Reformulation of ensemble averages via coordinate mapping

Andrew J. Schultz,[†] Sabry G. Moustafa,[†] Weisong Lin,[†] Steven J. Weinstein,[‡] and David A. Kofke[*,†]

[†]*Department of Chemical and Biological Engineering, University at Buffalo, The State University of New York, Buffalo, New York 14260-4200, USA*

[‡]*Department of Chemical Engineering, Rochester Institute of Technology, Rochester, NY 14623, USA*

E-mail: kofke@buffalo.edu

Abstract

A general framework is established for reformulation of the ensemble averages commonly encountered in statistical mechanics. This "mapped-averaging" scheme allows approximate theoretical results that have been derived from statistical mechanics to be reintroduced into the underlying formalism, yielding new ensemble averages that represent exactly the error in the theory. The result represents a distinct alternative to perturbation theory for methodically employing tractable systems as a starting point for describing complex systems. Molecular simulation is shown to provide one appealing route to exploit this advance. Calculation of the reformulated averages by molecular simulation can proceed without contamination by noise produced by behavior that has already been captured by the approximate theory. Consequently, accurate and precise values of properties can be obtained while using less computational effort—in favorable cases, many orders of magnitude less. The treatment is demonstrated using three

examples: (1) calculation of the heat capacity of an embedded-atom model of iron; (2) calculation of the dielectric constant of the Stockmayer model of dipolar molecules; and (3) calculation of the pressure of a Lennard-Jones fluid. It is observed that improvement in computational efficiency is related to the appropriateness of the underlying theory for the condition being simulated; the accuracy of the result is however not impacted by this. The framework opens many avenues for further development, both as a means to improve simulation methodology and as a new basis to develop theories for thermophysical properties.

Introduction

The basic framework of statistical mechanics was established over a century ago, and has provided the foundation for countless theories, methods, and applications. A key element in the formalism is the notion of an ensemble average, in which equilibrium properties are expressed in terms of averages over all relevant configuration variables for the chosen ensemble. Lacking in this framework is a systematic means to integrate estimated behaviors: often it is possible to characterize complex systems using approximate but tractable models that capture its essential features. Perturbation theory is the usual route to exploiting these models.[1] In a typical case, the complex behavior is expressed as a series expanded about the approximate form, in terms of a well-defined parameter that quantifies how much the complex system differs from its simplified form. Examples include the high-temperature expansion, taken with respect to a hard- or soft-core reference, with reciprocal temperature serving as the expansion variable; the virial equation, taken with respect to the ideal gas, with density as the expansion variable. Perturbation methods, however, have severe limitations: they become inaccurate when applied at conditions that are poorly described by the approximate behavior; series coefficients become increasingly difficult to derive and calculate with increasing series order; and the convergence of the series is often uncertain. The essential problem with such methods is that they "decorate" the approximate model,

rather than incorporate it more deeply in the formalism describing the system of interest.

Recently, we demonstrated[2] a different way to build on approximate knowledge of a statistical mechanical system, a way that reformulates the ensemble average itself. The approach that we presented applies specifically to monatomic crystalline materials, and exploits the approximate harmonic behavior exhibited by them. However, the basic framework that accomplished this transformation is general, and could in principle be applied to an endless variety of systems and properties. The element of our initial application that makes it specific to crystalline systems is a coordinate mapping, one having a form that was guided in an *ad hoc* manner by consideration of harmonic behavior. While this approach was shown to be extremely effective, it is difficult to transfer to other cases without a more systematic means to formulate the mapping. This is the problem that we address in the present work.

Formalism

In what follows, we will use bold symbols (e.g., $\mathbf{x}$) to represent dN-dimensional coordinate vectors, where d is the number of degrees of freedom associated with each of N molecules in the system. Also, we use greek letters to represent thermodynamic state variables (or more generally, parameters of the Hamiltonian), with λ representing a set of such parameters (e.g., $\lambda = (\mu, \nu)$, where μ and ν may be temperature, pressure, etc.).

We begin by reviewing the mapped averaging framework.[2] Mapped averaging is derived from the idea of targeted perturbation, which was suggested by Jarzynski[3] as a means to improve the calculation of free energy (FE) differences. With it, a bijective coordinate mapping $\mathbf{X} \to \mathbf{x} = \mathcal{X}(\mathbf{X}, \lambda)$ accompanies the parameter perturbation $\lambda_0 \to \lambda$. The mapping is meant to address a pernicious problem with FE perturbation, in which the $\lambda_0 \to \lambda$ perturbation produces a condition in which the coordinates $\mathbf{X}$ (which are importance-sampled at λ_0) are poorly suited for the perturbed state λ. The outcome is a degradation of the precision of the calculation, and even more damaging, the introduction of a bias in the computed FE

difference.[4] The mapping $\mathbf{X} \to \mathbf{x}$ aims to transform the coordinates such that they are more representative of those encountered in the λ state, and thereby attenuate these problems. The FE difference $\Delta\mathcal{A} = \mathcal{A}(\lambda) - \mathcal{A}(\lambda_0)$ is given by this targeted perturbation according to:[3]

$$\Delta\mathcal{A} = -\ln\left\langle Je^{-\Delta\mathcal{U}}\right\rangle. \tag{1}$$

Here, $J(\mathbf{X},\lambda) \equiv |\partial\mathbf{x}/\partial\mathbf{X}|$ is the Jacobian of the mapping, and

$$\begin{aligned}\Delta\mathcal{U}(\mathbf{X},\lambda) &= \mathcal{U}(\mathbf{X},\lambda) - \mathcal{U}(\mathbf{X},\lambda_0)\\ &\equiv u(\mathcal{X}(\mathbf{X},\lambda),\lambda) - u(\mathbf{X},\lambda_0),\end{aligned} \tag{2}$$

where $u(\mathbf{x},\lambda)$ is the energy of the configuration with coordinate vector $\mathbf{x}$; angle brackets $\langle\ldots\rangle$ in (1) represent an ensemble average in the λ_0 system. Physically, $\Delta\mathcal{U}$ represents the change in configurational energy resulting from the combined effects of the perturbation $\lambda_0 \to \lambda$ and the mapping $\mathbf{X} \to \mathbf{x}$. The symbols $\mathcal{U}$ and u indicate, respectively, Lagrangian and Eulerian representations of the same quantity: the arguments to $\mathcal{U}$ are the starting (unmapped) coordinate $\mathbf{X}$ and the transformation variable λ, with the mapping implicit in its definition; whereas u is the energy function as it is conventionally understood, so here it takes as arguments the current (already mapped) coordinate $\mathbf{x}$ and the corresponding value of λ. Eulerian versus Lagrangian representations are explained more explicitly further below. The energies $\mathcal{A}$, $\mathcal{U}$, and u each include in their definition a multiplication by $\beta \equiv (k_B T)^{-1}$, where T is the temperature and k_B is the Boltzmann constant, so they are all unitless. Targeted perturbation has been applied in a only a small number of cases (albeit quite successfully).[5–7] The likely reason for its limited impact is again the difficulty of formulating an effective mapping.

