JCTC Journal of Chemical Theory and Computation

Article

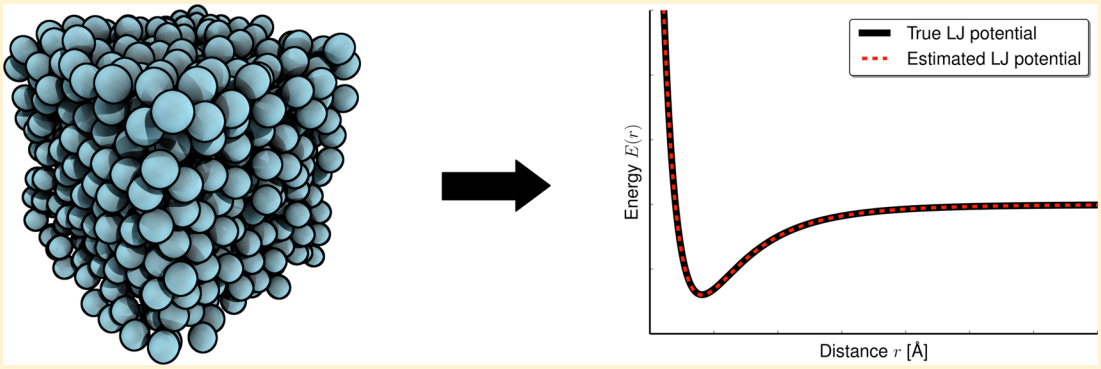
pubs.acs.org/JCTC

Estimation of Interaction Potentials through the Configurational Temperature Formalism

Martin Mechelke[‡,†] and Michael Habeck*[,†]

[†]Institute for Mathematical Stochastics, Georg August University Göttingen, 37077 Göttingen, Germany

[‡]Department of Protein Evolution, Max-Planck-Institute for Developmental Biology, 72076 Tübingen, Germany



ABSTRACT: Molecular interaction potentials are difficult to measure experimentally and hard to compute from first principles, especially for large systems such as proteins. It is therefore desirable to estimate the potential energy that underlies a thermodynamic ensemble from simulated or experimentally determined configurations. This inverse problem of statistical mechanics is challenging because the various potential energy terms can exhibit subtle indirect and correlated effects on the resulting ensemble. A direct approach would try to adapt the force field parameters such that the given configurations are highly probable in the resulting ensemble. But this would require a full simulation of the system whenever a parameter changes. We introduce an extension of the configurational temperature formalism that allows us to circumvent these difficulties and efficiently estimate interaction potentials from molecular configurations. We illustrate the approach for various systems including fluids and a coarse-grained protein model.

1. INTRODUCTION

Computational methods such as molecular dynamics and Monte Carlo simulation provide insights into biomolecular systems at a level of detail that is often unmatched by experimental techniques.[1] But the accuracy and reliability of computational results depends critically on the potential energy function. Small inaccuracies in the force field can bias simulations toward incorrect conformations.[2,3]

The design of potential functions is a complicated and time-consuming task because typically a large number of parameters needs to be adjusted. Although recent advances in computation hardware have made it possible to compare force fields in terms of their ability to fold a protein chain to its native state,[4,5] the problem still persists. Some of the difficulties are related to the tight coupling of the individual force field terms.[6,7]

Typically, the parameters of a potential function are chosen so as to reproduce experimental data such as, for example, vibrational spectra and vibrational frequencies of small representative systems.[8] When experimental data are not available, quantum mechanics calculations are used instead.[2,9] But the parametrization based on small systems may not be universally transferable.[10] Statistical potentials have been proposed as an alternative approach to directly learn energies from experimentally determined structures of biomolecules such as proteins.[11−13] But the resulting energy functions are at best potentials of mean force and generally differ from the potential energy function.

Here we aim to extract force field parameters directly from configurations drawn from the canonical ensemble:

$$p(x|\beta) = \frac{1}{Z(\beta)} \exp\{-\beta E(x)\} \tag{1}$$

where x is the configuration of the system, $E(x)$ is the system's potential energy, β is the inverse temperature, and $Z(\beta)$ is the partition function. The problem of estimating $E(x)$ from a set of configurations is the inverse problem of statistical mechanics. Because many experimental structures are available, it is desirable to devise a method that solves this inverse problem and thereby learn the underlying energy function.

The estimation of molecular potentials from configurations has a rich history dating back to work on liquids starting with the Ornstein−Zernike equations[14] and the Weeks−Chandler−Andersen perturbation theory.[15] Methods to estimate potential

Received: July 4, 2013
Published: November 15, 2013

   dx.doi.org/10.1021/ct400580p | J. Chem. Theory Comput. 2013, 9, 5685−5692

functions of complex systems such as biomolecules are often based on correlation and distribution functions. The most straightforward approach is direct Boltzmann inversion,[16] which approximates the potential by the negative logarithm of the radial distribution function. Direct Boltzmann inversion is applicable to uncoupled potentials but struggles with strongly coupled interactions. More sophisticated techniques that can deal with coupled interactions encompass iterative Boltzmann inversion,[17,18] reverse Monte Carlo[19] techniques, and inverse Monte Carlo methods.[20,21] These methods adjust the force field iteratively to explain the observed distribution functions of the molecular system but differ in how the correction of the force field parameters is computed. In each iteration, these methods require a full simulation of the system and the calculation of ensemble averages and distribution functions. Both methods are special cases of the relative entropy formalism[22] as they minimize an entropic divergence between the true and estimated potential function. Reweighting techniques can be used to improve the efficiency of iterative methods[23,24] by making use of the information from all simulations, not only the last simulation.

Methods to infer potential functions from distribution functions are also vital to coarse graining, where complex systems are represented by a reduced number of degrees of freedom. Besides correlation functions, coarse graining also uses atomistic potential functions as reference. Force matching[25] and multiscale coarse graining[26,27] infer potentials that approximate the forces of the reference potential. The generalized Yvon−Born−Green (gYBG) method developed by Mullinax and Noid[28] provides a powerful extension to multiscale coarse graining. It builds upon the works of Yvon[29] and Born and Green[30] for monatomic fluids. The gYBG method calculates the potential energy directly from structural correlation functions and similar to multiscale coarse graining does not involve repeated ensemble simulations.

