general expression for this quantity as it will depend on the definition of the observable of interest. Derivations for approximations to this term are provided below for radial distribution functions, three-body angle distributions, and hard-sphere insertions. Though this term is typically orders of magnitude smaller than the other terms in Eq. 27, it may still become significant when the difference between the other two terms is small. While we have found the last term in Eq. 27 to be small for RDFs and three-body angle distributions, it makes a non-negligible contribution for hard-sphere insertions. For the case of the tetrahedral order parameter, we do not present a derivation as we expect no contribution — interparticle distances do not affect the calculation of the three-body angle and are only used to determine the four nearest neighbors. The selected neighbors will not change with box scaling as all particle separations will be scaled by the same amount.

## 1. Radial distribution functions

Extrapolating an RDF entails extrapolating each set of bin counts of a certain interparticle radial distance  $r$ . The average fraction of particle pairs with a separation in bin  $k$  of fixed bin width  $d$  is the average number of bin counts divided by the

number of particle pairs and is given by

$$\begin{aligned}
 c_k &= \frac{2}{N(N-1)} \int \frac{e^{-\beta U}}{Z} \sum_i \Theta \left( r_i - \left( r_k - \frac{d}{2} \right) \right) \\
 &\quad \times \left( 1 - \Theta \left( r_i - \left( r_k + \frac{d}{2} \right) \right) \right) d\mathbf{r}^N \\
 &= \frac{2}{N(N-1)} \sum_i \int \frac{e^{-\beta U}}{Z} \Theta \left( r_i - \left( r_k - \frac{d}{2} \right) \right) \\
 &\quad \times \left( 1 - \Theta \left( r_i - \left( r_k + \frac{d}{2} \right) \right) \right) d\mathbf{r}^N \\
 &= \int \frac{e^{-\beta U}}{Z} \Theta \left( r_1 - \left( r_k - \frac{d}{2} \right) \right) \\
 &\quad \times \left( 1 - \Theta \left( r_1 - \left( r_k + \frac{d}{2} \right) \right) \right) d\mathbf{r}^N \quad (\text{A9})
 \end{aligned}$$

The summation runs over all pairs of particles with the  $i^{\text{th}}$  pairwise particle distance given by  $r_i$ . The product of Heaviside functions  $\Theta(x)$  ensures that the configuration is only counted if the interparticle separation is within bin  $k$ . In the last line, we have utilized the fact that all particle pairs are identical. It will be convenient to rewrite Eq. A9 explicitly in terms of cartesian coordinates and transform these coordinates such that  $\mathbf{r}'_i = \mathbf{r}_i - \mathbf{r}_1$ . For a homogeneous system, setting  $\mathbf{r}_1$  as the origin does not change the average since the potential energy only depends on relative particle positions. With this new coordinate system and selecting the first pair as particles 1 and 2, we may write

$$\begin{aligned}
 c_k &= \int \frac{e^{-\beta U}}{Z} \Theta \left( |\mathbf{r}'_2| - \left( r_k - \frac{d}{2} \right) \right) \\
 &\quad \times \left( 1 - \Theta \left( |\mathbf{r}'_2| - \left( r_k + \frac{d}{2} \right) \right) \right) d\mathbf{r}'^N \quad (\text{A10})
 \end{aligned}$$

where the distance between particles 1 and 2 is, in cartesian coordinates,  $|\mathbf{r}'_2| = (x_2'^2 + y_2'^2 + z_2'^2)^{1/2}$ . It is then clear that the simulation observable we are averaging is  $X = \Theta \left( |\mathbf{r}'_2| - \left( r_k - \frac{d}{2} \right) \right) \left( 1 - \Theta \left( |\mathbf{r}'_2| - \left( r_k + \frac{d}{2} \right) \right) \right)$ . All of the derivatives in last term of Eq. 27 will be zero except those with respect to the transformed coordinates of particle 2. The derivative with respect to a single cartesian coordinate of particle 2 is

$$\begin{aligned}
 \frac{\partial X}{\partial x_2'} &= \frac{\partial X}{\partial |\mathbf{r}'_2|} \frac{\partial |\mathbf{r}'_2|}{\partial x_2'} \\
 &= \frac{\partial X}{\partial |\mathbf{r}'_2|} \frac{x_2'}{|\mathbf{r}'_2|} \\
 &= \left[ \delta \left( |\mathbf{r}'_2| - \left( r_k - \frac{d}{2} \right) \right) \left( 1 - \Theta \left( |\mathbf{r}'_2| - \left( r_k + \frac{d}{2} \right) \right) \right) \right. \\
 &\quad \left. - \Theta \left( |\mathbf{r}'_2| - \left( r_k - \frac{d}{2} \right) \right) \delta \left( |\mathbf{r}'_2| - \left( r_k + \frac{d}{2} \right) \right) \right] \frac{x_2'}{|\mathbf{r}'_2|} \\
 &= \left[ \delta \left( |\mathbf{r}'_2| - \left( r_k - \frac{d}{2} \right) \right) \right. \\
 &\quad \left. - \delta \left( |\mathbf{r}'_2| - \left( r_k + \frac{d}{2} \right) \right) \right] \frac{x_2'}{|\mathbf{r}'_2|} \quad (\text{A11})
 \end{aligned}$$

Due to most derivatives being zero, the sum inside the expectation in the last term of Eq. 27 becomes

$$\begin{aligned}
 \sum_i \frac{\partial X}{\partial x_i} x_i &= \frac{\partial X}{\partial x_2'} x_2' + \frac{\partial X}{\partial y_2'} y_2' + \frac{\partial X}{\partial z_2'} z_2' \\
 &= \left[ \delta \left( |\mathbf{r}'_2| - \left( r_k - \frac{d}{2} \right) \right) \right. \\
 &\quad \left. - \delta \left( |\mathbf{r}'_2| - \left( r_k + \frac{d}{2} \right) \right) \right] |\mathbf{r}'_2| \quad (\text{A12})
 \end{aligned}$$

Taking the expectation over all system configurations yields

$$\begin{aligned}
 \left\langle \sum_i \frac{\partial X}{\partial x_i} x_i \right\rangle &= \int \frac{e^{-\beta U}}{Z} \left[ \delta \left( |\mathbf{r}'_2| - \left( r_k - \frac{d}{2} \right) \right) \right. \\
 &\quad \left. - \delta \left( |\mathbf{r}'_2| - \left( r_k + \frac{d}{2} \right) \right) \right] |\mathbf{r}'_2| d\mathbf{r}'^N \\
 &= \int P(\mathbf{r}'_2) \left[ \delta \left( |\mathbf{r}'_2| - \left( r_k - \frac{d}{2} \right) \right) \right. \\
 &\quad \left. - \delta \left( |\mathbf{r}'_2| - \left( r_k + \frac{d}{2} \right) \right) \right] |\mathbf{r}'_2| d\mathbf{r}'_2 \quad (\text{A13})
 \end{aligned}$$

By performing the integration over all coordinates except  $\mathbf{r}'_2$  we obtain a marginal probability  $P(\mathbf{r}'_2)$ . In order to integrate the above, we first transform from cartesian to spherical coordinates.

$$\begin{aligned}
 \left\langle \sum_i \frac{\partial X}{\partial x_i} x_i \right\rangle &= \int \frac{P(\mathbf{r}'_2)}{\sin(\theta_2) r_2^2} \\
 &\quad \times \left[ \delta \left( r_2 - \left( r_k - \frac{d}{2} \right) \right) - \delta \left( r_2 - \left( r_k + \frac{d}{2} \right) \right) \right] \\
 &\quad \times r_2 \sin(\theta_2) r_2^2 dr_2 d\theta_2 d\phi_2 \\
 &= \int P(r_2) \left[ \delta \left( r_2 - \left( r_k - \frac{d}{2} \right) \right) \right. \\
 &\quad \left. - \delta \left( r_2 - \left( r_k + \frac{d}{2} \right) \right) \right] r_2 dr_2 \\
 &= P \left( r_k - \frac{d}{2} \right) \left( r_k - \frac{d}{2} \right) \\
 &\quad - P \left( r_k + \frac{d}{2} \right) \left( r_k + \frac{d}{2} \right) \quad (\text{A14})
 \end{aligned}$$

By transforming to spherical coordinates, the probability density is divided by a term that cancels the Jacobian determinant. To arrive at the second line we have integrated over the angular coordinates to obtain the marginal probability of only the radial coordinate.  $P(r)$  represents the marginal probability density of observing any particle pair with an interparticle separation of  $r$ . If the bin width  $d$  is sufficiently small, then to a good approximation  $P(r = r_k - \frac{d}{2}) \approx P(r = r_k + \frac{d}{2})$ . Furthermore, these are also approximately equal to  $c_k$  divided by the width of the bin

$$\begin{aligned}
 \left\langle \sum_i \frac{\partial X}{\partial x_i} x_i \right\rangle &= \frac{c_k}{d} \left( r_k - \frac{d}{2} - r_k - \frac{d}{2} \right) \\
 &= -c_k \quad (\text{A15})
 \end{aligned}$$

In practice, we obtain  $c_k$  for each frame by counting the number of pairs of particles with a separation in bin  $k$  and dividing by the total number of unique pairs.

