


# Optimization of Coarse-Grained Interaction Potential: Inside Inherent Limitations of Coarse-Graining Methods

Piotr Kowalczyk,<sup>\*,†</sup> Piotr A. Gauden,<sup>‡</sup> and Alina Ciach<sup>§</sup>

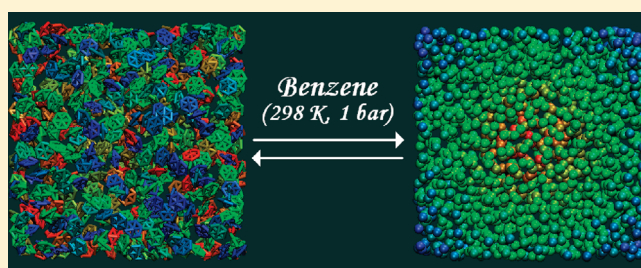
<sup>†</sup>Nanochemistry Research Institute, Department of Chemistry, Curtin University of Technology, P.O. Box U1987, Perth, 6845 Western Australia, Australia

<sup>‡</sup>Department of Chemistry, Physicochemistry of Carbon Materials Research Group, N. Copernicus University, Gagarin St. 7, 87-100 Torun, Poland

<sup>§</sup>Institute of Physical Chemistry, Polish Academy of Science, Kasprzaka Street 44/52, 01-224 Warsaw, Poland

 Supporting Information

**ABSTRACT:** We studied the inherent limitations of coarse-grained (CG) potentials within the recently developed approach (Kowalczyk et al. *J. Phys. Chem. B* **2009**, *113*, 12988–12998). For all studied fluids, the spherically symmetric CG potential constructed according to our scheme modified in this work balances the reproduction of various equilibrium properties (i.e., structural and thermodynamic) measured in CG simulations. The inherent loss of atomistic information at the CG level correlates with the contribution from short-range directional interactions. The highest loss of atomistic information at 298 K and 1 bar is reported for protic liquids (i.e., methanol and acetamide), while the best description at the CG level was obtained for molecular hydrogen and carbon dioxide. The investigated aprotic liquids (i.e., benzene, toluene, and acetone) can be CG into spherically symmetric interaction potentials with some loss of atomistic details. Interestingly, we show that the proposed optimal CG potential reproduces also the interfacial properties of vapor–liquid coexistence for aprotic benzene at 298 K. For all studied fluids, we find that one can easily reproduce structural properties without preserving their cohesive properties or vice versa. However, a general conclusion from our study is the following: an increase in the protic character of a fluid leads to an increase of inherent loss of atomistic details at the CG level.



## I. INTRODUCTION

Computer simulations that take into account all the microscopic forces, i.e., fully atomistic simulations, will in principle provide accurate descriptions of complex systems, such as soft materials or bimolecular systems.<sup>1–4</sup> However, they are strongly limited by the computational resources available that usually do not allow mesoscopically and macroscopically relevant size and time scales to be reached.<sup>5,6</sup> Coarse-graining (CG) of these complex systems is a fundamental problem of modern statistical mechanics for which a general solution has not been found yet.<sup>5–15</sup> Common to all CG methods is that the groups of atoms are clustered into new CG beads.<sup>5</sup> Consequently, interactions between single atoms are replaced by effective interactions between the CG beads. Reduction of molecular scale information during the CG process results in lower resolution models of the systems of interest. Since in any CG procedure some details of the underlying atomistic system are lost, it is crucial to answer the following key questions: *How to integrate out microscopic degrees of freedom to derive optimal effective interactions acting on mesoscopic scales? What are the inherent limitations of the CG methods?*

In the majority of papers connected with CG, the authors have concentrated on the reproduction of a particular equilibrium

property of the atomistic system (e.g., pair correlation function, average density, etc.)<sup>5,16,17</sup> at the CG level. An empirical CG potential with fitting parameters has been commonly used to integrate out microscopic degrees of freedom that are not of direct interest. All these CG schemes are somewhat *ad hoc*; they have the advantage of being computationally inexpensive and are straightforward to implement. However, the range of their applicability and accuracy is usually limited to the adjusted equilibrium properties and thermodynamic conditions. Because there is no rigorous basis for this class of CG methods, they must be used with care.

To the best of our knowledge, the force matching CG methods are the most rigorous and promising present-day approaches.<sup>18–24</sup> They employ a variational principle to determine an interaction potential for a CG model from simulations of an atomically detailed model of the same system.

It reproduces in the most optimal way the force, defined by the strict N-body potential of mean force. However, coarse-grained

**Received:** March 2, 2011

**Revised:** April 27, 2011

**Published:** May 10, 2011

potentials coming from the force matching do not reproduce accurately the pair correlation functions.

On the other hand, the CG potential computed from the pair potential of mean force is optimized in order to predict the correct structure in CG simulations (more precisely, exact-pair distribution functions).<sup>5</sup> At the same time, the exact and CG average energies per bead and their fluctuations are different (see the inverse Monte Carlo method<sup>5,11,12</sup>). In our previous study, we have developed a new methodology to improve the CG potential computed directly from the pair potential of mean force.<sup>25</sup> We showed that simple linear rescaling of the CG potential computed directly from orientation-averaged atomistic forces has a significant impact on various equilibrium properties measured at the CG level. In the current study, we present a definition of an optimal CG potential that balances the reproduction of an average potential energy per bead and density in the CG simulations. We discuss our simulation results with special attention to the role of short-range directional interactions in the CG of studied fluids. As we will show next, these short-range directional interactions are responsible for inherent limitations of CG methods.

## II. THEORY

**II.1. Coarse-Graining.** Details of our CG method are given elsewhere.<sup>25</sup> Here, we present only a brief summary. Similarly to other studies, we compute a pair potential of the orientation-averaged force between two molecules (say, A and B) in fully atomistic simulation. More precisely, the force between the two molecules is averaged over their orientations and over positions and orientations of the remaining molecules in the  $N, P, T$  ensemble. Then, the trial effective potential is given by

$$\Psi(r) = \int_{r_{\min}}^r f(\xi) d\xi \quad \text{for } r \leq r_{\text{cut}} \quad (1)$$

where  $r$  is the distance between the centers of mass of the molecules A and B,  $r_{\text{cut}}$  is the cutoff distance used for calculation of intermolecular potentials in atomistic fluid, and  $r_{\min}$  is the shortest distance between the molecules.  $f = \|f\|$  is the magnitude of the average force  $\mathbf{f}$  between the molecules A and B when the separation between their centers of mass is  $\xi$ , and  $\mathbf{f}$  is given by

$$\mathbf{f}(\xi) = -\left\langle \sum_{i \in A} \sum_{j \in B} \nabla u_{ij}(r_{ij}^{A,B}) \right\rangle_{\xi} \quad (2)$$

where  $\langle \dots \rangle_{\xi}$  denotes the conditional ensemble average, with fixed separation  $\xi$  between the centers of mass of the molecules A and B,  $r_{ij}^{A,B} = \|\mathbf{r}_i^A - \mathbf{r}_j^B\|$  is the distance between the  $i$ th atom in the molecule A and the  $j$ th atom in the molecule B, and  $u_{ij}(x)$  denotes the distance dependent site–site intermolecular potential expressed by (12,6) Lennard-Jones or Columbic potential (see Kowalczyk et al.<sup>25</sup> for more details). The function  $\mathbf{f}(\xi)$  corresponds to the net force between two molecules when their centers are separated by the distance  $\xi$ . The average force given by eq 2 is simply collected in atomistic simulation. For this purpose, the distance  $0 \leq r \leq r_{\text{cut}}$  is divided into bins with a width of 0.02 Å. After an initial equilibration of the atomistic fluid, the equilibrium configurations are swapped with regular frequency. In these configurations, each pair of molecules separated by the distance  $\xi$  was used for the accumulation of the average force between the beads given by eq 2. The trial spherically symmetric potential between the beads is computed from eq 1 by numerical

integration. Previously, we found that the trial effective CG potential computed from eq 1 is characterized by correct asymptotic properties and reproduces the pair distribution function reasonably well, as expected by construction of this effective potential (Kowalczyk et al.<sup>25</sup> for more details). However, it underestimates the average potential energy and its fluctuations for all studied fluids measured in CG simulations.<sup>25</sup> This is because the distribution function for positions of the centers of mass is obtained by integrating the Boltzmann factor over orientational degrees of freedom, while the average energy is obtained by integrating over orientational degrees of freedom the Boltzmann factor times energy (see Kowalczyk et al.<sup>25</sup> for more details).