Here, we use the configurational temperature formalism[31,32] to estimate the parameters of the potential energy function. We treat the force constants and other force field parameters as multiple temperatures and derive a system of linear equations to estimate them from a set of molecular configurations. We study different options for setting up the configurational temperature equations and discuss means to regularize the optimization problem. We demonstrate the efficacy and accuracy of the approach for simple systems such as Lennard-Jones fluids as well as a coarse-grained protein model.

2. METHODS

2.1. Configurational Temperature. The configurational temperature has been introduced by Rugh[31] to estimate the temperature of a system without knowledge of the kinetic energy. Jepps et al.[32] generalized Rugh's equations to calculate the temperature of an equilibrium system from configurations only. The fundamental identity derived by Jepps et al. is

$$\langle B(\Gamma)\cdot\nabla_\Gamma h(\Gamma)\rangle_\Gamma = \langle \nabla_\Gamma\cdot B(\Gamma)\rangle_\Gamma \tag{2}$$

where Γ and ∇_Γ denote the vector of phase space variables and their derivatives and h is the reduced Hamiltonian (i.e., $h(\Gamma) = -\log\, p(\Gamma) + \text{const}$, where p is the statistical ensemble). Equation 2 holds for many ensembles including the micro- and macrocanonical ensembles and for a variety of phase space fields $B(\Gamma)$. The conditions on $B(\Gamma)$ are that both ensemble averages in eq 2 exist and are finite, $0 < |\langle \nabla_\Gamma h(\Gamma)\cdot B(\Gamma)\rangle_\Gamma| < \infty$, $0 < |\langle \nabla_\Gamma B(\Gamma)\rangle_\Gamma| < \infty$, and that $\langle \nabla_\Gamma h(\Gamma) B(\Gamma)\rangle_\Gamma$ grows more slowly than e^N in the thermodynamic limit, where N is the number of degrees of freedom.[32]

Because we typically lack knowledge of the momenta, we are interested in applying (2) to configurations only. For the special choice that B is zero in all momentum directions and independent of the momenta, the averages in eq 2 reduce to configuration space averages:

$$\langle B\cdot\nabla v\rangle = \langle \nabla\cdot B\rangle \tag{3}$$

where now $v(x)$ is the reduced potential energy, $B(x)$ is a configuration dependent vector field in configuration space, ∇ is the gradient with respect to the configurational degrees of freedom, and $\nabla\cdot B$ is the divergence of the vector field. In eq 3 and in the following, the average $\langle\cdot\rangle$ denotes the configuration average:

$$\langle f\rangle = \frac{1}{Z}\int f(x) e^{-v(x)}\,dx \quad \text{where} \quad Z = \int e^{-v(x)}\,dx$$

For the special case $v(x) = \beta E(x)$ with $\beta = (k_B T)^{-1}$ eq 3 gives us a general expression for the configurational temperature:[32]

$$\beta = \frac{\langle \nabla\cdot \mathbf{B}\rangle}{\langle \mathbf{B}\cdot\nabla E\rangle} \tag{4}$$

One definition of the multiple expressions for the configurational temperature is obtained by choosing $\mathbf{B}(x) = \nabla E(x)$, in which case we have

$$\frac{1}{k_B T_{\text{config}}} = \frac{\langle \nabla\cdot\nabla E\rangle}{\langle \nabla E\cdot\nabla E\rangle}$$

This expression has been derived by Jepps et al.[32]

2.2. Systems and Simulations. *2.2.1. Liquid Argon.* We demonstrate our approach for two systems. The first is a monatomic Lennard-Jones fluid simulated with the Molecular Modelling Toolkit (MMTK).[33] This simulation is based on an example provided together with MMTK. We used an *NVT* ensemble of liquid argon comprising 864 atoms in a box of edge length 34.8 Å and a step size of 10 fs. In the initial phase, the system was equilibrated at a temperature of 86 K by rescaling the velocities for 2000 time steps. Next we simulated the system for 2000 more steps. Out of the simulated configurations we retained every 100th to minimize the autocorrelation of the samples.

2.2.2. Minimalist Protein Model. The second system is a minimalist protein force field developed by Honeycutt and Thirumalai[34] and later refined by Sorenson and Head-Gordon.[35] This model represents the polypeptide chain as beads on a string. Each bead corresponds to an amino acid that is hydrophobic (B), hydrophilic (L), or neutral (N). An attractive Lennard-Jones potential between the hydrophobic beads drives the compaction of the polypeptide chain, whereas the nonbonded interactions between all other beads are purely repulsive. The chain geometry and secondary structure are stabilized by a harmonic potential on the bond lengths and angles as well as through a dihedral angle potential. The potential energy of a conformation x is given by

$$E(x) = \sum_{r\in\text{bonds}} k_{\text{bonds}}(r - r_0)^2 + \sum_{\theta\in\text{bond angles}} k_{\text{angles}}(\theta - \theta_0)^2 + \sum_{\varphi\in\text{dihedrals}} [a(1 + \cos(\varphi)) + b(1 + \cos(3\varphi))] + \sum_{\text{nonbonded}\, r_{ij}} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right]$$

We chose a barrel forming protein with sequence $B_9N_3(LB)_4N_3B_9N_3(LB)_5$. Simulations were carried out using GROMACS 4.0[36,37] with stochastic dynamics as a thermostat. The step size was 0.01 ns and every 500th sampled conformation was retained for further analysis. The system was simulated at reduced units with all particles having unity mass and temperature of $T = 0.28$. As initial configuration for the simulation, we used an extended conformation. After an equilibration phase of 50 000 steps, we kept every 500th conformation for further analysis.

2.2.3. Assessment of the Potential. We assess the quality of the estimated potentials by comparing the derivatives of the estimated energy function $E_{\text{est}}(y)$ to that of the true energy function $E_{\text{true}}(y)$ where y is a geometric feature such as a distance or an angle. The mean squared error (MSE) of the resulting forces provides a shift invariant similarity measure that is independent of the choice of basis functions and is given by

$$\text{MSE}[E'_{\text{est}}] = \int [E'_{\text{true}}(y) - E'_{\text{est}}(y)]^2\, \text{d}y \approx \frac{1}{n}\sum_{i}^{n} [E'_{\text{true}}(y_i) - E'_{\text{est}}(y_i)]^2 \tag{5}$$

where y_i are n equidistant support points used to approximate the MSE integral. Another figure of merit for distance dependent energies is the radial distribution function.

