

# Coexistence Calculation Using the Isothermal-Isochoric Integration Method

S. Mostafa Razavi

*Department of Chemical and Biomolecular Engineering, The University of Akron, Akron, Ohio 44325, USA*

Richard A. Messerly

*Thermodynamics Research Center, National Institute of Standards and Technology, Boulder, Colorado 80305, USA*

J. Richard Elliott\*

*Department of Chemical and Biomolecular Engineering, The University of Akron, Akron, Ohio 44325, USA*

---

## Abstract

In this work, the isothermal-isochoric integration (ITIC) method is demonstrated as a viable method for vapor-liquid coexistence calculation by molecular simulation. Several tests are carried out to validate the method. The first group of tests utilizes self-consistent NIST REFPROP values to demonstrate that, in the absence of simulation uncertainties, the ITIC method yields coexistence values with less than 1 % deviation for reduced temperatures of less than 0.85. The impact of various simulation specifications are then compared. Following our recommended simulation methodology, consistent results are achieved between the ITIC method, Gibbs Ensemble Monte Carlo (GEMC) method, and Grand Canonical Monte Carlo (GCMC) method for reduced temperatures of 0.6-0.85. The ITIC method proves to be much more effective compared to GEMC and GCMC methods for vapor-liquid coexistence calculations at reduced temperatures of 0.45-0.6, which are important for practical applications. It is shown that computational efficiency is often served best by applying the ITIC method for the entire temperature range rather than applying Monte Carlo (MC) methods for part of the range. Furthermore, the ITIC method lends itself to application with molecular dynamics (MD) as well as MC, advancing the prospect of simulation results that are quantitatively consistent across software platforms.

**Keywords:** Vapor Pressure, Vapor Liquid Equilibria, Phase Diagram, Liquid Density

---

## 1. Introduction

Phase coexistence determination is important when characterizing the physical properties of a chemical compound. Both the vapor pressure ( $P^{\text{sat}}$ ) and saturation liquid and vapor density ( $\rho_{\text{liq}}$  and  $\rho_{\text{vap}}$ ) provide sensitive measures of the quality provided by a particular force field. In principle, the computation of phase coexistence is a simple matter of equating pressures, temperatures, and chemical potentials between the coexisting phases. Nevertheless, accurate computation of phase coexistence by molecular simulation has posed challenges over the years.

The most straightforward method to compute phase transition in molecular simulation is to simply define an  $NVT$  system (constant number of molecules, volume, and temperature) of sufficient size and overall density that an explicit interface is encountered. However this method often results in imprecise results [1]. First order phase transitions exhibit a considerable free energy barrier between two phases due to interfacial free energies. For systems with large interfaces, this energy barrier

increases. This often results in hysteresis, and phase transformation irreversibly proceeds beyond the coexistence point. [1]

Alternatively, there are methods for calculating phase coexistence while avoiding explicit interfaces. The Gibbs Ensemble Monte Carlo (GEMC) method [2] is one of the most popular phase coexistence determination methods [3]. GEMC requires particle exchange between two phases which leads to its major drawback, i.e. insertion of particles in dense phases for large molecules. Histogram reweighting Monte Carlo and Transition-matrix Monte Carlo methods are two closely related methods used for calculation of phase equilibrium based on Grand Canonical Monte Carlo (GCMC) simulations. These methods also involve the problematic insertion/deletion moves. This problem is especially exacerbated at low temperatures and for large and branched molecules. The lowest temperatures that are available in the literature rarely extend below a reduced temperature ( $T_r = T/T_c$ , where  $T_c$  is the critical temperature) of 0.6 [4, 5]. However, common methods for industrial applications treat the temperature range from triple point (which may be as low as  $T_r = 0.3$ ) to the critical point. The Peng-Robinson equation of state, for example, is valid for reduced temperatures as low as 0.45 [6]. To provide fundamental physical models that address issues with industrial applications, molecular simulations must address the entire temperature range of interest.

---

\*Corresponding author

Email addresses: sr87@uakron.edu (S. Mostafa Razavi), richard.messerly@nist.gov (Richard A. Messerly), elliot1@uakron.edu (J. Richard Elliott)

As another alternative, Kofke [7] developed a method called Gibbs-Duhem integration which makes use of the Clapeyron equation to numerically integrate and proceed along the saturation line starting from one single coexistence point. This method has been extensively used to extend the lower temperature limit of coexistence calculations much below the normal boiling point [8] [9] [10] [11] [12]. The Gibbs-Duhem method can solve the insertion problem, but it relies on a second method to obtain the initial coexistence point. Ahunbay et al. [8], for example, have applied this approach in conjunction with GEMC to obtain the initial coexistence point, *NPT* simulations (constant number of molecules, pressure, and temperature) to estimate saturation liquid densities, and parallel tempering method [13] to increase the efficiency of low temperature simulations. Their implementation is tested at reduced temperatures above 0.45 for several compounds. Note that in Gibbs-Duhem method, vapor density can also be obtained from the ideal gas equation or through separate *NPT* simulations at the computed saturation pressure and temperature.

As one more alternative, thermodynamic integration (TI) can provide a reliable solution for free energy calculation. In this method, a series of simulations are performed along a path that connects the state of interest to a system for which the free energy is known. One should be careful that the path does not include any type of phase change [9]. Elliott et al. [14] applied an isochoric integration method to calculate vapor-liquid equilibria (VLE) of square well spheres, demonstrating deficiencies in the preceding MC results. For square well spheres, a convenient starting isotherm is the hard sphere limit, for which the thermodynamics are well represented by the Carnahan-Starling equation [15]. In this work, however, results are sought for soft potential models of arbitrary molecular shape, for which the infinite temperature limit is not convenient.

In the proposed method, *NVT* simulation points are used across an isotherm and along several isochores, hence the name isothermal-isochoric (ITIC) [16]. Because the approach to low temperatures proceeds along an isochore initialized at the high temperature (supercritical) isotherm, the insertion move problem is alleviated. In principle, there is no low temperature limit for the applicability of this method, except the triple point.

Also, the data generated along the paths of integration are valuable on their own merits, providing distinct insights about how well the molecular model is performing under conditions of high temperature and pressure. In other words, ITIC provides greater quality of characterization than other vapor-liquid coexistence calculations at low saturation temperatures and non-saturation temperatures. In addition, with combination of derivative properties from the work of Lustig et al [17], it should be possible to utilize these data in order to generate high accuracy multi-parameter equations of state [18, 19, 20, 21, 22, 23, 24, 25].

As another advantage, vapor pressure calculation using molecular dynamics (MD) is often perceived as an impractical approach [26]. In this work, we show the possibility of using molecular dynamics in the context of ITIC integration as a viable tool for VLE calculation. The ITIC method provides a cross-platform solution to the problem of VLE calculation with-

out a significant increase in computational effort compared to the typical GEMC or GCMC approaches [16, 27].

The presentation is initiated in Section 2 with a review of the thermodynamics underlying the integration method and some practical implementation aspects. The approach is validated in Section 3 by applying it to coexistence data available from the National Institute of Standards and Technology (NIST). Section 4 describes the details of simulations that are found in Sections 5-7. Section 5 provides recommendations for the most convenient and effective methods for characterizing the virial coefficients by molecular simulation. Section 6 investigates the impact of system size, fixed or flexible bonds, and Monte Carlo or molecular dynamics when simulating at low densities. Section 7 demonstrates applications to molecular simulations and comparison between the ITIC method and other methods of phase coexistence calculation. Finally, Section 8 summarizes the conclusions from this work.

## 2. The Isothermal-Isochoric Integration Method

### 2.1. Derivation of the vapor pressure and density formulas

In this section, the dimensionless departure functions defined below are used to represent different forms of energy

$$\begin{aligned} A^{\text{dep}}(T, \rho) &= \frac{A(T, \rho) - A_{\text{ig}}(T, \rho)}{RT} \\ U^{\text{dep}}(T, \rho) &= \frac{U(T, \rho) - U_{\text{ig}}(T, \rho)}{RT} \\ H^{\text{dep}}(T, \rho) &= \frac{H(T, \rho) - H_{\text{ig}}(T, \rho)}{RT} \end{aligned} \quad (1)$$

where  $T$ ,  $\rho$ ,  $P$ , and  $R$  are temperature, density, pressure, and the gas constant (8.3144598 J/(mol K)), respectively.  $A^{\text{dep}}$ ,  $U^{\text{dep}}$ , and  $H^{\text{dep}}$  represent Helmholtz energy, internal energy, and enthalpy departure functions, respectively. The subscript “ig” denotes the ideal gas.

For a single component system, Eq. (2) must be satisfied at vapor-liquid phase equilibrium

$$\begin{aligned} T_{\text{liq}} &= T_{\text{vap}} \\ P_{\text{liq}} &= P_{\text{vap}} \\ G_{\text{liq}} &= G_{\text{vap}} \end{aligned} \quad (2)$$

where  $G$  represents Gibbs energy and subscripts “vap” and “liq” denote the saturated vapor and liquid phases, respectively. The Gibbs energy criterion of phase equilibrium can be written in terms of Helmholtz energy

$$\frac{A_{\text{liq}}}{RT_{\text{liq}}} + \frac{P_{\text{liq}} V_{\text{liq}}}{RT_{\text{liq}}} = \frac{A_{\text{vap}}}{RT_{\text{vap}}} + \frac{P_{\text{vap}} V_{\text{vap}}}{RT_{\text{vap}}} \quad (3)$$

where  $V$  represents volume. This equation which is derived based on the fundamental relationship between Gibbs and Helmholtz energies ( $G \equiv A + PV$ ) can be simplified as

$$\frac{A_{\text{liq}}}{RT^{\text{sat}}} - \frac{A_{\text{vap}}}{RT^{\text{sat}}} = Z_{\text{vap}} - Z_{\text{liq}} \quad (4)$$

where  $T^{\text{sat}}$  and  $Z$  represent saturation temperature and compressibility factor, respectively, and  $A_{\text{liq}} = A(T^{\text{sat}}, \rho_{\text{liq}})$ ,  $A_{\text{vap}} = A(T^{\text{sat}}, \rho_{\text{vap}})$ . Calculating the vapor-liquid phase equilibrium requires solving Eq (4).

Determining the left-hand side of Eq (4) involves connecting the liquid phase to vapor phase, however this pathway cannot pass through the two-phase region. Since  $A$  is a state function, in order to avoid the two-phase region the following path is taken: 1) The temperature of the liquid is increased to a supercritical temperature ( $T^{\text{IT}}$ ). 2) The density of the supercritical fluid ( $\rho_{\text{liq}}$ ) is decreased to zero. 3) The temperature of the resulting supercritical fluid is decreased back to saturation temperature. 4) The density is increased from zero to the density of the vapor phase ( $\rho_{\text{vap}}$ ).

$$(T^{\text{sat}}, \rho_{\text{liq}}) \xrightarrow{1} (T^{\text{IT}}, \rho_{\text{liq}}) \xrightarrow{2} (T^{\text{IT}}, 0) \xrightarrow{3} (T^{\text{sat}}, 0) \xrightarrow{4} (T^{\text{sat}}, \rho_{\text{vap}})$$

Hence, the alternative pathway includes two isothermal and two isochoric paths. In the ITIC method, several separate simulations in the canonical ensemble (NVT) are needed to construct the first three paths (see Section 2.3), and the fourth path is characterized by virial coefficient expansion (see Section 5).

The Helmholtz energy change along an isochoric path is calculated by using Eq. (5) [28].

$$\frac{A(T_2, \rho)}{RT_2} - \frac{A(T_1, \rho)}{RT_1} = \int_{T_1}^{T_2} \frac{U(T, \rho)}{R} d\left(\frac{1}{T}\right) \quad (5)$$

where  $\rho$  represents the density of the isochore. The Helmholtz energy change along an isothermal path is calculated using

$$\frac{A(T, \rho_2)}{RT} - \frac{A(T, \rho_1)}{RT} = \int_{\rho_1}^{\rho_2} \frac{Z(T, \rho)}{\rho} d\rho \quad (6)$$

where  $T$  represents the isotherm temperature. Using Eq. (5) and Eq. (6), the Helmholtz energy change along the four paths mentioned above are derived as follows, respectively.

$$\frac{A(T^{\text{sat}}, \rho_{\text{liq}})}{RT^{\text{sat}}} - \frac{A(T^{\text{IT}}, \rho_{\text{liq}})}{RT^{\text{IT}}} = \int_{T^{\text{IT}}}^{T^{\text{sat}}} \frac{U(T, \rho_{\text{liq}})}{R} d\left(\frac{1}{T}\right) \quad (7)$$

$$\frac{A(T^{\text{IT}}, \rho_{\text{liq}})}{RT^{\text{IT}}} - \frac{A(T^{\text{IT}}, 0)}{RT^{\text{IT}}} = \int_0^{\rho_{\text{liq}}} \frac{Z(T^{\text{IT}}, \rho)}{\rho} d\rho \quad (8)$$

$$\frac{A(T^{\text{IT}}, 0)}{RT^{\text{IT}}} - \frac{A(T^{\text{sat}}, 0)}{RT^{\text{sat}}} = \int_{T^{\text{sat}}}^{T^{\text{IT}}} \frac{U(T, 0)}{R} d\left(\frac{1}{T}\right) \quad (9)$$

$$\frac{A(T^{\text{sat}}, 0)}{RT^{\text{sat}}} - \frac{A(T^{\text{sat}}, \rho_{\text{vap}})}{RT^{\text{sat}}} = \int_{\rho_{\text{vap}}}^0 \frac{Z(T^{\text{sat}}, \rho)}{\rho} d\rho \quad (10)$$

Adding Eq. (7-10) results in the following equation, noting that the left-hand side of Eq. (4) can be written as the summation of the left-hand sides of Eq. (7-10).

