# An advanced Gibbs-Duhem integration method: Theory and applications

A. van 't Hof, C. J. Peters, and S. W. de Leeuw[a)]

*Physical Chemistry and Molecular Thermodynamics, Faculty of Applied Sciences, Delft University of Technology, Julianalaan 136, 2628 BL Delft, The Netherlands*

The conventional Gibbs-Duhem integration method is very convenient for the prediction of phase equilibria of both pure components and mixtures. However, it turns out to be inefficient. The method requires a number of lengthy simulations to predict the state conditions at which phase coexistence occurs. This number is not known from the outset of the numerical integration process. Furthermore, the molecular configurations generated during the simulations are merely used to predict the coexistence condition and not the liquid- and vapor-phase densities and mole fractions at coexistence. In this publication, an advanced Gibbs-Duhem integration method is presented that overcomes above-mentioned disadvantage and inefficiency. The advanced method is a combination of Gibbs-Duhem integration and multiple-histogram reweighting. Application of multiple-histogram reweighting enables the substitution of the unknown number of simulations by a fixed and predetermined number. The advanced method has a retroactive nature; a current simulation improves the predictions of previously computed coexistence points as well. The advanced Gibbs-Duhem integration method has been applied for the prediction of vapor-liquid equilibria of a number of binary mixtures. The method turned out to be very convenient, much faster than the conventional method, and provided smooth simulation results. As the employed force fields perfectly predict pure-component vapor-liquid equilibria, the binary simulations were very well suitable for testing the performance of different sets of combining rules. Employing Lorentz-Hudson-McCoubrey combining rules for interactions between unlike molecules, as opposed to Lorentz-Berthelot combining rules for all interactions, considerably improved the agreement between experimental and simulated data. © *2006 American Institute of Physics.*
[DOI: 10.1063/1.2137706]

## I. INTRODUCTION

The synthesis of a particular chemical component usually requires subsequent purification of the product by the removal of by-products and/or a solvent. Sometimes, separation itself is the main function of an entire process. As separation processes usually account for the larger part of the operational and capital investment costs,[1] there is a large incentive for research directed towards cost-reducing improvements of separation processes and the development of new ones. Obviously, the selection of a proper separation process and the design of the separation equipment require knowledge of the thermodynamic data behind the process.[1–3] Many pure-component physical properties can be retrieved from literature. Although pure-component properties are surely interesting, industry often works with mixtures. Unfortunately, mixture properties are seldom available at the conditions of interest. Thus, these properties need to be measured in some way. In case one is able to prepare a representative sample and to devise a proper experimental setup, properties can be measured in the laboratory. However, experiments can be quite time consuming and experiments at extreme conditions require expensive equipment.

Molecular simulation can be an attractive alternative at extreme physical conditions as computers are insensitive to the circumstances of the system that is simulated. It can also conveniently be employed with mathematically simple force fields when a large number of candidate molecules or solvents needs to be screened for a particular application. As computer power has increased considerably over the last years, the size and complexity of the systems that can be tackled with molecular simulation have increased as well.

This work focuses on Monte Carlo molecular simulation of vapor-liquid equilibria of binary mixtures. Well-known methods that can be used to predict phase equilibria of mixtures are the $Np\beta$+test-molecule (TM) method,[4–7] simulations in the $Np\beta$ Gibbs ensemble,[8,9] simulations in the grand-canonical ensemble combined with multiple-histogram reweighting,[10,11] and the Gibbs-Duhem integration method.[12,13] The $Np\beta$+TM method relies on the accurate computation of the liquid- and vapor-phase chemical potentials at a reference state point.[4–7] The method gets computationally demanding in case one of the phases has a high density as the computation of the chemical potential is obstructed by the large number of test-molecule insertions with a negligible Boltzmann factor.[14,15] Gibbs ensemble simulations can be very convenient as they directly provide a coexistence point from a single simulation. However, applications of the pure-component $NV\beta$ Gibbs ensemble are always limited to phases at moderate densities, as the results heavily depend on the possibility to insert molecules into the coexisting phases.[16,17] Although the capability of the $Np\beta$

a)Electronic mail: s.w.deleeuw@tnw.tudelft.nl

Gibbs ensemble to predict vapor-liquid equilibria still depends on the ability to successfully insert molecules, the application to mixtures brings some ease; instead of performing insertions and removals of molecules for all components, it is convenient to only perform these moves for the smallest component and to carry out trial identity changes for the larger molecules.[9] Like the $Np\beta$+TM method and the $Np\beta$ Gibbs ensemble method, the combination of grand-canonical simulations and multiple-histogram reweighting needs successful insertions and removals of molecules.

Among above-mentioned methods, the Gibbs-Duhem integration (GDI) method has the advantage of not being dependent on insertions and/or removals of molecules. This independence makes the method especially suitable for the computation of phase equilibria including a solid phase.[18–20] In case a good estimate of an initial coexistence point is known, the method is appropriate for the prediction of complete phase diagrams.[18,21,22]

In this work, the GDI method is in the center of interest. A brief description of conventional GDI is given in Sec. II A. In Sec. II A, we also comment on the disadvantages and inefficiency of the conventional GDI method. The discussion of the conventional method is followed by the introduction of an advanced GDI method. This advanced method implies a combination of GDI and multiple-histogram reweighting. The idea of combining GDI and multiple-histogram reweighting is not completely new;[23] however, this publication describes the first successful application. The principle of histogram reweighting is explained in Sec. II B. The mathematical framework of advanced GDI is presented in Sec. II C. Details of the simulations and simulation results are presented and discussed in Sec. III. Section IV contains a summary and concluding remarks.

## II. METHODOLOGY

### A. Conventional Gibbs-Duhem integration

A Gibbs-Duhem equation describes the mutual dependence of state variables in an individual phase. Combination of the Gibbs-Duhem equations of coexisting phases enables one to derive a so-called Clapeyron equation that describes monovariant phase coexistence. The original Clapeyron equation describes how pressure must change with temperature for two pure-component phases to remain at coexistence. GDI implies the numerical integration of such a Clapeyron equation. The method is easily extended to binary or multicomponent mixtures[24] and multiphase systems[25] by deriving the appropriate Clapeyron equation. The Clapeyron equation that is considered in this work describes isothermal vapor ($V$)-liquid ($L$) equilibria in a binary mixture[24]

$$\left(\frac{\partial p}{\partial \xi_2}\right)_{\beta,\sigma} = \frac{x_2^L - x_2^V}{\xi_2(1-\xi_2)\beta(v^L - v^V)} = F(p,\xi_2), \quad (1)$$

where $\xi_2$ is the fugacity fraction of component two,[24,26,27] $x_2$ is the mole fraction of component two, $\nu$ is the volume per molecule, $p$ is the pressure, and $\beta = 1/(k_B T)$ is the reciprocal temperature, $k_B$ being Boltzmann's constant. The subscript $\sigma$ indicates that the derivative is taken along the saturation line. The dependent and independent integration variables in

Eq. (1) correspond to an isobaric semigrand-canonical ensemble.[24,26] The right-hand side of Eq. (1), $F(p,\xi_2)$, will be called "slope" in what follows.

In conventional numerical integration, the evaluation of the slope yields its exact value. In GDI, the slope is determined with molecular simulation, and the longer the simulation proceeds, the better the estimate for the slope gets

$$\left(\frac{\partial p}{\partial \xi_2}\right)_{\beta,\sigma} = \frac{1}{\xi_2(1-\xi_2)\beta} \frac{\langle x_2\rangle^L_{N^L p\beta\xi_2} - \langle x_2\rangle^V_{N^V p\beta\xi_2}}{\langle v\rangle^L_{N^L p\beta\xi_2} - \langle v\rangle^V_{N^V p\beta\xi_2}}, \quad (2)$$

where $\langle\cdots\rangle_{N^\gamma p\beta\xi_2}$ represents a semigrand-canonical ensemble average in phase $\gamma$, and $N^\gamma$ is the number of molecules. Since molecular simulation is a computationally demanding technique, it is important to choose a numerical integration method that needs as little evaluations of the slope as possible. Kofke[12,13] and Mehta and Kofke[24] suggested the use of predictor-corrector equations. In this work, higher-order variable-step-size predictor-corrector equations are used.[15]

To start the numerical integration of a differential equation, initial conditions are needed. A convenient choice for Clapeyron equation (2) is a pure-component coexistence point along with the corresponding initial value for the slope. Different simulation techniques that can be used to predict the initial point and the initial slope are described in previous work.[14,15] The numerical integration of Clapeyron equation (2) comprises a number of steps.

**Initial condition**.
(1) Compute the saturation pressure $p_0^{sat}$ of pure component one ($[\xi_2]_0 = 0$), along with the value of the slope $F(p_0^{sat},[\xi_2]_0 = 0)$, both at the temperature of interest.

**Prediction**.
(2) Increment the fugacity fraction, $[\xi_2]_i = [\xi_2]_{i-1} + [\Delta\xi_2]_i$, and estimate the saturation pressure $p_i^{sat}([\xi_2]_i)$ from the predictor equation corresponding to integration step $i$.
(3) Collect samples of the volumes and compositions in isobaric liquid- and vapor-phase semigrand-canonical simulations and compute the slope $F(p_i^{sat},[\xi_2]_i)$.

**Correction**.
(4) Apply the corrector equation that corresponds to integration step $i$ to update the value of $p_i^{sat}$ using the slope computed in step (3) and previously estimated slopes.
(5) Repeat steps (3) and (4) until the specified convergence criterion has been met.

**Production**.
(6) Perform isobaric liquid- and vapor-phase semigrand-canonical simulations at the estimated saturation pressure so as to obtain the final value of $F(p_i^{sat},[\xi_2]_i)$ and other relevant thermodynamic properties at phase coexistence.

**Proceed with next integration step (if required)**.
(7) Go to step (2).

A serious disadvantage of the conventional GDI scheme is that the number of correction simulations [steps (4) and (5)] that is needed to obtain convergence of the saturation pressure is not known beforehand. Furthermore, the samples

taken during the corrector simulations are merely used to let the saturation pressure converge in a mathematical sense. The samples do not contribute to the ensemble averages of the densities and mole fractions of the final coexistence point [production phase, step (6)]. This makes the conventional GDI method inefficient. A method that circumvents aforementioned inefficiency is presented in Sec. II C. Some necessary background is given in Sec. II B.

## B. Multiple-histogram reweighting

The molecular configurations generated during a simulation contain a large amount of information. It is a waste of data and time not to extract as much information as possible from those configurations. Histogram-reweighting methods are examples of data analysis tools that increase the amount of information extracted from a simulation. Histogram reweighting provides an estimate of the probability distribution of the ensemble of interest at thermodynamic state variables close to the ones the simulation was performed at.[28-33] From a reweighted probability distribution one can compute ensemble averages at nearby state variables.

Suppose one would like to know the canonical probability distribution at reciprocal temperature $\beta_0$. All samples of the energy that are obtained during the course of a canonical simulation are collected into a histogram $H_{\beta_0}(U)$. Energy level $j$ includes all energies $U$ within the following interval

$$U_j - \frac{1}{2}\Delta U \leq U < U_j + \frac{1}{2}\Delta U, \qquad (3)$$

where $\Delta U$ is the bin width of the energy levels. As the molecular configurations of the system are generated in accordance with their Boltzmann weights, the histogram can serve as an estimate $\pi^{\text{est}}$ of the true canonical probability distribution at reciprocal temperature $\beta_0$

$$\pi^{\text{est}}(U;\beta_0) = \frac{H_{\beta_0}(U)}{N_{\beta_0}} = \frac{\Omega^{\text{est}}(N,V,U)\exp(-\beta_0 U)}{\sum_{U'} \Omega^{\text{est}}(N,V,U')\exp(-\beta_0 U')}, \qquad (4)$$

where $N_{\beta_0}$ is the sum of the entries of histogram $H_{\beta_0}(U)$. The degeneracy $\Omega^{\text{est}}$ is estimated from

$$\Omega^{\text{est}}(N,V,U) = \frac{H_{\beta_0}(U)}{N_{\beta_0}}\exp[+\beta_0(U - A_{NV\beta_0})], \qquad (5)$$

where $A_{NV\beta_0}$ is the Helmholtz energy. From now on, the subscript $\beta_0$ will be replaced by 0. The canonical probability distribution at any reciprocal temperature $\beta$ can be estimated using Eqs. (4) and (5), by reweighting the histogram collected at $\beta_0$

$$\pi^{\text{est}}(U;\beta) = \frac{H_0(U)/N_0\exp[-(\beta-\beta_0)U]}{\sum_{U'} H_0(U')/N_0\exp[-(\beta-\beta_0)U']}. \qquad (6)$$

Every histogram obtained from a finite simulation is subject to statistical errors. Furthermore, the range of energies sampled during a finite simulation at a single temperature $\beta$ is relatively small. Thus, a single simulation only provides a reliable estimate of the degeneracy over a limited range of energies. If temperature $\beta$ differs too much from $\beta_0$, the peak of the reweighted histogram will be shifted to the tails of the measured histogram where the statistical uncertainty is high. This means that the extrapolative power of Eq. (6) is limited to a relatively narrow range of temperatures around $\beta$.[34-38] Nevertheless, there are methods that can be used to obtain a degeneracy that is valid over a broad range of energies.

