

A Temperature-Dependent Length-Scale for Transferable Local Density Potentials

Ryan J. Szukalo and W. G. Noid^{a)}

*Department of Chemistry, Penn State University, University Park,
Pennsylvania 16802 USA*

(Dated: 30 August 2023)

Recent coarse-grained (CG) models have often supplemented conventional pair potentials with potentials that depend upon the local density around each particle. In this work we investigate the temperature-dependence of these local density (LD) potentials. Specifically, we employ the multiscale coarse-graining (MS-CG) force-matching variational principle to parameterize pair and LD potentials for one-site CG models of molecular liquids at ambient pressure. The accuracy of these MS-CG LD potentials depends quite sensitively upon the length-scale, r_c , that is employed to define the local density. When the local density is defined by the optimal length-scale, r_c^* , the MS-CG potential often accurately describes the reference state point and can provide reasonable transferability across a rather wide range of temperatures. At ambient pressure, the optimal LD length-scale varies linearly with temperature over a very wide range of temperatures. Moreover, if one adopts this temperature-dependent LD length-scale, then the MS-CG LD potential appears independent of temperature, while the MS-CG pair potential varies linearly across this temperature range. This provides a simple means for predicting pair and LD potentials that accurately model new state points without performing additional atomistic simulations. Surprisingly, at certain state points, the predicted potentials provide greater accuracy than MS-CG potentials that were optimized for the state point.

^{a)}Electronic mail: wnoid@chem.psu.edu

I. INTRODUCTION

By representing systems in reduced detail, coarse-grained (CG) models provide much greater computational efficiency than conventional all-atom (AA) models.^{1–8} Unfortunately, this reduced representation also introduces new challenges. Because they result from averaging over atomic interactions, the effective interaction potentials that govern CG models will generally depend upon the thermodynamic conditions for which they were parameterized.⁹ Consequently, a potential that has been optimized to accurately describe a particular thermodynamic state point may provide a relatively poor description at other state points. Thus, one expects that CG potentials may demonstrate relatively limited transferability. Moreover, coarse-graining also complicates the treatment of thermodynamic properties because the atomic details that give rise to energetic and entropic contributions are no longer explicitly present in the model.^{9–15}

While one expects that these general considerations are relevant for all CG models, they are particularly transparent and often quite severe in bottom-up approaches that attempt to formally integrate out AA degrees of freedom.^{5,16–18} In these bottom-up approaches, a (usually) linear mapping, $\mathbf{M}$, represents each AA configuration, $\mathbf{r}$, for n atoms with a CG configuration, $\mathbf{R}$, for $N(< n)$ sites. The “exact” bottom-up potential is the many-body potential of mean force (PMF),^{19–23} $W(\mathbf{R})$, which may be defined by the total Boltzmann weight of AA configurations, $\mathbf{r}$, that map to the CG configuration, $\mathbf{R}$,

$$\exp[-\beta W(\mathbf{R}; V, T)] = V^{-(n-N)} \int_{V^n} d\mathbf{r} \exp[-\beta u(\mathbf{r})] \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}). \quad (1)$$

If the PMF is known as a function of both configuration and also thermodynamic state point, then a CG model can perfectly reproduce the mapped configuration distribution (i.e., the distribution that is determined by the atomic model and the CG mapping) at every state point. Clearly, achieving this perfect transferability requires accounting for how the PMF varies as a function of, e.g., temperature and density. Moreover, while the exact PMF can perfectly reproduce all thermodynamic properties of the AA model, calculations of thermodynamic properties must account for this state-point-dependence.²⁴ For instance, one must account for the density-dependence of the PMF in order to reproduce the internal pressure of the AA model.^{20,25–29} Similarly, because Eq. (1) defines the

PMF as the excess Helmholtz potential for the CG coordinates, the PMF incorporates both energetic and entropic contributions.^{10,21,30} Specifically, $W(\mathbf{R}) = E_W(\mathbf{R}) - TS_W(\mathbf{R})$, where $E_W(\mathbf{R}) \equiv \langle u(\mathbf{r}) \rangle_{\mathbf{R}}$ is the conditioned average of the AA potential evaluated over the subensemble of AA configurations that map to $\mathbf{R}$, while $S_W(\mathbf{R})$ corresponds to the configurational entropy associated with this subensemble.^{18,31} Consequently, W cannot be used to directly calculate energetic properties. Indeed, Louis warned that these “representability” issues will generally arise when computing thermodynamic properties with effective potentials, such as the PMF, that depend upon thermodynamic conditions.^{32,33}

In practice, bottom-up CG models usually approximate the exact PMF, W , with a relatively simple potential, U , that, e.g., describes intermolecular interactions with central pair potentials.^{2,7,23} In some cases a single approximate potential, $U(\mathbf{R})$, can provide good transferability and a reasonable description of thermodynamic properties over a relatively wide range of conditions.^{34–39} More generally, though, one expects that the “optimal” approximate potential, $U(\mathbf{R}; V, T) \approx W(\mathbf{R}; V, T)$, will vary with thermodynamic state point. Moreover, this state-point-dependence may be important for accurately reproducing thermodynamic properties and for improving the transferability of bottom-up CG models.^{18,24,29,31–33,40,41}

Consequently, many prior studies have investigated the influence of temperature-, density-, and composition-dependence of bottom-up pair potentials.^{42–61} In particular, many studies suggest that bottom-up pair potentials vary linearly with temperature and density over a rather wide range of thermodynamic conditions.^{30,33,59,62–64} This linearity motivated a simple “dual” approach for predictively treating the temperature-dependence of bottom-up pair potentials based upon simulations at a single state point.^{65–67} This dual approach employs conventional bottom-up approaches, e.g., iterative Boltzmann inversion⁶⁸ or force-matching,^{69,70} to determine an effective potential, U , that accurately approximates the configuration-dependence of the PMF. The dual approach repurposes an energy-matching variational principle⁷¹ to determine an operator, $E(\mathbf{R})$, that approximates the energetic component of the PMF, $E_W(\mathbf{R})$. While simulations with the bottom-up effective potential, U , accurately reproduce the mapped ensemble, atomic energetics can be estimated by evaluating the energetic operator, E , for the sampled configurations. Moreover, the dual approach approximates the entropic component of the PMF by $S(\mathbf{R}) \equiv (E(\mathbf{R}) - U(\mathbf{R})) / T \approx S_W(\mathbf{R})$, which provides a predictive estimate for the temperature-dependence of the approximate potential, i.e., $\partial U(\mathbf{R}) / \partial T \approx \partial W(\mathbf{R}) / \partial T = -S_W(\mathbf{R}) \approx -S(\mathbf{R})$. Despite its simplicity,

initial studies suggest that this dual approach is surprisingly accurate and predictive for several molecular liquids, including water, methanol, and ortho-terphenyl (OTP).^{65,67} More recently, Pretti and Shell have introduced a complementary approach for addressing the temperature-dependence of the PMF by approximating a microcanonical partition function, $\Omega(\mathbf{R}, E)$, for CG coordinates.⁷²

While these prior studies have focused on pair-additive potentials, a growing number of CG models have adopted local density (LD) potentials, U_ρ , that depend upon the density, ρ_I , around each CG site. By defining the local density, ρ_I , with pair-additive contributions from neighboring molecules within a given radius, r_c , LD potentials generate pair-additive forces and, thus, provide the same computational scaling as conventional pair-additive potentials.⁷³ These LD potentials were first introduced by Pagonabarraga and Frenkel to describe non-ideal thermodynamic properties with relatively simple top-down dissipative particle dynamics (DPD) models.⁷³ Independently, Papoian et al. employed LD potentials to describe water-mediated interactions in CG models for protein structure prediction.⁷⁴ In particular, the many-body DPD method of Pagonabarraga and Frenkel employs LD potentials to associate a temperature-dependent, local free energy with each particle.^{73,75} Similarly, several related DPD variations assign a temperature to each particle and employ thermodynamic models to describe this local free energy or the transfer of internal energy between particles.^{76–82}

More recently, bottom-up CG models have employed LD potentials to model many-body solvent effects,^{57,83–85} liquid/liquid phase separation,⁸⁶ liquid-vapor interfaces,^{55,87–90} polymeric systems,^{60,91–93} and even shock-waves in explosive materials.⁹⁴ These LD potentials can improve the description of many-body structural correlations and also provide a much improved description of the internal pressure. Moreover, they provide remarkable transferability between bulk and interfacial environments, which has proven quite challenging with pair-additive potentials.^{95,96}

One expects that bottom-up LD potentials should also vary with temperature in order to model the temperature-dependence of the exact PMF. However, few studies have carefully investigated this. To our knowledge, the work of Shahidi et al. provides the most thorough prior investigation of temperature-dependent bottom-up LD potentials.⁶⁰ Shahidi et al. employed the Inverse Monte Carlo method^{97,98} to parameterize pair and LD potentials for polyisoprene melts at three different temperatures. Interestingly, they defined the local

density based upon nearest-neighbors. In particular, they employed the solid-angle based nearest-neighbor algorithm of van Meel et al.⁹⁹ to determine the length-scale and weighting function for the local density at each temperature. The resulting models quite accurately reproduced conformational and thermodynamic properties of the underlying united atom model for polyisoprene.

In this work, we investigate the temperature-dependence of bottom-up LD potentials for 1-site CG models of molecular liquids, such as methanol and OTP. We first investigate the influence of the local density radius, r_c , upon the potentials obtained via the multiscale coarse-graining (MS-CG) force-matching (FM) variational principle^{69,70,100,101} at a given reference state point. When r_c is properly chosen, the MS-CG model quite accurately describes the AA pair structure and pressure-density equation of state near this reference state point. These MS-CG potentials also provide relatively good transferability across a rather wide temperature range. However, if one fixes the LD length scale, r_c , and recalculates the MS-CG potentials for other temperatures, the resulting MS-CG models perform quite poorly. Interestingly, though, if one varies the LD length scale with temperature, then a single LD potential accurately describes the liquid pair structure and pressure-density equation of state with nearly quantitative accuracy over this temperature range. Furthermore, this temperature-dependent length scale, $r_c(T)$, and the corresponding pair potentials, $U_2(T)$, can be accurately predicted via simple linear interpolation between two reference state points. Surprisingly, we find that these predicted potentials describe certain state points more accurately than MS-CG potentials that were specifically parameterized for the state points.

The remainder of this work is organized as follows. Section II summarizes the relevant properties of LD potentials. Sections III and IV describe our computational methods and the results of our numerical studies, respectively. Section V summarizes our findings, while Sec. VI provides closing comments. The Supplementary Material (SM) provides additional computational details and analysis.

II. LOCAL DENSITY POTENTIALS

In this work we consider CG models for one-component molecular liquids and represent each molecule with a single site. We adopt CG potentials of the following form:

$$U(\mathbf{R}) = U_{\text{pair}}(\mathbf{R}) + U_{\text{LD}}(\mathbf{R}), \quad (2)$$

where U_{pair} is a sum of central pair potentials, U_2 , that depend upon the distance, R_{IJ} , between each pair of molecules,

$$U_{\text{pair}}(\mathbf{R}) = \sum_{(I,J)} U_2(R_{IJ}), \quad (3)$$

and U_{LD} is a sum of local density (LD) potentials, U_ρ , that depend upon the local density, ρ_I , around each molecule,

$$U_{\text{LD}}(\mathbf{R}) = \sum_{I=1}^N U_\rho(\rho_I). \quad (4)$$

We have recently introduced potentials that also depend upon the square gradient, $|\nabla_I \rho_I|^2$, of the local density.⁹⁰ However, we do not consider these square gradient potentials in this work.

The local density around molecule I is defined by a sum of distance-dependent contributions from its neighbors:

$$\rho_I(\mathbf{R}) = \sum_{J(\neq I)} \bar{w}(R_{IJ}), \quad (5)$$

where $\bar{w}$ is a scaled weighting function. While prior studies have considered various weighting functions,^{55,73,74,83,88,91} we find the Lucy function to be particularly convenient. The Lucy function is defined

$$w(r) = (1 - r/r_c)^3 (1 + 3r/r_c) \Theta(1 - r/r_c), \quad (6)$$

where $\Theta(x)$ denotes the Heaviside function. The Lucy function, $w(r) = w(r; r_c)$, depends on a single parameter, r_c , that determines the length-scale defining the local density. As illustrated in Fig. 1, the Lucy function equals 1 when $r = 0$, monotonically decays to 0 over the interval $0 < r < r_c$, and vanishes for $r \geq r_c$. The scaled weighting function in Eq. (5) is

defined

$$\bar{w}(r) = w(r)/[w], \quad (7)$$

where $[w] = \int dr 4\pi r^2 w(r) = 16\pi r_c^3/105$.

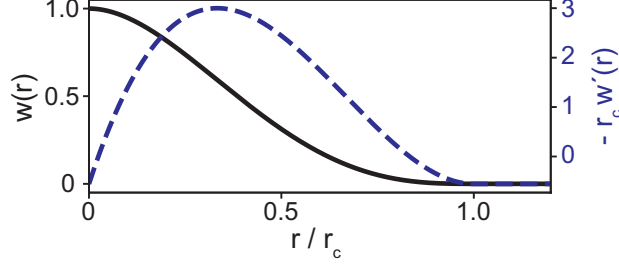


FIG. 1. Plot of the Lucy function, $w(r)$, (solid curve) and its (scaled) derivative, $-r_c w'(r)$, (dashed curve) as a function of r/r_c .

In contrast to our previous studies,^{87–89} this definition of the local density excludes the “self-term,” i.e., the $J = I$ term in Eq. (5). This self-term simply shifts the definition of local density and has no physical significance. However, we find that excluding the self-term simplifies our analysis of the LD length scale, r_c . Moreover, Eq. (5) has the appealing property that $\rho_I \rightarrow 0$ in the gas phase.

Because the local density, ρ_I , is defined by pair-additive contributions that depend upon pair distances, R_{IJ} , Eq. (2) generates pair additive forces

$$\mathbf{F}_I(\mathbf{R}) \equiv -\nabla_I U(\mathbf{R}) = \sum_{J(\neq I)} F_{IJ}(\mathbf{R}) \hat{\mathbf{R}}_{IJ}, \quad (8)$$

where $\hat{\mathbf{R}}_{IJ} = (\mathbf{R}_I - \mathbf{R}_J)/R_{IJ}$ is the unit vector pointing from site J to site I . The magnitude of the total pair force is

$$F_{IJ}(\mathbf{R}) \equiv F_2(R_{IJ}) + [F_\rho(\rho_I) + F_\rho(\rho_J)] \bar{w}'(R_{IJ}), \quad (9)$$

where $F_2(R) = -dU_2(R)/dR$ is the pair force and $F_\rho(\rho_I) = -dU_\rho(\rho_I)/d\rho_I$ is the corresponding LD force.

Equation (9) indicates that LD forces between a given pair (I, J) of particles is proportional to $w'(R_{IJ})$, which is plotted in Fig. 1. If one treats $F_\rho(\rho_I)$ and $F_\rho(\rho_J)$ as fixed, then the LD force between the pair is largest when the two are separated by a distance

$R_{IJ} = r_c/3$. Moreover, note that the scaled Lucy weighting function, $\bar{w}$, is proportional to r_c^{-3} . Therefore, one expects that the magnitude of the LD force between pairs will generally decrease as r_c increases and the LD forces are distributed over a larger length scale. (This assumes that $F_\rho(\rho_I)$ and $F_\rho(\rho_J)$ vary relatively slowly with r_c .)

Equations (8) and (9) imply that the LD potential generates pair additive contributions to the internal pressure,

$$P_{\text{LD}}(\mathbf{R}, V) \equiv - \left(\frac{\partial U_{\text{LD}}}{\partial V} \right)_{\hat{\mathbf{R}}, T} = \frac{1}{3V} \sum_{I=1}^N F_\rho(\rho_I) \sum_{J(\neq I)} \bar{w}'(R_{IJ}) R_{IJ}. \quad (10)$$

Here the partial derivative is performed with respect to volume, while holding fixed the scaled coordinates, $\hat{\mathbf{R}}_I = V^{-1/3} \mathbf{R}_I$. Note that Eq. (10) assumes that r_c is independent of V and is consistent with the temperature-dependent length-scale, $r_c = r_c(T)$, that we introduce in Sec. IV A 3.

Equations (8) and (9) also imply the existence of a non-trivial one-parameter family of pair and LD potentials that are equivalent.⁸⁸ Specifically, consider a given potential of the form given by Eqs. (2)-(4) with both pair and LD terms, $U = \{U_2, U_\rho\}$. For any constant, c , the force transformation

$$F_2(R) \rightarrow \tilde{F}_2(R; c) = F_2(R) + c\bar{w}'(R) \quad (11)$$

$$F_\rho(\rho) \rightarrow \tilde{F}_\rho(\rho; c) = F_\rho(\rho) - c/2 \quad (12)$$

defines new pair and LD potentials, $\tilde{U}(c) = \{\tilde{U}_2(c), \tilde{U}_\rho(c)\}$, that leave Eqs. (8) and (9) unchanged.⁸⁸ In particular, by employing Eqs. (11) and (12) to simultaneously vary both $\tilde{U}_2(c)$ and $\tilde{U}_\rho(c)$, one can choose c to make the LD potential either purely repulsive or purely attractive without altering the equilibrium configuration distribution. Consequently, although they correspond to very different pair, $\tilde{U}_2(c)$, and LD, $\tilde{U}_\rho(c)$, potentials, all members of this one-parameter family give rise to identical structural and thermodynamic properties. This invariance explains prior observations that pair and LD potentials can compensate for each other.^{60,86,88}

The LD potential, U_ρ , often appears parabolic with a single minima at a characteristic local density, ρ_0 . In this case, $F_\rho(\rho_I) > 0$ drives the system to higher local densities when $\rho_I < \rho_0$. Because $\bar{w}'(R_{IJ}) \leq 0$, this leads to attractive pair forces,

$F_\rho(\rho_I)\bar{w}'(R_{IJ})\hat{\mathbf{R}}_{IJ} = -|F_\rho(\rho_I)\bar{w}'(R_{IJ})|\hat{\mathbf{R}}_{IJ}$, and negative contributions to the internal pressure, $\frac{1}{3V}F_\rho(\rho_I)\sum_{J(\neq I)}\bar{w}'(R_{IJ})R_{IJ} < 0$. Conversely, when $\rho_I > \rho_0$, $F_\rho(\rho_I) < 0$ drives the system to lower local densities via repulsive pair forces and positive contributions to the internal pressure.

