A general and efficient Monte Carlo method for sampling intramolecular degrees of freedom of branched and cyclic molecules

Jindal K. Shah' and Edward J. Maginn

Citation: The Journal of Chemical Physics **135**, 134121 (2011); doi: 10.1063/1.3644939
View online: http://dx.doi.org/10.1063/1.3644939
View Table of Contents: http://aip.scitation.org/toc/jcp/135/13
Published by the American Institute of Physics

A general and efficient Monte Carlo method for sampling intramolecular degrees of freedom of branched and cyclic molecules

Jindal K. Shah[1,2,a)] and Edward J. Maginn[1]
[1]*Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, Indiana 46556, USA*
[2]*The Center for Research Computing, University of Notre Dame, Notre Dame, Indiana 46556, USA*

(Received 5 July 2011; accepted 11 September 2011; published online 7 October 2011)

A simple and easily implemented Monte Carlo algorithm is described which enables configurational-bias sampling of molecules containing branch points and rings with endocyclic and exocyclic atoms. The method overcomes well-known problems associated with sequential configurational-bias sampling methods. A "reservoir" or "library" of fragments are generated with known probability distributions dependent on stiff intramolecular degrees of freedom. Configurational-bias moves assemble the fragments into whole molecules using the energy associated with the remaining degrees of freedom. The methods for generating the fragments are validated on models of propane, isobutane, neopentane, cyclohexane, and methylcyclohexane. It is shown how the sampling method is implemented in the Gibbs ensemble, and validation studies are performed in which the liquid coexistence curves of propane, isobutane, and 2,2-dimethylhexane are computed and shown to agree with accepted values. The method is general and can be used to sample conformational space for molecules of arbitrary complexity in both open and closed statistical mechanical ensembles. [doi:10.1063/1.3644939]

I. INTRODUCTION

Atomistic Monte Carlo simulations are an important tool for computing the thermodynamic properties and phase equilibria of pure compounds and their mixtures. A number of sophisticated Monte Carlo techniques, including Gibbs ensemble[1] and grand canonical ensemble with histogram reweighting,[2,3] have been developed for this purpose. Application of these techniques to complex molecules typically requires some type of biasing strategy to properly sample the many coupled degrees of freedom and to enable moves to overcome large free energy barriers.

One such biasing strategy is configurational-bias Monte Carlo (CBMC),[4–7] a method commonly used to generate conformations of chain molecules. In CBMC, molecules are "grown" sequentially with each increment added in a biased fashion. The efficiency of the method decreases rapidly as the chain length increases, primarily due to the fact that a small perturbation in the interior of the molecule leads to a large displacement of the end sections. Pant and Theodorou[8] invented an internal rebridging scheme referred to as concerted rotation (ConRot) that allows conformational rearrangement of interior segments. Their approach focused on perturbing two "driver" dihedral angles at the end of a segment separated by five to eight pseudoatoms. Deem and Bader combined the CBMC technique with the ConRot formalism to efficiently generate conformations of linear and cyclic peptides.[9,10] In a subsequent publication, Wu and Deem showed that an analytical solution exists that has at most 16 solutions to the rebridging problem and applied the methodology to study cis/trans isomerization in proline containing cyclic peptides.[11] Many other enhancements to CBMC have been reported over the years.

The development of CBMC and other specialized sampling methods address, to a large extent, conformational flexibility associated with soft degrees of freedom such as dihedral angles. However, the introduction of bond angle flexibility in all-atom models and branched molecules with united-atom representations present additional challenges in sampling of intramolecular degrees of freedom. In the case of linear or branched portions of a molecule, the generation of correct bond angle distributions around a branch point, i.e., around a site that is bonded to more than two sites (Fig. 1) is difficult due to the interdependence of bond angles. This is precisely the reason sequential selection of individual bond angles by the Boltzmann rejection technique fails to produce a correct bond angle distribution for a branch point.[12] Simultaneous angle generation with Boltzmann rejection can give the correct distribution, but at an enormous computational cost. As shown below, it is virtually impossible to use Boltzmann rejection techniques on a species like that shown in Fig. 1.

To overcome these limitations, two classes of algorithms have been developed. The first is the "coupled–decoupled" scheme proposed by Martin and Siepmann.[13] It is based on the concept of having a separation between "hard" and "soft" degrees of freedom. The energy of hard degrees of freedom such as bond angles is very sensitive to small perturbations in atom position, while the energy of soft degrees of freedom such as dihedral angles and intramolecular nonbonded interactions is a much weaker function of atom position. In this scheme, some hard and soft degrees of freedom are "decoupled" and chosen independently while others are "coupled" and selected simultaneously during a CBMC move. The

a)Author to whom correspondence should be addressed. Electronic mail: jshah@nd.edu.

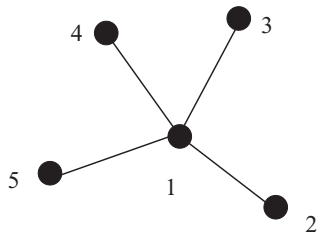

FIG. 1. A schematic representation of a "branch point" fragment containing a central atom directly bonded to four atoms. Such a fragment represents a united-atom model of neopentane, or it might represent the CH_3–C group of an all-atom alkane, with the central atom being the methyl carbon.

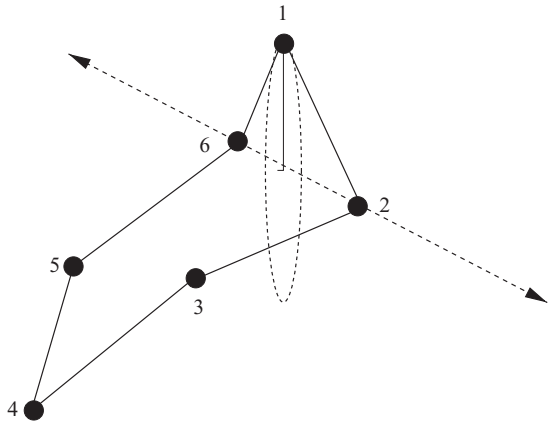

FIG. 2. Schematic representation of the crank-shaft move. Atom 1 is rotated about the axis passing through atoms 2 and 6. The move results in perturbation of angles 561 and 123, and all the dihedral angles in the molecule.

second type of approach, and the focus of the present study, also decouples hard and soft degrees of freedom, but differs from the Martin-Siepmann algorithm in that it utilizes a library of conformations having a predetermined probability distribution based on the hard degrees of freedom. In the so-called "branch point sampling" (BPS) scheme introduced by Macedonia and Maginn,[14] conformations of chain molecules are sampled by first decomposing the molecules into fragments that can be reassembled via CBMC to reconstitute the molecule. Fragments are defined such that each one is connected to another fragment via a bond central to a single dihedral angle. A library of fragments having a specified probability distribution of hard degrees of freedom is created and stored for use during a simulation. The fragments are reassembled via a CBMC algorithm using the soft degrees of freedom to establish the overall molecule conformation. This procedure ensures that the sampling of hard degrees of freedom such as those arising from angle bending is totally decoupled from the soft degrees of freedom. Errington and Panagiotopoulos[15] developed a similar method called "reservoir Monte Carlo," in which a library of whole molecules in the ideal gas phase is generated prior to a simulation and then inserted when needed. Essentially, each molecule is treated as a single fragment. In all-atom models that treat some bond angles as flexible and others as rigid, the procedure shown in the Monte Carlo study by Chen *et al*. may be adopted.[16] This model treated the C–C–C bond angle as flexible while the C–C–H and H–C–H bond angles were constrained. The sampling of intramolecular degrees of freedom in this model was carried out by placing the backbone carbon atoms using a CBMC methodology. The positions of hydrogen atom were subsequently assigned using a simple geometric procedure.[16]