In our previous paper,<sup>25</sup> we showed that equilibrium properties of CG fluid can be adjusted by introduction of factor  $C$

$$U_c(r) = C \cdot \Psi(r) \quad (3)$$

The unknown factor  $C$  was found from the solution of the following variational functional

$$\min[\langle U_a \rangle - \langle U_c \rangle^2] \quad (4)$$

where  $\langle U_a \rangle$  and  $\langle U_c \rangle$  are the average potential energy per bead computed from the atomistic and CG molecular simulation, respectively. For an unspecified shape of the effective potential  $U_c(r)$ , the result of the procedure given by eq 4 is not unique, but with the assumption given by eq 3, the unknown factor  $C$  can be determined in a unique way.

As we showed previously,<sup>25</sup> for aprotic CG liquids, eq 4 significantly improved the description of atomistic fluids at the CG level. In contrast, for protic CG liquids, we found that adjustment of the average potential energy per bead in CG simulations significantly overpredicted correlation functions measured at the atomistic level. For instance, the CG water looks like ice.<sup>25</sup> The difference between protic and aprotic liquids results from the action of short-range directional interactions. To improve the CG description of various fluids, we reconsider and modify the variational functional given by eq 4.

Let us consider the fundamental properties of liquids. The local structure in liquid is described by the radial distribution function. The local structure depends on density, because the latter depends on the average separation between the nearest neighbors, which is given by the positions of the first peak of the radial distribution function (RDF). The smaller this separation, the larger the density is. The density depends also on the average number of particles in the first shell around the test particle, which in turn determines the height of the first peak of the radial distribution function. One should remember, however, that for dense gases two systems may have the same RDF but different densities and different coordination numbers, because the RDF is normalized by densities.<sup>26</sup> For optimization of the CG procedure, we suggest to consider the quantity which is a function of the factor  $C$

$$\Omega(C) = w \frac{[\langle U_c(C) \rangle - \langle U_a \rangle]^2}{[\langle U_c(C_d) \rangle - \langle U_a \rangle]^2} + (1 - w) \frac{[\langle \rho_c(C) \rangle - \langle \rho_a \rangle]^2}{[\langle \rho_c(C_U) \rangle - \langle \rho_a \rangle]^2} \quad (5)$$

where the weight is  $w \in [0, 1]$ . In eq 5,  $\langle U_c(C) \rangle$  and  $\langle \rho_c(C) \rangle$  are the average potential energy per bead and density measured at the CG level for selected values of  $C$  in eq 3, respectively,  $\langle U_a \rangle$  and  $\langle \rho_a \rangle$  are the average potential energy per bead and density measured in

**Table 1.** Comparison of the Thermodynamic Properties of the Studied Fluids Computed from the  $N,P,T$  CG and Atomistic Simulations<sup>a</sup>

molecule	parameter	atomistic simulation	CG (density) simulation	CG (op) simulation	CG (energy) simulation
acetone	$E_{\text{pot}}$ (kcal/mol)	−6.15	−3.52	−4.83	−6.14
	$\Delta H_{\text{vap}}$ (kcal/mol)	6.74	4.11	5.42	6.73
	$C_p$ (cal/(mol K))	19.0	10.57	10.58	10.46
	$\rho$ (g/cm <sup>−3</sup> )	0.7	0.7	0.75	0.78
benzene	$E_{\text{pot}}$ (kcal/mol)	−6.8	−5.21	−6.07	−6.93
	$\Delta H_{\text{vap}}$ (kcal/mol)	7.39	5.81	6.66	7.52
	$C_p$ (cal/(mol K))	16.26	10.92	11.47	11.32
	$\rho$ (g/cm <sup>−3</sup> )	0.85	0.85	0.86	0.87
carbon dioxide	$E_{\text{pot}}$ (kcal/mol)	−2.32	−1.74	−2.03	−2.31
	$\Delta H_{\text{vap}}$ (kcal/mol)	2.91	2.33	2.63	2.91
	$C_p$ (cal/(mol K))	26.8	14.9	14.1	12.9
	$\rho$ (g/cm <sup>−3</sup> )	0.85	0.85	0.92	0.97
hydrogen	$E_{\text{pot}}$ (kcal/mol)	−0.027	−0.038	−0.032	−0.027
	$\Delta H_{\text{vap}}$ (kcal/mol)	0.62	0.63	0.62	0.62
	$C_p$ (cal/(mol K))	8.18	6.07	6.08	6.09
	$\rho$ (g/cm <sup>−3</sup> )	0.014	0.014	0.014	0.014
toluene	$E_{\text{pot}}$ (kcal/mol)	−7.72	−2.65	−5.15	−7.74
	$\Delta H_{\text{vap}}$ (kcal/mol)	8.31	3.24	5.74	8.34
	$C_p$ (cal/(mol K))	18.6	7.99	8.66	9.38
	$\rho$ (g/cm <sup>−3</sup> )	0.84	0.84	0.84	0.84
acetamide	$E_{\text{pot}}$ (kcal/mol)	−9.76	−3.48	−6.15	−9.83
	$\Delta H_{\text{vap}}$ (kcal/mol)	10.36	4.07	6.74	10.42
	$C_p$ (cal/(mol K))	23.04	11.4	11.38	11.93
	$\rho$ (g/cm <sup>−3</sup> )	0.93	0.93	1.05	1.13
methanol	$E_{\text{pot}}$ (kcal/mol)	−8.05	−2.46	−4.81	−7.98
	$\Delta H_{\text{vap}}$ (kcal/mol)	8.64	3.05	5.4	8.57
	$C_p$ (cal/(mol K))	16.7	11.74	9.91	9.78
	$\rho$ (g/cm <sup>−3</sup> )	0.8	0.8	0.99	1.1

<sup>a</sup> Three CG potentials were studied: CG(density) - CG potential optimized for average density, CG(op) - optimal CG potential computed from eq 5 for  $w = 0.5$ , and CG(energy) - CG potential optimized for average potential energy per bead. Abbreviations:  $E_{\text{pot}}$  - potential energy per bead,  $\Delta H_{\text{vap}}$  - heat of vaporization,  $C_p$  - isobaric heat capacity, and  $\rho$  - density.

atomistic simulation, respectively, and  $C_U$  and  $C_d$  denote factors corresponding to optimized average potential energy per bead and density measured in CG simulations, respectively.  $C = C_d$  corresponds to the minimum of  $[\langle \rho_c(C) \rangle - \langle \rho_a \rangle]^2$ , and for  $C = C_U$ ,  $[\langle U_c(C) \rangle - \langle U_a \rangle]^2$  assumes the minimum. In eq 5, the two terms should be comparable, since for  $C_d \leq C \leq C_U$  they are normalized such that for  $w = 1$  the first term decreases from 1 and for  $w = 0$  the second term increases to 1 for  $C$  increasing from  $C_d$  to  $C_U$ . With this equation, we have a precise definition of the quantity that should be minimized, for each chosen  $w$ . If  $w = 0$ , then the average potential energy per bead is unimportant and the average density is optimized. It means that we simply recover the potential of mean force CG method.<sup>25</sup> If  $w = 1$ , the average potential energy per bead in CG simulation is optimized and the average density is unimportant. How to select  $w$ ? Clearly, if  $w = 0.5$ , we equally balance the reproduction of the average potential energy per bead and density in CG simulations. It is a straightforward and reasonable selection that we consider in all our calculations.