In the following, we determine the optimal CG potential, $U = \{U_2, U_\rho\}$, by minimizing the MS-CG force-matching (FM) functional:^{69,70,100,101}

$$\chi^2[U] = \frac{1}{3N} \left\langle \sum_{I=1}^N |\mathbf{F}_I(\mathbf{M}(\mathbf{r})) - \mathbf{f}_I(\mathbf{r})|^2 \right\rangle, \quad (13)$$

where $\mathbf{f}_I(\mathbf{r})$ is the net force on CG site I in the atomic configuration $\mathbf{r}$.^{100,102} The MS-CG variational principle simultaneously determines both U_2 and U_ρ such that $U = \{U_2, U_\rho\}$ optimally approximates the configuration-dependence of the exact PMF.^{23,100,101} As discussed above, the local density, ρ_I , depends upon the length scale, r_c , that defines the Lucy weighting function, $w(r; r_c)$. Consequently, the calculated MS-CG pair and LD potentials will both depend upon this length-scale, i.e., $U = U(r_c) = \{U_2(r_c), U_\rho(r_c)\}$.

III. COMPUTATIONAL METHODS

A. Atomistic Models

In the following we report constant NPT simulations of AA models for liquid methanol and liquid OTP. We described both systems with the OPLS/AA force field.¹⁰³ We truncated all non-bonded interactions at 1.3 nm, while switching the corresponding forces to zero between 1.0 nm and 1.3 nm. We modeled electrostatic interactions with the particle-mesh Ewald method,¹⁰⁴ while employing a short-range cut-off of 1.3 nm and a grid spacing of 0.08 nm. We did not employ dispersion corrections for modeling the internal energy or pressure of methanol, but did employ these dispersion corrections in the reported simulations of OTP. The SM demonstrates that these dispersions corrections had relatively little impact upon the simulated AA models or the parameterized CG models. In particular, the SM demonstrates that the treatment of these dispersion corrections does not impact the accuracy of the optimized or predicted CG models that are reported in Section IV B.

We performed all AA simulations with the GROMACS 2019.6 simulation package,^{105,106}

while employing three-dimensional periodic boundary conditions. These simulations propagated the equations of motion with a leap-frog integrator, while employing a 1 fs time step and updating the neighbor list every 10 time steps. The reported AA simulations employed the velocity-re-scaling thermostat¹⁰⁷ with a relaxation time of 0.5 ps and an isotropic Parrinello-Rahman barostat¹⁰⁸ with a relaxation time of 5 ps and a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$ to control energy and volume fluctuations, respectively.

TABLE I. Temperature, T , and equilibrium bulk density, $\bar{\rho}_b$, of simulated state points in units of K and nm^{-3} , respectively.

Methanol		OTP	
T	$\bar{\rho}_b$	T	$\bar{\rho}_b$
220	16.24	400	2.70
240	15.84	500	2.49
260	15.44	575	2.32
280	15.04	650	2.14
300	14.62		

We report AA constant NPT simulations of 4407 methanol molecules at 1 bar external pressure and the five temperatures indicated in Table I. We generated an initial condition for each state point by employing the GROMACS insert-molecules function to randomly place molecules inside a box with dimensions of $5 \text{ nm} \times 5 \text{ nm} \times 10 \text{ nm}$. Following energy minimization, we performed constant volume simulations of the system for 2 ns at the corresponding target temperature. We used the final configuration from this constant NVT simulation as the initial configuration for a subsequent 2 ns simulation at 1 bar external pressure, while employing the Berendsen thermostat¹⁰⁹ to control the volume fluctuations. We employed the resulting configuration as the initial configuration for a 32 ns production simulation in the constant NPT ensemble at 1 bar external pressure. After discarding the first 2 ns as additional equilibration, we sampled the remainder of the simulation every 1 ps in order to obtain 30,000 independent samples.

We also report AA constant NPT simulations of 800 OTP molecules at 1 bar external pressure and the four temperatures indicated in Table I. We generated an equilibrated initial condition for OTP using a similar protocol, but employed longer equilibration times, as reported in our previous work.⁶⁴ After equilibrating the AA OTP model at each temperature, we performed an 80 ns production simulation in the constant NPT ensemble. We sampled this production simulation every 1 ps to obtain 80,000 independent samples.

For each state point, we calculated the pressure-density equation of state and the isothermal compressibility, κ_T , from the set, $\{(V, P_{\text{int}})\}$, of volumes and internal pressures sampled from the corresponding constant NPT production simulation. We calculated the equation of state by histogramming the sampled densities and determining the mean internal pressure for the configurations assigned to each bin. We calculated the compressibility, κ_T , by employing linear regression to determine $(\partial P_{\text{int}}(V)/\partial V)_T$ from the sampled volumes and internal pressures. This procedure for calculating κ_T appeared very robust and agreed well with the fluctuation formula, $\kappa_T = \sigma_V^2/Vk_B T$.¹¹⁰ Conversely, estimating $(\partial P_{\text{int}}(V)/\partial V)_T$ from the calculated equation of state appeared much more sensitive to statistical uncertainty in estimating the internal pressure at each density.

B. CG Model Parameterization

We parameterized CG models for liquid methanol and liquid OTP at each state point indicated in Table I. In both cases, the CG model represented each molecule by a single site with coordinates corresponding to its molecular mass center. We employed the Bottom-up Open-source Coarse-graining Software (BOCS) package¹¹¹ to map each AA trajectory to the CG representation and to determine the MS-CG potential that minimized the force-matching functional, $\chi^2[U]$, in Eq. (13). The MS-CG potential, $U = \{U_2, U_\rho\}$, was defined by a pair potential, U_2 , and a LD potential, U_ρ . We represented the pair force, $F_2(R)$, with cubic spline functions on a uniform grid with spacing $\Delta R = 0.01$ nm that was truncated at 2.0 nm. We represented the LD force, $F_\rho(\rho)$, with cubic spline functions on a uniform grid with spacing $\Delta\rho = 0.01$ nm⁻³ and 0.001 nm⁻³ for methanol and OTP, respectively.

We determined the optimal coefficients for these cubic spline basis functions by matching the AA forces in a least squares sense via singular value decomposition (SVD), while employing a numerical threshold, $\epsilon = 10^{-8}$, to distinguish meaningful singular values from statistical noise. The calculated MS-CG force functions were essentially independent of the SVD threshold over the range $10^{-7} \geq \epsilon \geq 10^{-10}$.

The BOCS code employs a default parameter, $\gamma = 10^{-8}$, to eliminate basis functions that are not adequately sampled. The calculated force functions appeared to depend upon this γ parameter. However, the force functions, $\{F_2(\gamma), F_\rho(\gamma)\}$, obtained for different values of γ corresponded to different members of the one-parameter family described by Eqs. (11) and

(12). Consequently, the calculated MS-CG potentials appear quite robust with respect to the parameters, γ and ϵ , employed in numerically minimizing χ^2 .

In this work, we parameterized and simulated hundreds of CG potentials. In order to perform stable simulations with the calculated MS-CG potentials, it is necessary to sensibly extrapolate these potentials into regions that were not sampled by the AA simulations. In particular, one expects that $F_2(r) > 0$ and, moreover, $dF_2(r)/dr < 0$ in the “hard-core” region as $r \rightarrow 0$. Similarly, one expects that $F_\rho(\rho) > 0$ and $dF_\rho(\rho)/d\rho < 0$ in the low-density regime, $\rho \rightarrow 0$. Conversely, one expects that $F_\rho(\rho) < 0$ and $dF_\rho(\rho)/d\rho > 0$ in the high-density regime, $\rho \rightarrow \infty$.

It is standard and relatively straight-forward to extrapolate calculated pair potentials into the hard-core region. However, it is somewhat less straight-forward to extrapolate the calculated LD potentials, $F_\rho(\rho)$, for $\rho \rightarrow 0$ and $\rho \rightarrow \infty$. Consequently, we developed a systematic high-throughput method for automatically extrapolating calculated LD forces. Briefly, we first fit the AA LD distribution to a Gaussian distribution in order to determine a “confidence region” for which the calculated LD force does not suffer from poor statistics. We extrapolated the LD forces into poorly sampled regimes with a quadratic polynomial that we determined by fitting the calculated LD forces near the boundaries of the confidence region. This procedure proved very successful for methanol. However, this procedure failed to treat the $\rho \rightarrow 0$ limit for OTP. Consequently, we set the LD force for OTP to a constant in the regime where $\rho \rightarrow 0$ and determined this constant from the boundary of the confidence region. The SM describes our numerical procedure for extrapolating the LD forces in greater detail.

C. CG Simulations

We simulated all CG models with a version of LAMMPS^{112,113} that we previously modified to support LD potentials.⁸⁸ These simulations controlled energy fluctuations with a Nosé-Hoover thermostat,^{114,115} while employing a damping parameter of 100 fs and the default chain length ($n = 3$).¹¹⁶ The simulations controlled volume fluctuations with the Martyna-Tuckerman-Tobias-Klein barostat,^{117,118} while employing a damping parameter of 1 ps. The initial condition for each CG simulation was obtained by mapping to the CG resolution the final configuration of the corresponding AA simulation. This configuration was equilibrated

for 1 ns and then simulated for 20 ns in the constant NPT ensemble at 1 bar external pressure and the corresponding temperature, while employing a 1 fs timestep. We calculated the pressure-density equation of state and compressibility of the CG model via the procedure described above for the AA model.

IV. RESULTS

A. Methanol

1. Local Density Definition

We first consider an AA constant NPT simulation of 4407 methanol molecules in the liquid phase at the reference temperature, $T_{\text{ref}} = 260$ K, and an external pressure of 1 bar. The equilibrium bulk density of this AA simulation, $\bar{\rho}_{b;\text{AA}} = 15.44 \text{ nm}^{-3}$ (821.8 kg m^{-3}), agrees quantitatively with experimental measurements, $\bar{\rho}_{b;\text{exp}} = 15.46 \text{ nm}^{-3}$ (822.7 kg m^{-3}), for the density of liquid methanol in this state point. Figure 2 characterizes the mapped ensemble that is obtained by representing each molecule by its mass center.

Figure 2a presents the resulting mapped radial distribution function (RDF). The RDF vanishes at distances less than $r_0 \approx 0.28$ nm. The first peak of the RDF at $r_{\text{HB}} \approx 0.34$ nm corresponds to molecules that are forming hydrogen bonds. The second peak at $r_{\text{VDW}} \approx 0.45$ nm corresponds to molecules that are forming van der Waals contacts. We consider the second minimum near $r_2 \approx 0.62$ nm to indicate the end of the first solvation shell. For comparison, we also employed the solid-angle based nearest-neighbor (SANN) algorithm of van Meel et al.⁹⁹ to determine the radius, R_{NN} , of the nearest-neighbor shell. The SM presents this analysis and determines that the nearest-neighbor radius is $R_{\text{NN}} \approx 0.60$ nm, which is very close to the second minimum of the mapped RDF.

Figure 2b presents the Lucy function, $w(r) = w(r; r_c)$, that is used to define the local density, ρ_I , around each molecule for several different values of the LD length-scale, r_c . Figure 2c presents the corresponding distributions of local densities, $P(\rho_{\text{LD}}) = P(\rho_{\text{LD}}; r_c)$, for molecules in the mapped ensemble. As might be expected, the LD distributions are quite Gaussian.

As r_c increases and the local density is defined over longer length-scales, individual molecules generally contribute less to the local density. Moreover, since Eq. (5) defines

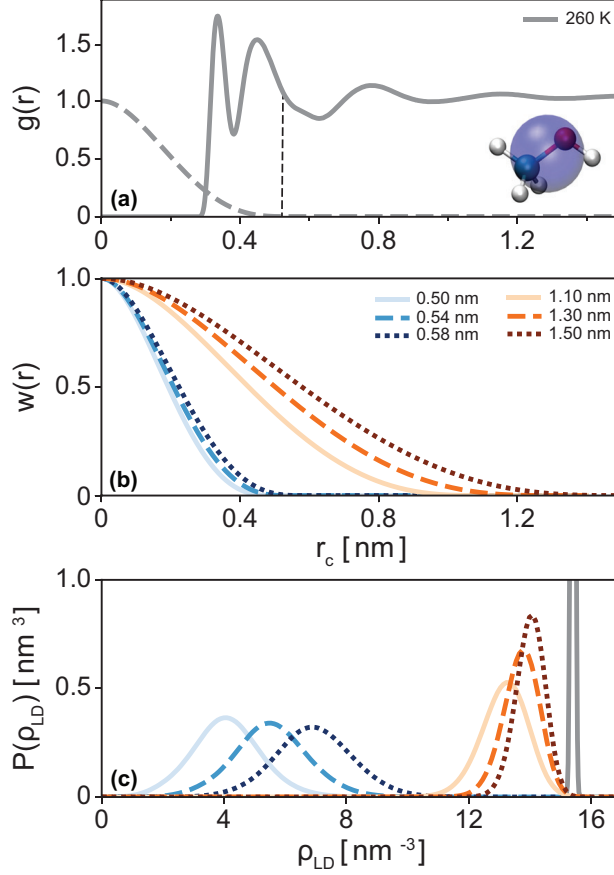


FIG. 2. Analysis of the reference mapped ensemble obtained from an AA constant NPT simulation of methanol at the reference temperature, $T_{\text{ref}} = 260$ K, and external pressure, $P_{\text{ext}} = 1$ bar. Panel (a) presents the mapped RDF, as well as the mapped representation of each methanol molecule. In panel (a), the dashed vertical line indicates the optimal LD length-scale, r_c^* , while the dashed grey curve presents the corresponding Lucy function, $w(r; r_c^*)$. Panel (b) presents the Lucy function, $w(r) = w(r; r_c)$, for several different values of the LD length-scale, r_c . Panel (c) presents the corresponding LD distributions, $P(\rho_{LD}) = P(\rho_{LD}; r_c)$, calculated from the reference mapped ensemble for the various LD length-scales. In panels (b) and (c), the blue and orange curves indicate relatively short and long LD length-scales, respectively, while the intensity of the color increases as r_c increases. The solid grey curve in panel (c) presents the distribution of bulk densities, ρ_b , sampled by the constant NPT simulation.

the local density, ρ_I , around site I without a “self-contribution” from molecule I , the local density generally increases as r_c increases. For sufficiently large r_c , the local density eventually converges upon the global density, ρ_b .

The solid, dashed, and dotted blue curves of increasing intensity consider local densities over increasing length-scales $r_c = 0.50$ nm, 0.54 nm, and 0.58 nm, respectively. For these relatively short length-scales, the local density is determined by only the nearest 8-12 molecules. As r_c increases from 0.50 nm to 0.58 nm, molecules sense significantly larger local densi-

ties. Moreover, the distribution of local densities also slightly broadens as molecules sense a broader range of environments. The LD distribution, $P(\rho_{\text{LD}}; r_c)$, continues to broaden as r_c increases until $r_c \approx 0.7$ nm, which is slightly greater than the radius of the first solvation shell, $r_2 \approx 0.62$ nm.

The solid, dashed, and dotted orange curves of increasing intensity consider local densities over increasing length-scales $r_c = 1.10$ nm, 1.30 nm, and 1.50 nm, respectively. For these relatively long length-scales, the local density is determined by approximately 81, 141, and 216 molecules, respectively. As r_c increases from 1.10 nm to 1.50 nm, the distribution of local densities in the mapped ensemble slightly shifts to larger local densities. Moreover, as r_c increases in this regime, the distribution of local densities sharpens and eventually converges upon the distribution of bulk densities, which is indicated by the solid grey distribution in Fig. 2c.

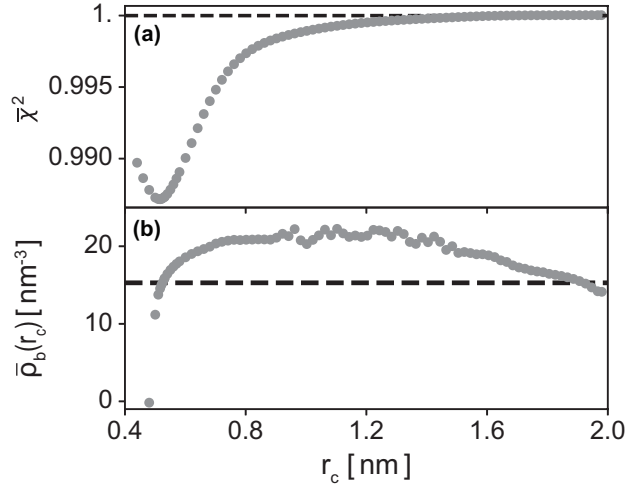


FIG. 3. The impact of the LD length-scale, r_c , upon the MS-CG potential, $U(r_c) = \{U_2(r_c), U_\rho(r_c)\}$, calculated for the one-site methanol model from the reference mapped ensemble at $T_{\text{ref}} = 260$ K and $P_{\text{ext}} = 1$ bar. Panel (a) presents the scaled FM error, $\bar{\chi}^2(r_c) = \chi^2[U(r_c)]/\chi^2[U_\phi]$, where $U_\phi = \{U_{2\phi}\}$ indicates the pair-additive MS-CG potential obtained by performing the FM variational calculation without a LD potential. The dashed horizontal line in panel (a) indicates the limit in which the LD potential does not reduce the FM error, i.e., $\bar{\chi}^2 = 1$. Panel (b) presents the equilibrium bulk density, $\bar{\rho}_b(r_c)$, for the MS-CG potential, $U(r_c)$, which was determined by simulating $U(r_c)$ in the constant NPT ensemble at the reference state point, $T_{\text{ref}} = 260$ K and $P_{\text{ext}} = 1$ bar. The dashed horizontal line in panel (b) indicates the equilibrium bulk density, $\bar{\rho}_{b;\text{AA}}$, of the AA model at this reference state point.