The treatment of rings is also difficult within a CBMC framework. Conventional CBMC algorithms are not suitable for rings because closure cannot always be guaranteed. Early attempts focused on the implementation of a crank-shaft move similar to that proposed by Baumgärtner and Binder[17] for sampling freely jointed polymer chains with excluded volume interactions. The move consists of rotating one of the randomly selected interior segments of a chain and cranking the segment around a "shaft" formed by the segments directly connected to the interior segment, as shown in Fig. 2. Such a move was implemented by Pertsin *et al*.[18] to sample conformers of even-numbered ring cycloalkanes, C_{12}, C_{14}, C_{16}, C_{22}, and C_{24} and was also used by Errington and Panagiotopoulos in their reservoir method.[15] Neubauer *et al*.[19] employed this methodology to achieve thermal equilibration and increase acceptance rates for exchange of cyclic alkanes in a Gibbs ensemble Monte Carlo simulation. Bourasseau *et al*.[20] combined the crank-shaft move with the BPS technique[14] to compute equilibrium properties of cyclic alkanes. United-atom models were used in the studies by Neubauer *et al*.[19] and Bourasseau *et al*.;[20] it is not clear how an all-atom representation of a cyclic molecule would be sampled using these approaches.

Wick and Siepmann[21] have proposed an alternative technique known as self-adapting fixed endpoint (SAFE) for efficient sampling of intramolecular degrees of freedom of long polymer chains and cyclic molecules. With this method it is possible to rearrange any number of interior segments of a polymer network thereby overcoming the limitation of the concerted rotation algorithm to rearrange 5–8 interior segments in a molecule. In addition, the SAFE algorithm allows rebuilding of interior regions with more than two endpoints thereby efficiently equilibrating, for example, conformations of cross-linked polymers in which branch points are connected by linear segments. The implementation of SAFE approach requires *a priori* knowledge of guiding probabilities, an initial guess for which can be obtained from a short simulation of a non-cyclic molecule. During an actual simulation, these probabilities are further refined and used to determine placement of atoms such that the last atom can be positioned to satisfy the connectivity constraint. SAFE, however, cannot handle fixed endpoints bonded to more than three atoms. Chen and Escobedo[22] invented a "rebridging configurational-bias" approach to efficiently simulate inner sections of linear and cyclic molecules. The methodology was shown to successfully sample different conformations of small cyclic alkanes such as cyclooctane as well as a large cyclic molecule containing 100 carbon atoms. Similar to the SAFE methodology, the rebridging scheme also utilizes guiding probabilities to connect the growing segment with the fixed end-point.

Sklenar *et al*.[23] devised a ring breakage and closure algorithm and provided Jacobians associated with the forward and reverse Monte Carlo moves. The authors formulated

the ring breakage/closure equations in the torsion and bond angle space and showed that the equations possess only two solutions.

Molecular dynamics simulations could also be used to generate a library of fragments with a known probability distribution, as was done by Nath and Khare.[24] This approach has two major drawbacks. First, it requires the use of finite time step integrators and the computation of forces, neither of which are part of a normal Monte Carlo algorithm. Second, it is difficult to implement such a scheme with models having fixed bond lengths; satisfying bond length constraints with sufficient accuracy can require excessive computational time. Thus a method that relies solely on Monte Carlo sampling to generate library fragments is highly desirable.

The objective of the present study is to develop algorithms and methods to enable a reservoir-type simulation to be performed within a CBMC framework for molecules of arbitrary complexity. Both branch point fragments of the type that occur in chain molecules and cyclic fragments that occur in ring structures having both endocyclic and exocyclic atoms will be treated. The method requires no knowledge of guiding probabilities and can be implemented without recourse to molecular dynamics. The approach is fully compatible with CBMC and other standard Monte Carlo sampling techniques. The article is organized as follows. First, a simple and efficient Monte Carlo algorithm to generate bond angles for branch point fragments is described. Second, it is shown how this sampling technique may be combined with a crank-shaft move to sample intramolecular degrees of freedom of cyclic molecules. Following this, acceptance rules are developed for applying these methods to the Gibbs ensemble, and then a series of validation simulations are conducted to demonstrate the validity and application of the sampling methods.

II. FRAGMENT LIBRARY GENERATION AND USE IN CONFIGURATIONAL-BIAS MONTE CARLO

In the following, the term "fragment" is used to denote both branch points and ring moieties with and without exocyclic atoms, such that a molecule may be represented as a collection of fragments connected by a bond shared between two adjacent fragments. Figure 1 is an example of a fragment containing a central atom with four bonded atoms and six coupled bond angles. An all-atom model of *n*-butane is composed of two terminal fragments (a carbon atom connected to three hydrogen atoms and one other carbon atom) and two internal fragments (a carbon atom connected to two hydrogen atoms and two carbon atoms) as shown in Fig. 3(a). Each fragment shares two common atoms with the fragment that it is directly bonded to. United-atom models of linear chains can also be modeled using the concept of fragments. In such instances, a fragment is made up of three contiguous united-atom sites with a bond angle at the central site. For example, a united-atom model of *n*-butane contains two identical fragments (Fig. 3(b)) (a united-atom CH_2 connected to one united-atom CH_3 and one united-atom CH_2). The representation of a molecule as a collection of fragments allows separation of hard degrees of freedom such as bond angles that can

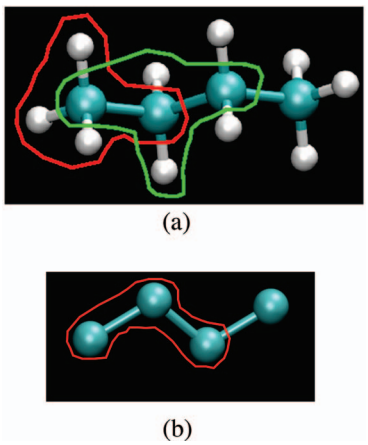

FIG. 3. (a) Fragment representation of an all-atom model of *n*-butane. The molecule can be represented by two fragments: one terminal fragment containing a central carbon atom bonded to three hydrogen atoms and one methylene carbon (bounded by the red curve) and an internal fragment that contains one central carbon bonded to two carbon atoms and two hydrogen atoms (enclosed by the green curve). (b) Fragment representation of united-atom model of *n*-butane. The molecule can be represented by two identical fragments (bounded by the red curve).

be sampled prior to the selection of soft degrees of freedom such as dihedral angles.