Jarzynski developed targeted perturbation with the aim to improve FE calculations for finite perturbations $\lambda_0 \to \lambda$, but the framework may be more valuable when applied using differential perturbations, such that $\lambda = \lambda_0 + d\lambda$ — not to describe the λ state; rather, to

better characterize the λ_0 state itself. Expansion of Eq. (1) to second order in $d\lambda$ yields expressions for the first and second derivatives of the FE, which, according to thermodynamics, represent the interesting physical properties of the λ_0 system. Using subscripts ν and μ to denote derivatives with respect to elements of λ, we have:[2]

$$\mathcal{A}_\nu = -\langle J_\nu \rangle + \langle\, \mathcal{U}_\nu \rangle \tag{3a}$$

$$\begin{aligned} \mathcal{A}_{\mu\nu} = &-\langle J_{\mu\nu} - J_\mu J_\nu \rangle + \langle\, \mathcal{U}_{\mu\nu} \rangle \\ &- \mathrm{Cov}\left[J_\mu - \mathcal{U}_\mu, J_\nu - \mathcal{U}_\nu\right] \end{aligned} \tag{3b}$$

where $\mathrm{Cov}\left[\mathrm{X}, \mathrm{Y}\right] \equiv \langle \mathrm{XY} \rangle - \langle \mathrm{X} \rangle \langle \mathrm{Y} \rangle$. The derivatives of $\mathcal{A}$ have a straightforward interpretation, but those involving $\mathcal{U}$ and J warrant some discussion. In moving to the limit of a differential perturbation, the mapping takes the form of a velocity-like field on the space of $\mathbf{x}$, with λ playing the role normally held by time t. Accordingly, it becomes useful to invoke concepts familiar to classical field theory, and in particular the notion of Lagrangian versus Eulerian specifications of the field.[8]

Let $\mathbf{x}$ be a coordinate reached via a trajectory that follows the mapping field, initiated at the point $\mathbf{X}$ at $\lambda = \lambda_0$; we represent this trajectory in Lagrangian form as $\mathbf{x} = \mathcal{X}(\mathbf{X}, \lambda)$, such that $\mathcal{X}(\mathbf{X}, \lambda_0) = \mathbf{X}$ (cf. Eq. (2)). Any Lagrangian function $\Psi(\mathbf{X}, \lambda)$ may be converted to its Eulerian counterpart $\psi(\mathbf{x}, \lambda)$ by inverting the Lagrangian trajectory as $\mathbf{X} = \mathcal{X}^{-1}(\mathbf{x}, \lambda)$, and writing $\Psi(\mathbf{X}, \lambda) = \Psi(\mathcal{X}^{-1}(\mathbf{x}, \lambda), \lambda) \equiv \psi(\mathbf{x}, \lambda)$. The distinction is important when taking derivatives with respect to λ: Lagrangian derivatives are performed at constant initial coordinate $\mathbf{X}$, while the Eulerian derivatives are performed at constant current coordinate $\mathbf{x}$. Relating them requires an explicit accounting of the mapping, i.e., the dependence of $\mathbf{x}$ on λ. Following the standard relationships between Lagrangian and Eulerian perspectives used in continuum mechanics,[8] the Lagrangian and Eulerian derivatives are related as:

$$\Psi_\nu = \psi_\nu + \mathbf{v}^\nu \cdot \nabla \psi, \tag{4}$$

where ∇ is the gradient operator in the space of $\mathbf{x}$. In Eq. (4), $\mathbf{v}^\nu$ is an analog to an Eulerian velocity, defined in terms of its Lagrangian equivalent $\mathbf{V}^\nu(\mathbf{X},\lambda)$ (remembering $\nu \in \lambda$):

$$\frac{\partial}{\partial\nu}\mathcal{X}(\mathbf{X},\lambda) \equiv \mathbf{V}^\nu(\mathbf{X},\lambda) = \mathbf{V}^\nu(\mathcal{X}^{-1}(\mathbf{x},\lambda),\lambda) \equiv \mathbf{v}^\nu(\mathbf{x},\lambda) \tag{5}$$

One should note that the Eulerian cross derivatives are not necessarily equal (see Supplemental Material (SM)[9]):

$$\mathbf{v}^\mu_\nu \equiv \left(\frac{\partial\mathbf{v}^\mu}{\partial\nu}\right)_{\mathbf{x},\mu} \neq \left(\frac{\partial\mathbf{v}^\nu}{\partial\mu}\right)_{\mathbf{x},\nu} \equiv \mathbf{v}^\nu_\mu$$

Following the notation in Eq. (4), we will use upper-case symbols to represent variables in the Lagrangian specification, and lower-case symbols to indicate the Eulerian specification.

Based on its definition given in Eq. (2), $\mathcal{U}$ is Lagrangian, and accordingly its derivatives are expressed:

$$\mathcal{U}_\nu = u_\nu - \mathbf{v}^\nu \cdot \mathbf{f} \tag{6a}$$

$$\begin{aligned}\mathcal{U}_{\mu\nu} = u_{\mu\nu} &- \left(\mathbf{v}^\nu_\mu + \mathbf{v}^\mu \cdot \nabla\mathbf{v}^\nu\right) \cdot \mathbf{f} + \mathbf{v}^\mu \cdot \boldsymbol{\phi} \cdot \mathbf{v}^\nu \\ &- \left(\mathbf{v}^\nu \cdot \mathbf{f}_\mu + \mathbf{v}^\mu \cdot \mathbf{f}_\nu\right),\end{aligned} \tag{6b}$$

where $\mathbf{f} \equiv -\nabla u$ is the force vector and $\boldsymbol{\phi} \equiv \nabla\nabla u$ is the force-constant matrix (each implicitly multiplied by β, as with $\mathcal{U}$ and u); we see again that u, which was first invoked in Eq. (2), is the Eulerian representation of $\mathcal{U}$. What we accomplish with Eq. (6) is to express the Lagrangian derivatives required by Eq. (3) in terms of the more conventional (Eulerian) derivatives, which is the only type of derivative that appears on the right-hand side of Eq. (6).