We employed the FM variational principle to determine the MS-CG potential, $U(r_c) = \{U_2(r_c), U_\rho(r_c)\}$, for each LD length-scale, r_c . Figure 3a presents the FM error for the

MS-CG potential, $U(r_c)$, as a function of r_c . Specifically, we report the scaled FM error, $\bar{\chi}^2(r_c) = \chi^2[U(r_c)]/\chi^2[U_{2\phi}]$, where $U_{2\phi}$ is the pair potential obtained from the FM variational calculation when the MS-CG potential only treats pair-additive interactions. Equation (9) demonstrates that the LD potential generates pair-additive forces that are proportional to $w'(r)$, which vanishes for $r \geq r_c$. Consequently, for $r_c \leq r_0 \approx 0.28$ nm, the LD potential has no impact upon the CG model. As r_c increases past r_0 , $\bar{\chi}^2$ rapidly decreases because the LD potential generates forces on a growing number of molecules. Consequently, $\bar{\chi}^2(r_c)$ attains a minimum at $r_c \approx 0.515$ nm. However, even at this minimum, the LD potential reduces $\bar{\chi}^2$ by less than 1.5%. As r_c increases further, the LD forces generally grow weaker and contribute progressively less to $\bar{\chi}^2$. Figure 3b presents the equilibrium bulk density, $\bar{\rho}_b(r_c)$, for each MS-CG potential, $U(r_c)$, which was determined by a constant NPT simulation with $U(r_c)$ at the reference temperature, $T_{\text{ref}} = 260$ K, and an external pressure of 1 bar. The dashed horizontal line in Fig. 3b indicates the equilibrium bulk density, $\bar{\rho}_{b,\text{AA}}$, of the AA model at this state point.

Figure 4 presents the MS-CG pair (left column) and LD (right column) potentials calculated for the LD length-scales, r_c , that are indicated in Fig. 2. As previously observed,^{59,70} the MS-CG pair potential for the 1-site methanol model demonstrates a narrow minimum near the hydrogen-bonding distance, $r_{\text{HB}} \approx 0.34$ nm, as well as a minimum that is broader and deeper around the vdW contact distance, $r_{\text{VDW}} \approx 0.45$ nm. The SM demonstrates that these pair potentials are significantly more attractive than the MS-CG pair potential, $U_{2\phi}$, that is obtained in the absence of a LD potential. Conversely, the MS-CG LD potentials appear rather parabolic with a minimum at a local density, $\rho_{\text{LD}} = \rho_0$, that is near the maximum of the corresponding LD distribution, $P(\rho_{\text{LD}}; r_c)$.

The top row of Fig. 4 presents MS-CG potentials for the relatively short LD length-scales, $r_c = 0.50$ nm, 0.54 nm, and 0.58 nm. As r_c increases in this regime, the MS-CG potential, $U(r_c)$, becomes significantly more attractive. Specifically, the pair potential deepens, while the LD potential dramatically shifts to larger local densities. Figure 3b demonstrates that the equilibrium bulk density, $\bar{\rho}_b(r_c)$, of the MS-CG potential, $U(r_c)$, dramatically increases as r_c increases from 0.50 nm to 0.58 nm. While for $r_c < 0.49$ nm the model vaporizes, the MS-CG model quantitatively reproduces the AA bulk density at $r_c^* = 0.5234$ nm, which is slightly larger than the length-scale, $r_c \approx 0.515$ nm, that minimizes $\bar{\chi}^2$. As r_c increases further, the CG potential becomes increasingly attractive and the bulk density continues to

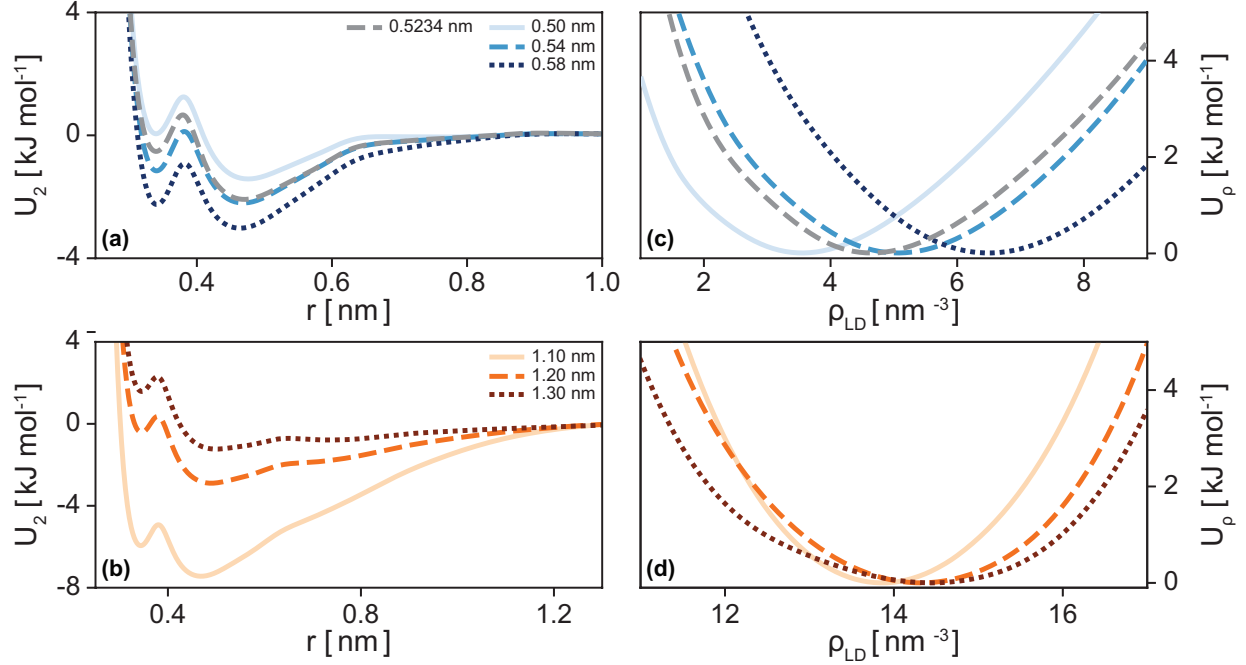


FIG. 4. The impact of the LD length-scale, r_c , upon the MS-CG potentials, $U(r_c) = \{U_2(r_c), U_\rho(r_c)\}$, calculated for the one-site methanol model at the reference state point, $T_{\text{ref}} = 260$ K and $P_{\text{ext}} = 1$ bar. The left panels (a and b) present the MS-CG pair potentials, $U_2(r_c)$, while the right panels (c and d) present the MS-CG LD potentials, $U_\rho(r_c)$. As in Fig. 2, the blue and orange curves correspond to relatively short and long LD length-scales, respectively, while the intensity of the color increases as r_c increases. The dashed grey curve in panels (a) and (b) indicate the MS-CG pair and LD potential, respectively, calculated for the optimal LD length-scale, $r_c^* = 0.5234$ nm.

increase until it reaches a maximum of $\bar{\rho}_{b;\text{max}} = 22.57 \text{ nm}^{-3}$ when $r_c = r_{c;\text{max}} \equiv 1.10$ nm.

The bottom row of Figure 4 presents calculated potentials for the relatively long length-scales, $r_c = 1.10$ nm, 1.30 nm, and 1.50 nm. As r_c increases past 1.10 nm, the MS-CG potential becomes significantly less attractive. In particular, the pair potential becomes much more repulsive. Moreover, while the minimum of the LD potential very slightly shifts to higher densities, the LD potential more dramatically tilts towards lower local densities. Consequently, Fig. 3b demonstrates that the equilibrium density, $\bar{\rho}_b(r_c)$, of the MS-CG model decreases as r_c increases in this regime. When $r_c = 1.842$ nm the CG model again reproduces the bulk density of the AA model.

In the following, we adopt the shortest length-scale, $r_c^* = 0.5234$ nm, for which the MS-CG potential, $U(r_c)$, reproduces the bulk density. The dashed vertical line in Fig. 2a indicates this optimal LD length-scale. Rather strikingly, the optimal LD length-scale, r_c^* , corresponds very closely to the shoulder in the second peak of the mapped RDF. Conversely, the optimal

LD length-scale is considerably shorter than the first solvation shell that is defined by either the second minimum of the mapped RDF or by the SANN algorithm of van Meel et al.⁹⁹ This short length-scale not only maximizes the simulated efficiency of the MS-CG model, but also allows the model to more accurately describe the local structure of inhomogeneous interfacial environments.⁸⁸

2. Accuracy and Transferability

We next assess the accuracy of the optimal MS-CG potential $U(r_c^*)$ at the state point for which it was parameterized. The solid grey curve in Fig. 5a reproduces the mapped RDF from Fig. 2a. The solid grey curve in Fig. 5b presents the pressure-density equation of state (eos) obtained from the AA constant NPT simulation at the reference state point, $T_{\text{ref}} = 260$ K and $P_{\text{ext}} = 1$ bar. The dashed grey curves in Figs. 5a and 5b present the RDF and pressure-density eos obtained from a constant NPT simulation of the optimized MS-CG potential, $U(r_c^*)$, at this reference state point. The MS-CG model reproduces the mapped AA RDF with nearly quantitative accuracy. While the CG model reproduces the bulk density by construction, it also quite accurately reproduces the AA pressure-density equation of state. In particular, Table II demonstrates that the MS-CG model overestimates the AA compressibility by approximately 30%. For comparison, IBI models that employ a pressure-correction can overestimate the AA compressibility by more than 300%.^{119,120} Thus, the optimized MS-CG potential, $U(r_c^*)$, quite accurately models the state point for which it was parameterized.

Figure 5 also assesses the transferability of the MS-CG potential, $U(r_c^*)$, that was optimized for the reference temperature, $T_{\text{ref}} = 260$ K. The solid blue and red curves in Fig. 5 present the results of AA constant NPT simulations at $P_{\text{ext}} = 1$ bar for $T_{\text{lo}} = 220$ K and $T_{\text{hi}} = 300$ K, respectively. In particular, Fig. 5a demonstrates that the first peak of the AA RDF slightly shifts to larger distances and also slightly broadens as T increases at constant external pressure. Similarly, Fig. 5b and Table I demonstrate that the equilibrium bulk density of the AA methanol model decreases from 16.24 nm^{-3} to 14.62 nm^{-3} as the temperature increases from 220 K to 300 K at 1 bar pressure.

The dashed blue and red curves in Fig. 5 present the results of CG constant NPT simulations with the fixed MS-CG potential, $U(r_c^*)$, at the corresponding state points. Simulations

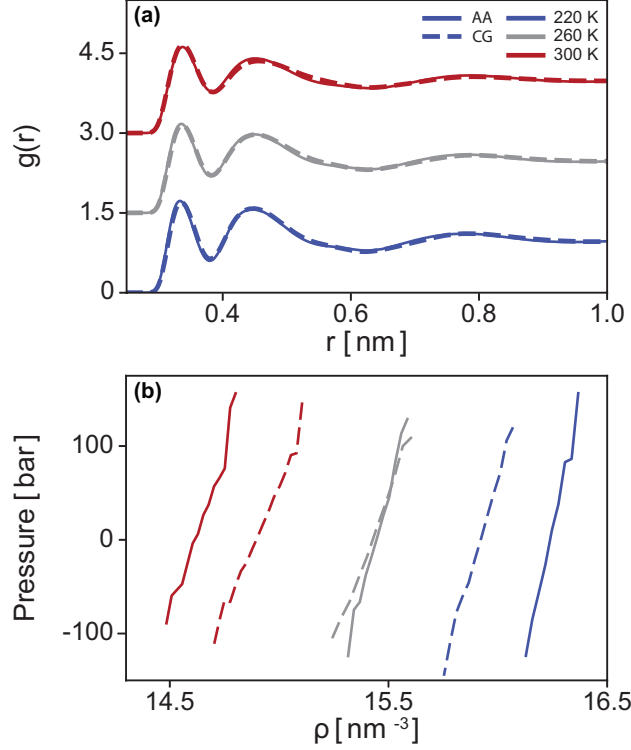


FIG. 5. Accuracy and transferability of the fixed MS-CG potential, $U(r_c^*) = \{U_2(r_c^*), U_\rho(r_c^*)\}$, calculated for the one-site methanol model at the reference state point, $T_{\text{ref}} = 260$ K and $P_{\text{ext}} = 1$ bar, while employing the optimal LD length-scale, $r_c^* = 0.5234$ nm. Panel (a) compares the mapped AA RDF (solid curve) with the RDF obtained from simulations with the MS-CG potential (dashed curve). Panel (b) compares the pressure-density equations of state obtained from constant NPT simulations at an external pressure $P_{\text{ext}} = 1$ bar with the AA (solid curve) and MS-CG model (dashed curve). The grey curves present results for constant NPT simulations with $P_{\text{ext}} = 1$ bar at the reference temperature, $T_{\text{ref}} = 260$ K, while the blue and red curves present results for constant NPT simulations with $P_{\text{ext}} = 1$ bar at $T = 220$ K and 300 K, respectively.

with the fixed MS-CG potential quite accurately reproduce the AA RDF at both temperatures. These simulations also reasonably describe the AA pressure-density eos at both temperatures. In particular, constant NPT simulations with the fixed MS-CG potential, $U(r_c^*)$, underestimate the AA bulk density at $T_{\text{lo}} = 220$ K by approximately 2.0%, while overestimating the AA bulk density at $T_{\text{hi}} = 300$ K by approximately 2.2%. Moreover, Table II demonstrates that the fixed MS-CG potential overestimates the AA compressibility at $T_{\text{lo}} = 220$ K by approximately 40%, and overestimates the AA compressibility at $T_{\text{hi}} = 300$ K by approximately 21%. Thus, the CG potential $U(r_c^*)$ provides rather good, although not perfect, transferability across a temperature range of 80 K at ambient pressure.

One approach for improving the transferability of bottom-up potentials is to explicitly

TABLE II. Equilibrium bulk density, $\bar{\rho}_b$, and isothermal compressibility, $\kappa_T = -\bar{V}^{-1}(\partial\bar{V}/\partial P)_T$, obtained from constant NPT simulations of methanol with AA and CG models at the temperature, T , and $P_{\text{ext}} = 1$ bar. The “MS-CG” column presents results for the fixed potential, $U(r_c^*)$, that is optimized for $T_{\text{ref}} = 260$ K in Sec. IV A 1 and assessed at each temperature in Sec IV A 2. The “Pred” column presents results for the predicted potential, $U_{\text{pred}}(T)$, that is defined and assessed in Sec. IV A 3. Temperatures are reported in units of K, densities are reported in units of nm^{-3} , and compressibilities are reported in units of 10^{-5} bar^{-1} .

T	$\bar{\rho}_b$			κ_T		
	AA	MS-CG	Pred	AA	MS-CG	Pred
220	16.24	15.92	16.21	5.46	7.60	5.57
260	15.44	15.43	-	7.32	9.53	-
300	14.62	14.94	14.63	10.39	12.59	11.06

model their state-point dependence. Because structure-based pair potentials often vary linearly with both temperature and density,^{30,33,59,62–67} we investigated whether LD potentials might also vary linearly with temperature. Specifically, we considered constant pressure ensembles with $P_{\text{ext}} = 1$ bar at $T_{\text{lo}} = 220$ K and $T_{\text{hi}} = 300$ K, which correspond to the blue and red curves, respectively, in Fig. 5. We performed independent FM variational calculations to determine the MS-CG potential, $U(T, r_c^*) = \{U_2(T, r_c^*), U_\rho(T, r_c^*)\}$, for modeling methanol at each additional state point. Importantly, we determined $U(T_{\text{lo}}, r_c^*)$ and $U(T_{\text{hi}}, r_c^*)$ using the same LD length-scale, $r_c^* = 0.5234$ nm, that was optimized for $T_{\text{ref}} = 260$ K.

Figure 6a presents the resulting pair potentials, $U_2(T) = U_2(T, r_c^*)$. As the temperature increases (at constant external pressure), the H-bonding minimum of $U_2(T)$ slightly decreases, while the vdW minimum more noticeably increases. Figure 6b presents the corresponding finite difference, $\Delta U_2/\Delta T$, where $\Delta U_2 \equiv U_2(T_{\text{hi}}) - U_2(T_{\text{lo}})$, while $\Delta T \equiv T_{\text{hi}} - T_{\text{lo}}$.