Strictly speaking, the definition of the radial distribution function involves normalizing by a term depending on the volume in order to account for the fractional volume of the bins.

$$g(r_k) = \frac{c_k}{(v_k/V)} \quad (\text{A16})$$

The volume of bin  $k$  is given by  $v_k = \frac{4}{3}\pi(r_k^3 - r_{k-1}^3)$  and  $V$  is the total volume. Through the chain rule, the appropriate full derivative for the radial distribution function becomes

$$\frac{dg(r_k)}{dV} = \frac{c_k}{v_k} + \frac{V}{v_k} \frac{dc_k}{dV} \quad (\text{A17})$$

This is only the first-order derivative, but will suffice for the present work.

## 2. Three-body angle distributions

While an RDF classifies the translational order in water, the angular order may be explored by using the distribution of three-body angles.<sup>28</sup> Given a central water oxygen, the three-body angle is that between vectors connecting the central water to two neighbors within its first hydration shell. The hydration shell is simply defined through a cutoff distance  $r_c$ , with the resulting probability distribution being largely insensitive to this cutoff as long as it does not extend significantly beyond the first minimum in the RDF.<sup>28</sup> As with RDFs, we may consider all triples of particles indistinguishable and average over only one of them. As such, the observable that we will average is

$$X = (1 - \Theta(r_{1,2} - r_c))(1 - \Theta(r_{1,3} - r_c)) \times \delta \left( \cos^{-1} \left[ \frac{\mathbf{r}_2 - \mathbf{r}_1}{r_{1,2}} \cdot \frac{\mathbf{r}_3 - \mathbf{r}_1}{r_{1,3}} \right] - \theta \right) \quad (\text{A18})$$

Particle 1 has been chosen as the central particle with distances to the two other particles defined by  $r_{1,j}$ . When averaged over all configurations, this counts the number of times that a triple is observed with two particles within the cutoff and forming an angle  $\theta$ . Technically, the angle  $\theta$  is binned and so the delta function should be the product of two Heaviside functions. However, the term within the inverse cosine does not depend on the box size due to the normalizations of the vectors between particles and so treatment of this term will not impact the analysis below. This can be seen by imagining that we reference all particles to the coordinates of the central particle and transform to spherical coordinates. The term in the delta function will then no longer depend on the radial coordinates but only the angles. Angular coordinates will have no dependence on the system size and thus not enter into the sum over degrees of freedom in the last term of Eq. 27. By performing the coordinate transform mentioned above, we arrive at the derivative with respect to the radial coordinate of a

non-central particle  $j$

$$\frac{\partial X}{\partial r_j} = -\delta(r_j - r_c)(1 - \Theta(r_k - r_c)) \times \delta \left( \cos^{-1} \left[ \frac{\mathbf{r}'_j \cdot \mathbf{r}'_k}{r_j r_k} \right] - \theta \right) \quad (\text{A19})$$

where  $k$  represents the other non-central particle forming the angle. Multiplying by the interparticle separation followed by summing and averaging yields

$$\begin{aligned} \left\langle \sum_i \frac{\partial X}{\partial r_i} r_i \right\rangle &= -2 \int \frac{e^{-\beta U}}{Z} r_2 \delta(r_2 - r_c) \\ &\quad \times (1 - \Theta(r_3 - r_c)) \\ &\quad \times \delta \left( \cos^{-1} \left[ \frac{\mathbf{r}'_2 \cdot \mathbf{r}'_3}{r_2 r_3} \right] - \theta \right) \mathbf{d}\mathbf{r}'^N \\ &= -2 \int P(\mathbf{r}'_2, \mathbf{r}'_3) r_2 \delta(r_2 - r_c) \\ &\quad \times (1 - \Theta(r_3 - r_c)) \\ &\quad \times \delta \left( \cos^{-1} \left[ \frac{\mathbf{r}'_2 \cdot \mathbf{r}'_3}{r_2 r_3} \right] - \theta \right) \mathbf{d}\mathbf{r}'_2 \mathbf{d}\mathbf{r}'_3 \\ &= -2P(\theta, r_2 = r_c, r_3 \leq r_c) r_c \end{aligned} \quad (\text{A20})$$

In the first line, we have asserted that terms from particles 2 and 3 will be identical, resulting in factor of 2. Integrating over all degrees of freedom except those of particles 2 and 3, we arrive at the second line where  $P(\mathbf{r}'_2, \mathbf{r}'_3)$  is a marginal probability density. The final line indicates that further integration results in the probability that the angle  $\theta$  is formed with particle 3 within the cutoff and particle 2 exactly at a radial distance of  $r_c$  from the central particle. Configurations contributing to this probability are a subset of those contributing to the full probability of an angle with both non-central particles at or below the cutoff. As such, we may write

$$P(\theta, r_2 = r_c, r_3 \leq r_c) = P(\theta, r_2 \leq r_c, r_3 \leq r_c) \times P(r_2 = r_c | \theta, r_2 \leq r_c, r_3 \leq r_c) \quad (\text{A21})$$

The first probability on the right-hand side is equal to  $\langle X \rangle$  which is proportional to the number of counts. The second is the probability of  $r_2$  being at a distance of  $r_c$  given that all other conditions are met. In other words, this represents the fraction of the number of counts of an angle  $\theta$  where one non-central particle is exactly at the cutoff. If we approximate  $P(r_2 = r_c | \theta, r_2 \leq r_c, r_3 \leq r_c)$  to be uniform along  $r_2$ , then

$$\begin{aligned} P(r_2 = r_c | \theta, r_2 \leq r_c, r_3 \leq r_c) &\approx \frac{4\pi r_c^2}{\frac{4}{3}\pi r_c^3} \\ &= \frac{3}{r_c} \end{aligned} \quad (\text{A22})$$

Combining this with Eqs. A20 and A21 we obtain

$$\left\langle \sum_i \frac{\partial X}{\partial r_i} r_i \right\rangle = -6 \langle X \rangle \quad (\text{A23})$$

In practice,  $\langle X \rangle$  is the average number of counts over all configurations where a triple of particles exhibits an angle  $\theta$  and

the non-central molecules are within the cutoff. Presented three-body angle distributions represent the probability of an angle given that the distance cutoff is satisfied, which involves a normalization by the total number of bin counts. Note that the three-body angle distribution with discrete bins is a probability mass function with the sum of the bins equalling one. While we directly extrapolate bin counts, we always normalize the result so that multiple such distributions may be compared on the same scale.

### 3. Hard-sphere insertions

For hard-spheres, the potential energy change upon adding the particle is infinite if any molecules overlap and zero otherwise, leading to a simplification in the expression for the excess chemical potential

$$e^{-\beta\mu_{\text{ex}}^{\text{HS}}} = \langle e^{-\beta\Delta U} \rangle = P(N=0) \quad (\text{A24})$$

Hence the average quantity of interest is the probability that no atoms lie within the defined volume of a hard-sphere particle placed randomly within the system. In the case of most water models, we only check water oxygens as the hydrogens do not participate in LJ or other repulsive core interactions. We can write  $P(N=0)$  for a hard-sphere centered at the origin with a distance of closest-approach to water oxygens of  $R$  as

$$P(N=0) = \int \frac{e^{-\beta U}}{Z} \prod_i \Theta(r_i - R) d\mathbf{r}^N \quad (\text{A25})$$

$r_i$  is the radial coordinate (or distance from the origin) of particle  $i$  with the product running over all particles in the system (oxygens only for water). The product takes the value of 1 only if no particles are within  $R$  of the hard-sphere particle centered at the origin. In this way, the total probability of having no particles in the hard-sphere volume is totaled by integrating over all configurations weighted by their configurational weights. From this it is clear that, in our usual notation,  $\langle X \rangle = P(N=0)$  and  $X = \prod_i \Theta(r_i - R)$ . We have that

$$\frac{\partial X}{\partial r_i} = \delta(r_i - R) \prod_{j \neq i} \Theta(r_j - R) \quad (\text{A26})$$

Under an integral, the above ensures that no particles  $j \neq i$  are within the hard-sphere volume and that particle  $i$  is exactly at  $R$ .

$$\begin{aligned} \left\langle \sum_i \frac{\partial X}{\partial r_i} r_i \right\rangle &= \int \frac{e^{-\beta U}}{Z} \sum_i \delta(r_i - R) r_i \prod_{j \neq i} \Theta(r_j - R) d\mathbf{r}^N \\ &= NP(r_1 = R, r_j \geq R \forall j > 1)R \\ &= NP(N=0, r_1 = R)R \\ &= NP(r_1 = R | N=0)P(N=0)R \\ &\approx NP(r_1 = R)P(N=0)R \\ &= \frac{4\pi R^3}{V} NP(N=0) \end{aligned} \quad (\text{A27})$$

The approximation on the fifth line assumes that the probability of a particle residing at  $R$  is negligibly affected by the fact that no other particles are inside the hard-sphere volume, which for a homogeneous system with uniform particle density leads to the expression on the last line. This is certainly not true if no particles are inside  $R$ , in which case the probability of finding a particle at  $R$  is increased and is termed the ‘‘contact value.’’<sup>53</sup> More accurate expressions for this quantity may be found in the context of scaled particle theories,<sup>53</sup> but are not implemented here.