The degeneracy is independent of state variables and can in principle be estimated from every histogram. As histograms collected at nearby temperatures overlap, a series of histograms collected at nearby temperatures covers a large part of the energies available to the system. This series of histograms can be combined with the multiple-histogram reweighting (MHR) method of Ferrenberg and Swendsen[28,29,39] and Swendsen.[37] The resulting equation for the degeneracy, estimated from histograms collected at $R$ state points, is given by

$$\Omega^{\text{est}}(N,V,U) = \frac{\sum_{i=1}^{R} H_i(U)/g_i}{\sum_{j=1}^{R} N_j/g_j \exp[-\beta_j(U - A_{NV\beta_j})]}, \qquad (7)$$

where the parameter $g_j$ depends on the correlation time of successive configurations collected at state point $j$.[29,37,39] The Helmholtz energies in Eq. (7) are estimated from

$$\exp(-\beta_i A_{NV\beta_i}) = \sum_{U} \Omega^{\text{est}}(N,V,U)\exp(-\beta_i U). \qquad (8)$$

Equation (7) and the Helmholtz energies at the $R$ state points form a dependent set of equations that can be solved self-consistently with a Newton-Raphson method.[39,40] The optimization of the Helmholtz energies is only possible when histograms obtained from different simulations overlap sufficiently. As Eqs. (7) and (8) determine the Helmholtz energies only within an additive constant, it is convenient to set one of the Helmholtz energies to zero and compute the other values with respect to this reference state. In case the histograms are extended to a thermodynamic state point with known Helmholtz energy, all absolute Helmholtz energies at other state points can be computed.

Now, we will derive the MHR relations for the binary isobaric semigrand-canonical ensemble. In this ensemble, the total energy $U$, volume $V$, and number of molecules of component two $N_2$ fluctuate, while the temperature $\beta$, pressure $p$, fugacity fraction $\xi_2$, and total number of molecules $N$ are fixed. The isobaric semigrand-canonical partition function is given by[15]

$$Y^*_{Np\beta\xi_2} = \int dV \frac{[Vq_1(\beta)]^N \exp(-\beta pV)}{V(N-1)!} \sum_{\text{iden}} \left[\frac{C_1(\beta)\xi_2}{C_2(\beta)\xi_1}\right]^{N_2}$$

$$\times \prod_{j=1}^{2}\prod_{i=1}^{N_j} \int J_{ij}(\mathbf{q}_i^{n_j}) d\mathbf{q}_i^{n_j} d\mathbf{s}_i^3$$

$$\times \exp[-\beta U^{\text{tot}}(\mathbf{s}^{3N}, \mathbf{q}^{N_1 n_1 + N_2 n_2})], \tag{9}$$

where $U^{\text{tot}}$ is the sum of the inter- and intramolecular interaction energies, $J_{ij}$ is the Jacobian of transformation from Cartesian to generalized coordinates of molecule $i$ which is of type $j$, $q_1(\beta)$ is the kinetic contribution of a molecule of component one, $\mathbf{s}_i^3$ represents the vector of overall translational coordinates of molecule $i$, scaled by the length of the cubic simulation box, and $\mathbf{q}_i^{n_j}$ represents the configuration vector of molecule $i$ which is of type $j$ and has got $n_j+3$ degrees of freedom. The summation indicated by "iden" represents a sum over all possible identities of all molecules.[24] The asterisk in Eq. (9) indicates that the partition function correctly counts volume states.[15,41–43] The thermodynamic potential belonging to partition function (9) is the effective full semigrand energy $\beta\mu_1 N$. The purely temperature-dependent variables $C_i(\beta)$ in Eq. (9) are given by

$$C_i(\beta) \int J_i(\mathbf{q}^{n_i}) d\mathbf{q}^{n_i} d\mathbf{s}_i^3 \exp(-\beta U^{\text{intra}}), \tag{10}$$

where $U^{\text{intra}}$ comprises all intramolecular contributions to the interaction energy and $J_i$ is the Jacobian of transformation from Cartesian to generalized coordinates of a molecule of type $i$. To simplify the notation, the factor $J_{ij}(\mathbf{q}_i^{n_j}) d\mathbf{q}_i^{n_j} d\mathbf{s}_i^3$ will be indicated by $d\mathbf{q}'_{ij}$ in what follows. In order to introduce the degeneracy into the isobaric semigrand-canonical partition function, a total energy $U_0^{\text{tot}}$, volume $V_0$, and number of molecules of component two $N_2^0$, along with Dirac $\delta$ functions and a Kronecker delta, are introduced in Eq. (9)

$$Y^*_{Np\beta\xi_2} = \sum_{N_2^0} \int dV_0 \int dU_0^{\text{tot}} \int dV \delta(V-V_0)$$

$$\times \frac{[Vq_1(\beta)]^N \exp(-\beta pV)}{V(N-1)!} \sum_{\text{iden}} \delta_{N_2,N_2^0} \left[\frac{C_1(\beta)\xi_2}{C_2(\beta)\xi_1}\right]^{N_2}$$

$$\times \prod_{j=1}^{2}\prod_{i=1}^{N_j} \int d\mathbf{q}'_{ij} \delta(U^{\text{tot}}-U_0^{\text{tot}}) \exp(-\beta U^{\text{tot}}). \tag{11}$$

Now, partition function (11) is written as the summation over a product of the degeneracy and a pseudo-Boltzmann factor. Keeping all parts that depend on temperature, pressure, and fugacity fraction out of the degeneracy results in

$$Y^*_{Np\beta\xi_2} = \sum_{N_2^0}\sum_{V_0}\sum_{U_0^{\text{tot}}} \Omega^*(N_2^0, V_0, U_0^{\text{tot}})$$

$$\times \exp\left\{-\beta U_0^{\text{tot}} - \beta p V_0 + N\ln\left[q_1(\beta)V_0 \int d\mathbf{q}'_1\right] \right.$$

$$\left. + N_2^0 \ln\left[\frac{\xi_2 C_1(\beta)\int d\mathbf{q}'_2}{\xi_1 C_2(\beta)\int d\mathbf{q}'_1}\right]\right\}, \tag{12}$$

where the degeneracy $\Omega^*(N_2^0, V_0, U_0^{\text{tot}})$ is given by

$$\Omega^*(N_2^0, V_0, U_0^{\text{tot}}) = \frac{\Delta V_0 \Delta U_0^{\text{tot}}}{(N-1)!}\int dV \frac{\delta(V-V_0)}{V}$$

$$\times \sum_{\text{iden}} \delta_{N_2,N_2^0} \frac{\prod_{j=1}^{2}\prod_{i=1}^{N_j}\int d\mathbf{q}'_{ij}\delta(U^{\text{tot}}-U_0^{\text{tot}})}{\prod_{j=1}^{2}\prod_{i=1}^{N_j}\int d\mathbf{q}'_{ij}}, \tag{13}$$

where $\Delta V_0$ and $\Delta U_0^{\text{tot}}$ are bin widths of the histograms. Over the course of a simulation performed at thermodynamic state point $i$, a histogram is collected containing the energy-volume-composition combinations the system encounters

$$H_i(N_2, V, U^{\text{tot}}) \equiv H(N_2, V, U^{\text{tot}}; [\xi_2]_i, p_i, \beta_i). \tag{14}$$

Simulations performed at $R$ different thermodynamic state points are combined with the MHR method in order to obtain an estimate of the degeneracy $\Omega^*(N_2, V, U^{\text{tot}})$ that is valid over a large range of temperatures, pressures, and fugacity fractions

$$\Omega^{*\text{ est}}(N_2, V, U^{\text{tot}})$$

$$= \frac{\sum_{i=1}^{R} H_i(N_2, V, U^{\text{tot}})/g_i}{\sum_{j=1}^{R}(N_j/g_j)\exp[-\beta_j U^{\text{tot}} - \beta_j p_j V + N\ln(q_1 V \int d\mathbf{q}'_1) + N_2\ln(C_1\int d\mathbf{q}'_2[\xi_2]_j/C_2\int d\mathbf{q}'_1[\xi_1]_j) + \beta_j[\mu_1]_j N]}. \tag{15}$$

Combination of the effective full semigrand energy and several nonfluctuating purely temperature-dependent contributions in the pseudo-Boltzmann factor in Eq. (15) gives the effective configurational semigrand energy

$$N \ln(\hat{f}_1/p^{\text{ref}}) = \beta\mu_1 N + N \ln\left(\frac{q_1 C_1}{\beta p^{\text{ref}}}\right),$$ (16)

where $p^{\text{ref}}$ is a reference pressure, the value of which is trivial in this work, and $\hat{f}_1$ is the fugacity of component one in the binary mixture. The estimate of the degeneracy and the effective configurational semigrand energies are found self-consistently by iterating Eqs. (17) and (19)

$$\Omega^{*\,\text{est}}(N_2,V,U^{\text{tot}}) = \frac{\displaystyle\sum_{i=1}^{R} H_i(N_2,V,U^{\text{tot}})/g_i}{\displaystyle\sum_{j=1}^{R} N_j/g_j \exp[-\beta_j U^{\text{tot}} - \beta_j p_j V + \kappa(\beta_j,N_2,V) + N_2 \ln([\xi_2]_j/[\xi_1]_j) + N \ln([\hat{f}_1]_j/p^{\text{ref}})]}.$$ (17)

The variable $\kappa(\beta_j,N_2)$ in Eq. (17) is defined by

$$\kappa(\beta_j,N_2,V) \equiv N \ln(V\beta_j p^{\text{ref}}) + N_1 \ln\left[\frac{\int d\mathbf{q}_1'}{C_1(\beta_j)}\right]$$
$$+ N_2 \ln\left[\frac{\int d\mathbf{q}_2'}{C_2(\beta_j)}\right].$$ (18)

The effective configurational semigrand energy at thermodynamic state point $j$ is given by

$$N \ln([\hat{f}_1]_j/p^{\text{ref}}) = -\ln\left\{ \sum_{N_2}\sum_{V}\sum_{U^{\text{tot}}} \Omega^*(N_2,V,U^{\text{tot}}) \right.$$
$$\times \exp\left[ -\beta_j U^{\text{tot}} - \beta_j p_j V + \kappa(\beta_j,N_2,V) \right.$$
$$\left. \left. + N_2 \ln\left(\frac{[\xi_2]_j}{[\xi_1]_j}\right) \right] \right\}.$$ (19)

Finally, the probability to observe the system with a particular volume-energy-composition combination at temperature $\beta$, fugacity fraction $\xi_2$, and pressure $p$ is given by

$$\pi(N_2,V,U^{\text{tot}};\beta,\xi_2,p) = \frac{\Omega^*(N_2,V,U^{\text{tot}})\exp[-\beta U^{\text{tot}} - \beta p V + \kappa(\beta,N_2,V) + N_2 \ln(\xi_2/\xi_1)]}{\displaystyle\sum_{N_2'}\sum_{V'}\sum_{U'^{\text{tot}}} \Omega^*(N_2',V',U'^{\text{tot}})\exp[-\beta U'^{\text{tot}} - \beta p V' + \kappa(\beta,N_2',V') + N_2' \ln(\xi_2/\xi_1)]}.$$ (20)

A simplified version of probability distribution (20) will be combined with predictor-corrector methods in Sec. II C to predict vapor-liquid equilibria of binary mixtures.

### C. Advanced Gibbs-Duhem integration

The search for phase-equilibrium conditions requires the frequent adjustment of state variables and the subsequent computation of ensemble averages at those state variables. As MHR is a perfect means to obtain ensemble averages over a range of state variables from only a few simulations, it is very convenient in phase-equilibrium calculations.

To actually compute phase coexistence from MHR, there must be a connection between the liquid and the vapor phase; for example, overlap of liquid- and vapor-phase histograms. If such a connection can be realized and MHR is applied to all liquid- and vapor-phase histograms simultaneously, the optimized ensemble probability distribution of the order parameter close to vapor-liquid phase coexistence is double peaked. Examples of the order parameter are the density in a pure-component system and the composition in a

binary system. Each of the two peaks in the bimodal distribution corresponds to a single phase. The coexistence condition is given by those state variables for which the areas under the two peaks are equal. This is the so-called criterion of equal peak weight.[44,45] In case the bimodal probability distribution of the order parameter can be measured during a single simulation in a single simulation box, phase coexistence can be simply obtained by reweighting a single histogram. However, accurate prediction of the relative peak areas in a single simulation is often difficult as the free-energy barrier for condensation or vaporization, which is high at low temperatures, decreases the frequency of sampling of both phases. This problem may be solved by artificially lowering the free-energy barrier with multicanonical sampling methods.[45–49]

In case direct overlap cannot be realized or is computationally too demanding, an indirect connection between the liquid and the vapor phase must be sought for. The connection may be realized by performing a simulation at a state point in the vicinity of the vapor-liquid critical point.[49–51]

The free-energy barrier for condensation or vaporization close to the vapor-liquid critical point is so low that a single simulation will frequently sample both liquid and vaporlike configurations. Additional histograms are collected at subcritical liquid- and vapor-phase conditions. Application of MHR to all liquid- and vapor-phase histograms simultaneously results in an estimate of the degeneracy from which one can extract the bimodal probability distribution of the order parameter over a range of state variables close to coexistence conditions. Clearly, application of above-mentioned methods requires some knowledge about the position of the vapor-liquid critical point and the coexistence line before a series of simulations is started.

As pointed out in Sec. II A, the conventional GDI scheme starts with the prediction of the saturation pressure at an incremented value of the fugacity fraction. An isobaric semigrand-canonical simulation at the estimated saturation pressure provides an estimate of the slope of the Clapeyron equation. The first estimate of the saturation pressure is then updated by applying a corrector equation. Subsequently, a new simulation at the updated saturation pressure is performed to again compute the slope of the Clapeyron equation. For sufficiently small integration steps, the prediction and correction simulations are performed at only slightly different state points. For that reason, application of the histogram-reweighting method to histograms obtained from a single liquid- and vapor-phase simulation at the predicted saturation pressure can be conveniently used to replace all simulations belonging to the prediction, correction, and production steps.

