This is the accepted manuscript made available via CHORUS. The article has been published as:

Coarse-grained lattice Monte Carlo simulations with continuous interaction potentials

Xiao Liu, Warren D. Seider, and Talid Sinno

Phys. Rev. E **86**, 026708 — Published 16 August 2012

DOI: 10.1103/PhysRevE.86.026708

Coarse-Grained Lattice Monte Carlo Simulations with Continuous Interaction Potentials

Xiao Liu, Warren D. Seider and Talid Sinno[1]
Department of Chemical and Biomolecular Engineering
University of Pennsylvania
Philadelphia, PA 19104

A coarse-grained lattice Metropolis Monte Carlo (CG-MMC) method is presented for simulating fluid systems described by standard molecular force-fields. First, a thermodynamically consistent coarse-grained interaction potential is obtained numerically and automatically from a continuous force-field such as Lennard-Jones. The coarse-grained potential then is used to drive CG-MMC simulations of vapor-liquid equilibrium in Lennard-Jones, square-well, and SPC-water systems. The CG-MMC predicts vapor-liquid phase envelopes, as well as the particle density distributions in both the liquid and vapor phases, in excellent agreement with full-resolution Monte Carlo simulations at a fraction of the computational cost.

PACS numbers: 02.70.-c, 05.10.-a

[1] Email: talid@seas.upenn.edu

1. INTRODUCTION

While full-resolution, discrete-particle simulation methods such as standard Metropolis Monte Carlo or molecular dynamics are becoming increasingly powerful as a result of more available computational resources, generally they remain limited to nanoscale lengths and times. As a result, accurate, simple, and broadly applicable coarse-graining (CG) methods are increasingly sought to expand the range of non-equilibrium phenomena that can be probed with atomic and/or molecular simulations. Most generally, coarse-graining refers to a transformation in which degrees-of-freedom are eliminated, increasing computational efficiency while sacrificing some information. A popular example of coarse-graining is the united atom representation, in which groups of atoms are combined into single particles that obey a new, coarse-grained potential function [1]. The aim then is to generate a coarse-grained inter-particle potential that embeds the enthalpic and entropic contributions of the discarded degrees-of-freedom, and also to ensure that the dynamical fluctuations of interest are preserved [2-3]. Such approaches have been used to study a broad range of materials including polymers [1, 4], proteins [3], and ionic liquids [5].

A somewhat different approach for degree-of-freedom reduction is to map the problem onto a fixed lattice, starting with block-spin renormalization group theory [6]; here, we refer to this type of transformation as *spatial* coarse-graining as opposed to the *topological* coarse-graining described above. In spatial coarse-graining of polymeric systems, for example, chains are placed on fixed grids and allowed to evolve subject to discretized moves; one example is the bond fluctuation method [7]. Recently, spatial coarse-graining has been applied extensively to Metropolis and kinetic Monte Carlo simulations of Ising-type systems in which the aim is to transform one (high-resolution) lattice problem onto a lower-resolution lattice by grouping together lattice sites into coarse "cells". Notable examples include the work of Katsoulakis, Vlachos and coworkers [8-10] and Ismail et al. [11-12]. A key element of these methods is the closure rule, which dictates how processes on the fine-grid lattice

are averaged to generate consistent processes on the coarse-cell grid. This may be accomplished using analytical approximations [8, 13] or numerical averaging [12, 14-15].

Here, we extend spatial coarse-graining to a more general situation in which a continuous system of particles, subject to an arbitrary interaction potential, is mapped onto a rigid lattice of variable scale that can then be evolved with Metropolis Monte Carlo using an appropriate coarse-grained potential. Most significantly, we seek a coarse-grained potential that is thermodynamically consistent with the microscopic potential, i.e., that the entropy associated with the missing degrees-of-freedom is properly embedded into the coarse-grained potential. This latter issue has not been addressed in prior spatial coarse-graining approaches [16-18], which, while successful for simple interaction models, have not been tested for realistic potentials under near-equilibrium conditions.