For present purposes, we require all derivatives in Eqs. (3) and (6) to be evaluated at $\lambda = \lambda_0$ (at which $\mathbf{x} = \mathbf{X}$). We will consider the J derivatives in the next section, as they are closely related to the definition of the mapping.

Mapping

A perfect mapping will be one for which the integrand in Eq. (1) is constant, such that for all $\mathbf{X}$, the mapping assures that

$$J(\mathbf{X},\lambda)\frac{P(\mathbf{X},\lambda)}{P(\mathbf{X},\lambda_0)} = \frac{q(\lambda)}{q(\lambda_0)}, \tag{7}$$

or

$$J(\mathbf{X},\lambda)\frac{P(\mathbf{X},\lambda)}{q(\lambda)} = \frac{P(\mathbf{X},\lambda_0)}{q(\lambda_0)}; \tag{8}$$

we define $P \equiv \exp(-\mathcal{U})$, $q \equiv \exp(-\mathcal{A}) = \int p(\mathbf{x},\lambda)\mathrm{d}\mathbf{x}$ with $p \equiv \exp(-u)$ (we denote this q rather than Q here because, following convention, it is defined via an integral of an Eulerian quantity). Physically, (8) means that the normalized weight of the points, $P(\mathbf{X},\lambda)/q(\lambda)$, is unchanged along the trajectory. This "conservation rule" for P/q can be expressed more usefully if we shift from a Lagrangian to an Eulerian perspective. Details are provided in the SM, with the result:

$$\frac{\partial}{\partial\nu}\left(\frac{p}{q}\right) + \nabla\cdot\left(\frac{p}{q}\mathbf{v}^\nu\right) = 0. \tag{9}$$

This formula allows, at least in principle, for direct solution of the mapping field $\mathbf{v}^\nu$ given p and q, which is far more convenient than arriving at the same via Eq. (8).

The SM provides an analysis of the evolution of the Jacobian in the Eulerian specification. The results developed there yield expressions for the Jacobian derivatives required by Eq. (3) for a mapping field satisfying Eq. (9). Our interest is for $\lambda = \lambda_0$, in which case $J \equiv 1$, and when evaluated here the necessary derivatives are:

$$J_\nu = \nabla\cdot\mathbf{v}^\nu, \tag{10a}$$

$$J_{\mu\nu} - J_\mu J_\nu = \nabla\cdot\mathbf{v}^\nu_\mu + \mathbf{v}^\mu\cdot\nabla(\nabla\cdot\mathbf{v}^\nu). \tag{10b}$$

Again, all derivatives on the right-hand side of Eq. (10) are Eulerian.

For any problem of practical interest, it will not be possible to determine $q(\lambda)$ to set up Eq. (9), let alone solve it for the necessary mapping, which is represented by $\mathbf{v}^\nu$. Nevertheless, it can be very useful to accomplish this even approximately — not by relaxing Eq. (9), but by devising approximate Eulerian forms of $p(\mathbf{x}, \lambda)$ and $q(\lambda)$ to use in it. In this manner, we redirect results of approximate theories toward the formulation of mapping fields and, via Eqs. (3), (6) and (10), to new, formal expressions for FE derivatives and the properties they describe.

The choice of the approximation will depend on the system, but a strategy that may prove broadly useful will represent p as a product of N one-body terms of identical form p_1, each depending on the coordinate $\mathbf{x}_i$ of a single molecule i:

$$p(\mathbf{x}, \lambda) = \prod_{\imath=1}^{N} p_1(\mathbf{x}_i, \lambda), \tag{11}$$

Then, it happens that Eq. (9) can be satisfied by requiring that $\mathbf{v}_i^\nu$, p_1, and $q_1(\equiv \int p_1 d\mathbf{x}_i)$ obey the same form as (9):

$$\frac{\partial}{\partial \nu}\left(\frac{p_1(\mathbf{x}_i, \lambda)}{q_1(\lambda)}\right) + \nabla_{\mathbf{x}_i} \cdot \left(\frac{p_1(\mathbf{x}_i, \lambda)}{q_1(\lambda)} \mathbf{v}_i^\nu\right) = 0. \tag{12}$$

It must be emphasized that no matter how severe or poor the approximation used for p and/or q to enable solution of Eq. (9), the expressions derived from them for the FE derivatives given by Eq. (3) will be rigorous. The consequence of a poor approximation to p, q is a less-than-minimal variance of the integrand of Eq. (1). This may affect the precision of calculations based on the treatment, but it will not affect the accuracy. In practice, we find that even simple approximations can produce very good results.

Example applications

We now consider several examples to demonstrate the effectiveness and versatility of this framework for providing high-precision averages from molecular simulation. Comparison of mapped and conventional averaging is made via reference to the *difficulty*,[10] $D \equiv t^{1/2}\sigma$, where t is the amount of CPU time required to produce a result with uncertainty σ. The *square* of the ratio of difficulties for two methods gives the ratio of CPU times needed to produce a result of given precision from them. In all cases described below, mapped and conventional averages were computed in separate simulations, and the individual timings to complete a given run were recorded and used to evaluate the difficulty ratios presented here.

Classical heat capacity of a crystal

Crystalline systems are often well represented by some form of harmonic treatment. Lattice dynamics approximates the pair interactions via a quadratic form, which can be decomposed into independent phonon modes that can be treated via Eq. (11). Such a detailed procedure can be useful when developing mapped averages for the bulk and shear moduli, wherein a quasi-harmonic approach is required to capture the relevant effects. For the present example, we will consider only thermal properties, and the heat capacity in particular, for which an even simpler approximation is suitable. More specifically, we approximate the system as a classical Einstein crystal, in which each molecule vibrates independently about its lattice site. Then we identify $\mathbf{x}_i$ as $\Delta\mathbf{r}_i$, the vector from the fixed lattice site of molecule i to its current position, and $p_1 = \exp\left(-\beta c|\Delta\mathbf{r}_i|^2\right)$, $q_1 = (\pi/\beta c)^{3/2}$, where c is a constant. The mapping parameter is the reciprocal temperature: $\mu = \nu \equiv \beta$. Then, solution of (12) in spherical coordinates with boundary condition $\mathbf{v}_i^\beta = 0$ at $\Delta\mathbf{r}_i = 0$ yields (see SM for details):

$$\mathbf{v}_i^\beta = -\Delta\mathbf{r}_i/2\beta. \tag{13}$$

Physically, this describes a mapping in which each molecule is displaced from its lattice site in proportion to $(\beta_0/\beta)^{1/2}$ for the reciprocal-temperature perturbation $\beta_0 \to \beta$, so the molecules would be moved farther from their lattice sites with a perturbation that increases temperature, and closer with a temperature decrease (see SM).