$$\int_{T^{\text{IT}}}^{T^{\text{sat}}} \frac{U(T, \rho_{\text{liq}})}{R} d\left(\frac{1}{T}\right) + \int_{T^{\text{sat}}}^{T^{\text{IT}}} \frac{U(T, 0)}{R} d\left(\frac{1}{T}\right) + \int_0^{\rho_{\text{liq}}} \frac{Z(T^{\text{IT}}, \rho)}{\rho} d\rho + \int_{\rho_{\text{vap}}}^0 \frac{Z(T^{\text{sat}}, \rho)}{\rho} d\rho = Z_{\text{vap}} - Z_{\text{liq}} \quad (11)$$

Since the integrands ( $\frac{Z}{\rho}$ ) in Eq. (8) and Eq. (10) are infinity at  $\rho = 0$ , using these two equations is not practical. A workaround for this problem is to obtain the ideal gas equivalents of Eq. (8) and Eq. (10) as follows

$$\frac{A_{\text{ig}}(T^{\text{IT}}, \rho_{\text{liq}})}{RT^{\text{IT}}} - \frac{A_{\text{ig}}(T^{\text{IT}}, 0)}{RT^{\text{IT}}} = \int_0^{\rho_{\text{liq}}} \frac{1}{\rho} d\rho \quad (12)$$

$$\frac{A_{\text{ig}}(T^{\text{sat}}, 0)}{RT^{\text{sat}}} - \frac{A_{\text{ig}}(T^{\text{sat}}, \rho_{\text{vap}})}{RT^{\text{sat}}} = \int_{\rho_{\text{vap}}}^0 \frac{1}{\rho} d\rho \quad (13)$$

noting that  $Z_{\text{ig}} = 1$ . Eq. (12-13) are then added to and subtracted from Eq. (11) to obtain

$$\int_{T^{\text{IT}}}^{T^{\text{sat}}} \frac{U(T, \rho_{\text{liq}}) - U_{\text{ig}}(T, \rho_{\text{liq}})}{R} d\left(\frac{1}{T}\right) + \int_0^{\rho_{\text{liq}}} \frac{Z(T^{\text{IT}}, \rho) - 1}{\rho} d\rho + \int_{\rho_{\text{vap}}}^0 \frac{Z(T^{\text{sat}}, \rho) - 1}{\rho} d\rho + \int_{\rho_{\text{vap}}}^{\rho_{\text{liq}}} \frac{1}{\rho} d\rho = Z_{\text{vap}} - Z_{\text{liq}} \quad (14)$$

where  $U_{\text{ig}}(T, \rho_{\text{liq}}) = U(T, 0)$ . The main advantage of this equation over Eq. (11) is that the integrands of isothermal integrals have the  $\frac{(Z-1)}{\rho}$  form with a finite value at  $\rho = 0$ . In addition, the last integral of Eq. (14) which appeared as a result of adding right-hand sides of Eq. (12-13) has non-zero limits.

The first two integrals of Eq. (14) represent the Helmholtz energy departure of the liquid phase ( $A_{\text{liq}}^{\text{dep}}$ )

$$A_{\text{liq}}^{\text{dep}} = \frac{A(T^{\text{sat}}, \rho_{\text{liq}})}{RT^{\text{sat}}} - \frac{A_{\text{ig}}(T^{\text{sat}}, \rho_{\text{liq}})}{RT^{\text{sat}}} = \int_{T^{\text{IT}}}^{T^{\text{sat}}} \frac{U - U_{\text{ig}}}{R} d\left(\frac{1}{T}\right) + \int_0^{\rho_{\text{liq}}} \frac{Z - 1}{\rho} d\rho \Big|_{T=T^{\text{IT}}} \quad (15)$$

that can be derived by adding Eq. (7-9) and subtracting Eq. (12) evaluated at  $T^{\text{sat}}$ , noting that  $A(T, 0) = A_{\text{ig}}(T, 0)$ . The above two integrals are calculated numerically by using NVT simulation data (see Section 2.3).

Since the density of the vapor phase is low at most conditions of interest, the value for  $Z_{\text{vap}}$  in Eq. (14) can be approximated using a virial expansion truncated at the  $B_3$  term

$$Z_{\text{vap}} = 1 + B_2 \rho_{\text{vap}} + B_3 \rho_{\text{vap}}^2 \quad (16)$$

which is substituted in the third integral of Eq. (14) to obtain

$$\int_{\rho_{\text{vap}}}^0 \frac{Z(T^{\text{sat}}, \rho) - 1}{\rho} d\rho = -B_2 \rho_{\text{vap}} - \frac{1}{2} B_3 \rho_{\text{vap}}^2 \quad (17)$$

which is equal to  $-A_{\text{vap}}^{\text{dep}}$ . Finally, substituting Eq. (15-17) in Eq. (14) results in

$$A_{\text{liq}}^{\text{dep}} + Z_{\text{liq}} - 1 + \ln\left(\frac{\rho_{\text{liq}}}{\rho_{\text{vap}}}\right) = 2B_2 \rho_{\text{vap}} + \frac{3}{2} B_3 \rho_{\text{vap}}^2 \quad (18)$$

where the  $\ln\left(\frac{\rho_{\text{liq}}}{\rho_{\text{vap}}}\right)$  term appeared by solving the last integral of Eq. (14). Rearranging this equation in terms of  $\rho_{\text{vap}}$  results in

$$\rho_{\text{vap}} = \rho_{\text{liq}} \exp\left(A_{\text{liq}}^{\text{dep}} + Z_{\text{liq}} - 1 - 2B_2 \rho_{\text{vap}} - \frac{3}{2} B_3 \rho_{\text{vap}}^2\right) \quad (19)$$

Since  $\rho_{\text{vap}}$  appears on both sides of Eq. (19), this equation can be solved iteratively, as explained in detail in Section 2.2. Finally, assuming  $\rho_{\text{vap}}$ ,  $T^{\text{sat}}$ , and virial coefficients at  $T^{\text{sat}}$  are known, vapor pressure ( $P^{\text{sat}}$ ) can be calculated using

$$P^{\text{sat}} = P_{\text{vap}} = Z_{\text{vap}} \rho_{\text{vap}} RT^{\text{sat}} = (1 + B_2 \rho_{\text{vap}} + B_3 \rho_{\text{vap}}^2) \rho_{\text{vap}} RT^{\text{sat}} \quad (20)$$

## 2.2. ITIC algorithm

Implementation of the ITIC procedure starts with determining the temperatures and densities of the ITIC points shown in Figure 1. In ITIC method,  $\rho_{\text{liq}}$  is imposed and  $T^{\text{sat}}$ ,  $P^{\text{sat}}$ , and  $\rho_{\text{vap}}$  are calculated. This involves estimating saturation temperatures ( $T^{\text{sat}}_{\text{est}}$ ), i.e.  $T$  at points 11, 13, 15, 17, and 19. Once ITIC state points are determined,  $NVT$  simulations are performed at the obtained conditions using either MD or MC, and the ensemble averages of  $Z$  and  $U^{\text{dep}}$  are calculated from simulation output.  $A^{\text{dep}}_{\text{liq}}$  is then calculated as explained in Section 2.3.

The first iteration starts by setting  $Z_{\text{liq}}$  to an initial value of zero (After the first iteration,  $Z_{\text{liq}}$  is updated with the value from previous iteration). Next,  $T^{\text{sat}}$  is calculated by inter- or extrapolating the isochoric simulation results, specifically, we fit  $Z$  vs reciprocal temperature to a second order polynomial and solve for  $T^{\text{sat}}$  such that  $Z(T^{\text{sat}}) = Z_{\text{liq}}$ . It should be noted that iterations of  $T^{\text{sat}}$  do not require further simulations because the isochoric integration can be performed with interpolated values of the compressibility factor, noting the smooth behavior of  $Z$  vs. reciprocal temperature.

Since the value of  $T^{\text{sat}}$  has changed,  $U^{\text{dep}}_{\text{liq}}$  and  $A^{\text{dep}}_{\text{liq}}$  should be recomputed.  $U^{\text{dep}}_{\text{liq}}$  is updated by inter- or extrapolating  $U^{\text{dep}}T$  vs  $1/T$  (i.e.  $(U - U^{\text{ig}})/R$  vs.  $1/T$  in Eq. 15) using a first order polynomial along the isochore, such that  $U^{\text{dep}}_{\text{liq,new}} = U^{\text{dep}}(T^{\text{sat}}_{\text{new}})$ . Because  $U^{\text{dep}}T$  vs.  $1/T$  has a linear shape in the vicinity of  $T^{\text{sat}}$ , trapezoid rule is used to update  $A^{\text{dep}}_{\text{liq}}$  according to the following equation

$$A^{\text{dep}}_{\text{liq,new}} = A^{\text{dep}}_{\text{liq,old}} + \left( \frac{1}{T^{\text{sat}}_{\text{new}}} - \frac{1}{T^{\text{sat}}_{\text{old}}} \right) \frac{U^{\text{dep}}_{\text{liq,new}} T^{\text{sat}}_{\text{new}} + U^{\text{dep}}_{\text{liq,old}} T^{\text{sat}}_{\text{old}}}{2} \quad (21)$$

where subscript ‘‘old’’ denotes the corresponding property at  $T^{\text{sat}}_{\text{old}}$  and subscript ‘‘new’’ denotes the property at updated  $T^{\text{sat}}_{\text{new}}$  (For the first iteration,  $T^{\text{sat}}_{\text{old}} = T^{\text{sat}}_{\text{est}}$ ).

At this point, all terms on the right-hand side of Eq. (19) are known except  $\rho_{\text{vap}}$ . For the first iteration,  $\rho_{\text{vap}}$  on the right-hand side of Eq. (19) is set to zero and the new vapor density ( $\rho^{\text{new}}_{\text{vap}}$ ) is calculated. After the first iteration,  $\rho_{\text{vap}}$  on the right-hand side is updated with  $\rho^{\text{old}}_{\text{vap}}$  from previous iteration.

Then,  $P^{\text{sat}}$  is computed by substituting the new  $\rho_{\text{vap}}$  in Eq. (20) and the new value of  $Z_{\text{liq}}$  is calculated using

$$Z_{\text{liq}} = \frac{P^{\text{sat}}}{\rho_{\text{liq}} R T^{\text{sat}}} \quad (22)$$

where  $\rho_{\text{liq}}$  is equal to the isochoric density.

Having new value for  $Z_{\text{liq}}$ , the next iteration starts by calculating the new  $T^{\text{sat}}$ . The iterations are stopped when the difference in two consecutive  $\rho_{\text{vap}}$  values are less than a defined tolerance (We use 0.1 % deviation as the stopping criterion. Using smaller tolerances does not significantly improve the accuracy.) This procedure is repeated for other isochores. Figure 2 illustrates a detailed step-by-step algorithm starting from determining ITIC state points to obtaining VLE properties. As shown

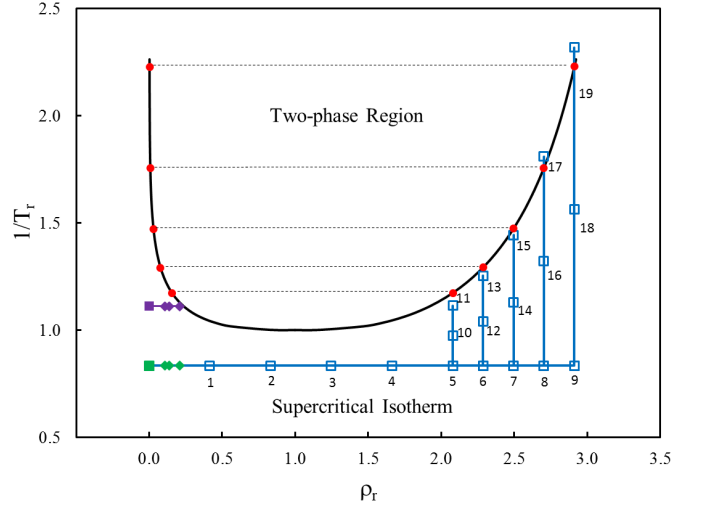


Figure 1: A schematic plot of the ITIC state points. Squares and diamonds represent  $NVT$  state points obtained from simulation. Square points 11, 13, 15, 17, and 19 represent the initial estimated  $T^{\text{sat}}$  values, therefore they do not necessarily match the coexistence curve. Green and purple diamonds show the state points required for  $B_2$  calculation at isothermal temperature and  $T_r = 0.9$ , respectively. Red points show the saturated vapor and liquid obtained from ITIC method.

in this figure, the iterations discussed above do not involve performing additional molecular simulations, rather the iterative solution of the ITIC equations is performed post-simulation.

The algorithm in Figure 2 also includes a step for obtaining second virial coefficient along isotherm ( $B_2^{\text{IT}}$ ) and saturation second virial coefficients ( $B_2^{\text{sat}}$ ) which is explained in Section 5.

## 2.3. ITIC state points and integration schemes

Figure 1 provides a schematic plot of ITIC state points. The minimal number of data points on each path required to achieve reliable results is determined. In order to estimate VLE data to a reduced temperature of 0.45 with precision comparable to that of GEMC and GCMC methods, one needs at least 9 data points on the isotherm and three data points on the isochore for each saturation point (see Section 3). The highest temperature state points, however, can serve on both isochores and isotherm, hence these points are only simulated once. The isotherm is constructed at a supercritical temperature. A reduced temperature of  $T_r \approx 1.2$  was the default value for the isotherm. For some compounds, owing to lack of NIST REFPROP values at high temperatures and the desire to compare isochores and isotherms to NIST REFPROP values when possible, a lower reduced temperature ( $T_r \approx 1.05$ ) is chosen for the isotherm. Validations against NIST REFPROP values show that any reduced temperature from 1.05 to 1.2 led to acceptable accuracy (results not presented here).