2. COARSE-GRAINING METHODOLOGY

Consider a three-dimensional system of N particles within a cubic simulation cell of length L subject to periodic boundary conditions and evolving within the canonical ensemble (constant NVT). Within the coarse-grained representation, the overall domain is discretized into m^3 cubic coarse cells, each with length $L_{cell} = L/m$ and volume $V_{cell} = {L_{cell}}^3$. Each coarse cell can contain multiple particles, which are assumed to always exist in local equilibrium [8]. The coarse-graining transformation is derived by first considering the system-wide canonical partition function

$$Q(N,V,T) = \frac{1}{\Lambda^{3N} N!} \int \mathrm{d}\mathbf{r}^N \exp\left[-\frac{U(\mathbf{r}^N)}{k_B T} \right], \tag{1}$$

where $U(\mathbf{r}^N)$ is a specified interaction potential function. Rewriting the partition function in terms of sub-integrals over coarse cells, whereby the N particles in the system are distributed over the $M (= m^3)$ coarse cells, gives

$$Q(N,V,T) = \frac{1}{\Lambda^{3N} N!} \sum_k \int \mathrm{d}\mathbf{r}^{n_1} \exp\left[-\frac{U(\mathbf{r}^N)}{k_B T} \right] \ldots \int \mathrm{d}\mathbf{r}^{n_M} \exp\left[-\frac{U(\mathbf{r}^N)}{k_B T} \right], \tag{2}$$

where $\mathbf{n} \equiv (n_1, n_2, \cdots, n_M)$ is an *M*-dimensional vector that defines the cell occupancy within the domain, and the sum index *k* runs over all possible ways of distributing the *N* particles over the *M* coarse cells. Each of the sub-integrals in eq. (2) is related to a local Helmholtz free energy so that

$$Q(N,V,T) = \frac{1}{\Lambda^{3N} N!} \sum_k \exp\left(-\frac{A_1}{k_B T} \right) \ldots \exp\left(-\frac{A_M}{k_B T} \right), \tag{3}$$

where $A_i \equiv -k_B T \ln \int d\mathrm{r}^{n_i} \exp\left(-U(\mathrm{r}^N) / k_B T \right)$. Finally, defining a coarse-grained system-wide free energy, $A_{CG}(\mathbf{n}) = \sum_{i=1}^{M} A_i$, leads to

$$Q(N,V,T) = \frac{1}{\Lambda^{3N} N!} \sum_k \exp\left(-\frac{A_{CG}(\mathbf{n})}{k_B T} \right). \tag{4}$$

Eq. (4) suggests that a valid coarse-grained Metropolis Monte Carlo (CG-MMC) simulation proceeds identically to one on the original continuous system, except that the move acceptance criterion would be based on $\Delta A_{CG}(\mathbf{n})$ rather than $\Delta E(\mathbf{r}^N)$, i.e.,

$$P_{CG}^{acc}(1 \to 2) = \min\left[\exp\left(-\frac{\Delta A_{CG}}{k_B T} \right), 1 \right], \tag{5}$$

where P_{CG}^{acc} is the acceptance probability for moving from state 1 to state 2.

To compute $\Delta A_{CG}(\mathbf{n})$ for use in a CG-MMC simulation, a coarse-grained potential function must be calculated from the original inter-particle potential. Generally, the free energy change within a coarse cell due to the addition of one particle is a function of the local number density ρ, the local surrounding density distribution ρ_{env}, and temperature *T*, i.e.,

$$\Delta A = \Delta A_{id}\left(\rho, T\right) + \Delta A_{ex}\left(\rho, \rho_{env}, T\right), \tag{6}$$

where the "CG" subscript has been dropped for clarity and the free energy has been separated into ideal and excess contributions. Note that this free energy difference corresponds to the chemical potential in the limit of large particle number.

The ideal contribution to the free energy difference on a coarse cell is given analytically as

$$\Delta A_{id}\left(\rho, T\right) = k_B T \ln \Lambda^3 + k_B T \ln \rho, \tag{7}$$

where $\rho = n / V_{cell}$. The excess portion must be computed by ensemble averaging under the influence of the interaction potential. Here, we choose to employ the standard Widom particle insertion method [19] (although any other method for free energy estimation also can be applied):

$$\Delta A_{ex}(n \to n+1) = -k_B T \ln \int \mathrm{d}\mathbf{r}^{n+1} \left\langle \exp\left(-\frac{\Delta U}{k_B T} \right) \right\rangle_n . \tag{8}$$

As indicated in eq. (6), the cell excess chemical potential is expected to depend on the surrounding particle density, and possibly the spatial distribution of that density. In order to simplify the closure approximation, we assume here that the spatial distribution within the surrounding environment is relatively unimportant and only consider the average density in neighboring cells.