3. RESULTS AND DISCUSSION

3.1. Estimation of Interaction Potentials. We assume that configurations were generated from the Boltzmann distribution:

$$p(x|\lambda) = \frac{1}{Z(\lambda)} \exp\{-v(f(x);\lambda) - f_0(x)\} \qquad Z(\lambda) = \int \exp\{-v(f(x);\lambda) - f_0(x)\}\, \text{d}x \tag{6}$$

where the reduced interaction potential v depends on the configuration through a set of K features f_k, $k = 1, ..., K$ (e.g., angles or distances) and λ is the vector of parameters specifying the shape of the potential; $Z(\lambda)$ is the normalization constant or partition function. $f_0(x)$ is an arbitrary reference energy that we assume to be known; for example, we could let $f_0(x)$ be constant, or $f_0(x)$ could be a reference distribution that already accounts for some interactions. In the following, we will assume that the potential function depends linearly on the parameters: $v(f(x);\lambda) = \lambda^T f(x) = \sum_{k=1}^{K}\lambda_k f_k(x)$. That is, we expand the potential energy into different features, which in general will be correlated. This choice is not restrictive but allows us to treat all of the standard force field terms with our formalism. The canonical ensemble (1) is the simplest version involving a single feature that is identical to a known force field.

We use the configurational temperature formalism to estimate the parameters λ_k. Equation 3 becomes the identity

$$\langle B\cdot\nabla(\lambda^T f + f_0)\rangle_\lambda = \langle\nabla\cdot B\rangle_\lambda$$

where $\langle\cdot\rangle_\lambda$ is an average over the ensemble (6) and the subscript indicates the dependence on λ. The choice of the configuration space field $B(x)$ is arbitrary except for some weak conditions.[32] We choose a series of vector fields B_k, one for each expansion coefficient λ_k, to obtain

$$\sum_{l=1}^{K} \lambda_l\langle B_k\cdot\nabla(f_l - f_0)\rangle_\lambda = \langle\nabla\cdot B_k\rangle_\lambda \qquad k = 1, ..., K \tag{7}$$

Each of the B_k has to fulfill the conditions $0 < |\langle B_k\cdot v\rangle_\lambda| < \infty$ and $0 < |\langle\nabla\cdot B_k\rangle_\lambda| < \infty$, and $\langle B_k\cdot v\rangle_\lambda$ must not grow faster than e^N in the thermodynamic limit. This system of linear equations is a multitemperature generalization of the configurational temperature relation 4 and determines the parameters λ by solving

$$\sum_{l=1}^{K} A_{kl}\lambda_l = b_k \quad \text{with} \qquad A_{kl} = \langle B_k\cdot\nabla f_l\rangle_\lambda \qquad b_k = \langle\nabla\cdot B_k - B_k\cdot\nabla f_0\rangle_\lambda \tag{8}$$

If we choose B_k along the direction of the gradient of the kth feature (see eq 9 below), the matrix A_{kl} measures how strongly features f_k and f_l are correlated: If both features result in similar forces, their correlation will be high and A_{kl} will be large, whereas features whose forces are orthogonal will lead to small A_{kl}. Therefore, the off-diagonal elements of A_{kl} account for double counting of different energy terms. By the same argument, in case we have a nonzero reference potential $f_0(x)$, the right-hand side of eq 8 accounts for contributions that are already explained by the reference force field. We can solve the system of linear eq 8 by standard methods such as LU or singular value decomposition. Here we consider three different choices for B_k:

$$B_k(x) = \nabla f_k(x)/\|\nabla f_k(x)\|^n \qquad n = 0, 1, 2 \tag{9}$$

where $\|\cdot\|$ indicates the Euclidean norm of configuration space vectors. To evaluate the configurational temperature equation (8), we need to compute the divergence of B_k:

$$\nabla\cdot B_k(x) = \frac{\Delta f_k(x)}{\|f_k(x)\|^n} - n\frac{\sum_{i,j}(\partial_i f_k(x))(\partial_i\partial_j f_k(x))(\partial_j f_k(x))}{\|\nabla f_k(x)\|^{n+2}} \tag{10}$$

where indices i, j enumerate all configurational degrees of freedom and ∂_i indicates a partial derivative along the ith coordinate (i.e., $\partial_i\partial_j f_k(x)$ is the Hessian matrix of the kth feature); Δ is the Laplace operator. To estimate the force field parameters λ from T sampled configurations $x^{(t)}$, we replace the ensemble averages in eq 8 with sample averages:

$$A_{kl} \approx \frac{1}{T}\sum_{t=1}^{T} B_k(x^{(t)})\cdot\nabla f_l(x^{(t)}) \qquad b_k \approx \frac{1}{T}\sum_{t=1}^{T} \nabla\cdot B_k(x^{(t)}) - B_k(x^{(t)})\cdot\nabla f_0(x^{(t)}) \tag{11}$$

Depending on the choice of the vector fields B_k, we can draw connections to related methods in statistical physics and machine learning. If we choose $B_k = \nabla f_k$, the resulting least-squares problem is identical to the score matching procedure

