Elsevier Editorial System(tm) for Fluid Phase Equilibria

Manuscript Draft

Manuscript Number: FPE-D-18-00726R1

Title: Coexistence Calculation Using the Isothermal-Isochoric Integration Method

Article Type: VSI:Molecular Simulation

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

Corresponding Author: Dr. J. Richard Elliott,

Corresponding Author's Institution: The University of Akron

First Author: S. Mostafa Razavi

Order of Authors: S. Mostafa Razavi; Richard A Messerly, PhD; J. Richard Elliott, PhD

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 which results in less than 1 \% deviation from NIST REFPROP values for reduced temperatures of less than 0.85. Consistency is achieved between the ITIC method, the Gibbs Ensemble Monte Carlo (GEMC) method, and the 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, where relative deviations from experimental data are often quite large but 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.

S. Mostafa Razavi
The University of Akron
sr87@zips.uakron.edu

9/15/2018

Dear Editor,

We wish to submit an original research article entitled “Coexistence Calculation Using the Isothermal-Isochoric Integration Method” for consideration by Special Issue of Fluid Phase Equilibria. We confirm that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere.

This article is intended to demonstrate a method in the context of thermodynamic integration for molecular simulation of vapor-liquid equilibria for pure fluids. Gibbs Ensemble Monte Carlo (GEMC) and Grand Canonical Monti Carlo (GCMC) involve molecular insertions that become slow or infeasible below reduced temperature of 0.55. The proposed method theoretically has no lower temperature limit except the triple point. Furthermore, the proposed method can be applied with consistency across all platforms, including molecular dynamics and Monte Carlo simulators. Multiplatform inconsistency has raised important criticism lately, that could undermine the credibility of molecular simulation generally without a rigorous and robust method as proposed here.

We believe that this manuscript in appropriate for publication by Fluid Phase Equilibrai (Special Issue), because it can contribute in providing an overall view of the state of the art in computer simulation with respect to vapor-liquid equilibria calculation.

Should you select our manuscript for peer review, we would like to suggest the following potential reviewers/referees because they would have the requisite background to evaluate our findings and interpretation objectively.

- Dr. Jeffery Potoff (Wayne State University, jpotoff@wayne.edu)
- Dr. Michael Shirts (University of Colorado Boulder, michael.shirts@colorado.edu)
- Dr. Edward Maginn (University of Notre Dame, ed@nd.edu)
- Dr. Bennett Marshall (ExxonMobil Research and Engineering, bennettd1980@gmail.com)
- Dr. David Kofke (University of Buffalo, kofke@buffalo.edu)
- Dr. Andrew Schultz (University of Buffalo, ajs42@buffalo.edu)

Thank you for your consideration of this manuscript.

Sincerely,

S. Mostafa Razavi

Responses to Reviewer Comments for: FPE-D-18-00726

Coexistence Calculation Using the Isothermal-Isochoric Integration Method

S. Mostafa Razavi, Richard A. Messerly, J. Richard Elliott

Dear Ioannis,

Thanks for your extension of the time for us to respond to reviewers' comments. They were more extensive than expected and some of them required further simulations to prove the points. The reviewers' raised a number of valid points and we thank them for their thoroughness. I am never quite sure how much detail of the classical thermodynamics to include in a paper like this. Mostafa dealt with a number of these issues in his thesis but his previous analyses were based on flexible bonds whereas we have now reached the conclusion that rigid bonds are to be preferred. Therefore, we needed to reconstruct a lot of simulations in order to present a consistent perspective throughout this manuscript. We still feel that some of these details may be distracting for readers, so we added a few of them to the Supplementary Material, while moving a few into the main text that were in supporting information previously. At least one reviewer mentioned that they could not find the supporting information file, so I hope more detailed instructions can be provided to ensure that all reviewers have access to all the necessary files. If we have erred again and placed too much information in the Supplementary Material, please advise. It will be a simple matter to move whatever is wanted into the main text.

J. Richard Elliott

Response to Reviewer #1

General

This paper presents a new procedure to compute coexistence curves based on the integrating isothermal and isochoric paths. In addition to the undeniable technical merit, I especially like the fact that the presentation begins with the overall description of the method, followed by a discussion of its many details. This facilitates the reading, while providing enough information to a reader who might be interested in implementing the method. The comparisons with results from other methods show that the proposed procedure works as intended. I am definitely in favor of this paper's publication but I do have a few minor issues, as follows:

We appreciate that the reviewer is in favor of this manuscript being published with only minor revisions.

Comment #1
Section 2.4 contains the sentence: "In order to reach the densities of interest and maintain an acceptable accuracy of VLE data to a reduced temperature of 0.45, one needs at least 9 data points on the isotherm and three data points on the isochore for each saturation point." In this and other points of the manuscript, there is not a clear statement of "acceptable accuracy" (or similar terms) means? What are the tolerances used in different parts of the procedure?

We agree that this sentence is ambiguous. We modified this sentence to:

> 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. The highest temperature state points, however,

Comment 2#
Integrations along the isotherm and isochores are performed using Simpson's rule: what stops you from using, say, Gaussian quadrature, which should be more accurate for the same number of sampled points?

We use Simpson's rule mostly due to its simplicity. It is also important to the method that a fixed step size be used for the integration so that simulations can be re-used for different isochores. Gaussian quadrature varies the density states depending on the range of integration. We clarified this in the second paragraph of Section 2.4

> The integrations along the isotherm and isochores are performed using Simpson's rule [25], due 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. (18) and Eq. (19) are articu-

Comment 3#
Why use fixed-point iterations to solve equation (14)? Why not some quasi-Newton iteration, which should be potentially faster?

We would like to point out that speed of fixed-point iteration is not necessarily a deciding factor because these iterations do not require additional simulations. We mainly chose fixed-point iterations for simplicity. We now modified the sentence after Equation 14 to:

> Eq. (14) can be solved for ρ_{vap} using a root-solving approach. For simplicity, we solve Eq. (14) with the fixed-point iteration method [17]. Finally, vapor pressure is calculated using Eq. (15)

Comment 4#
Elsevier journals provide a lot of flexibility for references in manuscripts but there is a mix of capitalization or not in paper and journal titles that may lead to mistakes in the final version. I suggest a full review of the references. For example, I call attention to reference [23], in which GPU, Gibbs, Monte Carlo, and Lennard-Jonesium are not capitalized.

We thank the reviewer for thorough review of the references. We fixed reference 23 and modified several references to improve their consistency.

Comment 5#
Right after Eq. (1), it should be "… where G represents the molar Gibbs free energy, …".

We now use molar in the mentioned sentence.

Comment 6#
There is a missing "." at the end of the first paragraph of subsection 3.3.

We have fixed this problem.

Comment 7#
In the first line of the second paragraph of subsection 3.3, it should be "… method does not …"

We appreciate the reviewer pointing out this issue. We now fixed this issue.

As mentioned, these are minor points. I recommendation publication after the authors address them.

We appreciate that the reviewer recommends publication of the revised manuscript.

Response to Reviewer #2

General Comments

This manuscript proposes a new integration method for solving liquid vapor coexistence of a pure compound along a continuous path of monophasic fluid states circumventing the critical point. The method makes use of isochoric integration in the liquid state from the initial estimate to a supercritical temperature, then an isothermal expansion from liquid to vapor state and finally an isochoric temperature decrease down to the saturated vapor. An iterative loop on the investigated temperature is implemented to satisfy the requirements of phase coexistence. The method makes a combined use of molecular simulations, equations of state based on system-specific truncated viral expansions, and empirical formulae for critical pressure and acentric factor. Application examples on methane, ethane, isobutane , isohexane, dodecane, and water are presented in good agreement with earlier results from the literature.

The data shown do not allow a complete assessment of the manuscript (the Supplementary Information file mentioned by the authors was not available for this review). Information is indeed lacking about initialization of saturation temperature in the examples investigated. The origin of the equation used for updating the departure of free energy (eq. 17) when iterating on temperature is unclear. It does not seem consistent with the fundamental equation (6). It is unsatisfactory that increasing systematic deviations are observed when initial temperature estimate (Figure 7). As the purpose of the article is the method, it is necessary that for each coexistence state point computed with the ITIC method, the initial conditions (liquid density and temperature) are provided. It is also necessary that the article documents not only the iteration steps on vapor density as in figure 5, but also the iteration steps on temperature. When the coexistence curve is unknown (e.g. new compounds), the computational load required by the ITIC method is probably underestimated by the authors as they do not seem to account for the temperature loop.

Another important concern about the proposed method is its heterogeneity. At times the integration is using NIST Refprop properties to perform integration (e.g. n-dodecane in Figure 3) while in other cases a forcefield (TraPPE-UA, Mie, TIP4P-2005) is used. Some properties are obtained by simulation (energy changes along liquid isochore and supercritical isotherm) but others are obtained from a virial equation truncated to the second or third order. The critical pressure and the acentric factor are fitted (equations 27 and 28) without applying the definition of the acentric factor, but an empirical correlation developed by Lee-Kesler. It is also likely that gas pressure is evaluated with different routes in MC (molecular virial route) and in MD with flexible bonds (atomic virial). Unfortunately the authors do not document this point, which is needed to assess convergence problems in the low density limit (Figure 9).

The proposed method does not involve transfer moves and this is a significant advantage over of the Gibbs ensemble Monte Carlo (GEMC) which is delicate to code and display low acceptance rates at low temperature. However, the application range of the proposed method is rather narrow : Tr=0.45 to Tr=0.8. The performance in near critical conditions (Tr=0.8 to Tr=1) appears uncertain, because the virial-based expansions are approximate in high density vapor. As a consequence, obtaining critical coordinates by near-critical scaling is subject to important uncertainties. Application to reduced temperatures lower than 0.45 is neither considered, the main reason being probably the limitations of the packing software , Packmol, to produce liquid configurations at the desired density without molecular overlaps. It is unacceptable that the authors compare their method with GEMC only or

GCMC only. This referee is aware of numerous published studies showing that GEMC + TI cover a larger range of coexistence conditions (Tr=0.3 to 0.95) as needed in applications. The citation of Ahunbay et al (2004) by the authors is just one of these articles. A more appropriate reference would be the earlier article of the same team (Ungerer et al, JCP 2000) in which ethane and dodecane are considered, two compounds studied in the proposed manuscript.

For these reasons, the manuscript requires significant additional checks and modifications along the points listed below, before publication can be considered.

Responses to General Comments

> The data shown do not allow a complete assessment of the manuscript (the Supplementary Information file mentioned by the authors was not available for this review).

The Supplementary Information was provided as a link in the generated PDF file. We made sure that this time the reviewers have access to Supplementary Material.

> Information is indeed lacking about initialization of saturation temperature in the examples investigated.

We used the DIPPR database to get initial estimates of T^{sat}. Tables 1-10 in Supplementary Material contain the temperatures and densities of NVT simulations, as well as the simulation results. The following sentence was added to the Example Simulation section to clarify the T^{sat} initializations:

> plicity. Initial T^{sat} values for all the compounds shown in Figures 10 and 11 were obtained from the DIPPR [46] database. Complete information about chosen ITIC state points as well as the results of NVT simulations is included in Supplementary Materials.

> The origin of the equation used for updating the departure of free energy (eq. 17) when iterating on temperature is unclear. It does not seem consistent with the fundamental equation (6).

Discussed below in “detailed comments.”

> It is unsatisfactory that increasing systematic deviations are observed when initial temperature estimate (Figure 7).

According to Figure 7 even 10 % error in the T^{sat} initial guess causes only 0.1-0.4 % deviation in Psat and 0.2-0.4 % deviation in rhoL. Most force fields provide an accurate representation of Tsat. For example, TraPPE-UA and Mie-UA force fields that are used in this study, predict rhoL with less than 1 % deviation from experimental data. Therefore, in this study experimental data is used to estimate the initial Tsat values. For compounds for which experimental Tsat vs. rhoL data is not available, accurate estimation of T^{sat} is possible at each density by using a linear interpolating or extrapolating of compressibility factor vs reciprocal temperature. An example of such situation was added to the Example Simulations section.

> As the purpose of the article is the method, it is necessary that for each coexistence state point computed with the ITIC method, the initial conditions (liquid density and temperature) are provided.

The initial conditions are now provided in Supplementary Material.

> It is also necessary that the article documents not only the iteration steps on vapor density as in figure 5, but also the iteration steps on temperature. When the coexistence curve is unknown (e.g. new compounds), the computational load required by the ITIC method is probably underestimated by the authors as they do not seem to account for the temperature loop.

Using Tsat instead of rhoV as the stopping criterion resulted similar convergence behavior and computational load. It should be noted that iterations of Tsat 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. 1/T. We clarify that these iterations do not require additional simulations in the second and last paragraph of Section 2.3.

> Another important concern about the proposed method is its heterogeneity. At times the integration is using NIST Refprop properties to perform integration (e.g. n-dodecane in Figure 3) while in other cases a force field (TraPPE-UA, Mie, TIP4P-2005) is used. Some properties are obtained by simulation (energy changes along liquid isochore and supercritical isotherm) but others are obtained from a virial equation truncated to the second or third order.