The configuration of the Widom insertion simulations for computing the excess free energy changes is described schematically in Fig. 1. A central coarse cell is embedded within a shell domain representing the surrounding environment. Only the environment shell's outer boundaries are subject to periodic boundary conditions – the inner boundary between the center cell and environment is impermeable to mass (thus constraining the cell density) but does allow cross-boundary interactions. Test particle insertions are performed only within the center cell although the corresponding potential energy change is computed over all particles. About 100 test insertions are performed every 2500 MMC moves (depending on the particle count). The procedure is repeated for a range of center cell

and environment densities and temperatures; the final result is a multidimensional surface of excess free energy differences that, along with the ideal contribution (eq. 7), defines the coarse-grained potential.

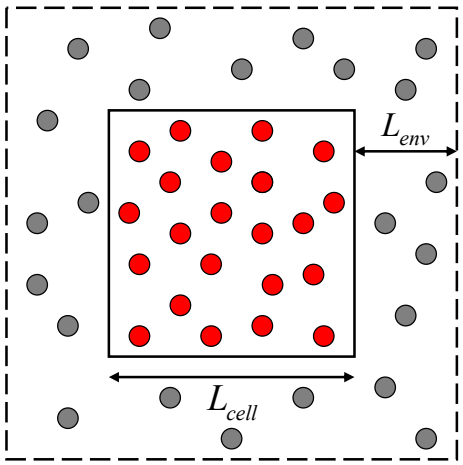


Fig. 1. (Color online) Cell setup for computing coarse-grained interaction potential. An inner cell (solid line) is surrounded by an "environment" shell subject to periodic boundary conditions (dashed line). Particles are not allowed to move between the two regions during free energy sampling but do interact across the partition.

3. RESULTS AND DISCUSSION

3.1 Coarse-Grained Potential Calculations

Three interatomic potential examples are used to generate coarse-grained interaction functions. The first is the Lennard-Jones (LJ) potential for argon ($\sigma = 3.405$ Å, $\varepsilon / k_B = 119.8$ K):

$$U(r_{ij}) = \begin{cases} 4\varepsilon\left[\left(\dfrac{\sigma}{r_{ij}}\right)^{12} - \left(\dfrac{\sigma}{r_{ij}}\right)^{6}\right], & r_{ij} \le r_c \\ 0, & r_{ij} > r_c \end{cases} . \qquad (9)$$

The second is a square-well (SW) potential with two different parameterizations,

$$U(r_{ij}) = \begin{cases} \infty, & r_{ij} < \sigma \\ -\varepsilon, & \sigma \le r_{ij} < \lambda\sigma \\ 0, & r_{ij} > \lambda\sigma \end{cases}, \tag{10}$$

with $r_c = 2.5\sigma$, and $\lambda = 1.5$ or 1.25. Finally, we consider the more general case of water as modeled by a spherically-truncated version of the SPC potential [20-21] which demonstrates the applicability of CG-MMC to molecular systems. The SPC potential is given by

$$U(r_{ij}) = \begin{cases} 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j e^2}{r_{ij}}, & r_{ij} \le r_c \\ 0, & r_{ij} > r_c \end{cases}, \tag{11}$$

with $r_c = 7.75$ Å, $\sigma = 3.1655$ Å, $\varepsilon = 0.15542$ kcal/mol, $q_H = 0.41$ and $q_O = -0.82$ electrons. The LJ portion in eq. (11) applies only between oxygen atoms. While spherical truncation leads to various deficiencies in the description of water, the VLE curve predicted by this potential, which will be used to validate the CG-MMC method, is in good agreement with results obtained using full Ewald summation [21]. We also note that although Ewald summation is not compatible with the simulation cell structure shown in Fig. 1a, other methods exist for including long-range electrostatic interactions into a truncated potential, namely the reaction field [22] and Wolf summation [23] methods. The latter, in particular, has recently been shown to offer advantages over other methods because of its computational efficiency and applicability to inhomogeneous and finite systems.