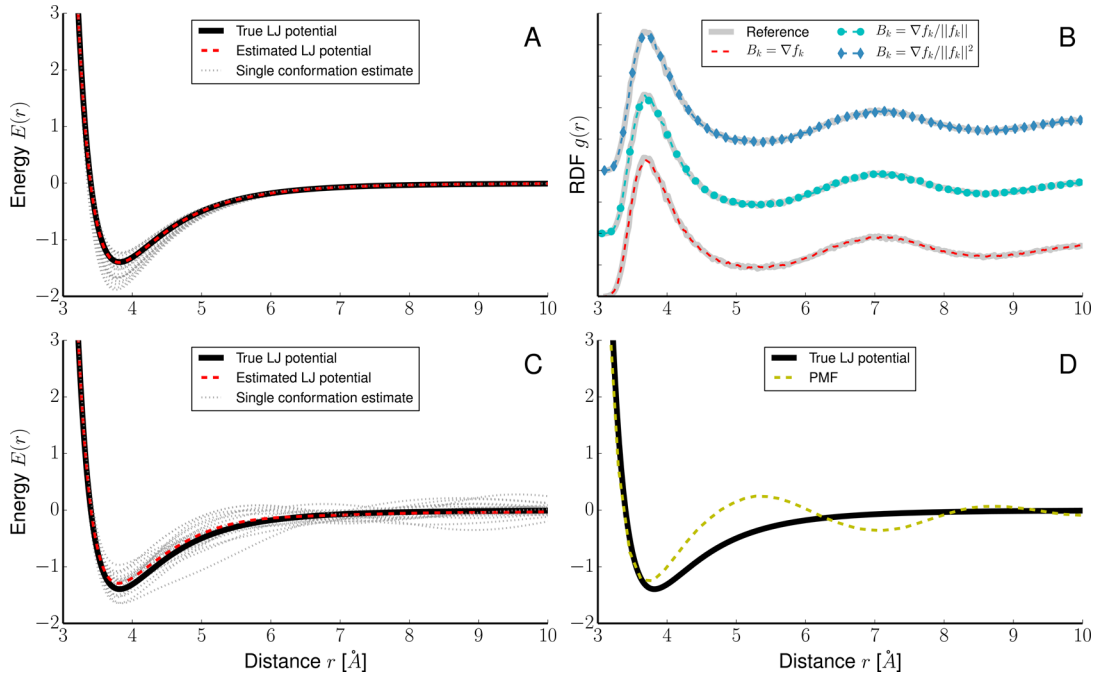


Figure 1. Lennard-Jones potential estimated from a simulation of liquid argon. (A) Configurational temperature estimates using the correct features of the Lennard-Jones potential. (B) Radial distribution function of simulations with estimated parameters compared to the original simulation. The radial distribution functions are shifted for clarity. (C) Configurational temperature estimate based on a Laguerre representation of the potential using the 20 first Laguerre polynomials. (D) Potential of mean force obtained by the inverse Boltzmann law applied to the radial distribution function. The black solid line indicates the true potential used in the simulation of the argon fluid. In panels A and B the gray dashed curves indicate the potentials recovered from single configurations, and the red solid line is the result based on all 20 structures.

introduced by Hyvärinen.[38] Score matching aims to estimate continuous probability density functions with intractable normalization constants or partition functions by minimizing the Fisher information. For the same choice of B_k, the configurational temperature eq 8 can be derived from the generalized Yvon−Born−Green equations (gYBG).[28,39] It is also interesting to compare our approach with force matching.[25] In force matching (FM), the true forces of every configuration $y^{(t)} = \nabla v(x^{(t)})$ are assumed to be known. The true forces are approximated with the forces resulting from our linear ansatz $\nabla(\lambda^T f(x^{(t)})) = F^{(t)}\lambda$ where the gradients of every feature are the columns of the force matrix $F^{(t)} = [\nabla f_1(x)^{(t)}, ..., \nabla f_K(x)^{(t)}]$. FM minimizes the residual

$$\chi_{\mathrm{FM}}{}^2(\lambda) = \frac{1}{2T}\sum_t \|\nabla v(x^{(t)}) - \nabla\lambda^T f(x^{(t)})\|^2 = \frac{1}{2T}\sum_t \|y^{(t)} - F^{(t)}\lambda\|^2 \tag{12}$$

Setting the gradient of $\chi_{\mathrm{FM}}{}^2$ to zero results in a system of linear equations that determines the optimal λ:

$$\tilde{A}\lambda = \tilde{b} \quad \text{with} \quad \tilde{A} = \frac{1}{T}\sum_t [F^{(t)}]^T F^{(t)} \qquad \tilde{b} = \frac{1}{T}\sum_t [F^{(t)}]^T y^{(t)} \tag{13}$$

Comparison with the system of equations derived with the configurational temperature approach (eq 8) shows that for the choice $B_k = \nabla f_k$ matrices A and $\tilde{A}$ are identical. The difference is in the right-hand side of eqs 8 and 13: In FM, we need to know the true forces to evaluate $\tilde{b}$. This is possible for coarse graining of atomic potential functions but not in the application we are interested in here. Therefore, our approach could be viewed as a generalization of FM to situations where the true forces that we aim to approximate are unknown.

3.2. Lennard-Jones Fluid. As a first application of our approach, we consider a monatomic Lennard-Jones fluid. We simulated the *NVT* liquid argon system at 86 K just below the boiling temperature with periodic boundary conditions. The interactions in this system are described by a distance dependent Lennard-Jones potential:

$$E_{\text{L-J}}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{14}$$

where σ determines the location of the minimum and ϵ its depth. The total potential energy of the system is $E(x) = \Sigma_{i<j}E_{\text{L-J}}(\|x_i - x_j\|)$ where x_i is the three-dimensional position of the ith atom. Our goal is to estimate the parameters of the Lennard-Jones potential, ϵ and σ. If we use r^{-6} and r^{-12} as basis functions, we have the following features:

$$f_1(x) = \sum_{i<j}\|x_i - x_j\|^{-6} \quad \text{and} \quad f_2(x) = \sum_{i<j}\|x_i - x_j\|^{-12} \tag{15}$$

with the corresponding parameters

$$\lambda_1 = -4\beta\epsilon\sigma^6 \quad \text{and} \quad \lambda_2 = 4\beta\epsilon\sigma^{12} \tag{16}$$

By estimating λ_1 and λ_2, we obtain ϵ and σ assuming that we know the temperature β of the system.

We estimated the parameters ϵ and σ from 20 configurations using the configurational temperature equations (8) and (11). Figure 1A shows the estimated potential obtained from the single configurations and from a joint analysis of all configurations. The latter recovers the true potential very