The integrations along the isotherm and isochores are performed using Simpson’s rule [29], due to its simplicity compared with other integration schemes as well as the possibility to use fixed step size for the integration so that simulations can be re-used for different isochores. Eq. (23) and Eq. (24) are

articulations of Simpson's rule used for numerical integration along the isotherm and isochores to calculate  $A^{\text{dep}}$  by Eq. (15).

$$\int_a^b f(x)dx \approx \frac{b-a}{6} \left[ f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] \quad (23)$$

$$\int_a^b f(x)dx \approx \frac{b-a}{8} \left[ f(a) + 3f\left(\frac{b-a}{3}\right) + 3f\left(\frac{2(b-a)}{3}\right) + f(b) \right] \quad (24)$$

The  $A^{\text{dep}}$  values at points 2, 4, 7, and 9 in Figure 1 are sequentially calculated using Eq. (23) in which the value of function at three equidistant points on the x-axis are needed. The  $A^{\text{dep}}$  values at points 3, 5, and 8 are sequentially calculated using Eq. (24) in which the value of the function at four equidistant points on the x axis are needed. The  $A^{\text{dep}}$  at point 6 is equal to the integration value from point 6 to point 8 (from Eq. (23)) subtracted from the  $A^{\text{dep}}$  value at point 8. The  $A^{\text{dep}}$  value at point 1 is equal to the integration value from point 1 to point 4 (from Eq. (24)) subtracted from  $A^{\text{dep}}$  value at point 4.

Second virial coefficient at a given temperature can be estimated by extrapolating  $(Z - 1)/\rho$  to  $\rho = 0$  at that temperature. The green and purple diamonds in Figure 1 are used to obtain the intercept of  $(Z - 1)/\rho$  vs. density at isothermal temperature and  $T_r = 0.9$ , respectively (represented by green and purple squares). These values are used to obtain a  $B_2$  correlation as a function of temperature. The green square is also used to obtain  $A^{\text{dep}}$  at point 2. In Section 5, a more detailed discussion of obtaining virial coefficient is provided.

#### 2.4. Enthalpy of Vaporization Calculation

In the ITIC method, enthalpy of vaporization ( $\Delta H_v$ ) is calculated using

$$\Delta H_v = (H_{\text{vap}}^{\text{dep}} - H_{\text{liq}}^{\text{dep}})RT \quad (25)$$

where  $H_{\text{vap}}^{\text{dep}}$  and  $H_{\text{liq}}^{\text{dep}}$  are unitless enthalpy departure functions of saturated vapor and liquid, respectively.  $H_{\text{vap}}^{\text{dep}}$  and  $H_{\text{liq}}^{\text{dep}}$  can be calculated by subtracting the ideal gas contribution from both sides of  $H = U + PV$

$$H_{\text{liq}}^{\text{dep}} \approx U_{\text{liq}}^{\text{dep}} + Z_{\text{liq}} - 1 \quad (26)$$

$$H_{\text{vap}}^{\text{dep}} \approx U_{\text{vap}}^{\text{dep}} + Z_{\text{vap}} - 1 \quad (27)$$

The value of  $U_{\text{liq}}^{\text{dep}}$  is calculated at each iteration based on the updated value of  $T^{\text{sat}}$ .  $U_{\text{vap}}^{\text{dep}}$ , on the other hand, is calculated by taking the derivative of  $B_2$  with respect to  $\beta$  as shown in Eq. (28).

$$U_{\text{vap}}^{\text{dep}} \approx \rho_{\text{vap}} \beta \frac{dB_2}{d\beta} \quad (28)$$

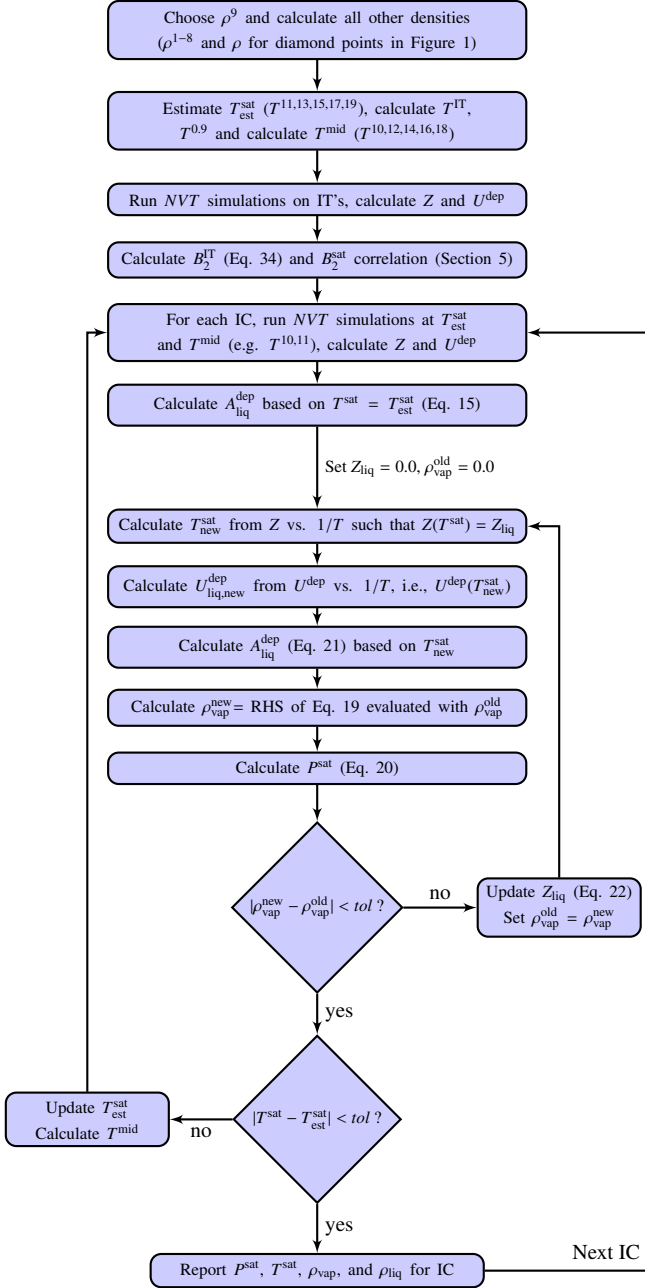


Figure 2: Algorithm to obtain VLE from NVT simulations using ITIC method

### 2.5. Critical Point Calculation

We follow the standard approach used in GEMC and GCMC methods to calculate the critical point. The critical temperature and critical density ( $\rho_c$ ) are estimated by the law of rectilinear diameter [30] and the density scaling law [31]

$$\frac{\rho_{\text{liq}} + \rho_{\text{vap}}}{2} = \rho_c + A(T_c - T^{\text{sat}}) \quad (29)$$

$$\rho_{\text{liq}} - \rho_{\text{vap}} = B(T_c - T^{\text{sat}})^{0.325} \quad (30)$$

where  $\rho_c$ ,  $T_c$ ,  $A$  and  $B$  are fit to simulation data.

The critical pressure ( $P_c$ ) is computed in two steps. The first step is to fit the ITIC saturation temperatures ( $T^{\text{sat}}$ ) and pressures ( $P^{\text{sat}}$ ) to the Antoine equation [32]

$$\log_{10}(P^{\text{sat}}) = a_0 + \frac{a_1}{a_2 + T^{\text{sat}}} \quad (31)$$

where  $a_i$  are fitting constants and  $a_1$  is constrained to be negative. The second step is to evaluate Eq.(31) with the optimal values of  $a_i$  for  $T^{\text{sat}} = T_c$ .

Some caution should be exercised when extrapolating Eqs.(29-31) because ITIC data are typically not available near the critical point. For example, we recommend excluding low temperature ITIC data ( $T_r < 0.6$ ) when fitting Eqs.(29-31) as these equations are typically not reliable over the entire temperature range. The results in Section 7 demonstrate that, although ITIC is generally limited to  $T_r < 0.85$ , careful application of Eqs.(29-31) provides reasonable estimates for all three critical constants.

## 3. ITIC Validation

### 3.1. Validation using NIST REFPROP

A first test to validate the ITIC method is to use a database that provides precise and self-consistent saturation properties and isochoric/isothermal properties. NIST Reference Fluid Properties software (REFPROP) provides such values [33], which were used to validate the ITIC method. The following comparisons are solely based on NIST REFPROP equations, therefore the lack of statistical noise inherent to molecular simulation enables an accurate evaluation of the numerical integration. Reproducing the REFPROP  $P^{\text{sat}}$  values using ITIC state points obtained from REFPROP is effectively a test of the spacing of the quadrature points since all the REFPROP thermodynamics derive exactly from their analytical equation of state.

Figure 3 shows the ITIC validation results for *n*-dodecane when virial expansion in Eq. (16) includes or excludes the  $B_3$  term. The deviations of calculated  $P^{\text{sat}}$ ,  $\rho_{\text{liq}}$ ,  $\rho_{\text{vap}}$ , and  $\Delta H_v$  from the data obtained directly from NIST REFPROP [34] are plotted in Figure 4. According to this figure, by including the  $B_3$  term, one can reach a reduced saturation temperature ( $T_r^{\text{sat}}$ ) of 0.9 with less than 1 % error in all saturation properties. The ITIC method is not able to calculate accurate saturation properties when  $T_r^{\text{sat}} > 0.9$ . If the  $B_3$  term is excluded from Eq. (16), the ITIC method does not converge for  $T_r^{\text{sat}} > 0.85$ . When  $T_r^{\text{sat}} < 0.85$ , excluding the  $B_3$  term provides less than 1 % deviation in  $P^{\text{sat}}$ ,  $\rho_{\text{liq}}$ , and  $\Delta H_v$ , and less than 2.5 % deviation in  $\rho_{\text{vap}}$ .

Figure 5(a) illustrates the convergence paths taken by a the ITIC iterations to calculate  $\rho_{\text{vap}}$ , with  $B_3$  included in Eq (16). Figure 5(b) shows the same plot, except  $B_3$  is excluded, i.e. virial expansion in Eq. (16) is truncated at  $B_2$  term. Using  $B_3$  corrects the curve representing the right-hand side of Eq. (19) in such a way that the iterations converge.

### 3.2. Vapor Pressure Sensitivity to Virial Coefficients

In this section, we investigate the sensitivity of vapor pressure to accuracy of  $B_2$  and  $B_3$  at saturation temperatures as well as  $B_2$  at supercritical isothermal temperature. This is important because  $B_2$  and  $B_3$  are often unknown for a given force field.

In order to estimate the required accuracy of the second virial coefficient at the isothermal temperature, Figure 6(a) is generated by changing the REFPROP  $B_2$  and calculating the corresponding deviations in *n*-dodecane vapor pressure. For example, a 5 % change in  $B_2$  results in around 2 % deviation in *n*-dodecane  $P^{\text{sat}}$ . This shows that it is imperative to use an accurate  $B_2$  value at the isothermal temperature.

Vapor pressure estimate is very weakly influenced at low saturation temperatures by accuracy of the second and third virial coefficients in Eq. (19) and Eq. (20). Figure 6(b) shows the  $P^{\text{sat}}$  sensitivity to  $B_2$  at various reduced temperatures, each representing one isochore. The first three lowest temperatures, are fairly insensitive to  $B_2$  precision such that even 50 % error in  $B_2$  results in less than 1 % deviation in  $P^{\text{sat}}$ . However, a relatively accurate  $B_2$  is required to obtain accurate  $P^{\text{sat}}$  when  $T_r^{\text{sat}} > 0.75$ . Similarly, when  $T_r^{\text{sat}} \approx 0.84$   $B_2$  deviations must not be greater than 2 % in order to have less than 1 % error in vapor pressure.

Figure 6(c) was plotted similar to Figure 6(b), with respect to  $B_3$ . It is worth mentioning that in order to truly understand the influence of  $B_3$ , exact values of  $B_2$  were used from NIST REFPROP. Even though adding  $B_3$  improves the overall behavior of the ITIC iteration in terms of convergence (Figure 5), the sensitivity of  $P^{\text{sat}}$  to  $B_3$  is negligible when  $T_r^{\text{sat}}$  is less than 0.85. This supports the idea of setting the  $B_3$  term to zero without significant loss of precision at such temperatures.

Similar to Figure 6, the effect of changing virial coefficients on  $\rho_{\text{vap}}$  and  $T^{\text{sat}}$  was considered. The sensitivity of  $\rho_{\text{vap}}$  to virial coefficients is similar to  $P^{\text{sat}}$ , and  $T^{\text{sat}}$  was found to be insensitive to virial coefficient deviations.

### 3.3. Sensitivity to Estimated Saturation Temperature

As mentioned in Section 2, saturation liquid densities in the ITIC method are fixed values at which we compute other saturation properties. This requires an initial estimate for saturation temperatures at the densities of interest.