**II.2. Simulation Methodology on the Atomistic and CG Level.** The potential models, computational methodology, and their validation against known experimental data for studied fluids are documented in our previous study.<sup>25</sup> The parameters of

the potential (i.e., Lennard-Jones dispersion parameters and point charges) for all investigated molecules are taken from different force fields to ensure that the presented analysis does not depend on the particular force field but only on intrinsic properties of the studied fluids. We used the following force fields: benzene and toluene (Wick et al.<sup>27</sup> force field for arenes), acetone, acetamide, and methanol (Jorgensen et al.<sup>28,29</sup> OPLS force field), carbon dioxide (Nguyen et al.<sup>30,31</sup> force field), and molecular hydrogen (Belof et al.<sup>32</sup> force field).

The interfacial properties of liquid–vapor coexistence for benzene were computed directly from CG and atomistic simulations. The slab of liquid benzene consisting of 800 atomistic molecules (or beads at CG level) equilibrated at 298 K by the  $N, P, T$  Monte Carlo method was immersed in a rectangular box of size  $L_x L_y L_z$ , where  $L_x = L_y = 38$  Å and  $L_z = 234$  Å. Periodic boundary conditions were applied in the three coordinate directions.<sup>33–36</sup> We focus our attention on the localized interface in the closed system. In the canonical ensemble, the fluctuations of the interface position resulting from fluctuating  $N$  are suppressed. For this reason, we choose the  $N, V, T$  ensemble. All studied systems were equilibrated for  $10^6$  steps in the  $N, V, T$  ensemble by the Monte Carlo method, with  $N$  chosen such that the interface is created in the system. The liquid–vapor interface

may affect energy and local density through the whole simulated volume, and the properties of the interface in the  $N, P, T$  ensemble can be different. Density profiles and pressure tensor components were collected in a production run consisting of an additional  $10^8$  steps. The surface tension was computed from the Irving–Kirkwood mechanical definition<sup>37,38</sup>

$$\gamma = \frac{1}{2} \int_{-L_z/2}^{L_z/2} dz [p_N(z) - p_T(z)] \quad (6)$$

where  $p_N(z) \equiv p_{zz}(z)$  and  $p_T(z) \equiv (1/2)[p_{xx}(z) + p_{yy}(z)]$  are the normal and tangential components of the pressure with respect to the planar interface, respectively. Notice that  $p_N(z)$  and  $p_T(z)$  are the same in the bulk phase because the structure is isotropic in any direction, and they are different from each other only near the interface because the structure can be very anisotropic. The components of the pressure tensor in eq 6 were calculated for each thin slab (i.e.,  $L_x = L_y = 38 \text{ \AA}$ ,  $\Delta z = 0.2 \text{ \AA}$ ) of the simulation box during the Monte Carlo simulation as a canonical average by the Irving–Kirkwood mechanical definition<sup>39,40</sup>

$$p_{\alpha\beta}(z) = \langle \rho(z) \rangle k_B T \mathbf{I} + \frac{1}{A_s} \left\langle \sum_{i=1}^{N-1} \sum_{j>i}^N (\mathbf{r}_{ij})_{\alpha} (\mathbf{F}_{ij})_{\beta} \frac{1}{|z_{ij}|} \cdot \theta \left( \frac{z - z_i}{z_{ij}} \right) \cdot \theta \left( \frac{z_j - z}{z_{ij}} \right) \right\rangle \quad (7)$$

where  $\langle \dots \rangle$  denotes the ensemble average,  $\mathbf{I}$  is the unit tensor,  $k_B T$  is the thermal energy,  $\rho(z) (=n_z/L_x L_y \Delta z)$  denotes the density computed for a slab at  $z$ ,  $A_s \equiv L_x L_y$  is the surface area normal to the  $z$  axis,  $N$  is the number of particles, and  $z_{ij}$  is the distance between two molecular centers of mass.  $\alpha$  and  $\beta$  represent the  $x$ ,  $y$ , and  $z$  directions.  $\theta(x)$  is the unit step function defined by  $\theta(x) = 0$  when  $x < 0$  and  $\theta(x) = 1$  when  $x \geq 0$ .  $\mathbf{F}_{ij}$  in eq 7 is the intermolecular force between molecules  $i$  and  $j$ . It is expressed as the sum of all the site–site forces acting between these two molecules:<sup>39,40</sup>

$$\mathbf{F}_{ij} = \sum_{a=1}^{N_i} \sum_{b=1}^{N_j} (\mathbf{f}_{iajb}) = - \sum_{a=1}^{N_i} \sum_{b=1}^{N_j} \frac{\mathbf{r}_{iajb} dU(r_{iajb})}{r_{iajb}} \quad (8)$$

For the spherically symmetric CG potential, the normal and tangential components of the pressure tensor can be directly computed from the simplified equations<sup>41,42</sup>

$$p_N(z) = \langle \rho(z) \rangle k_B T - \frac{1}{V_s} \left\langle \sum_{ij} \frac{z_{ij}^2}{r_{ij}} \frac{dU(r_{ij})}{dr_{ij}} \right\rangle \quad (9)$$

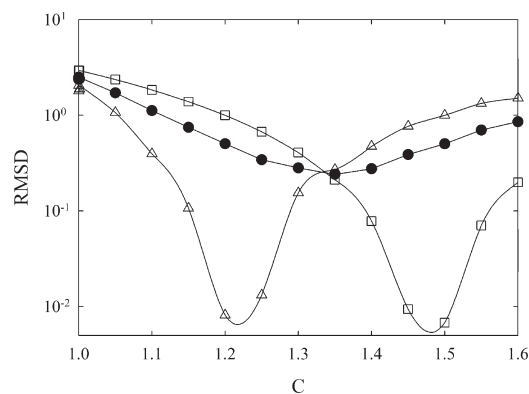
$$p_T(z) = \langle \rho(z) \rangle k_B T - \frac{1}{V_s} \left\langle \sum_{ij} \frac{x_{ij}^2 + y_{ij}^2}{2r_{ij}} \frac{dU(r_{ij})}{dr_{ij}} \right\rangle \quad (10)$$

where  $V_s (=L_x L_y \Delta z)$  denotes the slab volume.

Density profiles calculated from both atomistic and CG simulations were fitted to the following functions:<sup>39,40</sup>

$$\rho(z) = \frac{1}{2}(\rho_l + \rho_v) - \frac{1}{2}(\rho_l - \rho_v) \tanh \left[ \frac{2 \cdot (z - z_0)}{d} \right] \quad (11)$$

$$\rho(z) = \frac{1}{2}(\rho_l + \rho_v) - \frac{1}{2}(\rho_l - \rho_v) \operatorname{erf} \left[ \frac{\sqrt{\pi} \cdot (z - z_0)}{d} \right] \quad (12)$$



**Figure 1.** Variation of the root-mean-square deviation (rmsd) for optimized average density (open triangles) and potential energy per bead (open squares) with the factor  $C$  in eq 3 for liquid benzene at 298 K and 1 bar. The optimal CG potential (black circles) is computed from eq 5 for  $w = 0.5$ . Note that the CG potential minimum at  $C = 1.35$  balances density and potential energy error measured in the CG simulations.

where the fitting parameter  $\rho_l$  is the bulk liquid density,  $\rho_v$  is the bulk vapor density,  $z_0$  is the position of the Gibbs dividing surface, and  $d$  denotes the average thickness of the liquid–vapor interface.