We agree that this study incorporates different types of data (simulation, experimental, virial coefficient, etc.), however, the ITIC method itself relies only on simulation data. We do use DIPPR data to initialize Tsat, but in the absence of such database, those initial Tsat values can come from simple inter- or extrapolation of two points on isochore (Z vs. 1/T). In Example Simulations section, we have added a case study for an unfamiliar molecule where no experimental data are available. We used REFPROP data only to validate the ITIC method and discuss its sensitivities.

> The critical pressure and the acentric factor are fitted (equations 27 and 28) without applying the definition of the acentric factor, but an empirical correlation developed by Lee-Kesler.

Discussed below in “detailed comments.”

> It is also likely that gas pressure is evaluated with different routes in MC (molecular virial route) and in MD with flexible bonds (atomic virial). Unfortunately the authors do not document this point, which is needed to assess convergence problems in the low density limit (Figure 9).

Discussed below in “detailed comments.”

> The proposed method does not involve transfer moves and this is a significant advantage over of the Gibbs ensemble Monte Carlo (GEMC) which is delicate to code and display low acceptance rates at low temperature. However, the application range of the proposed method is rather narrow : Tr=0.45 to Tr=0.8. The performance in near critical conditions (Tr=0.8 to Tr=1) appears uncertain, because the virial-based expansions are approximate in high density vapor. As a consequence, obtaining critical coordinates by near-critical scaling is subject to important uncertainties.

In the absence of B3, calculating critical properties in ITIC method is indeed subject to important uncertainties. We now modified the following paragraph in the Section 7 (Example Simulations)

> 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 $T_r^{sat} \approx 0.85$, 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. [60]. Accurate estimation of critical points by molecular simulation requires a careful and deliberate effort that includes accounting for system size effects [61]. That effort goes beyond the scope of the current manuscript.

> Application to reduced temperatures lower than 0.45 is neither considered, the main reason being probably the limitations of the packing software, Packmol, to produce liquid configurations at the desired density without molecular overlaps.

Discussed below in “detailed comments.”

> It is unacceptable that the authors compare their method with GEMC only or GCMC only. This referee is aware of numerous published studies showing that GEMC + TI cover a larger range of coexistence conditions (Tr=0.3 to 0.95) as needed in applications. The citation of Ahunbay et al (2004) by the authors is just one of these articles. A more appropriate reference would be the earlier article of the same team (Ungerer et al, JCP 2000) in which ethane and dodecane are considered, two compounds studied in the proposed manuscript.

Discussed below in “detailed comments.”

Responses to Detailed Comments of Reviewer 2:

Section I. Introduction

> 1.1 There is no reason that coexistence calculations must extend to a reduced temperature of 0.45 only. The industrial needs extend over the whole liquid-vapor coexistence curve down to the triple point and this may be down to Tr=0.30 for many compounds (eg isooctane, toluene, heavy iso-alkanes, …). The Peng-Robinson EOS loses accuracy between Tr=0.3 and Tr=0.45. The current text is misleading. The text must indicate that prediction methods are required down to Tr=0.3 and that the PR EOS is not accurate in the range Tr=0.3 to 0.45.

We modified the text to emphasize that the reduced triple point temperature may go as low as 0.3 in some cases.

[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.

> 1.2 The authors state about Gibbs-Duhem integration implemented by Ahunbay et al. [FPE, 2004, vol 224, page 73] that it does not predict vapor density. This statement is untrue, as vapor density is always an output of GEMC calculations above the normal boiling point. Below the normal boiling point, the GEMC +TI method considers that vapor density can be approximated by the ideal gas law. If a more accurate vapor density is desired, NPT simulation of the vapor phase may also provide vapor density a posteriori, using saturation pressures obtained by TI. The text must acknowledge that GEMC + TI allows computation of vapor density.

We changed the text to avoid this confusion:

tion line starting from one single coexistence point. 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] have applied this approach in conjunction with GEMC to obtain the initial coexistence point, *NPT* simulations to estimate saturation liquid densities, and parallel tempering method [9] 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.

> 1.3 The citation of Ahunbay et al (2004) by the authors is just one of the numerous articles applying the GEMC +TI method. A more appropriate reference would be the earlier article of the same team (Ungerer et al, JCP, 112, (2000), 5499) in which ethane and dodecane are considered, two compounds studied in the proposed manuscript. The text must acknowledge that GEMC + thermodynamic integration has been applied to the prediction of coexistence of large alkanes (eg dodecane) in a large range of conditions (Tr= 0.45 to Tr=0.9) using several forcefields (including TraPPE-UA).

We thank the reviewer for the Ungerer2000 reference. This paper does have TraPPE-UA dodecane, but it does not have TraPPE-UA ethane. We now included TraPPE-UA dodecane data from Ungerer2000 to Figure 10 (unfilled black triangles). Reasonable agreement between ITIC and GEMC+TI is observed in Figure 10.

Section II - Integration method

> 2.1 Figure 1 is drawn as if the coexistence curve is already known, i.e. a case when most of the problem is solved. This is not the general case as the proposed method includes an iterative process to

find the saturation temperature. The figure must be modified to show the general case when the initial estimate of the saturation temperature is different from the effective saturation temperature.

The purpose of this figure is to schematically show the location of ITIC points relative to the coexistence curve. The reviewer is correct that the saturation temperature is not known a priori and, therefore, the simulation temperature is not identical to the actual coexistence temperature. To clarify this point, the caption now reads, “circle points 11, 13, 15, 17, and 19 represent the initial estimate of saturation temperature, therefore they do not necessarily match the coexistence curve.” To further clarify this, we have added to the Example Simulation section, a new case study for a rare aromatic molecule for which no experimental information is available.

> 2.2 In equation 8, the variable Aig is the free energy of the ideal gas state. The text must specify that Aig is defined in the same temperature and volume conditions as the liquid. It is thus an imaginary state.

We added the subscript TV to the left hand side of Eq(8)

$$\left(\frac{A-A_{ig}}{RT}\right)_{TV}=\int\frac{U-U_{ig}}{R}\mathrm{d}\left(\frac{1}{T}\right) \qquad (8)$$

> 2.3 The origin of the equation (17) used for updating the departure of free energy when iterating on temperature is unclear. Equation (17) does not seem consistent with the fundamental equation (6). The authors must check or correct the derivation of this equation. If Equation (17) is corrected, the authors must redo a significant fraction of their coexistence calculations and investigate whether the behavior of Figure 7 is still observed.

Eq. 17 is using trapezoid rule to integrate $U^{dep}T$ vs. 1/T (i.e. $(U-U^{ig})/R$ vs. 1/T) according to Eq. 8 and calculate A^{dep} at the new T^{sat}. Since $U^{dep}T$ vs. 1/T is linear (See Fig 1-10 in Supplementary Material) within proximity of T^{sat}, using trapezoid rule results in good approximation of A^{dep}. This explanation is now included before Eq.(17).

of T^{sat} has changed, A^{dep} should also be recomputed. Trapezoid rule is used to integrate $U^{dep}T$ vs. $1/T$ (i.e. $(U-U^{ig})/R$ vs. $1/T$ in Eq. 8) due to the linear shape of $U^{dep}T$ vs. $1/T$ in the vicinity of T^{sat}.

$$A^{dep}_{new}=A^{dep}_{old}+\left(\frac{1}{T^{sat}_{new}}-\frac{1}{T^{sat}_{old}}\right)\frac{U^{dep}_{new}T^{sat}_{new}+U^{dep}_{old}T^{sat}_{old}}{2} \qquad (17)$$

> 2.5 Equations 27 and 28 may not be consistent with the fundamental definition of the acentric factor from the force field considered ; the parameters of equation 28 have been derived by Lee and Kesler

based on a compilation of experimental data about different fluids comprising light n-alkanes. The authors must recall the usual definition of the acentric factor and explain that equations 27-28 are an approximation that may not be relevant for every compound.

We admit that predicting Pc with the Lee-Kesler equation is not common in simulation studies. The standard approach, for practically any system, is to fit simulation data to the Antoine equation. For example, even Pc for TIP4P/2005 is well represented with the Antoine equation (J. Chem. Phys. 125, 034503 (2006); doi: 10.1063/1.2215612). For this reason, we reprocessed all of the ITIC data using the more traditional Antoine equation approach. Interestingly, we found that the Antoine Pc values were statistically indistinguishable from their Lee-Kesler counterparts for the systems studied here, with the exception of TIP4P/2005.

Figure 10 has been updated with the Antoine Pc values while Equations 27 and 28 have been replaced by the Antoine equation. This section also includes the following recommendation for computing critical constants with ITIC.

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

$$\log_{10}(P^{sat}) = a_0 + \frac{a_1}{a_2 + T^{sat}} \quad (27)$$

where a_i are fitting constants and a_1 is constrained to be negative. The second step is to evaluate Eq.(27) with the optimal values of a_i for $T^{sat} = T_c$.

Some caution should be exercised when extrapolating Eq.(25-27) as 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 Eq.(25-27) 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 Eq.(25-27) provides reasonable estimates for all three critical constants.

Section III - ITIC validation

> 3.1 In the case of n-dodecane illustrated by Figures 3 , 4 and 5, the manuscript must provide the initial estimates of the saturation temperature for the state points considered.

T_{est}^{sat} values are now listed in Supplementary material (Table XXI and XXII)

> 3.2 In Figure 4, the saturation temperature seems identical for the two series (2nd order virial and 3d order virial). However very different solutions are found for vapor density and vaporization enthalpy. In other words there is a satisfactory convergence of saturation temperature in both series while vapor density and vaporization enthalpy diverge for the 2nd order virial series. Is there an explanation to this

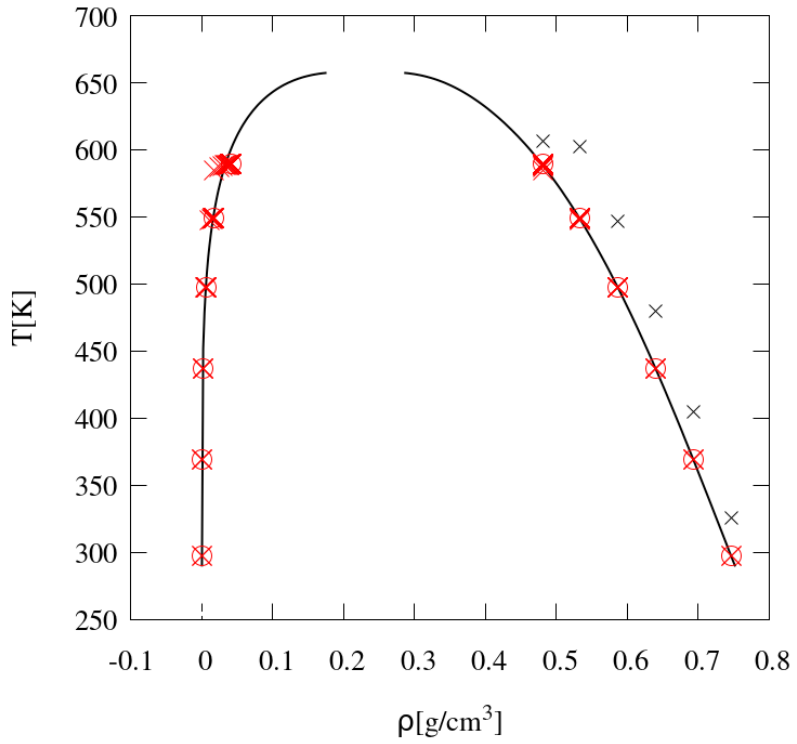
behavior ? The authors must indicate the iterative sequence of saturation temperature followed in both series.

As mentioned in paragraph 2 of Section 2.3, Tsat is determined by extrapolating or interpolating Z along the isochore, so it is fairly insensitive of the virial coefficients. The only impact the virials have on Tsat is the value of Zliq (from Psat and rhovap), but this dependence is very small as Zliq is always around 0. We added a paragraph to the end of Section 3.2, to explain our observation of Tsat sensitivity to virial coefficients.

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

> 3.3 In Figure 5, the iterative process on vapor density is abundantly illustrated but not the iterative process on saturation temperature. This figure must be complemented by a specific graph showing how saturation temperature converges to the solution when initialized some distance away from the coexistence curve.

The below figure is a modified version of Figure 3(b) by initializing the ITIC calculation away from coexistence curve. As can be seen, Tsat nearly converges in one iteration and the sequence of (rhoV,Tsat) represented by red X symbols are practically identical even when initial Tsat estimate (black X) is deviated from true Tsat by ~10 %. We added this figure to the Supplementary Material to show how Tsat converges when it is initialized away from coexistence curve.



> 3.4 In Figure 6 showing deviations on saturation pressure, how is it possible that round values of Tsat are mentioned in legend ? As Tsat is the result of an iterative process it is unlikely that round values are obtained. What is the explanation of the authors of this ? The legend must be changed to indicate the imposed variable (liquid density). The manuscript must also indicate whether saturation temperature is significantly changed when the virial coefficients are modified.