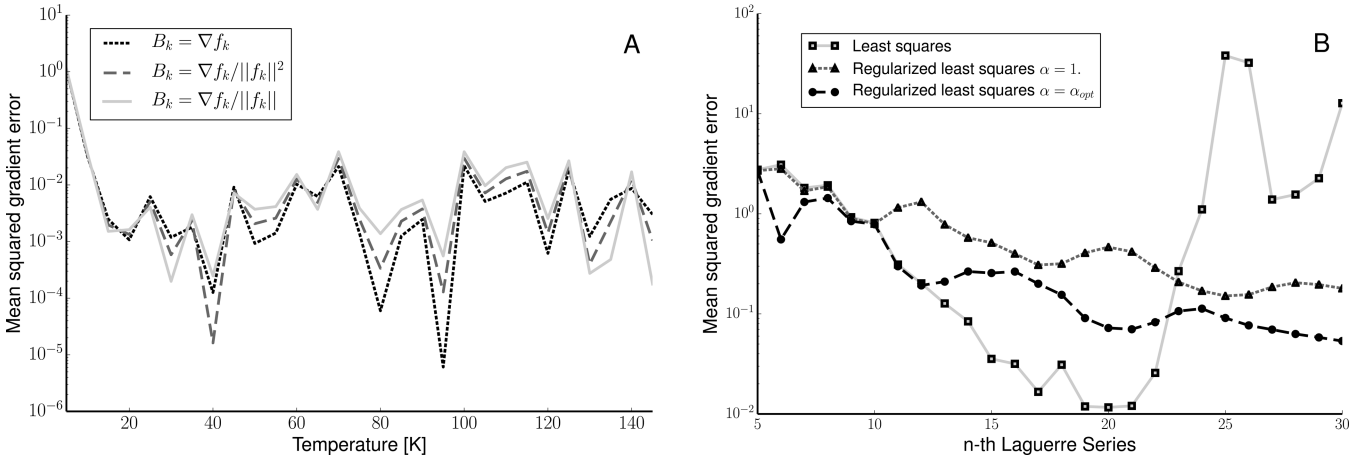


Figure 2. Prediction accuracy dependence on system temperature. The left panel shows the MSE between the derivatives of the true and approximate potential function in the range 3−10 Å for r^{-6} and r^{-12} basis functions. The right panel shows the influence of a regularization term on the accuracy of the Laguerre basis.

accurately. The different choices for the vector fields B_k (eq 9) result in very similar estimates. However, for reasons of clarity we only show the results for $B_k = \nabla f_k$. Already a single configuration is sufficient to obtain a rough estimate of λ_1 and λ_2 (Figure 1A). As an additional validation of our method, we ran simulations based on the estimated Lennard-Jones parameters and compared the radial distribution functions (RDF) of the original and new simulations. The results are detailed in Figure 1B. For each of the three choices of the vector fields B_k (9), the reference RDF is reproduced with high accuracy.

So far we used the same basis functions in the simulation of the system and for the estimation of the energy function. But often the functional form of the potential is unknown, and we have to choose a set of basis functions that do not necessarily coincide with the true underlying energy function. To illustrate this situation, we also expanded the Lennard-Jones potential into a weighted sum of Laguerre polynomials. Similar to the Fourier basis, Laguerre polynomials form an orthonormal system. Because their support is the non-negative axis, they are suited to represent distance dependent energy terms. The Laguerre polynomials $L_k(r)$ are defined as

$$L_k(r) = \frac{e^r}{k!}\frac{d^k}{dr^k}(e^{-r}r^k)$$

We choose $f_k(r) = e^{-r/2}L_k(r)$ because then the features will be orthonormal $\int_0^\infty f_k(r)\, f_l(r)\, dr = \delta_{kl}$.

We used features based on the $K = 20$ first Laguerre polynomials to analyze the argon simulation. The full interaction energy in this representation is given by

$$E_{\text{Laguerre}}(x) = \sum_{k=1}^{K} \lambda_k \{ \sum_{i<j} L_k(\|x_i - x_j\|) e^{-\|x_i - x_j\|/2} \} \quad (17)$$

The estimated potentials are shown in Figure 1C and reconstruct the Lennard-Jones potential accurately. However, the potentials estimated from the single configurations show higher variability. This is caused by the higher flexibility of the basis functions and the larger number of free parameters that we have to estimate from the same simulation.

Figure 1D contrasts our results with the potential of mean force (PMF) approach, which is often used in biomolecular simulations to learn "knowledge-based potentials".[12,40] PMFs are often used as potential functions, although they technically represent the mean energy of changing a single particle in a multiparticle system.[41] As the PMF is computed through inversion of the radial distribution function, it exhibits subordinate minima, which are the result of second shell interactions.

3.3. Impact of Simulation Temperature. Next, we investigate how the simulation temperature influences the accuracy of the estimated potential. Again, we use monatomic argon as our model system and run MD simulations at temperatures ranging from 5 to 140 K. We used 20 conformations per temperature to estimate the potential energy. Figure 2A depicts the MSE between the derivatives of the true and estimated potential for the range of simulation temperatures and different choices of B_k (eq 9) evaluated at 1000 linearly spaces points ranging from 3 to 10 Å. The choice of B_k does not have a strong impact on the estimation accuracy. As suggested by Figures 2A and 1B, all three vector fields perform equally well. We therefore use $B_k = \nabla f_k$ in the remainder of paper because it is easier to implement and numerically more stable.

There is no significant dependence of the magnitude of the error on the temperature (Figure 2A). Only for very low simulation temperatures does the accuracy of the reconstructed potential decrease, which is to be expected because the system only samples a reduced range of the energy function and is more difficult to equilibrate. Furthermore, we note that the phase transition of the system does not affect the accuracy of the method.

3.4. Impact of Basis Functions. The use of Laguerre polynomials raises additional challenges. With increasing order of the Laguerre expansion the linear equation (eq 8) tends to become more and more ill-conditioned: we risk to overfit the observed data and start to model noise. But if we break off the series expansion too early, the approximation of the energy function will be inherently limited. What is the optimal number of features K? How can we guarantee that the matrix inversion is well conditioned?

The system of the linear equation (eq 8) specifies the solution of the least-squares optimization problem min $\|A\lambda - b\|^2$. Therefore, to improve the conditioning of the force field estimation, we can resort to standard regularization techniques. We introduce additional restraints to prevent overfitting and