The right panels of Fig. 6 consider the corresponding LD potentials. Figure 6c presents the distribution of local densities, $P(\rho_{\text{LD}}) = P(\rho_{\text{LD}}; T, r_c^*)$, that are sampled by the AA model at each state point. The distribution of local densities broadens and shifts to lower densities with increasing temperature (at constant external pressure). Given that the local density is defined by molecules within the first solvation shell, i.e., $r \leq r_c^* = 0.5234$ nm, it is notable that $P(\rho_{\text{LD}})$ quite closely mirrors the shift in the global density. The SM presents the ratio, $\eta(T) \equiv \bar{\rho}_{\text{LD}}(T)/\bar{\rho}_{b;\text{AA}}(T)$, of the average local density, $\bar{\rho}_{\text{LD}}(T)$, and the equilibrium bulk density, $\bar{\rho}_{b;\text{AA}}(T)$, that are sampled by the AA model. The SM demonstrates that this ratio is almost independent of temperature for methanol. (However, the SM also demonstrates

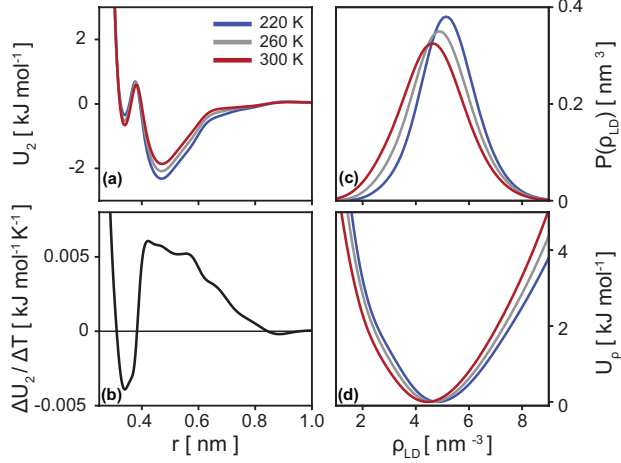


FIG. 6. Temperature-dependence of the MS-CG potentials, $U(T) = U(T, r_c^*)$, calculated from AA constant NPT simulations at ambient pressure, $P_{\text{ext}} = 1$ bar, and various temperatures, T , while employing the fixed LD length-scale, $r_c^* = r_c^*(T_{\text{ref}}) = 0.5234$ nm, that was optimized for simulations at $T_{\text{ref}} = 260$ K. Panels (a) and (d) present the calculated pair potentials, $U_2(T) = U_2(T, r_c^*)$, and calculated LD potentials, $U_\rho(T) = U_\rho(T, r_c^*)$, respectively. Panel (b) presents the finite difference in the calculated pair potentials, $\Delta U_2 / \Delta T$, where $\Delta U_2 = U_2(T_{\text{hi}}) - U_2(T_{\text{lo}})$ and $\Delta T = T_{\text{hi}} - T_{\text{lo}}$, while $T_{\text{hi}} = 300$ K and $T_{\text{lo}} = 220$ K. Panel (c) presents the local density distributions, $P(\rho_{\text{LD}}) = P(\rho_{\text{LD}}; T, r_c^*)$, calculated from the mapped ensembles for each temperature, while employing the fixed LD length-scale, r_c^* . The grey curves present results for constant NPT simulations with $P_{\text{ext}} = 1$ bar at the reference temperature, $T_{\text{ref}} = 260$ K, while the blue and red curves present results for constant NPT simulations with $P_{\text{ext}} = 1$ bar at $T = 220$ K and 300 K, respectively.

that this trend does not hold for OTP.)

Figure 6d presents the calculated LD potentials, $U_\rho(T, r_c^*)$. The LD potentials shift to lower local densities with increasing temperature, following the shift in $P(\rho_{\text{LD}})$. Thus, the pair and LD potentials both become increasingly repulsive with increasing temperature (at constant external pressure).

Figures 7a and 7b report the RDF and pressure-density eos, respectively, obtained from a constant NPT simulation with the MS-CG potential, $U(T, r_c^*)$, at each temperature. As before, the MS-CG potentials quite accurately reproduce the mapped RDF for each temperature. However, the MS-CG potentials $U(T_{\text{lo}}, r_c^*)$ and $U(T_{\text{hi}}, r_c^*)$ do not so accurately reproduce the AA pressure-density eos at $T_{\text{lo}} = 220$ K and $T_{\text{hi}} = 300$ K. In particular, these temperature-dependent MS-CG potentials overestimate the equilibrium density at $T_{\text{lo}} = 220$ K by 5.2% and underestimate the equilibrium density at $T_{\text{hi}} = 300$ K by 2.9%. In comparison, constant NPT simulations with the fixed MS-CG potential $U(T_{\text{ref}}, r_c^*)$ only underestimated the bulk density at T_{lo} by 2.0% and only overestimated the bulk density at

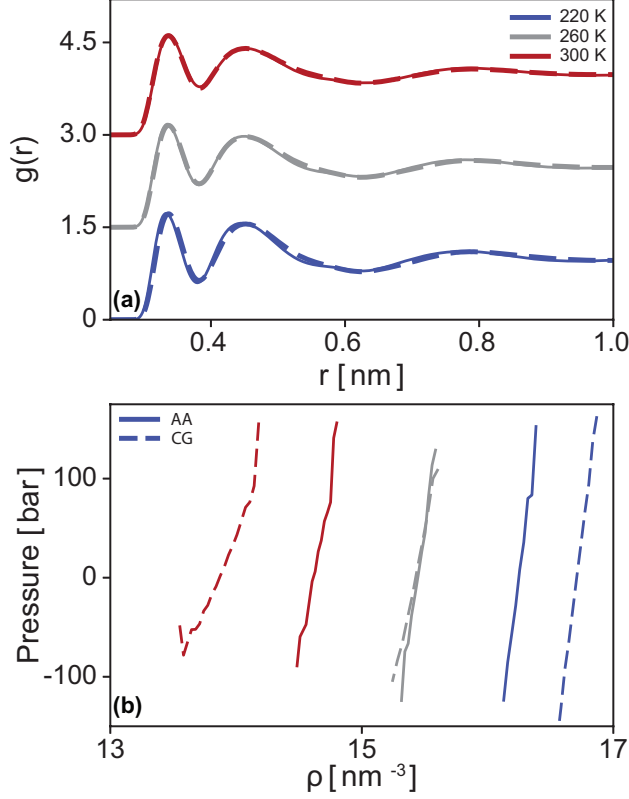


FIG. 7. Assessment of the temperature-dependent MS-CG potentials, $U(T) = U(T, r_c^*)$, that were independently calculated for each temperature, while employing the fixed LD length-scale, $r_c^* = r_c^*(T_{\text{ref}}) = 0.5234$ nm. Panels (a) and (b) present the mass-center RDF and pressure-density equations of state obtained from constant NPT simulations at ambient external pressure, $P_{\text{ext}} = 1$ bar, and the temperature, T . The solid curves present the results of AA simulations and correspond to the solid curves in Fig. 5. The dashed curves present the results of simulations with the temperature-dependent MS-CG potentials, $U(T) = U(T, r_c^*)$. The grey curves present results for the reference temperature, $T_{\text{ref}} = 260$ K, while the blue and red curves present results for $T = 220$ K and 300 K, respectively.

T_{hi} by 2.2%. Thus, independently calculating the MS-CG potential for each temperature actually reduced the accuracy of the model when r_c^* was held fixed. This presumably occurs because the LD length-scale, r_c^* , that accurately described the reference state point, $T_{\text{ref}} = 260$ K, does not accurately describe $T_{\text{lo}} = 220$ K and $T_{\text{hi}} = 300$ K.

3. Temperature-dependent length-scale

Accordingly, we next investigated the temperature-dependence of the optimal LD length-scale. We considered the constant NPT ensemble with $P_{\text{ext}} = 1$ bar for the four additional temperatures indicated in Table I, i.e., $T = 220$ K, 240 K, 280 K, and 300 K. At each of these

state points, we employed the FM variational principle to determine the MS-CG potential $U(T, r_c) = \{U_2(T, r_c), U_\rho(T, r_c)\}$ for a wide range of LD length-scales, r_c . We simulated each MS-CG potential, $U(T, r_c)$, in the constant NPT ensemble at the temperature, T , and external pressure, $P_{\text{ext}} = 1$ bar. At each temperature, T , we determined the equilibrium bulk density, $\bar{\rho}_b(T, r_c)$, of the MS-CG potential, $U(T, r_c)$, as a function of r_c , as illustrated in Fig. 3b for the reference temperature, $T_{\text{ref}} = 260$ K. We define the optimal LD length-scale, $r_c^*(T)$, for each T as the shortest length-scale for which the corresponding MS-CG potential, $U(T, r_c^*(T))$, accurately reproduces the equilibrium bulk density of the AA model, $\bar{\rho}_{b;\text{AA}}(T)$. We consider $U(T) = U(T, r_c^*(T))$ to be the optimal MS-CG potential for the temperature, T .

Figure 8a presents the optimal LD length-scale, $r_c^*(T)$, as a function of temperature. Over this 80 K temperature range, the optimal length-scale varies by only 0.0015 nm, i.e., $\approx 3\%$. While this variation appears quite small, Fig. 3b demonstrates that the simulated bulk density of the CG model varies rapidly with r_c near its optimal value. As might be expected, $r_c^*(T)$ increases as methanol expands with increasing T at constant $P_{\text{ext}} = 1$ bar. The temperature-dependent LD length-scale, $\{r_c^*(T)\}$, is described almost perfectly ($R^2 = 0.9998$) by the line:

$$f(T) = m(T - T_{\text{ref}}) + r_{c;0}^*, \quad (14)$$

where $m = 1.865 \times 10^{-4}$ nm/K, $T_{\text{ref}} = 260$ K, and $r_{c;0}^* = r_c^*(T_{\text{ref}}) = 0.5234$ nm. For comparison, the SM demonstrates that the nearest-neighbor shell distance, $R_{\text{NN}}(T)$, also increases with temperature, but at a different rate. Moreover, the SM demonstrates that $m/r_{c;0}^*$ is quite distinct from the coefficient of thermal expansion, $\alpha = V^{-1}(\partial V/\partial T)_P$, of the AA model.

Figure 8b presents the LD distributions obtained by mapping each AA constant NPT ensemble to the CG resolution, while using the optimal LD length-scale, $r_c^*(T)$, for each state point. Because a different length-scale, $r_c^*(T)$, is used for each T , a different local density is being computed for each state point, i.e., $\rho_I(\mathbf{R}; T) = \rho_I(\mathbf{R}; r_c^*(T))$ and $P(\rho_{\text{LD}}; T) = P(\rho_{\text{LD}}; T, r_c^*(T))$. As expected, the LD distribution broadens with increasing temperature. However, we find it rather remarkable that the LD distribution, $P(\rho_{\text{LD}}; T)$, does not visibly shift with T when using the optimal LD length-scale, $r_c^*(T)$, for each T . This suggests that the mapped ensemble for methanol may be characterized by some underlying temperature-

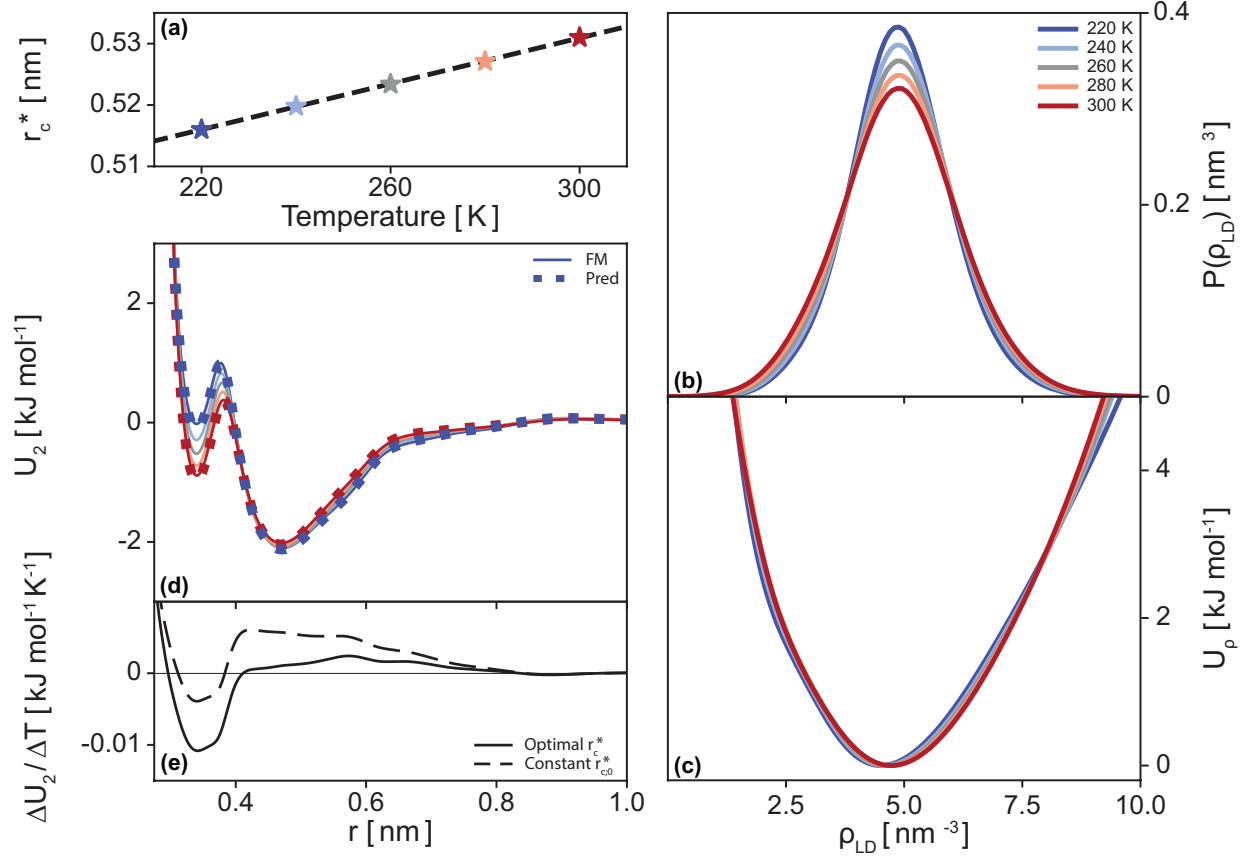


FIG. 8. Analysis of the optimal LD length-scales, $r_c^*(T)$, determined from constant NPT simulations of methanol at ambient pressure, $P_{\text{ext}} = 1$ bar, and various temperatures, T . Panel (a) plots the optimal LD length-scale, $r_c^*(T)$, as a function of temperature, as well as the corresponding best fit line. Panel (b) presents the mapped AA LD distribution, $P(\rho_{\text{LD}}) = P(\rho_{\text{LD}}; T, r_c^*(T))$, calculated for each temperature, T , while employing the optimal LD length-scale, $r_c^*(T)$. Panels (c) and (d) present the LD potential, $U_\rho(T) = U_\rho(T, r_c^*(T))$, and pair potential, $U_2(T) = U_2(T, r_c^*(T))$, calculated independently for each T , while employing the corresponding optimal length-scale, $r_c^*(T)$. The solid curve in panel (e) presents the finite difference in these pair potentials, $\Delta U_2 / \Delta T$, where $\Delta U_2 = U_2(T_{\text{hi}}) - U_2(T_{\text{lo}})$ and $\Delta T = T_{\text{hi}} - T_{\text{lo}}$, while $T_{\text{hi}} = 300$ K and $T_{\text{lo}} = 220$ K. For comparison, the dashed curve in panel (e) reproduces from Fig. 6b the finite difference in the temperature-dependent pair potentials obtained with a single fixed LD length-scale, $r_{c;0}^* = r_c^*(T_{\text{ref}}) = 0.5234$ nm. The grey curves present results for the reference temperature, $T_{\text{ref}} = 260$ K, while the blue, light blue, orange, and red curves present results for $T = 220$ K, 240 K, 280 K and 300 K, respectively.

dependent length scale. (Our results for OTP indicate that this trend does not always hold.)

The remainder of Fig. 8 analyzes the optimal MS-CG potential, $U(T) = \{U_2(T), U_\rho(T)\}$, that was calculated for each T , while employing the optimal LD length-scale, $r_c^*(T)$. Figure 8c presents the optimized MS-CG LD potentials, $U_\rho(T) = U_\rho(T, r_c^*(T))$, for the five state points. Remarkably, these five LD potentials are almost indistinguishable. They very

slightly shift to higher local densities and also appear to demonstrate slightly greater curvature with increasing T . Nevertheless, these variations are much smaller than the variations observed in Fig. 6d among the LD potentials, $U_\rho(T, r_{c;0}^*)$, that were calculated with the fixed LD length-scale, $r_{c;0}^* = r_c^*(T_{\text{ref}}) = 0.5234$ nm.

The solid curves in Fig. 8d present the optimized MS-CG pair potentials, $U_2(T) = U_2(T, r_c^*(T))$, for the five state points. As T increases (at constant $P_{\text{ext}} = 1$ bar), the H-bonding minimum decreases, while the vdW minimum slightly rises and sharpens. The solid curve in Fig. 8e quantifies this temperature dependence with the finite difference $\Delta U_2/\Delta T$, where $\Delta U_2 \equiv U_2(T_{\text{hi}}, r_c^*(T_{\text{hi}})) - U_2(T_{\text{lo}}, r_c^*(T_{\text{lo}}))$ and $\Delta T \equiv T_{\text{hi}} - T_{\text{lo}}$, as before. For comparison, the dashed curve presents the corresponding finite difference for the potentials, $U_2(T, r_{c;0}^*)$, in Fig. 6a, which employed a fixed LD cut-off, $r_{c;0}^*$. The calculated pair potentials demonstrate qualitatively similar, but greater temperature-dependence, when calculated with the optimal temperature-dependent LD length-scales, $r_c^*(T)$.

The dashed curves in Fig. 9 assess the accuracy of the optimized MS-CG potentials, $U(T) = U(T, r_c^*(T))$, for modeling $T = T_{\text{lo}}, T_{\text{ref}}$, and T_{hi} . The optimized MS-CG potentials reproduce the AA RDF and pressure-density eos with nearly quantitative accuracy at each state point. The SM demonstrates that the optimized MS-CG potentials provide similar accuracy at the two other temperatures considered, i.e., $T = 240$ K and 280 K.