Consider a CBMC move from state m to state n. This could involve the growth of a new molecule in a new phase or the regrowth of a molecule in an existing phase. The detailed balance condition states that

$$acc(m \to n)\alpha_{mn}\pi_m = acc(n \to m)\alpha_{nm}\pi_n, \tag{1}$$

where $acc(m \to n)$ is the one step transition probability of moving from state m to n, α_{mn} is the probability of attempting such a move and π_m is the probability of being in state m. Biased Monte Carlo methods rely upon the development of creative ways to construct α_{mn} to achieve better sampling. In the current method, the term α_{mn} (and by symmetry, α_{nm}) is a compound probability comprised of terms arising from hard and soft degrees of freedom. That is, the probability of generating a new trial conformation of the molecule for state n is the product of hard terms that govern the intramolecular conformation of each fragment of the molecule and soft terms that govern the relative arrangement of the individual fragments,

$$\alpha_{mn} = P_{mn}^{hard} P_{mn}^{soft}, \tag{2}$$

P_{mn}^{soft} is determined using any of the standard CBMC moves that have been developed over the years. P_{mn}^{hard}, on the other hand, is the product of the probabilities associated with each of the fragments selected from the library. We assume in the following that a fixed bond length model is used, so the hard degrees of freedom of a fragment are the collection of bond angles $\{\theta\}$. A particular fragment i appears in the library with probability,

$$P_i(\{\theta\}) = \frac{\exp\left(-\beta E_i\left\{\theta\right\}\right) J(\{\theta\}) d\{\theta\}}{Z_i}, \tag{3}$$

where Z_i is the canonical partition function of the isolated fragment i, $J(\{\theta\})$ is the Jacobian of transformation from the

Cartesian coordinates to the set of internal coordinates represented by $\{\theta\}$, and $d\{\theta\}$ is the differential volume associated with the set of generalized coordinates. The energy of the particular conformation is $E_i\{\theta\}$ while β is the inverse temperature. A large number of fragments N_f are generated and placed into the library such that the fragment probability can be approximated as

$$P_i^{frag}(\{\theta\}) \approx \frac{\exp(-\beta E\{\theta\}) J(\{\theta\})\delta\{\theta\}}{\sum_{j=1}^{N_f} \exp(-\beta E_j\{\theta_j\}) J(\{\theta_j\})\delta\{\theta_j\}}, \quad (4)$$

where $\delta\{\theta\}$ is the discrete volume element associated with the set of internal coordiantes. If the molecule requires the placement of M different fragments, then the total probability associated with hard degrees of freedom of the molecule is

$$P_{mn}^{hard} = \prod_{i=1}^{M} P_i^{frag}(\{\theta\}). \quad (5)$$

A. Bond angle sampling around a fragment with atom displacement algorithm

Consider the generation of a library of fragments consisting of "simple" fragments without rings. In principle, Eq. (4) could be used to determine the probability of each fragment and then used to compute P_{mn}^{hard} in Eq. (2). However, this requires *a priori* knowledge of the Jacobian of transformation, which can be very difficult to evaluate, even for the simple case of a fragment containing three bond angles.[25] This difficulty can be overcome, however, by a transformation of generalized coordinates to a different coordinate system for which the Jacobian is easier to evaluate or need not be computed at all. For fixed bond length models, spherical coordinates (l, ψ_1, ψ_2) are very convenient. Figure 4 defines the spherical coordinate system used here, where l is the distance of a site from the origin (chosen to coincide with the central site of the fragment and hence l corresponds to the bond length), ψ_1 is the polar angle of the vector connecting the central site to the branched site, and ψ_2 is the azimuthal angle. The angles are measured with respect to a lab frame of reference fixed at the central site. The partition function, in the spherical coordinate representation of each of the fragments, can then be expressed as

$$Z(\{\mathbf{l}\}, \{\psi_\mathbf{1}\}, \{\psi_\mathbf{2}\}) = \int \exp(-\beta E) \prod_{i=1}^{N} l_i^2 dcos(\psi_{1,i}) d\psi_{2,i}, \quad (6)$$

where N refers to the number of sites connected to the central site of the fragment. In practice the integration is replaced by a summation over a large number of fragments, as was done in Eq. (4). Note that the Jacobian associated with the transformation from the Cartesian coordinates to polar coordinates is contained in the differential volume $d\cos(\psi_1)$ and is exactly known irrespective of the number of atoms involved in the fragment. Thus, for fixed bond length models, the probability of observing a set of polar and azimuthal angles for fragment i is simply

$$P_i^{frag}(\{\psi_1\}, \{\psi_2\}) = \frac{\exp(-\beta E_i) \prod_{j=1}^{N} dcos(\psi_{1,j}) d\psi_{2,j}}{Z_i}. \quad (7)$$

This probability can then be used in Eq. (5) in place of $P_i^{frag}(\{\theta\})$. A Monte Carlo atom displacement sampling scheme is used to generate fragment conformations consistent with the probability given in Eq. (7). To do this, a trial move is performed to create a new fragment conformation n starting from old fragment conformation m. This is done by randomly selecting a polar angle on a cosine distribution from $(-\Delta\cos(\psi_1), \Delta\cos(\psi_1))$ and azimuthal angle from $(-\Delta\psi_2, \Delta\psi_2)$, i.e.,

$$\cos(\psi_1)|_n = \cos(\psi_1)|_m + (2 * \tau_1 - 1) * \Delta\cos(\psi_1), \quad (8)$$

$$\psi_2|_n = \psi_2|_m + (2 * \tau_2 - 1) * \Delta\psi_2, \quad (9)$$

where τ_1 and τ_2 are random numbers uniformly distributed in the interval [0, 1]. Moves that result in $\cos(\psi_1)|_n > 1$ or $\cos(\psi_1)|_n < -1$ are automatically rejected. The attempt probability for the forward move is

$$\alpha_{mn} = \frac{d\cos(\psi_1)}{2\Delta\cos(\psi_1)} \frac{d\psi_2}{2\Delta\psi_2}. \quad (10)$$

The attempt probability of the reverse move is identical to that of the forward move and cancels out in the final acceptance rule (see Eq. (1)), thereby yielding the following acceptance rule for generating fragment conformation n starting from state m,

$$acc(m \to n) = \min(1, \exp(-\beta\Delta E)). \quad (11)$$

Note that the acceptance rule does not require evaluation of the Jacobian for the forward or reverse move, thus enabling sampling of fragments with an arbitrary number of sites. In this algorithm, displacement of one site leads to a change in $(N-1)$ bond angles. Although the present formalism attempts to perturb a single site, there is no restriction on the number of sites that can be displaced. As with standard center-of-mass (COM) translation algorithms, the maximum displacement widths can also be tuned to achieve a target acceptance rate. An alterative algorithm is provided in Ref. 26.

To generate a library of simple fragments prior to a simulation, one starts with a particular fragment and carries out a Monte Carlo "pre-simulation" using the above algorithm and the acceptance rule given by Eq. (11). Conformations are saved at regular intervals, such that the probability of a given fragment is given by Eq. (7). This probability can then be used to compute P_{mn}^{hard} via Eq. (5) and used in the overall attempt probability given by Eq. (2).

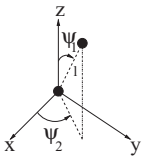

FIG. 4. Spherical coordinate representation of a site in a fragment.

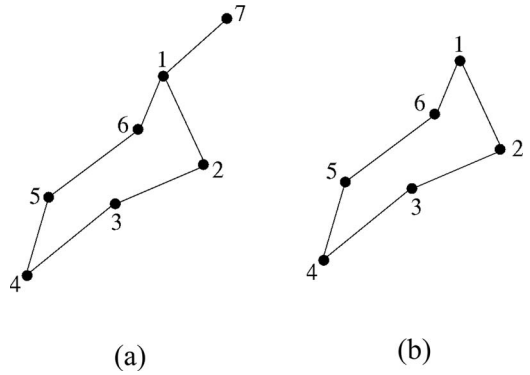

FIG. 5. Schematic representation of (a) united-atom methylcyclohexane and (b) united-atom cyclohexane.