An example excess chemical potential field for $L_{cell} = 3\sigma$ and $L_{env} = 1.5\sigma$ is shown in Fig. 2 for the LJ potential at a reduced temperature ($T^* = k_B T / \varepsilon$) of 0.8. The value of L_{env} was chosen based on a compromise between convergence with respect to the environment shell thickness and computational expediency. Each of the spherical symbols in Fig. 2 corresponds to a single Widom insertion simulation; the color field surface is a 3rd-order (per dimension) polynomial fit to the data.

While the overall computational effort associated with the pre-calculation of the coarse-grained potential can be significant, it is trivially distributable over an arbitrary number of processing units.

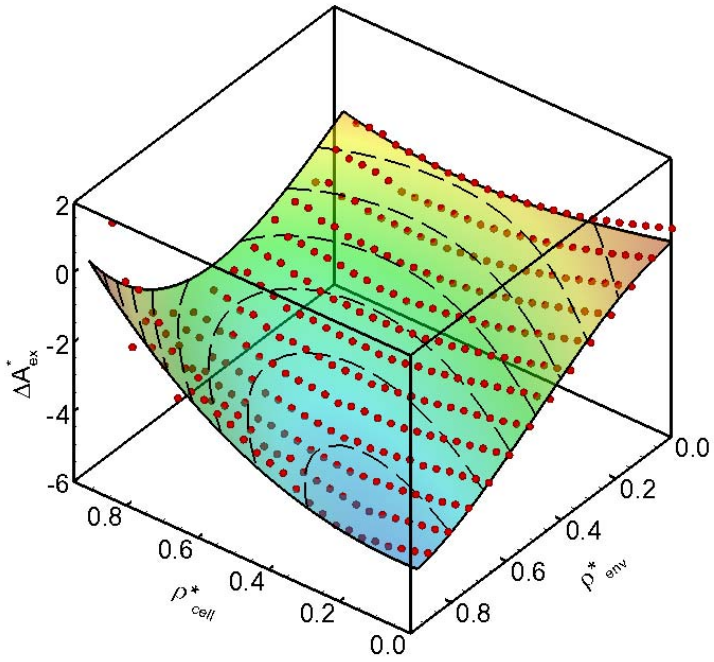


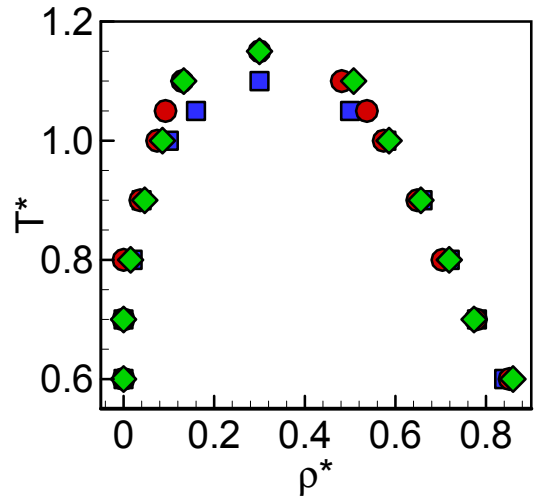
Fig. 2. (Color online) Excess chemical potential computed for LJ potential at $T^* \equiv 0.8$ as a function of cell and environment reduced number densities, $\rho^*_{cell} = \rho_{cell}\sigma^3$ and $\rho^*_{env} = \rho_{env}\sigma^3$, respectively. Coarse cell length (L_{cell}) is 3σ, environment shell thickness (L_{env}) is 1.5σ. Symbols – Widom insertion; color field contours – polynomial interpolation.