- <sup>1</sup>J. C. Palmer and P. G. Debenedetti, ‘‘Recent advances in molecular simulation: A chemical engineering perspective,’’ *AIChE Journal* **61**, 370–383 (2015).
- <sup>2</sup>S. A. Hollingsworth and R. O. Dror, ‘‘Molecular Dynamics Simulation for All,’’ *Neuron* **99**, 1129–1143 (2018).
- <sup>3</sup>V. Spiwok, Z. Sucur, and P. Hosek, ‘‘Enhanced sampling techniques in biomolecular simulations,’’ *Biotechnology Advances* **33**, 1130–1140 (2015).
- <sup>4</sup>Y. I. Yang, Q. Shao, J. Zhang, L. Yang, and Y. Q. Gao, ‘‘Enhanced sampling in molecular dynamics,’’ *The Journal of Chemical Physics* **151**, 070902 (2019).
- <sup>5</sup>W. M. Brown, P. Wang, S. J. Plimpton, and A. N. Tharrington, ‘‘Implementing molecular dynamics on hybrid high performance computers – short range forces,’’ *Computer Physics Communications* **182**, 898–911 (2011).
- <sup>6</sup>W. M. Brown, A. Kohlmeier, S. J. Plimpton, and A. N. Tharrington, ‘‘Implementing molecular dynamics on hybrid high performance computers – Particle–particle particle–mesh,’’ *Computer Physics Communications* **183**, 449–459 (2012).
- <sup>7</sup>P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L.-P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, and V. S. Pande, ‘‘OpenMM 7: Rapid development of high performance algorithms for molecular dynamics,’’ *PLOS Computational Biology* **13**, e1005659 (2017).
- <sup>8</sup>A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand, and R. C. Walker, ‘‘Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born,’’ *Journal of Chemical Theory and Computation* **8**, 1542–1555 (2012).
- <sup>9</sup>C. Kutzner, S. Páll, M. Fechner, A. Esztermann, B. L. Groot, and H. Grubmüller, ‘‘More bang for your buck: Improved use of GPU nodes for GRO-MACS 2018,’’ *Journal of Computational Chemistry* **40**, 2418–2431 (2019).
- <sup>10</sup>R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, and R. C. Walker, ‘‘Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald,’’ *Journal of Chemical Theory and Computation* **9**, 3878–3888 (2013).
- <sup>11</sup>D. E. Shaw, J. Grossman, J. A. Bank, B. Batson, J. A. Butts, J. C. Chao, M. M. Deneroff, R. O. Dror, A. Even, C. H. Fenton, A. Forte, J. Gagliardo, G. Gill, B. Greskamp, C. R. Ho, D. J. Ierardi, L. Iserovich, J. S. Kuskin, R. H. Larson, T. Layman, L.-S. Lee, A. K. Lerer, C. Li, D. Killebrew, K. M. Mackenzie, S. Y.-H. Mok, M. A. Moraes, R. Mueller, L. J. Nociolo, J. L. Peticolas, T. Quan, D. Ramot, J. K. Salmon, D. P. Scarpazza, U. B. Schafer, N. Siddique, C. W. Snyder, J. Spengler, P. T. P. Tang, M. Theobald, H. Toma, B. Towles, B. Vitale, S. C. Wang, and C. Young, ‘‘Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer,’’ in *SC14: International Conference for High Performance Computing, Networking, Storage and Analysis* (IEEE, New Orleans, LA, USA, 2014) pp. 41–53.
- <sup>12</sup>A. Purohit, A. J. Schultz, and D. A. Kofke, ‘‘Force-sampling methods for density distributions as instances of mapped averaging,’’ *Molecular Physics* **117**, 2822–2829 (2019).
- <sup>13</sup>A. J. Schultz and D. A. Kofke, ‘‘Alternatives to conventional ensemble averages for thermodynamic properties,’’ *Current Opinion in Chemical Engineering* **23**, 70–76 (2019).
- <sup>14</sup>A. J. Schultz, S. G. Moustafa, W. Lin, S. J. Weinstein, and D. A. Kofke, ‘‘Reformulation of Ensemble Averages via Coordinate Mapping,’’ *Journal of Chemical Theory and Computation* **12**, 1491–1498 (2016).
- <sup>15</sup>N. A. Mahynski, J. R. Errington, and V. K. Shen, ‘‘Multivariable extrapolation of grand canonical free energy landscapes,’’ *The Journal of Chemical Physics* **147**, 234111 (2017).