We believe that using equal sign in the legend adds to confusion. We also agree that it is less confusing if the legend includes the imposed densities, however we have based our comparisons on Tr values, because the readers are more familiar with temperature than density. We now changed the legend and used "≈" instead of "=" to emphasize that Tsat values are not imposed. We also plotted rhoV and Tsat sensitivity vs B2, B3, and B2_IT deviation. We added this paragraph to explain our observations:

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

> 3.5 In Figure 7 it is unsatisfactory that increasing systematic deviations are observed when the initial saturation temperature is estimated at increasing distance from the true solution. This behavior may reveal a problematic convergence of the iterative scheme on Tsat (see point 2.3).

Figure 7 presents the error incurred by deviations in the simulated Tsat^est. This figure is not related to the iterative scheme for computing Tsat that is mentioned in point 2.3.

The deviations in Figure 7 are mainly due to lack of extrapolation precision when Tsat is away from real Tsat values. As we discussed in general comments, according to Figure 7 even 10 % error in Tsat initial guess causes only 0.1-0.4 % deviation in Psat and 0.2-0.4 % deviation in rhoL. This amount of error is satisfactory compared to typical (GEMC/GCMC) simulation uncertainty. Tsat can be accurately estimated for most force fields by using experimental values, because typical force fields characterize Tsat-rhoL relationship within less than 1 % error. If experimental values cannot be found for the molecule of interest a simple two point extrapolations on corresponding Z vs. 1/T plot can give an accurate estimation for Tsat at a given density (See last two paragraphs of Section 7)

> 3.6 Did he authors investigate other accuracy requirements on vapor density than 0.1% ? Is it possible to reach a better accuracy ?

We use 0.1 % as the stopping criterion for iteration. In our experience setting this number to a even smaller number does not improve the accuracy. We now mention this in Section 2.3.

Section IV - Simulation details

> 4.1 It is surprising to read that "the MC method with fixed bond lengths is favored due to lower fluctuations at low density" . Indeed fluctuations are characteristic of the NVT ensemble (e.g. energy fluctuations are related with heat capacity) and should not depend on the simulation engine used to generate the ensemble. The statement must be removed or carefully reformulated.

We show in Figure 9 that at low densities on IT, MC has smaller pressure uncertainties than MD and fixed-bond MD has smaller uncertainties than harmonic-bond MD. We changed "fluctuations" to "uncertainties" to avoid confusion.

> 4.2 In the paragraph devoted to the computational cost of the ITIC method, how is it possible that the authors apply the ITIC method at imposed Tsat = 0.65, 0.75 and 0.85 in reduced units, as the method consists in imposing liquid density and finding Tsat ? This sentence must be reformulated.

We agree that this sentence is not clear enough. We are using Tsat as a basis to compare ITIC range and GEMC range, because it is easy to understand for the reader. We are also aware that Tsat in ITIC is calculated and Tsat in GEMC is imposed. Therefore, we changed

> In order to approximate the computational cost of 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 0.65, 0.75, and 0.85. This

> 4.3 When evaluating the computational cost of ITIC versus GEMC the authors seem to underestimate the computation time required by successive iterations on saturation temperature in ITIC.

We would like to emphasize that NVT simulations are simulated only once, and the Tsat iterations do not require new simulations. When the simulations are done, the method converges quickly after 3-50 iterations, therefore we did not discuss the negligible computational cost of ITIC analysis. This explanation is now added to the second and last paragraph of Section 2.3.

> 4.4 the final statement "for reduced temperatures of 0.45 and 0.55 … GEMC is not feasible for n-dodecane" is misleading, because simple monophasic NPT simulations allow to extend GEMC coexistence curve calculations down to reduced temperatures of 0.45 for this compound (see point 1.3 above). Comparing ITIC with GEMC only is irrelevant. The authors must compare instead with GEMC + thermodynamic integration.

In this paragraph we are trying to prove that ITIC is comparable to GEMC with respect to computational cost. In this study, we do not intend to compare ITIC to Gibbs-Duhem. Furthermore, Gibbs-Duhem can be used for ITIC as well, so it seems to be fair to compare ITIC and GEMC methods. However, it is fair to briefly mention that GEMC can be extended to lower Tsat's if Gibbs-Duhem is used. So we changed this paragraph to:

> 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 ≈ 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.

> 4.5 The initialization of NVT simulations of the liquid phase is a delicate task that the authors achieve with the Packmol software. What is the lower reduced temperature allowing safe initialization by Packmol and subsequent simulation by Cassandra ? The authors must discuss briefly this point.

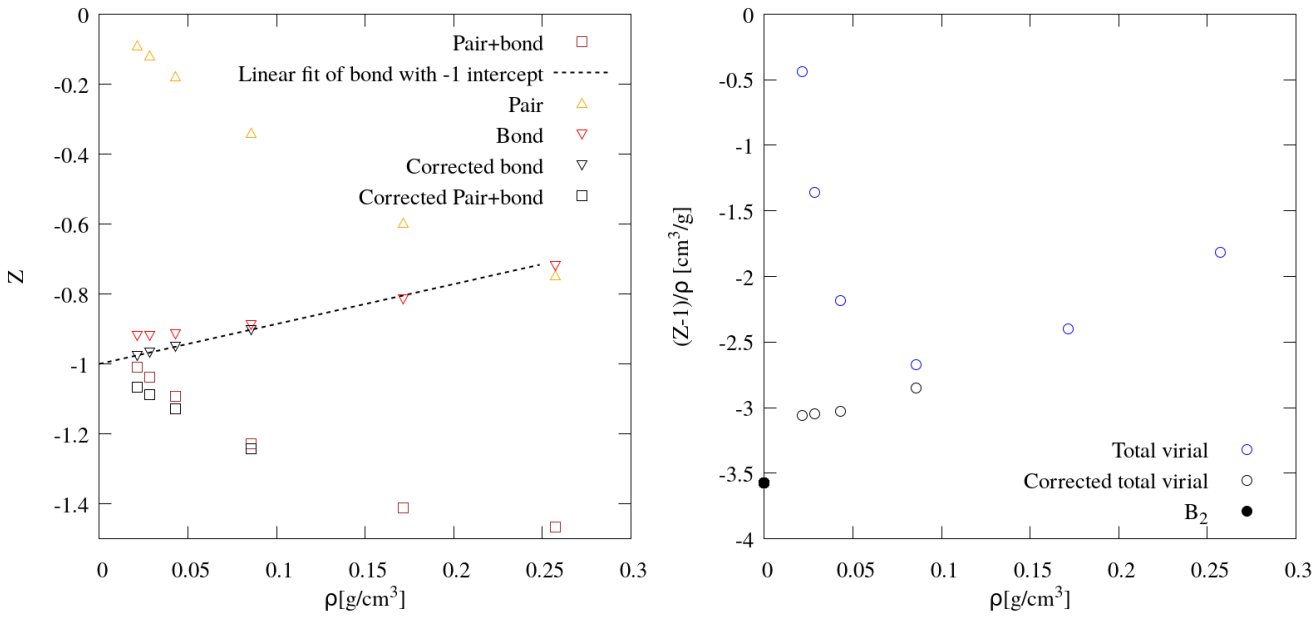
As mentioned in Simulation Details, the Packmol was used only for GOMC and LAMMPS simulations. Cassandra and Gromacs have a a powerful internal tool that performs initialization.

> The Packmol [30] software is used to create the initial configurations for LAMMPS and GOMC simulations, while Cassandra simulations were initialized using internal capabilities of this software. The simulation boxes contain 1200 sites ex-

Section V Calculation of virial coefficients Section VI Finite size effects

> 6.1 It seems to this referee that Figure 9 is illustrating the divergence of the pressure determination in the zero density limit when the molecule has internal degrees of freedom. This behavior is frequent when the atomic virial route is used for pressure determination because the atomic virial expresses as $PV = nkT + W$, where n is the number of atoms (not the number of molecules!) and the virial term W contains contributions from internal forces and from intermolecular forces. Depending on the numerical approximations in computing intra- and intermolecular forces in the atomic virial, Z may not tend toward 1 in the zero density limit. MC predicts better the compressibility factor Z in low density phases, because MC codes are generally using the molecular virial route in which $PV = NkT + W'$ where N is the number of molecules (not the number of atoms) and where the virial W' contains contributions from intermolecular forces only. Thus Z tends toward 1 in the low density limit. If this interpretation is correct, there is no reason to use a large number of molecules for vapor phases in MC. The authors should discuss these points more thoroughly. They may also consider merging the section "Finite size effects" with "Simulation details" .

In light of the reviewer’s comments, we examined the virial contributions for ethane in greater detail, as shown in the figure below. It is clear from the figure that the intermolecular contribution to the virial is approaching zero as expected, but the bonded contribution deviates significantly from its steady approach to -1 as the density approaches zero. The discrepancy is magnified in the computation of $(Z-1)/\rho$ owing to the division by density. The bonded contribution to the virial is a quantity reported by LAMMPS with little recourse for user intervention. Presumably, the problem is the small number of intermolecular collisions at low density relative to the large number of intramolecular collisions, inhibiting the equilibration of the various components of momentum. To illustrate one manner of correcting for this deficiency, we used a linear extrapolation of the intramolecular virial, enforcing a value of -1 at zero density. Then we recomputed the total virial. As shown below, this procedure reduces the problem, but requires considerably more effort than switching to fixed bond lengths.



We include this figure with a detailed explanation in the Supplementary Material. We also briefly mention the nature of the correction in Section 6.

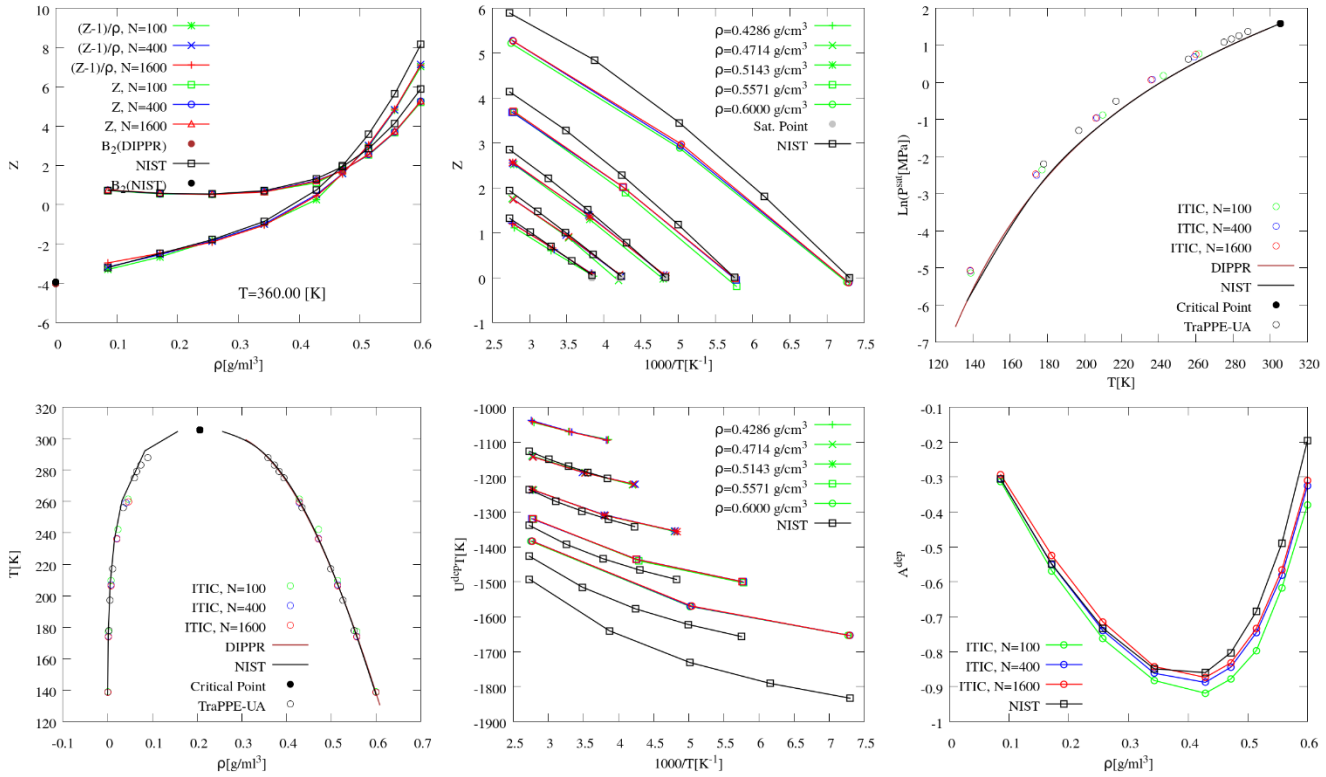
Furthermore, we agree with the reviewer that the section name is not appropriate, since we also investigate MC vs MD and fixed vs flexible bonds. For this reason, we have modified the section name to, “Accurate Low Density Simulations.”

The effect of using flexible bonds is shown in Figure 9(b). The systematic discrepancy from rigorous values (solid black line) as well as large uncertainties suggests not using flexible bonds at very low densities. The divergence of pressure at low densities is presumably due to the small number of intermolecular collisions at low density relative to the large number of intramolecular collisions, inhibiting the equilibration of the various components of momentum. In Supplementary Materials, we illustrate one manner of correcting for this deficiency by modifying the intramolecular virial when different virial contributions are separately accessible.