The isochoric heat capacity is given by $C_V/k_\mathrm{B} = -\beta^2 \mathcal{A}_{\beta\beta}$, and the conventional method to evaluate this is in terms of fluctuations of the energy:

$$C_V/k_\mathrm{B} = \mathrm{Var}\left[u\right]. \tag{14}$$

With the mapping defined by Eq. (13), Jacobian derivatives given by Eq. (10), and FE derivatives given by Eqs. (3b) and (6), the corresponding mapped average is:[2]

$$\begin{aligned} C_V/k_\mathrm{B} = \frac{3(N-1)}{2} - \frac{1}{4}\left\langle \Delta\mathbf{r}\cdot\mathbf{f} + \Delta\mathbf{r}\cdot\boldsymbol{\phi}\cdot\Delta\mathbf{r}\right\rangle \\ + \mathrm{Var}\left[u + \frac{1}{2}\Delta\mathbf{r}\cdot\mathbf{f}\right]. \end{aligned} \tag{15}$$

(remembering that u, $\mathbf{f}$ and $\boldsymbol{\phi}$ each include an implicit multiplication by β). The first term on the right-hand side is the exact expression for a classical harmonic system, hence the averages given by the remaining terms constitute a direct measurement of the anharmonic contributions to the properties. The advantage gained is that any fluctuations associated with harmonic behavior are removed from the averages, allowing a much more precise evaluation of the anharmonic effects.

Crystalline systems were the focus of our original work[2] on mapped averaging, where we first presented it as an approach to calculation of thermodynamic properties. There we proposed a temperature-dependent mapping that was formulated by an appeal to the dependence of the root-mean-square deviation of a harmonic oscillator from its minimum-energy position. We return to this application here to demonstrate that the systematic approach now being proposed does indeed yield the same mapping (Eq. (13)) that we developed previously, a mapping that was shown to be highly effective. We will in addition now apply

this "harmonically-mapped averaging" (HMA) to a more realistic model for a crystalline material (rather than the simple Lennard-Jones (LJ) system examined previously), showing in particular that the approach is effective for non-pairwise-additive potentials.

Specifically, we demonstrate the performance of HMA for calculation of C_V of hcp-Fe. Molecular modeling of iron has attracted much interest in recent years due to its geological significance as the primary component of the Earth's core. The conditions there are too extreme to allow easy study by experiment, and models sufficiently accurate to allow predictive calculations have begun to emerge. We use the model of Belonoshko et al.,[11] which is an embedded-atom model that has been well tested against melting, elastic constants, and speed of sound results obtainable from first-principles calculations.

Isochoric molecular dynamics (MD) simulations using this model were performed of crystalline Fe in the hcp structure with c/a = 1.6. The interaction potential was cut off at separations beyond $r_c = 3.5 \times (v/7.0)^{1/3}$, where v is the molar volume (thus, for v=7, 8, and 9.5 Å^3/atom we have r_c = 3.5, 3.659, and 3.875 Å, respectively), with the same r_c used for both pairwise and embedding potentials. No correction for long-range interactions was applied. An MD time step of 1 fs was employed, with an Andersen thermostat invoked at every 100 MD steps to control the temperature. The simulated system comprised $N = 250$ atoms, arranged on $5 \times 5 \times 5$ hcp cells, with 2 atoms/basis. A total of 2×10^4 MD steps were performed for equilibration, and 1×10^5 MD steps were employed to collect averages, with samples taken every N steps. Uncertainties were computed as one standard deviation of the mean (68% confidence) of 100 independent runs, each conducted in this manner.

The performance of HMA is illustrated in Figures 1 and 2, which present data for three densities as a function of temperature, from $T = 0$ to the melting point for each density. Figure 1 presents C_V for the highest of these densities, as computed using conventional and HMA methods, each using roughly the same amount of computation time. The difference in precision of the results is stunning: the uncertainty in the results from the conventional method overwhelms the variation in the measured value over the temperature range, while

the HMA calculations provide data that allow the variation with T to be clearly discerned. The degree of improvement is clarified in Fig. 2, which describes the relative performance for all three densities in terms of the difficulty ratio. At the melting density—which is where the system is least harmonic and HMA should be least effective—values of C_V of a given precision are obtained 100 times faster with the mapped average, in comparison to conventional averaging; at the lowest temperature shown, conventional averaging would require almost 100,000 times more computational effort to obtain results of the precision seen in the mapped average.

Approximate harmonic behavior arises in a variety of circumstances apart from crystalline systems, and it should be possible to formulate mapped averages to exploit this behavior in those cases as well. For example, path-integral methods are often applied to capture nuclear quantum effects in atomistic systems, wherein each atom is represented by a ring of atoms joined by harmonic bonds. Formulation of a mapped average for these systems, in a manner analogous to that shown for crystals, yields a result that is equivalent to the centroid method, which has been shown to be superior to conventional averaging for these calculations.[12] In another direction, it may be advantageous to apply HMA to glasses, using the inherent structure in lieu of the crystalline lattice to define the atom displacements $\Delta\mathbf{r}$. Such an application will need to deal with significant complications in comparison to crystalline systems, but it may prove worthwhile if effective at pulling out precise values of anharmonic contributions to behavior.