3.2 Coarse-Grained Vapor-Liquid Equilibrium

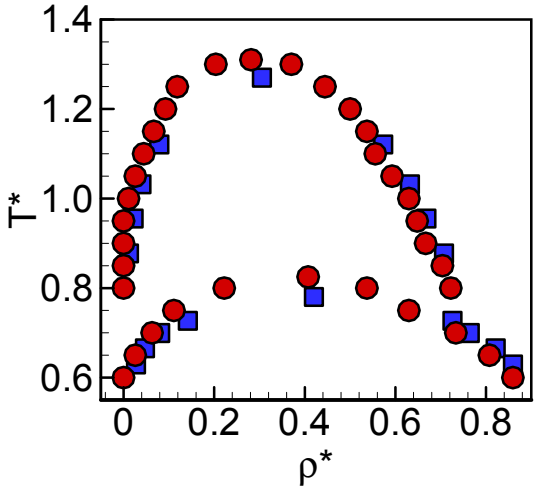
The vapor-liquid equilibrium phase diagram was used as a test-bed for evaluating the thermodynamic consistency of the CG-MMC framework for all three potentials discussed in the previous section. The system size for each case consisted of 10^3 cells, each initialized with the same number of particles to provide a prescribed overall reduced number density, $\rho^* = 0.3$. CG-MMC simulation proceeded by first choosing a random cell and random neighboring destination cell and

then computing the system free energy difference for a particle move using the data in Fig. 2 (for LJ, and similar plots – not shown – for the other potentials) and eq. (7). Equation (5) was then used to determine move acceptance. Equilibration of the CG system was assessed by monitoring the total coarse-grained free energy. Picking cells randomly to execute particle moves satisfies detailed balance and thus generates the correct equilibrium distribution. Unlike in full-resolution MMC, detailed balance in CG-MMC is enforced by move rejections in the direction of increasing chemical potential. However, the CG-MMC "dynamics" as the system evolves towards equilibrium are not completely consistent with full resolution Metropolis Monte Carlo within this scheme. A more detailed discussion of non-equilibrium CG-MMC will be provided in a future publication.

Shown in Fig. 3a and 3b are $T-\rho$ VLE envelopes obtained for the LJ and SW potentials using CG-MMC with cells of size $L_{cell}=3\sigma$. Also shown in Fig. 3a are LJ results using larger cells ($L_{cell}=4\sigma$). The CG-MMC liquid and vapor phase densities at each temperature were obtained from the equilibrium density distributions by locating the two peaks in the distribution that correspond to the liquid and vapor, respectively (see Fig. 4). For each case, the corresponding VLE envelope predicted by full-resolution simulation based on the Gibbs-ensemble Monte Carlo (GEMC) method also is shown [24-25]. The agreement in each case is generally excellent, with some deviation observed near the critical points. The LJ prediction also appears to be largely insensitive to cell size (see Fig. 3a), which suggests that the deviations near the critical point are not due primarily to the small cell size. In fact, the main source of error likely is the statistical uncertainty in the excess chemical potential values and the resulting interpolated surface (Fig. 2). Generally, near the critical point the excess chemical potential surface becomes quite flat, and the scatter in the individual data points there leads to greater uncertainty in the fitted polynomial function. Moreover, the polynomial fitting itself leads to some systematic error, particularly when small changes in curvature lead to large shifts in density. A more quantitative error analysis will require improvements in excess chemical potential calculation; e.g. using bias methods [24].



(a)



(b)

Fig. 3. (Color online) VLE phase envelopes for (a) Lennard-Jones argon, (b) square-well potentials (top - $\lambda = 1.5$, bottom - $\lambda = 1.25$). Red circles – CG-MMC with $L_{cell} = 3\sigma$, green diamonds – CG-MMC with $L_{cell} = 4\sigma$, blue squares – full resolution GEMC.

The VLE curve for SPC water is shown in Fig. 4 along with literature values taken from GEMC simulations and experiments for the continuous system [21]. The agreement is also generally excellent, with similar deviations once again found near the critical point. The calculation of the

coarse-grained potential for molecular systems proceeds in essentially the same manner as described above for point particles, except that some additional care must be exercised at the wall separating the inner cell and the environment shell. Restriction of entire molecules to either the inner cell or environment leads to an artificial configurational penalty at the wall. In order to remove this artifact, the wall was modified to allow free exchange of the hydrogen atoms, while restricting the (point) position of the oxygen atoms to either side of the wall. While this choice of point constraint is convenient for the specific case of water, a more general (but equivalent) choice is to simply use the molecular center-of-mass. In this way, it is anticipated that essentially any type of molecular entity can be readily considered within the CG-MMC framework.