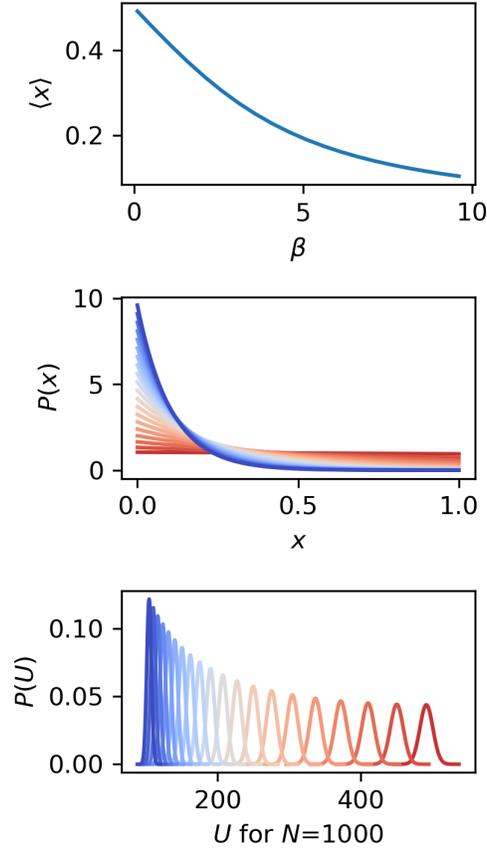
- <sup>16</sup>N. A. Mahynski, M. A. Blanco, J. R. Errington, and V. K. Shen, "Predicting low-temperature free energy landscapes with flat-histogram Monte Carlo methods," *The Journal of Chemical Physics* **146**, 074101 (2017).
- <sup>17</sup>N. A. Mahynski, J. R. Errington, and V. K. Shen, "Temperature extrapolation of multicomponent grand canonical free energy landscapes," *The Journal of Chemical Physics* **147**, 054105 (2017).
- <sup>18</sup>M. Witman, N. A. Mahynski, and B. Smit, "Flat-Histogram Monte Carlo as an Efficient Tool To Evaluate Adsorption Processes Involving Rigid and Deformable Molecules," *Journal of Chemical Theory and Computation* **14**, 6149–6158 (2018).
- <sup>19</sup>H. W. Hatch, S. Jiao, N. A. Mahynski, M. A. Blanco, and V. K. Shen, "Communication: Predicting virial coefficients and alchemical transformations by extrapolating Mayer-sampling Monte Carlo simulations," *The Journal of Chemical Physics* **147**, 231102 (2017).
- <sup>20</sup>N. A. Mahynski, S. Jiao, H. W. Hatch, M. A. Blanco, and V. K. Shen, "Predicting structural properties of fluids by thermodynamic extrapolation," *The Journal of Chemical Physics* **148**, 194105 (2018).
- <sup>21</sup>Z. A. Piskulich, O. O. Mesele, and W. H. Thompson, "Removing the barrier to the calculation of activation energies: Diffusion coefficients and reorientation times in liquid water," *The Journal of Chemical Physics* **147**, 134103 (2017).
- <sup>22</sup>Z. A. Piskulich, O. O. Mesele, and W. H. Thompson, "Activation Energies and Beyond," *The Journal of Physical Chemistry A* **123**, 7185–7194 (2019).
- <sup>23</sup>Z. A. Piskulich and W. H. Thompson, "The dynamics of supercooled water can be predicted from room temperature simulations," *The Journal of Chemical Physics* **152**, 074505 (2020).
- <sup>24</sup>E. Brini, C. J. Fennell, M. Fernandez-Serra, B. Hribar-Lee, M. Lukšič, and K. A. Dill, "How Water's Properties Are Encoded in Its Molecular Structure and Energies," *Chemical Reviews* **117**, 12385–12414 (2017).
- <sup>25</sup>P. Gallo, K. Amann-Winkel, C. A. Angell, M. A. Anisimov, F. Caupin, C. Chakravarty, E. Lascaris, T. Loerting, A. Z. Panagiotopoulos, J. Russo, J. A. Sellberg, H. E. Stanley, H. Tanaka, C. Vega, L. Xu, and L. G. M. Pettersson, "Water: A Tale of Two Liquids," *Chemical Reviews* **116**, 7463–7500 (2016).
- <sup>26</sup>M. Agarwal, M. P. Alam, and C. Chakravarty, "Thermodynamic, Diffusional, and Structural Anomalies in Rigid-Body Water Models," *The Journal of Physical Chemistry B* **115**, 6935–6945 (2011).
- <sup>27</sup>J. R. Errington and P. G. Debenedetti, "Relationship between structural order and the anomalies of liquid water," *Nature* **409**, 318–321 (2001).
- <sup>28</sup>J. I. Monroe and M. S. Shell, "Decoding signatures of structure, bulk thermodynamics, and solvation in three-body angle distributions of rigid water models," *The Journal of Chemical Physics* **151**, 094501 (2019).
- <sup>29</sup>Z. A. Piskulich and W. H. Thompson, "On the temperature dependence of liquid structure," *The Journal of Chemical Physics* **152**, 011102 (2020).
- <sup>30</sup>R. W. Zwanzig, "High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases," *The Journal of Chemical Physics* **22**, 1420–1426 (1954).
- <sup>31</sup>J. G. Kirkwood, "Statistical Mechanics of Fluid Mixtures," *The Journal of Chemical Physics* **3**, 300–313 (1935).
- <sup>32</sup>G. Hummer, "Fast-growth thermodynamic integration: Error and efficiency analysis," *The Journal of Chemical Physics* **114**, 7330–7337 (2001).
- <sup>33</sup>A. D. D. Craik, "Prehistory of Faà di Bruno's Formula," *The American Mathematical Monthly* **112**, 119–130 (2005).
- <sup>34</sup>L. Comtet, *Advanced Combinatorics* (D. Reidel Publishing Company, Dordrecht, Holland, 1974).
- <sup>35</sup>N. Lu and D. A. Kofke, "Understanding and Improving Free Energy Calculations in Molecular Simulations: Error Analysis and Reduction Methods," in *Free Energy Calculations*, edited by C. Chipot and A. Pohorille (Springer, 2007) pp. 199–247.
- <sup>36</sup>C. H. Bennett, "Efficient estimation of free energy differences from Monte Carlo data," *Journal of Computational Physics* **22**, 245–268 (1976).
- <sup>37</sup>K. A. Beauchamp, J. D. Chodera, L. N. Naden, and M. R. Shirts, "Pymbar,".
- <sup>38</sup>K. Kamat and B. Peters, "Diabat Interpolation for Polymorph Free-Energy Differences," *The Journal of Physical Chemistry Letters* **8**, 655–660 (2017).
- <sup>39</sup>P. V. Klimovich, M. R. Shirts, and D. L. Mobley, "Guidelines for the analysis of free energy calculations," *Journal of Computer-Aided Molecular Design* **29**, 397–411 (2015).
- <sup>40</sup>M. R. Shirts and J. D. Chodera, "Statistically optimal analysis of samples from multiple equilibrium states," *The Journal of Chemical Physics* **129**, 124105 (2008).
- <sup>41</sup>M. R. Shirts and V. S. Pande, "Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration," *The Journal of Chemical Physics* **122**, 144107 (2005).
- <sup>42</sup>S. Bruckner and S. Boresch, "Efficiency of alchemical free energy simulations. II. Improvements for thermodynamic integration," *Journal of Computational Chemistry* **32**, 1320–1333 (2011).
- <sup>43</sup>D. A. Kofke, "Free energy methods in molecular simulation," *Fluid Phase Equilibria* **228-229**, 41–48 (2005).
- <sup>44</sup>N. Lu and D. A. Kofke, "Optimal intermediates in staged free energy calculations," *The Journal of Chemical Physics* **111**, 4414–4423 (1999).
- <sup>45</sup>A. Pohorille, C. Jarzynski, and C. Chipot, "Good Practices in Free-Energy Calculations," *The Journal of Physical Chemistry B* **114**, 10235–10253 (2010).
- <sup>46</sup>J. L. F. Abascal and C. Vega, "A general purpose model for the condensed phases of water: TIP4P/2005," *The Journal of Chemical Physics* **123**, 234505 (2005).
- <sup>47</sup>M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX* **1-2**, 19–25 (2015).
- <sup>48</sup>G. Hummer, S. Garde, A. E. Garcia, A. Pohorille, and L. R. Pratt, "An information theory model of hydrophobic interactions," *Proceedings of the National Academy of Sciences* **93**, 8951–8955 (1996).
- <sup>49</sup>S. N. Jamadagni, R. Godawat, and S. Garde, "Hydrophobicity of Proteins and Interfaces: Insights from Density Fluctuations," *Annual Review of Chemical and Biomolecular Engineering* **2**, 147–171 (2011).
- <sup>50</sup>V. Molinero and E. B. Moore, "Water Modeled As an Intermediate Element between Carbon and Silicon <sup>†</sup>," *The Journal of Physical Chemistry B* **113**, 4008–4016 (2009).
- <sup>51</sup>P.-L. Chau and A. J. Hardwick, "A new order parameter for tetrahedral configurations," *Molecular Physics* **93**, 511–518 (1998).
- <sup>52</sup>A. Chaimovich and M. S. Shell, "Tetrahedrality and structural order for hydrophobic interactions in a coarse-grained water model," *Physical Review E* **89**, 022140 (2014).
- <sup>53</sup>H. S. Ashbaugh and L. R. Pratt, "*Colloquium* : Scaled particle theory and the length scales of hydrophobicity," *Reviews of Modern Physics* **78**, 159–178 (2006).
- <sup>54</sup>A. P. Willard and D. Chandler, "Instantaneous Liquid Interfaces," *The Journal of Physical Chemistry B* **114**, 1954–1958 (2010).
- <sup>55</sup>A. J. Patel, P. Varilly, S. N. Jamadagni, M. F. Hagan, D. Chandler, and S. Garde, "Sitting at the Edge: How Biomolecules use Hydrophobicity to Tune Their Interactions and Function," *The Journal of Physical Chemistry B* **116**, 2498–2503 (2012).
- <sup>56</sup>P. G. Debenedetti, *Metastable Liquids: Concepts and Principles*, Physical Chemistry: Science and Engineering (Princeton University Press, Princeton, NJ, 1997).
- <sup>57</sup>C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, Adaptive Computation and Machine Learning (MIT Press, Cambridge, MA, 2006).
- <sup>58</sup>D. A. McQuarrie, *Statistical Mechanics* (University Science Books, Sausalito, CA, 2000).
- <sup>59</sup>M. S. Shell, *Thermodynamics and Statistical Mechanics* (Cambridge University Press, Cambridge, United Kingdom, 2015).