B. Crank-shaft move coupled with atom displacement algorithm

Although the atom displacement method described above works for fragments that make up chain molecules, the generation of a library of conformations containing cyclic groups poses a special problem for Monte Carlo sampling, particularly when the model contains fixed bond lengths. In the case of cyclic molecules, sites in a ring and any exocyclic sites directly bonded to the ring sites are collectively treated as one fragment. For example, in united-atom methylcyclohexane (Fig. 5(a)), all the sites are part of the ring fragment. In such systems, endocylic bond angles are sampled by a crank-shaft move[17,18] that perturbs a randomly selected ring site, as shown in Fig. 2. The selected site is given an angular rotation around the axis formed by the two sites directly attached to the site being perturbed. In the crank-shaft move, if a ring site is connected to exocyclic sites, for example, hydrogens attached to ring carbon atoms in an all-atom model of benzene, the ring site and exocyclic sites are treated as a rigid body during the rotation. The angular displacements of the ring sites sample both the endocylic bond angles and dihedral angles. It has been shown that the Jacobian associated with the crank-shaft moves is identical for the forward and reverse attempt[27] yielding the acceptance rule for the move that is identical to Eq. (11). In the case where multiple exocyclic atoms are present, sampling of the exocyclic bond angles is achieved by applying the atom displacement algorithm described earlier to the sites external to the ring. As before, Eq. (11) provides an appropriate acceptance criterion for such a move. It is to be noted that this strategy places no restriction, as is the case in the SAFE implementation,[21] on the number of sites that are connected to a common ring site.

This algorithm does not require solution of a ring closure problem.[23] It also does not depend on *a priori* knowledge of guiding probabilities to ensure ring closure.[21,22] Due to the local nature of the crank-shaft and atom displacement moves, they are likely to be more successful than ring closure algorithms for conformational changes in dense systems where the probability of ring closure drops dramatically due to the stiffness of bond angles. Moreover, a completely new ring fragment conformation can be generated as the methodology devised in the present work allows for the flexibility of rings to be sampled in the gas phase.

The library generation scheme for fragments containing rings follows the same procedure as that for "simple" fragments. Starting from an existing fragment, new fragments are generated during a Monte Carlo presimulation using crank-shaft and atom displacement moves. New conformations are accepted according to Eq. (11) and coordinates are periodically stored in the library. These library fragments occur with a probability P_i^{frag} and are used to construct new molecular conformations for which the overall probability of "hard" degrees of freedom is given by the product of fragment probabilities as in Eq. (5).

III. FRAGMENT LIBRARY USE IN GIBBS ENSEMBLE MONTE CARLO

Fragment libraries created using the algorithms described above can be used for biased insertion and deletion in open ensembles such as the Gibbs, grand canonical, and osmotic ensembles. They can also be used to equilibrate conformations of molecules via rearrangement moves. The derivation of acceptance rules for molecular transfer in the Gibbs ensemble requires that, for a fragment denoted by i, a ratio of attempt probabilities $\alpha_{nm}(i)/\alpha_{mn}(i)$ be calculated (see Eq. (1)). Here m and n represent the old and new states of the system, respectively. The ratio is then used in the acceptance rule to correct for the bias introduced during the sampling. For particle exchange or a complete regrowth, a molecule with M fragments will require computation of the following total attempt probability ratio,

$$\frac{\alpha_{nm}}{\alpha_{mn}} = \prod_{i=1}^{i=M} \frac{\alpha_{nm}(i)}{\alpha_{mn}(i)} = \frac{P_{nm}^{hard} P_{nm}^{soft}}{P_{mn}^{hard} P_{mn}^{soft}}. \tag{12}$$

To derive the ratio of attempt probabilities, we consider a transfer move in the canonical Gibbs ensemble. Derivation of acceptance rules for other ensembles and other methods follows a similar procedure.[14] Consider a system consisting of N molecules distributed between two boxes such that one box contains N_1 molecules and the other box, $N_2 = N - N_1$ molecules. A new state of the system is generated by randomly selecting a box to be the donor box and transferring a molecule from this box to the receiving box. Details of growth of the molecule in a biased fashion and calculation of attempt probabilities for addition of each of the fragments are outlined below. Once again, it is assumed that a fixed bond length model is used and the hard degrees of freedom for fragments are the collection of intramolecular bond angles $\{\theta\}$.

1. One of M possible fragments is chosen at random to initiate insertion of the molecule into the receiving box. The probability ($P_{mn}(1)$) associated with the selection is

$$P_{mn}(1) = \frac{1}{M}. \tag{13}$$

2. Next the conformation of the initial fragment is chosen. This may be kept identical to that of an equivalent

fragment in the donor box. However, in the current implementation, a new set of intramolecular degrees of freedom is selected by picking one of the fragments of the chosen type from the reservoir. The probability of selecting one of the fragments from the library (and hence that of selecting a set of intramolecular degrees of freedom) is given by

$$P_{mn,1}(\{\theta\}) = \frac{\exp(-\beta E(\{\theta\}))\, d\Omega_1}{Z_1}, \tag{14}$$

where $d\Omega_1$ and Z_1 denote the differential configurational volume and canonical partition function of the fragment. This sets the "hard" degrees of freedom of this fragment.

3. The fragment is given a random orientation and κ_{ins} trial positions for the fragment are generated in the receiving box, out of which one is selected with a probability,

$$P_{mn,1}\left(\{\mathbf{r_p}\}\right) = \frac{\exp(-\beta E(\{\mathbf{r_p}\}))}{\sum_{i=1}^{\kappa_{ins}} \exp(-\beta E(\{\mathbf{r_{p_i}}\}))}, \tag{15}$$

where $\{\mathbf{r_{p_i}}\}$ represents the set of position vectors of the atoms of the fragment for the ith trial. This sets the "soft" degrees of freedom of the fragment.

The forward attempt probability for the placement of the first fragment of the molecule is then given by

$$\alpha_{mn}(1) = P_{mn}(1) * P_{mn,1}(\{\theta\}) * P_{mn,1}(\{\mathbf{r_p}\}). \tag{16}$$

Note that additional "soft" biasing strategies such as rotational bias may be applied at this stage to improve insertion efficiency. Such biasing schemes may be especially useful for large ring fragments.

After placing the first fragment of the molecule, all the fragments connected to this fragment are placed in a random order. The sequence in which multiple fragments are added to a given fragment must be identical in the reverse move to satisfy microscopic reversibility. For simplicity, consider the case in which the molecule is composed of a linear chain of fragments. For the addition of the kth fragment, a two step procedure is adopted. First, a conformation of the fragment is selected from the reservoir with a probability

$$P_{mn,k}(\{\theta\}) = \frac{\exp(-\beta E(\{\theta\}))\, d\Omega_k}{Z_k} = P_{mn,k}^{hard}, \tag{17}$$

where $\{\theta\}$ is the set of bond angles for the fragment. The next step is to allow rotation of the fragment around the central atoms connecting the two fragments for selection of dihedral angles. During the rotation, the atoms of the kth fragment are rotated as a rigid body so that the predetermined selection of bond angles is preserved. For each rotation τ, the contribution due to the complete set of dihedrals $\{\phi\}$ involving all the atoms of the kth fragment is computed and one set of dihedrals out of κ_ϕ is selected with the probability

$$P_{mn,k}(\{\phi\}) = \frac{\exp(-\beta E(\{\phi\}))\Delta\tau}{\sum_{i=1}^{\kappa_\phi} \exp(-\beta E(\{\phi\}_i))\Delta\tau} = P_{mn,k}^{soft}. \tag{18}$$

The energy that is used in biasing the selection involves energy due to all the dihedral angles of the fragment and any intra and intermolecular nonbonded interactions experienced by the sites of the added fragment. Thus the total forward probability of adding the kth fragment is

$$\alpha_{mn}(k) = P_{mn,k}(\{\theta\}) * P_{mn,k}(\{\phi\}) = \frac{\exp(-\beta E(\{\theta\}))\, d\Omega_k}{Z_k} * \frac{\exp(-\beta E(\{\phi\}))\Delta\tau}{\sum_{i=1}^{\kappa_\phi} \exp(-\beta E(\{\phi\}))\Delta\tau}. \tag{19}$$

Growth of the entire molecule is thus accomplished by adding the remaining fragments in a similar biased fashion. The total probability of generating a conformation of the molecule in the receiving phase is the product of the attempt probability of each of the fragments,

$$\alpha_{mn} = \prod_{i=1}^{i=M} \alpha_{mn}(i). \tag{20}$$

The reverse move consists of deleting the molecule from the receiving phase and performing a biased insertion of the molecule in the donor phase. Calculation of the attempt probability for this move is carried out by mimicking the growth of the molecule in the donor phase with a configurational-bias scheme exactly as above, except that one of the trial positions is selected to be the original position. The expression for the reverse probability is identical to that obtained in Eq. (20) with corresponding energies being calculated in the donor box. Note that attempt probabilities associated with deletion of sites in each phase are identical for the forward and reverse moves and cancel out when the ratio in Eq. (12) is computed as

$$\frac{\alpha_{nm}}{\alpha_{mn}} = \frac{\prod_{i=1}^{i=M} \alpha_{nm}(i)}{\prod_{i=1}^{i=M} \alpha_{mn}(i)}. \tag{21}$$

An example of the construction of an all-atom model of linear butane is shown in Fig. 6.