Static dielectric constant

The dielectric constant is a notoriously difficult quantity to compute by molecular simulation, as its behavior is determined by fluctuations involving long-range electrostatic interactions. The specific manner in which the fluctuations determine the static dielectric constant depends on the larger medium in which the sample is embedded.[13] In any case, the key quantity making the connection to molecular phenomena is the variance of the total dipole moment

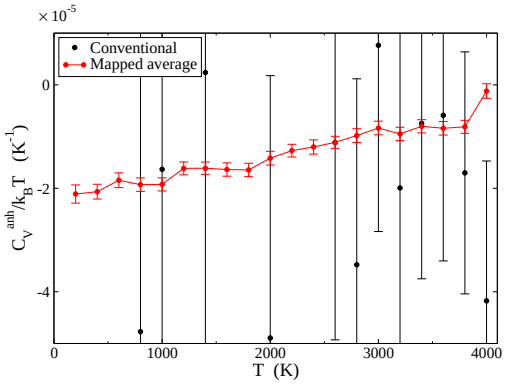

Figure 1: Anharmonic contribution to the heat capacity of hcp-Fe along the 7.0 Å^3/atom isochore, as determined by the conventional average (Eq. (14)) and mapped average (Eq. (15)). Error bars represent 68% confidence limits.

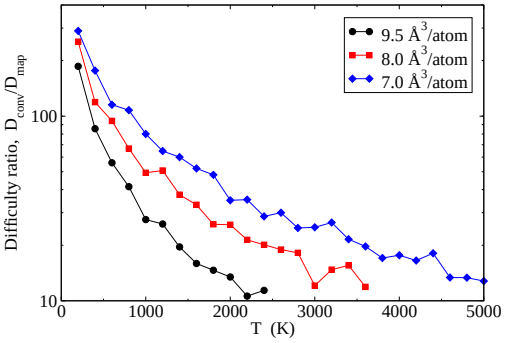

Figure 2: Difficulty ratio[10] for calculation of the anharmonic contribution to the heat capacity of hcp-Fe. The difficulty $D \equiv t^{1/2}\sigma$, where t is the amount of CPU time required to produce a result with uncertainty σ. Hence, the square of the plotted value is the ratio of CPU time required to obtain a result of given precision using conventional versus mapped averaging. The pressures are roughly 275, 150, and 60 GPa for the 7.0, 8.0, and 9.5 Å^3/atom isochores, respectively. The upper-limit temperature of each line is approximately at its melting point.

$\mathbf{M}$, obtained by summing all molecular dipoles: $\mathbf{M} = \sum_i \boldsymbol{\mu}_i$; we consider non-polarizable molecules, so $|\boldsymbol{\mu}_i| \equiv \mu_D = \text{constant}$. The variance relates to the second-order FE derivative with respect to an external electric field $\mathbf{E}$:

$$\mathcal{A}_{\mathbf{EE}} = -\beta^2 \left(\langle \mathbf{MM} \rangle - \langle \mathbf{M} \rangle \langle \mathbf{M} \rangle \right). \tag{16}$$

This is evaluated in the limit $\mathbf{E} \to 0$. Assuming the system has no permanent dipole moment ($\langle \mathbf{M} \rangle = 0$), and is isotropic and homogeneous, then only the diagonal elements of the fluctuation tensor are non-zero. Hence, we may take the derivative of interest as the trace of the tensor or, in 3D,

$$\mathcal{A}_{EE} = -\beta^2 \left(\left\langle M_x^2 \right\rangle + \left\langle M_y^2 \right\rangle + \left\langle M_z^2 \right\rangle \right). \tag{17}$$

A mapped average can be formed for each of these terms, and summed to yield the required derivative. As they differ only in the coordinate direction, we use $-\beta^2 \langle M_z^2 \rangle = \mathcal{A}_{E_z E_z}$ as a prototype for all three. For the mapping coordinate, we define $z_i = \boldsymbol{\mu}_{z,i}/\mu_D$, the z-component of the unit dipole vector of molecule i as oriented in a given configuration; $z_i \in [-1, 1]$. So, in terms of the general formalism introduced above, $\mathbf{x}_i \equiv z_i, \mu = \nu \equiv E_z$.

An approximate $p(\boldsymbol{\mu}, \mathbf{E})$ can be formed by appeal to ideal dipoles interacting with an electric field. Specifically, we identify $p_1(z_i, E_z) = \exp(\beta \mu_D E_z z_i)$, for which $q_1(E_z) = \sinh(\beta \mu_D E_z) / (\beta \mu_D E_z)$. Solution of Eq. (12) with boundary condition $v_i^{E_z} = 0$ for $z = 1$ yields (for $E_z \to 0$):

$$v_i^{E_z} = \frac{1}{2} \beta \mu_D (1 - z_i^2). \tag{18}$$

As $v_i^{E_z} \geq 0$ for all z_i, the mapping moves the dipoles toward better alignment with the field. The remainder of the derivation is provided in the SM. Upon summing over all coordinate

directions, the mapped FE derivative needed for the dielectric constant is:

$$\mathcal{A}_{EE} = -N\beta^2\mu_D^2 + \frac{\beta^4}{4}\left\langle\left|\sum_{i=1}^{N}\boldsymbol{\tau}_i\times\boldsymbol{\mu}_i\right|^2\right\rangle - \frac{\beta^3}{4}\left\langle\sum_{i=1}^{N}\sum_{j=1}^{N}(\nabla_{\boldsymbol{\Omega}_i}\cdot\boldsymbol{\tau}_j)(\boldsymbol{\mu}_i\cdot\boldsymbol{\mu}_j) - \boldsymbol{\mu}_j\cdot\nabla_{\boldsymbol{\Omega}_i}\boldsymbol{\tau}_j\cdot\boldsymbol{\mu}_i\right\rangle. \tag{19}$$

Here, $\nabla_{\boldsymbol{\Omega}_i}$ is the molecule-i orientation gradient, and $\boldsymbol{\tau}_i$ is the torque on i. The first term in (19) is the Clausius-Mossotti-Debye result,[14] obtained by considering fluctuations of independent dipoles. Thus, as found with C_V, we see that the mapped average of the dipole fluctuations separates out the behavior characterized by the approximate theory, leaving a residual in the form of an ensemble average that can be evaluated by molecular simulation.

We applied this mapped average toward calculation of $\mathcal{A}_{EE}$ for the Stockmayer potential, which is the LJ model with a point dipole of magnitude μ_D. NVT Monte Carlo (MC) simulations were performed for $N = 256$ Stockmayer atoms. A total of 1×10^8 MC trials were performed, where each trial is an attempted displacement or rotation (selected with equal probability) of a single atom selected at random. Data from the run were divided into 1000 block averages, which were used to determine the 68% confidence limits (accounting for block correlations if needed[15]).