Figure 9(c) shows the low density *NVT* state points simulated using GOMC [25]. 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.

> 6.2 Finite size effects on the compressibility factor Z and on the internal energy U are not discussed for dense (liquid) phases. Packing effects are sometimes important in such conditions. The authors should briefly mention the steps taken in their method to avoid significant finite size effects in the simulation of liquid phases.

It is true that finite size effects can impact dense phases. We investigated this impact and found no clear trend to report. We include this figure in Supplementary Material but recommend a more thorough investigation in future work.



Section VII Example simulations

> 7.1 : Figures 10 and 11 : The agreement of ITIC with previous simulation results on methane, ethane, isobutane , isohexane, dodecane and TIP4P/2005 water is very good ; In order to asses the relevance of this comparison it is necessary to mention how far the initialization (vapor density and saturation temperature) was from the final coexistence curve. The authors must add this precision in text or in tables. A few cases of significant differences (e.g. TIP4P water at 473K) merit specific explanation from the authors.

All ITIC results of example simulations are initialized at DIPPR Tsat values. In Supplementary Material, we include a set of ethane simulations with deviated Tsat values which shows that final coexistence curve is weakly influenced by initial Tsat values.

> 7.2 Figure 11 : Dodecane has been already investigated by GEMC + thermodynamic integration with TraPPE-UA [Ungerer et al. , JCP, 112, (2000), 5499] in a larger temperature range than the GEMC results shown in this graph. The authors must include the related GEMC+TI results in Figure 11 a.

We added Ungerer2000 results for TraPPE-UA n-dodecane to Figure 10 (See unfilled black triangles), which show that ITIC is in relatively good agreement with GEMC+TI for n-dodecane at low temperatures.

Section 8. Conclusions

> 8.1 In the conclusion as in the rest of the manuscript, the authors are considerably underestimating the application range of "classical" molecular simulation methods (Tr=0.3 to 0.95) , which is in fact larger than the ITIC proposed here (0.45 to 0.85). Restricting comparison to GEMC alone or GCMC alone should be avoided in the conclusion as in the rest of the article.

The conclusion has been modified to say,

> 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

> 8.2 The authors may insist on a very significant advantage of their method, i.e. it avoids the implementation of transfer moves in Gibbs Ensemble Monte Carlo simulations, which is always a delicate task involving statistical bias and configurational bias.

We added a statement in the Conclusions to emphasize the need for transfer/insertion moves in conventional methods.

> simulate systems at temperatures as low as $T_r = 0.45$. 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. The ITIC method, hence, outperforms GEMC and GCMC when T_r is less than 0.6. This method,

Furthermore, an important advantage of the ITIC method is the possibility of calculating saturation properties using MD. We emphasize on the cross-platform feature of ITIC in the Conclusions.

Response to Reviewer #3

General

The manuscript describes a method predict phase behavior using NVT simulations along two isotherms and several isochores based on previous work to do the same for a square-well fluid. The method utilizes estimates of virial coefficients to characterize the vapor phase. The manuscript is well structured and provides comparisons with GEMC and GCMC predictions of phase behavior. The method has potential to be a useful route to obtain coexistence, but there are a number of issues that need to be addressed before publication.

Comment #1

> A general comment is that the manuscript often notes how much results differ from REFPROP values ("reproduces NIST REFPROP vapor pressure within 1 % deviation" in the conclusions), with the implication that this indicates the method is working well. This isn't quite right since REFPROP data describes the behavior of real molecules and may differ (by more than 1%) from the properties of the models considered. Validating the method is then supported by comparing against GCMC and GEMC while agreement with REFPROP primarily indicates the quality of the model.

Section 3 (ITIC validation) focuses on understanding the convergence behavior of the ITIC method, sensitivity analysis, and more importantly assessing the accuracy of the formulas discussed in Section 2. Reproducing the REFPROP Psat values using ITIC state points obtained from REFPROP is really a test of the spacing of the quadrature points since all the REFPROP thermodynamics are derived from NIST's analytical equation of state. It does not necessarily mean that the method will work for simulation data as well because the simulation data include uncertainties whereas the REFPROP equation of state does not. Therefore, extensive simulation results are provided in Section 7, in order to complement the validations. To clarify, the following sentence was added to Section 3.1 on Page 5

"Reproducing the REFPROP Psat 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."

Comment #2

> The authors rightly point out that the method has several advantages over GEMC and GCMC (working at low temperature, not requiring insertions and being possible with MD). However, GDI does not have any of those restrictions, only requiring the coexistence properties at one point. If a point is known (perhaps from GEMC/GCMC), does the method ITIC approach offer advantages over using GDI to get the rest of the coexistence points?

We have not compared the performance of ITIC against GDI from a simulation cost perspective. Assuming that these two methods have similar costs, one advantage of ITIC method is that it generates valuable data along the paths of integration, as mentioned in the Introduction. This information could be particularly valuable in the process of force field development. Also, with combination of derivative properties as in the work of Lustig et al (2015), it should be feasible to make use of this data to generate high accuracy multi parameter equations of state, similar to the work of Thol et al (2015). Another advantage of the ITIC method is that it is more homogeneous compared to GDI, as it only uses one type of ensemble (NVT). We added a sentence to the Introduction to explain these advantages.

temperatures. In addition, with combination of derivative properties from the work of Lustig et al [13], it should be possible to utilize these data in order to generate high accuracy multi-parameter equations of state [14, 15, 16, 17, 18, 19, 20, 21].

Comment #3

> The manuscript states in the introduction that it is necessary to avoid a phase transition in the path between coexisting phase. While it is not a bad idea to avoid such transitions where it is practical, the GCMC approach which the authors compare against does exactly this by design. The authors may not wish to pursue such options, but it doesn't seem necessary to dismiss them entirely unless they have actually tested it and found it to fail.

The IT at Tr<1 would be very scattered between rhoVsat and rhoLsat, and unlikely to be very accurate. We thank the reviewer for this suggestion, but exploring this option would be beyond the scope of the current manuscript.

Comment #4

> The derivation of the working equations for vapor and liquid free energies and for phase coexistence seems to be unnecessarily complicated, a bit confusing and perhaps not quite correct. The primary issue seems to be due to trying to write everything in terms of departure functions (deviations from ideal gas properties).

We address this comment in combination with comment #5.

Comment #5

> Eq. 2 is correct, but only if both G_ig correspond to the same coexistence pressure (rather than the corresponding liquid and vapor densities). This is unclear from the equation itself although it might be inferred from the discussion of eq 3. Still, it is not easy to get to eq. 3 from 2.

We modified Eq.(2) to emphasize the constant T and P conditions. The derivation of Eq.(3) from Eq.(2) is explained in Ref[16]. It should also be noted that the adjustment to go from the actual volume reference to the actual pressure is a term of ln(Z) ie. $(G\text{-}G^{ig})_{TP} = (G\text{-}G^{ig})_{TV} - RT\ln(Z)$. Since that term appears on both sides of the equation relating liquid fugacity to vapor, it can be expressed as $\ln(Z_V/Z_L) = \ln(\rho_L/\rho_V)$ because the T and P are the same in numerator and denominator. This reverts that pressure dependency into the term ρ_V, since ρ_L is exactly specified.

$$\left(\frac{G_{\text{liq}} - G_{\text{ig}}}{RT}\right)_{T,P} = \left(\frac{G_{\text{vap}} - G_{\text{ig}}}{RT}\right)_{T,P} \qquad (2)$$

Comment #6

> Eq. 7 and 8 are written as indefinite integrals, but they not useful in this form. Eq. 8 is used to construct Eq. 10, where it becomes a definite integral. It would be simpler to write them as definite integrals from the start.

I would suggest that the derivation would be simpler and less confusing if the derivation targets full free energy

A(rho,T^IT) = A_id(rho,T^IT) + integral[rho'=0,rho (d(A-A_ig)/drho) drho']
beta A(rho,T) = beta A(rho,T^IT) + integral[beta'=beta^IT,beta (d betaA/dbeta) dbeta']

Splitting the free energy into ideal and non-ideal parts is necessary for density integration (because the full integrand would diverge at rho=0), but it's unclear what it accomplishes for the temperature integration, except perhaps to make the results incorrect: the authors later (in section 4.1) describe subtracting out the intramolecular energy (taking U_ig = U_bond + U_intra).

$$U^{\text{dep}} = \frac{E^{\text{tot}} - E^{\text{bonded}} - E^{\text{intra}}}{NRT} \qquad (29)$$

This seems wrong. The U_bond and U_intra contributions in a simulation at some finite density are not the energy an ideal gas would experience. The ideal gas contribution might instead be determined from a simulation of a single molecule where U_ig = U_tot = U_bond + U_intra.

Stated another way, the path from coexisting vapor to coexisting liquid might be represented as

$$(\rho_V, T^{\text{sat}}) \quad ==1==> \quad (0, T^{\text{sat}}) \quad ==2==> \quad (0, T^{\text{IT}}) \quad ==3==> \quad (\rho_L, T^{\text{IT}}) \quad ==4==> \quad (\rho_L, T^{\text{sat}})$$

Rather than explicitly accounting for the change in ideal gas free energy in stage 2, the authors instead attempt to avoid it entirely by subtracting it out in stage 4. But the energy used for this purpose seems not to be the ideal gas energy (as would have been seen in stage 2), but the intramolecular (including bonding) energy seen in stage 4. I don't see how that can be right. If the authors are computing ideal gas energy for stage 4 (using a single molecule in a box), the they should state this explicitly. And, in practice, the whole thing would be simpler if the ideal gas contribution to the full free energy were included along (in stage 2) instead of trying to subtract the same from stage 4.

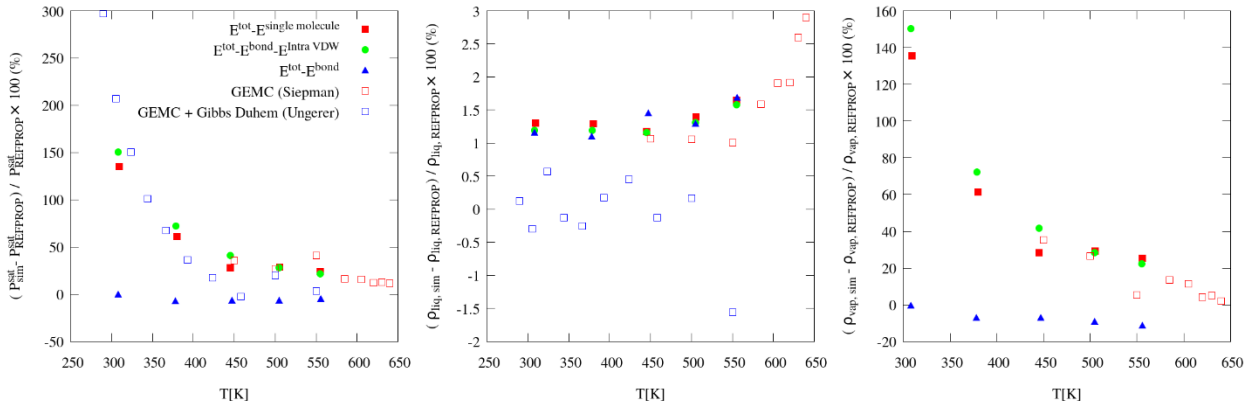
$$\left(\frac{A_{\text{liq}} - A_{\text{ig}}}{RT}\right)_{T,V} = \int_0^{\rho_{\text{liq}}} \frac{Z-1}{\rho}\, d\rho\,|_{T^{\text{IT}}} + \int_{T^{\text{IT}}}^{T^{\text{sat}}} \frac{U - U_{\text{ig}}}{R}\, d\left(\frac{1}{T}\right)|_{\rho_{\text{liq}}} \qquad (10)$$

The reviewer raises a point that has worried us for several years. We agree wholeheartedly with the reviewer's statement that "U_bond and U_intra contributions in a simulation at some finite density are not the energy an ideal gas would experience." After following the reviewer's recommendation of performing single molecule simulations to determine the ideal gas contribution, we did find some significant improvement for large molecules while smaller molecules were unaffected. This makes intuitive sense as a large molecule in a condensed phase cannot sample the same intramolecular configurations as a large molecule at zero density (single molecule simulation), while a smaller molecule samples similar intramolecular configurations at both finite and zero density.