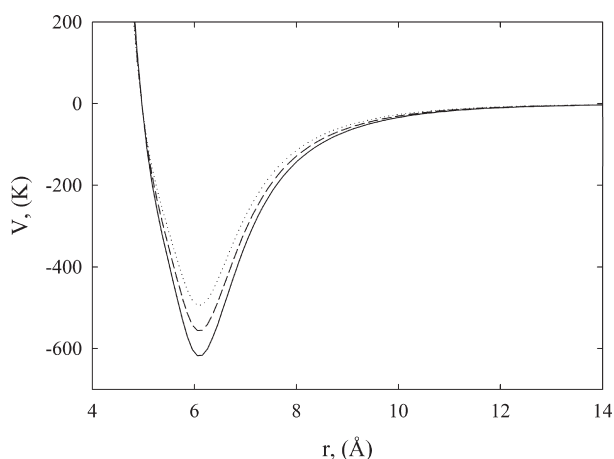
### III. RESULTS AND DISCUSSION

At the beginning of this section, we want to emphasize that we will consider the agreement between the spectrum of equilibrium properties computed in atomistic and CG simulations of selected fluids.

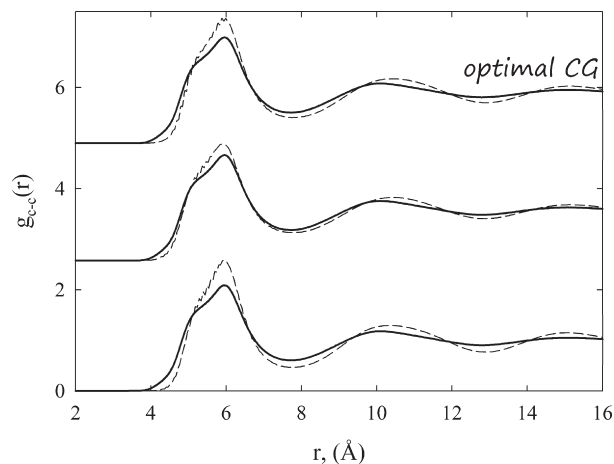
We begin a discussion of the results by considering atomistic and CG simulations of molecular hydrogen, carbon dioxide, benzene, toluene, and acetone at 1 bar and 298 K, as is presented in Table 1. We start the analysis of the thermodynamic properties computed for molecular hydrogen at 298 K and 1 bar. As could be expected, CG of molecular hydrogen by the spherically symmetric CG potential is a very effective method (see Figures 1S–4S in the Supporting Information). At the studied point of the phase diagram, molecular hydrogen molecules behave as low-density gas of noninteracting particles. Due to small molecular mass and frequent rotations, hydrogen dumbbell-shaped molecules behave as spherical particles interacting via isotropic potential. Indeed, the isotropic Silvera–Goldman potential for para-hydrogen has been successfully used for simulations of molecular hydrogen under various operating conditions.<sup>43–47</sup>

Now we would like to consider CG of liquid carbon dioxide. Table 1 shows that the optimal spherically symmetric CG potential slightly overestimates density and significantly underestimates isobaric heat capacity computed from atomistic simulations. What is more interesting, all studied CG potentials significantly washed out the spectrum of fluctuations in atomistic carbon dioxide (see Figures 5S–8S in the Supporting Information). We argue that the observed disagreement is not associated with the CG method used here, but it results from inadequacy of any spherically symmetric potential. Notice that the studied operating temperature is very close to the liquid–gas critical one for carbon dioxide, i.e., 304.3 K. Isotropic CG potential is in principle not able to reproduce a broad range of fluctuations expected near critical temperature, as is shown in





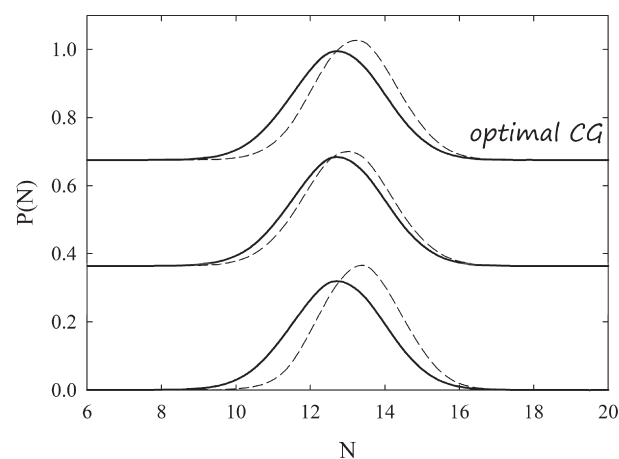
**Figure 2.** CG potentials computed for liquid benzene at 298 K and 1 bar. Abbreviations: solid line - CG potential optimized for average potential energy per bead, dashed line - optimal CG potential computed from eq 5 for  $w = 0.5$ , and dotted line - CG potential optimized for density.



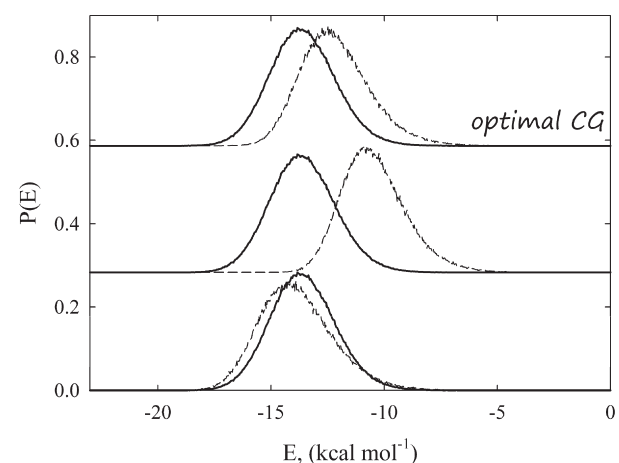
**Figure 3.** Radial distribution functions (mass-center-to-mass-center) for liquid benzene computed from CG (dashed lines) and atomistic (solid lines) simulation at 298 K and 1 bar. The bottom panel corresponds to the CG potential optimized for average potential energy per bead, whereas the central panel corresponds to CG potential optimized for average density. The upper optimal CG potential is computed from eq 5 for  $w = 0.5$ .

Figure 8S in the Supporting Information. We believe that this observation is not limited to phase transitions in simple fluids, but the same effect is expected for other phenomena affected by a broad spectrum of fluctuations, such as helix–coil transition in proteins, etc. Therefore, the use of CG methods for a location of the phase transitions cannot be precise.

Now, we would like to discuss the influence of the CG method on the equilibrium properties of aprotic benzene liquid at 298 K and 1 bar. As other aprotic molecules, benzene ones are not able to form H-bonds by themselves. Figure 1 depicts the variation of the root-mean-square deviation (rmsd) for optimized average density and potential energy per bead with factor  $C$  in eq 3. As we argued in our previous study, the rmsd minima for average density and potential energy per bead correspond to different CG potentials, as is displayed in Figure 2. Even for aprotic