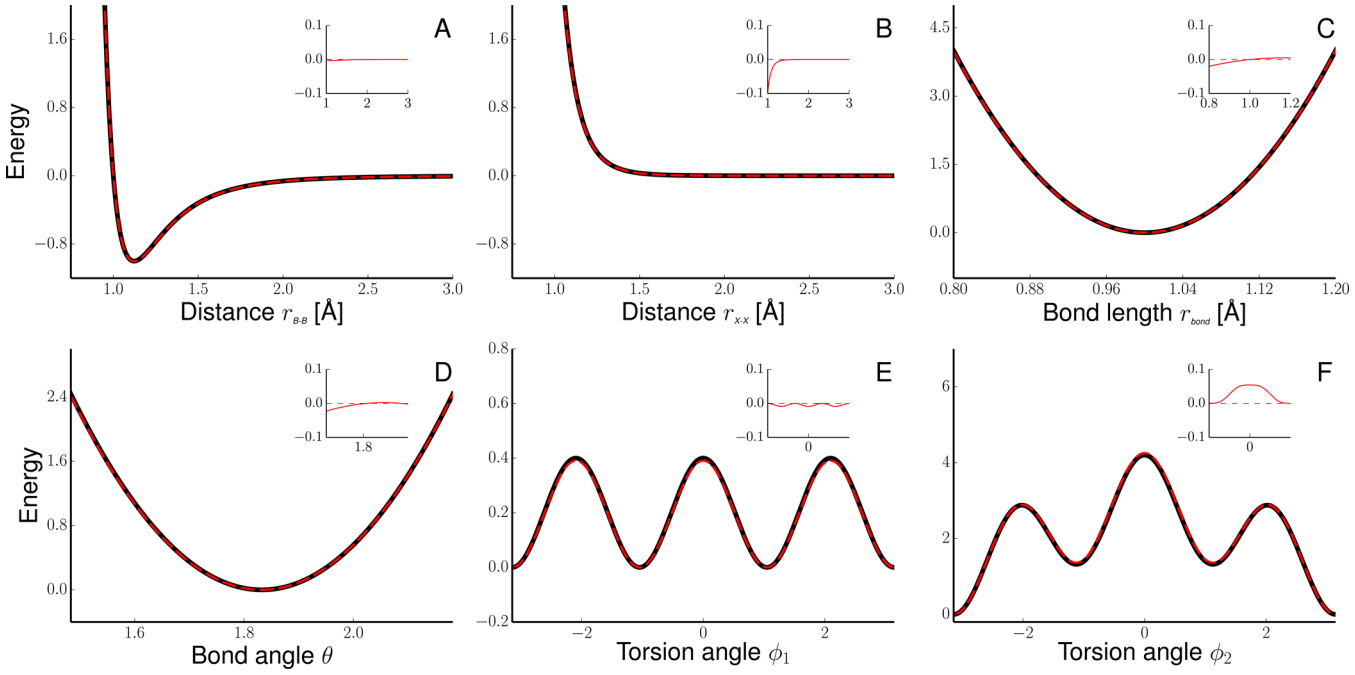


Figure 3. Estimated potentials (red broken line) compared to the true potentials (black line) with insets showing the error.

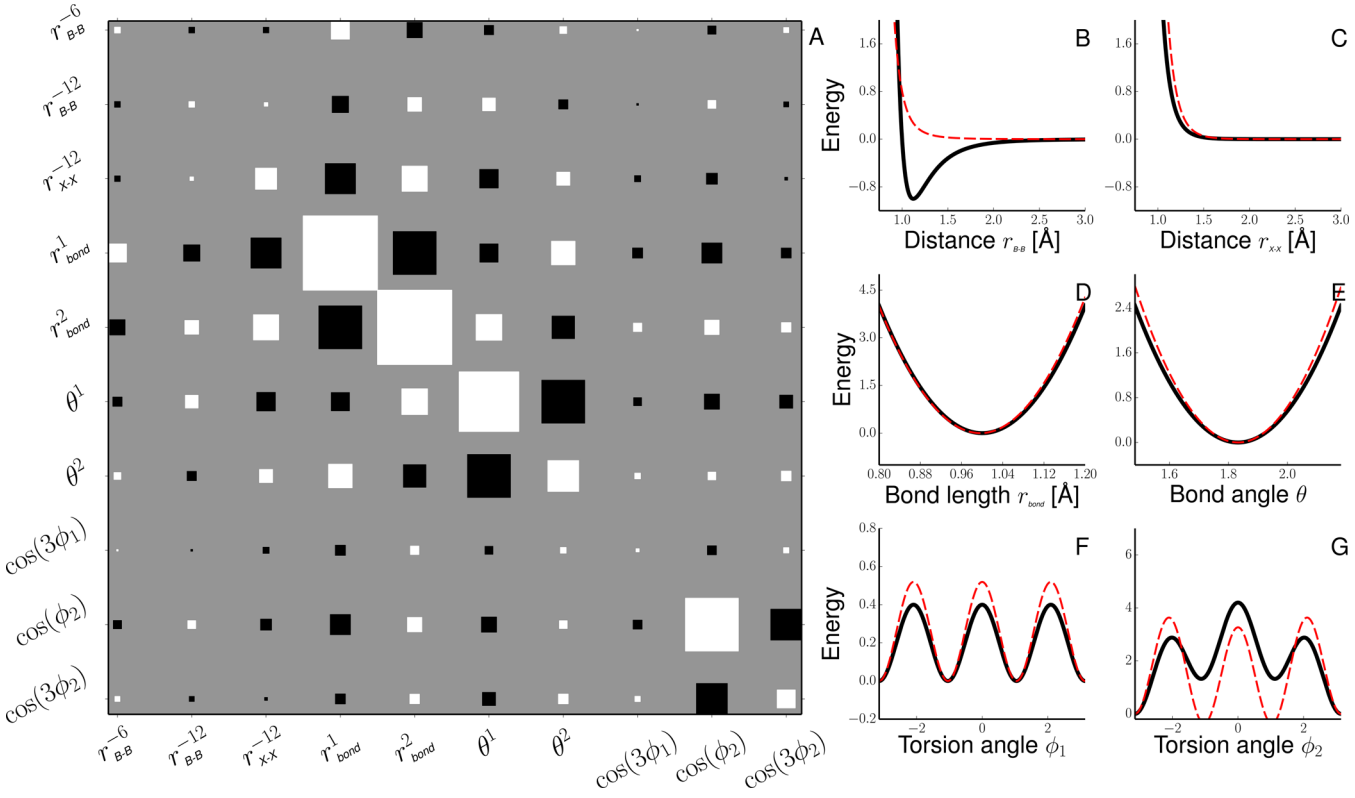


Figure 4. (A) Hinton diagram showing the covariances $(AA^T)^{-1}$ between the force field terms. The color (black/white) indicates the sign (negative/positive) of the covariance and the area of the rectangles is proportional to the magnitude. The features are indicated as axis labels. (B)−(G) Reconstructed potentials that ignore the correlations between the features.