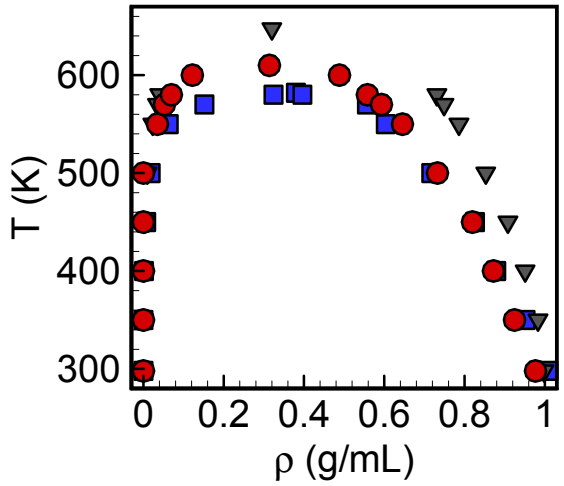


Fig. 4. (Color online) VLE phase envelopes for SPC water. Red circles – CG-MMC with $L_{cell} = 3\sigma$, blue squares – full resolution GEMC, gray deltas – experimental data (see text).

Snapshots of example equilibrium configurations at three different overall system densities of the LJ system are shown for CG-MMC and full-resolution MMC simulations in Fig. 5. The coarse-grained simulations show very clearly the liquid-vapor phase boundaries, and also explicitly highlight

the density fluctuations within each phase. The CG-MMC simulations reach equilibrium in far fewer Monte Carlo move attempts per particle than the full resolution simulations, primarily because the (successful) moves in the coarse-grained system are much larger. Moreover, the computational cost associated with each move, on a per-particle basis, is lower in the CG-MMC case. Overall, for $L_{cell} = 3\sigma$, the CG-MMC simulation reaches equilibrium about 10^6 times faster than a similarly-sized full-resolution system.

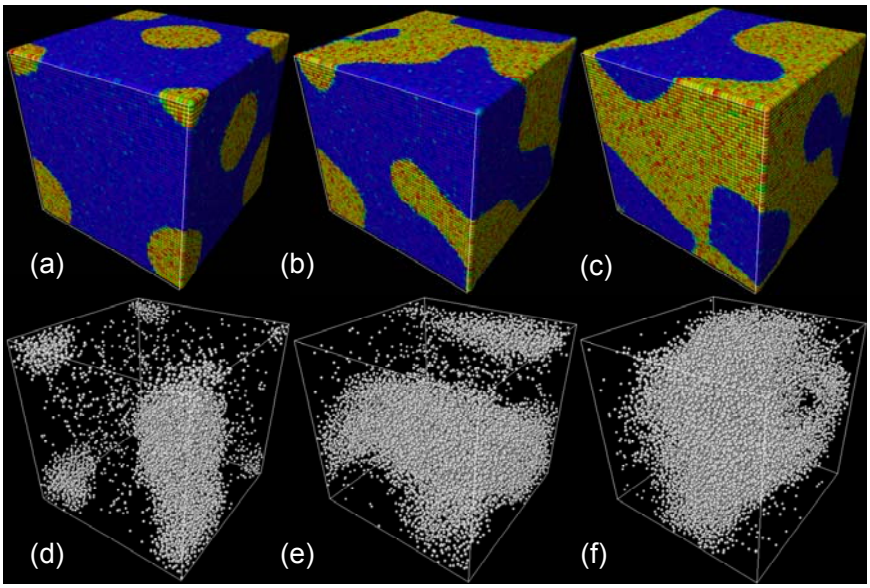


Fig. 5. (Color online) Top row: Equilibrium snapshots of LJ CG-MMC simulation configurations at (a) $\rho^* = 0.15$, (b) $\rho^* = 0.3$, and (c) $\rho^* = 0.45$. System size is $180\times180\times180\ \sigma^3$, ($L_{cell} = 3\sigma$). Cell color denotes particle number that ranges from zero (dark blue) to 24 (red). Bottom row (d-f): Corresponding full resolution MMC simulation snapshots for a system size of $36\times36\times36\ \sigma^3$.