IV. FORCE FIELD

Tests of the algorithm were performed on model systems. The TraPPE united atom force field was used for all the species.[13,28] The force field contains harmonic angle bending terms

$$E_\theta = \frac{K_\theta}{2}(\theta - \theta_0)^2, \tag{22}$$

where K_θ is the harmonic force constant and θ_0 is the nominal bond angle. The torsion potential governing the motion of a dihedral angle is represented in the OPLS functional form (Eq. (23)),

$$E_\phi = c_0 + c_1(1 + \cos(\phi)) + c_2(1 - \cos(2\phi)) + c_3(1 + \cos(3\phi)). \tag{23}$$

Nonbonded interactions between atoms i and j belonging to different molecules or separated by more than three bonds in the same molecules is described by Lennard-Jones 12-6

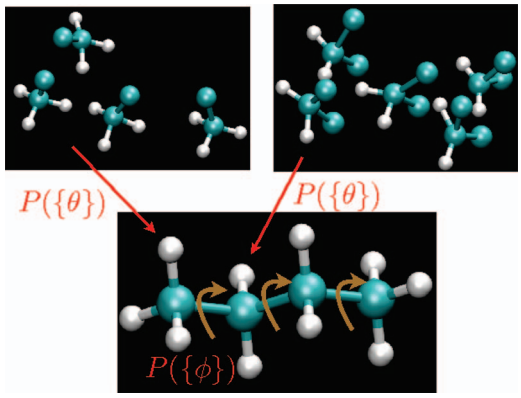

FIG. 6. Procedure for generation of a conformation of an all-atom model of linear butane. The upper panel indicates the gas phase reservoirs for the two fragments shown in Fig. 3(a). A conformation for a given fragment is chosen randomly. The probability associated with the selection of a fragment is given by $P(\{\theta\})$. Except for the first fragment, subsequent fragments are given a series of rigid body rotations around the atoms connecting two fragments. The probability associated with the selection of one of the rotations is indicated by $P(\{\phi\})$. It is assumed here that the molecule is constructed by placing the leftmost fragment first and subsequently adding the remaining three fragments, although in practice the choice of initial fragment and the order of each connecting fragment is random.

potential,

$$E_{LJ} = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], \tag{24}$$

where cross parameters ϵ_{ij} and σ_{ij} are obtained by standard Lorentz-Berthelot combining rules. Force field parameters for various terms were obtained from published literature.[13,28]

V. SIMULATION DETAILS

A. Single molecule gas phase simulations

Initial validation studies of the fragment generation algorithms were performed to ensure that P_{mn}^{hard} was being determined properly. For the atom displacement algorithm, simulations were conducted for united-atom propane, isobutane, and neopentane. The systems were chosen so as to rigorously test the algorithm for fragments with two, three, and four branches. Single molecule simulations were carried out at 300 K in the gas phase to populate fragment libraries. Additional gas phase simulations were conducted for isobutane at 1000 K; the use of an elevated temperature ensures that incorrect bond angle distributions are more apparent, as Vlugt *et al.*[12] showed for a sequential growth algorithm. Each simulation was performed for 1 M MC steps and conformations were saved every 10 steps for later analysis of the bond angle distributions. The maximum width of azimuthal angle (ψ_2 in Fig. 4) sampling was set at 10 while that for the cosine of the polar angle (ψ_1 in Fig. 4) was 0.2. Brute force simulations employing a simultaneous angle generation algorithm with Boltzmann rejection were also attempted to gauge the efficiency of the proposed algorithm and to test the accuracy of the algorithm. These simulations consisted of simultaneous but random generation of positions of fragment atoms on a sphere. A given configuration was accepted or rejected based on the Boltzmann weight of the bond angle energy. While this method guarantees the correct set of angles will be generated, the efficiency becomes very low as additional coupled angles are sampled, as shown below.

Cyclohexane (Fig. 5(b)) and 1-methylcyclohexane (Fig. 5(a)) were simulated to test the ring sampling part of the algorithm. The crank-shaft move was applied to sample bond angle distributions in cyclohexane. In the case of 1-methylcyclohexane, the atom displacement algorithm was applied to the exocyclic methyl moeity (denoted by 7 in Fig. 5) in addition to the crank-shaft move. As mentioned earlier, during the crank-shaft move about the axis formed by groups 2 and 6, both groups 1 and 7 were rotated simultaneously. Single molecule gas phase simulations of the ring molecules were carried out at 450 K. Each simulation was carried out for 5 M MC steps and conformations were saved every 100 steps. The maximum rotational width for the crank-shaft move was set to 15°, while widths for the atom displacement move were fixed to the same values as before.

B. Gibbs ensemble Monte Carlo simulations

In addition to single molecule gas phase simulations to test the fragment generation algorithms, Gibbs ensemble Monte Carlo simulations in the canonical ensemble were performed to predict vapor-liquid equilibria of propane, isobutane, and 2,2-dimethylhexane at four different temperatures spanning at least ~100 K range for each of the molecules investigated. The wide range of temperatures allowed us to validate our implementation of the fragment sampling scheme in the Gibbs ensemble against published results obtained using the coupled-decoupled Monte Carlo scheme of Martin and Siepmann.[13]

For each state point, at least three independent simulations were conducted starting with a different random seed. The total number of molecules used in the simulation ranged between 350 and 400. Initially, the number of molecules in the liquid box was set to be less than that expected for the equilibrium liquid density. Correspondingly, the vapor phase contained more molecules than expected for the equilibrium vapor density. At least 20 M MC steps were performed at each state point. The types of Monte Carlo moves consisted of COM translation, rotation about the COM, configurational biased molecule exchange between the two boxes, volume displacements, and configurational-bias regrowths in the case of 2,2-dimethylhexane. For the placement of the first fragment and selection of dihedral angles for the subsequent fragments, the parameters κ_{ins} and κ_{ϕ} were both set to 12. The energy between the growing fragment sites and those within 5.5 Å of these sites was used for biasing along with any energy that resulted from the dihedral angles due to the placement of the fragment. Additionally, for all the systems but propane, the atom displacement move was carried out for terminal atoms, i.e., atoms bonded to only one other atom. The move was included to allow for any differences in the bond angle distributions between those of the fragment library and

those observed in the fluid phase due to the presence of other molecules.

Propane and isobutane were modeled as a single fragment while 2,2-dimethylhexane was built from four fragments. For each of the fragments, the gas phase reservoir was obtained at each temperature of interest and contained 100 000 conformations generated using the atom displacement algorithm. The fragment library was used during the course of a simulation for the selection of "hard" degrees of freedom of a given fragment.