alleviate the ill-posedness of the parameter estimation problem. From a probabilistic point of view, the regularizer can be interpreted as a prior distribution over the force field parameters. We can use the prior to incorporate problem specific knowledge and constraints such as positivity of some parameters. Here, we explore the use of a regularizer enforcing the smoothness of the reconstructed potential energy. With increasing order k, the Laguerre features show finer details analogous to the Fourier basis at higher frequencies. Smoothness can be imposed by a regularizer based on the Laplace operator. In analogy to the Fourier basis we therefore minimize

dx.doi.org/10.1021/ct400580p | J. Chem. Theory Comput. 2013, 9, 5685−5692

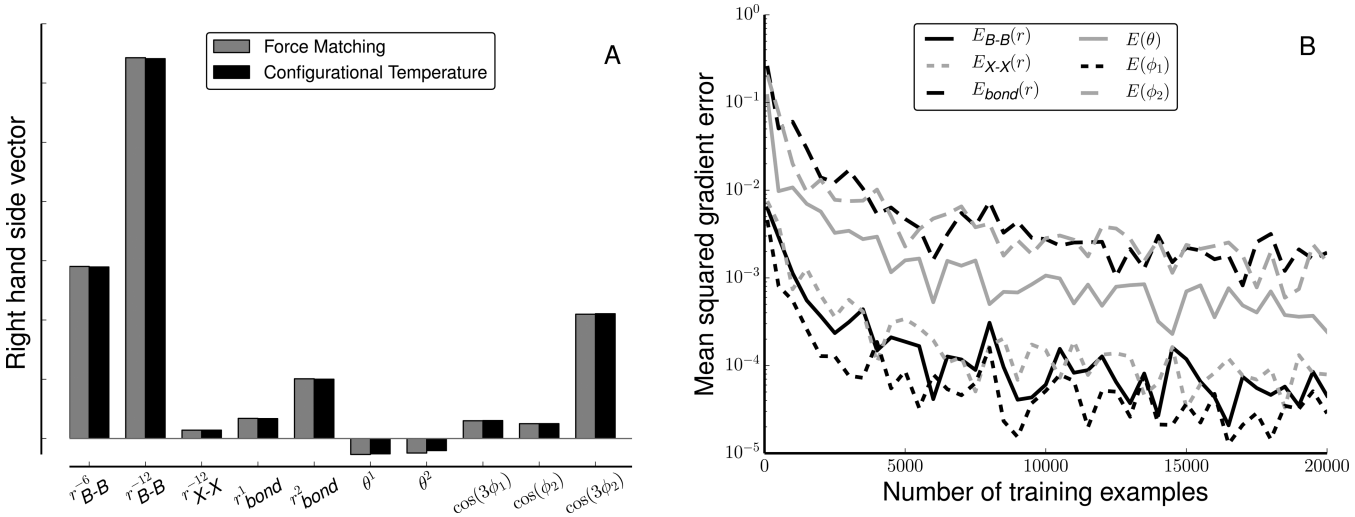


Figure 5. (A) Elements of the right-hand side vectors obtained in the force matching and configurational temperature approach. (B) Mean squared gradient error of the estimated individual potentials depending on sample size averaged over five nonoverlapping sets of structures.

$$\|A\lambda - b\|^2 + \alpha \sum_{k=1}^{K} k^2 {\lambda_k}^2$$

where α is the strength of the regularizer. The effect of the regularizer is that higher-order features are more and more suppressed because their expansion coefficients λ_k are kept close to zero with a force proportional to k^2.

We applied this regularization scheme to the argon simulation and estimated Laguerre approximations of different orders with and without regularization based on 20 argon configurations. The accuracy is again measured by the MSE of the forces evaluated at 1000 linearly spaced points ranging from 3 to 10 Å, and the results are shown in Figure 2. The accuracy of the reconstructed Lennard-Jones potential strongly depends on the maximum number of Laguerre features. Without regularization, the problem becomes ill-conditioned if more than 22 Laguerre features are used and the approximation breaks down. With the help of the regularizer, we can prevent this effect but pay in terms of accuracy. As common in parameter estimation problems, we have a trade-off between model complexity and goodness-of-fit. Instead of choosing α ad hoc, we use an iterative scheme[42] in which we cycle through conditional updates of λ and α. For fixed α we use the LSQR-Algorithm to update λ; for fixed λ we treat α as the precision of a Gaussian distribution and calculate the update analytically. Figure 2 compares the accuracy of the optimized regularization scheme denoted by α_{opt} to the arbitrary choice of $\alpha = 1$. The optimization of α increases the accuracy with little computational cost.

3.5. Coarse-Grained Protein Model. We also tested our estimation method to recover a minimal protein model. Although the Honeycutt−Thirumalai model uses a reduced number of force field terms, it has the characteristics of full-atom force fields and a complex energy landscape with many interdependent contributions.[43] To estimate the force field parameters with the configurational temperature framework, we need to rewrite the energy as a sum of basis functions. For the Lennard-Jones potential for hydrophobic beads we use two features, r^{-6} and r^{-12}, for all other nonbonded interaction only r^{-12}. The gradient of the features is only nonzero for the correct interaction type, or if the interaction beads are less than three bonds apart. Bond and bend terms are modeled with the features r^2, r and θ^2, θ, respectively. We obtain the original parameters by $k_{bonds} = \lambda_{r^2}$ and $r_0 = -\lambda_r/2\lambda_{r^2}$. The parameters k_{angles} and θ_0 are treated in the same manner. For the parameters for the torsion angle potential, a and b, no transformation is necessary; $\cos(\varphi)$ and $\cos(3\varphi)$ can directly be used as features.

We generated 5000 conformations of the β-barrel-like structures comprising 76 beads. The estimated and true potential together with the differences between both are shown in Figure 3. Overall, the estimated potentials provide a very close fit, with only minimal error in sparsely sampled regions of the torsion potential.

To actually learn the potential function, we solve $A\lambda = b$ by least-squares minimization. From a probabilistic perspective this is similar to fitting a multivariate normal distribution with a covariance matrix $(AA^T)^{-1}$. The covariance matrix provides insight into the dependence of the force field terms. A graphical representation of this matrix can be found in Figure 4. It shows that there is considerable coupling between the features. Especially the bond and bend potentials are coupled to the Lennard-Jones features.

We test the importance of the covariances by a simple experiment; for recovering the individual potentials we use the part of the matrix that corresponds to the parameters of the potential. For example, to estimate the B−B potential we use only the r^{-6}_{B-B} and r^{-12}_{B-B} entries of A and b, which reduces our problem to a system of linear equations given by a 2×2 matrix and a vector of size 2. The potential function estimated with reduced interdependence shows systematic errors in the nonbonded potential (Figure 4). Also, the location of the minima of the bond length and bond angle terms are affected as well as the strength of the dihedral potentials.

As discussed in section 3.1, there is a close relation between the configurational temperature approach and force matching. The main difference is that force matching requires knowledge of the true gradient whose energy will be approximated by a parametrized potential. Our configurational temperature approach, on the other hand, works for configurations only and is therefore more broadly applicable. Both approaches result in a system of linear equations for the force field parameters and differ only in the right-hand side (eqs 8 and

13). In Figure 5A we show the right-hand side vectors for the HT model. It is evident that both approaches will give almost identical estimates for λ because the RHS differs only very slightly.