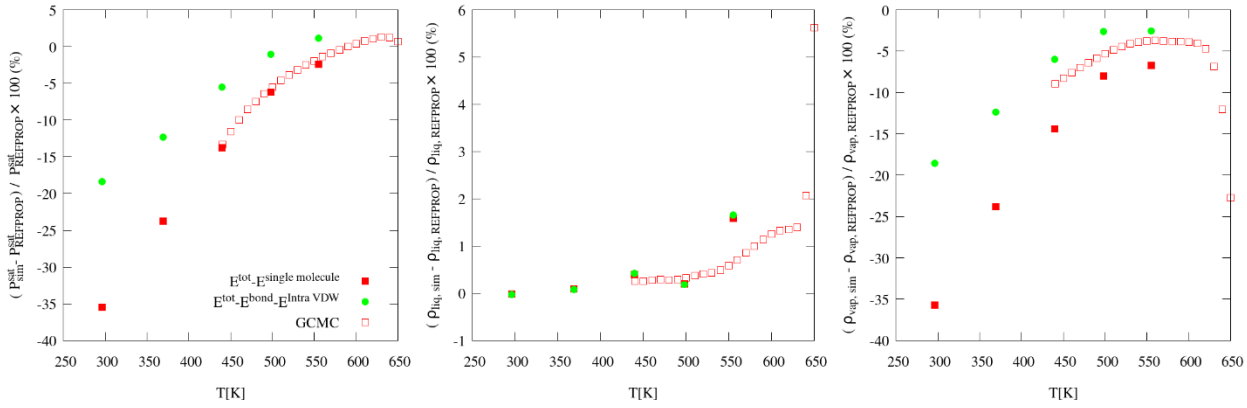
The following figures (included in Supplementary Material) help demonstrate the need for this subtle (but important) correction. These figures compare the ITIC results (filled symbols) with GEMC results from the literature (open symbols). The different shapes/colors for ITIC correspond to different ways

for computing Udep. These figures compare deviations from REFPROP values as a baseline to make the magnitudes of the discrepancies more clear. In other words, the best ITIC method is the one that agrees with the open symbols (GEMC), not the one that has lower deviation from REFPROP.

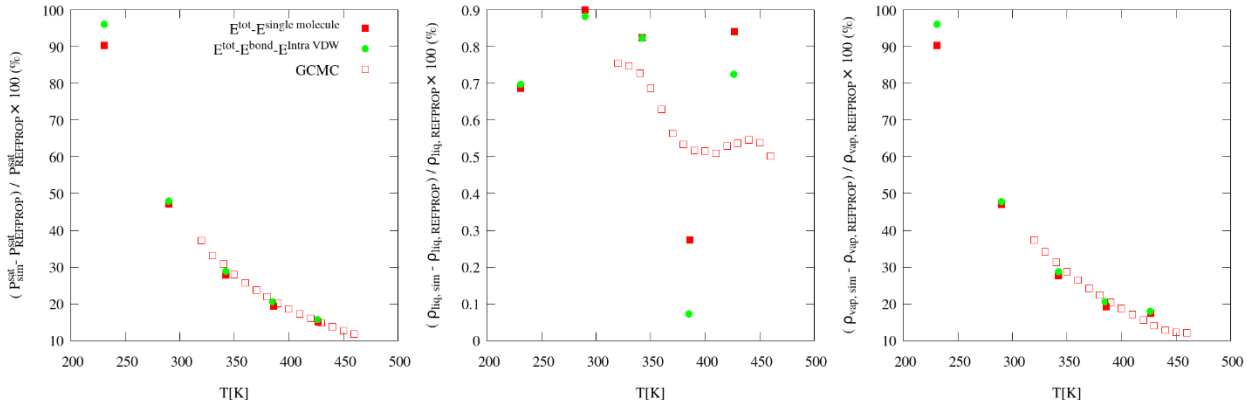
TraPPE n-dodecane:



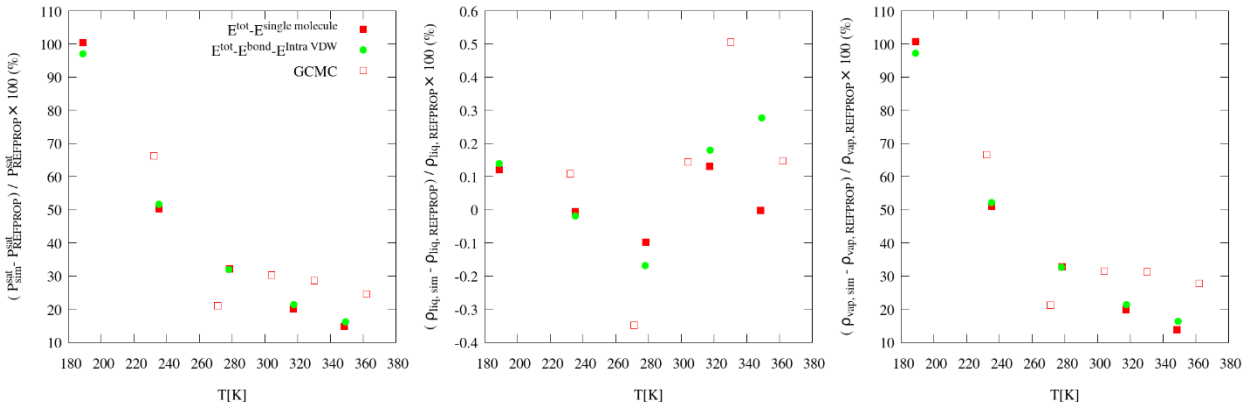
Mie n-dodecane:



TraPPE isohexane



TraPPE isobutane



The red filled squares (ITIC single molecule method) provide the best agreement in Psat and rhov for n-dodecane (C12). However, the green filled circles (ITIC original intra method) provide indistinguishable values for smaller molecules, e.g., isobutane (iC4). The blue filled triangles (ITIC without subtracting intra method) is simply wrong and will not be discussed further.

The difference between uDep calculated using the single molecule approach is on average around 1.7 % for C12. This small difference causes a significant deviation in Psat and rhov which increases with decreasing temperature, while the rhol values are essentially the same. Psat differences for C12 and iC6 are 10-15 % for Tr=0.45 and 1-4 % for Tr=0.85.

The improvement with the single molecule method is most evident for the Mie-C12 results, where the single molecule method completely resolves the discrepancy between the ITIC and GCMC Psat values. However, note that the single molecule method did not reduce the deviation between the ITIC and GCMC rhov values.

It is also important that the difference between the single molecule method and the original (intra) method for TraPPE-C12 is of a similar magnitude as the difference between GEMC and GEMC+Gibbs Duhem from the literature. Therefore, the deviations introduced by our original assumption are still less than the statistical uncertainty in the simulation data.

In brief, the single molecule method is clearly the most rigorous approach. However, the original (intra) method is not obsolete as it has some benefits for smaller molecules compared to the single molecule method. For example, although the additional single molecule simulations are extremely fast, this adds to the complexity of the ITIC method. Furthermore, single molecule simulations are ill-suited for traditional molecular dynamics simulations where a thermostat couples many degrees of freedom to a single bath, e.g., Nosé-Hoover. Stochastic dynamics (which is available in some molecular dynamics packages, e.g., GROMACS) is better suited for single molecule simulations, but this again adds complexity that is not needed for smaller molecules. For this reason, we did not remove the original (intra) method from the manuscript.

We have made the following modifications to address the reviewers well-founded concerns. Section 4.1 now recommends the single molecule method but presents the original (intra) method as an alternative for smaller molecules. A more detailed comparison is included in Supplementary Material. The C12 and iC6 ITIC results in Figures 10 and 11 and the 1-naphthalenyl,4-phenanthrenyl butane results in

Figure 12 have been recomputed with the single molecule method. The figure captions and discussion in Section 7 clearly denote this distinction.

4.1. Internal Energy Departure Function Calculation

Computing the internal energy departure function in Eq. (10) requires estimating the ideal gas energy (U_{ig}) at each temperature simulated along a given isochore. The most rigorous approach to determine U_{ig} is by simulating a single molecule system at the corresponding temperature, such that

$$U^{dep} = \frac{E^{tot} - N \times E^{N=1}}{NRT} \tag{28}$$

where E^{tot} and $E^{N=1}$ are the total potential energy and the single molecule potential energy, respectively. We recommend Monte Carlo or stochastic dynamics to compute $E^{N=1}$. 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 [31].

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_{ig} is approximately equal to the intramolecular energy at the isochoric state point, such that

$$U^{dep} = \frac{E^{tot} - E^{bonded} - E^{intra}}{NRT} \tag{29}$$

where E^{bonded} and E^{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^{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^{intra} causes a significant error in vapor pressure.

We emphasize that Eq. (29) 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^{sat} are between 10% and 20% when applying Eq. (29) to larger molecules, where the deviation increases with decreasing T^{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^{dep} (i.e., Eq. (28)-(29)) is discussed in Supplementary Material.

Comment #7

> The manuscript states that eq. 14 "can be solved for ρ_V using" using an iterative technique. It does not seem to make sense to say that eq. 14 is solved, but rather eq. 14-17, as described later.

The iterative technique for solving equations 14-17 is different than the root-solver used for solving Eq 14 for rhov. A root-solver is necessary because rhov is found in both the left-hand-side and the right-hand-side of Eq 14.

Comment #8

> Eq. 14 imposes equality of Gibbs free energy between the phases while eq. 16 (indirectly) imposes equality of pressure. It would be helpful to identify that it accomplishes this and perhaps relabel P^sat to P_vap from eq. 15 to make it clearer.

We agree that using P_{vap} in Eq.(15) makes the equation more clear. Instead of removing P^{sat} we now have both:

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

Comment #9
> In describing how T^sat is updated, the manuscript states that Z vs. 1000/T is extrapolated or interpolated. This is confusing because it's unclear what Z refers to (it seems to be the liquid simulation data) and it's unclear how using 1000/T could be meaningfully different from using 1/T. It's also unclear what scheme was used to interpolate or extrapolate the data.

Z refers to compressibility factor from simulation data on liquid isochores. We obtain the same Tsat for either 1/T or 1000/T, but to avoid any confusion we have modified the sentence to read 'reciprocal temperature' and converted 1000/T to 1/T.

Comment #10
> The manuscript describes using 3 points on an isochore for temperature integration. This seems rather sparse. Do the isotherms lack curvature (linear)? Or are they reasonably approximated by a quadratic? From the discussion, it seems that whatever errors are introduced by this are commensurate with other sources of error in the overall approach, but it would be helpful to understand what sort of approximation this is. Would it be better to run shorter simulations at more temperatures (same computational effort) and have more confidence in the behavior? The interpolation of Z (as mentioned previously) and U faces similar issues.

The plot of $U^{dep}T$ vs 1/T is usually slightly curved, as seen in figures in Section III of Supplementary Material. Using 3 points on isochores is enough for the densities of interest (Tr>0.45), but as the density increases the curvature of the isochore increase. Since we are using three point Simpson's rule for integrating along isochore, this curvature is captured. We are also using a 2^{nd} order polynomial along the Z vs. 1/T to update Tsat (See our response to Comment 9). For high density isochores it is not recommended to run short simulations, because those are the hardest to equilibrate, especially at lower temperatures.

Comment #11
> The figure 5 caption states that B3 helps the iterations converge, but the figure actually shows that convergence fails without B3 (as described in the main text). Is that correct?

Using B_3 term helps the iteration to converge. Figure 5a (with B_3) converges and Figure 5b (without B_3) diverges. As this statement already seems to be consistent, we did not modify the caption.

Comment #12
> In section 3.2, the manuscript states that "Vapor pressure precision is very weakly influenced...". It seems this is actually referring to accuracy (tendency to deviate by a particular amount) rather than precision (tendency to fluctuate due to stochastic effects).

To avoid confusions we changed the sentence to:

“Vapor pressure estimate is very weakly influenced ...”

Comment #13
> The figure 8 caption states that "Diamond points are very low density simulations that are not used in B_2 and B_3 calculations." It's unclear why they should be excluded. They are the most important data to estimate virial coefficients (because virial coefficients describe behavior in the limit of low density). Their uncertainties are larger than other densities, but this should have the effect of diminishing their impact on the virial coefficient fits so long as the fits account for uncertainty.

In the proposed ITIC procedure, we show that simulating circles in Figure 8(a) is enough to obtain reliable B_2 values, and we do not need to simulate diamonds points. Since they are unnecessary, we do not recommend performing these simulations as part of the method. We simply show them in Figure 8(a) for illustration. We now clarify this point in the caption of Figure 8.

Comment #14
> The authors attribute the deviation of the fitted B3 from the rigorous values due to being "effective" coefficients that capture effects from B4 or higher. If true, then shouldn't B4 be included as a fitting parameter. Are the simulation data too uncertain to justify a 3-parameter fit? The manuscript does state that B3eff was better able to describe the simulation data than rigorous B3 and B4 and refers to figure 9 to support that. The figure has curves for the VEOS using rigorous B2 & B3 and also using rigorous B2, B3 & B4. Without a curve for B3eff, I'm not sure how the figure supports the claim in the main text.

As the reviewer points out, the simulation data shows too much uncertainty to justify a 3-parameter fit. We apologize that the curve for B3eff has been overlooked. We now include this curve in Figure 9(c) (purple dashed line).

Comment #15
> Does the vapor description by virial coefficients work for strongly interacting models like water as it does for more weakly interacting models like ethane?

The study of TIP4P water in the Example Simulations section shows that, second virial coefficients can be obtained using Eq.(30) for this model. We added a sentence to Example Simulations section to emphasize on this point.

Comment #16
> The error bars in figure 8a seem to be too large; they are much larger than the scatter in the data and the fitted lines go right through most of the points.