Nonbonded interactions were truncated at 14 Å and standard long range corrections were applied.[29] A hard inner cutoff was also used such that any move bringing two sites closer than 1 Å was immediately rejected.

VI. RESULTS AND DISCUSSION

A. Gas phase calculations: Alkanes

The correctness of the atom displacement algorithm was verified by simulating united-atom models of propane, isobutane, and neopentane in the gas phase and checking the resulting bond angle distributions. For propane modeled with a harmonic bond angle, the bond angle distribution is analytic and direct validation of the simulation results is possible. For isobutane and neopentane, the bending potentials employed for all the bond angles are identical and therefore the resulting bond angle distributions must be identical if the algorithm is correct. To test this, the probability of observing a given bond angle was computed and compared to a reference bond angle in the molecule. A ratio of one indicates the bond angle distributions are identical for a given value of θ. Since the sequential selection of bond angles in a CBMC move results in equivalent bond angles having different distributions,[12] this is a sensitive test of the atom displacement algorithm. For statistically meaningful results, the ratio was computed only for the cases in which the probability of observing both the bond angles exceeded 0.001.

The ratio of the bond angle distribution obtained from Monte Carlo simulations for propane with the atom displacement algorithm to that calculated for the analytical distribution is displayed in Fig. 7. The ratio is close to unity over the relevant range of bond angles, indicating that the algorithm correctly reproduces the analytical bond angle distribution in propane. The deviations approach 5–10% at angles less than 105° and greater than 120° due to the high energy of these states and the statistical uncertainty associated with infrequent sampling of these angles.

For isobutane, the ratios of different bond angle distributions to that of a reference bond angle (arbitrarily selected to be angle 2-1-3) are shown in Fig. 8(a). The ratios are close to unity over the range of θ values for which the probability exceeds 0.001, confirming that the algorithm results in identical bond angle distributions in isobutane and that there is no bias in sampling of any of the bond angles. Figure 8(b) shows that at 1000 K where even small differences become evident,[12] the bond angle distributions are close to unity over the relevant range of θ values.

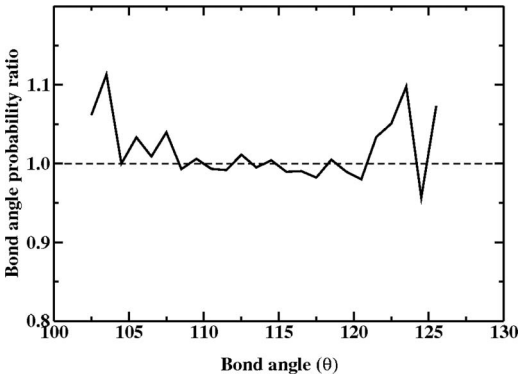

FIG. 7. Ratio of bond angle probability distribution in propane obtained from Monte Carlo simulation to the analytical bond angle distribution at 300 K.

An additional way of testing the correctness of the atom displacement method is to compare the resulting bond angle distributions against those obtained using a Boltzmann rejection scheme in which all three angles of isobutane are chosen simultaneously, a procedure known to generate unbiased angle distributions albeit at much greater computational cost. Figure 9 compares the bond angle distributions at 300 K from the atom displacement algorithm and the Boltzmann rejection technique for the three angles of isobutane. It is clear that both methods yield identical distributions. However, the Boltzmann rejection method requires almost two orders of

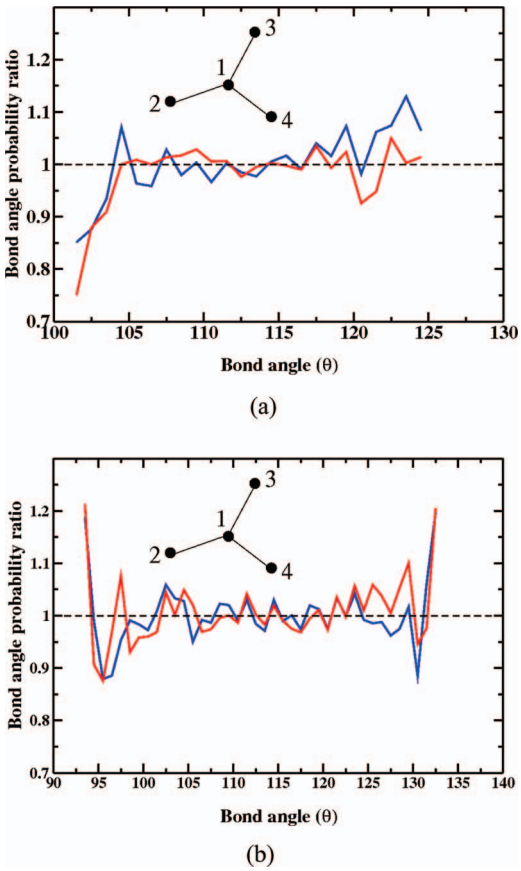

FIG. 8. Ratios of the bond angle probability distributions obtained with the atom displacement method. The ratio is defined with respect to the probability distribution for the bond angle 2-1-3 as shown in the schematic. Bond angle 2-1-4 (blue line) and bond angle 3-1-4 (red line). (a) 300 K and (b) 1000 K.

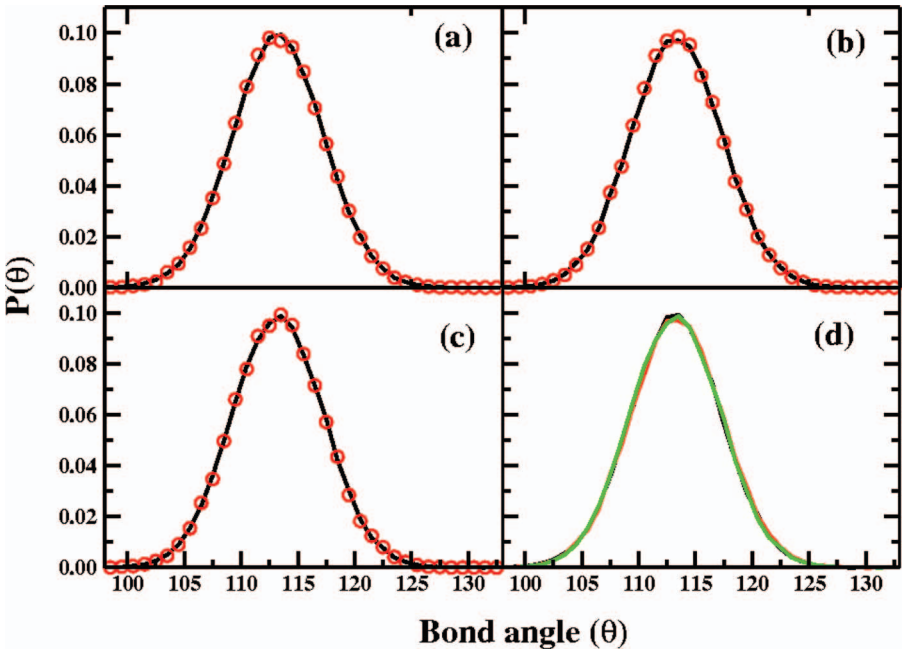

FIG. 9. Comparison of the bond angle probability distributions obtained with the atom displacement algorithm (solid line) with those obtained from the Boltzmann rejection scheme (circles) for isobutane at 300 K. (a) Bond angle 2-1-3, (b) bond angle 2-1-4, and (c) bond angle 3-1-4. (d) Bond angle probability distributions obtained with the atom atom displacement algorithm showing that they are nearly identical. Bond angle 2-1-3 (black line), bond angle 2-1-4 (red line), and bond angle 3-1-4 (green line).

magnitude more MC steps to generate the bond angle distributions shown in Fig. 9. The higher efficiency of the atom displacement scheme is related to the restriction of sampling to the important region of bond angle phase space, while no such restriction is placed on the Boltzmann rejection method. The acceptance rate for atom displacement moves was found to be at least 25%. A high acceptance rate coupled with very little computational overhead associated with the move (≈30 s/ 1 M MC steps) means that the fragment reservoir can be computed or refreshed during a simulation "on the fly" if one desires. This feature is particularly attractive for reduction in the memory storage requirement in simulations involving large fragments or those requiring large numbers of fragments, for example, in biomolecular systems.