The condition for phase coexistence used in this work is implied in the corrector equation and in previously determined coexistence points. Thus, the liquid- and vapor-phase histograms neither need direct nor indirect overlap, which is clearly an advantage over other MHR methods. The combination of GDI and histogram reweighting will be called "advanced Gibbs-Duhem integration" in what follows.

As already pointed out, the advanced GDI method strictly requires only one liquid- and vapor-phase simulation per coexistence point. However, the extrapolative power of a single histogram is limited to only a narrow range of state variables. Therefore, the single-histogram reweighting approach is only appropriate in case the integration step is very small. In order to make the method applicable for larger integration steps, a number of simulations is performed at state points that cover the region nearby the coexistence condition that is estimated with the predictor equation. This enables one to compute phase coexistence from MHR by interpolation instead of extrapolation to nearby state conditions.

Equations (17) and (20) can be simplified considerably by remembering that all simulations in this work are performed isothermally. Thus, reweighting to other temperatures is not needed to compute phase coexistence. The conjugate density of the reciprocal temperature $\beta$ is the energy $U$. It is possible to perform the summation over the energy levels beforehand. The pseudo-Boltzmann factor in Eq. (17) is referred to a certain reference state point, say, $(p_0, [\xi_2]_0, \beta_0 = \beta)$. Division of the denominator and numerator of Eq. (17) by the pseudo-Boltzmann factor at the reference state point gives

$$\Omega^{*\,\mathrm{est}}(N_2, V, U^{\mathrm{tot}}) = \frac{\displaystyle\sum_{i=1}^{R} H_i(N_2, V, U^{\mathrm{tot}})/g_i \exp[+\beta U^{\mathrm{tot}} + \beta p_0 V - \kappa(\beta, N_2, V) - N_2 \ln([\xi_2]_0/[\xi_1]_0)]}{\displaystyle\sum_{j=1}^{R} N_j/g_j \exp[-\beta(p_j - p_0)V + N_2 \ln([\xi_2]_j[\xi_1]_0/[\xi_1]_j[\xi_2]_0) + N \ln([\hat{f}_1]_j/p^{\mathrm{ref}})]}. \tag{21}$$

It is convenient to define a temperature-dependent pseudodegeneracy

$$\Theta^{*\,\mathrm{est}}(N_2, V) = \frac{\displaystyle\sum_{i=1}^{R} H_i(N_2, V)/g_i}{\displaystyle\sum_{j=1}^{R} N_j/g_j \exp[-\beta(p_j - p_0)V + N_2 \ln([\xi_2]_j[\xi_1]_0/[\xi_1]_j[\xi_2]_0) + N \ln([\hat{f}_1]_j/p^{\mathrm{ref}})]}. \tag{22}$$

Using Eq. (22), probability distribution (20) simplifies to

$$\pi(N_2, V; \xi_2, p) = \frac{\Theta^{*\,\mathrm{est}}(N_2, V)\exp[-\beta(p - p_0)V + N_2 \ln(\xi_2[\xi_1]_0/\xi_1[\xi_2]_0)]}{\displaystyle\sum_{N_2'}\sum_{V'} \Theta^{*\,\mathrm{est}}(N_2', V')\exp[-\beta(p - p_0)V' + N_2' \ln(\xi_2[\xi_1]_0/\xi_1[\xi_2]_0)]}. \tag{23}$$

The total procedure of advanced GDI comprises a number of steps.

**Initial condition**.

(1) Compute the saturation pressure $p_0^{sat}$ of pure component one ($[\xi_2]_0=0$), along with the value of the slope $F(p_0^{sat}, [\xi_2]_0=0)$, both at the temperature of interest.

**Prediction**.

(2) Increment the fugacity fraction, $[\xi_2]_i=[\xi_2]_{i-1}+[\Delta\xi_2]_i$, and estimate the corresponding saturation pressure $p_i^{sat}([\xi_2]_i)$ from the predictor equation corresponding to this integration step. In order to cover the region of interest, $m-1$ further thermodynamic state points, $(p_j, [\xi_2]_i; j=1\cdots m-1)$, are chosen with values close to the estimated saturation pressure at the imposed fugacity fraction.

(3) During the course of the simulations at integration step $i$, $m$ liquid-phase and $m$ vapor-phase histograms $H_{ij}(N_2, V)$ are collected. The histograms of the liquid phase are combined with the method of Ferrenberg and Swendsen so as to obtain an estimate of the probability distribution $\pi^L(N_2, V; \xi_2, p)$ that is valid over the range of pressures and fugacity fractions covered by the simulations. The same procedure applies to the vapor phase.

**Correction**.

(4) The estimated slope of Clapeyron equation (2), $(p_i^{sat}, [\xi_2]_i; \beta)$, is obtained from

$$F_{Np_i\beta[\xi_2]_i} = \frac{1}{[\xi_2]_i(1-[\xi_2]_i)}\frac{1}{\beta p_i}$$
$$\times \frac{\langle x_2\rangle^L_{N^Lp_i\beta[\xi_2]_i} - \langle x_2\rangle^V_{N^Vp_i\beta[\xi_2]_i}}{\langle v\rangle^L_{N^Lp_i\beta[\xi_2]_i} - \langle v\rangle^V_{N^Vp_i\beta[\xi_2]_i}}. \quad (24)$$

The properties within brackets are computed from

$$\langle O\rangle^\alpha_{N^\alpha p_i\beta[\xi_2]_i} = \sum_{N_2^\alpha}\sum_{V^\alpha} O(N_2^\alpha, V^\alpha)\pi^\alpha(N_2^\alpha, V^\alpha; p_i, [\xi_2]_i). \quad (25)$$

(5) Apply the relevant corrector equation to update the value of $p_i^{sat}$, using the slope computed in step (4) along with previously computed slopes.

(6) Repeat steps (4) and (5) until the specified convergence criterion has been met. Ensemble averages of relevant properties at coexistence conditions are again obtained from Eq. (25).

**Proceed with next integration step (if required)**.

(7) Go to step (2).

In the procedure described above, the probability distribution $\pi^\alpha(N_2, V; \xi_2, p)$ is estimated by combining $m$ histograms at a single integration step. However, in case the integration steps are not too large, histograms obtained at different fugacity fractions (integration steps) have considerable overlap as well. Combining the histograms of several integration steps extends the region of validity of $\pi^\alpha(N_2, V; \xi_2, p)$. This procedure enables the computation of a whole coexistence line instead of only coexistence points at predetermined fugacity fractions. This coexistence line can be represented as a polynomial of saturation pressure expressed in powers of the fugacity fraction. Such a polyno-

mial, along with the probabilities $\pi^\alpha(N_2, V; \xi_2, p)$, provides the possibility to compute a coexistence point corresponding to a particular mole fraction or density of either phase. This is advantageous as, from an experimental point of view, mole fractions and densities are more interesting properties than a fugacity fraction is.

As histograms collected at the current integration step also improve the estimate of that part of the degeneracy that is important at prior integration steps, the current simulation improves the predictions at previously computed coexistence points as well. This retroactive behavior is an advantage of the advanced GDI method over the conventional method. The fixed number of simulations, as opposed to an unknown number, is another advantage of the advanced GDI method. It is expected that the retroactive nature and the fixed number of simulations make the advanced method faster than the conventional method for similar statistical errors of the results.

## D. Error analysis of advanced Gibbs-Duhem integration

In this section, the major sources of error in advanced GDI are discussed: the propagation of statistical errors via the corrector equations and the influence of the error in the initial point on the deviation of the predicted coexistence curve from the true coexistence curve. As pointed out in Sec. II B, the extrapolative power of histogram reweighting is limited due to the poorly sampled tails of the histograms.[35,36] However, in this work interpolation between histograms (state points) is used instead of extrapolation to regions where few data have been collected. Furthermore, it is checked whether histograms at nearby thermodynamic state points have sufficient overlap. For that reason, the errors due to bad sampling of the tails of the histograms and due to insufficient overlap of histograms will not be considered.

### 1. Propagation of statistical errors

The slopes of the Clapeyron equation at successive integration points are computed from isobaric semigrand-canonical Monte Carlo simulations. In this example, it is assumed that the integration variable is the fugacity fraction although it can be the saturation pressure as well. As the successive estimates of the saturation pressures are coupled via corrector equations,[15] the variance of the saturation pressure at a particular thermodynamic state point depends on the variances of formerly determined saturation pressures/slopes of the Clapeyron equation. The set of corrector equations for subsequent integration steps forms a linear system. After successive substitution of corrector equations into one another, one obtains the influence of the saturation pressure on predetermined slopes[15]

$$\mathbf{P} = [\mathbf{\Omega}]\mathbf{F} + p_0, \quad (26)$$

where $p_0$ is the saturation pressure corresponding to the estimate of the initial coexistence point, and $[\mathbf{\Omega}]$ is a matrix that contains parameters from corrector equations. The elements of the vectors $\mathbf{P}$ and $\mathbf{F}$ respectively, contain the saturation pressures and slopes of successive integration steps.

The variance $\sigma_{p_n}^2$ of the saturation pressure at integration step $n$ is given by

$$\sigma_{p_n}^2 = \mathbf{\Omega}_n \mathbf{\Sigma}(\mathbf{F}) \mathbf{\Omega}_n^T, \qquad (27)$$

where $\mathbf{\Sigma}(\mathbf{F})$ is the variance-covariance matrix of the slopes, and $\Omega_n$ is the $n$th row vector of matrix $[\mathbf{\Omega}]$. The slopes at successive integration points are statistically dependent as they are computed from a degeneracy that has been optimized to samples collected at multiple state points simultaneously. Thus, covariances in $\mathbf{\Sigma}(\mathbf{F})$ are generally nonzero.

Instead of collecting data at thermodynamic state point $i$ into a single histogram, data are collected on a regular basis into $M$ subhistograms. These subhistograms are more or less statistically independent. The variances and covariances in the matrix $\mathbf{\Sigma}(\mathbf{F})$ and other ensemble averages at coexistence are computed as follows.

(1) From the $M$ subhistograms collected at thermodynamic state point $i$ in phase $\alpha$ randomly pick $M$ subhistograms, not trying to avoid multiple sampling. The sum of these selected subhistograms forms a bootstrap-sample[52] histogram, $H_i^{*,\alpha}(N_2, V)$, of $H_i^\alpha(N_2, V)$.

(2) Repeat step (1) for every histogram that takes part in MHR.

(3) Perform MHR with the bootstrap samples of the histograms. This results in bootstrap-sample probabilities $\pi^{*,\alpha}(N_2, V; [\xi_2]_i, p_i)$ for both phases.

(4) Application of Eqs. (24) and (25) and the relevant corrector equations provides bootstrap samples of the saturation pressures $p_i^*$, densities $(\rho_i^{*,L}, \rho_i^{*,V})$ and mole fractions $([x_2^{*,L}]_i, [x_2^{*,V}]_i)$ at coexistence, and slopes $F_i^*$ at integration step $i$.

(5) Repeat steps (1)–(4) a large number of times, say, $N_B$ times, and compute the bootstrap estimate of the variance of variable $Y_i$ at thermodynamic state point $i$ and the covariance between variables $Y_i$ and $Y_k$ at thermodynamic state points $i$ and $k$ from

$$\sigma_{Y_i}^2 = \frac{1}{N_B - 1} \sum_{j=1}^{N_B} (Y_{ij}^* - \langle Y_i \rangle)^2, \qquad (28)$$

$$\sigma_{Y_i Y_k}^2 = \frac{1}{N_B - 1} \sum_{j=1}^{N_B} (Y_{ij}^* - \langle Y_i \rangle)(Y_{kj}^* - \langle Y_k \rangle). \qquad (29)$$

### 2. Influence of the error in the initial point on deviation from the true coexistence line

Gibbs-Duhem integration commences from an estimate of the initial coexistence point. The saturation pressure at the initial coexistence point has a statistical uncertainty and will deviate from the true, unknown, saturation pressure at the temperature of interest. As a consequence, the liquid-phase effective configurational semigrand energy per molecule, $\ln(\hat{f}_1/p^{\rm ref})$, will be different from the vapor-phase value. Thus, in general GDI will trace a curve of nonzero free-energy difference between the phases.

Kofke[53] analyzed the influence of the error in the initial coexistence point on the difference between the true outcome and the result obtained from GDI. Such a comparison is only possible in case the true coexistence line is known. Unfortunately, this is not the case in this work. The analysis goes as follows.[53] The variable that is fixed throughout the series of integration steps is the temperature. The fugacity fraction is the independent integration variable and the pressure the dependent one. The Taylor-series expansion of $\ln(\hat{f}_1/p^{\rm ref})$ up to first order in the pressure expanded about the true saturation pressure $p^{\rm sat}$ at the imposed fugacity fraction $[\xi_2]_0$ and reciprocal temperature $\beta$ is given by

$$\ln\left\{\frac{\hat{f}_1}{p^{\rm ref}}\right\}(p; [\xi_2]_0, \beta) \approx \ln\left\{\frac{\hat{f}_1}{p^{\rm ref}}\right\}(p^{\rm sat}; [\xi_2]_0, \beta) + \beta\langle v\rangle_{[\xi_2]_0, p^{\rm sat}, \beta}(p - p^{\rm sat}). \qquad (30)$$

Subtraction of the truncated Taylor-series expansion for the liquid phase from the expansion for the vapor phase gives

$$\Delta^{LV} \ln\left\{\frac{\hat{f}_1}{p^{\rm ref}}\right\}(p; [\xi_2]_0, \beta) \approx \beta \Delta^{LV}\langle v\rangle_{[\xi_2]_0, p^{\rm sat}, \beta}(p - p^{\rm sat}). \qquad (31)$$

Assume that the Gibbs-Duhem integration method keeps the difference $\Delta^{LV} \ln(\hat{f}_1/p^{\rm ref})$ unchanged. This assumption results in the following relation for the deviation of the predicted saturation pressure from the true saturation pressure at coexistence point $i$.[53]

$$(p_i^{\rm sat,est} - p_i^{\rm sat}) = \frac{\Delta^{LV}\langle v\rangle_{[\xi_2]_0, p_0^{\rm sat}, \beta}}{\Delta^{LV}\langle v\rangle_{[\xi_2]_i, p_i^{\rm sat}, \beta}}(p_0^{\rm sat,est} - p_0^{\rm sat}). \qquad (32)$$

Obviously, it makes sense to perform the integration into the direction of increasing $\Delta^{LV} v$ so as to weaken the influence of the error in the estimate of the initial saturation pressure.[53] However, this effect is only important if the difference $\Delta^{LV} v$ changes significantly over the range of fugacity fractions.