Figure 8 suggests a framework for predicting CG potentials, $U_{\text{pred}}(T) = \{U_{2;\text{pred}}(T), U_{\rho;\text{pred}}(T)\}$, that accurately model new temperatures, T , without performing additional AA simulations. Specifically, Fig. 8a suggests that the optimal LD length-scale, $r_c^*(T)$, can be accurately predicted by the dashed best fit line, $r_{c;\text{pred}}(T) = f(T)$, that is given by Eq. (14). Figure 8c suggests that if the local density is defined with this predicted length-scale, $\rho_I(\mathbf{R}; T) = \rho_I(\mathbf{R}; r_{c;\text{pred}}(T))$, then the same LD potential can be used to model the constant NPT ensemble at $P_{\text{ext}} = 1$ bar and any temperature, T . More explicitly, we predict the potential, $U_{\text{LD};\text{pred}}(\mathbf{R}; T)$, by

$$U_{\text{LD};\text{pred}}(\mathbf{R}; T) = \sum_{I=1}^N U_\rho(\rho_I(\mathbf{R}; T); T_{\text{ref}}), \quad (15)$$

where $U_\rho(\rho_I; T_{\text{ref}}) = U_\rho(\rho_I; T_{\text{ref}}, r_{c;0}^*)$ is the LD potential originally optimized for T_{ref} , while $\rho_I(\mathbf{R}; T)$ indicates the local density calculated in configuration $\mathbf{R}$ with the temperature-dependent length-scale, $r_c^*(T)$. For simplicity, we denote this prediction by $U_{\rho;\text{pred}}(T)$. Fi-

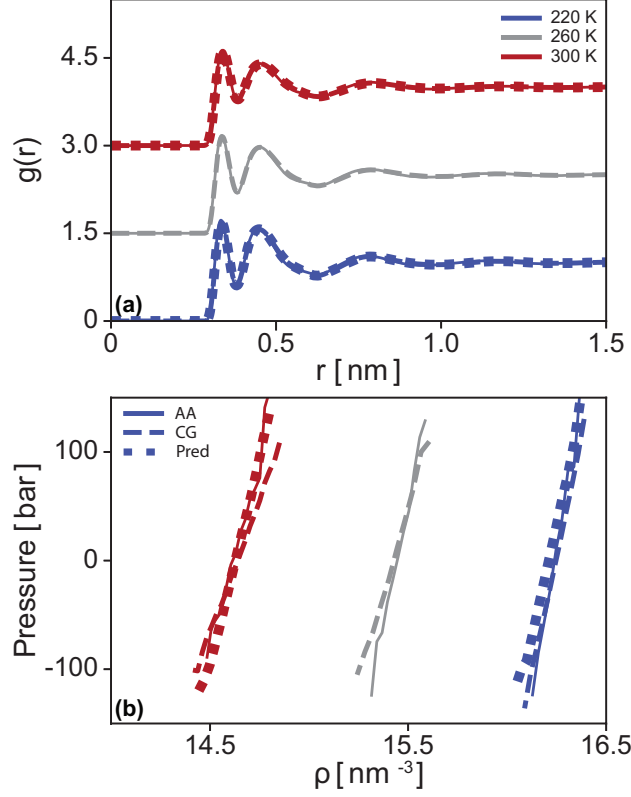


FIG. 9. Analysis of CG potentials, $U(T) = U(T, r_c^*(T))$, with temperature-dependent LD length-scales, $r_c^*(T)$. Panels (a) and (b) present the mass-center RDF and pressure-density eos obtained from a constant NPT simulation of methanol at ambient external pressure, $P_{\text{ext}} = 1$ bar, and each temperature, T . The solid curves present the results of AA simulations and correspond to the solid curves in Fig. 5. The dashed curves present the results of simulations with temperature-dependent MS-CG potentials, $U(T) = U(T, r_c^*(T))$, that were independently optimized for each temperature. The dotted curves present the results of simulations with the predicted CG potential, $U_{\text{pred}}(T) = \{U_{2;\text{pred}}(T), U_{\rho;\text{pred}}(T)\}$, for each temperature. The grey curves present results for the reference temperature, $T_{\text{ref}} = 260$ K, while the blue and red curves present results for $T = 220$ K and 300 K, respectively.

nally, Fig. 8d and 8e suggest that the optimal pair potential, $U_2(T, r_c^*(T))$, can be predicted from simple linear extrapolation:

$$U_{2;\text{pred}}(T) = U_{2;0} + (T - T_{\text{ref}})\Delta U_2/\Delta T. \quad (16)$$

where $U_{2;0} = U_2(T_{\text{ref}}, r_{c;0}^*)$ is the pair potential optimized for the reference state point, and $\Delta U_2/\Delta T$ is indicated by the solid curve in Fig. 8e. The dotted curves in Fig. 8d demonstrate that the predicted pair potentials accurately match the optimized MS-CG pair potentials for both T_{lo} and T_{hi} . The SM demonstrates that this linear extrapolation also accurately

predicts the optimal MS-CG pair potentials for the intermediate temperatures, $T = 240$ K and 280 K.

The dotted curves in Fig. 9 demonstrate the accuracy of the predicted potentials, $U_{\text{pred}}(T) = \{U_{2;\text{pred}}(T), U_{\rho;\text{pred}}(T)\}$, in constant NPT simulations with $P_{\text{ext}} = 1$ bar at both T_{lo} and T_{hi} . The predicted potentials reproduce the AA RDF and pressure-density eos with nearly quantitative accuracy. In particular, Table II demonstrates that the predicted potentials more accurately reproduce the bulk density and, especially, the compressibility of the AA model than the MS-CG potential, $U(T_{\text{ref}})$, that was optimized for the reference state point. The SM demonstrates that the predicted potentials also provide similar accuracy for modeling $T = 240$ K and 280 K.

B. Ortho-Terphenyl

The preceding section demonstrated that, if one defines the local density with a temperature-dependent length-scale, then one-site CG models for methanol can employ a single LD potential across a rather wide range of temperatures. In order to investigate the generality of these findings, we now consider OTP. As illustrated by the inset of Fig. 10a, OTP is a nearly planar molecule consisting of three connected benzene rings. Consequently, the one-site CG model for OTP employs a much lower resolution than the one-site CG model for methanol. While each CG site corresponded to two heavy atoms in the one-site methanol model, each site corresponds to 18 heavy atoms in the one-site OTP model. Moreover, the OTP liquid phase is stable over a much larger temperature range than the methanol liquid phase.

Figure 10 characterizes the mapped ensemble obtained from AA constant NPT simulations of 800 OTP molecules at $P_{\text{ext}} = 1$ bar and the reference temperature, $T_{\text{ref}} = 575$ K. Our prior study indicated that the simulated AA model for OTP undergoes a glass transition near $T_g \approx 340$ K.⁶⁴ Consequently, we expect that the AA OTP model behaves as a well-equilibrated liquid at this reference state point. It is worth noting that this AA model overestimates experimental measurements for the density of liquid OTP by approximately 2-4% and, moreover, significantly underestimates the observed glass transition temperature.^{64,121}

Figure 10a presents the relatively featureless mapped RDF. The mapped RDF vanishes for $r < 0.4$ nm, exhibits a very broad first peak with a maximum near $r \approx 0.8$ nm, and an almost imperceptible second peak near $r \approx 1.4$ nm. It is perhaps worth noting that

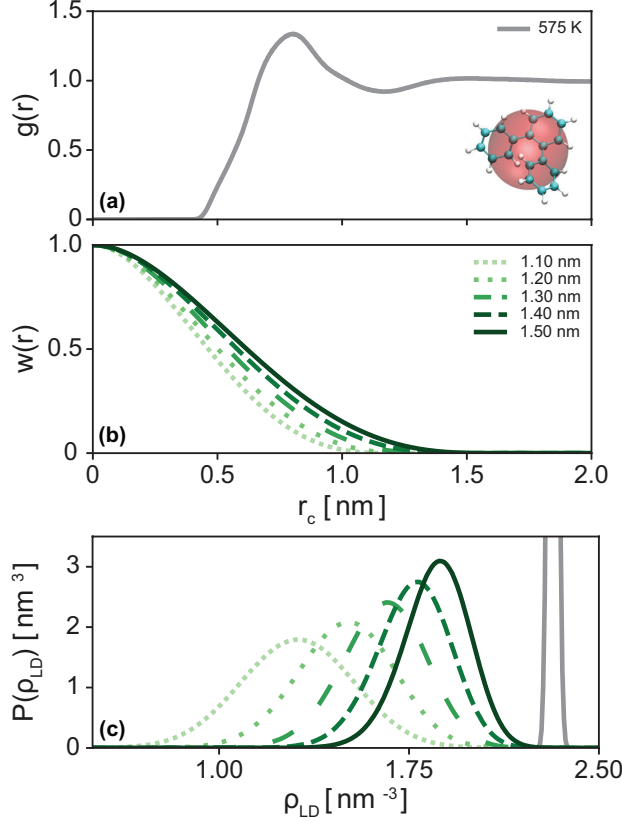


FIG. 10. Analysis of the reference mapped ensemble obtained from an AA constant NPT simulation of OTP at the reference temperature, $T_{\text{ref}} = 575$ K, and the external pressure, $P_{\text{ext}} = 1$ bar. Panel (a) presents the mapped RDF, as well as the mapped representation of each OTP molecule. Panel (b) presents the Lucy function, $w(r) = w(r; r_c)$, for several different values of the LD length-scale, r_c . Panel (c) presents the corresponding LD distributions, $P(\rho_{\text{LD}}) = P(\rho_{\text{LD}}; r_c)$, calculated from the reference mapped ensemble for the various LD length-scales. In panels (b) and (c), the intensity of the colored curves increases as r_c increases. The solid grey curve in panel (c) presents the distribution of bulk densities, ρ_b , sampled by the AA constant NPT simulation.

contacting OTP molecules interact via short-ranged interactions between benzene rings that are quite distant from the location of the CG sites used to represent the molecules.

Figure 10b presents the Lucy function, $w(r) = w(r; r_c)$, used to define the local density for five different length-scales, r_c , ranging from $r_c = 1.10$ nm to 1.50 nm. Figure 10c presents the corresponding mapped LD distribution, $P(\rho_{\text{LD}}) = P(\rho_{\text{LD}}; r_c)$, for each length-scale. The grey curve in Fig. 10c presents the distribution of bulk densities sampled by the AA constant NPT simulation. As r_c increases from 1.10 nm to 1.50 nm, the LD distribution shifts toward the bulk density distribution and also sharpens.

We employed the FM variational principle to determine the MS-CG potential, $U(T, r_c) = \{U_2(T, r_c), U_\rho(T, r_c)\}$, for a wide range of LD length-scales, r_c , at the reference temperature,

$T_{\text{ref}} = 575$ K, and also at three additional temperatures, $T = 400$ K, 500 K, and 650 K. We performed constant NPT simulations (with $P_{\text{ext}} = 1$ bar) with each MS-CG potential, $U(T, r_c)$, at the corresponding temperature. Figure 11 presents the equilibrium density, $\bar{\rho}_b(r_c) = \bar{\rho}_b(T, r_c)$, for each MS-CG potential, $U(T, r_c)$, as a function of r_c at each temperature. The dashed horizontal curves in Fig. 11 present the equilibrium bulk density, $\bar{\rho}_{b;\text{AA}}(T)$, of the AA model at the corresponding state point.

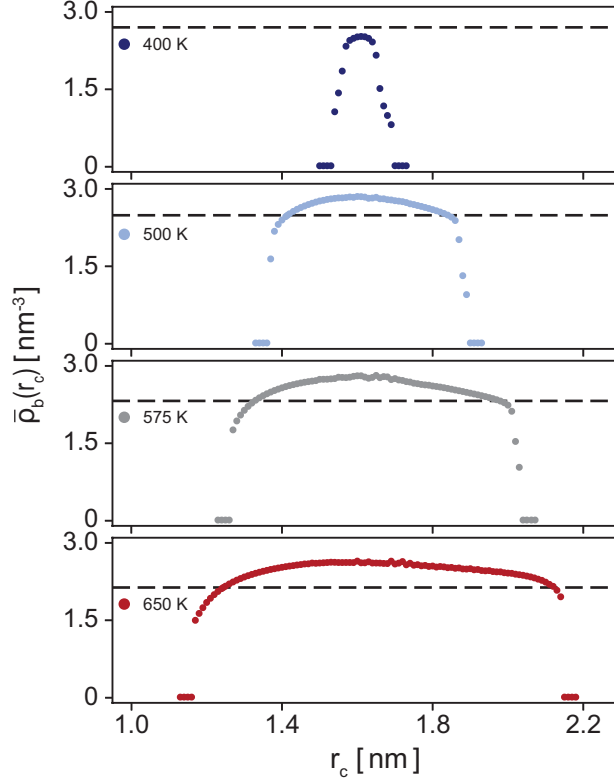


FIG. 11. Impact of the LD length-scale, r_c , upon the equilibrium bulk density of one-site MS-CG potentials, $U(r_c) = \{U_2(r_c), U_\rho(r_c)\}$, for OTP. Each point indicates the equilibrium bulk density, $\bar{\rho}_b(r_c)$, determined from a constant NPT simulation with the CG potential, $U(r_c)$, at ambient pressure, $P_{\text{ext}} = 1$ bar. The horizontal dashed line indicates the equilibrium bulk density, $\bar{\rho}_{b;\text{AA}}$, of the AA model for OTP at the same state point. Each panel corresponds to the indicated fixed temperature.

Figure 11 recapitulates many of the features observed in Fig. 3b for methanol. When r_c is smaller than some threshold value, r_{c1} , the LD potential does not act on sufficient molecules and the CG model vaporizes. As r_c increases past this threshold, the bulk density of the CG model rapidly increases. As r_c increases further, the bulk density of the CG model increases less rapidly before eventually plateauing near a maximum density, $\bar{\rho}_{b;\text{max}}$, that occurs when $r_c = r_{c;\text{max}}$. Beyond this point, the bulk density begins to gradually decrease until r_c again

approaches another threshold value, beyond which the bulk density very rapidly decreases. When r_c is greater than some critical value, r_{c2} , the MS-CG potential again vaporizes.

Figure 11 also indicates several additional interesting trends. As T decreases from 650 K, r_{c1} increases, while r_{c2} decreases. Consequently, the window of LD length-scales, r_c , for which the MS-CG potential, $U(T, r_c)$, stabilizes a condensed phase shrinks as T decreases. Moreover, as T decreases, the gap, $\Delta\bar{\rho}_{\max}(T) = \bar{\rho}_{b;\max}(T) - \bar{\rho}_{b;AA}(T)$, between the maximum density of the CG model, $\bar{\rho}_{b;\max}(T)$, and the target AA density, $\bar{\rho}_{b;AA}(T)$, also shrinks. In particular, we find that no calculated MS-CG potential, $U(T, r_c)$, reproduces the AA bulk density at $T_x = 400$ K.

For $T = 500$ K, 575 K, and 650 K, we define the optimal LD length-scale, $r_c^*(T)$, as the shortest length-scale for which the MS-CG potential, $U(T, r_c^*(T))$, accurately reproduces the AA bulk density. For these temperatures, we consider $U(T) = U(T, r_c^*(T))$ to be the optimal MS-CG potential for OTP at the temperature, T . However, the MS-CG potentials, $U(T_x, r_c)$, for $T_x = 400$ K always underestimate the equilibrium density of the AA model at $T_x = 400$ K. In this case, we define the optimal LD length scale to be $r_{c;\max}(T_x)$ and the optimal MS-CG potential to be the potential, $U(T_x) = U(T_x, r_{c;\max}(T_x))$, that most closely reproduces the AA bulk density at $T_x = 400$ K.