It is important that the ITIC method does not depend strongly on the accuracy of  $T_{\text{est}}^{\text{sat}}$ . Figure 7 demonstrates that deviations in  $P^{\text{sat}}$  are less than 1 % and deviations in  $\rho_{\text{liq}}$  are less than 0.25 % when errors in  $T_{\text{est}}^{\text{sat}}$  are within 10 %. Deviations in  $\rho_{\text{vap}}$  are less than 1 %, except for the high temperature point for which deviations are smaller than 3 %. This can be improved by including  $B_3$ . Typical errors in  $T_{\text{est}}^{\text{sat}}$  are less than 1 %, in which case the deviations for all three properties are nearly indistinguishable

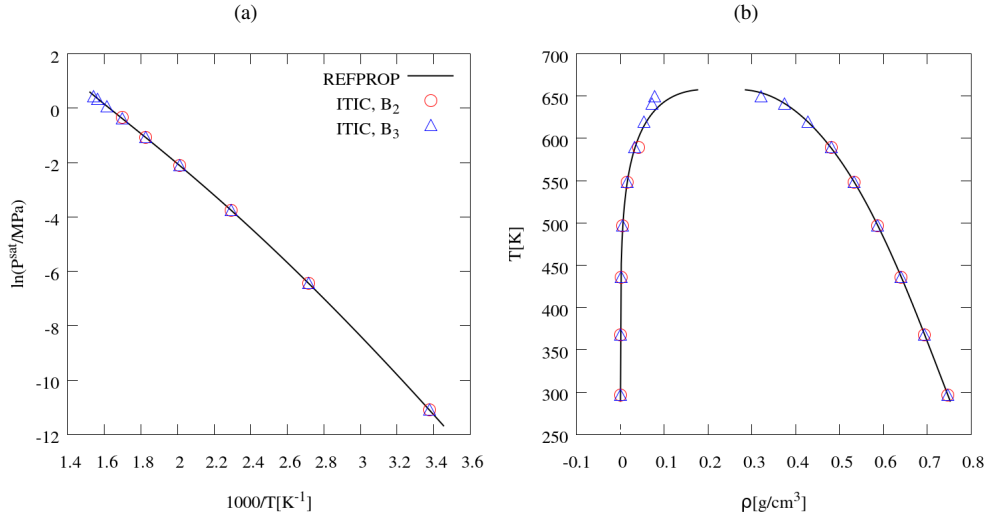


Figure 3: *n*-Dodecane Clausius-Clapeyron and coexistence curves. ITIC results (symbols) are obtained using NIST REFPROP [34] values for  $U^{\text{dep}}$  and  $Z$ . Circles and triangles represent ITIC results when virial expansion in Eq. (16) is truncated at  $B_2$  and  $B_3$  terms, respectively. Solid line represents true NIST REFPROP VLE data.

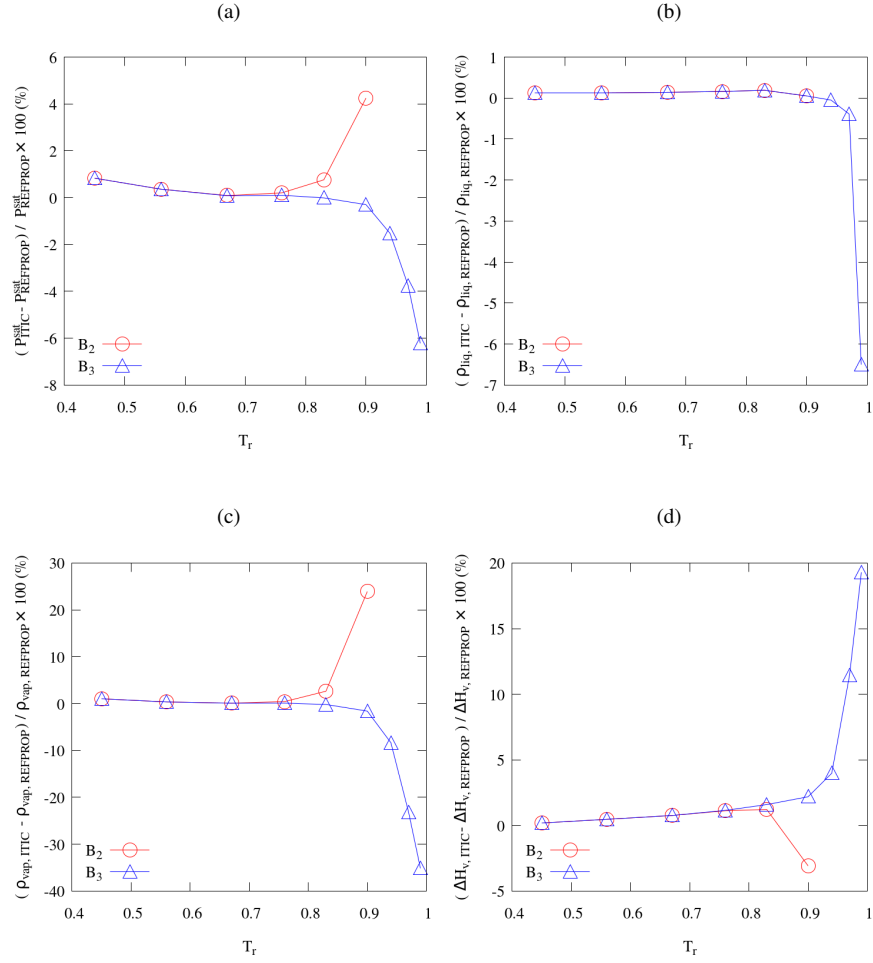


Figure 4: Accuracy of the ITIC method for *n*-dodecane when Eq. (16) is truncated at  $B_2$  and  $B_3$  term. If  $B_3$  is excluded the ITIC iteration does not converge, when  $T_r^{\text{sat}} > 0.9$ . Y-axis represents deviations calculated using  $\frac{\text{ITIC} - \text{REFPROP}}{\text{REFPROP}} \times 100$ , which compares ITIC results using REFPROP state points with REFPROP coexistence data.

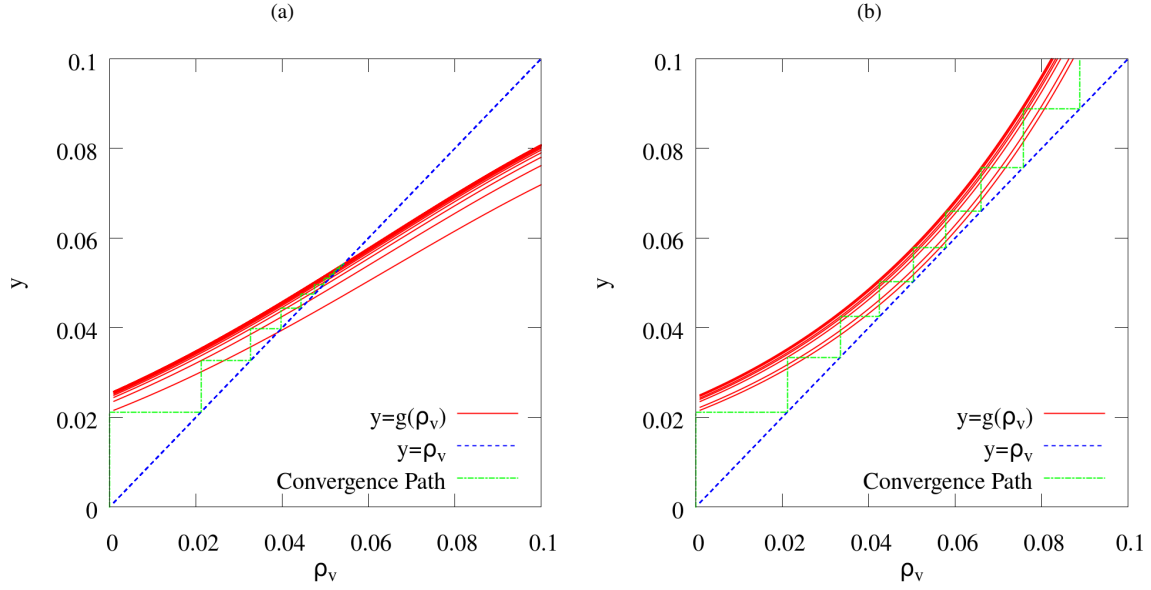


Figure 5: The ITIC iteration and convergence path for *n*-dodecane for the isochore corresponding to  $\rho_{\text{liq}} = 0.4269 \text{ g/cm}^3$  ( $T_r^{\text{sat}} \approx 0.94$ ). The *y*-axis represents the right-hand side of Eq. (19). At each iteration, the *y* curve is plotted based on a new set of  $T_r^{\text{sat}}$ ,  $P^{\text{sat}}$ ,  $A^{\text{dep}}$ , and  $Z_{\text{liq}}$ . Iteration starts with a low initial guess for  $\rho_{\text{vap}}$  and stops when absolute percent deviation between two consecutive  $\rho_{\text{vap}}$  values is less than a small tolerance, e.g. 0.1 %. The blue line represents the 45-degree line. **a)**  $B_3$  term is used in Eq. (16), **b)**  $B_3$  term is excluded from Eq. (16). Using  $B_3$  helps iteration to converge for high  $T^{\text{sat}}$  values.

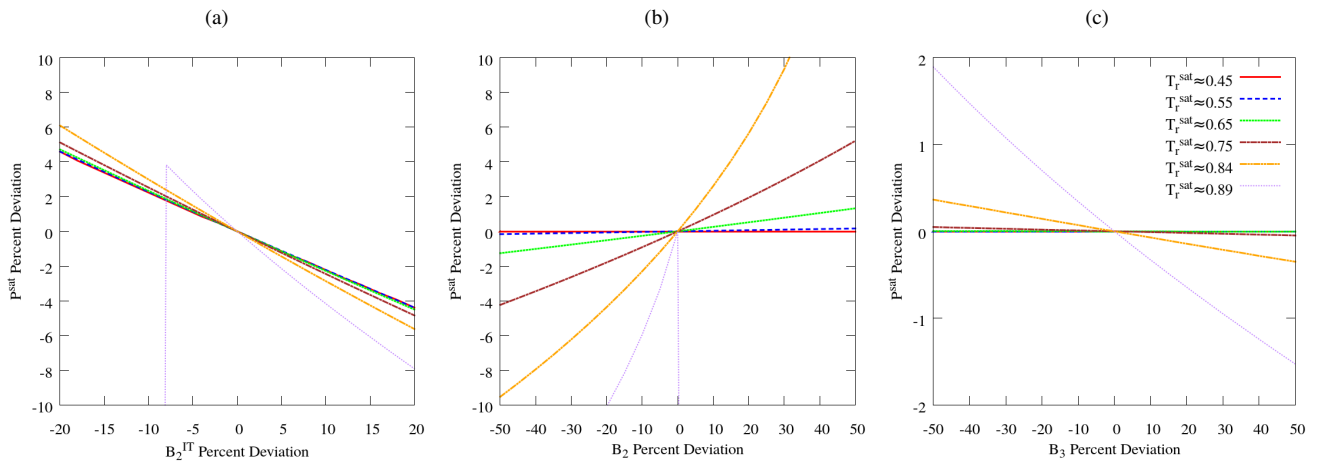


Figure 6: **a)** *n*-Dodecane  $P^{\text{sat}}$  sensitivity to isotherm  $B_2$ , **b)**  $P^{\text{sat}}$  sensitivity to second virial coefficient used in Eq. (20), and **c)**  $P^{\text{sat}}$  sensitivity to third virial coefficient used in Eq. (20)



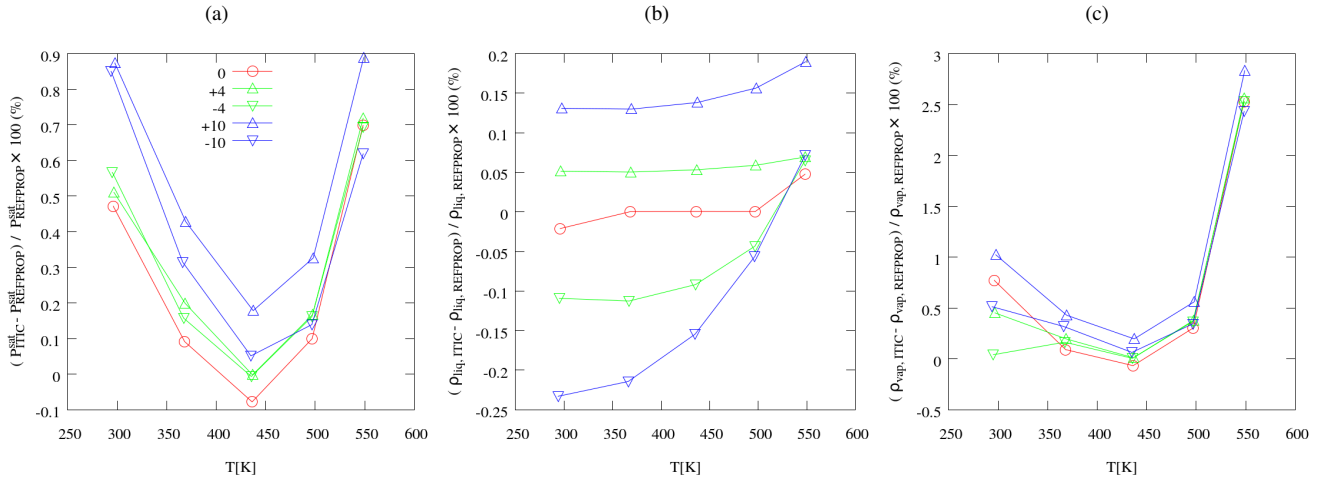


Figure 7: Sensitivity of the ITIC method to  $T_{\text{sat}}^{\text{sat}}$  for  $n$ -dodecane. The ITIC method is applied using REFPROP values of  $Z$  and  $U^{\text{dep}}$ . The y-axis represents percent deviation of the corresponding property calculated using ITIC compared with REFPROP saturation data. Triangles pointing up or down represent ITIC results when  $T_{\text{sat}}^{\text{sat}}$  is increased or decreased by the percentage shown in the legend, respectively.  $B_3$  is not included in Eq. (19).

from the deviations resulting from numerical integration alone, i.e., those for 0 % error in  $T_{\text{sat}}^{\text{sat}}$ .

Note that in Figure 7, isochoric/isothermal properties ( $Z$  and  $U^{\text{dep}}$ ) used for determining saturation properties are obtained from REFPROP. If  $T_{\text{sat}}^{\text{sat}}$  percent deviation shown in the legend of Figure 7(a) is less than zero (i.e.  $T_{\text{sat}}^{\text{sat}} < T_{\text{REFPROP}}^{\text{sat}}$ ), some of the ITIC points are in metastable state. Since the REFPROP database does not provide  $Z$  and  $U^{\text{dep}}$  for such points, a linear extrapolation of REFPROP data was used to approximate  $Z$  and  $U^{\text{dep}}$ . A similar sensitivity analysis using simulation results is included in supplementary material.