## III. RESULTS AND DISCUSSION

### A. Simulation details

The force field that is used in this work is a Lennard-Jones 12-6 potential with Coulombic charges on the interaction sites wherever necessary. The Coulombic energy is computed with the Ewald summation method.[54] The Ewald screening parameter and the upper bound in the reciprocal space were fixed at $\alpha = 5.6$ and $k_{\max} = 6$. The total intermolecular potential is given by

$$U^{\rm inter} = 4 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \varepsilon_{\gamma\delta}\left\{\left(\frac{\sigma_{\gamma\delta}}{r_{i\gamma j\delta}}\right)^{12} - \left(\frac{\sigma_{\gamma\delta}}{r_{i\gamma j\delta}}\right)^{6}\right\}$$
$$+ \frac{1}{4\pi\varepsilon_0} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \frac{q_i^\gamma q_j^\delta}{r_{i\gamma j\delta}}, \qquad (33)$$

where $\varepsilon_0$ is the vacuum permittivity, $q_i^\gamma$ is the partial charge $\gamma$ on molecule $i$, $n_i$ is the number of partial charges on molecule $i$, $r_{i\gamma j\delta}$ is the distance between interaction sites $i\gamma$ and

TABLE I. Vapor-liquid coexistence points of pure components along with simulation results for ratios of solvent's fugacity over solute's Henry's constant at coexistence.

| $T$ (K) | $p^{\text{sim}}$ (MPa) | $\rho_L^{\text{sim}}$(kg/m³) | $\rho_V^{\text{sim}}$(kg/m³) | $\left(\dfrac{f_{\text{C}_2\text{H}_6}}{H_{\text{CH}_4}}\right)_L$ | $\left(\dfrac{f_{\text{C}_2\text{H}_6}}{H_{\text{CH}_4}}\right)_V$ | Comb. rule unlike molecules | Simulation method |
|---|---|---|---|---|---|---|---|
| 250 | 1.313(7) | 443.5(4) | 23.3(2) | 0.108(6) | 0.83(2) | LHMcC | DM+OlD |

| $T$ (K) | $p^{\text{sim}}$ (MPa) | $\rho_L^{\text{sim}}$(kg/m³) | $\rho_V^{\text{sim}}$(kg/m³) | $\left(\dfrac{f_{\text{C}_3\text{H}_8}}{H_{\text{CH}_4}}\right)_L$ | $\left(\dfrac{f_{\text{C}_3\text{H}_8}}{H_{\text{CH}_4}}\right)_V$ | Comb. rule unlike molecules | Simulation method |
|---|---|---|---|---|---|---|---|
| 270 | 0.433(9) | 527.2(3) | 9.3(2) | 0.32(5) | 0.89(2) | LHMcC | DM+OlD |

| $T$ (K) | $p^{\text{sim}}$ (MPa) | $\rho_L^{\text{sim}}$(kg/m³) | $\rho_V^{\text{sim}}$(kg/m³) | $\left(\dfrac{f_{\text{C}_2\text{H}_6}}{H_{\text{CO}_2}}\right)_L$ | $\left(\dfrac{f_{\text{C}_2\text{H}_6}}{H_{\text{CO}_2}}\right)_V$ | Comb. rule unlike molecules | Simulation method |
|---|---|---|---|---|---|---|---|
| 207 | 0.297(5) | 506.6(4) | 5.5(1) | 0.20(1) | 0.97(1) | LHMcC | DM+OlD |
| 250 | 1.313(7) | 443.5(4) | 23.3(2) | 0.32(1) | 0.89(1) | LHMcC | DM+OlD |
| 263.15 | 1.85(1) | 419.8(5) | 33.2(2) | 0.34(2) | 0.87(2) | LHMcC | DM+OlD |
| 207 | 0.297(5) | 506.6(4) | 5.5(1) | 0.252(9) | 0.961(8) | LB | DM+OlD |
| 250 | 1.313(7) | 443.5(4) | 23.3(2) | 0.37(3) | 0.89(2) | LB | DM+OlD |
| 263.15 | 1.85(1) | 419.8(5) | 33.2(2) | 0.40(3) | 0.88(3) | LB | DM+OlD |

| $T$ (K) | $p^{\text{sim}}$ (MPa) | $\rho_L^{\text{sim}}$(kg/m³) | $\rho_V^{\text{sim}}$(kg/m³) | $\left(\dfrac{f_{\text{C}_3\text{H}_8}}{H_{\text{CO}_2}}\right)_L$ | $\left(\dfrac{f_{\text{C}_3\text{H}_8}}{H_{\text{CO}_2}}\right)_V$ | Comb. rule unlike molecules | Simulation method |
|---|---|---|---|---|---|---|---|
| 230 | 0.096(4) | 576.1(4) | 2.3(1) | 0.029(4) | 0.96(2) | LHMcC | DM+OlD |
| 270 | 0.433(9) | 527.2(3) | 9.3(2) | 0.088(9) | 0.90(1) | LHMcC | DM+OlD |
| 294.26 | 0.88(1) | 493.7(8) | 18.7(2) | 0.093(1) | 0.85(3) | LHMcC | DM+OlD |

$j\delta$, and $\sigma_{\gamma\delta}$ and $\varepsilon_{\alpha\gamma}$, respectively, are the Lennard-Jones collision-diameter and well-depth parameters for the interaction between site $\gamma$ and site $\delta$. The usual Lennard-Jones long-range corrections were applied[54] with a potential cutoff of at least three times the largest Lennard-Jones collision-diameter $\sigma$ in the simulated system. The cutoff was checked against half the box length after every successful volume change.

The Lennard-Jones parameters for interactions between unlike interaction sites were computed from so-called combining rules. The most widely used combining rules for Lennard-Jones parameters are the Lorentz rule for the collision diameters and the Berthelot rule for the well depths[54]

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2}, \quad \varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}. \tag{34}$$

The Lorentz rule is exact for mixtures of hard spheres. The Berthelot combining rule is a simplification of the combining rule that can be derived from the London theory of dispersion forces.[55,56] It is known to overestimate the interactions between unlike molecules.[55,57] Several researchers tried to correct for this systematic deviation by the inclusion of binary-interaction parameters in the Lorentz/Berthelot (LB) combining rules[4,5,7,58–61] with values close to zero.[4,60–62]

Over the years, a large variety of combining rules has been developed.[55,63,64] A combining rule for the well depths, which weakens the strength of the interactions between unlike groups with respect to the Berthelot rule, was given by Hudson and McCoubrey (HMcC)[56]

$$\varepsilon_{ij}^{\text{HMcC}} \approx 2\frac{\sqrt{I_i I_j}}{I_i + I_j}\left(\frac{2\sqrt{\sigma_{ii}\sigma_{jj}}}{\sigma_{ii} + \sigma_{jj}}\right)^6 \sqrt{\varepsilon_{ii}\varepsilon_{jj}}, \tag{35}$$

where $I_j$ is the ionization potential belonging to the molecule that contains interaction site $j$. Although superior to the Ber-thelot combining rule, the HMcC combining rule still tends to overestimate the dispersion energy between unlike molecules.[55] For molecules with equal collision diameters and ionization potentials, Eq. (35) reduces to the Berthelot combining rule.

Khalaf Al-Matar and Rockstraw[65] state that combining rules are meant to reduce the number of parameters in a force field. In their opinion, incorporation of additional parameters such as ionization potentials should be avoided whenever possible. Of course, it makes sense to exclude parameters derived from experimental mixture data. However, incorporation of ionization potentials or polarizabilities, which are pure-component properties that are available for a lot of molecules, is not against the purpose of combining rules.

The simulations in this work were either performed with the LB combining rules for all interactions or with the LB combining rules for interactions between like molecules and Lorentz/Hudson-McCoubrey (LHMcC) combining rules for interactions between unlike molecules. The pure-component interactions are to be computed with the LB combining rules as force-field parameters for the components of interest are optimized with these combining rules. For that reason, the HMcC combining rule was only used for interactions between unlike molecules.

Simulation results for several binary mixtures are presented in Secs. III B–III E. The force-field parameters for the molecules of interest were taken from literature. Methane, ethane, and propane were modeled with the TraPPE-EH force field,[66] and carbon dioxide was modeled with the TraPPE-AA force field.[10] These all-atom force fields predict pure-component vapor-liquid equilibria in close agreement with experimental data.[15,66] Thus, simulation of the vapor-
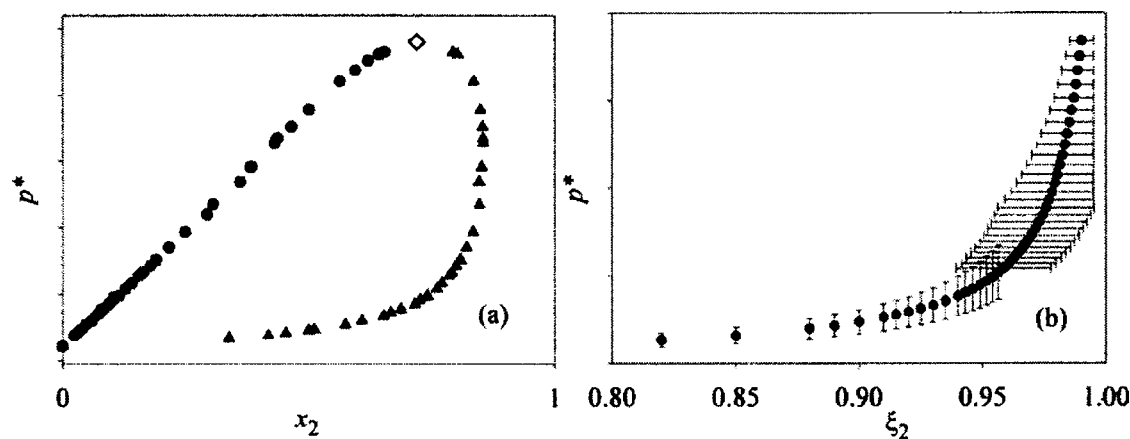
FIG. 1. Schematic pictures of a system in which component two is supercritical. (a) Vapor-liquid equilibria. (b) Dimensionless saturation pressure $p^*$ vs fugacity fraction. The symbols between vertical bars represent saturation pressures obtained from the predictor equation at an imposed fugacity fraction. On the other hand, the symbols between horizontal bars represent fugacity fractions obtained from the predictor equation at imposed saturation pressures. The ends of the bars represent the chosen state points in the neighborhood of the estimated coexistence point.

liquid equilibria of mixtures of above-mentioned components is a perfect means to judge the capabilities of different combining rules.

Intramolecular flexibility in the TraPPE-EH force field[66] comprises bond-angle bending and rotation about torsion angles. The intramolecular potential, without Lennard-Jones and Coulombic contributions, is given by

$$U^{\text{intra}} = \frac{k_\theta}{2}(\theta - \theta_0)^2 + c_1[1 + \cos \phi] + c_2[1 - \cos 2\phi]$$
$$+ c_3[1 + \cos 3\phi], \tag{36}$$

where $\theta$ is a bond angle, $\theta_0$ is the expectation value of bond angle $\theta$, and $\phi$ is a torsion angle. The types of molecules that are used in this work neither have intramolecular Lennard-Jones nor intramolecular Coulombic interactions.

Liquid phases contained $N^L = 300$ molecules and vapor phases $N^V = 100$ molecules. After 2500 equilibration Monte Carlo (MC) cycles, 10 000 production MC cycles were performed. Every MC cycle consisted of $0.7N^i$ translations, $0.3N^i$ rotations, $N^i$ identity changes, a change of the volume, and $0.05N^i$ parallel-tempering configuration switches,[15]

where $i$ represents the liquid or the vapor phase. The parallel-tempering configuration switches were only performed between phases with neighboring state points.

Maximum translations, rotations, and volume changes were adjusted during the equilibration period of the simulation to yield a fraction of accepted MC trial moves of about 50%. Different maximum displacements were used for the liquid and the vapor phases. The usual cubic periodic-boundary conditions with the minimum-image convention were employed.[54] Initially, the statistical uncertainty of the ensemble averages was computed with the block-averaging method of Flyvbjerg and Petersen.[67] This analysis revealed the minimum block length needed to obtain uncorrelated blocks of samples. It turned out that blocks of a few thousands MC cycles were uncorrelated. The statistical errors given in this work have been computed from the application of the bootstrap method[52] to regularly saved blocks of 2500 MC cycles.