3.3 Density Distributions in Coarse-Grained Simulations

The VLE curve is a highly coarse-grained measure in that it does not provide a view into microscopic density fluctuations. To further analyze the fidelity of the CG-MMC approach, we compare the equilibrium probability distribution of particle number density predicted by LJ CG-MMC simulations and that obtained from *post facto* coarse-graining of full-resolution MMC configurations at $T^* = 0.8$. For the latter, ~250 equilibrated (phase separated) configurations were captured and gridded into coarse cell lattices which were then used to collect density distribution data. Each configuration was gridded 100 times using a randomly selected origin to improve the density distribution statistics. As shown in Fig. 6, the density distributions for $L_{cell} = 3\sigma$ cells obtained from CG-MMC and *post facto* coarse-grained full resolution configurations are in excellent quantitative agreement for the LJ potential, demonstrating that spatial fluctuations are fully captured at the finest resolvable scale within the CG-MMC simulations. The one area of discrepancy between the two distributions appears at intermediate densities ($0.2 \le \rho^* \le 0.5$). The source of this discrepancy is the tendency for the CG-MMC atoms to align with the lattice, particularly at the liquid-vapor boundaries. In other words, the CG-MMC simulation will tend to group particles along the interface so that cells are either mostly full or mostly empty in order to minimize the free energy. This natural alignment can be regarded as unavoidable "pixelation" that is inherent to a grid-based representation.

Finally, the influence of coarse cell size on the density distribution is shown in the inset of Fig. 6 for LJ simulations of a homogeneous liquid phase ($T^* = 1.1$ and $\rho^* = 0.6$). As the coarse cell size increases the distribution become more tightly centered about the mean density value, indicating, as expected, that the magnitude of density fluctuations within the single liquid phase becomes smaller. Gaussian fits to the density distributions demonstrate that the variance of the distribution decreases as $1/V_{cell}$.

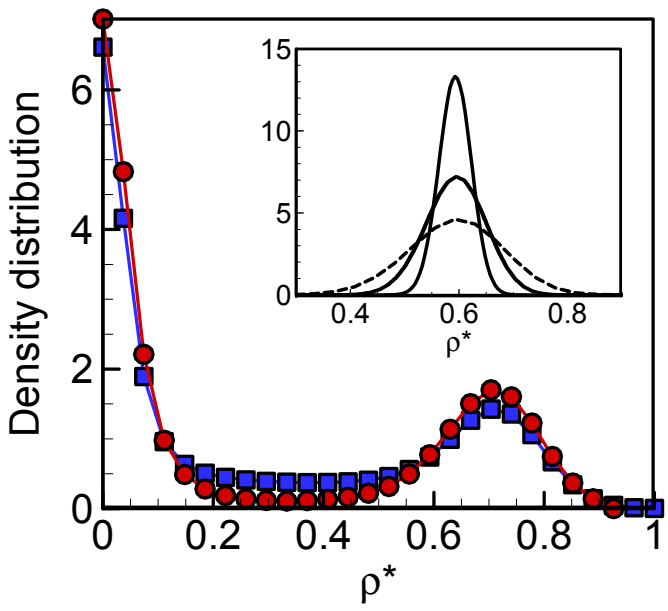


Fig. 6. (Color online) Normalized LJ density distribution functions obtained at $T^* = 0.8$ and $\rho^* = 0.3$. Red circles – CG-MMC, blue squares – *post facto* coarse-grained full-resolution MMC. Coarse cell size for both cases is $L_{cell} = 3\sigma$. Inset shows broadening of density distribution in a homogeneous liquid with decreasing cell size at $T^* = 1.1$ and $\rho^* = 0.6$. Data is shown for $L_{cell} = 3\sigma$ (dashed line), 4σ (dotted line), and 6σ (solid line). Distributions from CG-MMC, *post facto* coarse-grained full-resolution MMC, and Gaussian fits are indistinguishable.