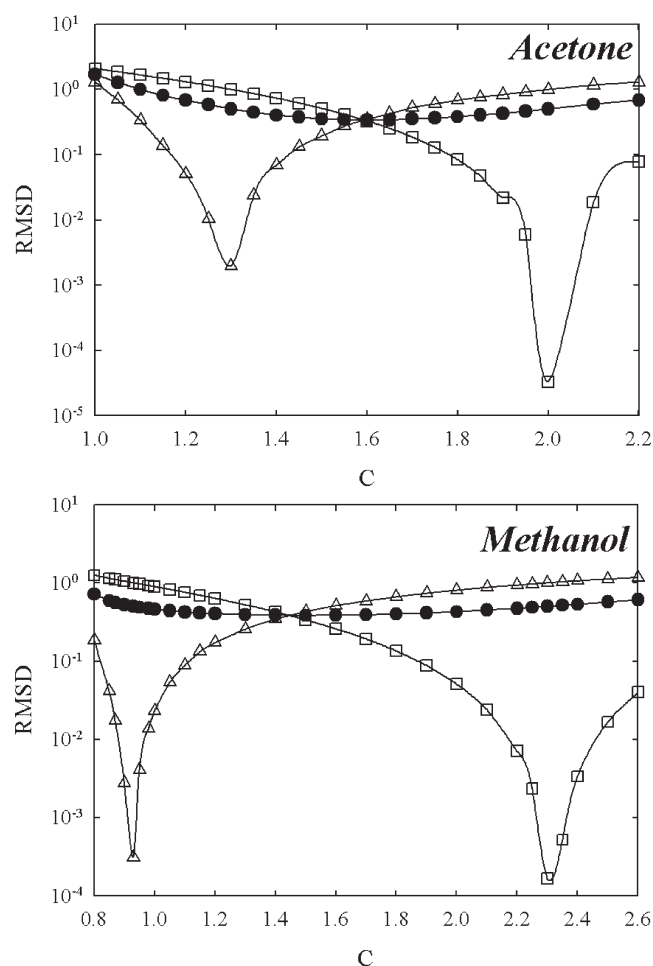


**Figure 4.** Probability distributions of the number of benzene molecules for the first solvation shell computed from CG (dashed lines) and atomistic (solid lines) simulation of liquid benzene at 298 K and 1 bar. The bottom panel corresponds to the CG potential optimized for average potential energy per bead, whereas the central panel corresponds to the CG potential optimized for average density. The upper optimal CG potential is computed from eq 5 for  $w = 0.5$ .



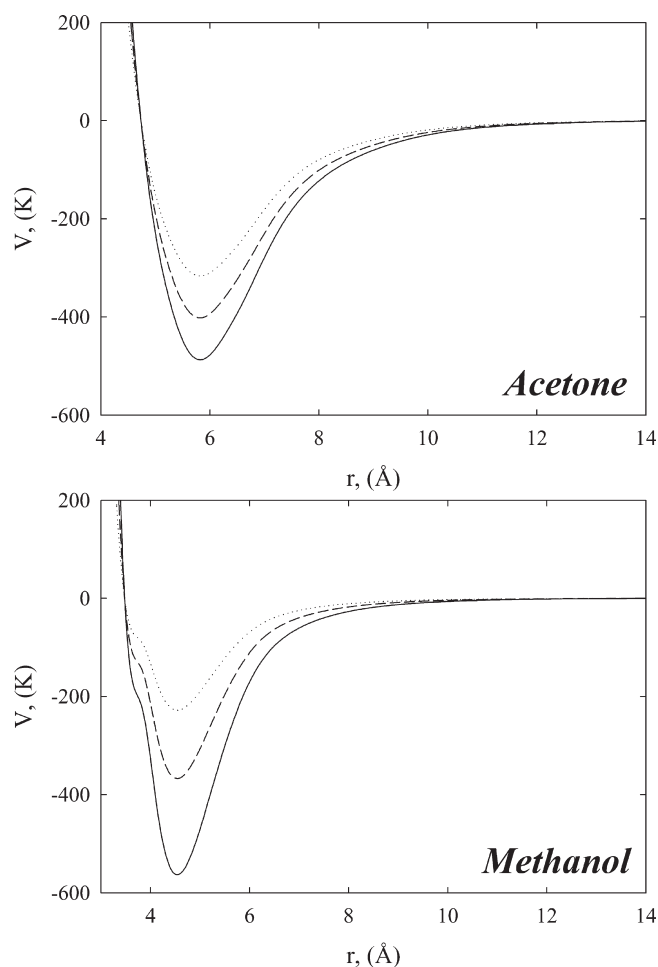
**Figure 5.** Distributions of total bonding energies for monomers in liquid benzene computed from CG (dashed lines) and atomistic (solid lines) simulation at 298 K and 1 bar. The bottom panel corresponds to the CG potential optimized for average potential energy per bead, whereas the central panel corresponds to the CG potential optimized for density. The upper optimal CG potential is computed from eq 5 for  $w = 0.5$ .

benzene molecules, it is not possible to simultaneously describe average potential energy per bead and density in CG simulations with high precision. In dense liquid, benzene molecules are not interacting via spherically symmetric interaction potential, which fundamentally limits the accuracy of any CG scheme. As expected, the CG potential optimized for average potential energy per bead is characterized by a deeper well-depth and wider attractive range. This increases both the heat of vaporization and the isobaric heat capacity computed in CG simulations (see Figure 1). To preserve the average density in atomic/CG simulations, one needs to reduce the well-depth and attractive range of the CG potential (see Figure 2). The optimal CG potential computed from eq 5 corresponds to  $C = 1.35$ . By construction, it balances the reproduction of an average potential



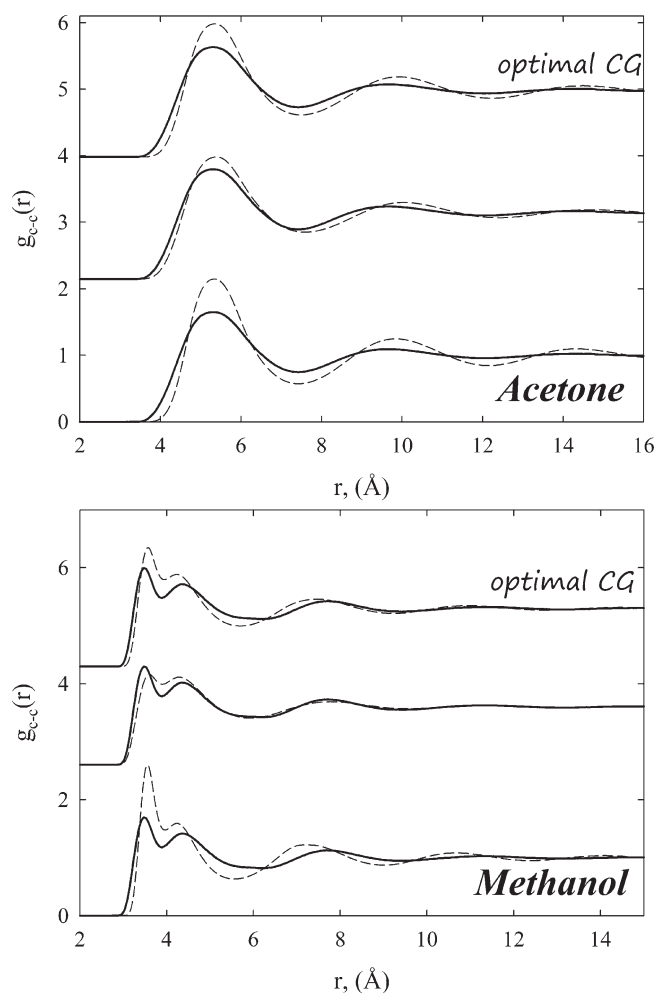
**Figure 6.** Variation of the rmsd for optimized average density (open triangles) and potential energy per bead (open squares) with factor C in eq 3 for liquid acetone and methanol at 298 K and 1 bar. The optimal CG potential (black circles) is computed from eq 5 for  $w = 0.5$ .

energy per bead and density in CG simulations. However, what is more interesting, this CG potential minimizes deviations for the spectrum of equilibrium properties computed in CG and atomistic simulations of liquid benzene. Let us discuss this important feature in detail. Figure 3 displays radial distribution functions for liquid benzene computed from three CG potentials and atomistic simulation at 298 K and 1 bar. By construction, the CG potential optimized for average density reproduces the local structure of atomistic liquid benzene almost quantitatively. As the average energy per CG bead increases, the structure in radial distribution functions gradually increases. We found that at short and longer distances the packing of CG benzene beads is higher in comparison to atomistic ones. This results in a higher density of the CG beads, as is shown in Table 1. Figure 4 illustrates the impact of the CG potential on the probability distributions of the number of benzene molecules for the first solvation shell. As expected, the local structure of benzene liquid computed from CG potential optimized for average density is almost indistinguishable from that computed from atomistic simulations. At higher average energy per bead, the number of CG beads in the first solvation shell increases. This observation is consistent with the results presented in Figure 3 and Table 1. Figure 5 depicts distributions of total bonding energies for monomers in liquid