#### 4. Simulation Details

In principle, both Monte Carlo and molecular dynamics methods can be used to simulate the  $NVT$  state points required to construct the isothermal and isochoric paths in the ITIC method. In this study, the MC method is favored due to smaller uncertainties at low density. The Cassandra [35] and GOMC (GPU Optimized Monte Carlo) [36] packages are used to simulate several molecules in  $NVT$  ensemble using united-atom potential models. In united-atom force fields, interaction sites may consist of a group of atoms, which is centered on the main atom of the group for the TraPPE-UA method [37]. In the TraPPE-UA model van der Waals interactions are truncated at 1.4 nm and standard analytical long-range corrections are applied to compensate for truncation effects on energy and pressure[38]. Furthermore, the bond lengths are considered fixed and the bond energy is zero. This approximation results in smaller pressure fluctuations at low densities, but we note that the MC results at high densities were consistent with MD results simulated in Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [39] and GROMACS [40] within their uncertainties.

For each compound, 26  $NVT$  points are simulated in order to obtain 5 saturation points, as illustrated in Figure 1. The den-

sity of the isochore with highest density is chosen to match the experimental liquid density at the minimum reduced saturation temperature ( $T_r^{\text{min}}$ ) of 0.45. Densities and temperatures of all simulated state points are listed in the supplementary material along with average pressures and average energies.

The Packmol [41] software is used to create the initial configurations for LAMMPS and GOMC simulations, while Cassandra and GROMACS simulations were initialized using internal capabilities of this software. These initialization methods were adequate for generating the highest density liquid phase configurations of the most complex molecules studied. However, for even more challenging systems where these techniques might fail, we recommend initializing at a lower density and slowly compressing the box size to the desired density.

The simulation boxes typically contain 1200 sites except for the four simulations required for estimating  $B_2$  at the isotherm temperature (the purple and green diamonds in Figure 1) for which simulation boxes contain 4800 sites. Standard Periodic Boundary Conditions (PBCs) are used. Simulations are run for 10 million Monte Carlo steps, and the last 5 million MC steps are used for calculating the properties which are stored every 50,000 MC steps.

In order to approximate the computational cost of the ITIC method,  $n$ -dodecane coexistence points obtained at reduced temperatures of 0.65, 0.75, and 0.85 using GEMC are compared with the ITIC coexistence points obtained at liquid densities corresponding to reduced temperatures of approximately 0.65, 0.75, and 0.85. This temperature range is chosen because both GEMC and ITIC are reliable within  $0.65 < T_r < 0.85$ . Cassandra package is used for both simulations using an AMD 1.3 GHz processor. 3,500 and 13,500 MC cycles are used for equilibration and production of GEMC simulation, using five block averages to characterize the uncertainty. The average run-time of the three GEMC simulations is 15.8 hours with each simulation running on a separate CPU core, resulting in an average  $P^{\text{sat}}$  uncertainty of 7.5 % (relative standard error). On the other

hand, running ITIC state points for 500 MC cycles of equilibration and 2000 MC cycles of production with four block averages reproduced the GEMC results with 7.1 %  $P^{\text{sat}}$  uncertainty. The ITIC simulations were run on 20 CPU core. The maximum run-time (highest density and lowest temperature) of ITIC simulations is 18.9 hours. Therefore, when the number of available CPU cores is not limited, the ITIC method is approximately 20% slower than GEMC. If ITIC coexistence points at reduced temperatures of 0.45 and 0.55 are to be included, the maximum run-time slightly increases (i.e. 22.3 hours  $\approx$  40 % additional computational time compared to GEMC), whereas GEMC is not feasible for *n*-dodecane, unless a thermodynamic integration approach (i.e. Gibbs-Duhem) is used to extend the lower temperature limit of GEMC.

#### 4.1. Internal Energy Departure Function Calculation

Computing the internal energy departure function in Eq. (15) requires estimating the ideal gas energy ( $U_{\text{ig}}$ ) at each temperature simulated along a given isochore. The most rigorous approach to determine  $U_{\text{ig}}$  is by simulating a single molecule system at the corresponding temperature, such that

$$U^{\text{dep}} = \frac{E^{\text{tot}} - N \times E^{N=1}}{NRT} \quad (32)$$

where  $E^{\text{tot}}$  and  $E^{N=1}$  are the total potential energy and the single molecule potential energy, respectively. Some caution should be exercised when computing  $E^{N=1}$ . For example, molecular dynamics simulations with a single molecule are ill-advised when performed with standard thermostats that couple many degrees of freedom to a single bath [42]. For this purpose, we recommend using either chains of thermostats, Monte Carlo simulations, or stochastic (Langevin) dynamics [43] to compute  $E^{N=1}$ .

The additional single molecule simulations equilibrate very quickly, such that the extra CPU time incurred to compute  $E^{N=1}$  is not significant. However, a slightly simpler approach is to assume that  $U_{\text{ig}}$  is approximately equal to the intramolecular energy at the isochoric state point, such that

$$U^{\text{dep}} = \frac{E^{\text{tot}} - E^{\text{bonded}} - E^{\text{intra}}}{NRT} \quad (33)$$

where  $E^{\text{bonded}}$  and  $E^{\text{intra}}$  are, respectively, the bonded energy (bond, angle, and dihedral) and intramolecular pairwise energy (Coulombic and van der Waals) from the isochoric simulation. If the molecular simulation package does not provide an internal way of estimating  $E^{\text{intra}}$ , a post-processing code is required to calculate this quantity. For example, LAMMPS simulations require this post-processing. Note that failure to subtract  $E^{\text{intra}}$  causes a significant error in vapor pressure.

We emphasize that Eq. (33) should only be applied with small molecules (i.e., 5 or fewer atoms along the backbone), where the molecular configurations are similar for the ideal gas and condensed phases. Typical errors in  $P^{\text{sat}}$  are between 10% and 20% when applying Eq. (33) to larger molecules, where the deviation increases with decreasing  $T^{\text{sat}}$ .

Section 7 utilizes both methods for demonstrative purposes, where the single molecule method is implemented for larger molecules. A more detailed comparison between the two methods of calculating  $U^{\text{dep}}$  (i.e., Eq. (32)-(33)) is discussed in Supplementary Material.

#### 4.2. Bootstrapping Method for Uncertainty Calculation

Bootstrapping is used to estimate the statistical uncertainties [44] in  $T^{\text{sat}}$ ,  $P^{\text{sat}}$ ,  $\rho_{\text{vap}}$ , and  $\Delta H_v$ . Note that the ITIC method determines saturation conditions at a fixed value of  $\rho_{\text{liq}}$  (equal to the isochoric density) and, therefore, the bootstrap uncertainty in  $\rho_{\text{liq}}$  is zero. Four series of independent *NVT* simulations are performed for each compound using different random number generator seeds. Each series of *NVT* simulations comprises 26 state points including four at  $T_r = 0.9$  and three on isotherm used to estimate  $B_2$ . In order to increase the number of samples, the production data in each run is divided into two blocks. The ITIC analysis is performed using *NVT* state points randomly selected from the eight blocks, and saturation properties are computed. This process is repeated 500 times resulting in 500 sample values for each saturation property. The 95 % confidence intervals are then calculated from the resulting 500 ITIC outputs.

The resulting uncertainties are used to generate error bars represented in Figure 11, however in most cases the error bars are smaller than the symbols. For greater clarity, uncertainties are tabulated in supplementary material for each property at each saturation condition.

### 5. Calculation of Virial Coefficients

Mayer Sampling is a common approach for estimating virial coefficients of a given force field [45]. For example, Kofke and Schultz implement this approach for *n*-alkanes [46]. Efforts like those of Kofke and Schultz could make simulations on the lower end of supercritical isotherm unnecessary [47], in addition to facilitating computations just below the critical temperature. Furthermore, the higher virial coefficients play a significant role on their own merits in the development of high accuracy equations of state for simulation models [48].

In this study, we use a simpler approach that is amenable for developing a correlation for  $B_2$  and  $B_3$  with respect to temperature. Second virial coefficients are estimated by calculating the intercept of  $(Z - 1)/\rho$  with respect to  $\rho$ . In principle, the slope of this line at zero density also gives the third virial coefficient. Figure 8 shows the accuracy of this method when used at various temperatures. The blue lines in Figure 8(a) are obtained from Eq. (34)

$$\frac{Z - 1}{\rho} = B_2 + B_3\rho \quad (34)$$

where the intercept ( $B_2$ ) and slope ( $B_3$ ) correspond to Schultz's values [46].

Figure 8(a) shows that for TraPPE-UA ethane at temperatures above  $T_r^{\text{sat}} = 0.85$ ,  $B_2$  values calculated using this method

agree with values reported by Schultz to within 2 %. In Figure 8(a),  $Z$  at  $T_r = 0.8$  and  $\rho \approx 0.085 \text{ g/cm}^3$  is not consistent with the Schultz values due to proximity with the two-phase region. Therefore, using the lowest three densities would give a more accurate estimate for those temperatures. It is worth considering that the value of accurate virial coefficient characterization is enhanced in the context of the ITIC method.

According to Eq. (19), it is important to have a correlation for  $B_2$  and  $B_3$  with respect to temperature, because  $T^{\text{sat}}$  estimates change after each iteration and updated values for  $B_2$  and  $B_3$  are needed. Recently, a Taylor Series expansion with Mayer Sampling approach has been utilized to generate correlations with respect to temperature for  $B_2$  and  $B_3$  [50]. In this study, we obtain such a correlation from the formula used in the DIPPR [51] database, except the last term is omitted to decrease the number of parameters and avoid overfitting, as shown in Eq. (35)

$$B_2 = A + \frac{B}{T} + \frac{C}{T^3} \quad (35)$$

Eq. (36) and Eq. (37) are obtained by inserting  $B_2$  values extrapolated using Eq. (34) and their corresponding temperatures into Eq. (35).

$$B_2(T_{\text{IT}}) = A + \frac{B}{T_{\text{IT}}} + \frac{C}{T_{\text{IT}}^3} \quad (36)$$

$$B_2(T_{0.9}) = A + \frac{B}{T_{0.9}} + \frac{C}{T_{0.9}^3} \quad (37)$$

where  $T_{0.9}$  is the temperature corresponding to reduced temperature of 0.9 and  $T_{\text{IT}}$  represents the isothermal temperature.

Subtracting Eq. (37) from Eq. (36) gives

$$B_2(T_{\text{IT}}) - B_2(T_{0.9}) = B \left( \frac{1}{T_{\text{IT}}} - \frac{1}{T_{0.9}} \right) + C \left( \frac{1}{T_{\text{IT}}^3} - \frac{1}{T_{0.9}^3} \right) \quad (38)$$

Taking the derivative of  $B_2$  with respect to  $\beta$ , as shown in Eq. (28) leads to the internal energy departure function

$$\rho \frac{\partial B_2}{\partial \beta} = \rho \left( \frac{B}{T} + \frac{3C}{T^3} \right) = \frac{U - U_{\text{ig}}}{RT} \quad (39)$$

Therefore, the intercept of  $\frac{U - U_{\text{ig}}}{\rho RT}$  with respect to  $\rho$  gives the value of  $\beta \frac{\partial B_2}{\partial \beta}$ . Having this value can lead to Eq. (40) with two unknowns ( $B$  and  $C$ )

$$\beta \frac{\partial B_2}{\partial \beta} = \frac{B}{T_{0.9}} + \frac{3C}{T_{0.9}^3} \quad (40)$$

Solving three equations (Eq. (36)/(37), (38), and (40)) with three unknowns gives the values of  $A$ ,  $B$ , and  $C$ , hence a correlation for  $B_2$  with respect to temperature is derived. Figure 8(b) shows a correlation obtained by this method which is in good agreement with Schultz's simulation results [46].

$B_3$  calculation using Eq. (34) deviates significantly from the rigorous values, as shown in Figure 8(c). Therefore,  $B_3$  values estimated this way are useful but should be regarded as

“effective” values that indirectly account for  $B_4$  and higher order terms. On the other hand, more accurate representation of  $(Z - 1)/\rho$  is achieved using the  $B_3^{\text{eff}}$  from Eq. (34) than using the rigorous  $B_3$  or the rigorous  $B_3$  and  $B_4$  in combination, as shown in Figure 9 (compare solid and dashed black lines with green line). Note that in the ITIC applications considered in this study, only second virial coefficient is used, therefore obtaining accurate  $B_3$  is not a concern.

## 6. Accurate Low Density Simulations

According to Figure 6(a),  $P^{\text{sat}}$  accuracy is sensitive to accuracy of  $B_2$  values at the isothermal temperature ( $B_2^{\text{IT}}$ ). This sensitivity is due to accumulation of errors when integrating along isotherm, such that an error in  $B_2^{\text{IT}}$  affects the  $A^{\text{dep}}$  values at all other points along the isotherm and isochores. Similarly, one would expect a significant influence from low density points, i.e. points 1, 2, and 3 in Figure (1). Therefore, it is important to investigate the factors affecting the accuracy of  $Z$  at low densities on the isotherm. The factors considered in this study are system size, the choice of MD or MC, and the choice of fixed bonds or flexible bonds.

The system size effects at low densities are demonstrated in Figure 9. Plot of  $(Z - 1)/\rho$  with respect to  $\rho$  shown in Figure 9(a) demonstrates the effect of varying number of molecules in MD simulation of ethane with rigid bonds. Figure 9(a) demonstrates that, for the simulation times chosen, 3200 ethane molecules are necessary to obtain reliable estimates of  $B_2$  and to match the  $B_2 + B_3\rho + B_4\rho^2$  line, which represents the rigorous values of  $(Z - 1)/\rho$ . In the MD simulations shown in Figure 9(a) C-C bonds are held constant using the SHAKE algorithm [52].