The dashed blue lines in Figure 8(a) are actually representing rigorous B2 and B3 values of TraPPE-UA ethane, as expressed in the caption of Figure 8(a) “Blue dashed lines represent Schultz values of B_2 (intercept) and B_3 (slope)”

Comment #17

> The manuscript states that "we need 3200 ethane molecules to achieve low enough uncertainties..." It's unclear why the authors have coupled the uncertainty to the system size. The authors have apparently chosen to run each system size for the same number of MD steps, but this is unfair and especially so if the MC simulations were also run for the same number of MC steps regardless of system size. MD CPU time per step increases with system size, but the simulation will yield increasing precision with system size (commensurate with the additional effort). In MC, the CPU time and uncertainty are both relatively independent of system size for a given number of steps. It would make for a fairer comparison to run the MD simulations for a number of steps that is inversely proportional to system size. Although the number of steps for MC is described in section 4, the number of MD steps does not seem to be specified.

First, we would like to emphasize that the statement "we need 3200 ethane molecules to achieve low enough uncertainties..." only refers to MD rigid. We modified the second paragraph in Section 6 to avoid this confusion.

Furthermore, in light of the reviewer's comment, we replotted Figure 9a and 9b by modifying the number of MD steps depending on the number of molecules. The number of MD steps used in calculating the averages is now inversely proportional to the number of molecules, as specified in the caption of Figure 9. We also increased the number of samples by splitting each run into two blocks, hence eight blocks of data are used to calculate error bars based on 95 % confidence intervals. The first paragraph of Section 4.2 reflects the new approach to calculate the uncertainties.

> Bootstrapping is used to estimate the statistical uncertainties [31] in T^{sat}, P^{sat}, ρ_{vap}, and ΔH_v. Note that the ITIC method determines saturation conditions at a fixed value of ρ_{liq} (equal to the isochoric density) and, therefore, the bootstrap uncertainty in ρ_{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.

Comment #18

> I find the behavior in figure 9b and the attribution to lack of equilibration a bit disturbing. United-atom ethane is a very simple molecule and the bond length should not need much time to equilibrate in the vapor phase. Have the authors been able to collect better data using longer equilibration times? Does the divergence at low density at least decrease with increasing equilibration length? This should not be hard for the smallest system size (N=120). If not, then I suspect something is just broken.

Running long simulations revealed that the behavior in Figure 9b is not due to lack of equilibration. Therefore, we removed the sentence in the manuscript that attributed this behavior lack of

equilibration. As discussed in our response to Comment 6.1 of Reviewer 2, the divergence of pressure is due to inaccuracies in the intramolecular bonded virial at low densities. In Supplementary Material, we illustrate a method to correct for these inaccuracies if the simulation package provides different virial contributions, i.e. pair virial, bond virial, kinetic virial, etc. Below are the paragraphs containing the above explanation.

> The effect of using flexible bonds is shown in Figure 9(b). The systematic discrepancy from rigorous values (solid black line) as well as large uncertainties suggests not using flexible bonds at very low densities. The divergence of pressure at low densities is presumably due to the small number of intermolecular collisions at low density relative to the large number of intramolecular collisions, inhibiting the equilibration of the various components of momentum. In Supplementary Materials, we illustrate one manner of correcting for this deficiency by modifying the intramolecular virial when different virial contributions are separately accessible.
>
> Figure 9(c) shows the low density *NVT* state points simulated using GOMC [25]. 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.

Comment #19

> The manuscript describes the relative standard deviation between runs (STD2) being much smaller than the average standard deviation within a run (STD1). Is STD1 then the standard deviation of the samples (how much Z fluctuates from configuration to configuration)? Or is it based on block averages? And STD2 shows how much the average value (from a simulation) fluctuates when the simulation is repeated? It's unclear why these values are being reported in this way, or why STD1 is reported at all. The authors chose to use STD2 for uncertainty (error bars), but this is inappropriate if the value being plotted is the average across the runs. Please compute a proper uncertainty ("standard deviation of the mean", "standard error", etc.) for error bars. Also, computing a standard deviation from 4 points is not likely to yield an accurate result. Uncorrelated block averages are more likely to yield a good estimate of the uncertainty.

We agree with the reviewer that 4 samples is not enough for computing reliable uncertainties, however we believe that block averaging can also be problematic when the block lengths are too small that the data remain correlated. Therefore, we now combine the two approaches, i.e. calculating uncertainties from replicate simulations and from block averaging. In the current text, uncertainties are calculated based on 95 % confidence intervals from eight decorrelated samples, i.e. two blocks on four separate runs. Note that, since most of the data in the table that contained STD1 and STD2 data is plotted in Figure 9, we decided to remove this table from the manuscript.

Comment #20
> In some cases, the data shown in figure 10 shows agreement between ITIC and GEMC/GCMC methods. In other cases, the agreement is not so good. The level of agreement or disagreement ought to be discussed in the text, since this (rather than agreement with REFPROP) is what determines if the method is working. Is it due to stochastic effects (where are the error bars? are they smaller than the symbols?)? Or is there systematic disagreement? If so, which results are inaccurate? Also, figure 10a appears to show that the vapor pressure may deviate from REFPROP data by up to 100%, certainly more than 1% as asserted in the conclusions. Is the statement in the conclusion referring to a different type of comparison?

It is challenging to add error bars for ITIC results in Figure 10, because we have uncertainties in both Tsat and Psat. Also, adding error bars to Figure 10 makes it very hard to read. Therefore we only include the error bars in Figure 11.

Note that in Figure 10, we calculate the deviation between ITIC and REFPROP and the deviation between literature data and REFPROP, then we compare these two deviations together. Therefore, REFPROP serves as a baseline for comparison, such that the discrepancies are more visible than on a logarithmic plot. Only the difference between the two deviations matter, not the deviation values themselves.

In the Conclusion section, it is shown that the REFPROP Psat value is in agreement with ITIC Psat values obtained using REFPROP state points. This agreement validates the ITIC formula and the choice of ITIC points.

The level of agreement between ITIC and literature data shown in Figure 10, is now discussed in a more specific manner in the first paragraph of Section 7.

Comment #21
> Additional discussion of the method's limitations is needed to help readers if and how they could use the method for their own purposes.

One limitation of ITIC method is the need for reasonable initial estimates for Tsat. This limitation and our approach to address it is included in last two paragraphs of Section 7.

Comment #22
> The inability of ITIC to yield coexistence properties at a given temperature or pressure receives little attention in the manuscript; liquid density is specified and other properties are determined. Interest in a given temperature or pressure is far more likely and inability to specify temperature makes it difficult to compare results against other methods (GDI can do T or P; GCMC and GEMC can do T). The authors are able to make comparisons by subtracting REFPROP values from the predictions of each method, but this is not always possible and may not be adequate. It would be helpful if the authors could suggest an approach (Clapeyron equation, interpolation, running one more simulation, etc.) that would allow estimates to be obtained for a condition that is near one that was actually solved for.

We added a sentence in Section 2.3 to emphasize that in ITIC method, rhoL is imposed and the rest of saturation properties including Tsat are calculated.

Implementation of the ITIC procedure starts with determining the temperatures and densities of the ITIC points shown in Figure 1. In ITIC method, ρ_{liq} is imposed and T^{sat}, P^{sat}, and ρ_{liq} are calculated. This involves estimating saturation temperatures (T^{sat}_{est}), i.e. T at points 11, 13, 15, 17, and 19. Once ITIC state points are determined, NVT simulations are performed at the

Comment #23

> Finally, the manuscript addresses sensitivity of the results to the guess for T^sat, but it is unclear how to identify appropriate saturation temperatures for a given model. If the coexistence behavior is unknown, how can a researcher determine these temperatures? Is it necessary to use another method (GDI, GCMC and GEMC do not require such knowledge) to get an estimate coexistence point for each liquid density? Or should we be unconcerned with getting it right and add temperatures as necessary (above or below the initial guess)

For an unknown molecule, crude Tsat estimation at each rhoL is possible by obtaining two points on each isochore (Z vs. 1/T) and interpolating or extrapolating to Z=0.0. An example of such situation (1-naphthalenyl, 4-phenanthrenyl butane) is added to the Example Simulations section to explain the necessary steps taken to use ITIC method when no Tsat-rhoL information is known. This adds to the cost of the ITIC method, since extra simulations are performed at each isochore, but one can use that additional information to improve the integration accuracy and the ultimate knowledge of the equation of state for the molecule.

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^{sat}) and saturation liquid and vapor density (ρ_{liq} and ρ_{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 ($T_{\text{r}} = T/T_{\text{c}}$, 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 $T_{\text{r}} = 0.3$) 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. 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] have applied this approach in conjunction with GEMC to obtain the initial coexistence point, NPT simulations to estimate saturation liquid densities, and parallel tempering method [9] 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 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[10]. Elliott et al. [11] 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[12]. 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). 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 [13], it should be possible to utilize these data in order to generate high accuracy multi-parameter equations of state [14, 15, 16, 17, 18, 19, 20, 21].

As another advantage, vapor pressure calculation using molecular dynamics (MD) is often perceived as an impractical approach [22]. 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 without a significant increase in computational effort compared to the typical GEMC or GCMC approaches.

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. Free energy departure function derivation

For a single component system, Eq. (1) must be satisfied at vapor-liquid phase equilibrium

$$\begin{aligned} T_{\text{vap}} &= T_{\text{liq}} \\ P_{\text{vap}} &= P_{\text{liq}} \\ G_{\text{vap}} &= G_{\text{liq}} \end{aligned} \tag{1}$$

where G represents the molar Gibbs free energy, T is the temperature, P is pressure, and subscripts "vap" and "liq" denote the vapor and liquid phases, respectively.

The Gibbs energy criterion of phase equilibrium can be rewritten in departure function form

$$\left(\frac{G_{\text{liq}} - G_{\text{ig}}}{RT}\right)_{T,P} = \left(\frac{G_{\text{vap}} - G_{\text{ig}}}{RT}\right)_{T,P} \tag{2}$$

where the subscript "ig" denotes the ideal gas, and R is the gas constant (8.3144598 J/(mol K)). This equation can be rearranged in terms of the Helmholtz energy (A) departure function

$$\begin{aligned} &\left(\tfrac{A_{\text{liq}} - A_{\text{ig}}}{RT}\right)_{T,V} + Z_{\text{liq}} - 1 - \ln(Z_{\text{liq}}) \\ &= \left(\tfrac{A_{\text{vap}} - A_{\text{ig}}}{RT}\right)_{T,V} + Z_{\text{vap}} - 1 - \ln(Z_{\text{vap}}) \end{aligned} \tag{3}$$

where Z is the compressibility factor ($Z \equiv \frac{P}{\rho RT}$). Note that the $\ln(Z)$ terms are introduced when converting Gibbs energy departure function at constant T, P in Eq. (2) to Helmholtz energy departure function at constant T, V in Eq. (3) [23].

Calculating free energy of a system using the ITIC integration method involves connecting the state of interest to a state of known free energy. In the case of phase equilibrium calculation, the state of interest is saturated liquid and the state of known free energy is an ideal gas. In order to connect saturated liquid to low densities we should not cross the two phase region. Instead, the temperature of saturated liquid is increased to a supercritical temperature through an isochoric path, then the supercritical state is connected to low densities through an isothermal path.

The first stage involves integrating internal energy, U, according to Eq. (8) which is derived using Eqs. (4-7). Eq. (8) is

obtained from Eq. (7) by taking the difference between the real and ideal gas states (where the constants of integration cancel).

$$U = kT^2 \left(\frac{\partial \ln Q}{\partial T} \right)_{N,V} \tag{4}$$

$$A\left(N,V,T\right) = -kT \ln Q\left(N,V,T\right) \tag{5}$$

$$U = kT^2 \left(\frac{\partial(-A/kT)}{\partial T} \right)_{N,V} = \frac{\partial \beta A}{\partial \beta} \tag{6}$$

$$\beta A = \int U \mathrm{d}\beta + \text{constant} \tag{7}$$

$$\left(\frac{A - A_{\mathrm{ig}}}{RT} \right)_{TV} = \int_{T^{\mathrm{IT}}}^{T^{\mathrm{sat}}} \frac{U - U_{\mathrm{ig}}}{R} \mathrm{d}\left(\frac{1}{T} \right) \tag{8}$$

where V represents volume, k is the Boltzmann constant, $\beta \equiv \frac{1}{kT}$, and Q is the canonical partition function.

The second stage consists in performing the following integration

$$\frac{A - A_{\mathrm{ig}}}{RT} = \int_0^{\rho} \frac{Z-1}{\rho} \, \mathrm{d}\rho \tag{9}$$

where ρ represents density.

Therefore, the Helmholtz energy departure function of the liquid phase is calculated using Eq. (10)

$$\begin{aligned} &\left(\tfrac{A_{\mathrm{liq}} - A_{\mathrm{ig}}}{RT} \right)_{T,V} \\ &= \int_0^{\rho_{\mathrm{liq}}} \tfrac{Z-1}{\rho} \, \mathrm{d}\rho \,|_{T^{\mathrm{IT}}} + \int_{T^{\mathrm{IT}}}^{T^{\mathrm{sat}}} \tfrac{U - U_{\mathrm{ig}}}{R} \, \mathrm{d}\left(\tfrac{1}{T} \right)|_{\rho_{\mathrm{liq}}} \end{aligned} \tag{10}$$

where the first integral is performed along the supercritical isotherm ($T^{\mathrm{IT}} \approx 1.2\, T_{\mathrm{c}}$) from zero density to the saturated liquid density of interest, and second integral is performed along the saturated liquid density isochore from $T_{\mathrm{r}} \approx 1.2$ to saturation temperature T^{sat}. Several separate simulations in the canonical ensemble (NVT) are needed to provide enough data points on the paths for numerical integration of Eq. (10) to be accurate.