Results are presented in Fig. 3, where we plot the difficulty ratio for conventional versus mapped averaging, as a function of density for three temperatures and two values of the dipole strength $\mu_{\rm D}$ (note that $\mu_D^2 \approx 10$ for water). The performance improvement with mapped averaging is greatest at conditions where the noninteracting-dipole approximation to p is most appropriate: higher temperature, lower density, and smaller dipole strengths; the behavior correlates well with the dipolar coupling strength, μ^2/T.In the most favorable case presented (at very low density), the reduction of CPU is approaching a factor of 1×10^8; the least-favorable case corresponds to a liquid-like density and temperature, where the advantage of mapped averaging (in this particular formulation, based on non-interacting dipoles) is negligible.

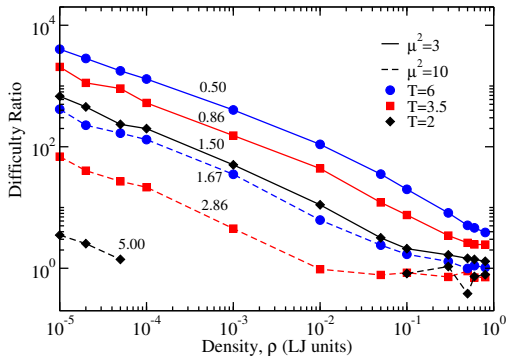

Figure 3: Difficulty ratio (see Fig. 2 for definition) for mapped versus standard averaging of the dielectric-constant derivative $\mathcal{A}_{EE}$, computed for the Stockmayer potential. Data for $\mu^2 = 10$, $T = 2$ are not shown in its vapor-liquid coexistence region. Value next to each line is the dipolar coupling strength, μ^2/T. All units are such that the LJ parameters $\sigma_{\rm LJ}$ and $\epsilon_{\rm LJ}/k_{\rm B}$ are unity (LJ units). For scale, note that the critical temperatures are approximately 1.8 and 3.6 for $\mu^2 = 3$ and 10, resp.[16]

Pressure of a fluid

For our third example, we consider a mapped average for the pressure, based on an approximate model that goes beyond the independent-molecule approach of the previous examples, and includes pair interactions. Let us assume a spherically-symmetric, pairwise-additive potential $u_2(r_{ij})$, where r_{ij} is the separation of the ij pair. We introduce for each pair a scaling parameter s_{ij}, such that the total energy U has the form:

$$U = \sum_{i<j} u_2\left(s_{ij} r_{ij}\right). \tag{20}$$

The derivative of the FE with respect to the scale factor for the ij pair is

$$\mathcal{A}_{s_{ij}} = -\beta \left\langle r_{ij} u_2'(r_{ij}) \right\rangle, \tag{21}$$

where the prime on u_2 indicates a derivative. Comparison to the usual expression obtained for the pressure via differentiation with respect to volume (for a pairwise potential)[17] shows

that

$$\beta P = \rho + \frac{1}{3V} \sum_{i<j} \mathcal{A}_{s_{ij}}. \tag{22}$$

Our aim now is to use mapped averaging for each of the s_{ij} derivatives. We will use a single pair (labeled 1 and 2) as the prototype for all of them, and for notational clarity dispense with the ij subscripts when not needed to distinguish the two molecules of the pair.

For $\mathcal{A}_s$, we identify the relevant coordinates $\mathbf{x}$ to be the positions $(\mathbf{r}_1, \mathbf{r}_2)$ of the two molecules, and λ is identified as the scaling parameter s. We introduce an approximation for $p(\mathbf{x}, \lambda)$ that assumes that the weight associated with the separation of the pair depends on their coordinates and theirs alone, and therefore excludes multibody correlations. Accordingly, we write:

$$p(\mathbf{r}_1, \mathbf{r}_2, s) = p(sr) = e^{-\beta v_2(sr)}, \tag{23}$$

such that p depends only on the pair separation $r = |\mathbf{r}_1 - \mathbf{r}_2|$, via the spherically-symmetric function v_2, introduced here. The normalization q in the volume V can be written

$$\begin{aligned} q(s) &= V^2 + V s^{-3} \int_0^\infty (p(r) - 1) 4\pi r^2 dr \\ &\equiv V \left(V + \hat{q} s^{-3} \right); \end{aligned} \tag{24}$$

this defines $\hat{q}$, which is just a constant (given $p(r)$). Solution of Eq. (9) yields the mapping derivative for r:

$$v^s(r) = -r + 3 \int_0^r d\hat{r} \left(\frac{\hat{r}}{r} \right)^2 \frac{p(\hat{r})}{p(r)}, \tag{25}$$

which has used an approximation based on $V >> \hat{q}$ (see SM).

Expressions for J_s and $\mathcal{U}_s$ are given in the SM, and when combined with Eqs. (3a) and

(22), we find the mapped-average expression for the pressure:

$$\beta P = \rho - \frac{\hat{q}}{2}\rho^2 + \frac{\beta}{3V}\sum_{i<j}\left\{ \langle r\,[v_2'(r) - u_2'(r)]\rangle + \left\langle v^s(r)\left[v_2'(r) + \frac{1}{2}(\mathbf{F}_i - \mathbf{F}_j)\cdot\hat{\mathbf{r}}_{ji}\right]\right\rangle \right\}. \tag{26}$$

Here, $\mathbf{F}_1$ and $\mathbf{F}_2$ are the total forces on molecules 1 and 2, respectively; $\hat{\mathbf{r}}_{12}$ is the unit vector from 1 to 2, such that the mapping causes 2 to move in the direction of $\hat{\mathbf{r}}_{12}$, and 1 opposite to this, each by an amount $v^s/2$.

We demonstrate via application to a LJ fluid, considering performance along three isotherms. NVT MC simulations of 1000 LJ atoms were performed over 1×10^8 MC atom-displacement trials (in addition to 1×10^7 steps of equilibration) at each thermodynamic state condition. Uncertainties were estimated based on the standard deviation of 1000 block averages. The potential was truncated at 4 σ_{LJ} with no correction applied for neglect of longer-range interactions. Correction of the mapped average for truncation is required for the $w_2(r)$ contribution, and this was applied as described in the SM. Values of $v^s(r)$ needed for the mapped average were interpolated from a precomputed 1×10^6-bin grid, in constant steps of $\ln(r+1)$ to allow the use of narrower bins for small r, where variation of v^s with r is greater.