Figure 10 shows a comparison plot of five of the bond angle distributions in neopentane referenced to the sixth angle. The distributions are essentially identical over the relevant range, indicating that the method works for fragments having three branches. Note that it was not possible to compare against the Boltzmann rejection method for neopentane, because this procedure failed to produce a smooth bond angle distribution even after 2×10^9 MC moves. Based on the observations here, it can be concluded that the atom displacement algorithm can be extended to handle higher order fragments such as those encountered in molecules such as SF_6 or ions such as $[PF_6]^-$.

B. Gas phase calculations: Cyclic molecules

Bond angle distributions were computed for a united-atom model of cyclohexane using the crank-shaft algorithm. A comparison plot of the various angle distributions normalized against the angle formed between sites 1, 2, and 3 (see Fig. 5) is shown in Fig. 11(a). Each ratio is close to unity, indicating that the crank-shaft move does not introduce any bias in the sampling of the bond angles and that the move was properly implemented in our code.

In the case of a united-atom model of 1-methylcyclohexane, the atom displacement algorithm was applied to the exocyclic atom (denoted by site 7 in Fig. 5) while the crank-shaft move was performed to sample endocylic angles. Due to the presence of an exocylic site, all the endocylic bond angles are not equivalent. However, there exists a symmetry in the molecule such that the bond angle pairs 712-716, 123-165, and 234-654 are equivalent (see Fig. 5). The ratios of the bond angle distributions of these pairs are

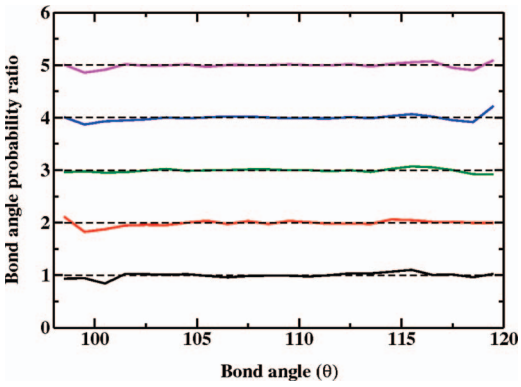

FIG. 10. Ratios of the bond angle probability distributions in united-atom neopentane. The bond angle 2-1-3 is defined to be the reference bond angle for calculating the ratios. Note that the ratios of distributions are shifted for clarity. Bond angle 2-1-4 (black line), bond angle 2-1-5 (red line), bond angle 3-1-4 (green line), bond angle 3-1-5 (blue line), and bond angle 4-1-5 (pink line). Simulation was carried out at 300 K. The numbering scheme is the same as that provided in Fig. 1.

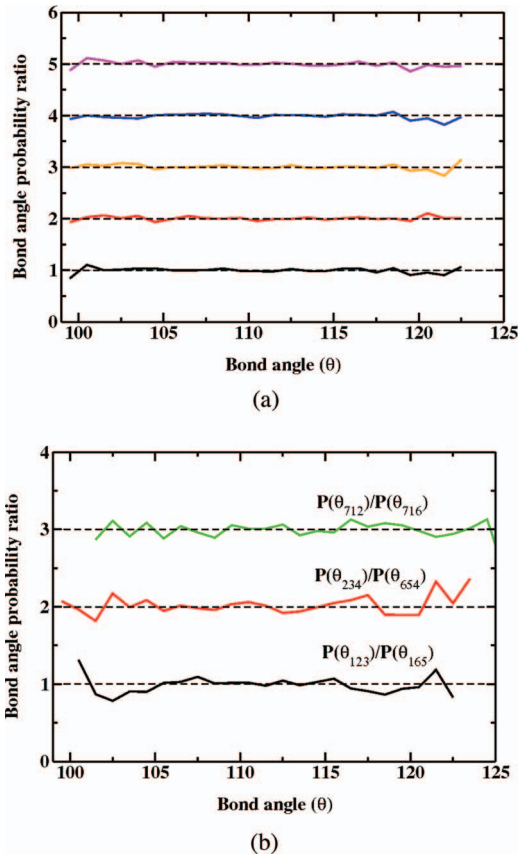

FIG. 11. Ratios of the bond angle probability distributions to the reference bond angle probability distribution in cyclohexane and methycyclohexane. The bond angle distributions were determined from Monte Carlo simulations using the crank-shaft move to sample endocyclic bond angles for cyclohexane and methylcycolhexane. For methylcyclohexane, the atom displacement algorithm was applied to sample exocyclic bond angles. (a) Cyclohexane: the reference bond angle is 1-2-3. Bond angle 2-3-4 (black line), bond angle 3-4-5 (red line), bond angle 4-5-6 (orange line), bond angle 5-6-1 (blue line), and bond angle 6-1-2 (pink line). (b) Methylcyclohexane: reference bond angles are chosen such that they form a symmetric pair for comparison. Simulations were conducted at 450 K. Refer to Fig. 5 for the numbering scheme.

shown in Fig. 11(b) and indicate no bias in the algorithm. Based on these results, it can be concluded that the atom displacement algorithm can be implemented in conjunction with the crank-shaft move to sample intramolecular degrees of freedom of cyclic molecules containing endocyclic and exocyclic sites, such as will occur in all-atom models of benzene or more complicated ring molecules.

C. Gibbs ensemble Monte Carlo simulations

Given that the fragment sampling method yields the correct structures of molecules, the Gibbs ensemble implementation was tested. To do this, vapor-liquid coexistence curves were computed for propane, isobutane, and 2,2-dimethylether. Results are shown in Fig. 12 and compared against literature results obtained using the coupled-decoupled algorithm.[13,28] Numerical values are given in Ref. 26. The two different methods give vapor-liquid equilibria that are identical within numerical uncertainty. This is an encouraging result, as it helps confirm that both methods give correct results. It also suggests that the "reservoir" method described here, coupled

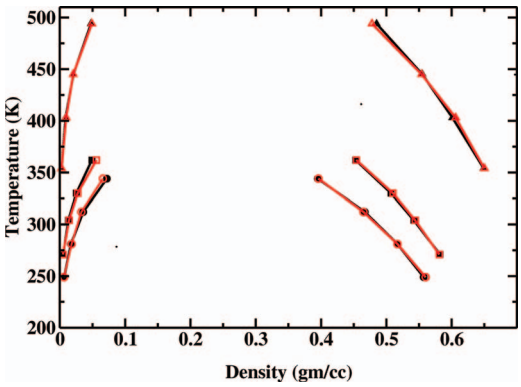

FIG. 12. Computed vapor-liquid equilibria from the Gibbs ensemble simulations utilizing fragment sampling approach proposed in this work (black curve). The results are compared with those published by Martin and Siepmann (Refs. 13 and 28) (red curve). Propane (circles), isobutane (squares), 2,2-dimethylhexane (triangles). Lines connecting the data points are provided as a guide to the eye.

with algorithms that sample the library of fragments using atom displacement and/or crank-shaft moves, is a viable alternative to the coupled-decoupled algorithm and can be used to simulate molecules of varying complexity.