Because the density of vapor phase is low at most conditions of interest, the value for Z_{vap} in Eq. (3) can be approximated using a virial expansion truncated at the B_3 term

$$Z_{\mathrm{vap}} = 1 + B_2 \rho_{\mathrm{vap}} + B_3 \rho_{\mathrm{vap}}^2 \tag{11}$$

which is substituted in Eq. (9) to obtain free energy departure function of vapor phase

$$\begin{aligned} &\left(\tfrac{A_{\mathrm{vap}} - A_{\mathrm{ig}}}{RT} \right)_{T,V} = \int_0^{\rho_{\mathrm{vap}}} \tfrac{Z-1}{\rho} \mathrm{d}\rho \\ &= \int_0^{\rho_{\mathrm{vap}}} \tfrac{B_2 \rho_{\mathrm{vap}} + B_3 \rho_{\mathrm{vap}}^2}{\rho_{\mathrm{vap}}} \mathrm{d}\rho_{\mathrm{vap}} = B_2 \rho_{\mathrm{vap}} + \tfrac{1}{2} B_3 \rho_{\mathrm{vap}}^2 \end{aligned} \tag{12}$$

2.2. Vapor pressure and density calculation

We have characterized the Helmholtz energy departure function in vapor and liquid phases. Substituting Eq. (10) and Eq. (12) in Eq. (3) yields

$$\left(\frac{A_{\mathrm{liq}} - A_{\mathrm{ig}}}{RT} \right)_{T,V} + Z_{\mathrm{liq}} - 1 + \ln\left(\frac{\rho_{\mathrm{liq}}}{\rho_{\mathrm{vap}}} \right) = 2B_2 \rho_{\mathrm{vap}} + \frac{3}{2} B_3 \rho_{\mathrm{vap}}^2 \tag{13}$$

which can be rearranged in terms of ρ_{vap}

$$\begin{aligned} &\rho_{\mathrm{vap}} = \\ &\rho_{\mathrm{liq}} \exp\left(\left(\tfrac{A_{\mathrm{liq}} - A_{\mathrm{ig}}}{RT} \right)_{T,V} + Z_{\mathrm{liq}} - 1 - 2B_2 \rho_{\mathrm{vap}} - \tfrac{3}{2} B_3 \rho_{\mathrm{vap}}^2 \right) \end{aligned} \tag{14}$$

Eq. (14) can be solved for ρ_{vap} using a root-solving approach. For simplicity, we solve Eq. (14) with the fixed-point iteration method [24]. Finally, vapor pressure is calculated using Eq. (15)

$$P^{\mathrm{sat}} = P_{\mathrm{vap}} = Z_{\mathrm{vap}} \rho_{\mathrm{vap}} R T^{\mathrm{sat}} = (1 + B_2 \rho_{\mathrm{vap}} + B_3 \rho_{\mathrm{vap}}^2) \rho_{\mathrm{vap}} R T^{\mathrm{sat}} \tag{15}$$

2.3. 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, ρ_{liq} is imposed and T^{sat}, P^{sat}, and ρ_{liq} are calculated. This involves estimating saturation temperatures ($T^{\mathrm{sat}}_{\mathrm{est}}$), 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 compressibility factor and internal energy departure function ($U^{\mathrm{dep}} = \frac{U - U_{\mathrm{ig}}}{RT}$) are calculated from simulation output. Helmholtz energy departure function (A^{dep}) is then calculated as explained in Section 2.4, i.e. Eq (18-19). At this point, T^{sat}, ρ_{vap}, and P^{sat} are solved iteratively for each isochore. First, ρ_{vap} and Z_{liq} are set to an initial near-zero value. Then, ρ_{vap} and P^{sat} are recomputed using Eq. (14) and Eq. (15), respectively. Then, Z_{liq} is updated as

$$Z_{\mathrm{liq}} = \frac{P^{\mathrm{sat}}}{\rho_{\mathrm{liq}} R T^{\mathrm{sat}}} \tag{16}$$

where ρ_{liq} is equal to the isochoric density.

Next, T^{sat} is updated by inter- or extrapolating the isochoric simulation results, specifically, we fit Z vs reciprocal temperature to a second order polynomial and solve for T^{sat} such that $Z(T^{\mathrm{sat}}) = Z_{\mathrm{liq}}$. It should be noted that iterations of T^{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^{sat} has changed, A^{dep} should also be recomputed. Trapezoid rule is used to integrate $U^{\mathrm{dep}}T$ vs. $1/T$ (i.e. $(U - U^{\mathrm{ig}})/R$ vs. $1/T$ in Eq. 8) due to the linear shape of $U^{\mathrm{dep}}T$ vs. $1/T$ in the vicinity of T^{sat}.

$$A^{\mathrm{dep}}_{\mathrm{new}} = A^{\mathrm{dep}}_{\mathrm{old}} + \left(\frac{1}{T^{\mathrm{sat}}_{\mathrm{new}}} - \frac{1}{T^{\mathrm{sat}}_{\mathrm{old}}} \right) \frac{U^{\mathrm{dep}}_{\mathrm{new}} T^{\mathrm{sat}}_{\mathrm{new}} + U^{\mathrm{dep}}_{\mathrm{old}} T^{\mathrm{sat}}_{\mathrm{old}}}{2} \tag{17}$$

where subscript "old" denotes the corresponding property at $T^{\mathrm{sat}}_{\mathrm{old}}$ and subscript "new" denotes the property at updated $T^{\mathrm{sat}}_{\mathrm{new}}$ (For the first iteration, $T^{\mathrm{sat}}_{\mathrm{old}} = T^{\mathrm{sat}}_{\mathrm{est}}$). Note that $U^{\mathrm{dep}}_{\mathrm{new}}$ is calculated by inter- or extrapolating U^{dep} vs $1/T$ along the isochore, i.e., $U^{\mathrm{dep}}_{\mathrm{new}} = U^{\mathrm{dep}}(T^{\mathrm{sat}}_{\mathrm{new}})$, where $U^{\mathrm{dep}}T$ vs. $1/T$ is fit to a first order polynomial.

Having new values for T^{sat}, Z_{liq}, ρ_{vap}, and A^{dep} the next iteration starts from Eq. (14). The iterations are stopped when the

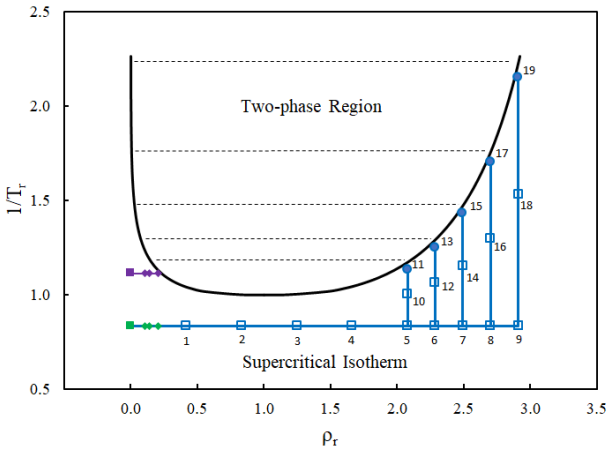


Figure 1: A schematic plot of the pathways taken in the ITIC method. Squares and diamonds represent non-saturated state points obtained from simulation, while blue filles circles represent estimated T^{sat} values. Green and purple diamonds show the state points required for B_2 calculation at isothermal temperature and $T_{\text{r}} = 0.9$, respectively. Note that circle points 11, 13, 15, 17, and 19 represent the initial estimate of saturation temperature, therefore they do not necessarily match the coexistence curve.

difference in two consecutive ρ_{vap} (or T^{sat}) 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 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^{IT}) and saturation second virial coefficients (B_2^{sat}) which is explained in Section 5.

2.4. 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 $T_{\text{r}} \approx 1.2$ 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 ($T_{\text{r}} \approx 1.05$) 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).

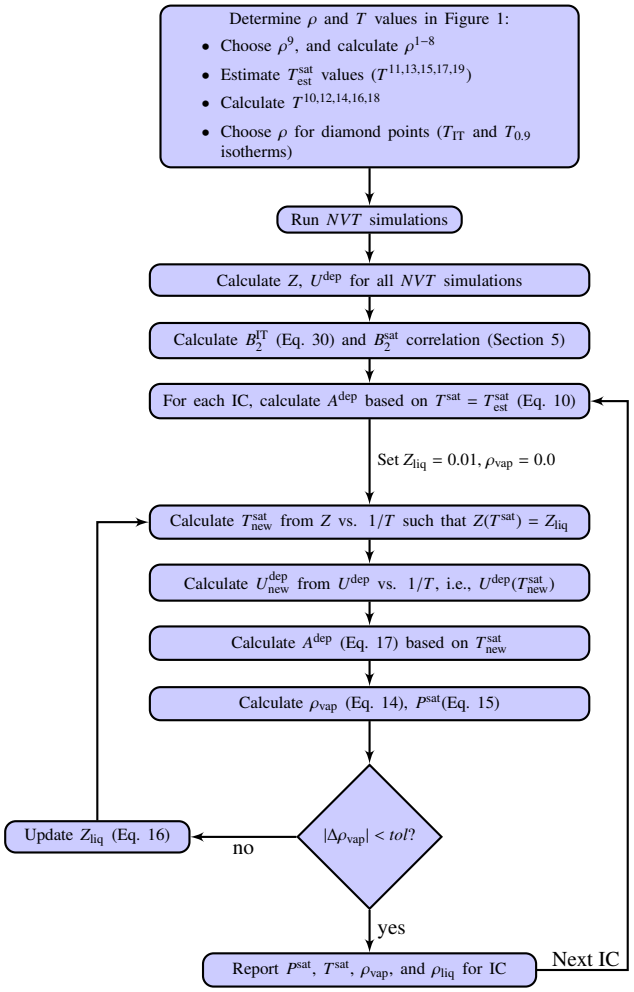


Figure 2: Algorithm to obtain VLE from NVT simulations using ITIC method

The integrations along the isotherm and isochores are performed using Simpson's rule [25], 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. (18) and Eq. (19) are articulations of Simpson's rule used for numerical integration along the isotherm and isochores to calculate A^{dep} by Eq. (10).

$$\int_a^b f(x)\mathrm{d}x \approx \frac{b-a}{6}\left[f(a)+4f\left(\frac{a+b}{2}\right)+f(b)\right] \tag{18}$$

$$\int_a^b f(x)\mathrm{d}x \approx \frac{b-a}{8}\left[f(a)+3f\left(\frac{b-a}{3}\right)+3f\left(\frac{2(b-a)}{3}\right)+f(b)\right] \tag{19}$$

The A^{dep} values at points 2, 4, 7, and 9 in Figure 1 are sequentially calculated using Eq. (18) in which the value of function at three equidistant points on the x-axis are needed. The A^{dep} values at points 3, 5, and 8 are sequentially calculated using Eq. (19) in which the value of the function at four equidistant points on the x axis are needed. The A^{dep} at point 6 is equal to the integration value from point 6 to point 8 (from Eq. (18)) subtracted from the A^{dep} value at point 8. The A^{dep} value at point 1 is equal to the integration value from point 1 to point 4 (from Eq. (19)) subtracted from A^{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 $T_{\text{r}} = 0.9$, 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^{dep} at point 2. In Section 5, a more detailed discussion of obtaining virial coefficient is provided.

2.5. Enthalpy of Vaporization Calculation

In the ITIC method, enthalpy of vaporization (ΔH_{v}) is calculated using

$$\Delta H_{\text{v}} = (H_{\text{vap}}^{\text{dep,sat}} - H_{\text{liq}}^{\text{dep,sat}})RT \tag{20}$$

where $H_{\text{vap}}^{\text{dep,sat}}$ and $H_{\text{liq}}^{\text{dep,sat}}$ are unitless enthalpy departure functions of saturated vapor and liquid, respectively, defined as

$$H^{\text{dep}} = \frac{H - H^{\text{ig}}}{RT} \tag{21}$$

$H_{\text{vap}}^{\text{dep,sat}}$ and $H_{\text{liq}}^{\text{dep,sat}}$ can be calculated by subtracting the ideal gas contribution from both sides of $H = U + PV$

$$H_{\text{liq}}^{\text{dep,sat}} \approx U_{\text{liq}}^{\text{dep,sat}} + Z_{\text{liq}} - 1 \tag{22}$$

$$H_{\text{vap}}^{\text{dep,sat}} \approx U_{\text{vap}}^{\text{dep,sat}} + Z_{\text{vap}} - 1 \tag{23}$$

The value of $U_{\text{liq}}^{\text{dep,sat}}$ is calculated at each fixed-point iteration based on the updated value of T^{sat}. $U_{\text{vap}}^{\text{dep,sat}}$, on the other hand, is calculated by taking the derivative of B_2 with respect to β as shown in Eq. (24).

$$U_{\text{vap}}^{\text{dep,sat}} \approx \rho_{\text{vap}}\beta\frac{dB_2}{d\beta} \tag{24}$$

2.6. Critical Point Calculation

We follow the standard approach used in GEMC and GCMC methods to calculate the critial point. The critical temperature and critical density (ρ_{c}) are estimated by the law of rectilinear diameter [26] and the density scaling law [27]

$$\frac{\rho_{\text{liq}} + \rho_{\text{vap}}}{2} = \rho_{\text{c}} + A(T_{\text{c}} - T^{\text{sat}}) \tag{25}$$

$$\rho_{\text{liq}} - \rho_{\text{vap}} = B(T_{\text{c}} - T^{\text{sat}})^{0.325} \tag{26}$$

where ρ_{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^{sat}) and pressures (P^{sat}) to the Antoine equation [28]

$$\log_{10}(P^{\text{sat}}) = a_0 + \frac{a_1}{a_2 + T^{\text{sat}}} \tag{27}$$

where a_i are fitting constants and a_1 is constrained to be negative. The second step is to evaluate Eq.(27) with the optimal values of a_i for $T^{\text{sat}} = T_{\text{c}}$.