The initial pure-component coexistence points needed to start GDI were computed in the $NV\beta$ Gibbs ensemble. The saturation pressures of the initial points have been computed from volume perturbations.[68] Initial slopes of the Clapeyron
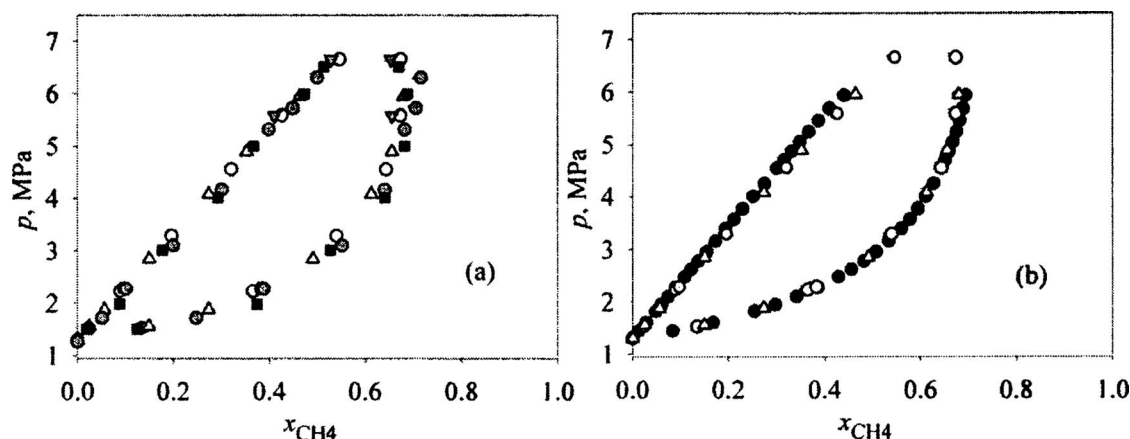


FIG. 2. $pxy$-data of the binary mixture ethane/methane at $T=250$ K. ○; experimental data published by Davalos *et al.*,[80] △; experimental data published by Wei *et al.*[81] (a) gray circles, ■, and gray triangles respectively represent simulation results published by Vrabec and Fischer,[60] by Liu and Beck,[61] and by Zhang and Duan.[62] (b) ●; our simulation results.
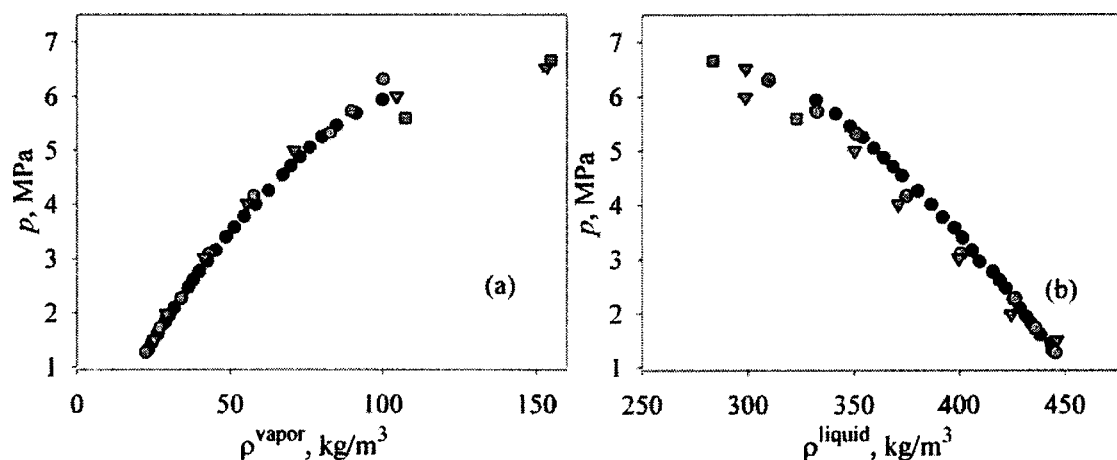
FIG. 3. Liquid and vapor densities at coexistence of the binary mixture ethane/methane at $T=250$ K. Gray triangles, gray circles, and gray squares respectively represent simulation results published by Liu and Beck,[61] by Vrabec and Fischer,[60] and by Zhang and Duan.[62] ●; our simulation results. (a) Vapor density. (b) Liquid density.

equation have been computed from the overlapping-distributions difference method (DM+OlD). A description of these methods is given in previous work.[15]

In all isobaric semigrand-canonical simulations, molecular translations and rotations were, respectively, force biased and torque biased.[14,15,69,70] The bias strengths for force-biased translations and torque-biased rotations were fixed at 0.5. Energy-biased trial identity changes with bias strengths $\gamma=-0.5$ and $\omega=0$ were employed.[15] Higher-order variable-step-size predictor-corrector equations were used.[15] The number of state points per integration step was fixed at three. Thus, the computation of one coexistence point required three liquid and three vapor-phase isobaric semigrand-ensemble simulations.

The bin widths $\Delta V$ for the volume dimensions of the liquid- and vapor-phase histograms were determined during the equilibration phase of the simulation at the very first integration step. The bin widths were given such values so as to result in approximately 40 volume bins. As a large number of histograms at different thermodynamic state points is combined, the fluctuating volume and composition may be very different in successive integration steps. This is no problem for the compositions, since the number of bins for

$N_2$ cannot exceed the total number of molecules. However, the change of volume can be more problematic. Consider, for example, the vapor-liquid equilibria of a binary mixture in which component two is supercritical [see Fig. 1(a)]. The numerical integration starts at pure component one. As the very first simulation takes place at relatively low pressures, the difference between the liquid- and vapor-phase molar volumes can be huge. Since the initial numbers of volume bins for the liquid- and the vapor phase are equal, the vapor-phase volume bin width is much larger than the volume bin width of the liquid phase. Now, the integration proceeds to higher fugacity fractions and higher pressures. The mixture approaches the vapor-liquid critical point, resulting in a decreasing difference between the molar volumes of the vapor and the liquid phase. Thus, the number of volume bins in the liquid phase grows. On the other hand, the number of volume bins in the vapor phase decreases.

In order to avoid an unbridled increase of liquid-phase volume bins or a low resolution of the volume distribution in the vapor phase, we used a logarithmic volume bin width $\Delta \ln V$. The lower (−) and upper (+) limits of the volumes that contribute to bin $i$ in phase $\alpha$ are computed from
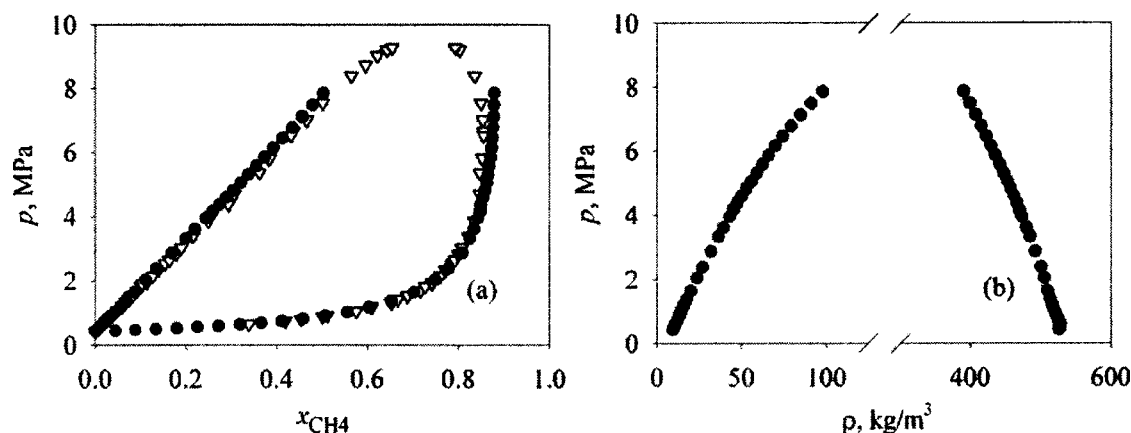


FIG. 4. Vapor-liquid equilibria of the binary mixture propane/methane at $T=270$ K. $\triangledown$, experimental data from Webster and Kidnay (Ref. 82), and ●; our simulation results computed with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. (a) $pxy$ data. (b) Liquid and vapor densities at coexistence.
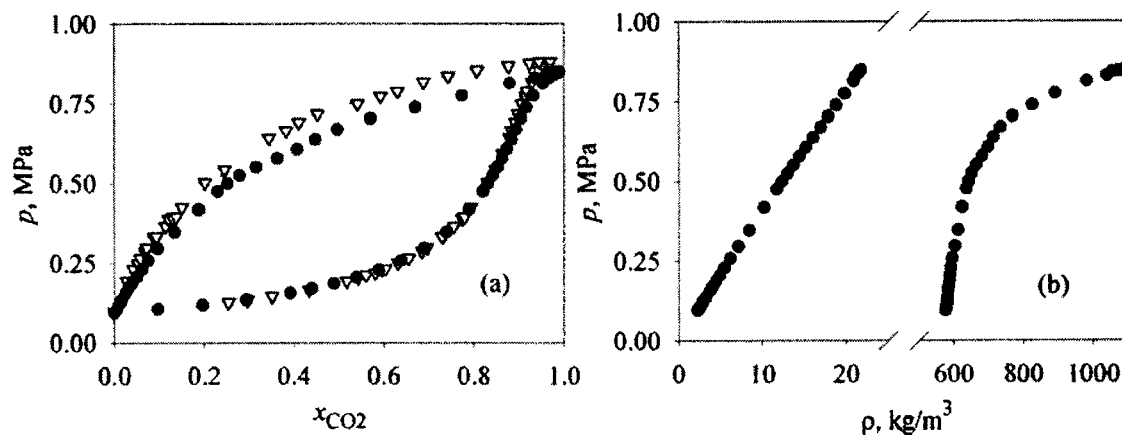
FIG. 5. Vapor-liquid equilibria of the binary mixture propane/carbon dioxide at $T=230$ K. $\nabla$, experimental data from Webster and Kidnay (Ref. 82) and $\bullet$; our simulation results computed with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. (a) $pxy$ data. (b) Liquid and vapor densities at coexistence.

$$V_i^{\alpha,-} = \exp\{\ln(V_{\text{ref}}^\alpha) + i\Delta \ln V^\alpha\},$$

$$V_i^{\alpha,+} = \exp\{\ln(V_{\text{ref}}^\alpha) + (i+1)\Delta \ln V^\alpha\}, \tag{37}$$

where $V_{\text{ref}}^\alpha$ is the reference volume of phase $\alpha$. The discretized volume that belongs to bin $i$ is given by

$$V_i^\alpha = \frac{V_i^{\alpha,-} + V_i^{\alpha,+}}{2}. \tag{38}$$

Figure 1(b) shows how the three thermodynamic state points at successive integration points are chosen. In case the slope of the Clapeyron equation is moderate, the fugacity fraction is incremented with an amount that makes sure that the composition dimensions of the histograms at state point $i$ sufficiently overlap. As already mentioned, the saturation pressure at the incremented fugacity fraction is estimated from the predictor equation. Two other state points that only differ in pressure are chosen, one pressure 10% lower and one pressure 10% higher than the predicted saturation pressure. As the pressure changes only slightly with the fugacity fraction, the volume dimensions of the histograms will have enough overlap.

From Fig. 1(b), it is clear that the slope can reach very large values in case component two is supercritical. This means that a tiny increment in the fugacity fraction results in a large change of the predicted saturation pressure. It is questionable whether the histograms collected at the predicted saturation pressure will have enough overlap with the volume dimensions of histograms collected at former thermodynamic state points. Therefore, it is convenient to increment the saturation pressure instead of the fugacity fraction. The fugacity fraction at the incremented saturation pressure is then estimated from a polynomial fit to former coexistence points. Two other state points are chosen that only differ in their fugacity fractions, one 3% lower and one 3% higher than the predicted fugacity. A lack of overlap between histograms collected at nearby state points is easily detected from an increasing number of MHR optimization iterations, or earlier during the simulation, from a low percentage of accepted parallel-tempering configuration switches. Like other authors do either implicitly[51,71,72] or explicitly,[73] the correlation between subsequent samples is neglected in the application of MHR to the collected histograms. This means that the $g_j$'s in Eq. (22) are put to one. The influence of this approxi-
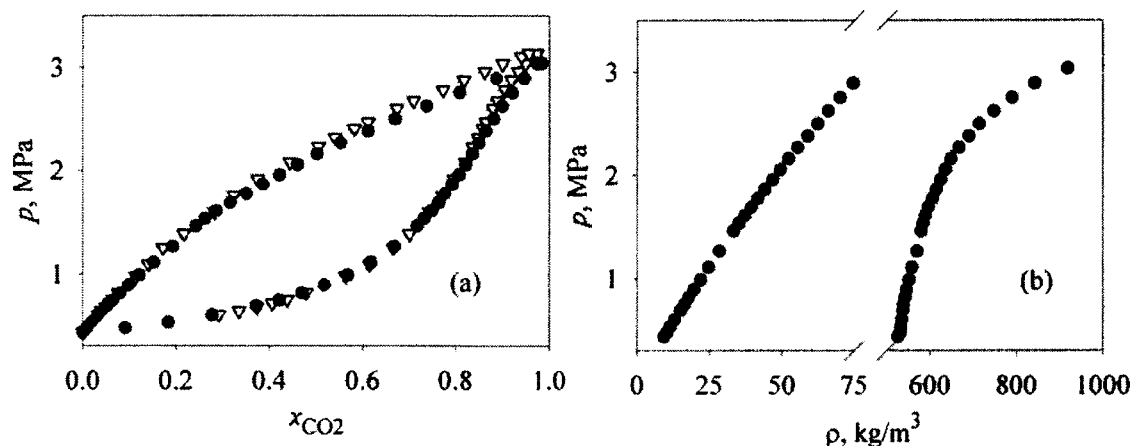


FIG. 6. Vapor-liquid equilibria of the binary mixture propane/carbon dioxide at $T=270$ K. $\nabla$, experimental data from Webster and Kidnay (Ref. 82) and $\bullet$; our simulation results computed with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. (a) $pxy$ data. (b) Liquid and vapor densities at coexistence.

mation has not been investigated but seems to be small.[37] Simulation data in Secs. III B–III E are presented in graphical form. The corresponding data have been deposited in tabular form as an electronic file at EPAPS.[74]

## B. Binary system methane/ethane

The advanced Gibbs-Duhem integration method was used to simulate the vapor-liquid equilibria of the binary system methane/ethane at $T=250$ K. Simulation results for the initial point (pure ethane) and the initial slope of the Clapeyron equation in the limit of pure ethane are given in Table I. The LB combining rules were used for the interactions between like molecules and the LHMcC combining rules for interactions between unlike molecules. The difference between the Hudson-McCoubrey and Berthelot combining rules is negligible for this binary mixture as ethane and methane have similar ionization potentials[75] and almost identical collision diameters.[66] Simulation results for the vapor-liquid equilibria of the binary mixture are given in Figs. 2 and 3. Figures 2 and 3 also include simulation results published by other authors. These results have been obtained with different force fields.