The effect of using flexible bonds is shown in Figure 9(b). Two key limitations exist for obtaining reliable estimates with flexible bond MD simulations. First, the MD flexible uncertainties are considerably larger than the MD rigid uncertainties. Note that the same number of production steps were used to compute the uncertainties for both MD flexible and MD rigid, for a given value of  $N$ . Second, while all of the MD rigid simulations equilibrated in less than 1 ns, the MD flexible systems at the lowest densities required equilibration times between 24-47 ns (with 24 ns corresponding to  $N = 120$ ). The slow equilibration for MD flexible simulations is presumably due to the small number of intermolecular collisions at low density relative to the large number of intramolecular collisions.

Figure 9(c) shows the low density  $NVT$  state points simulated using GOMC [36]. This plot shows that the MC method gives more reliable results than MD for low density  $NVT$  state points. Therefore, we recommend using MC when simulating these low density points. The choice of MD or MC for other high density state points in ITIC method is less important, because they generally agree with each other within their uncertainties.

## 7. Example Simulations

The TraPPE-UA [4, 60, 54], Mie-UA [5, 56, 61, 62], and TIP4P/2005 [63] models were chosen for the purpose of testing

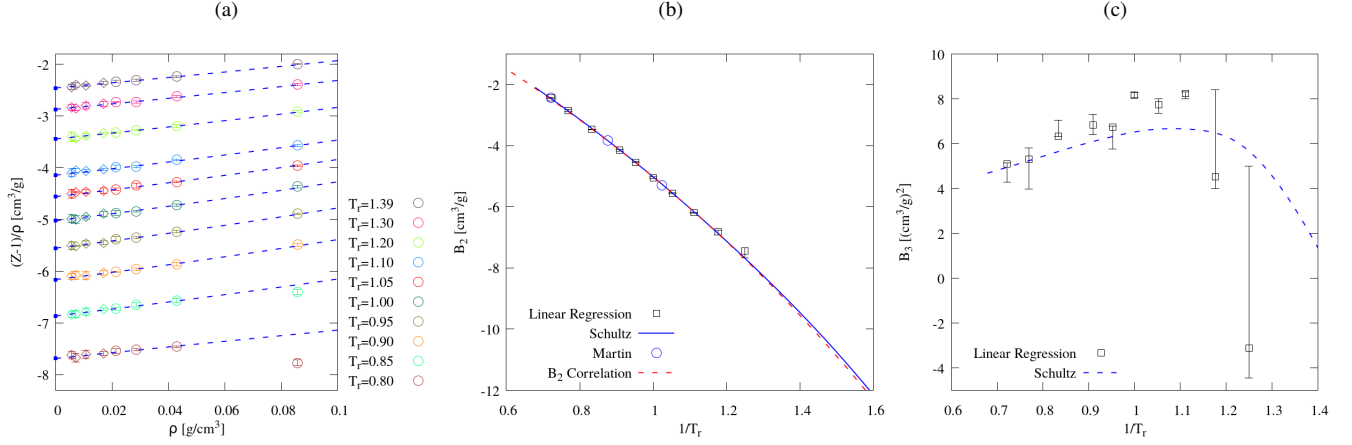


Figure 8: Panel **a**) shows the plot of  $(Z-1)/\rho$  with respect to  $\rho$  for TrAPPE-UA ethane. Blue solid squares represent Schultz's  $B_2$  values. [46]. Blue dashed lines represent Schultz values of  $B_2$  (intercept) and  $B_3$  (slope). Circles and diamonds are  $NVT$  state points simulated with the GOMC package [36]. Using circles in **a**) suffices to obtain accurate virial coefficients, therefore diamond points are not used in  $B_2$  and  $B_3$  calculations. In panel **b**) and **c**) black squares represent the median of  $B_2$  and  $B_3$  when circle points  $\{1,2,3\}$ ,  $\{1,2,3,4\}$ ,  $\{2,3,4\}$ , and  $\{1,2,4\}$  from panel **a**) were used in linear regression according to Eq.(34). Error bars represent bootstrapped uncertainties. The blue circles in panel **b**) represent  $B_2$  values obtained by Martin and Siepmann [4] using a Monte Carlo method [49].

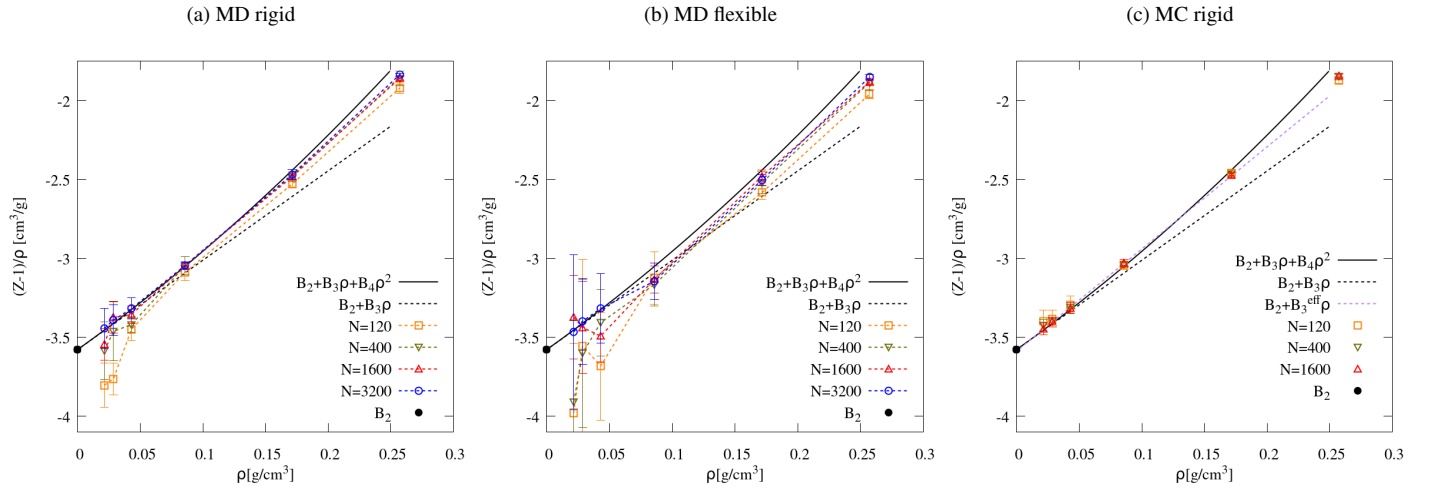


Figure 9: Effect of number of ethane molecules on compressibility factor at low densities at  $T^{\text{IT}} = 360$ . Panel **a**) is from MD with rigid bonds (using SHAKE algorithm [52] in LAMMPS), panel **b**) is from MD with flexible bonds in LAMMPS (harmonic potential using multiple-time-step algorithm RESPA [53] algorithm), and panel **c**) is from MC with rigid bonds (GOMC). Solid black lines represent  $B_2 + B_3\rho + B_4\rho^2$  curve where  $B_{2-4}$  are obtained from Schultz's work [46]. Dashed black lines represent  $B_2 + B_3\rho$  line where  $B_{2-3}$  are obtained from Schultz's work. Dashed purple lines represent  $B_2 + B_3^{\text{eff}}\rho$  where  $B_2$  and  $B_3^{\text{eff}}$  are fit to four lowest density simulation values. Note that the increasing deviation between black line and simulation points at higher densities is due to truncation of virial equation at  $B_4$ . The error bars illustrate the 95 % confidence intervals calculated from four separate runs each consisting two blocks (8 samples). Number of MD steps used to calculate averages is inversely proportional to  $N$  with 6 million timesteps for  $N=120$ . Equilibration times were 1 ns for MD rigid and between 24–47 ns for MD flexible.

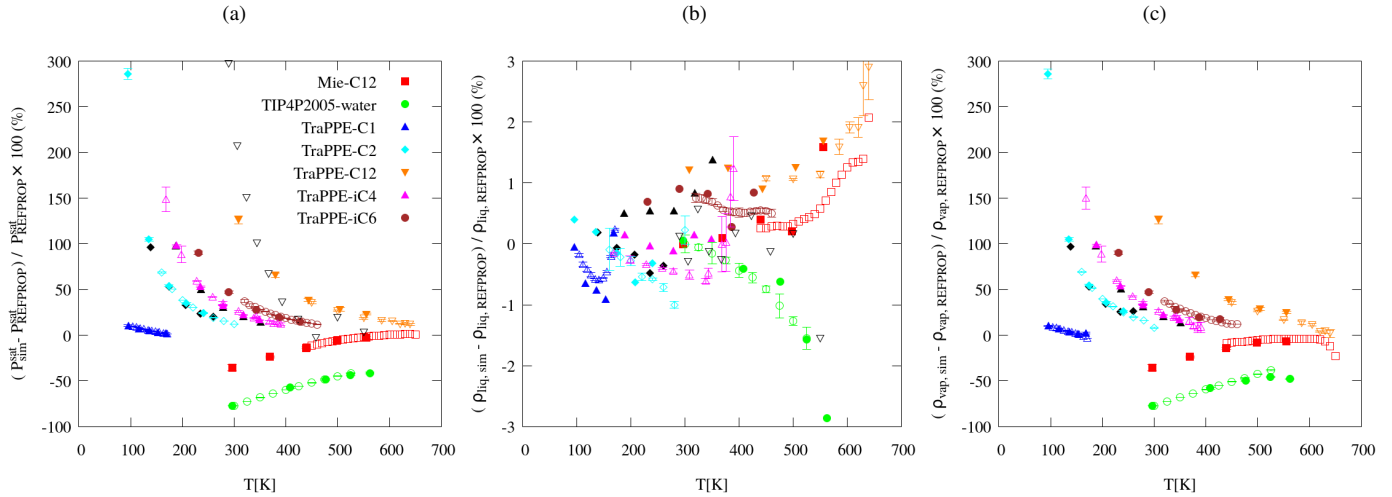


Figure 10: Comparison between ITIC (filled) with respect to GEMC and GCMC methods (unfilled). The y-axis represents deviation from REFPROP data. GEMC: TraPPE-UA *n*-dodecane (orange) [4], TraPPE-UA ethane (cyan) [4], TraPPE-UA isobutane (purple) [54]; GCMC: TraPPE methane[55] (blue), Mie-UA *n*-dodecane (red) [5], TIP4P/2005 water (green) [55], and TraPPE-UA isohexane (brown) [56]. Black filled symbols represent ITIC results simulated with GROMACS [40]. Black unfilled triangles represent TraPPE-UA *n*-dodecane from Ref. [10].  $U^{\text{dep}}$  is computed with Eq. (32) for *n*-dodecane and isohexane by performing single molecule simulations, while  $U^{\text{dep}}$  is computed with Eq. (33) for other compounds. Error bars for ITIC data (typically smaller than symbol sizes) denote 95 % confidence intervals assuming zero variation in  $T^{\text{sat}}$ . Due to some ambiguity regarding whether literature uncertainties correspond to standard deviations or 95 % confidence intervals, error bars for literature values represent uncertainties as reported by the respective authors.

the ITIC method due to the availability of literature vapor-liquid coexistence simulation results. Figures 10 and 11 demonstrate good agreement between traditional GEMC/GCMC methods and the ITIC approach using both MC and MD. The ITIC and literature values (GEMC/GCMC) typically agree to within less than 0.5 % for  $\rho_{\text{liq}}$  and within a few percent for  $P^{\text{sat}}$  and  $\rho_{\text{liq}}$ .

Figure 10 provides a quantitative comparison between ITIC method and traditional vapor-liquid coexistence Monte Carlo methods. This figure compares the deviations of ITIC and MC methods from REFPROP values as a baseline to make the magnitudes of the discrepancies more clear. Figure 11 shows the Clausius-Clapeyron and coexistence curves for all example simulations compared to MC methods. In all ITIC calculations,  $B_2$  is included in Eq. (19) and Eq. (20), while  $B_3$  is set to zero for simplicity. Initial  $T^{\text{sat}}$  values for all the compounds shown in Figures 10 and 11 were obtained from the DIPPR [51] database. Complete information about chosen ITIC state points as well as the results of  $NVT$  simulations is included in Supplementary Materials.

Saturation points calculated using the ITIC method are compared against TraPPE-UA results obtained using GEMC which are available from the TraPPE website [64]. Figure 10, 11(a), and 11(b) compare the ITIC and GEMC results for pure ethane, *n*-dodecane, and isobutane systems.  $NVT$  simulations at ITIC state points are performed using the Cassandra Monte Carlo [35] and GROMACS molecular dynamics [40] packages.

In Figure 10, the ITIC results of TraPPE-UA *n*-dodecane are also compared with the Gibbs-Duhem integration method [10]. Both methods provide a similar lower temperature limit and they are in relatively good agreement. Note that ITIC results for TraPPE-UA ethane are obtained near the triple point (i.e.,  $T^{\text{sat}} \approx 95\text{K}$ ). This demonstrates that ITIC can be applied

for  $T_r^{\text{sat}} \approx 0.3$ .

The ITIC method was also compared to histogram-reweighting GCMC provided in Figure 11(c) and Figure 11(d). GCMC results for Mie-UA *n*-dodecane are not available below a minimum reduced temperature ( $T_r^{\text{min}}$ ) of 0.67, however the ITIC method allowed us to calculate vapor pressure and liquid densities for reduced temperatures as low as 0.45.

In order to validate the ITIC method for polar molecules, the results of the ITIC method using TIP4P/2005 water simulated in Cassandra were compared against TIP4P/2005 data from the NIST Standard Reference Simulation Website [55] simulated using grand-canonical Wang-Landau/Transition-matrix Monte Carlo and histogram re-weighting. Figure 10, 11(c), and 11(d) show the agreement between the two methods for TIP4P/2005 water. The absolute average percent deviation between vapor pressure calculated using ITIC method and GCMC method for TIP4P/2005 water shown in Figure 10(a) is less than 1 %. This agreement shows that second virial coefficient can be obtained using Eq. (34) for TIP4P/2005 model.