Figure 5B shows the evolution of the mean squared gradient error of the estimated individual potentials for an increasing number of conformations averaged over five independent simulations. The potentials that exhibit little correlation in Figure 4 achieve a significantly smaller error. Also, convergence is reasonably fast: even with as few as 500 conformations we are already able to obtain an accurate approximation of the force field.

4. CONCLUSION

We present an extension of the configurational temperature formalism that allows us to estimate potential functions from configurations. Our method is applicable to a wide range of systems and does not involve additional simulations. It scales well with the number of configurations, is amenable to standard regularization techniques to avoid overfitting and ill-conditioning, and does not require repeated ensemble simulations. The method can easily be extended to optimize nonlinear parameters. Furthermore, it is possible to add and combine different vector fields B_k. For example, the vector field corresponding to the hyperconfigurational temperatures, as introduced by Han and Grier,[44] are more sensitive to the detailed structure of a potentials.

The gYBG equations can be derived as a special case of the configurational temperature formalism. An addition to the gYBG equations is that our method can also account for a known reference force field $f_0(x)$ and thus find "missing" interactions. We also found that for $B_k = \nabla f_k$ our configurational temperature approach is closely related to force matching, the main difference being that we do not require knowledge of the true forces whereas force matching does.

Future research will focus on the derivation of a nonbonded potential for proteins and its application to protein simulation and structure determination.

AUTHOR INFORMATION

Corresponding Author

*M. Habeck: e-mail, mhabeck@gwdg.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work has been supported by Deutsche Forschungsgemeinschaft (DFG) grant HA 5918/1-1, by the University of Göttingen, and by the Max Planck Society. We thank William Noid for providing code to simulate the HT model.

REFERENCES

(1) Lane, T. J.; Shukla, D.; Beauchamp, K. A.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2013**, *3*, 58−65.
(2) Freddolino, P. L.; Park, S.; Roux, B.; Schulten, K. *Biophys. J.* **2009**, *96*, 3772−3780.
(3) Wroblewska, L.; Skolnick, J. *J. Comput. Chem.* **2007**, *28*, 2059−2066.
(4) Lange, O. F.; Van der Spoel, D.; De Groot, B. L. *Biophys. J.* **2010**, *99*, 647−655.
(5) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. *PLoS ONE* **2012**, *7*, e32131.
(6) Wang, T.; Wade, R. C. *J. Chem. Theory Comput.* **2006**, *2*, 140−148.
(7) Best, R. B.; Hummer, G. *J. Phys. Chem. B* **2009**, *113*, 9004−9015.
(8) Mackerell, A. D. *J. Comput. Chem.* **2004**, *25*, 1584−1604.
(9) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(10) Feig, M. *J. Chem. Theory Comput.* **2008**, *4*, 1555−1564.
(11) Tanaka, S.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 945−950.
(12) Miyazawa, S.; Jernigan, R. L. *Macromolecules* **1985**, *18*, 534−552.
(13) Muñoz, V.; Serrano, L. *Nat. Struct. Mol. Biol.* **1994**, *1*, 399−409.
(14) Hansen, J.-P.; McDonald, I. R. *Theory of Simple Liquids, Third ed.*; Academic Press, 2006.
(15) Weeks, J. D.; Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1971**, *54*, 5237.
(16) Tschöp, W.; Kremer, K.; Batoulis, J.; Bürger, T.; Hahn, O. *Acta Polym.* **1998**, *49*, 61−74.
(17) Reith, D.; Pütz, M.; Müller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1624−1636.
(18) Soper, A. *Chem. Phys.* **1996**, *202*, 295−306.
(19) McGreevy, R.; Pusztai, L. *Mol. Simul.* **1988**, *1*, 359−367.
(20) Lyubartsev, A. P.; Laaksonen, A. *Phys. Rev. E* **1995**, *52*, 3730−3737.
(21) Savelyev, A.; Papoian, G. A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 20340−20345.
(22) Shell, M. S. *J. Chem. Phys.* **2008**, *129*, 144108.
(23) Carmichael, S. P.; Shell, M. S. *J. Phys. Chem. B* **2012**, *116*, 8383−8393.
(24) Wang, L.-P.; Chen, J.; Van Voorhis, T. *J. Chem. Theory Comput.* **2012**, *9*, 452−460.
(25) Ercolessi, F.; Adams, J. B. *Europhys. Lett.* **1994**, *26*, 583−588.
(26) Izvekov, S.; Parrinello, M.; Burnham, C. J.; Voth, G. A. *J. Chem. Phys.* **2004**, *120*, 10896.
(27) Izvekov, S.; Voth, G. A. 2005, *109*, 2469−2473.
(28) Mullinax, J.; Noid, W. *Phys. Rev. Lett.* **2009**, *103*, 198104.
(29) Yvon, J. *La théorie statistique des fluides et l'équation d'état*; Hermann & cie, 1935; Vol. *203*.
(30) Born, M.; Green, H. *Proc. R. Soc. London, Ser. A* **1946**, *188*, 10−18.
(31) Rugh, H. H. *Phys. Rev. Lett.* **1997**, *78*, 772−774.
(32) Jepps, O. G.; Ayton, G.; Evans, D. J. *Phys. Rev. E* **2000**, *62*, 4757−4763.
(33) Hinsen, K. *J. Comput. Chem.* **2000**, *21*, 79−85.
(34) Honeycutt, J. D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3526−3529.
(35) Sorenson, J. M.; Head-Gordon, T. *Proteins: Struct., Funct., Bioinf.* **2002**, *46*, 368−379.
(36) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(37) Berendsen, H. J.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(38) Hyvärinen, A. *J. Mach. Learn. Res.* **2005**, *6*, 695−709.
(39) Mullinax, J. W.; Noid, W. G. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19867−19872.
(40) Sippl, M. J. *Curr. Opin. Struct. Biol.* **1995**, *5*, 229−235.
(41) Chandler, D. *Introduction to modern statistical mechanics*; Oxford University Press, 1987; Vol. *1*.
(42) Besag, J. *J. R. Stat. Soc., Ser B* **1986**, 259−302.
(43) Brown, S.; Fawzi, N. J.; Head-Gordon, T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 10712−10717.
(44) Han, Y.; Grier, D. G. *Phys. Rev. Lett.* **2004**, *92*, 148301.