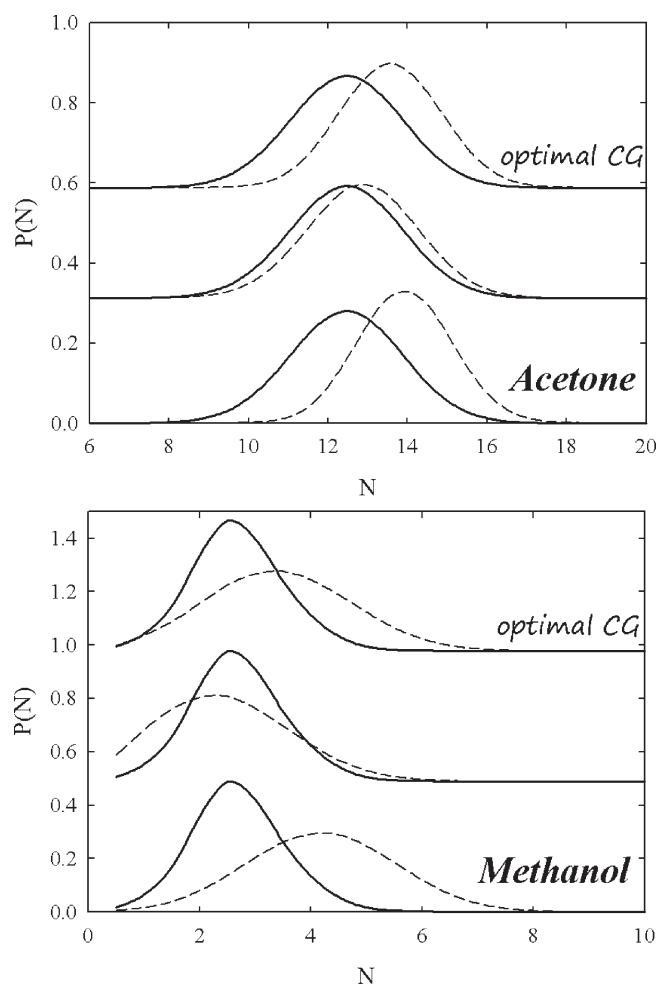
**Figure 7.** CG potentials computed for liquid acetone and methanol at 298 K and 1 bar. Abbreviations: solid line - CG potential optimized for average potential energy per bead, dashed line - optimal CG potential computed from eq 5 for  $w = 0.5$ , and dotted line - CG potential optimized for density.

benzene computed from CG and atomistic simulations at 298 K and 1 bar. In contrast to the reproduction of structure in CG simulations, we notice an opposite situation for energetic properties. By construction, CG potential optimized for average energy per bead reproduces the bonding energies in liquid benzene almost quantitatively (see Table 1 and Figure 5). Reproduction of average density in CG simulations results in gradual underestimation of bonding energies computed from atomistic simulations. This is accompanied by a large underestimation of heat of vaporization and isobaric heat capacity at the CG level, as is shown in Table 1. Taking into account results presented in Figures 3–5 and Table 1, we concluded that the CG potential constructed from eq 5 compromises the reproduction of studied equilibrium properties of liquid benzene in CG simulations. Here, we would like to stress that our theoretical results are general. They emphasize the deficiency of isotropic potential in description of atomistic molecules interacting via anisotropic potential. Now, we would like to test the CG potentials against more complex liquids, namely, acetone and methanol, at 298 K and 1 bar. As benzene, acetone molecules are aprotic, but methanol ones are protic with significant contribution of short-range directional H-bonds. Figure 6 shows the variation of rmsd



**Figure 8.** Radial distribution functions (mass-center-to-mass-center) for liquid acetone and methanol computed from CG (dashed lines) and atomistic (solid lines) simulation at 298 K and 1 bar. The bottom panel corresponds to the CG potential optimized for average potential energy per bead, whereas the middle panel corresponds to the CG potential optimized for average density. The upper optimal CG potential is computed from eq 5 for  $w = 0.5$ .

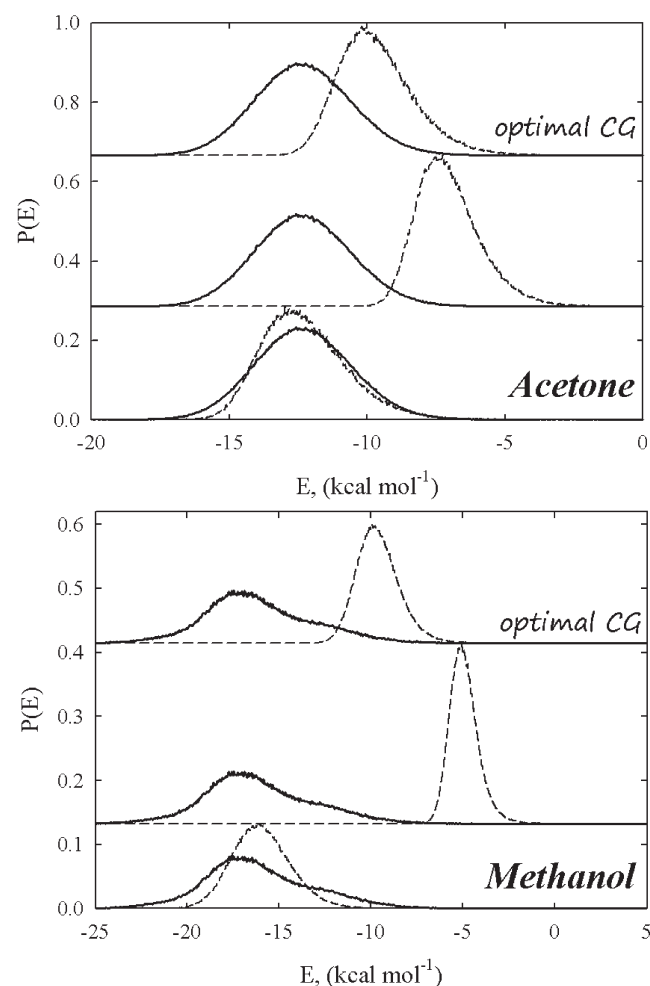
for optimized average density and potential energy per bead with factor  $C$  in eq 3. The generic features of Figure 6 are similar to those observed in Figure 1. We notice two well-separated minima corresponding to different values of  $C$  in eq 3. For both acetone and methanol, the rmsd corresponding to average density is minimal for low values of  $C$ . To preserve the average energy per bead in CG simulations, one needs to increase the well-depth and attractive range of the CG potential (see Figure 7). As for benzene, the optimal CG potential computed from eq 5 balances the reproduction of average density and energy per bead in CG simulations. Figure 8 displays radial distribution functions for liquid acetone and methanol computed from three CG potentials and atomistic simulation at 298 K and 1 bar. As previously, the CG potential optimized for average density reproduces the local structure of atomistic liquids almost quantitatively. As the average energy per CG bead increases, the structure in radial distribution functions gradually increases. The reader perhaps noticed that spherical beads tend to pack better than atomistic molecules. This results in higher density of the CG liquids in



**Figure 9.** Probability distributions of the number of acetone and methanol molecules for the first solvation shell computed from CG (dashed lines) and atomistic (solid lines) simulation of liquid benzene at 298 K and 1 bar. The bottom panel corresponds to the CG potential optimized for average potential energy per bead, whereas the middle panel corresponds to the CG potential optimized for average density. The upper optimal CG potential is computed from eq 5 for  $w = 0.5$ .