4. CONCLUSIONS

In summary, a spatial coarse-graining method was presented in which an arbitrary interparticle potential is numerically coarse-grained to enable MMC simulations of fluid systems on a rigid lattice. The coarse-grained potential naturally includes the degrees-of-freedom, and the corresponding entropy, that are lost when mapping a fully-resolved, continuous-space problem onto a coarse rigid lattice. The procedure thus ensures that the free energy landscape in the coarse-grained system is fully consistent with the underlying fully-resolved energy landscape. The resulting coarse-grained

representation is able to capture the full VLE characteristics of both atomic and molecular systems, a stringent test of the thermodynamic consistency of the approach and should greatly extend the length and "time" scales accessible to Monte Carlo simulations of non-equilibrium phenomena such as spinodal decomposition. The numerical averaging procedure, which does not require any specific physical insight, appears to be applicable to any (short-ranged) potential and only needs to be performed once before the MMC simulation(s). Although the CG potential pre-computation can be expensive, it can be trivially farmed out to an arbitrary number of compute nodes, limiting the bottleneck associated with this calculation.

ACKNOWLEDGEMENTS:

We gratefully acknowledge financial support from the National Science Foundation (CBET-0730971). We also thank John Crocker and Robert Riggleman for helpful discussions.

REFERENCES:

[1] K. Kremer, and F. Muller-Plathe, Mol Simulat **28**, 729 (2002).
[2] S. Izvekov, and G. A. Voth, J Phys Chem B **109**, 2469 (2005).
[3] J. W. Chu, and G. A. Voth, Biophys J **90**, 1572 (2006).
[4] C. F. Abrams, and K. Kremer, J Chem Phys **115**, 2776 (2001).
[5] Y. T. Wang, S. Izvekov, T. Y. Yan, and G. A. Voth, J Phys Chem B **110**, 3564 (2006).
[6] S. K. Ma, Phys Rev Lett **37**, 461 (1976).
[7] I. Carmesin, and K. Kremer, Macromolecules **21**, 2819 (1988).
[8] M. A. Katsoulakis, A. J. Majda, and D. G. Vlachos, P Natl Acad Sci USA **100**, 782 (2003).
[9] A. Chatterjee, D. G. Vlachos, and M. A. Katsoulakis, J Chem Phys **121**, 11420 (2004).
[10] A. Chatterjee, and D. G. Vlachos, J Comput-Aided Mater **14**, 253 (2007).
[11] A. E. Ismail, G. C. Rutledge, and G. Stephanopoulos, J Chem Phys **118**, 4414 (2003).
[12] A. E. Ismail, G. Stephanopoulos, and G. C. Rutledge, J Chem Phys **118**, 4424 (2003).
[13] M. A. Katsoulakis, and D. G. Vlachos, Phys Rev Lett **84**, 1511 (2000).
[14] S. D. Collins, A. Chatterjee, and D. G. Vlachos, J Chem Phys **129**, 184101 (2008).
[15] J. G. Dai, W. D. Seider, and T. Sinno, J Chem Phys **128**, 194705 (2008).
[16] J. G. Dai, J. M. Kanter, S. S. Kapur, W. D. Seider, and T. Sinno, Phys Rev B **72**, 134102 (2005).
[17] J. Dal, W. D. Seider, and T. Sinno, Mol Simulat **32**, 305 (2006).
[18] S. S. Kapur, M. Prasad, J. C. Crocker, and T. Sinno, Phys Rev B **72**, 014119 (2005).
[19] B. Widom, J Chem Phys **39**, 2808 (1963).
[20] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J Chem Phys **79**, 926 (1983).

[21] M. Mezei, Mol Simulat **9**, 257 (1992).
[22] I. G. Tironi, R. Sperb, P. E. Smith, and W. F. Vangunsteren, J Chem Phys **102**, 5451 (1995).
[23] D. Wolf, P. Keblinski, S. R. Phillpot, and J. Eggebrecht, J Chem Phys **110**, 8254 (1999).
[24] D. Frenkel, and B. Smit, *Understanding molecular simulation : from algorithms to applications* (Academic Press, San Diego, Calif., 2002), 2nd edn.
[25] J. R. Elliott, and L. G. Hu, J Chem Phys **110**, 3043 (1999).