When using a library of fragments, one concern is whether the library contains enough samples to ensure that the approximation in Eq. (4) is valid. To test this, the vapor-liquid equilibrium densities of isobutane at 362 K were computed using libraries containing 1000, 25 000 and 50 000 fragments. The results were indistinguishable from those obtained using 100 000 fragments; we therefore conclude that at least for this system, the approximation in Eq. (4) is valid.

VII. SUMMARY AND CONCLUSIONS

A Monte Carlo algorithm has been described and validated which enables configurational-bias sampling of molecules containing branch points and rings with either endocyclic or exocyclic atoms. The method overcomes problems associated with sequential configurational-bias sampling methods[12] and offers an alternative to the existing coupled-decoupled algorithm[13] that is relatively easy to implement and generally applicable to molecules of widely varying topologies. The method utilizes a "reservoir" or "library" of fragments that are generated either prior to the main simulation or periodically during a simulation. The fragments appear in the reservoir with a known probability distribution governed by "hard" degrees of freedom. The fragments are reassembled to reconstitute the molecules via configurational-bias techniques using the remaining "soft" degrees of freedom in the biasing step. To generate the library of fragments, two different methods are used. An atom displacement method utilizing spherical coordinates is used for simple fragments constituting linear or branched chain regions of a molecule. A crank-shaft move is used for sampling cyclic fragments. The atom displacement algorithm is combined with the crank-shaft move to sample endocyclic and exocyclic bond angle distributions in cyclic molecules. Acceptance rules for utilizing the methods in a Gibbs ensemble Monte Carlo simulation are described.

The methods for sampling fragments were validated by generating conformations of united-atom models of propane, isobutane, and neopentane in the gas phase and comparing bond angle distributions against analytic distributions (for propane) or against distributions obtained using an exact (but expensive) Boltzmann rejection method (for isobutane). It was found that the Boltzmann rejection method was unable to sample the neopentane bond angles even after 2×10^9 moves. Additional validation studies were conducted on isobutane and neopentane by ensuring that equivalent bond angles had the same angle distribution. Tests were conducted at 300 K but also at 1000 K for isobutane. To test the method for generating fragments containing rings, models of cyclohexane and methylcyclohexane were simulated in the gas phase. It was found that all equivalent bond angles had the same angle distribution.

The implementation of the method in the Gibbs ensemble was tested by computing vapor-liquid coexistence curves for united-atom models of propane, isobutane, and 2,2-dimethylhexane. The results of the simulations were found to be in excellent agreement with literature values generated using the coupled-decoupled algorithm.[13] We emphasize that the focus of the present paper is the algorithmic development for sampling of intramolecular degrees of freedom of branch points and all-atom models of ring molecules. However, it is recognized that the applicability of the reservoir technique in the prediction of vapor-liquid equilibria remains to be tested for large molecules such as triacontane,[30] and cyclic dodecane and octadecane.[31] We anticipate that the intramolecular degrees of freedom of these molecules can still be sampled with the algorithms presented in this work. However, additional biasing strategies such as continuous fractional component Monte Carlo[32] may be required for efficient insertions and deletions of large molecules. Though the methodology developed here was used for phase equilibria calculations in the Gibbs ensemble, it is applicable to configurational-bias insertion and deletion moves in any open ensemble. Moreover, it is a general sampling technique that can be applied to efficiently equilibrate and explore conformational space of a large number of systems in the canonical and isothermal-isobaric ensembles.

ACKNOWLEDGMENTS

This material is based upon work supported by the (U.S.) Air Force Office of Scientific Research (USAFOSR) under AFSOR Award No. FA9550-10-1-0244. Support for this work was also provided by the National Science Foundation (NSF) (CBET-0967458). Computing support was provided by Notre Dame's Center for Research Computing. J.K.S. gratefully acknowledges funding received from the Center for Research Computing. The authors would like to acknowledge the suggestion of Professor M. L. Greenfied to compare the bond angle probabilities in the form of ratios of probabilities. The authors also thank Professor J. I. Siepmann for a critical reading of the manuscript and providing invaluable feedback.

[1] A. Z. Panagiotopoulos, Mol. Phys. **61**, 813 (1987).
[2] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
[3] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).
[4] J. I. Siepmann and D. Frenkel, Mol. Phys. **75**, 59 (1992).
[5] G. C. A. M. Mooij, D. Frenkel, and B. Smit, J. Phys. Condens. Matter **4**, L255 (1992).
[6] J. J. de Pablo, M. Laso, and U. W. Suter, J. Chem. Phys. **96**, 2395 (1992).
[7] J. J. de Pablo, M. Laso, and U. W. Suter, J. Chem. Phys. **96**, 6157 (1992).
[8] P. V. K. Pant and D. N. Theodorou, Macromolecules **28**, 7224 (1995).
[9] M. W. Deem and J. S. Bader, Mol. Phys. **87**, 1245 (1996).
[10] M. G. Wu and M. W. Deem, Mol. Phys. **97**, 559 (1999).
[11] M. G. Wu and M. W. Deem, J. Chem. Phys. **111**, 6625 (1999).
[12] T. J. H. Vlugt, R. Krishna, and B. Smit, J. Phys. Chem. B **103**, 1102 (1999).
[13] M. G. Martin and J. I. Siepmann, J. Phys. Chem. B **103**, 4508 (1999).
[14] M. D. Macedonia and E. J. Maginn, Mol. Phys. **96**, 1375 (1999).
[15] J. R. Errington and A. Z. Panagiotopoulos, J. Chem. Phys. **111**, 9731 (1999).
[16] B. Chen, M. G. Martin, and J. I. Siepmann, J. Phys. Chem. B **102**, 2578 (1998).
[17] A. Baumgärtner and K. Binder, J. Chem. Phys. **71**, 2541 (1979).
[18] A. J. Pertsin, J. Hahn, and H. P. Grossmann, J. Comput. Chem. **15**, 1121 (1994).
[19] B. Neubauer, A. Boutin, B. Tavitian, and A. H. Fuchs, Mol. Phys. **97**, 769 (1999).
[20] E. Bourasseau, P. Ungerer, and A. Boutin, J. Phys. Chem. B **106**, 5483 (2002).
[21] C. D. Wick and J. I. Siepmann, Macromolecules **33**, 7207 (2000).
[22] Z. Chen and F. A. Escobedo, J. Chem. Phys. **113**, 11382 (2000).
[23] H. Sklenar, D. Wüstner, and R. Rohs, J. Comput. Chem. **27**, 309 (2006).
[24] S. K. Nath and R. Khare, J. Chem. Phys. **115**, 10837 (2001).
[25] M. L. Greenfield, Ph.D. dissertation (University of California, 1996).
[26] See supplementary material at http://dx.doi.org/10.1063/1.3644939 for details of an alternative algorithm that attempts to perturb Cartesian coordinates of sites in a fragment; numerical results.
[27] L. G. Peristeras, I. G. Economou, and D. N. Theodorou, Macromolecules **38**, 386 (2005).
[28] M. G. Martin and J. I. Siepmann, J. Phys. Chem. B **102**, 2569 (1998).
[29] M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids* (Oxford University Press, Oxford, 1989).
[30] N. D. Zhuravlev, M. G. Martin, and J. I. Siepmann, Fluid Phase Equilib. **202**, 307 (2002).
[31] J.-S. Lee, C. D. Wick, J. M. Stubbs, and J. I. Siepmann, Mol. Phys. **103**, 99 (2005).
[32] W. Shi and E. J. Maginn, J. Comput. Chem. **29**, 2520 (2008).