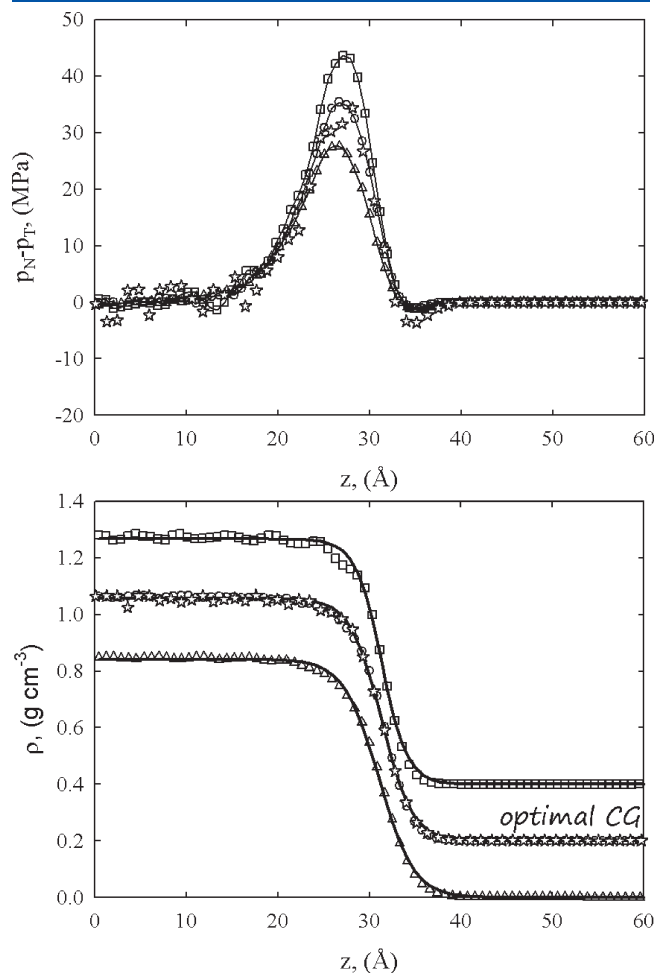
comparison to atomistic ones, as is presented in Table 1. Figure 9 illustrates an influence of CG potential on the distribution of acetone and methanol molecules for the first solvation shell. In contrast to liquid benzene, we found that an increase of factor  $C$  in eq 3 leads to a significant shift of the histogram to larger values of nearest neighbors. Moreover, for methanol, all peaks are broader. Such a difference in the behavior of probability distribution for the first solvation shell is a simple consequence of short-range directional interactions between protic molecules. Indeed, packing of acetone beads is not as much affected by the CG process as packing of methanol beads. Distributions of total bonding energies for monomers in liquid acetone and methanol computed from CG simulations are significantly affected by the CG process. Decrease of the well-depth and the range of attraction in the CG potential leads to underestimated potential energy per bead in the CG simulations in comparison to atomistic ones. What is more interesting, we notice a significant reduction in dispersion of peak for CG methanol, as is displayed in Figure 10. This indicates that a broad range of fluctuations in atomistic liquid methanol cannot be reproduced by spherically

symmetric CG potential. As a result, the isobaric heat capacity for protic liquids, such as methanol, is significantly underestimated (see Table 1). Here, we want to stress that any other CG method based on isotropic potential cannot reproduce the broad range of fluctuations expected for atomistic protic liquids. The presented CG method with function given by eq 5 is designed to reproduce the generic behavior (i.e., spectrum of thermodynamic and



**Figure 10.** Distributions of total bonding energies for monomers in liquid acetone and methanol computed from CG (dashed lines) and atomistic (solid lines) simulation at 298 K and 1 bar. The bottom panel corresponds to the CG potential optimized for average potential energy per bead, whereas the middle panel corresponds to the CG potential optimized for density. The upper optimal CG potential is computed from eq 5 for  $w = 0.5$ .

structural properties) of atomistic molecules by isotropic CG beads. As we have shown, this can be successfully achieved for low-density fluids or aprotic liquids, because the contribution from the short-range direction-dependent interactions is negligible. For protic liquids, any simple isotropic CG potential leads to a low resolution model. However, as we presented here, it is



**Figure 11.** Pressure tensor and density profiles along the vapor–liquid interface computed for benzene (filled stars) at 298 K and 1 bar. Abbreviations: filled stars correspond to full atomistic simulations, open triangles correspond to the CG potential optimized for density, open squares correspond to the CG potential optimized for average potential energy per bead, and open circles correspond to the optimal CG potential computed from eq 5 for  $w = 0.5$ . All density profiles obtained from our simulations were fitted using the hyperbolic tangent function (solid lines).

**Table 2.** Comparison of the Interfacial Vapor–Liquid Properties for Benzene at 298 K Computed from the  $N,V,T$  CG and Atomistic Simulations<sup>a</sup>

molecule	parameter	atomistic simulation	CG (density) simulation	CG (op) simulation	CG (energy) simulation
benzene	$\gamma$ (dyn/cm)	$32.9 \pm 1.6$	$25.6 \pm 0.4$	$32.0 \pm 0.7$	$37.4 \pm 1.5$
	$d$ (Å)	6.6 (7.0)	7.6 (8.1)	6.7 (7.1)	5.7 (6.0)
	$\rho_l$ (g/cm <sup>3</sup> )	0.85 (0.85)	0.84 (0.84)	0.86 (0.86)	0.87 (0.87)
	$\rho_v$ (g/cm <sup>3</sup> )	$1.6 \times 10^{-4}$ ( $6.3 \times 10^{-4}$ )	$0.0$ ( $1.5 \times 10^{-4}$ )	$0.0$ (0.0)	$0.0$ (0.0)

<sup>a</sup> Three CG potentials were studied: CG(density) - CG potential optimized for average density, CG(op) - optimal CG potential computed from eq 5 for  $w = 0.5$ , and CG(energy) - CG potential optimized for average potential energy per bead. Abbreviations:  $\gamma$  - surface tension,  $d$  - average thickness of the liquid–vapor interface,  $\rho_l$  - bulk liquid density, and  $\rho_v$  - bulk vapor density. All properties in parentheses were computed from eq 12, whereas the other properties, from eq 11.



possible to define an optimal CG potential that minimizes the differences between various equilibrium properties of atomistic liquids and isotropic CG beads.

Finally, we want to discuss an impact of CG potential on the vapor–liquid equilibrium of model aprotic liquid, i.e., benzene at 298 K. Because of the low contribution from short-range directional interactions in atomistic benzene liquid, we would expect that isotropic CG beads are able to approximate its properties in the vapor–liquid interfacial region. Further, we notice that our aim is to predict the interfacial properties of benzene from CG of pure liquid benzene molecules at 298 K. Table 2 and Figure 11 illustrate the impact of the CG potential on the interfacial properties computed from atomistic and CG simulations. We have found significant variation of interfacial properties, such as surface tension and thickness of the interface with factor  $C$  in eq 3. As  $C$  increases, the value of the surface tension increases, but the thickness of the interfacial layer decreases. Not surprisingly, the increase of the CG potential well-depth and attraction range (i.e., cohesive energy in CG liquids) lead to higher values of surface tension and progressive shirking of interfacial layer thickness. That is why the density profile is the steepest and the values of the pressure tensor through the vapor–liquid interface are the highest for the CG potential optimized for average potential energy per bead (see Figure 11). Interestingly, the optimal CG potential computed from eq 5 approximates the interfacial properties computed directly from atomistic simulations of vapor–liquid benzene coexistence reasonably well. Thus, we conclude that, for aprotic liquids, the interfacial properties computed from atomic simulations can be successfully reproduced by the spherically symmetric isotropic CG potential. Nevertheless, we want to point out that both the surface tension and the interfacial layer thickness are very sensitive to the details of the CG potential. Further investigation of our CG method for more complex interfacial systems will be a subject of future works.