Figure 4 shows that the mapped and conventional averages yield mutually consistent results—even at high-density conditions, where the theory underlying the mapping loses validity—and gives some indication of the relative uncertainties attainable by each for a given amount of computational effort. The difficulty ratio for conventional versus mapped averaging is presented in Fig. 5. The data show that the mapped average is computed much more efficiently at low density, where the approximation used to formulate the mapping is most appropriate. Performance diminishes with increasing density, up to liquid-like values, where the approach as formulated here offers little advantage. At the lowest density examined, the speed-up ranges from 25-fold (at T = 1.0) to 100-fold (at T = 2.0); at the

critical density, speed-up is by about a factor of 2 to 3. There are ways to build on this outcome to make the mapped averaged perform even better, and improve its effectiveness at high densities. First, we can perform a functional optimization of the method with respect to $v_2(r)$, and second, we can consider a formulation based on multibody (beyond pairwise) correlations.

We finish by demonstrating briefly how mapped averages can be applied to yield new formal representations of physical properties, which may then find use toward new theoretical treatments. Equation (26) is the form of the mapped average needed to guide implementation in a molecular simulation. This equation can be alternatively expressed in terms of the interaction of a single pair, using the radial distribution function $g_2(r)$ and the potential of mean force $w_2(r) = -k_{\rm B}T \ln g_2(r)$:

$$\beta P = \rho - \frac{1}{2}\hat{q}\rho^2 + \frac{\beta\rho^2}{6}\int_0^\infty dr 4\pi r^2 g_2(r) \times \left[r\left(v_2'(r) - u_2'(r)\right) + v^s(r)\left(v_2'(r) - w_2'(r)\right)\right]. \tag{27}$$

This expression is a generalized form of the familiar "pressure equation",[17] which is used in the statistical mechanics of fluids to connect $g_2(r)$ to the pressure; the conventional formula is recovered for $v_2(r) \equiv 0$. The generalization written here can be useful for deriving corrections to the mapped-average pressure due to long-range intermolecular interactions (see SM). It is also a convenient form for constructing a variational treatment, or alternatively an optimization, in which $v_2(r)$ is selected to minimize the variance in the mapped-average pressure when computed by simulation.

For the case where the weighting potential is given by the pair potential itself: $v_2(r) \equiv$

$u_2(r)$, then $\hat{q} = -2B_2$, where B_2 is the second virial coefficient, and we have:

$$\begin{aligned}\beta P =& \rho + B_2\rho^2 \\ &+ \frac{\beta\rho^2}{6}\int_0^\infty dr 4\pi r^2 g_2(r) v^s(r)\left(u_2'(r) - w_2'(r)\right).\end{aligned} \tag{28}$$

This formula is interesting because it shows how this mapped average builds upon the second-order virial equation of state. As with the previous examples, analytic terms corresponding to the approximate model appear as separate contributions, leaving residuals in the form of new ensemble averages. The connection to the second-order virial series suggests that we could build upon a higher-order virial equation by starting with three- or higher-body interactions in Eq. (23). Such an approach may run into difficulty when attempting to solve Eq. (9); nevertheless, it could be worthy of consideration.

Mapped averages that account for pair interactions can be developed for other properties as well. As with all implementations of mapped averaging, the procedure has no effect on sampling, and consequently many mapped averages can be evaluated at once during a single simulation.

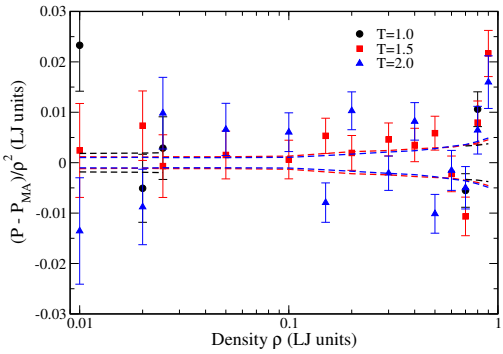

Figure 4: Difference in pressure as computed by conventional versus mapped averaging (MA). Data are for the LJ model for three temperatures: T = 1.0 (black, circles), 1.5 (red, squares), and 2.0 (blue, triangles). Points are the computed differences, with error bars indicating the uncertainty in the conventional average. Dashed lines on either side of the x-axis indicate the uncertainties in the mapped averages. The $T = 1.0$ point at the highest density (0.9) is off the top of the scale of the figure. All quantities are given in units such that the LJ parameters σ_{LJ} and $\epsilon_{\mathrm{LJ}}/k_{\mathrm{B}}$ are unity.

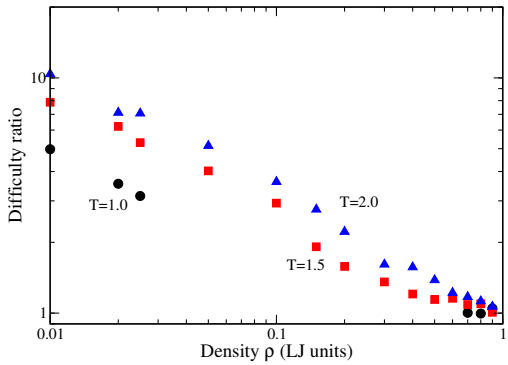

Figure 5: Difficulty ratio (see Fig. 2 for definition) for mapped versus standard averaging of the pressure of a LJ fluid. Data are shown for temperatures of 1.0 (black circles), 1.5 (red squares), and 2.0 (blue triangles). All quantities are given in units such that the LJ parameters σ_{LJ} and $\epsilon_{\mathrm{LJ}}/k_{\mathrm{B}}$ are unity.

Conclusions

We have proposed a general means to reformulate ensemble averages, and have shown that they can be used to vastly improve the precision of thermodynamic properties measured by molecular simulation. The effectiveness of a mapped average depends on how well the approximation to $p(\mathbf{x}, \lambda)$ is formulated, which can depend in part on the thermodynamic state. Our experience finds that it is not difficult to devise an approximate $p(\mathbf{x}, \lambda)$ such that over some range of conditions the computational effort is reduced by one to two orders of magnitude for first- and second-derivative properties, respectively. We find that in favorable conditions the computational savings can be many orders of magnitude greater than this, while in the worst cases the reformulated averages simply provide no benefit, and do not seem to be detrimental. Although our emphasis has been on computational advantage, we should re-emphasize that mapped averaging presents a promising route to develop statistical mechanical theories as well.