Supplementary Material: Extrapolation and  
interpolation strategies for efficiently estimating  
structural observables as a function of  
temperature and density

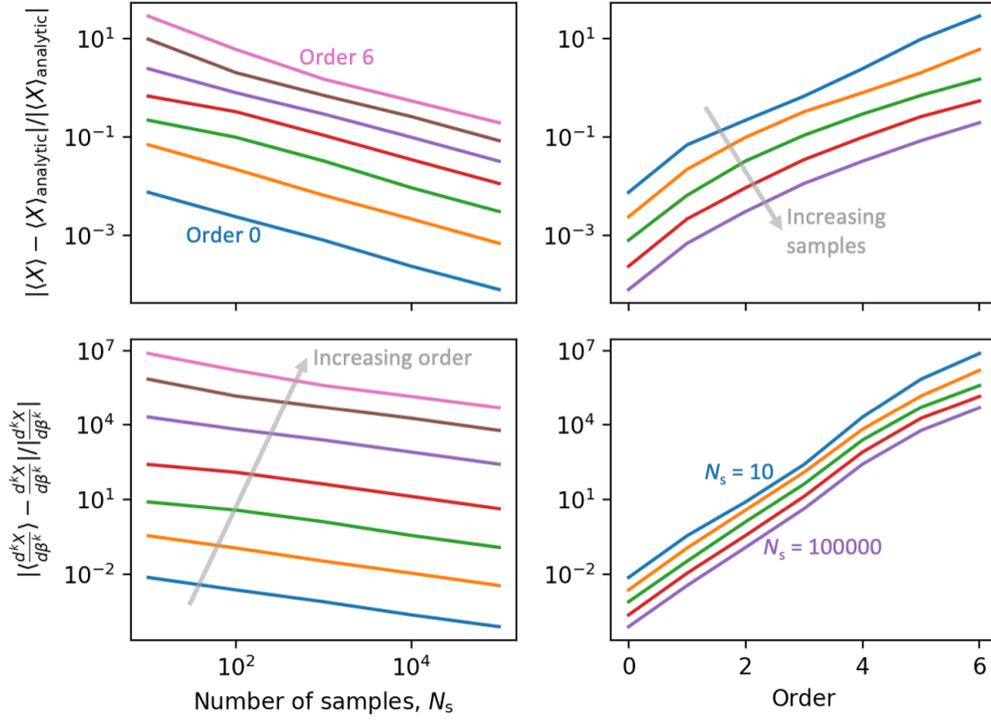
Jacob I. Monroe    Harold W. Hatch    Nathan A. Mahynski  
M. Scott Shell    Vincent K. Shen

September 3, 2020

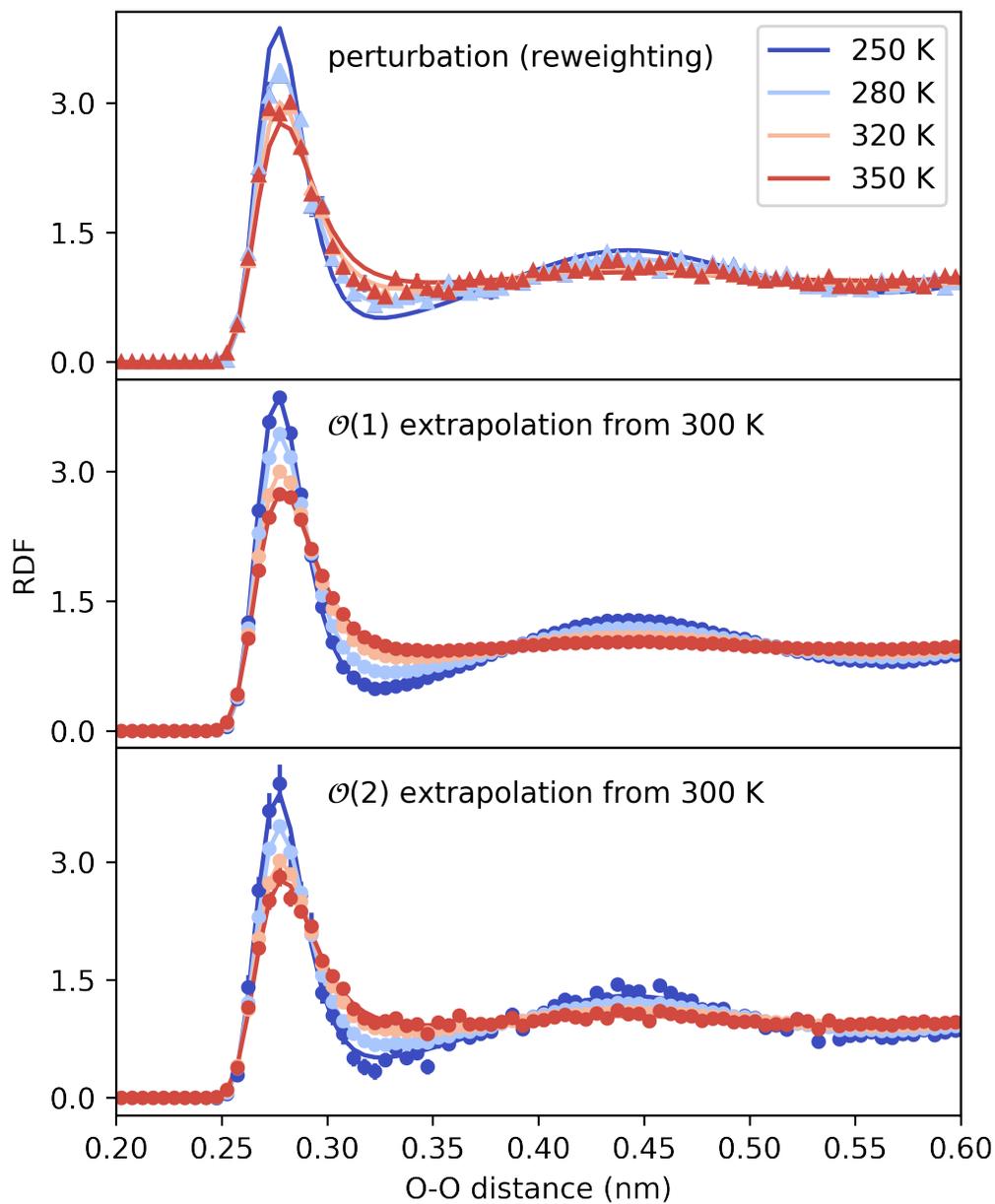
# 1 Supporting Figures



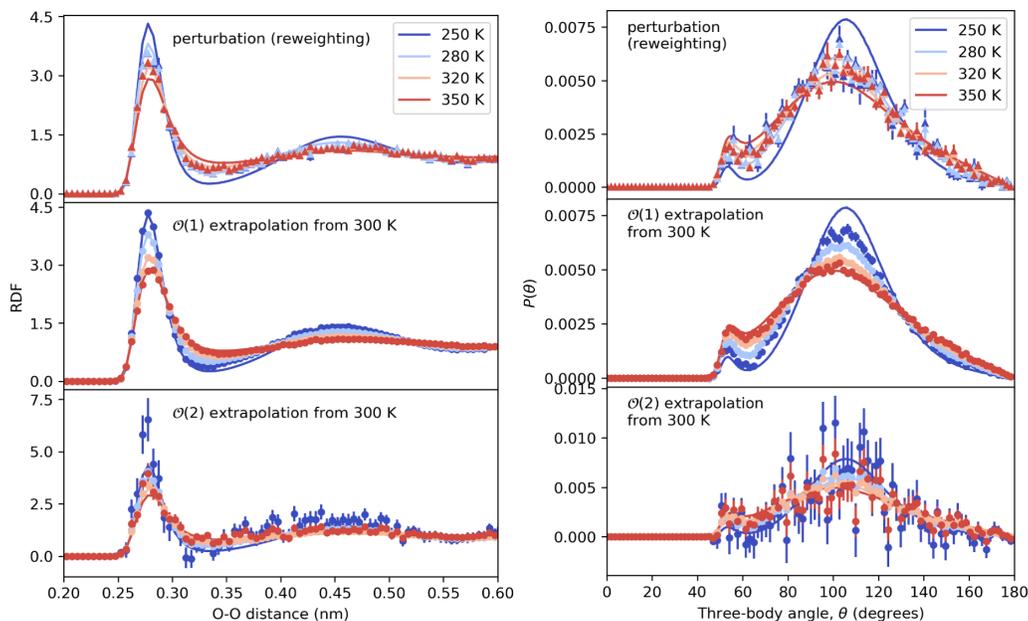
**Figure S1:** For the 1D ideal gas in an external field described in the text, a simple structural property of interest is the average  $x$  value, which is plotted in the top panel as a function of  $\beta$ . Changes in  $P(x)$  and  $P(U)$  with temperature are shown as well with coloring by temperature (blue is the lowest  $T$ , highest  $\beta$ ). Clearly the configurational distributions will not overlap in their important regions and neither will the potential energy distributions.