The simulation results presented in this section are much smoother than the simulation results published by other researchers. In general, they are also smoother than the results obtained from the conventional Gibbs-Duhem integration method. This is a result of the retroactive nature of MHR; histograms collected during a current simulation also improve the estimate of the degeneracy in the range that is important to previously computed coexistence points.

The simulation results obtained with advanced Gibbs-Duhem integration are in good agreement with experimental data. Unfortunately, there are no experimental density data available to compare our simulation results with. However, Vrabec and Fischer[60] compared their simulated liquid and vapor densities at phase coexistence with predictions from an equation of state (EOS). Their results were in good agreement with the EOS predictions, and the results presented in this section are in turn in good agreement with the results of Vrabec and Fischer. Thus, it is stated that the simple force fields of Vrabec and Fischer perform, at least at this temperature, equally well as the advanced TraPPE-EH force field. Nevertheless, the complexity of the TraPPE-EH force field is justified by the transferability of its parameters.

## C. Binary system propane/methane

Simulation results for the initial point (pure propane) and the initial slope of the Clapeyron equation in the limit of pure propane are given in Table I. The LB combining rules were used for the interactions between like molecules and the LH-McC combining rules for the interactions between unlike molecules. The difference between the Hudson-McCoubrey and the Berthelot combining rules is again negligible for this mixture as propane and methane have similar ionization potentials[75] and almost identical collision diameters.[66]

Simulation results for the vapor-liquid equilibria of the binary mixture are presented in graphical form in Fig. 4. The simulated vapor-liquid equilibria perfectly match the experimental data over a large range of fugacity fractions and pressures and are very smooth. The results in the vicinity of the critical point deviate from the experimental ones, probably due to finite-size effects. The simulation results for ethane/methane and for propane/methane show how well the advanced TraPPE-EH force field performs.

## D. Binary system propane/carbon dioxide

In this section, we present simulation results for the vapor-liquid equilibria of the binary system propane/carbon dioxide at $T=230$ K, $T=270$ K, and $T=294.26$ K. Vapor-liquid equilibria of this system have also been simulated by Moon and Moon with the $Np\beta$ Gibbs ensemble.[76] Potoff et al.[11] and Potoff and Siepmann[10] used grand-canonical simulations with multiple-histogram reweighting. In these studies, different force fields and combining rules were used. Potoff and Siepmann[10] used experimental vapor-liquid equilibria of this binary mixture at $T=294.26$ K to make the TraPPE-AA force field for carbon dioxide consistent with the TraPPE-EH force fields for $n$-alkanes. For this reason, it is not surprising that these force fields predict the vapor-liquid equilibria of this mixture, in particular, at $T=294.26$ K, very well.
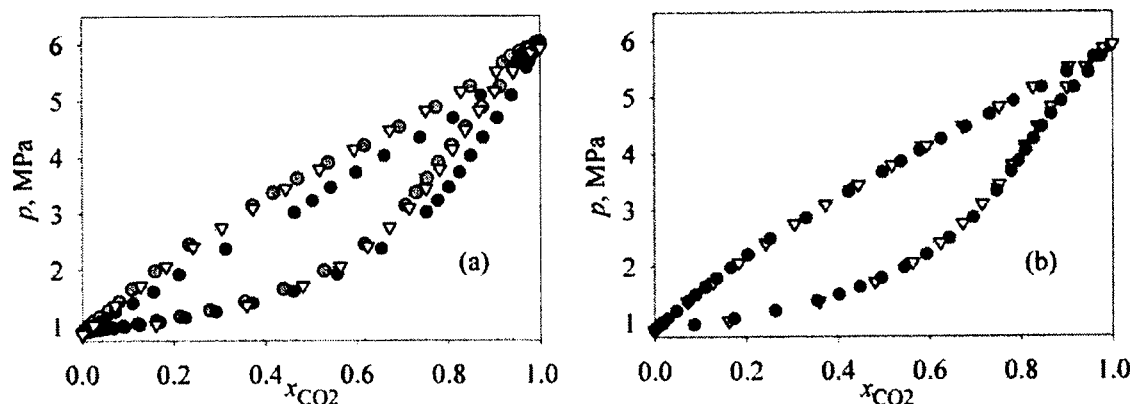


FIG. 7. $pxy$-data of the binary mixture propane/carbon dioxide at $T=294.26$ K. $\triangledown$; experimental vapor-liquid equilibria published by Reamer et al.[83] (a) $\bullet$; simulation results taken from Potoff et al.,[11] obtained with LB combining rules, gray circles; simulation results taken from Potoff et al.,[11] obtained with LB combining rules for interactions between like molecules and Kong combining rules for interactions between unlike molecules. (b) $\bullet$; our simulation results.
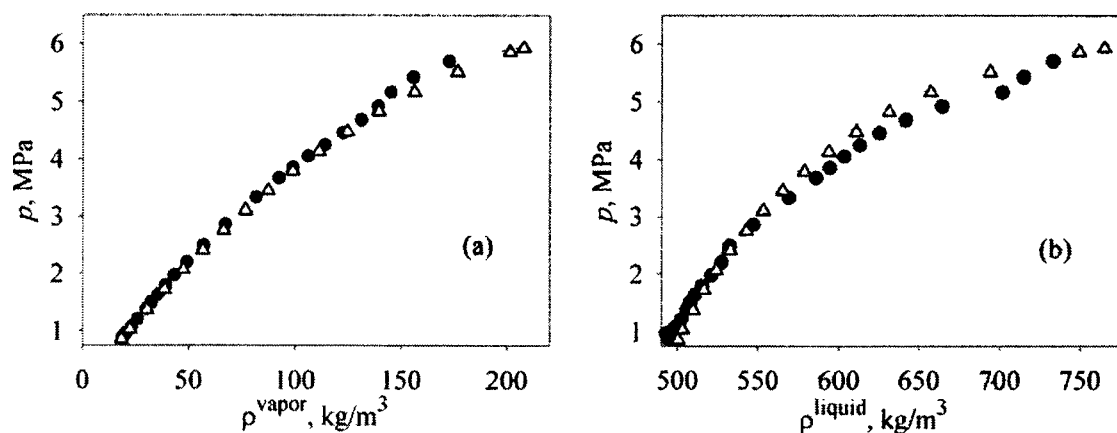
FIG. 8. Liquid and vapor densities at coexistence of the binary mixture propane/carbon dioxide at $T=294.26$ K. $\Delta$; experimental data published by Reamer *et al.* (Ref. 83) and $\bullet$; our simulation results obtained with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. (a) Vapor density. (b) Liquid density.

In contrast to the work of Potoff and Siepmann (all-atom TraPPE potential),[10] we employed the LB combining rules for interactions between like molecules and the LHMcC combining rules for interactions between unlike molecules. Simulation results for the initial point (pure propane) and the initial slope of the Clapeyron equation in the limit of pure propane are given in Table I. Simulation results for the vapor-liquid equilibria of the binary mixture are presented in graphical form in Figs. 5–8.

Figure 7 also includes simulation results published by Potoff *et al.*,[11] who performed two series of simulations: either the LB combining rules for all interactions or the LB combining rules for the interactions between like molecules and the Kong combining rules for the interactions between unlike molecules were used. They showed that the latter combination of combining rules considerably improved the agreement between experimental and simulated vapor-liquid equilibria of systems containing *n*-alkane/carbon dioxide mixtures. From Fig. 7, it is clear that the simulation results of Potoff *et al.*, performed with the LB combining rules for interactions between like molecules and the Kong combining rules for interactions between unlike molecules, agree very

well with the experimental data. Their results obtained with LB combining rules for all interactions are systematically low.

Potoff and Siepmann[10] showed that the simple LB combining rules can do a good job when predicting vapor-liquid equilibria of binary and ternary mixtures of *n*-alkanes, carbon dioxide, and nitrogen. In order to reach this result, pure-component force-field parameters of carbon dioxide were optimized to vapor-liquid equilibria of the binary mixture propane/carbon dioxide so as to find a good balance between the Lennard-Jones interactions and the Coulombic interactions. Simulation results published by Potoff and Siepmann,[10] at $T=294.26$ K, were not available in tabular form. However, from their simulation results in graphical form, it was obvious that the liquid- and vapor-phase compositions were only slightly too high.[10] The simulation results at $T=294.26$ K, as presented in this section, show even better agreement with experimental data than the results of Potoff *et al.*[11] and Potoff and Siepmann do.[10] From Fig. 8, it is seen that the vapor and liquid densities at coexistence are also in good agreement with experimental values. Unfortunately, the saturation pressures at $T=230$ K and $T=270$ K
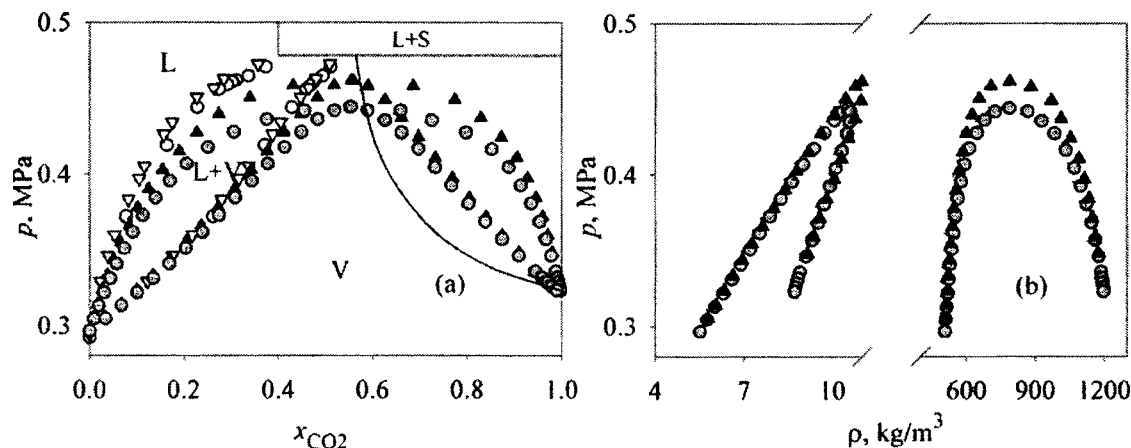


FIG. 9. Vapor-liquid equilibria of the binary mixture ethane/carbon dioxide at $T=207$ K. $\bigcirc$ and $\triangledown$ respectively represent experimental data published by Brown *et al.*[84] and by Wei *et al.*,[81] gray circles; our simulation results obtained with LB combining rules, $\blacktriangle$; our simulation results computed with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. (a) *pxy* data. Solid lines indicate the expected phase behavior.[79] (b) Liquid and vapor densities at coexistence.
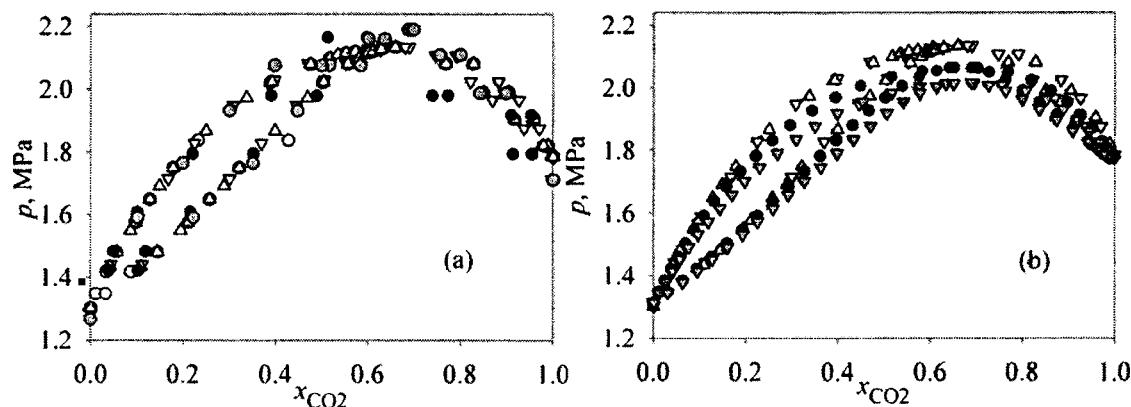
FIG. 10. *pxy*-data of the binary mixture ethane/carbon dioxide at $T=250$ K. $\bigcirc$, $\triangle$, and $\triangledown$ respectively represent experimental data published by Wei *et al.*,[81] by Brown *et al.*,[84] and by Davalos *et al.*[80] (a) $\bullet$ and gray circles are simulation results published by Liu and Beck[61] and by Vrabec and Fischer.[60] (b) gray triangles; our simulation results obtained with LB combining rules, $\bullet$; our simulation results obtained with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules.

are systematically low at high mole fractions of carbon dioxide. As the predictions of the binary vapor-liquid equilibria are qualitatively correct, further optimization of the force-field parameters of carbon dioxide to low-temperature data will improve the predictive power of the all-atom TraPPE force field.