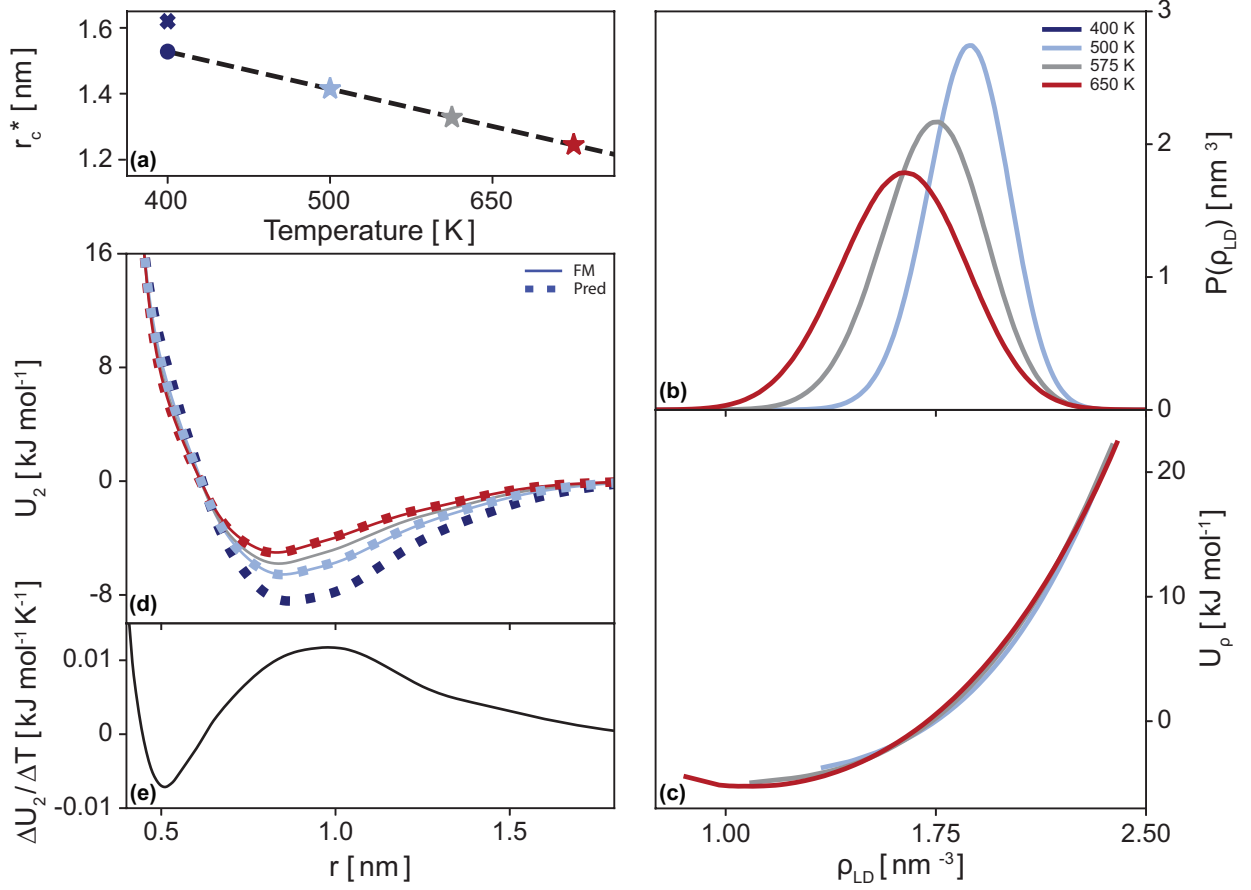


FIG. 12. Analysis of the optimal LD length-scales, $r_c^*(T)$, determined from constant NPT simulations of OTP at ambient pressure, $P_{\text{ext}} = 1$ bar, and various temperatures, T . The stars in panel (a) present the optimal LD length-scale, $r_c^*(T)$, for the indicated temperatures, while the dashed line indicates the best fit to these points. The dark blue ‘x’ in panel (a) presents the LD length-scale, $r_{c;\text{max}}(T_x)$, for the MS-CG potential, $U(T_x) = U(T_x, r_{c;\text{max}}(T_x))$, that most closely reproduces the AA density at the temperature, $T_x = 400$ K. The dark blue circle indicates the LD length-scale, $r_{c;\text{pred}}(T_x)$, predicted for T_x by the best fit line. Panel (b) presents the mapped AA LD distribution, $P(\rho_{\text{LD}}) = P(\rho_{\text{LD}}; T, r_c^*(T))$, calculated for each temperature, $T > T_x$, while employing the optimal LD length-scale, $r_c^*(T)$. Panels (c) and (d) present the LD potential, $U_\rho(T) = U_\rho(T, r_c^*(T))$, and pair potential, $U_2(T) = U_2(T, r_c^*(T))$, calculated independently for each $T > T_x$, while employing the corresponding optimal length-scale, $r_c^*(T)$. The solid curve in panel (e) presents the finite difference in these pair potentials, $\Delta U_2 / \Delta T$, where $\Delta U_2 = U_2(T_{\text{hi}}) - U_2(T_{\text{lo}})$ and $\Delta T = T_{\text{hi}} - T_{\text{lo}}$, while $T_{\text{hi}} = 650$ K and $T_{\text{lo}} = 500$ K. The dotted curves in panel (d) present the pair potentials that are obtained by extrapolating with this finite difference from the MS-CG pair potential, $U_2(T_{\text{ref}})$, that was optimized for the reference temperature, $T_{\text{ref}} = 575$ K. The grey curves present results for the reference temperature, $T_{\text{ref}} = 575$ K, while the dark blue, light blue, and red curves present results for $T = 400$ K, 500 K, and 650 K, respectively.

The stars in Fig. 12a present the optimal LD length-scales, $r_c^*(T)$, for $T = 500$ K, 575 K, and 650 K. As is clear from Fig. 11, the optimal LD length-scale for OTP decreases with temperature over this range. This trend runs counter to the more intuitive trend observed in Fig. 8 for methanol. Nevertheless, over the temperature-range $500 \text{ K} < T < 650 \text{ K}$, the optimal temperature-dependent LD length-scale for OTP is described almost perfectly ($R^2 = 0.9998$) by the line

$$f(T) = m(T - T_{\text{ref}}) + r_{c;0}^*, \quad (17)$$

where $m = -1.135 \times 10^{-3} \text{ nm/K}$ and $r_{c;0}^* = r_c^*(T_{\text{ref}}) = 1.3276 \text{ nm}$. The dark blue ‘X’ in Fig. 12a presents the optimal LD length-scale, $r_{c;\text{max}}(T_x)$, for the problematic temperature $T_x = 400 \text{ K}$. This length-scale deviates significantly from the best fit line for the length-scales, $r_c^*(T)$, that accurately reproduce the AA bulk density at $T = 500 \text{ K}$, 575 K, and 650 K. Consequently, we do not consider the temperature $T_x = 400 \text{ K}$ further in Fig. 12.

Figure 12b presents the mapped LD distribution, $P(\rho_{\text{LD}}; T) = P(\rho_{\text{LD}}; T, r_c^*(T))$, for $T = 500 \text{ K}$, 575 K, and 650 K. As in Fig. 8b for methanol, Fig. 12b employs the optimal cut-off, $r_c^*(T)$, at each temperature to define the local density, $\rho_I(\mathbf{R}; T) = \rho_I(\mathbf{R}; r_c^*(T))$. As in Fig. 8b, the mapped LD distribution broadens with increasing temperature. However, in contrast to Fig. 8b for methanol, the mapped LD distribution for OTP shifts to lower densities as T increases, mirroring the shift in the bulk density.

The remainder of Fig. 12 analyzes the optimized MS-CG potentials, $U(T) = U(T, r_c^*(T))$, for $T = 500 \text{ K}$, 575 K, and 650 K. These potentials were calculated using the optimal LD length-scale, $r_c^*(T)$, for each temperature. The SM demonstrates that the optimized MS-CG potential, $U(T_x) = U(T_x, r_{c;\text{max}})$, for $T_x = 400 \text{ K}$ deviates significantly from the MS-CG potentials that were optimized for these higher temperatures.

Figure 12c presents the optimized MS-CG LD potentials, $U_\rho(T) = U_\rho(T, r_c^*(T))$. Even though $P(\rho_{\text{LD}}; T)$ varies significantly across these state points, the optimized LD potentials again appear remarkably independent of T .

The solid curves in Fig. 12d present the optimized MS-CG pair potentials, $U_2(T) = U_2(T, r_c^*(T))$. In contrast to the optimized MS-CG pair potentials for methanol, the optimized pair potentials for OTP become slightly more repulsive as T increases (at constant external pressure, $P_{\text{ext}} = 1 \text{ bar}$). Fig. 12e presents the finite difference, $\Delta U_2 / \Delta T$, where $\Delta U_2 \equiv U_2(T_{\text{hi}}) - U_2(T_{\text{lo}})$ and $\Delta T = T_{\text{hi}} - T_{\text{lo}}$, while $T_{\text{hi}} = 650 \text{ K}$ and $T_{\text{lo}} = 500 \text{ K}$. The

dotted dark red and light blue curves in Fig. 12d present the linear extrapolation given by Eq. (16) for T_{hi} and T_{lo} , respectively. This linear extrapolation very accurately describes the optimized MS-CG pair potentials at both temperatures.

Figure 13 assesses the accuracy of the MS-CG models that have been optimized for OTP at each state point. The solid and dashed curves present the results of constant NPT simulations with the AA and optimized MS-CG models, respectively. In particular, the dark blue dashed curves present results with the optimized MS-CG potential, $U(T_x, r_{c;\text{max}})$, that most accurately reproduces the AA bulk density at the problematic temperature, $T_x = 400$ K.

TABLE III. Bulk density, $\bar{\rho}_b$ and isothermal compressibility, $\kappa_T = -\bar{V}^{-1}(\partial\bar{V}/\partial P)_T$ obtained for OTP from constant NPT simulations with the AA model and various CG models. The “MS-CG” column presents results for the optimal MS-CG potential, $U(T, r_c^*(T))$, at each temperature. The “Pred” column presents results for the potential, $U_{\text{pred}}(T)$, that is predicted for each temperature. Temperatures are reported in units of K, densities are reported in units of nm^{-3} , and compressibilities are reported in units of 10^{-5} bar^{-1} .

T	$\bar{\rho}_b$			κ_T		
	AA	MS-CG	Pred	AA	MS-CG	Pred
400	2.70	2.52	2.67	5.91	15.17	5.69
500	2.49	2.49	2.48	10.74	7.98	8.90
575	2.32	2.32	-	19.57	17.24	-
650	2.14	2.14	2.14	34.97	24.45	34.92

Figure 13a demonstrates that the optimized MS-CG models quite accurately reproduce the AA pair structure at each temperature. As the temperature decreases, the MS-CG model appears to become slightly more structured than the AA model. In particular, at $T_x = 400$ K, the MS-CG model slightly underestimates the AA RDF near $r \approx 0.5$ nm and also demonstrates slightly larger oscillations in the first two solvation shells. Nevertheless, the MS-CG models quite reasonably reproduce the AA RDF’s at all 4 temperatures.

Figure 13b demonstrates that the optimized MS-CG models quite accurately reproduce the AA pressure-density equations of state at $T = 500$ K, 575 K, and 650 K. The optimized MS-CG models quantitatively reproduce the bulk density by construction. Table III demonstrates that they underestimate the AA compressibility by only 12 - 30% for these three temperatures. Conversely, the dark blue dashed curve demonstrates that the optimized

MS-CG model at $T_x = 400$ K describes the AA pressure-density eos with considerably less accuracy. In particular, this MS-CG model underestimates the AA bulk density by 7.1% and badly overestimates the AA compressibility.

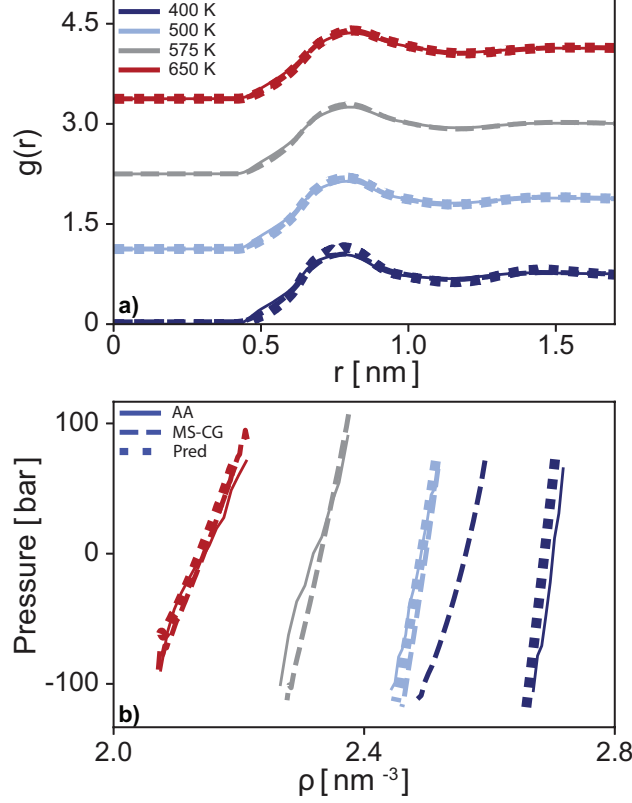


FIG. 13. Analysis of 1-site CG models for OTP with temperature-dependent LD length-scales. Panels (a) and (b) present the mass-center RDF and pressure-density eos obtained from constant NPT simulations of OTP at ambient external pressure, $P_{\text{ext}} = 1$ bar, and the indicated temperature, T . The solid curves present the results of AA simulations. The dashed and dotted curves present the results of simulations with optimized MS-CG potentials, $U(T)$, and predicted CG potentials, $U_{\text{pred}}(T)$, respectively. The grey curves present results for the reference temperature, $T_{\text{ref}} = 575$ K, while the dark blue, light blue, and red curves present results for $T = 400$ K, 500 K, and 650 K, respectively.

Finally, we predict CG potentials, $U_{\text{pred}}(T) = \{U_{2;\text{pred}}(T), U_{\rho;\text{pred}}(T)\}$, for OTP according to the approach introduced in Sec. IV A 3 and assessed for methanol in Figs. 8 and 9. Specifically, we first employ the dashed black line in Fig. 12 to predict an LD length-scale for OTP at each temperature, $r_{c;\text{pred}}(T) = f(T)$, according to Eq. (17). In particular, the dark blue dot in Fig. 12a indicates the predicted LD cut-off for $T_x = 400$ K, which is significantly shorter than the LD cut-off, $r_{c;\text{max}}(T_x)$, that was originally optimized for this temperature. In fact, Fig. 11a indicates that the MS-CG potential, $U(T_x, r_{c;\text{pred}}(T_x))$, for this predicted

length-scale vaporized. We then define the predicted LD potential for each T as the MS-CG LD potential that was optimized for the single fixed temperature $T_{\text{hi}} = 650$ K, while using the predicted LD length-scale, i.e., $U_{\rho;\text{pred}}(T) = U_{\rho}(T_{\text{hi}}, r_{c;\text{pred}}(T))$. In this case, we employ the high temperature simulation at $T_{\text{hi}} = 650$ K to determine this fixed LD potential because it determines U_{ρ} over the widest range of local densities. Finally, we define the predicted pair potential, $U_{2;\text{pred}}(T)$, for each temperature by employing Eq. (16) to extrapolate from the reference temperature, $T_{\text{ref}} = 575$ K. As already noted, the dotted red and light blue curves in Fig. 12d demonstrate that this linear extrapolation accurately predicts the MS-CG pair potentials that were independently optimized for $T = 500$ K and 650 K. Conversely, the dark blue dotted curve in Fig. 12d presents the predicted pair potential for $T_{\text{x}} = 400$ K. The SM demonstrates that this predicted pair potential deviates significantly from the MS-CG pair potential, $U_2(T_{\text{x}}) = U_2(T_{\text{x}}, r_{c;\text{max}}(T_{\text{x}}))$, that was originally optimized for this state point.

The dotted curves in Fig. 13 present results for the predicted potential, $U_{\text{pred}}(T)$, at each temperature. The predicted potentials are very similar to the MS-CG potentials originally optimized for $T = 500$ K, 575 K, and 650 K. Consequently, the red and light blue dotted curves demonstrate that the predicted potentials reproduce the AA RDF and pressure-density eos with similar accuracy to the optimized MS-CG potentials for these temperatures. In fact, Table III demonstrates that the predicted potentials reproduce the AA compressibilities with considerably better accuracy than the optimized MS-CG potentials. Conversely, the dark blue dotted curves present results for the predicted potential, $U_{\text{pred}}(T_{\text{x}})$, at $T_{\text{x}} = 400$ K, which significantly differs from the MS-CG potential, $U(T_{\text{x}})$, that was optimized for T_{x} . Figure 13a demonstrates that this predicted potential reproduces the AA RDF with similar accuracy to the MS-CG potential at $T_{\text{x}} = 400$ K. More importantly, the dark blue dotted curve in Fig. 13b demonstrates that the predicted potential for $T_{\text{x}} = 400$ K reproduces the AA pressure-density eos much more accurately than the original MS-CG potential. In particular, the potential, $U_{\text{pred}}(T_{\text{x}})$, which was predicted without any information from simulations at T_{x} , only underestimates the AA bulk density by 1.8 % and compressibility by 3.7 %. Thus, Fig. 13 demonstrates that this approach for predicting CG potentials accurately models OTP across a very wide range of temperatures. Even more remarkably, the predicted potentials can provide greater accuracy than the MS-CG potentials that we optimized for specific state points.

V. DISCUSSION

In the present work we have investigated the temperature-dependence of LD potentials for one-site MS-CG models of molecular liquids at a fixed external pressure, $P_{\text{ext}} = 1$ bar. We first considered models for methanol at a reference temperature, $T_{\text{ref}} = 260$ K. In this case, the MS-CG potential with a short LD length-scale, $r_{c;0}^* = 0.5234$ nm, very accurately described the mapped RDF and pressure-density eos of the AA model at the reference state point. Moreover, this single MS-CG potential accurately reproduced the mapped RDF and reasonably reproduced the pressure-density eos of the AA methanol across the entire temperature range from 220 K to 300 K. Thus, the fixed MS-CG potential not only accurately modeled the reference state point, but also provided rather satisfactory transferability for essentially the entire liquid phase (at 1 bar pressure).

We had anticipated that it would be possible to improve the transferability of the MS-CG potentials by predictively modeling the temperature-dependence of the pair and LD potentials, e.g., via the dual approach.⁶⁵ Consequently, we calculated an independent MS-CG potential, $U(T, r_{c;0}^*)$, for each temperature, T , while employing the same LD length-scale, $r_{c;0}^*$, that we had optimized for the reference state point. The resulting potentials did indeed vary linearly with temperature. However, these temperature-dependent potentials, $U(T, r_{c;0}^*)$, proved less transferable than the single fixed MS-CG potential, $U(T_{\text{ref}}, r_{c;0}^*)$, that we had originally optimized for T_{ref} . This reduced transferability reflects the high sensitivity of the MS-CG models to the LD length-scale.⁸⁸

Accordingly, we investigated the temperature-dependence of the optimal LD length-scale, $r_c^*(T)$. We observed that $r_c^*(T)$ increases linearly with temperature from 220 K to 300 K. Quite surprisingly, if one defines the local density at each temperature with the optimal length-scale, $r_c^*(T)$, then the LD distribution for methanol does not shift with temperature. Moreover, when employing this temperature-dependent optimal LD length-scale, $r_c^*(T)$, the MS-CG LD potential is essentially independent of temperature, while the MS-CG pair potential varies linearly with temperature. This suggests a simple predictive framework for parameterizing accurate CG models by predicting the temperature-dependent LD length-scale, $r_c^*(T)$, and pair potential, $U_2(T)$, based upon linear extrapolation between two state points, while employing a single fixed reference LD potential, $U_\rho(T_{\text{ref}})$. This approach accurately models methanol at all state points that we considered.