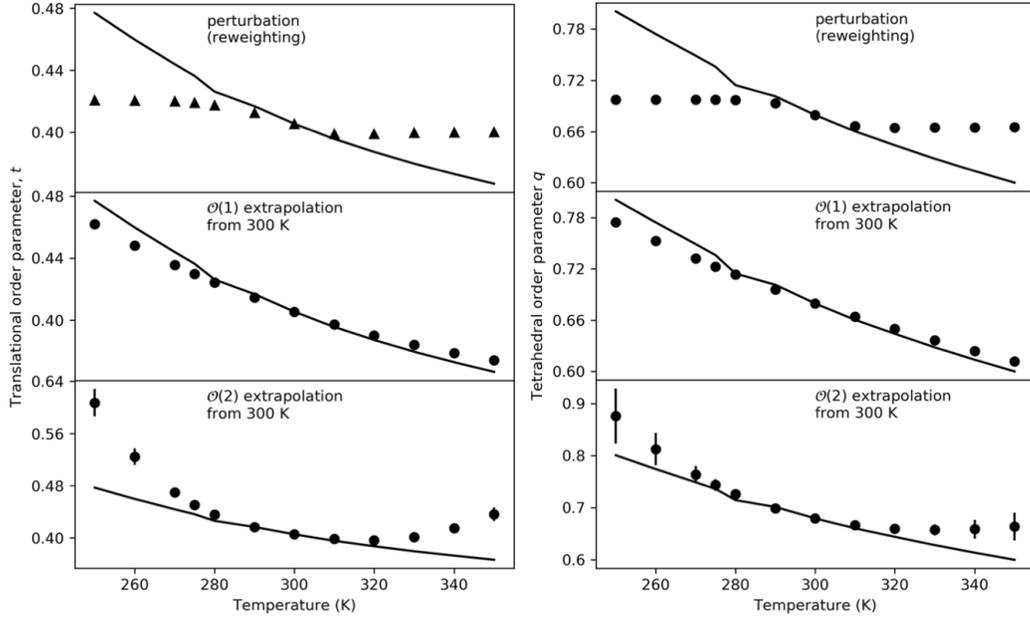
**Figure S2:** Relative unsigned error from extrapolation values at infinite sampling ( $\langle X \rangle_{\text{analytic}}$ ) for both observables (top) and derivatives (bottom) are shown as functions of the number of samples (left) and order of extrapolation or derivative (right). At low order, extrapolation results quickly converge to associated analytic values with few samples.



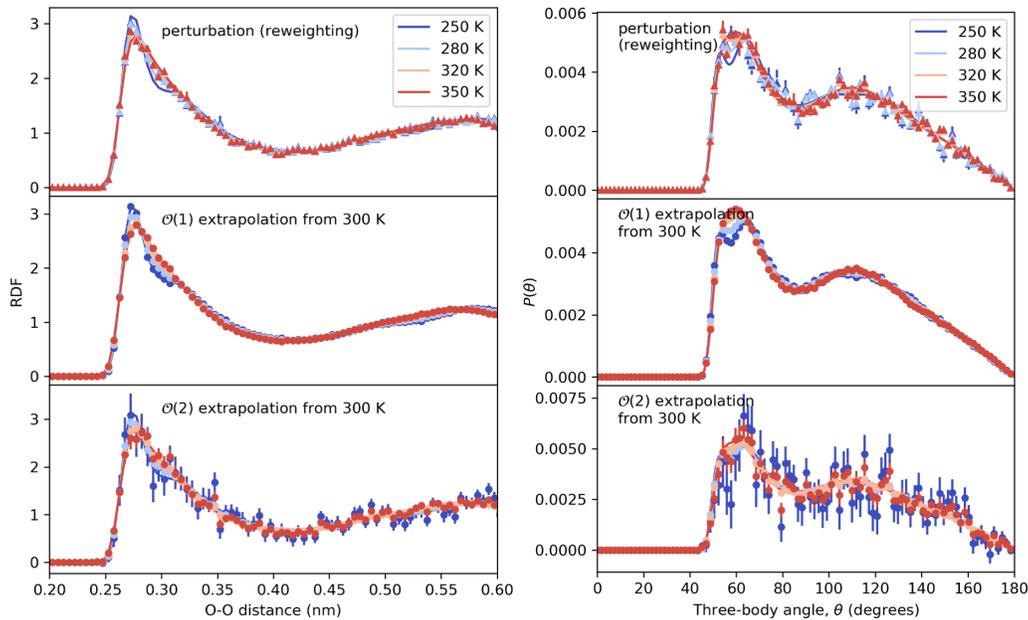
**Figure S3:** In all panels, lines represent the RDFs from direct simulation at each temperature while points are the estimates via extrapolation from 300 K (constant density of  $1.00 \text{ g/cm}^3$ ) with 50 ns of simulation (50000 snapshots saved every 1 ps). Error bars represent one standard deviation and are determined through 100 bootstrap resamples of the original data.



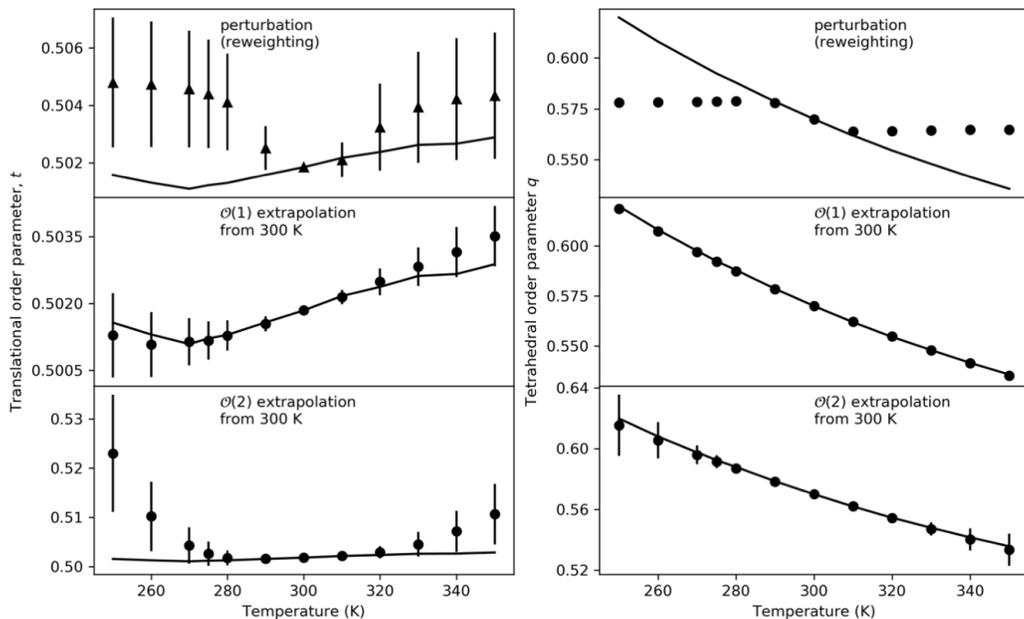
**Figure S4:** In all panels, lines represent the RDFs or three-body angle distributions from direct simulation at each temperature while points are the estimates via extrapolation from 300 K (constant density of  $0.87 \text{ g/cm}^3$ ) with 5 ns of simulation (5000 snapshots saved every 1 ps). Error bars represent one standard deviation and are determined through bootstrap resampling. The neighbor cutoff was set to 0.34 nm in computing three-body angles. When extrapolating three-body angle distributions to lower temperatures, the derivative is underestimated, indicating that at this density a linear extrapolation is not fully sufficient. Figure S5 also reveals that the slope of the tetrahedral order parameter with temperature increases at lower temperatures for this density.