### E. Binary system ethane/carbon dioxide

The vapor-liquid equilibria of the binary mixture ethane/carbon dioxide have been simulated by a number of researchers with a variety of simulation techniques, force fields, and combining rules,[11,60,61,77] the reason of which probably is the occurrence of maximum-pressure azeotropy. Furthermore, the system exhibits critical azeotropy,[78] however, this phenomenon has not been studied with molecular simulation. Maximum-pressure azeotropy occurs when component one is a bad solvent for component two and vice versa. Although azeotropy is not a very exciting phenomenon from a thermodynamic point of view, it is a challenge to predict the right azeotropic pressure and composition with molecular simulation.

In this section, simulation results of vapor-liquid equilibria at $T=207$ K, $T=250$ K, and $T=263.15$ K are presented.

Every GDI series has been performed twice. The first series of simulations has been performed with the LB combining rules for both interactions between like molecules and interactions between unlike molecules, while in the second case the LB combining rules were used for the interactions between like molecules and the LHMcC combining rules for interactions between unlike molecules.

Simulation results for the initial point (pure ethane) and the initial slope of the Clapeyron equation in the limit of pure ethane are given in Table I. Simulation results for the binary vapor-liquid equilibria are presented in graphical from in Figs. 9–13. Figures 10 and 11 also include simulation results published by Liu and Beck[61] and by Vrabec and Fischer.[60] Figure 12 also contains simulation results published by Potoff *et al.*[11] Details of the simulation results of Potoff *et al.* were given in Sec. III D.

The simulation results at $T=207$ K, as presented in Fig. 9, show that, though the predictions qualitatively match experimental data, even the combination of the advanced TraPPE-AA force field for carbon dioxide and the TraPPE-EH force field for ethane does not predict the right azeotropic pressure. Nevertheless, the azeotropic composition is predicted fairly well.
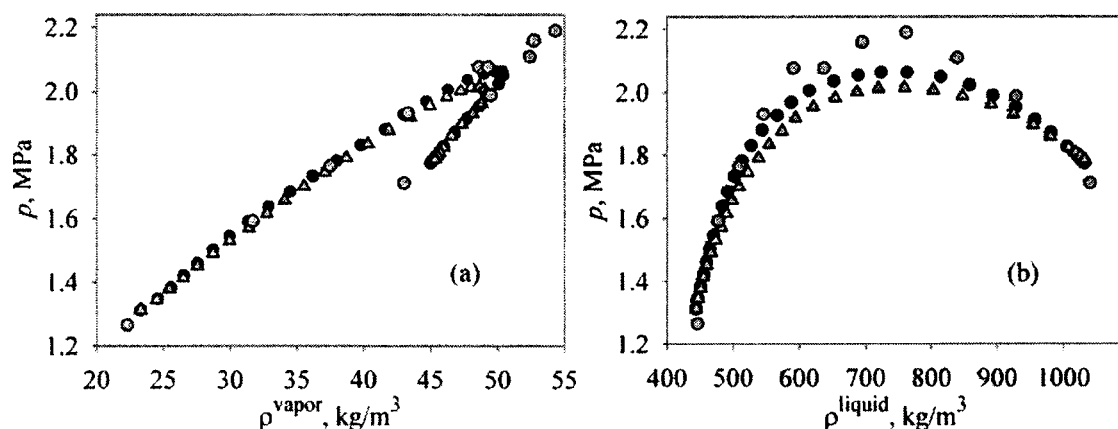


FIG. 11. Liquid and vapor densities at coexistence of the binary mixture ethane/carbon dioxide at $T=250$ K. Gray circles; simulation results published by Vrabec and Fischer,[60] gray triangles; our simulation results obtained with LB combining rules, $\bullet$; our simulation results obtained with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. (a) Vapor density. (b) Liquid density.
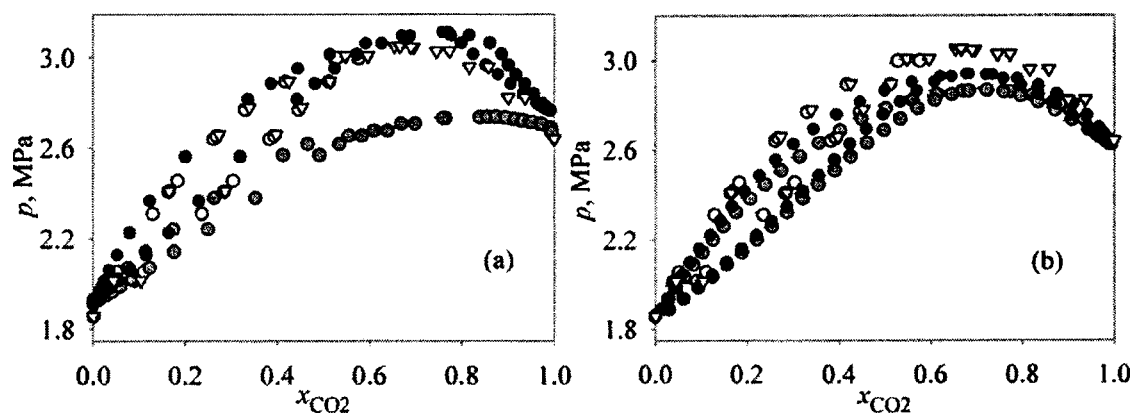
FIG. 12. *pxy*-data of the binary mixture ethane/carbon dioxide at $T=263.15$ K. $\circ$ and $\triangledown$ respectively represent experimental data published by Brown *et al.*[84] and by Fredenslund *et al.*[85] (a) Gray circles; simulation results taken from Potoff *et al.*,[11] computed with LB combining rules, $\bullet$; simulation results taken from Potoff *et al.*,[11] computed with LB combining rules for interactions between like molecules and Kong combining rules for interactions between unlike molecules. (b) Gray circles; our simulation results obtained with LB combining rules, $\bullet$; our simulation results obtained with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules.

It is known from experiment that pure carbon dioxide forms a stable solid phase at $T=207$ K.[79] However, the Gibbs-Duhem integration method traces the vapor-liquid coexistence line up to pure carbon dioxide. The expected phase behavior of the binary mixture at $T=207$ K, based on experimental points,[86] is indicated in Fig. 9. The uppermost vapor-liquid saturation pressures in Fig. 9 are close to liquid-vapor-solid (LVS)-three-phase equilibrium. The solid phase is expected to consist of pure carbon dioxide.[86] As the formation of a liquid-solid interface in a liquidlike simulation has a high physical and an artificial free-energy barrier, freezing of the liquid phase was not observed. It is also possible that the solid phase corresponding to the TraPPE-AA force field is not stable at the relevant physical conditions.

The simulation results obtained at $T=250$ K show the same trends as those at $T=207$ K. The results obtained with the advanced Gibbs-Duhem integration method are much smoother than those published by Vrabec and Fischer[60] and by Liu and Beck.[61] However, the simulation results taken from Vrabec and Fischer are closer to the experimental data.

Although their force fields are very simple, this is not surprising as Vrabec and Fischer used a binary interaction parameter that had been optimized to mixture data.

The simulation results obtained at $T=263.15$ K qualitatively agree with their experimental counterparts. Like the results obtained at $T=207$ K and $T=250$ K, the prediction at $T=263.15$ K is systematically lower than the experimental values. The simulation results published by Potoff *et al.*,[11] obtained with LB combining rules for interactions between like molecules and Kong combining rules for interactions between unlike molecules, have slightly higher saturation pressure than the experimental coexistence points. Their predictions with LB combining rules for all interactions have far too low saturation pressures. Notice that the predicted coexistence curves with and without Kong combining rules end up in rather different saturation pressures of pure carbon dioxide. This phenomenon is strange as the pure-component interactions for both series of simulations both depend on LB combining rules. It is concluded from Figs. 9, 10, and 12 that the use of the LHMcC combining rules, as opposed to the
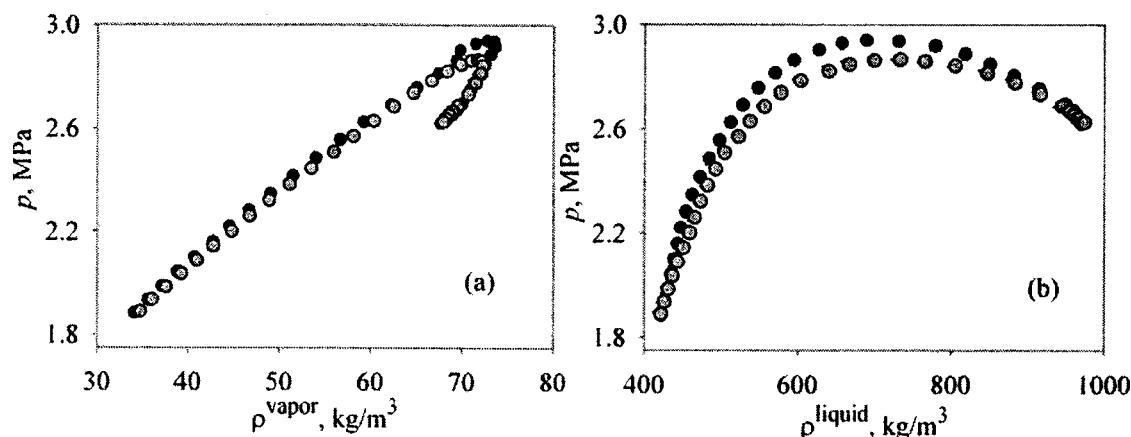


FIG. 13. Liquid and vapor densities at coexistence of the binary mixture ethane/carbon dioxide at $T=263.15$ K. Gray circles; our simulation results obtained with LB combining rules, $\bullet$; our simulation results obtained with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. (a) Vapor density. (b) Liquid density.

TABLE II. Propagation of statistical errors ($\sigma_p$) and influence of the absolute error in the initial point on the deviation ($\Delta p$) from the true coexistence curve.

| $\xi_{CH_4}$ (−) | $p$ (MPa) | $\Delta_p$ (MPa) | $\sigma_p$ (MPa) |
|---|---|---|---|
| 0.00 | 1.313 | 0.007 | 0.007 |
| 0.10 | 1.448 | 0.007 | 0.007 |
| 0.20 | 1.612 | 0.008 | 0.007 |
| 0.30 | 1.821 | 0.009 | 0.008 |
| 0.35 | 1.949 | 0.009 | 0.008 |
| 0.40 | 2.096 | 0.010 | 0.008 |
| 0.45 | 2.269 | 0.011 | 0.008 |
| 0.50 | 2.474 | 0.011 | 0.008 |
| 0.53 | 2.616 | 0.012 | 0.008 |
| 0.56 | 2.774 | 0.013 | 0.008 |
| 0.59 | 2.954 | 0.014 | 0.009 |
| 0.62 | 3.160 | 0.015 | 0.009 |
| 0.65 | 3.398 | 0.016 | 0.011 |
| 0.67 | 3.578 | 0.017 | 0.012 |
| 0.69 | 3.779 | 0.018 | 0.012 |
| 0.71 | 4.004 | 0.020 | 0.012 |
| 0.73 | 4.259 | 0.021 | 0.012 |
| 0.75 | 4.548 | 0.023 | 0.014 |
| 0.76 | 4.707 | 0.025 | 0.016 |
| 0.77 | 4.876 | 0.026 | 0.018 |
| 0.78 | 5.058 | 0.028 | 0.021 |
| 0.79 | 5.253 | 0.030 | 0.025 |
| 0.80 | 5.463 | 0.032 | 0.030 |
| 0.81 | 5.690 | 0.035 | 0.036 |

LB combining rules for interactions between unlike molecules, brings experimental data and simulation results much closer.

As pointed out in Sec. III D, Potoff and Siepmann[10] used experimental vapor-liquid equilibria of this binary mixture at $T=294.26$ K to make the TraPPE-AA force field for carbon dioxide consistent with the TraPPE-EH force fields for $n$-alkanes. Although the predictions for the binary mixture propane/carbon dioxide were rather quantitative, this is not the case for the binary mixture ethane/carbon dioxide. Thus, one can state that either optimization of the TraPPE-AA force field for carbon dioxide to a single binary alkane/carbon dioxide mixture is not enough or the LB combining rules are simply not good enough to capture the phase behavior of ethane/carbon dioxide mixtures.

### F. Error analysis

The propagation of the statistical uncertainties $\sigma_p$ of the predicted saturation pressures via the corrector equations in the binary mixture ethane/methane at $T=250$ K is given in Table II. It is the statistical uncertainty of the estimate of the initial saturation pressure that mainly determines the statistical error of the saturation pressures at successive coexistence points. Thus, in case the statistical error of the initial saturation pressure is small, the statistical error of the saturation pressure at successive integration steps will probably be small as well.