Figure 11 includes the critical points obtained using the method described in Section 2. The ITIC coexistence points shown in this figure do not exceed  $T_r^{\text{sat}} \approx 0.85$ , the estimated critical properties are subject to larger uncertainties and possible systematic deviations. In the case of TIP4P/2005 water, critical point calculation requires a more suitable method as shown in Ref. [11]. Accurate estimation of critical points by molecular simulation requires a careful and deliberate effort that includes accounting for system size effects [65]. That effort goes beyond the scope of the current manuscript.

One limitation of the ITIC approach is the need for a reasonable value of  $T_{\text{est}}^{\text{sat}}$ . The ITIC calculations shown so far are done for well-known molecules for which extensive experimen-

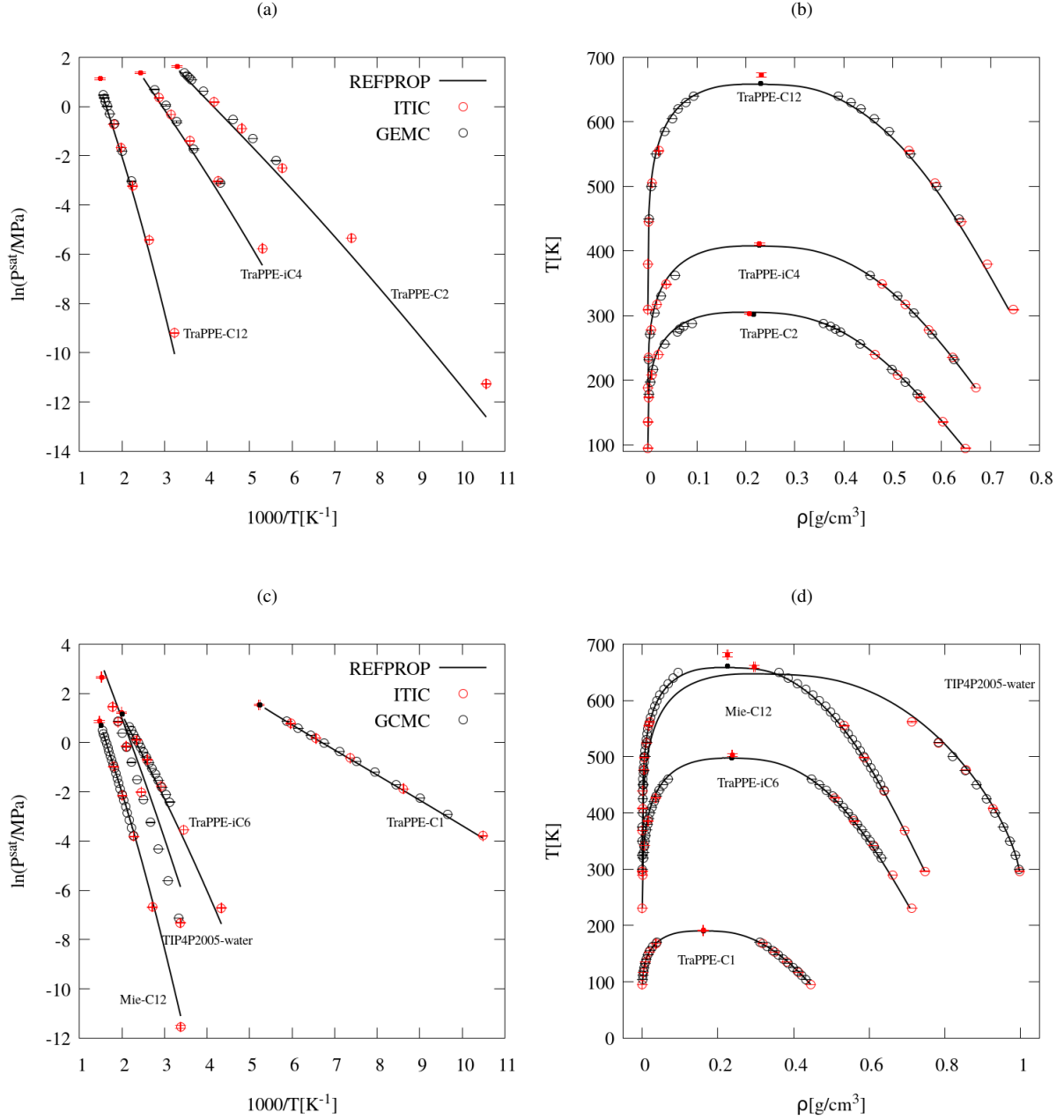


Figure 11: Comparison between ITIC method and Monte Carlo methods: GEMC (TraPPE-UA *n*-dodecane [4], TraPPE-UA ethane [4], and TraPPE-UA isobutane [57]), and GEMC (TraPPE-UA methane [55], Mie-UA *n*-dodecane [5], TIP4P/2005 water [55], and TraPPE-UA isohexane [56]).  $B_2$  values at saturation temperatures were obtained using low density simulations described in Section 5, except for TIP4P/2005 simulation where  $B_2$  correlation is obtained from Ref. [58, 59]. Black and red filled symbols show the critical points from literature and ITIC, respectively.

tal data are available. All force fields considered (TraPPE-UA, Mie-UA, TIP4P/2005) provide accurate representation of the  $T^{\text{sat}}$  vs.  $\rho_{\text{liq}}$  curve compared to experimental data. Therefore, the  $T_{\text{est}}^{\text{sat}}$  values for these compounds were obtained from DIPPR database. However, it is important to make sure that ITIC method also works for the molecules for which experimental data are not available. For example, 1-naphthalenyl, 4-phenanthrenyl butane is a large aromatic with 5 rings and 28 united-atom sites for which, to the best of our knowledge, no experimental data is available. In this case, a simple linear extrapolation using two points (black (x) symbols in Figure 12(a)) along the isochore is used to obtain  $T_{\text{est}}^{\text{sat}}$ , i.e. the temperature at which  $Z \approx 0$ . Once  $T_{\text{est}}^{\text{sat}}$  is calculated, the procedure in Figure 2 is followed. Figure 12(b-c) shows good agreement between the ITIC method with  $T_{\text{est}}^{\text{sat}}$  as above and both GCMC and GEMC results obtained using GOMC.

Alternatively, one can improve ITIC results when a reasonable estimate for  $T_{\text{est}}^{\text{sat}}$  is not available, by running the ITIC procedure iteratively, i.e. using  $T^{\text{sat}}$  from the previous iteration. It's worth mentioning that we do not need to resimulate the isotherms, and the isochoric information from previous iterations can potentially help improve the integration accuracy along the isochores.

The isothermal/isochoric plots of Helmholtz energy departure function, compressibility factor, and internal energy departure function as well as plots of second virial coefficient and heat of vaporization for all example simulations are included in supplementary material. Also included in supplementary material are tables containing ITIC results for all example simulations.

## 8. Conclusions

The ITIC method is shown to be a reliable alternative for phase equilibrium calculations. In the absence of simulation uncertainty, i.e. when REFPROP data was used as input, the vapor pressure calculated by the ITIC method with 9 points on isotherm and 3 points on isochore reproduces NIST REFPROP vapor pressure within 1 % deviation. In applications where simulation uncertainty is significant, ITIC is sensitive to low density  $NVT$  simulations, but the noise at low densities can be addressed by simulating larger systems, and preferring  $NVT$  Monte Carlo method when feasible.

It is important for engineering applications to be able to simulate systems near the triple point, i.e., at temperatures as low as  $T_r = 0.3$  to  $0.45$ . Monte Carlo methods such as GEMC and GCMC usually have a minimum reduced temperature limit of about  $0.6$ , due to the insertion/transfer moves, unless combined with thermodynamic integration. The ITIC method, hence, outperforms GEMC and GCMC when  $T_r < 0.6$ , while attaining a similar lower temperature limit as GEMC-TI ( $T_r = 0.3$ ). The ITIC method, on the other hand, is less favorable at high reduced temperatures, especially above  $T_r = 0.85$ , mostly due to lack of a convenient method to characterize the virial coefficients. In addition, ITIC requires high temperature high pressure property estimates, which provide additional information for force field characterization.

The presentation here has focused primarily on GEMC as a basis for comparison, owing to its common application to this purpose and reader familiarity. Our comparisons show that ITIC is moderately more computationally expensive than GEMC, but that the added expense is worthwhile because it provides simulation results throughout phase space and accesses lower reduced temperatures. Other approaches, such as GEMC combined with Gibbs-Duhem integration could address the lower temperatures, but the computational expense would then approach that of ITIC, but still not provide simulation results away from the coexistence curve.

In conclusion, ITIC can easily be implemented from  $T_r = 0.45$  to  $0.85$  with approximately 40 % additional computational time over the  $0.6$  to  $0.85$  temperature range using GEMC. If temperatures above  $0.85$  are required, it is recommended to approach the problem of coexistence calculation with a combination of Monte Carlo (GEMC or GCMC) and isothermal-isochoric integration. If a single method is preferred, or if MD is the preferred simulation method, ITIC has notable advantages. These advantages could be enhanced over time with the availability of rigorous higher order virial coefficients for broader ranges of molecular types.

## 9. Acknowledgments

This research was performed while R.A.M. held a National Research Council (NRC) Postdoctoral Research Associateship at the National Institute of Standards and Technology (NIST). Contribution of NIST, an agency of the United States government; not subject to copyright in the United States. We would like to thank the reviewers whose comments helped improve this manuscript as well as Mr. Mohammad Soroush Barhaghi for providing GCMC data used in Figure (12).

## 10. Supplementary Material

See supplementary material for the isothermal/isochoric plots of  $A^{\text{dep}}$ ,  $U^{\text{dep}}$ ,  $Z$ , and plots of  $\Delta H_v$ , and  $B_2$  for all example simulations. Also included are data shown in Figure 4 and figures of  $T_{\text{est}}^{\text{sat}}$  sensitivity study using simulation data. The ITIC code is available at <https://github.com/mostafa-razavi/ITIC>

## 11. References

- [1] D. Frenkel, B. Smit, *Understanding Molecular Simulation: from Algorithms to Applications*, Academic Press, San Diego, 1996.
- [2] A. Z. Panagiotopoulos, Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, *Molecular Physics* 61 (4) (1987) 813–826.
- [3] A. S. Paluch, V. K. Shen, J. R. Errington, Comparing the Use of Gibbs Ensemble and Grand-Canonical Transition-Matrix Monte Carlo Methods to Determine Phase Equilibria, *Industrial & Engineering Chemistry Research* 47 (13) (2008) 4533–4541. doi:10.1021/ie800143n.
- [4] M. G. Martin, J. I. Siepmann, Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes, *The Journal of Physical Chemistry B* 102 (97) (1998) 2569–2577. doi:10.1021/jp972543+.
- [5] J. J. Potoff, D. A. Bernard-Brunel, Mie potentials for phase equilibria calculations: Application to alkanes and perfluoroalkanes, *The Journal of Physical Chemistry B* 113 (44) (2009) 14725–14731.



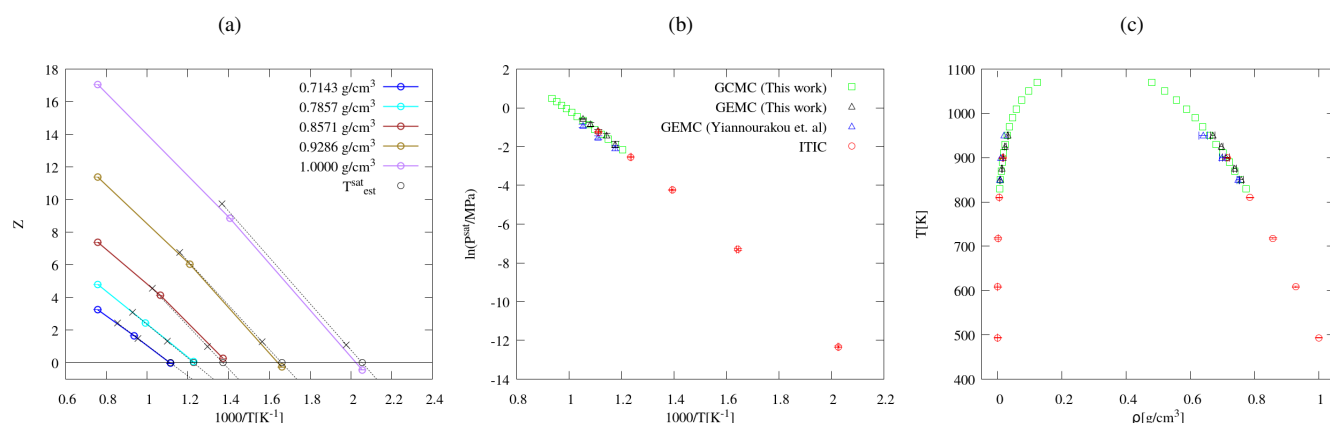


Figure 12: Simulation of 1-naphthalenyl,4-phenanthrenyl butane based on an extended TraPPE-UA model [12]. Black (x) symbols are simulated at two arbitrarily chosen temperatures to obtain  $T_{\text{est}}^{\text{sat}}$  (black circles) for each isochore. The ITIC, GCMC and GEMC results (black triangle) are simulated in this study using GOMC. Blue triangles are GEMC data obtained from Ref. [12].