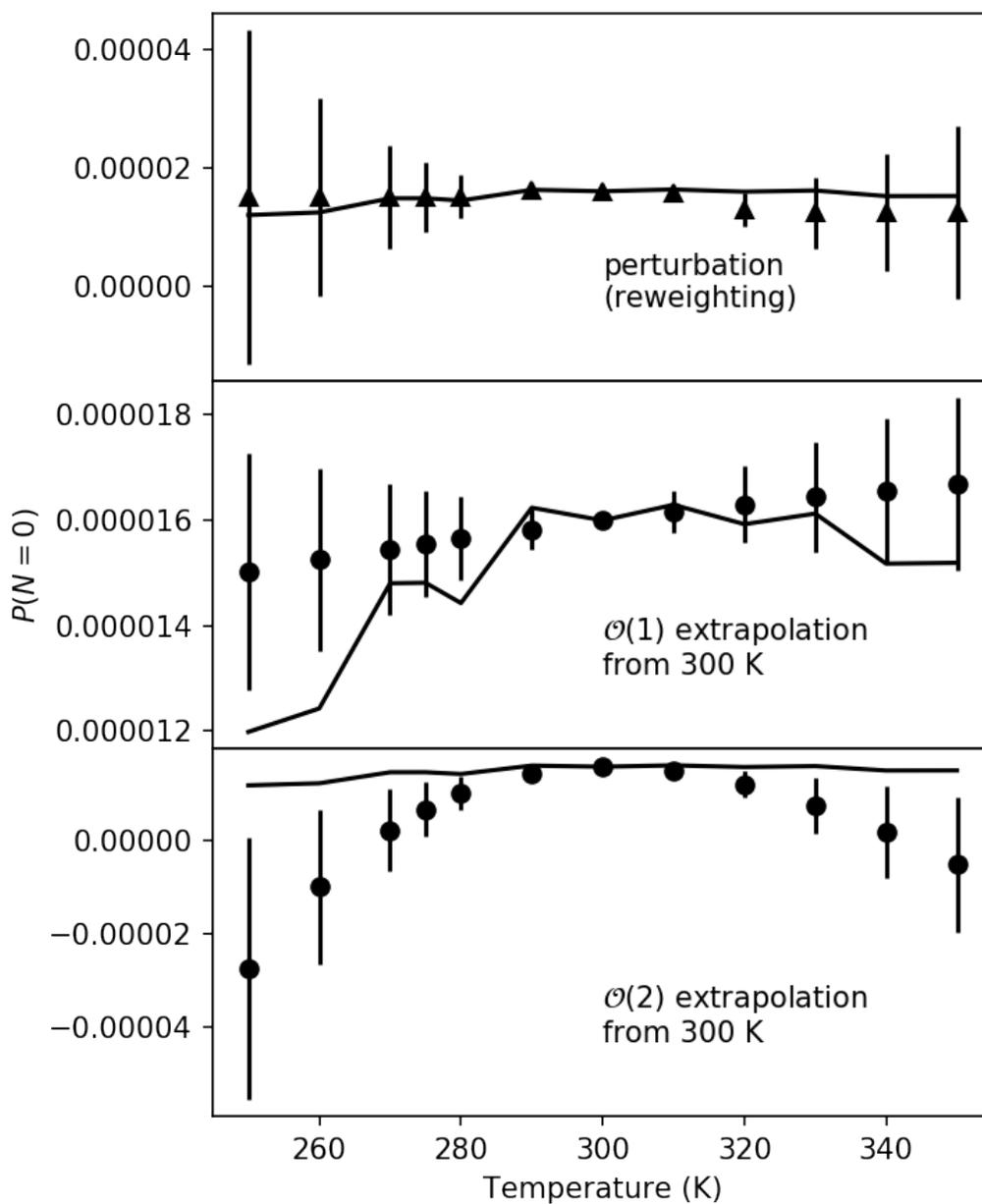
**Figure S5:** In all panels, lines represent the direct simulation translational or tetrahedral order parameter values while points are the estimates via perturbation or extrapolation from 300 K (constant density of  $0.87 \text{ g/cm}^3$ ) with 5 ns of simulation (5000 snapshots saved every 1 ps). Error bars represent one standard deviation and are determined through bootstrap resampling. For translational order parameters, this involves resampling to obtain uncertainties in RDF bins and propagating the uncertainty associated with application of Simpson's 1/3 rule.



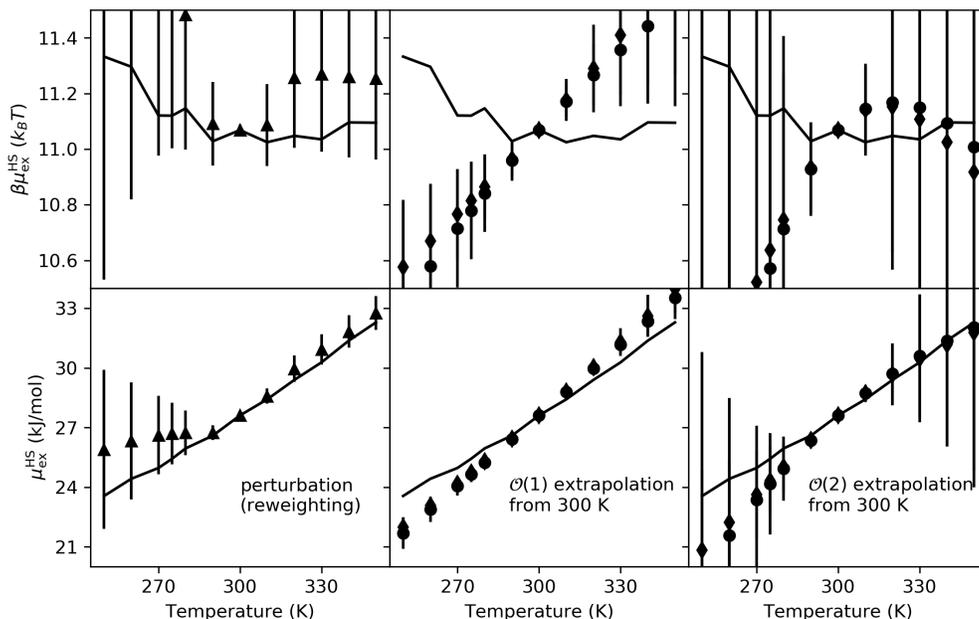
**Figure S6:** In all panels, lines represent the RDFs or three-body angle distributions from direct simulation at each temperature while points are the estimates via extrapolation from 300 K (constant density of  $1.40 \text{ g/cm}^3$ ) with 5 ns of simulation (5000 snapshots saved every 1 ps). Error bars represent one standard deviation and are determined through bootstrap resampling. The neighbor cutoff was set to 0.34 nm in computing three-body angles.



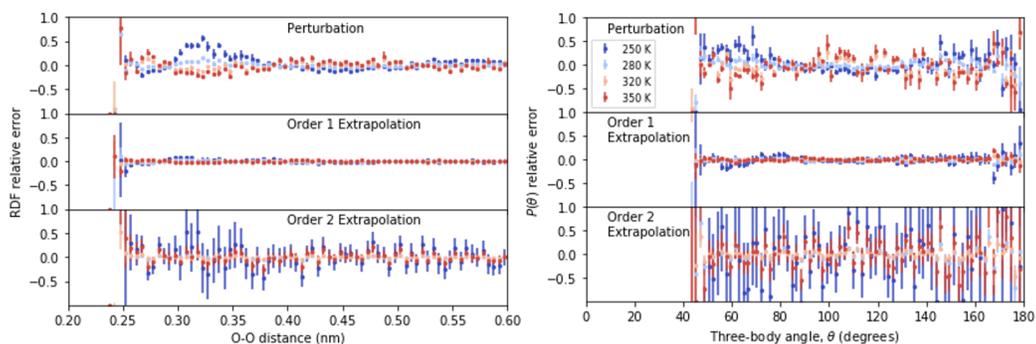
**Figure S7:** In all panels, lines represent the direct simulation translational or tetrahedral order parameter values while points are the estimates via perturbation or extrapolation from 300 K (constant density of  $1.40 \text{ g/cm}^3$ ) with 5 ns of simulation (5000 snapshots saved every 1 ps). Error bars represent one standard deviation and are determined through bootstrap resampling. Non-monotonic behavior in the translational order parameter at this density is captured because this quantity is not directly extrapolated. Successful first-order extrapolation of RDFs in Figure S6 leads to accurate translational order parameters.



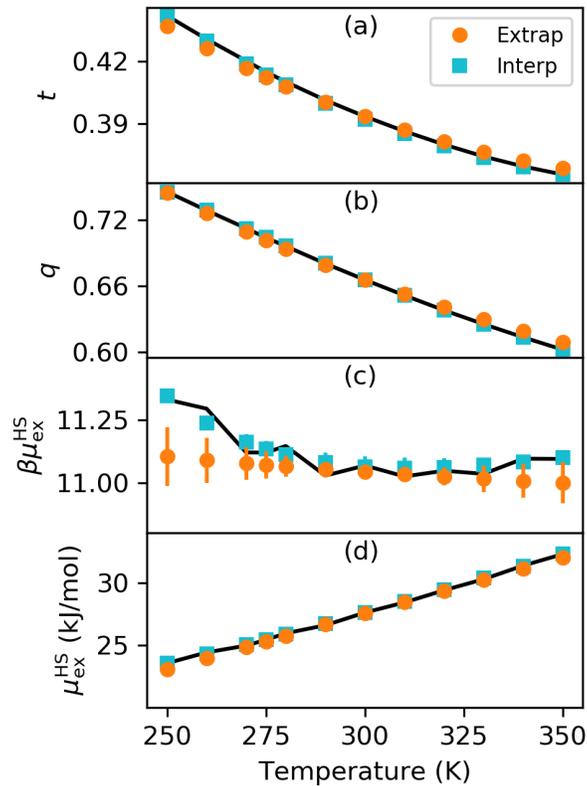
**Figure S8:** Extrapolation and perturbation results in temperature from reference simulations of 50 ns at 300 K are shown for the probability of having zero water oxygens within the hard-sphere probe volume at a constant density 1.00 g/cm<sup>3</sup>.