We next investigated the generality of these observations for OTP, which provides a very different model liquid. In particular, we employed the FM variational principle to calculate MS-CG potentials, $U(T, r_c)$, for a wide range of LD length-scales, r_c , at several different temperatures, T . None of the MS-CG potentials reproduced the equilibrium density of the AA model at 1 bar external pressure when $T_x = 400$ K. However, as T increased, we observed a growing range of LD length-scales for which the MS-CG potential, $U(T, r_c)$, stabilized the liquid phase. At these higher temperatures, we again observed an optimal LD length-scale, $r_c^*(T)$, for which the MS-CG potential, $U(T, r_c^*(T))$, very accurately described the corresponding mapped ensemble. In contrast to methanol, $r_c^*(T)$ decreased linearly and the corresponding LD distribution systematically shifted to lower densities with increasing temperature. More importantly, though, the optimal MS-CG pair potential, $U_2(T)$, varied linearly with temperature, while the optimal MS-CG LD potential, $U_\rho(T)$, appeared essentially temperature-independent. Remarkably, these trends allowed us to predict a CG potential, $U_{\text{pred}}(T_x)$, that accurately modeled the problematic state point, $T_x = 400$ K. Thus, this predictive framework not only predicts accurate MS-CG potentials, but also predicts novel potentials for state points that our MS-CG potentials did not accurately model.

We have focused on modeling homogeneous liquids at ambient pressure. Importantly, the predicted CG potentials quite accurately reproduce not only the RDF and bulk density, but also the compressibility of the AA models. In particular, Figs. 9b and 13b demonstrate that the predicted potentials quite accurately reproduce the AA equation of state for methanol and for OTP, respectively, over a pressure range from 0 bar up to approximately 100 bar. Consequently, one anticipates that these potentials should provide a reasonable description of these liquids at both lower and higher pressures. Our preliminary studies suggest that this is indeed the case.

In closing, it is worth noting that we have determined “optimal” CG potentials by minimizing the FM functional, $\chi^2[U]$, in Eq. (13).^{69,70} The FM approach employs AA forces to estimate conditioned mean forces that correspond to the gradients of the PMF.^{100,101} The FM variational principle then determines U in order to optimally match this configuration-dependence of the PMF. However, because the FM variational principle averages over the mapped ensemble, the resulting MS-CG potential is not guaranteed to reproduce the RDF or any other structural property of the mapped ensemble.^{122–127} Moreover, the FM variational principle does not guarantee that the CG model will reproduce the internal pres-

sure or any other thermodynamic property of the AA model. Consequently, the RDF and pressure-density eos determined by CG simulations are predictions of the CG model. The observed agreement between the mapped RDF and the simulated RDF, as well as between the pressure-density equations of state observed for the AA and MS-CG models, indicates that the MS-CG model accurately describes the mapped ensemble. Conversely, it may be possible that alternative approaches for parameterizing the CG potentials, e.g., by employing the inverse Monte Carlo method⁶⁰ or the relative entropy variational principle,^{85,91,128} may determine LD potentials that provide better accuracy and transferability than the MS-CG potentials that we obtained.

VI. CONCLUSIONS

In this study, we have systematically investigated the temperature-dependence of bottom-up LD potentials for one-site CG models of molecular liquids. As we have previously emphasized,^{88,90} the accuracy of MS-CG LD potentials depends very sensitively upon the length-scale employed to define the local density. In many - though not all - cases, it is possible to determine a LD length-scale, r_c^* , such that the optimized MS-CG potential very accurately describes the reference state point and also provides reasonably good transferability across a rather wide range of temperatures (at fixed external pressure). If one wishes to improve the transferability of the MS-CG potential, then it seems necessary to account for the temperature-dependence of the optimal LD length-scale, $r_c^*(T)$. Fortunately, the optimal LD length-scale can be very accurately predicted across a wide temperature range via simple linear extrapolation. Moreover, if one employs the optimal LD length-scale, $r_c^*(T)$, then a single LD potential can be employed to model this temperature range, while the optimal pair potential can be predicted via linear extrapolation. In the cases that we consider, these predicted potentials accurately model a very wide temperature range. Even more remarkably, the predicted potentials can accurately model problematic state-points that our MS-CG potentials did not accurately model.

The present work indicates several interesting directions for future study. Future studies should elucidate the physical basis for the temperature-dependent LD length-scale and temperature-independent LD potential. In particular, one wonders why the LD distribution appears temperature-independent for methanol - but not for OTP - when the local den-

sity is defined by the optimal length-scale, $r_c^*(T)$. It is worth noting that this observation is a property of the AA model and the CG mapping, but is otherwise independent of the CG model. Pragmatically, it would be very beneficial to determine methods for predicting $r_c^*(T)$ directly from the mapped ensemble, i.e., without parameterizing and testing families of CG potentials. While we have exclusively considered the Lucy function for modeling the local density, it would be highly beneficial to develop insight and computational methods to predictively optimize the LD weighting function for different state points, as illustrated by Shahidi et al.⁶⁰ One expects that this may be particularly important for transferring LD potentials between states points that are characterized by differences in the local intermolecular packing. Future studies should also investigate why our MS-CG LD potentials did not accurately model certain state points and, moreover, how the proposed approach predicted novel potentials that did accurately model these problematic state points. Furthermore, future work should extend this study to multicomponent systems and to more complex molecules, such as polymers or proteins. CG models for such systems will require parameterizing an increasing number of LD potentials with a corresponding number of LD length-scales, which may possibly vary with temperature and composition. Surmounting this daunting technical challenge may well require new physical insight and robust numerical methods for addressing the LD length-scale. Nevertheless, we hope that this study provides useful insight into the temperature-dependence of bottom-up LD potentials.

SUPPLEMENTARY MATERIAL

See the supplementary material for additional analysis of the local density and details of our computational methods, as well as for additional results of the CG models for methanol and OTP.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. Michael DeLyser for many helpful discussions throughout this project. The authors also gratefully acknowledge financial support from the National Science Foundation (Grant Nos. CHE-1856337 and CHE-2154433). Portions of this research were conducted with Advanced CyberInfrastructure computational resources pro-

vided by The Institute for Computational and Data Sciences at The Pennsylvania State University (<http://icds.psu.edu>). Additionally, parts of this research used the Expanse resource at the San Diego Supercomputer Center through allocation TG-CHE170062 from the Extreme Science and Engineering Discovery Environment (XSEDE),¹²⁹ which was supported by National Science Foundation grant number TG-CHE170062. This work also used Rockfish at Johns Hopkins University through allocation CHE170062 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296. Figures 2 and 10 employed VMD.¹³⁰ VMD is developed with NIH support by the Theoretical and Computational Biophysics group at the Beckman Institute, University of Illinois at Urbana-Champaign.

AUTHOR DECLARATIONS

Conflict of interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

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

REFERENCES

- ¹Marco Giulini, Marta Rigoli, Giovanni Mattiotti, Roberto Menichetti, Thomas Tarenzi, Raffaele Fiorentini, and Raffaello Potestio. From System Modeling to System Analysis: The Impact of Resolution Level and Resolution Distribution in the Computer-Aided Investigation of Biomolecules. Front. Mol. Biosci., 8:676976, June 2021.
- ²Satyen Dhamankar and Michael A. Webb. Chemically specific coarse-graining of polymers: Methods and prospects. Journal of Polymer Science, 59(22):2613–2643, 2021.
- ³Nicholas E. Jackson. Coarse-Graining Organic Semiconductors: The Path to Multiscale Design. The Journal of Physical Chemistry B, 125(2):485–496, 2021.
- ⁴Tiedong Sun, Vishal Minhas, Nikolay Korolev, Alexander Mirzoev, Alexander P. Lyubartsev, and Lars Nordenskiöld. Bottom-Up Coarse-Grained Modeling of DNA. Frontiers in Molecular Biosciences, 8:645527, March 2021.
- ⁵Jaehyeok Jin, Alexander J. Pak, Aleksander E. P. Durumeric, Timothy D. Loose, and Gregory A. Voth. Bottom-up Coarse-Graining: Principles and Perspectives. Journal of Chemical Theory and Computation, 18(10):5759–5791, October 2022.
- ⁶Siewert J. Marrink, Luca Monticelli, Manuel N. Melo, Riccardo Alessandri, D. Peter Tieleman, and Paulo C. T. Souza. Two decades of Martini: Better beads, broader scope. WIREs Computational Molecular Science, June 2022.
- ⁷Friederike Schmid. Understanding and Modeling Polymers: The Challenge of Multiple Scales. ACS Polymers Au, 3(1):28–58, February 2023.
- ⁸W. G. Noid. Perspective: Advances, challenges, and insight for predictive coarse-grained models. J. Phys. Chem. B, 127:4174–4207, 2023.
- ⁹Sereina Riniker, Jane R Allison, and Wilfred F van Gunsteren. On developing coarse-grained models for biomolecular simulation: a review. Phys. Chem. Chem. Phys., 14(36):12423–12430, Sep 2012.
- ¹⁰M. Muller, K. Katsov, and M. Schick. Biological and synthetic membranes: What can be learned from a coarse-grained description? Phys. Rep., 434(5-6):113–176, 2006.
- ¹¹Riccardo Baron, Alex H. de Vries, Philippe H. Hünenberger, and Wilfred F. van Gunsteren. Comparison of atomic-level and coarse-grained models for liquid hydrocarbons from molecular dynamics configurational entropy estimates. J. Phys. Chem. B, 110(16):8464–8473, 2006.

- ¹²Riccardo Baron, Alex H. de Vries, Philippe H. Hünenberger, and Wilfred F. van Gunsteren. Configurational entropies of lipids in pure and mixed bilayers from atomic-level and coarse-grained molecular dynamics simulations. J. Phys. Chem. B, 110(31):15602–15614, 2006.
- ¹³Riccardo Baron, Daniel Trzesniak, Alex H. de Vries, Andreas Elsener, Siewert J. Marrink, and Wilfred F. van Gunsteren. Comparison of thermodynamic properties of coarse-grained and atomic-level simulation models. ChemPhysChem, 8(3):452–461, 2007.
- ¹⁴Riccardo Baron and Valeria Molinero. Water-driven cavity–ligand binding: Comparison of thermodynamic signatures from coarse-grained and atomic-level simulations. J. Chem. Theory Comput., 8(10):3696–3704, 2012.
- ¹⁵Zack Jarin, James Newhouse, and Gregory A. Voth. Coarse-grained Force Fields from the Perspective of Statistical Mechanics: Better Understanding the Origins of a MARTINI Hangover. J. Chem. Theory Comput., 17(2):1170–1180, June 2020.
- ¹⁶W. G. Noid. Perspective: coarse-grained models for biomolecular systems. J. Chem. Phys., 139(9):090901, 2013.
- ¹⁷M. Guenza. Thermodynamic consistency and other challenges in coarse-graining models. Eur. Phys. J.: Spec. Top., 224:2177–2191, 2015.
- ¹⁸Katherine M Kidder, Ryan J. Szukalo, and W. G Noid. Energetic and entropic considerations for coarse-graining. Eur. Phys. J. B, 94:153, 2021.
- ¹⁹J. G. Kirkwood. Statistical mechanics of fluid mixtures. J. Chem. Phys., 3(5):300–313, 1935.
- ²⁰Reinier L. C. Akkermans and W. J. Briels. A structure-based coarse-grained model for polymer melts. J. Chem. Phys., 114(2):1020–1031, 2001.
- ²¹Christos N. Likos. Effective interactions in soft condensed matter physics. Phys. Rep., 348(4–5):267 – 439, 2001.
- ²²A. Liwo, C. Czaplewski, J. Pillardy, and H. A. Scheraga. Cumulant-based expressions for the multibody terms for the correlation between local and electrostatic interactions in the united-residue force field. J. Chem. Phys., 115:2323 – 2347, 2001.
- ²³W G Noid. Systematic methods for structurally consistent coarse-grained models. Methods Mol Biol, 924:487–531, 2013.
- ²⁴Nicholas J. H. Dunn, Thomas T. Foley, and William G. Noid. van der waals perspective on coarse-graining: progress toward solving representability and transferability problems.

- Acc. Chem. Res., 49(12):2832–2840, 2016.
- ²⁵Avisek Das and Hans C. Andersen. The multiscale coarse-graining method. V. Isothermal-isobaric ensemble. J. Chem. Phys., 132:164106, 2010.
- ²⁶Sergei Izvekov, Peter W. Chung, and Betsy M. Rice. The multiscale coarse-graining method: Assessing its accuracy and introducing density dependent coarse-grain potentials. J. Chem. Phys., 133:064109, 2010.
- ²⁷Nicholas J. H. Dunn and W. G. Noid. Bottom-up coarse-grained models that accurately describe the structure, pressure, and compressibility of molecular liquids. J. Chem. Phys., 143(24):243148, 2015.
- ²⁸Nicholas J. H. Dunn and W. G. Noid. Bottom-up coarse-grained models with predictive accuracy and transferability for both structural and thermodynamic properties of heptane-toluene mixtures. J. Chem. Phys., 144:204124, 2016.
- ²⁹Giuseppe D’Adamo, Andrea Pelissetto, and Carlo Pierleoni. Predicting the thermodynamics by using state-dependent interactions. J. Chem. Phys., 138(23):234107, Jun 2013.
- ³⁰Lanyuan Lu and Gregory A Voth. The multiscale coarse-graining method. VII. Free energy decomposition of coarse-grained effective potentials. J. Chem. Phys., 134(22):224107, Jun 2011.
- ³¹Thomas T. Foley, M. S. Shell, and W. G. Noid. The impact of resolution upon entropy and information in coarse-grained models. J. Chem. Phys., 143:243104, 2015.
- ³²A. A. Louis. Beware of density dependent pair potentials. J. Phys.: Condens. Matter, 14:9187–9206, 2002.
- ³³Margaret E. Johnson, Teresa Head-Gordon, and Ard A. Louis. Representability problems for coarse-grained water potentials. J. Chem. Phys., 126:144509, 2007.
- ³⁴Jayeeta Ghosh and Roland Faller. State point dependence of systematically coarse-grained potentials. Mol. Simu., 33:759–767, 2007.
- ³⁵E Brini and N F A van der Vegt. Chemically transferable coarse-grained potentials from conditional reversible work calculations. J. Chem. Phys., 137(15):154113, Oct 2012.
- ³⁶José G. Solano Canchaya, Alain Dequidt, Florent Goujon, and Patrice Malfreyt. Development of DPD coarse-grained models: From bulk to interfacial properties. J. Chem. Phys., 145(5):054107, August 2016.
- ³⁷Wenjie Xia, Jake Song, Nitin K Hansoge, Frederick R Phelan Jr, Sinan Keten, and Jack F Douglas. Energy renormalization for coarse-graining the dynamics of a model glass-

- forming liquid. J. Phys. Chem. B, 122(6):2040–2045, 2018.
- ³⁸Chenchen Hu, Teng Lu, and Hongxia Guo. Developing a transferable coarse-grained model for the prediction of thermodynamic, structural, and mechanical properties of polyimides at different thermodynamic state points. J. Chem. Inf. Model., 59(5):2009–2025, 2019.
- ³⁹Yiyang Li, Vipin Agrawal, and Jay Oswald. Systematic coarse-graining of semicrystalline polyethylene. J. Polym. Sci., Part B: Polym. Phys., 57(6):331–342, 2019.
- ⁴⁰Jacob W. Wagner, James F. Dama, Aleksander E. P. Durumeric, and Gregory A. Voth. On the representability problem and the physical meaning of coarse-grained models. J. Chem. Phys., 145(4):044108, 2016.
- ⁴¹Thomas Dannenhoffer-Lafage, Jacob W. Wagner, Aleksander E. P. Durumeric, and Gregory A. Voth. Compatible observable decompositions for coarse-grained representations of real molecular systems. J. Chem. Phys., 151(13):134115, 2019.
- ⁴²Marjolein Dijkstra, René van Roij, and Robert Evans. Direct simulation of the phase behavior of binary hard-sphere mixtures: Test of the depletion potential description. Phys. Rev. Lett., 82:117–120, Jan 1999.
- ⁴³Jörg Silbermann, Sabine H. L. Klapp, Martin Shoen, Naresh Channamsetty, Henry Block, and Keith E. Gubbins. Mesoscale modeling of complex binary fluid mixtures: Towards an atomistic foundation of effective potentials. J. Chem. Phys., 124:074105, 2006.
- ⁴⁴Thomas Vettorel and Hendrik Meyer. Coarse graining of short polyethylene chains for studying polymer crystallization. J. Chem. Theory Comput., 2:616–629, 2006.
- ⁴⁵Y. T. Wang, S. Izvekov, T. Y. Yan, and G. A. Voth. Multiscale coarse-graining of ionic liquids. J. Phys. Chem. B, 110:3564 – 3575, 2006.
- ⁴⁶Paola Carbone, Hossein Ali Karimi Varzaneh, Xioayu Chen, and Florian Müller-Plathe. Transferability of coarse-grained force fields: The polymer case. J. Chem. Phys., 128:064904, 2008.
- ⁴⁷V. Krishna, W. G. Noid, and G. A. Voth. The multiscale coarse-graining method. iv. transferring coarse-grained potentials between temperatures. J. Chem. Phys., 131(2):024103, Jul 2009.
- ⁴⁸Grigorios Megariotis, Antonia Vyrkou, Adrien Leygue, and Doros N. Theodorou. Systematic coarse graining of 4-cyano-4'-pentylbiphenyl. Ind. Eng. Chem. Res., 50:546–56, 2011.