The framework we propose leads methodically to the reformulated averages, but it does this while taking for granted a starting point that may require some creativity or physical insight to reach. First, it is necessary to formulate the property as a FE derivative. If the appropriate derivative is not already established, it should be possible to construct it

by defining a suitable generalized partition function, perhaps contriving an artificial field parameter and a corresponding FE derivative that yields the property of interest. Our development for the mapped-average pressure given in the third example above—where we introduced a parameter that scales the distance between a single pair of atoms—demonstrates how this may be done.

Second, one then needs an approximate model for how the microscopic structure depends on the field parameter, as encapsulated by $p(\mathbf{x}, \lambda)$ introduced with Eq. (8). The mapped average will be more effective to the degree that the approximate $p(\mathbf{x}, \lambda)$ is accurate. We have shown that quite simple ideal-particle approximations can be useful, but one might expect greater effectiveness if multibody correlations could be accommodated. This can be done, but only to the extent that it remains possible to evaluate the approximate partition function $q(\lambda)$ and solve Eq. (9) for the mapping $\mathbf{v}^\nu$. This problem is not terribly unlike those for which statistical mechanical theories have already been developed to solve, and the prospect for synthesizing those theoretical advances with molecular simulation in this manner is an appealing capability presented by the mapped averaging framework.

Our example for the pressure here led to a formula built on the second-order virial series. This result contrasts with the pressure formula given in our original application to crystals, which yielded instead an expression in reference to a quasi-harmonic approximation.[2] So one sees that the choices just summarized present different ways to arrive at mapped averages for the same property, with respective advantages that depend on the system and the state.

Finally, the prescription for the mapping given by Eq. (9) is just one equation for the N variables represented by $\mathbf{v}^\nu$. Additional constraints must be imposed to give a full specification for the mapping. We expect that in most cases these constraints will appeal to symmetry. Thus, for $p(\mathbf{x}, \lambda)$ approximated as a product of 1-body terms, as in Eq. (11), it is sensible to formulate identical equations for each component of $\mathbf{v}^\nu$, as in Eq. (12). Extension to treat multibody approximations to $p(\mathbf{x}, \lambda)$ in a likewise general manner will require further exploration.

Acknowledgements

This work was supported by the U.S. National Science Foundation Grant CHE-1464581.

Literature Cited

(1) Hansen, J.-P.; McDonald, I. *Theory of Simple Liquids*, 3rd ed.; Academic Press, London, 2006.

(2) Moustafa, S. G.; Schultz, A. J.; Kofke, D. A. Very fast averaging of thermal properties of crystals by molecular simulation. *Phys. Rev. E* **2015**, *92*, 043303.

(3) Jarzynski, C. Targeted free energy perturbation. *Phys. Rev. E* **2002**, *65*, 046122.

(4) Kofke, D. A. Free energy methods in molecular simulation. *Fluid Phase Equil.* **2005**, *228-229*, 41–48.

(5) Tan, T. B.; Schultz, A. J.; Kofke, D. A. Efficient calculation of temperature dependence of solid-phase free energies by overlap sampling coupled with harmonically targeted perturbation. *J. Chem. Phys.* **2010**, *133*, 134104.

(6) Tan, T. B.; Schultz, A. J.; Kofke, D. A. Efficient calculation of α- and β-nitrogen free energies and coexistence conditions via overlap sampling with targeted perturbation. *J. Chem. Phys.* **2011**, *135*, 044125.

(7) Paliwal, H.; Shirts, M. R. Multistate reweighting and configuration mapping together accelerate the efficiency of thermodynamic calculations as a function of molecular geometry by orders of magnitude. *J. Chem. Phys.* **2013**, *138*, 154108.

(8) Whitaker, S. *Introduction to Fluid Mechanics*; Krieger Publishing: Florida, 1992; note that Eulerian and Lagrangian specifications are referred to in this text as "spatial coordinates" and "material coordinates," respectively.

(9) See Supplemental Material at [URL will be inserted by publisher] for details of derivations of mapped-average formulas.

(10) Schultz, A. J.; Kofke, D. A. Quantifying Computational Effort Required for Stochastic Averages. *J. Chem. Theory Comput.* **2014**, *10*, 5229–5234.

(11) Belonoshko, A.; Ahuja, R.; Johansson, B. Quasi-Ab initio molecular dynamic study of Fe melting. *Phys. Rev. Lett.* **2000**, *84*, 3638–3641.

(12) Glaesemann, K. R.; Fried, L. E. An improved thermodynamic energy estimator for path integral simulations. *J. Chem. Phys.* **2002**, *116*, 5951.

(13) Neumann, M. Dipole moment fluctuation formulas in computer simulations of polar systems. *Mol. Phys.* **1983**, *50*, 841–858.

(14) Jepsen, D. W. Calculation of the Dielectric Constant of a Fluid by Cluster Expansion Methods. *J. Chem. Phys.* **1966**, *44*, 774.

(15) Kolafa, J. Autocorrelations and subseries averages in Monte Carlo Simulations. *Mol. Phys.* **1986**, *59*, 1035–1042.

(16) Stevens, M.; Grest, G. Phase coexistence of a Stockmayer fluid in an applied field. *Phys. Rev. E* **1995**, *51*, 5976–5984.

(17) McQuarrie, D. *Statistical Mechanics*; Harper & Row: New York, 1976.

Supplemental Material is provided with detailed derivations of the field equations and mapped-average formulas for the three examples presented in this work. This information is available free of charge via the Internet at http://pubs.acs.org.

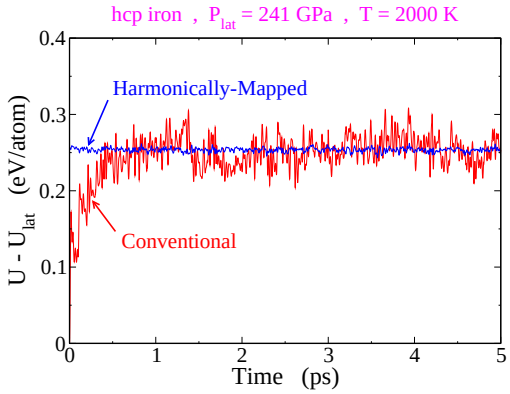

Table of Contents figure