The influence of the error in the saturation pressure corresponding to the initial coexistence point on the deviation

from the true coexistence curve is investigated as follows. Take, for example, the standard deviation of the estimate of the initial saturation pressure, $\sigma_{p0}$, as the absolute deviation from the true saturation pressure. It is assumed that the ratio of volume differences [see Eq. (32)] at the true saturation pressure can be approximated by its counterpart obtained from ensemble averages at the estimated saturation pressure.

The influence of the error in the estimate of the initial saturation pressure on the deviation $\Delta_p$ from the true coexistence line in the binary mixture ethane/methane at $T=250$ K is also given in Table II. As methane is supercritical at $T=250$ K, the difference $\Delta^{LV}v$ changes significantly over the integration range. Thus, the error in the estimate of the initial saturation pressure is magnified when the vapor-liquid critical point of the mixture is approached. From a theoretical point of view, the numerical integration should have been started from the methane side as the difference $\Delta^{LV}v$ increases when integrating from the methane side to pure ethane. However, it is questionable whether the result would have been different when we had started the integration from the supercritical side, as the computation of the saturation pressure of a mixture close to the vapor-liquid critical point involves a high statistical uncertainty or fails completely.

Clearly, the errors that arise during numerical integration can be substantial. Therefore, it is recommended to directly check the results of Gibbs-Duhem integration against the criterion of phase coexistence on a regular basis to be sure that the difference between the simulated and true coexistence curves remains small.

### IV. CONCLUSIONS

The conventional GDI method inefficiently exploits the molecular configurations that are generated during the simulations. Furthermore, the number of simulations in the correction step is not known from the outset of the simulation. In this work, an advanced GDI method has been introduced which overcomes these problems.

Combination of GDI and multiple-histogram reweighting enables one to replace the unknown number of simulations of the prediction, correction, and production steps by a fixed and predetermined number of simulations. The advanced method has a retroactive nature; a current simulation also improves the predictions of previously computed coexistence points, resulting in smooth simulation results.

The application of the advanced GDI method depends on the ability to compute the degeneracy and the semigrand-canonical probability distribution from histograms collected at different state points. The relevant relations for the degeneracy and semigrand-canonical probability distribution were derived. Application of the corrector equations, along with the semigrand-canonical probability distributions of the liquid and the vapor phase, enables the computation of a whole coexistence line from a limited number of simulations. Furthermore, it is possible to search for the coexistence point belonging to a particular liquid or vapor-phase mole fraction or density.

The advanced GDI has been used to simulate the vapor-liquid equilibria of a number of binary mixtures: ethane/

methane, propane/methane, ethane/carbon dioxide, and propane/carbon dioxide. Exploiting Lorentz/Hudson-McCoubrey combining rules for interactions between unlike molecules, as opposed to Lorentz/Berthelot combining rules for all interactions, considerably improved the agreement between experimental and simulated vapor-liquid equilibria. Nevertheless, the use of different sets of combining rules for interactions between like and unlike molecules is inconsistent. It deserves recommendation to optimize force-field parameters with a combining rule other than Berthelot. Although we did not present simulation results computed with the conventional GDI method, the advanced GDI method generally performed much faster than the conventional method did for approximately similar statistical uncertainties of the results. The gain in simulation speed depends on the system under study.

Although the all-atom TraPPE force field had been optimized for interactions between carbon dioxide and *n*-alkanes, it did not reproduce the experimental saturation pressures of the binary system ethane/carbon dioxide, and predictions of the binary mixture propane/carbon dioxide were systematically low at the carbon dioxide side. Thus, further optimization to other *n*-alkane/carbon dioxide systems is needed.

The errors that arise during numerical integration of the Clapeyron equation can be substantial. In order to weaken the influence of the error in the initial point, it is recommended to directly check the Gibbs-Duhem integration results against the conditions of phase coexistence on a regular basis.

[1] C. J. King, *Separation Processes*, Chemical Engineering Series, 2nd ed. (McGraw-Hill, New York, 1980).

[2] R. K. Sinnott, *Chemical Engineering Design*, Vol. 6, 2nd ed. (Pergamon, Oxford, Great Britain, 1993).

[3] J. R. Backhurst and J. H. Harker, *Particle Technology and Separation Processes*, Vol. 2, 4th ed. (Pergamon, Oxford, Great Britain, 1991).

[4] J. Vrabec and J. Fischer, AIChE J. **43**, 212 (1997).

[5] G. Kronome, I. Szalai, M. Wendland, and J. Fischer, J. Mol. Liq. **85**, 237 (2000).

[6] J. Vrabec, A. Lofti, and J. Fischer, Fluid Phase Equilib. **112**, 173 (1995).

[7] J. Vrabec and J. Fischer, Mol. Phys. **85**, 781 (1995).

[8] A. Z. Panagiotopoulos, N. Quirke, M. R. Stapleton, and D. J. Tildesley, Mol. Phys. **63**, 527 (1988).

[9] A. Z. Panagiotopoulos, Int. J. Thermophys. **10**, 447 (1989).

[10] J. J. Potoff and J. I. Siepmann, AIChE J. **47**, 1676 (2001).

[11] J. J. Potoff, J. R. Errington, and A. Z. Panagiotopoulos, Mol. Phys. **97**, 1073 (1999).

[12] D. A. Kofke, J. Chem. Phys. **98**, 4149 (1993).

[13] D. A. Kofke, Mol. Phys. **78**, 1331 (1993).

[14] A. van 't Hof, S. W. de Leeuw, C. K. Hall, and C. J. Peters, Mol. Phys. **102**, 301 (2004).

[15] A. van 't Hof, Ph.D. thesis, TU Delft, Delft, The Netherlands, 2005.

[16] A. Z. Panagiotopoulos, Mol. Phys. **61**, 813 (1987).

[17] A. Z. Panagiotopoulos, Mol. Simul. **9**, 1 (1992).

[18] M. R. Hitchcock and C. K. Hall, J. Chem. Phys. **110**, 11433 (1999).

[19] R. Agrawal, Mol. Phys. **85**, 43 (1995).

[20] R. Agrawal and D. A. Kofke, Mol. Phys. **85**, 23 (1995).

[21] M. R. Lamm and C. K. Hall, Fluid Phase Equilib. **194–197**, 197 (2002).

[22] M. R. Lamm and C. K. Hall, Fluid Phase Equilib. **182**, 37 (2001).

[23] A. Mori, B. B. Laird, Y. Kangawa, T. Ito, and A. Koukitu, Mater. Phys. Mech. **6**, 49 (2003).

[24] M. Mehta and D. A. Kofke, Chem. Eng. Sci. **49**, 2633 (1994).

[25] R. Agrawal, M. Mehta, and D. A. Kofke, Int. J. Thermophys. **15**, 1073 (1994).

[26] D. A. Kofke and E. D. Glandt, Mol. Phys. **64**, 1105 (1988).

[27] S. S. Leung and R. B. Griffiths, Phys. Rev. A **8**, 2670 (1973).

[28] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).

[29] A. M. Ferrenberg and R. H. Swendsen, Comput. Phys., **3**, 101 (1989).

[30] J. P. Valleau and D. N. Card, J. Chem. Phys. **57**, 5457 (1972).

[31] I. R. McDonald and K. Singer, Discuss. Faraday Soc. **43**, 40 (1967).

[32] D. A. Chesnut and Z. W. Salsburg, J. Chem. Phys. **38**, 2861 (1963).

[33] Z. W. Salsburg, J. D. Jacobson, W. Fickett, and W. W. Wood, J. Chem. Phys. **30**, 65 (1959).

[34] A. M. Ferrenberg and D. P. Landau, Phys. Rev. B **44**, 5081 (1991).

[35] A. M. Ferrenberg, D. P. Landau, and R. H. Swendsen, Phys. Rev. E **51**, 5092 (1995).

[36] M. E. J. Newman and R. G. Palmer, J. Stat. Phys. **97**, 1011 (1999).

[37] R. H. Swendsen, Physica A **194**, 53 (1993).

[38] I. R. McDonald, Mol. Phys. **23**, 41 (1972).

[39] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).

[40] D. M. Ferguson and D. G. Garrett, in *Monte Carlo Methods in Chemistry*, Advances in Chemical Physics Vol. 105, edited by D. M. Ferguson, J. I. Siepmann, and D. G. Truhlar (Wiley, New York, 1999).

[41] D. S. Corti and G. Soto-Campos, J. Chem. Phys. **108**, 7959 (1998).

[42] D. S. Corti, Phys. Rev. E **64**, 016128 (2001).

[43] D. S. Corti, Mol. Phys. **100**, 1887 (2002).

[44] C. Borgs and S. Kappler, Phys. Lett. A **171**, 37 (1992).

[45] N. B. Wilding, Am. J. Phys. **69**, 1147 (2001).

[46] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).

[47] B. A. Berg, Comput. Phys. Commun. **153**, 397 (2003).

[48] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).

[49] L. D. Gelb and K. E. Gubbins, Physica A **244**, 112 (1997).

[50] F. A. Escobedo, J. Chem. Phys. **110**, 11999 (1999).

[51] J. J. Potoff and A. Z. Panagiotopoulos, J. Chem. Phys. **109**, 10914 (1998).

[52] B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1982).

[53] D. A. Kofke, in *Monte Carlo Methods in Chemistry*, Advances in Chemical Physics Vol. 105, edited by D. M. Ferguson, J. I. Siepmann, and D. G. Truhlar (Wiley, New York, 1999).

[54] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).

[55] G. C. Maitland, M. Rigby, E. B. Smith, and W. A. Wakeham, *Intermolecular Forces, Their Origin and Determination*, International Series of Monographs on Chemistry Vol. 3 (Oxford University Press, New York, 1987).

[56] G. H. Hudson and J. C. McCoubrey, Trans. Faraday Soc. **56**, 761 (1960).

[57] A. J. Stone, *The Theory of Intermolecular Forces*, International Series of Monographs on Chemistry Vol. 32 (Oxford University Press, New York, 2000).

[58] W. Song, P. J. Rossky, and M. Maroncelli, J. Chem. Phys. **119**, 9145 (2003).

[59] D. Möller, J. Óprzynski, A. Müller, and J. Fischer, Mol. Phys. **75**, 363 (1992).

[60] J. Vrabec and J. Fischer, Int. J. Thermophys. **17**, 889 (1996).

[61] A. Liu and Th. L. Beck, J. Phys. Chem. B **102**, 7627 (1998).

[62] Z. Zhang and Z. Duan, Geochim. Cosmochim. Acta **66**, 3431 (2002).

[63] M. Diaz Peña, C. Pando, and J. A. R. Renuncio, J. Chem. Phys. **76**, 325 (1982).

[64] M. Diaz Peña, C. Pando, and J. A. R. Renuncio, J. Chem. Phys. **76**, 333 (1982).

[65] A. Khalaf Al-Matar and D. A. Rockstraw, J. Comput. Chem. **25**, 660 (2004).

[66] B. Chen and J. I. Siepmann, J. Phys. Chem. B **103**, 5370 (1999).

[67] H. Flyvbjerg and H. G. Petersen, J. Chem. Phys. **91**, 461 (1989).

[68] V. I. Harismiadis, J. Vorholz, and A. Z. Panagiotopoulos, J. Chem. Phys. **105**, 8469 (1996).

[69] C. Pangali, M. Rao, and B. J. Berne, Chem. Phys. Lett. **55**, 413 (1978).

[70] M. Rao, C. Pangali, and B. J. Berne, Mol. Phys. **37**, 1773 (1979).

[71] P. B. Conrad and J. J. de Pablo, Fluid Phase Equilib. **150–151**, 51 (1998).

[72] J. Chang and S. I. Sandler, J. Chem. Phys. **118**, 8390 (2003).

[73] W. Shi and J. K. Johnson, Fluid Phase Equilib. **187–188**, 171 (2001).

[74] See EPAPS Document No. E-JCPSA6-123-509545 for the computed coexistence densities and compositions in tabulated form for the vapor-liquid equilibria discussed in this paper. This document can be reached via a direct link in the online article's HTML reference section or via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html).

[75] D. R. Lide, *CRC Handbook of Chemistry and Physics*, 73rd ed. (CRC, Boca Raton, FL, 1992).

[76] S.-D. Moon and B. K. Moon, Bull. Korean Chem. Soc. **21**, 1133 (2000).

[77] J. Stoll, J. Vrabec, and H. Hasse, AIChE J. **49**, 2187 (2003).

[78] J. S. Rowlinson, *Liquids and Liquid Mixtures*, 2nd ed. (Butterworths, London, 1969).

[79] S. Angus, B. Armstrong, and K. M. Reuck, *IUPAC, International Thermodynamic Tables of the Fluid State Carbon Dioxide* (Pergamon, Oxford, 1976), Vol. 3.

[80] J. Davalos, W. R. Anderson, R. E. Phelps, and A. J. Kidnay, J. Chem. Eng. Data **21**, 81 (1976).

[81] M. S. W. Wei, T. S. Brown, A. J. Kidnay, and D. E. Sloan, J. Chem. Eng. Data **40**, 726 (1995).

[82] L. A. Webster and A. J. Kidnay, J. Chem. Eng. Data **46**, 759 (2001).

[83] H. H. Reamer, B. H. Sage, and W. N. Lacey, Ind. Eng. Chem. **43**, 2515 (1951).

[84] T. S. Brown, A. J. Kidnay, and D. E. Sloan, Fluid Phase Equilib. **40**, 169 (1988).

[85] A. Fredenslund and J. Mollerup, J. Chem. Soc., Faraday Trans. 1 **70**, 1653 (1974).

[86] R. H. Jensen and F. Kurata, AIChE J. **17**, 357 (1971).