- ⁴⁹Sergei Izvekov. Towards an understanding of many-particle effects in hydrophobic association in methane solutions. J. Chem. Phys., 134(3):034104, Jan 2011.
- ⁵⁰Sergei Izvekov, Peter W. Chung, and Betsy M. Rice. Particle-based multiscale coarse graining with density-dependent potentials: Application to molecular crystals (hexahydro-1,3,5-trinitro-s-triazine). J. Chem. Phys., 135(4):044112, 2011.
- ⁵¹Biswaroop Mukherjee, Luigi Delle Site, Kurt Kremer, and Christine Peter. Derivation of coarse grained models for multiscale simulation of liquid crystalline phase transitions. J. Phys. Chem. B, 116(29):8474–8484, 2012.
- ⁵²Jianguo Zhang and Hongxia Guo. Transferability of coarse-grained force field for ncb liquid crystal systems. J. Phys. Chem. B, 118(17):4647–4660, 2014.
- ⁵³Qiang Xiao and Hongxia Guo. Transferability of a coarse-grained atactic polystyrene model: the non-bonded potential effect. Phys. Chem. Chem. Phys., 18:29808–29824, 2016.
- ⁵⁴Marta Montes-Saralegui, Gerhard Kahl, and Arash Nikoubashman. On the applicability of density dependent effective interactions in cluster-forming systems. J. Chem. Phys., 146:054904, 2017.
- ⁵⁵Jacob W. Wagner, Thomas Dannenhoffer-Lafage, Jaehyeok Jin, and Gregory A. Voth. Extending the range and physical accuracy of coarse-grained models: Order parameter dependent interactions. J. Chem. Phys., 147(4):044113, 2017.
- ⁵⁶Gregor Deichmann, Marco Dallavalle, David Rosenberger, and Nico FA van der Vegt. Phase equilibria modeling with systematically coarse-grained models—a comparative study on state point transferability. J. Phys. Chem. B, 123(2):504–515, 2019.
- ⁵⁷David Rosenberger, Tanmoy Sanyal, M. Scott Shell, and Nico F. A. van der Vegt. Transferability of Local Density-Assisted Implicit Solvation Models for Homogeneous Fluid Mixtures. J. Chem. Theory Comput., 15(5):2881–2895, May 2019.
- ⁵⁸Gregor Deichmann, Marco Dallavalle, David Rosenberger, and Nico F. A. van der Vegt. Phase Equilibria Modeling with Systematically Coarse-Grained Models—A Comparative Study on State Point Transferability. J. Phys. Chem. B, 123(2):504–515, January 2019.
- ⁵⁹Kathryn M. Lebold and W. G. Noid. Systematic study of temperature and density variations in effective potentials for coarse-grained models of molecular liquids. J. Chem. Phys., 150(1):014104, 2019.

- ⁶⁰Nobahar Shahidi, Antonis Chazirakis, Vagelis Harmandaris, and Manolis Doxastakis. Coarse-graining of polyisoprene melts using inverse Monte Carlo and local density potentials. J. Chem. Phys., 152(12):124902, March 2020.
- ⁶¹Kevin Shen, Nicholas Sherck, My Nguyen, Brian Yoo, Stephan Köhler, Joshua Speros, Kris T Delaney, Glenn H Fredrickson, and M Scott Shell. Learning composition-transferable coarse-grained models: Designing external potential ensembles to maximize thermodynamic information. J. Chem. Phys., 153(15):154116, 2020.
- ⁶²A. Chaimovich and M. S. Shell. Anomalous waterlike behavior in spherically-symmetric water models optimized with the relative entropy. Phys. Chem. Chem. Phys., 11(12):1901–1915, 2009.
- ⁶³Karim Farah, Aoife Catherine Fogarty, Michael Christian Böhm, and Florian Müller-Plathe. Temperature dependence of coarse-grained potentials for liquid hexane. Phys. Chem. Chem. Phys., 13(7):2894–902, Feb 2011.
- ⁶⁴Ryan J Szukalo and WG Noid. Investigation of coarse-grained models across a glass transition. Soft Mater., 18:1–15, 2020.
- ⁶⁵Kathryn M. Lebold and W. G. Noid. Dual approach for effective potentials that accurately model structure and energetics. J. Chem. Phys., 150:234107, 2019.
- ⁶⁶Kathryn M. Lebold and W. G. Noid. Dual-potential approach for coarse-grained implicit solvent models with accurate, internally consistent energetics and predictive transferability. J. Chem. Phys., 151(16):164113, 2019.
- ⁶⁷Ryan J Szukalo and W G Noid. Investigating the energetic and entropic components of effective potentials across a glass transition. J. Phys.: Condens. Matter, 33(15):154004, Feb 2021.
- ⁶⁸F. Müller-Plathe. Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. ChemPhysChem, 3:754 – 769, 2002.
- ⁶⁹S. Izvekov and G. A. Voth. A multiscale coarse-graining method for biomolecular systems. J. Phys. Chem. B, 109:2469 – 2473, 2005.
- ⁷⁰S. Izvekov and G. A. Voth. Multiscale coarse graining of liquid-state systems. J. Chem. Phys., 123:134105, 2005.
- ⁷¹Gergely Tóth. Effective potentials from complex simulations: a potential-matching algorithm and remarks on coarse-grained potentials. J. Phys.: Condens. Matter, 19(33):335222, Aug 2007.

- ⁷²Evan Pretti and M. Scott Shell. A microcanonical approach to temperature-transferable coarse-grained models using the relative entropy. J. Chem. Phys., 155:094102, 2021.
- ⁷³I. Pagonabarraga and D. Frenkel. Dissipative particle dynamics for interacting systems. J. Chem. Phys., 115(11):5015–5026, 2001.
- ⁷⁴Garegin A. Papoian, Johan Ulander, Michael P. Eastwood, Zaida Luthey-Schulten, and Peter G. Wolynes. Water in protein structure prediction. Proc. Natl. Acad. Sci. U.S.A., 101(10):3352–3357, 2004.
- ⁷⁵S. Y. Trofimov, E. L. F. Nies, and M. A. J. Michels. Thermodynamic consistency in dissipative particle dynamics simulations of strongly nonideal liquids and liquid mixtures. J. Chem. Phys., 117(20):9383–9394, November 2002.
- ⁷⁶Pep Español and Patrick B. Warren. Perspective: Dissipative particle dynamics. J. Chem. Phys., 146(15):150901, App. Phys. Rev. 2017.
- ⁷⁷J Bonet Avalos and A D Mackie. Dissipative particle dynamics with energy conservation. Europhysics Letters (EPL), 40:141–146, 1997.
- ⁷⁸P Español. Fluid particle dynamics: A synthesis of dissipative particle dynamics and smoothed particle dynamics. Europhysics Letters (EPL), 39(6):605–610, September 1997.
- ⁷⁹Pep Español and Mariano Revenga. Smoothed dissipative particle dynamics. Physical Review E, 67(2):026705, February 2003.
- ⁸⁰Pep Español, Mar Serrano, Ignacio Pagonabarraga, and Ignacio Zúñiga. Energy-conserving coarse-graining of complex molecules. Soft Matter, 12(21):4821–4837, May 2016.
- ⁸¹Josep Bonet Avalos, Martin Lísal, James P. Larentzos, Allan D. Mackie, and John K. Brennan. Generalised dissipative particle dynamics with energy conservation: density- and temperature-dependent potentials. Physical Chemistry Chemical Physics, 21(45):24891–24911, 2019.
- ⁸²Martin Lísal, James P. Larentzos, Josep Bonet Avalos, Allan D. Mackie, and John K. Brennan. Generalized Energy-Conserving Dissipative Particle Dynamics with Reactions. Journal of Chemical Theory and Computation, 18(4):2503–2512, App. Phys. Rev. 2022.
- ⁸³Erik C. Allen and Gregory C. Rutledge. A novel algorithm for creating coarse-grained, density dependent implicit solvent models. J. Chem. Phys., 128:154115, 2008.
- ⁸⁴Erik C. Allen and Gregory C. Rutledge. Evaluating the transferability of coarse-grained, density-dependent implicit solvent models to mixtures and chains. J. Chem. Phys.,

130:034904, 2009.

- ⁸⁵Tanmoy Sanyal and M. Scott Shell. Transferable coarse-grained models of liquid-liquid equilibrium using local density potentials optimized with the relative entropy. J. Phys. Chem. B, 122:5678–5693, 2018.
- ⁸⁶Tanmoy Sanyal and M. Scott Shell. Transferable Coarse-Grained Models of Liquid-Liquid Equilibrium Using Local Density Potentials Optimized with the Relative Entropy. J. Phys. Chem. B, 122:5678–5693, February 2018.
- ⁸⁷M. R. DeLyser and William G. Noid. Extending pressure-matching to inhomogeneous systems via local-density potentials. J. Chem. Phys., 147:134111, 2017.
- ⁸⁸Michael R DeLyser and WG Noid. Analysis of local density potentials. J. Chem. Phys., 151(22):224106, 2019.
- ⁸⁹Michael DeLyser and WG Noid. Bottom-up coarse-grained models for external fields and interfaces. J. Chem. Phys., 153(22):224103, 2020.
- ⁹⁰Michael DeLyser and WG Noid. Coarse-grained models for local density gradients. J. Chem. Phys., 156:034106, 2022.
- ⁹¹Tanmoy Sanyal and M. Scott Shell. Coarse-grained models using local-density potentials optimized with the relative entropy: Application to implicit solvation. J. Chem. Phys., 145(3):034109, 2016.
- ⁹²Vipin Agrawal, Pedro Peralta, Yiyang Li, and Jay Oswald. A pressure-transferable coarse-grained potential for modeling the shock hugoniot of polyethylene. J. Chem. Phys., 145(10):104903, 2016.
- ⁹³Fabian Berressem, Christoph Scherer, Denis Andrienko, and Arash Nikoubashman. Ultra-coarse-graining of homopolymers in inhomogeneous systems. J. Phys.: Condens. Matter, 33(25):254002, June 2021.
- ⁹⁴Joshua D. Moore, Brian C. Barnes, Sergei Izvekov, Martin Lísal, Michael S. Sellers, DeCarlos E. Taylor, and John K. Brennan. A coarse-grain force field for rdx: Density dependent and energy conserving. J. Chem. Phys., 144(10):104501, 2016.
- ⁹⁵Mara Jochum, Denis Andrienko, Kurt Kremer, and Christine Peter. Structure-based coarse-graining in liquid slabs. J. Chem. Phys., 137(6):064102, 2012.
- ⁹⁶Cahit Dalgicdir, Ozge Sensoy, Christine Peter, and Mehmet Sayar. A transferable coarse-grained model for diphenylalanine: How to represent an environment driven conformational transition. J. Chem. Phys., 139(23), 2013.

- ⁹⁷A. P. Lyubartsev and A. Laaksonen. Calculation of effective interaction potentials from radial distribution functions: a reverse Monte Carlo approach. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics, 52:3730–3737, 1995.
- ⁹⁸A. Lyubartsev, A. Mirzoev, L. J. Chen, and A. Laaksonen. Systematic coarse-graining of molecular models by the newton inversion method. Faraday Discuss., 144:43–56, 2010.
- ⁹⁹Jacobus A. van Meel, Laura Filion, Chantal Valeriani, and Daan Frenkel. A parameter-free, solid-angle based, nearest-neighbor algorithm. The Journal of Chemical Physics, 136(23):234107, June 2012.
- ¹⁰⁰W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen. The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. J. Chem. Phys., 128:244114, 2008.
- ¹⁰¹W. G. Noid, P. Liu, Y. T. Wang, J.-W. Chu, G. S. Ayton, S. Izvekov, H. C. Andersen, and G. A. Voth. The multiscale coarse-graining method. II. Numerical implementation for molecular coarse-grained models. J. Chem. Phys., 128:244115, 2008.
- ¹⁰²Andreas Krämer, Aleksander P. Durumeric, Nicholas E. Charron, Yaoyi Chen, Cecilia Clementi, and Frank Noé. Statistically Optimal Force Aggregation for Coarse-Graining Molecular Dynamics, February 2023. arXiv:2302.07071 [physics, stat], doi: 10.48550/ARXIV.2302.07071.
- ¹⁰³W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. Development and testing of the OPLS All-Atom force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc., 118:11225–11236, 1996.
- ¹⁰⁴Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An $N \log(N)$ method for ewald sums in large systems. J. Chem. Phys., 98:10089–10092, 1993.
- ¹⁰⁵Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J. Chem. Theory Comput., 4(3):435–447, 2008.
- ¹⁰⁶Mark James Abraham, Teemu Murtola, Roland Schulz, Szilard Páll, Jeremy C. Smith, Berk Hess, and Erik Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX, 1-2:19 – 25, 2015.
- ¹⁰⁷Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. J. Chem. Phys., 126(1):014101, JAN 7 2007.

- ¹⁰⁸M. Parrinello and A. Rahman. Strain fluctuations and elastic constants. J. Chem. Phys., 76:2662–6, 1982.
- ¹⁰⁹H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. J. Chem. Phys., 81:3684–3690, 1984.
- ¹¹⁰T. L. Hill. An Introduction to Statistical Thermodynamics. Addison Wesley Publishing Company, 1987.
- ¹¹¹Nicholas J. H. Dunn, Kathryn M. Lebold, Michael R. DeLyser, Joseph F. Rudzinski, and W. G. Noid. BOCS: Bottom-up Open-source Coarse-graining Software. J. Phys. Chem. B, 122(13):3363–3377, App. Phys. Rev. 2018.
- ¹¹²S. Plimpton. Fast parallel algorithms for short-range molecular dynamics. J. Comp. Phys., 117:1 – 19, 1995.
- ¹¹³Aidan P. Thompson, H. Metin Aktulga, Richard Berger, Dan S. Bolintineanu, W. Michael Brown, Paul S. Crozier, Pieter J. in ’t Veld, Axel Kohlmeyer, Stan G. Moore, Trung Dac Nguyen, Ray Shan, Mark J. Stevens, Julien Tranchida, Christian Trott, and Steven J. Plimpton. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. Computer Physics Communications, 271:108171, February 2022.
- ¹¹⁴S. Nose. A molecular dynamics method for simulations in the canonical ensemble. Mol. Phys., 52:255–68, 1984.
- ¹¹⁵W. G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. Phys. Rev. A, 31:1695–7, 1985.
- ¹¹⁶Glenn J. Martyna, Michael L. Klein, and Mark Tuckerman. Nosé–hoover chains: The canonical ensemble via continuous dynamics. J. Chem. Phys., 97(4):2635–2643, 1992.
- ¹¹⁷Glenn J. Martyna, Douglas J. Tobias, and Michael L. Klein. Constant pressure molecular dynamics algorithms. J. Chem. Phys., 101(5):4177–4189, 1994.
- ¹¹⁸Glenn J. Martyna, Mark E. Tuckerman, Douglas J. Tobias, and Michael L. Klein. Explicit reversible integrators for extended systems dynamics. Mol. Phys., 87(5):1117–1157, 1996.
- ¹¹⁹Paola Carbone, Hossein Ali Karimi Varzaneh, Xiaoyu Chen, and Florian Müller-Plathe. Transferability of coarse-grained force fields: The polymer case. The Journal of Chemical Physics, 128(6):064904, February 2008.

- ¹²⁰H. Wang, C. Junghans, and K. Kremer. Comparative atomistic and coarse-grained study of water: What do we lose by coarse-graining? Eur. Phys. J. E: Soft Matter Biol. Phys., 28(2):221–229, 2009.
- ¹²¹Motosuke Naoki and Susumu Koeda. Pressure-volume-temperature relations of liquid, crystal, and glass of o-terphenyl: excess amorphous entropies, and factors determining molecular mobility. J. Chem. Phys., 93:948–955, 1989.
- ¹²²V. Rühle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko. Versatile object-oriented toolkit for coarse-graining applications. J. Chem. Theory Comput., 5(12):3211–3223, 2009.
- ¹²³Avisek Das, Lanyuan Lu, Hans C. Andersen, and Gregory A. Voth. The multiscale coarse-graining method. X. Improved algorithms for constructing coarse-grained potentials for molecular systems. J. Chem. Phys., 136(19):194115, 2012.
- ¹²⁴Joseph F Rudzinski and William G Noid. The role of many-body correlations in determining potentials for coarse-grained models of equilibrium structure. J. Phys. Chem. B, 116(29):8621–8635, Jul 2012.
- ¹²⁵Lanyuan Lu, James F. Dama, and Gregory A. Voth. Fitting coarse-grained distribution functions through an iterative force-matching method. J. Chem. Phys., 139(12):121906, 2013.
- ¹²⁶Joseph F. Rudzinski and William G. Noid. Investigation of coarse-grained mappings via an iterative generalized yvon-born-green method. J. Phys. Chem. B, 118(28):8295–8312, 2014.
- ¹²⁷J. F. Rudzinski and W. G. Noid. A generalized-yvon-born-green method for coarse-grained modeling. Eur. Phys. J.: Spec. Top., 224:2193–2216, 2015.
- ¹²⁸M. Scott Shell. The relative entropy is fundamental to multiscale and inverse thermodynamic problems. J. Chem. Phys., 129:144108, 2008.
- ¹²⁹J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott, and N. Wilkins-Diehr. Xsede: Accelerating scientific discovery. Computing in Science & Engineering, 16(5):62–74, Sept.-Oct. 2014.
- ¹³⁰William Humphrey, Andrew Dalke, and Klaus Schulten. VMD: Visual Molecular Dynamics. J. Mol. Graph., 14:33–38, 1996.