#### IV. CONCLUSIONS

We studied the inherent loss of atomistic information in CG simulations of selected fluids, e.g., molecular hydrogen, carbon dioxide, benzene, toluene, acetone, acetamide, and methanol, under normal conditions. For this purpose, we optimized the CG potential that balances the reproduction of various equilibrium properties (i.e., structural and thermodynamic) at the CG level. We show that inherent loss of atomistic information in CG simulations correlates with the contribution from short-range directional interactions. This explains the effective CG of molecular hydrogen and carbon dioxide and significant loss of microscopic details in CG simulations of the studied protic liquids near ambient conditions. For aprotic liquids, we find that balance between the average potential energy per bead (i.e., strength of attractive potential) and correlations in liquid (i.e., radial distribution function) leads to reasonable results.

#### ■ ASSOCIATED CONTENT

**S Supporting Information.** Atomistic and CG simulation results corresponding to molecular hydrogen and carbon dioxide at 298 K and 1 bar. This material is available free of charge via the Internet at <http://pubs.acs.org>.

#### ■ AUTHOR INFORMATION

##### Corresponding Author

\*Phone: +61 (08) 92667882. E-mail: [piotr.kowalczyk@curtin.edu.au](mailto:piotr.kowalczyk@curtin.edu.au).

#### ■ ACKNOWLEDGMENT

P.K. acknowledges the Curtin University of Technology for a senior research fellowship. P.A.G. acknowledges the use of the computer cluster at Poznan Supercomputing and Networking Centre as well as the Information and Communication Technology Centre of the Nicolaus Copernicus University (Torun, Poland). A.C. acknowledges partial support by the Polish Ministry of Science and Higher Education, Grant No. NN 202 006034. P.A.G. acknowledges partial support from grant NN 204 288634 (2008–2011). P.K. acknowledges partial support by the Office of Research & Development, Curtin University of Technology, Grant CRF10084.

#### ■ REFERENCES

- (1) Cheatham, T. E., III. *Annu. Rep. Comput. Chem.* **2005**, *1*, 75.
- (2) Levy, Y.; Onuchic, J. N. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389.
- (3) Okazaki, K.-I.; Koga, N.; Takada, S.; Onuchic, J. N.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11844.
- (4) Comer, J.; Dimitrov, V.; Zhao, Q.; Timp, G.; Aksimentiev, A. *Biophys. J.* **2009**, *96*, 593.
- (5) Voth, G. A. In *Coarse-Graining of Condensed Phase and Biomolecular Systems*; Voth, G. A., Ed.; CRC Press: Boca Raton, FL, 2008.
- (6) Lyubartsev, A.; Mirzoev, A.; Chen, L. J.; Laaksonen, A. *Faraday Discuss.* **2010**, *144*, 43.
- (7) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812.
- (8) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 2469.
- (9) Izvekov, S.; Parrinello, M.; Burnham, C. J.; Voth, G. A. *J. Chem. Phys.* **2004**, *120*, 10896.
- (10) Hooper, J. B.; Bedrov, D.; Smith, G. D. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2034.
- (11) Reith, D.; Meyer, H.; Müller-Plathe, F. *Comput. Phys. Commun.* **2002**, *148*, 299.
- (12) Reith, D.; Pütz, M.; Müller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1624.
- (13) Lyubartsev, A. P. *Eur. Biophys. J.* **2005**, *35*, 53.
- (14) Lyubartsev, A. P.; Karttunen, M.; Vattulainen, I.; Laaksonen, A. *Soft Mater.* **2003**, *1*, 121.
- (15) Tòth, G. *J. Phys.: Condens. Matter* **2007**, *19*, 335220.
- (16) Chaimovich, A.; Shell, M. S. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1901.
- (17) Kamio, K.; Moorthi, K.; Theodorou, D. N. *Macromolecules* **2007**, *40*, 710.
- (18) Ercolessi, F.; Adams, J. B. *Europhys. Lett.* **1994**, *26*, 583.
- (19) Ercolessi, F.; Parrinello, M.; Tosatti, E. *Philos. Mag. A* **1988**, *58*, 213.
- (20) Izvekov, S.; Voth, G. A. *J. Chem. Phys.* **2005**, *123*, 134105.
- (21) Izvekov, S.; Chung, P. W.; Rice, B. M. *J. Chem. Phys.* **2010**, *133*, 064109.
- (22) Noid, W. G.; Chu, J.-W.; Ayton, G. S.; Krishna, V.; Izvekov, S.; Voth, G. A.; Das, A.; Andersen, H. C. *J. Chem. Phys.* **2008**, *128*, 244114.
- (23) Noid, W. G.; Liu, P.; Wang, Y.; Chu, J.-W.; Ayton, G. S.; Izvekov, S.; Andersen, H. S.; Voth, G. A. *J. Chem. Phys.* **2008**, *128*, 244115.
- (24) Chu, J. W.; Ayton, G. S.; Izvekov, S.; Voth, G. A. *Mol. Phys.* **2007**, *105*, 167.
- (25) Kowalczyk, P.; Gauden, P. A.; Ciach, A. *J. Phys. Chem. B* **2009**, *113*, 12988.
- (26) Sakai, H.; Stillinger, F. H.; Torquato, S. *J. Chem. Phys.* **2002**, *117*, 297.
- (27) Wick, C. D.; Siepmann, J. I.; Klotz, W. L.; Schureb, M. R. *J. Chromatogr., A* **2002**, *954*, 181.
- (28) Jorgensen, W. L.; Jenson, C. J. *Comput. Chem.* **1998**, *19*, 1179.
- (29) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

- (30) Nguyen, T. X. Characterization of Nanoporous Carbons. Ph.D. Thesis, The University of Queensland, Brisbane, Australia, 2006.
- (31) Kowalczyk, P.; Furmaniak, S.; Gauden, P. A.; Terzyk, A. P. *J. Phys. Chem. C* **2010**, *114*, 21465.
- (32) Belof, J. L.; Stern, A. C.; Space, B. *J. Chem. Theory Comput.* **2008**, *4*, 1332.
- (33) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon: Oxford, U.K., 1987.
- (34) Frenkel, D.; Smit, B. *Understanding Molecular Simulation From Algorithms To Applications*; Academic Press: London, 1996.
- (35) Kowalczyk, P.; Ciach, A.; Neimark, A. V. *Langmuir* **2008**, *24*, 6603.
- (36) Kowalczyk, P.; Holyst, R.; Terrones, M.; Terrones, H. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1786.
- (37) Hill, T. L. *Introduction to Statistical Mechanics*; Dover: New York, 1986.
- (38) Jang, S. S.; Lin, S.-T.; Maiti, P. K.; Blanco, M.; Goddard, W. A.; Shuler, P.; Tang, Y. *J. Phys. Chem. B* **2004**, *108*, 12130.
- (39) Rowlinson, J. S.; Widom, B. *Molecular Theory of Capillarity*; Clarendon: Oxford, U.K., 1982.
- (40) Biscay, F.; Ghoufi, A.; Goujon, F.; Lachet, V.; Malfreyt, P. *J. Chem. Phys.* **2009**, *130*, 184710.
- (41) Kirkwood, J. G.; Buff, F. P. *J. Chem. Phys.* **1949**, *17*, 338.
- (42) Alejandre, J.; Tildesley, D. J.; Chapela, G. A. *J. Chem. Phys.* **1995**, *102*, 4574.
- (43) Silvera, I. F.; Goldman, V. J. *J. Chem. Phys.* **1978**, *69*, 4209.
- (44) Kowalczyk, P.; Holyst, R.; Terzyk, A. P.; Gauden, P. A. *Langmuir* **2006**, *22*, 1970.
- (45) Kowalczyk, P.; Tanaka, H.; Holyst, R.; Kaneko, K.; Ohmori, T.; Miyamoto, J. *J. Phys. Chem. B* **2005**, *109*, 17174.
- (46) Kowalczyk, P.; Gauden, P. A.; Terzyk, A. P. *J. Phys. Chem. B* **2008**, *112*, 8275.
- (47) Kowalczyk, P.; Gauden, P. A.; Terzyk, A. P.; Furmaniak, S. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11351.