- [6] D. Y. Peng, D. B. Robinson, A New Two-Constant Equation of State, *Industrial and Engineering Chemistry Fundamentals* 15 (1976) 59.
- [7] D. A. Kofke, Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line, *The Journal of Chemical Physics* 98 (5) (1993) 4149–4162. doi:10.1063/1.465023.
- [8] M. G. Ahunbay, S. Kranias, V. Lachet, P. Ungerer, Prediction of thermodynamic properties of heavy hydrocarbons by Monte Carlo simulation, *Fluid Phase Equilibria* 224 (1) (2004) 73–81. doi:10.1016/j.fluid.2004.06.053.
- [9] C. Vega, E. Sanz, J. L. F. Abascal, E. G. Noya, Determination of phase diagrams via computer simulation: Methodology and applications to water, electrolytes and proteins 20 (2008). arXiv:0901.1823, doi:10.1088/0953-8984/20/15/153101.
- [10] P. Ungerer, C. Beauvais, J. Delhommelle, A. Boutin, B. Rousseau, A. H. Fuchs, Optimization of the anisotropic united atoms intermolecular potential for n-alkanes, *Journal of Chemical Physics* 112 (12) (2000) 5499–5510. doi:10.1063/1.481116.
- [11] C. Vega, J. L. Abascal, I. Nezbeda, Vapor-liquid equilibria from the triple point up to the critical point for the new generation of TIP4P-like models: TIP4P/Ew, TIP4P/2005, and TIP4P/ice, *Journal of Chemical Physics* 125 (3). doi:10.1063/1.2215612.
- [12] M. Yiannourakou, P. Ungerer, V. Lachet, B. Rousseau, J.-M. Teuler, United atom forcefield for vapor-liquid equilibrium (VLE) properties of cyclic and polycyclic compounds from Monte Carlo simulations, *Fluid Phase Equilibria* doi:10.1016/j.fluid.2018.07.001.
- [13] Q. Yan, J. J. De Pablo, Hyper-parallel tempering Monte Carlo: Application to the Lennard-Jones fluid and the restricted primitive model, *Journal of Chemical Physics* 111 (21) (1999) 9509–9516. doi:10.1063/1.480282.
- [14] J. R. Elliott, L. Hu, Vapor-liquid equilibria of square-well spheres, *The Journal of Chemical Physics* 110 (6) (1999) 3043–3048.
- [15] N. F. Carnahan, Equation of State for Nonattracting Rigid Spheres, *The Journal of Chemical Physics* 51 (2) (1969) 635. doi:10.1063/1.1672048.
- [16] S. M. Razavi, Optimization of a transferable shifted force field for interfaces and inhomogeneous fluids using thermodynamic integration, Master's thesis, University of Akron (2016).
- [17] R. Lustig, G. Rutkai, J. Vrabec, Thermodynamic correlation of molecular simulation data, *Molecular Physics* 113 (9-10) (2015) 910–931. doi:10.1080/00268976.2015.1023752.
- [18] M. Thol, G. Rutkai, A. Köster, R. Lustig, R. Span, J. Vrabec, Equation of state for the Lennard-Jones fluid, *Journal of Physical and Chemical Reference Data* 45 (2) (2016) 023101. arXiv:https://doi.org/10.1063/1.4945000, doi:10.1063/1.4945000.
- [19] M. Thol, G. Rutkai, A. Köster, F. H. Dubberke, T. Windmann, R. Span, J. Vrabec, Thermodynamic properties of octamethylcyclotetrasiloxane, *Journal of Chemical & Engineering Data* 61 (7) (2016) 2580–2595. arXiv:https://doi.org/10.1021/acs.jced.6b00261, doi:10.1021/acs.jced.6b00261.
- [20] M. Thol, F. Dubberke, G. Rutkai, T. Windmann, A. Köster, R. Span, J. Vrabec, Fundamental equation of state correlation for hexamethyldisiloxane based on experimental and molecular simulation data, *Fluid Phase Equilibria* 418 (2016) 133 – 151, special Issue covering the Nineteenth Symposium on Thermophysical Properties. doi:https://doi.org/10.1016/j.fluid.2015.09.047.
- [21] M. Thol, G. Rutkai, A. Köster, S. Miroshnichenko, W. Wagner, J. Vrabec, R. Span, Equation of state for 1,2-dichloroethane based on a hybrid data set, *Molecular Physics* 115 (9-12) (2017) 1166–1185. arXiv:https://doi.org/10.1080/00268976.2016.1262557, doi:10.1080/00268976.2016.1262557.
- [22] G. Rutkai, M. Thol, R. Lustig, R. Span, J. Vrabec, Communication: Fundamental equation of state correlation with hybrid data sets, *The Journal of Chemical Physics* 139 (4) (2013) 041102. arXiv:https://doi.org/10.1063/1.4817203, doi:10.1063/1.4817203.
- [23] G. Rutkai, J. Vrabec, Empirical fundamental equation of state for phosphene based on molecular simulation data, *Journal of Chemical & Engineering Data* 60 (10) (2015) 2895–2905. arXiv:https://doi.org/10.1021/acs.jced.5b00266, doi:10.1021/acs.jced.5b00266.
- [24] M. Thol, G. Rutkai, A. Köster, M. Kortmann, R. Span, J. Vrabec, Fundamental equation of state for ethylene oxide based on a hybrid dataset, *Chemical Engineering Science* 121 (2015) 87 – 99, 2013 Danckwerts Special Issue on Molecular Modelling in Chemical Engineering. doi:https://doi.org/10.1016/j.ces.2014.07.051.
- [25] M. Thol, G. Rutkai, R. Span, J. Vrabec, R. Lustig, Equation of State for the Lennard-Jones Truncated and Shifted Model Fluid, *International Journal of Thermophysics* 36 (1) (2015) 25–43. doi:10.1007/s10765-014-1764-4.
- [26] C. Nieto-Draghi, G. Fayet, B. Creton, X. Rozanska, P. Rotureau, J. C. De Hemptinne, P. Ungerer, B. Rousseau, C. Adamo, A General Guidebook for the Theoretical Prediction of Physicochemical Properties of Chemicals for Regulatory Purposes, *Chemical Reviews* 115 (24) (2015) 13093–13164. doi:10.1021/acs.chemrev.5b00215.
- [27] R. Messerly, S. Razavi, M. Shirts, Configuration-Sampling-Based Surrogate Models for Rapid Parameterization of Non-Bonded Interactions, *Journal of Chemical Theory and Computation* 14 (6). doi:10.1021/acs.jctc.8b00223.
- [28] J. R. Elliott, C. T. Lira, Introductory chemical engineering thermodynamics, Prentice Hall PTR Upper Saddle River, NJ, 1999.
- [29] K. E. Atkinson, An introduction to numerical analysis, John Wiley & Sons, 2008.
- [30] J. S. Rowlinson, Liquids and Liquid Mixtures (1982).
- [31] J. S. Rowlinson, B. Widom, Molecular theory of capillarity, Courier Corporation, 2013.
- [32] J. S. Rowlinson, F. Swinton, Liquids and liquid mixtures: Butterworths



- monographs in chemistry, Butterworth-Heinemann, 2013.
- [33] E. W. Lemmon, M. L. Huber, M. O. McLinden, NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties-REFPROP, Version 9.1, National Institute of Standards and Technology (2013). doi:<http://dx.doi.org/10.18434/T4JS3C>.
  - [34] E. W. Lemmon, M. L. Huber, Thermodynamic properties of n-dodecane, *Energy & Fuels* 18 (4) (2004) 960–967. arXiv:<https://doi.org/10.1021/ef0341062>, doi:10.1021/ef0341062.
  - [35] J. K. Shah, E. Marin-Rimoldi, R. G. Mullen, B. P. Keene, S. Khan, A. S. Paluch, N. Rai, L. L. Romanielo, T. W. Rosch, B. Yoo, E. J. Maginn, Cassandra: An open source Monte Carlo package for molecular simulation, *Journal of Computational Chemistry* 38 (19) (2017) 1727–1739. doi:10.1002/jcc.24807.
  - [36] J. Mick, E. Hailat, V. Russo, K. Rushaidat, L. Schwiebert, J. Potoff, GPU-accelerated Gibbs Ensemble Monte Carlo simulations of Lennard-Jonesium, *Computer Physics Communications* 184 (12) (2013) 2662–2669.
  - [37] B. Smit, S. Karaborni, J. I. Siepmann, B. Smit, S. Karaborni, J. I. Siepmann, Computer simulations of vapor liquid phase equilibria of n-alkanes Computer simulations of vapor liquid phase equilibria of n-alkanes 2126 (1995).
  - [38] M. P. Allen, D. J. Tildesley, *Computer simulation of liquids*, Oxford university press, 2017.
  - [39] S. Plimpton, P. Crozier, A. Thompson, LAMMPS-large-scale atomic/molecular massively parallel simulator, Sandia National Laboratories 18.
  - [40] E. Lindahl, B. Hess, D. van der Spoel, GROMACS 3.0: A Package for Molecular Simulation and Trajectory Analysis, *J. Mol. Mod.* 7 (2001) 306–317.
  - [41] L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, PACKMOL: a package for building initial configurations for molecular dynamics simulations, *Journal of computational chemistry* 30 (13) (2009) 2157–2164.
  - [42] P. T. Merz, M. R. Shirts, Testing for physical validity in molecular simulations, *PLOS ONE* 13 (9) (2018) 1–22. doi:10.1371/journal.pone.0202764.
  - [43] N. Goga, A. J. Rzepiela, A. H. de Vries, S. J. Marrink, H. J. C. Berendsen, Efficient algorithms for langevin and dpd dynamics, *Journal of Chemical Theory and Computation* 8 (10) (2012) 3637–3649. doi:10.1021/ct3000876.
  - [44] B. Efron, Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods, *Biometrika* 68 (3) (1981) 589–599.
  - [45] J. K. Singh, D. A. Kofke, Mayer sampling: Calculation of cluster integrals using free-energy perturbation methods, *Physical review letters* 92 (22) (2004) 220601.
  - [46] A. J. Schultz, D. A. Kofke, Virial coefficients of model alkanes, *The Journal of Chemical Physics* 133 (10). doi:10.1063/1.3486085.
  - [47] N. S. Barlow, A. J. Schultz, S. J. Weinstein, D. A. Kofke, Communication: Analytic continuation of the virial series through the critical point using parametric approximants (2015).
  - [48] M. Thol, G. Rutkai, A. Köster, R. Lustig, R. Span, J. Vrabec, Equation of state for the Lennard-Jones fluid, *Journal of Physical and Chemical Reference Data* 45 (2) (2016) 23101.
  - [49] V. I. Harismiadis, I. Szeleifer, 2nd Virial-Coefficients Of Chain Molecules - A Monte-Carlo Study, *Molecular Physics* 81 (4) (1994) 851–866.
  - [50] H. W. Hatch, S. Jiao, N. A. Mahynski, M. A. Blanco, V. K. Shen, Communication: Predicting virial coefficients and alchemical transformations by extrapolating Mayer-sampling Monte Carlo simulations, *Journal of Chemical Physics* 147 (23) (2017) 1–5. doi:10.1063/1.5016165.
  - [51] DIPPR, Design Institute for Physical Property Data (2004).
  - [52] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes, *J. Comp. Phys.* 23 (1977) 327–341.
  - [53] M. Tuckerman, B. J. Berne, G. J. Martyna, Reversible multiple time scale molecular dynamics, *The Journal of chemical physics* 97 (3) (1992) 1990–2001.
  - [54] C. D. Wick, M. G. Martin, J. I. Siepmann, Transferable potentials for phase equilibria. 4. United-atom description of linear and branched alkenes and alkylbenzenes, *The Journal of Physical Chemistry B* 104 (33) (2000) 8008–8016.
  - [55] V. Shen, D. Siderius, W. Krelberg, H. E. Hatch, NIST Standard Reference Simulation Website (2008). doi:<http://doi.org/10.18434/T4M88Q>.
  - [56] J. R. Mick, M. Soroush Barhaghi, B. Jackman, L. Schwiebert, J. J. Potoff, Optimized mie potentials for phase equilibria: Application to branched alkanes, *Journal of Chemical & Engineering Data* 62 (6) (2017) 1806–1818.
  - [57] B. Eggimann, P. Bai, A. Bliss, C. Bunner, Q. Chen, R. DeJaco, E. Fetisov, D. Harwood, T. Josephson, R. Lindsey, M. Shah, H. Stern, K. Struksheats, J. Sung, A. Sunnarborg, B. Xue, J. Siepmann, T-ua no. 16 2-methylpropane, TraPPE Validation Database University of Minnesota: Minneapolis, MN. URL <http://chem-siepmann.oit.umn.edu/siepmann/trappe/validation.html>, accessed February 11, 2019
  - [58] K. M. Benjamin, J. K. Singh, A. J. Schultz, D. A. Kofke, Higher-order virial coefficients of water models, *Journal of Physical Chemistry B* 111 (39) (2007) 11463–11473. doi:10.1021/jp0710685.
  - [59] A. A. Chialvo, A. Bartók, A. Baranyai, On the re-engineered TIP4P water models for the prediction of vapor-liquid equilibrium, *Journal of Molecular Liquids* 129 (1-2) (2006) 120–124. doi:10.1016/j.molliq.2006.08.018.
  - [60] M. G. Martin, J. I. Siepmann, Novel configurational-bias Monte Carlo method for branched molecules. Transferable potentials for phase equilibria. 2. United-atom description of branched alkanes, *The Journal of Physical Chemistry B* 103 (21) (1999) 4508–4517.
  - [61] M. S. Barhaghi, J. R. Mick, J. J. Potoff, Optimised Mie potentials for phase equilibria: application to alkynes, *Molecular Physics* 115 (9-12) (2017) 1378–1388. doi:10.1080/00268976.2017.1297862.
  - [62] J. R. Mick, M. Soroush Barhaghi, B. Jackman, K. Rushaidat, L. Schwiebert, J. J. Potoff, Optimized Mie potentials for phase equilibria: Application to noble gases and their mixtures with n-alkanes, *The Journal of chemical physics* 143 (11) (2015) 114504.
  - [63] J. L. Abascal, C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005., *The Journal of chemical physics* 123 (23) (2005) 234505. doi:10.1063/1.2121687.
  - [64] B. L. Eggimann, A. J. Sunnarborg, H. D. Stern, A. P. Bliss, J. I. Siepmann, An online parameter and property database for the TraPPE force field, *Molecular Simulation* 40 (1-3) (2014) 101–105.
  - [65] G. Orkoulas, M. E. Fisher, A. Panagiotopoulos, Precise simulation of criticality in asymmetric fluids, *Physical Review E* 63 (5) (2001) 051507.