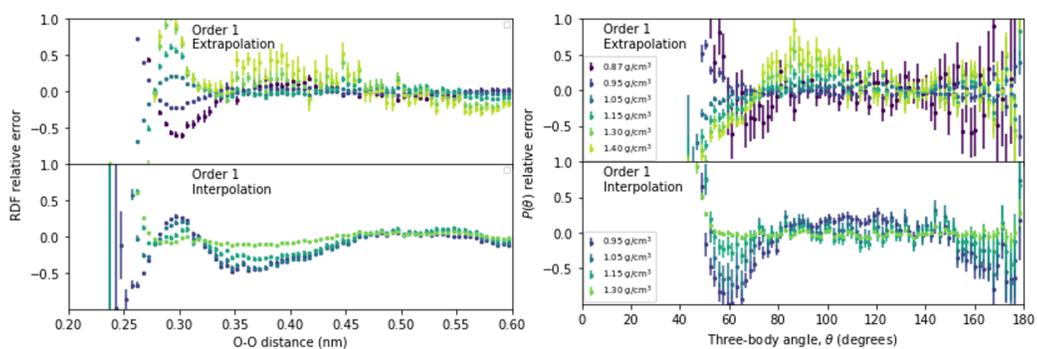
**Figure S9:** Extrapolation and perturbation of  $\beta\mu_{\text{ex}}^{\text{HS}}$  and  $\mu_{\text{ex}}^{\text{HS}}$  are shown as a function of temperature based on reference simulations of 5 ns at 300 K and 1.00 g/cm<sup>3</sup>. Triangles are used for perturbation results, while circles represent direct extrapolation in of  $\beta\mu_{\text{ex}}^{\text{HS}}$ . Diamonds are determined by extrapolating  $P(N = 0)$  and taking the negative of the natural logarithm and generally agree closely with direct extrapolations.



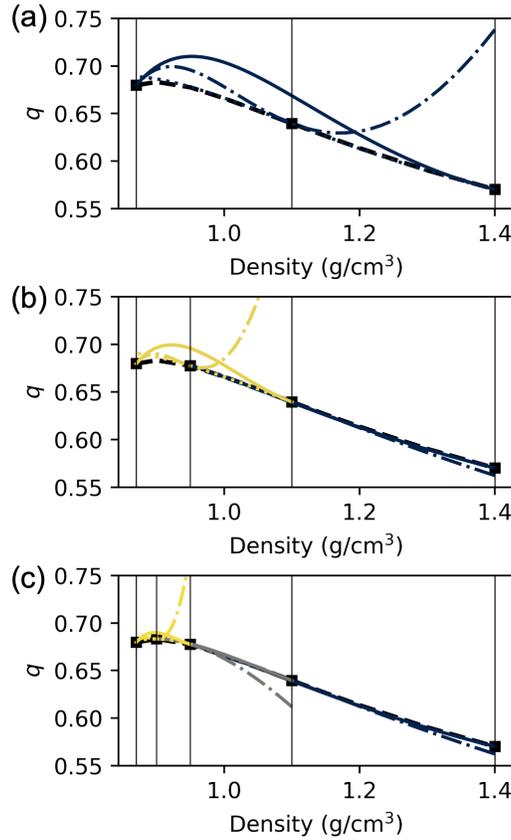
**Figure S10:** Relative error of RDF (left) and three-body angle (right) extrapolations from a reference temperature of 300 K are shown. The same data is shown in Figure 5, but here the RDFs calculated from simulations at each temperature are subtracted from extrapolations and then used to normalize this difference.



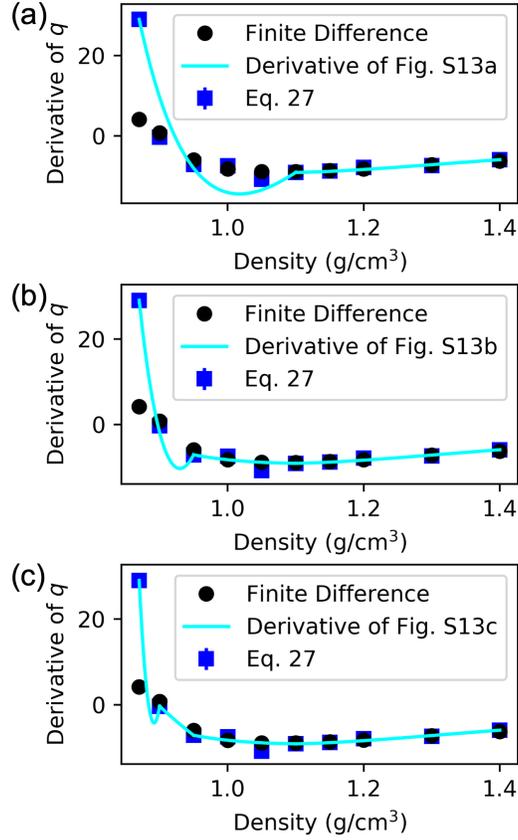
**Figure S11:** Direct simulation results (solid lines), first-order extrapolations from 300 K (circles), and interpolations from the most extreme temperatures (squares) are shown versus temperature (at constant density of  $1.00 \text{ g/cm}^3$ ) for (a) translational order parameters, (b) tetrahedral order parameters, (c)/(d) excess chemical potentials of 0.33 nm radii hard-spheres. Translational order parameters are computed from RDFs extrapolated in temperature from 5 ns simulations at 300 K and  $1.00 \text{ g/cm}^3$ , while all other quantities are extrapolated directly from the same reference state. Interpolations utilize two simulations, one from each of the most extreme densities, and hence twice as much data as presented extrapolations. Due to the inherently high ratio of noise to temperature variation for excess chemical potential estimates at this density, 50 ns of simulation are used for extrapolations and interpolations in (c) and (d).



**Figure S12:** Relative error of RDF (left) and three-body angle (right) extrapolations from a density of  $1.00 \text{ g/cm}^3$  or interpolations from the most extreme densities are shown. The same data is shown in Figure 7, but here the RDFs calculated from simulations at each density are subtracted from extrapolations and then used to normalize this difference.



**Figure S13:** An example implementation of the recursive algorithm and manual visual consistency check is presented for the tetrahedral order parameter over water densities at a fixed temperature of 300 K. As in Fig. 9, polynomial fits are shown over sliding sets of three points. For each set, a polynomial in the same color is plotted based on each pair of points (solid lines for the outermost points, dotted for the higher-density interval, and dash-dot lower-density). In step (a), we choose the point with the highest bootstrapped error over the initial interval defined by the two most extreme densities. The relative error from the results based on brute force simulations goes from a maximum of 6% to 3%. Since the consistency check is most poorly satisfied on the lower-density interval, we select the point of highest bootstrap error on this interval and add its data, lowering the relative error to less than 1% (b). Note that both the low- and high-density intervals now more fully satisfy the visual consistency check, validating our choice of selecting a new point in the low-density region. Finally, in (c) we select another point on the lowest-density interval, further improving visual consistency. The end result is within 0.01% error of all reference points (thick black dashed curve), but utilizes only half of the data. If, through intuition or reasoning, we had initially selected the point at the maximum of the polynomial predicted by the two most extreme density points (rather than the point with the maximum bootstrapped error), the relative error would drop to less than 1% with only 3 points, using only 30% of the original data.



**Figure S14:** Each panel displays the derivative of the piecewise interpolating polynomial described in the corresponding panels of Fig. S13 (light blue curves). By design, these curves pass exactly through sets of points (vertical lines in Fig. S13) at which the derivative of the tetrahedral order parameter with respect to density is estimated using Eq. 27 (blue squares). As a comparison, finite difference estimates of derivatives are shown as black circles, using central difference accounting for unequally spaced data for internal points and forward/backward difference for the leftmost/rightmost data points. It is clear that the curvature increases sharply at low density, with derivatives changing rapidly below  $0.95 \text{ g/cm}^3$ . Due to these rapid variations, estimates from Eq. 27 and finite differences begin to disagree. In particular, the lowest density finite difference value is determined based on information at higher densities where the derivative is not so steep, leading to an underestimate. Note that the derivatives of interpolations in Fig. S13 are not the best polynomial fits to solely the derivative data, but also include information concerning the observable values themselves.