Some caution should be exercised when extrapolating Eqs.(25-27) because ITIC data are typically not available near the critical point. For example, we recommend excluding low temperature ITIC data ($T_{\text{r}} < 0.6$) when fitting Eqs.(25-27) 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 $T_{\text{r}} < 0.85$, careful application of Eqs.(25-27) 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 [29], 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^{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. (11) includes or excludes the B_3 term. The deviations of calculated P^{sat}, ρ_{liq}, ρ_{vap}, and ΔH_{v} from the data obtained directly from NIST REFPROP [30] are plotted in Figure 4. According to this figure, by including the B_3 term, one can reach a reduced saturation temperature ($T_{\text{r}}^{\text{sat}}$) of 0.9 with less than 1 % error in all saturation properties. The ITIC method is not able to calculate accurate saturation properties when $T_{\text{r}}^{\text{sat}} > 0.9$. If the B_3 term is excluded from Eq. (11), the ITIC method does not converge for $T_{\text{r}}^{\text{sat}} > 0.85$. When $T_{\text{r}}^{\text{sat}} < 0.85$, excluding the B_3 term provides less than 1 % deviation in P^{sat}, ρ_{liq}, and ΔH_{v} , and less than 2.5 % deviation in ρ_{vap}.

Figure 5(a) illustrates the convergence paths taken by a fixed-point method to calculate ρ_{vap}, with B_3 included in Eq (11). Figure 5(b) shows the same plot, except B_3 is excluded, i.e. virial expansion in Eq. (11) is truncated at B_2 term. Using B_3 corrects the curve representing the right-hand side of Eq. (14) ($g(\rho_{vap})$) in such a way that fixed-point iteration converges.

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^{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. (14) and Eq. (15). Figure 6(b) shows the P^{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^{sat}. However, a relatively accurate B_2 is required to obtain accurate P^{sat} when $T_r^{sat} > 0.75$. Similarly, when $T_r^{sat} \approx 0.84$ 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 fixed-point iteration in terms of convergence (Figure 5), the sensitivity of P^{sat} to B_3 is negligible when T_r^{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 ρ_{vap} and T^{sat} was considered. The sensitivity of ρ_{vap} to virial coefficients is similar to P^{sat}, and T^{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_{est}^{sat}. Figure 7 demonstrates that deviations in P^{sat} are less than 1 % and deviations in ρ_{liq} are less than 0.25 % when errors in T_{est}^{sat} are within 10 %. Deviations in ρ_{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_{est}^{sat} are less than 1 %, in which case the deviations for all three properties are nearly indistinguishable from the deviations resulting from numerical integration alone, i.e., those for 0 % error in T_{est}^{sat}.

Note that in Figure 7, isochoric/isothermal properties (Z and U^{dep}) used for determining saturation properties are obtained from REFPROP. If T_{est}^{sat} percent deviation shown in the legend of Figure 7(a) is less than zero (i.e. $T_{est}^{sat} < T_{REFPROP}^{sat}$), some of the ITIC points are in metastable state. Since the REFPROP database does not provide Z and U^{dep} for such points, a linear extrapolation of REFPROP data was used to approximate Z and U^{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 with fixed bond lengths is favored due to smaller uncertainties at low density. The Cassandra [31] and GOMC (GPU Optimized Monte Carlo) [32] 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 [33]. 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[34]. 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) [35] and GROMACS [36] within their uncertainties.

For each compound, 26 NVT points are simulated in order to obtain 5 saturation points, as illustrated in Figure 1. The density of the isochore with highest density is chosen to match the experimental liquid density at the minimum reduced saturation temperature (T_r^{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 [37] 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. The simulation boxes 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 ar 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,

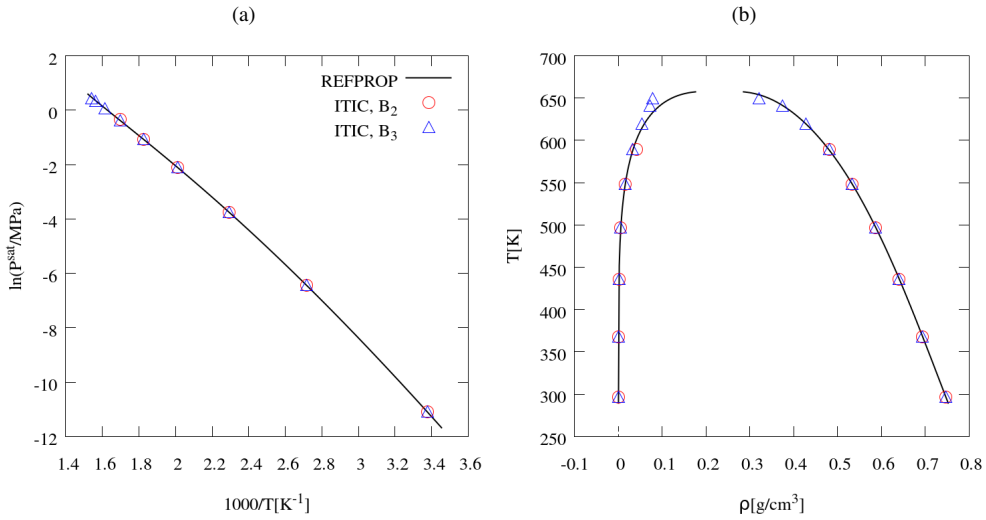


Figure 3: n-Dodecane Clausius-Clapeyron and coexistence curves. ITIC results (symbols) are obtained using NIST REFPROP [30] values for U^{dep} and Z. Circles and triangles represent ITIC results when virial expansion in Eq. (11) is truncated at B_2 and B_3 terms, respectively. Solid line represents true NIST REFPROP VLE data.

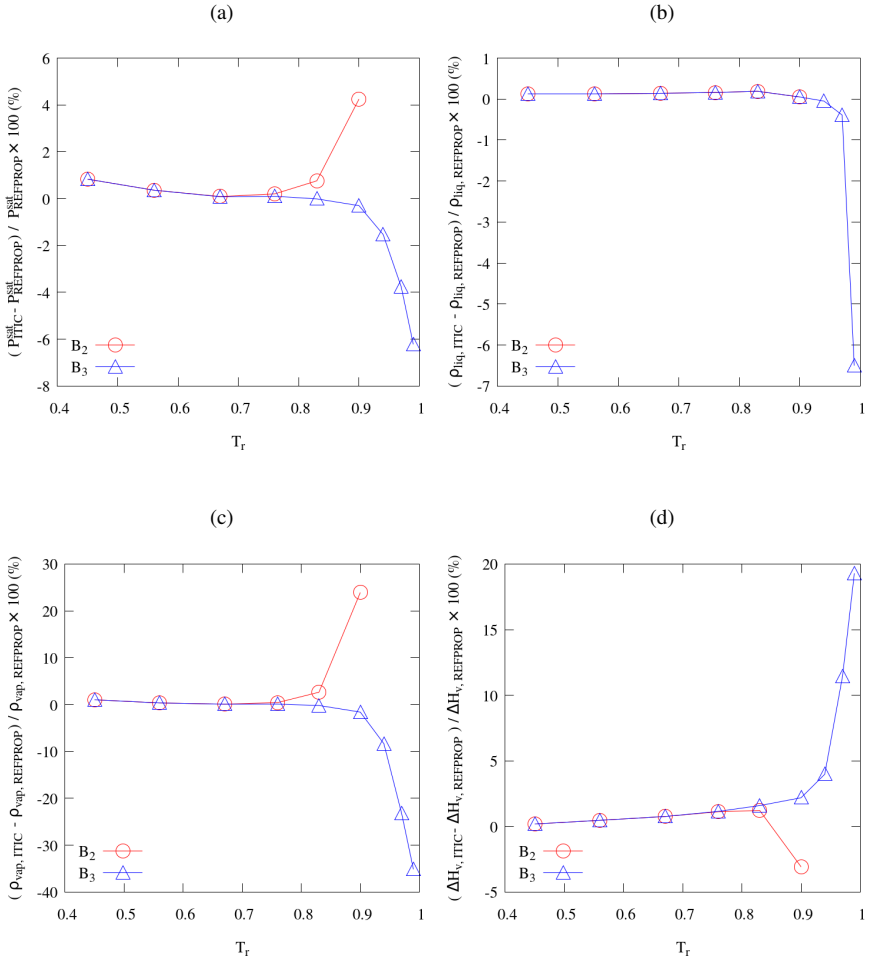


Figure 4: Accuracy of the ITIC method for n-dodecane when Eq. (11) is truncated at B_2 and B_3 term. If B_3 is excluded the fixed-point iteration does not converge, when $T_{\text{r}}^{\text{sat}} > 0.9$. Y-axis represents deviations calculated using $\frac{\text{ITIC}-\text{REFPROP}}{\text{REFPROP}} \times 100$, which compares ITIC results using REFPROP state points with REFPROP coesistence data.

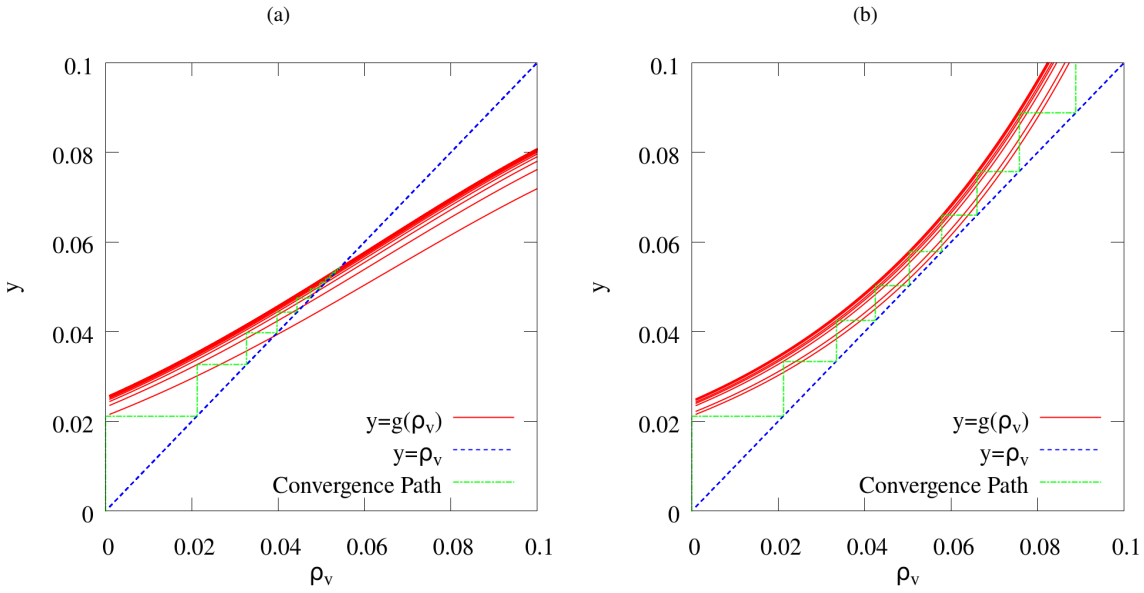


Figure 5: Fixed-point method iteration and convergence path for n-dodecane for the isochore corresponding to $\rho_{\text{liq}} = 0.4269\,\text{g/cm}^3$. Eq. (14) is summarized into $\rho_{\text{vap}} = g(\rho_{\text{vap}})$, i.e. the standard form of fixed-point method, where the $g(\rho_{\text{vap}})$ curves represent the right-hand side of Eq. (14). At each iteration, $g(\rho_{\text{vap}})$ is calculated based on a new set of T^{sat}, P^{sat}, A^{dep}, and Z_{liq}. Iteration starts with a low initial guess for ρ_{vap} and stops when absolute percent deviation between two consecutive ρ_{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. (11) , **b)** B_3 term is excluded from Eq. (11). Using B_3 helps iteration to converge for high T^{sat} values.

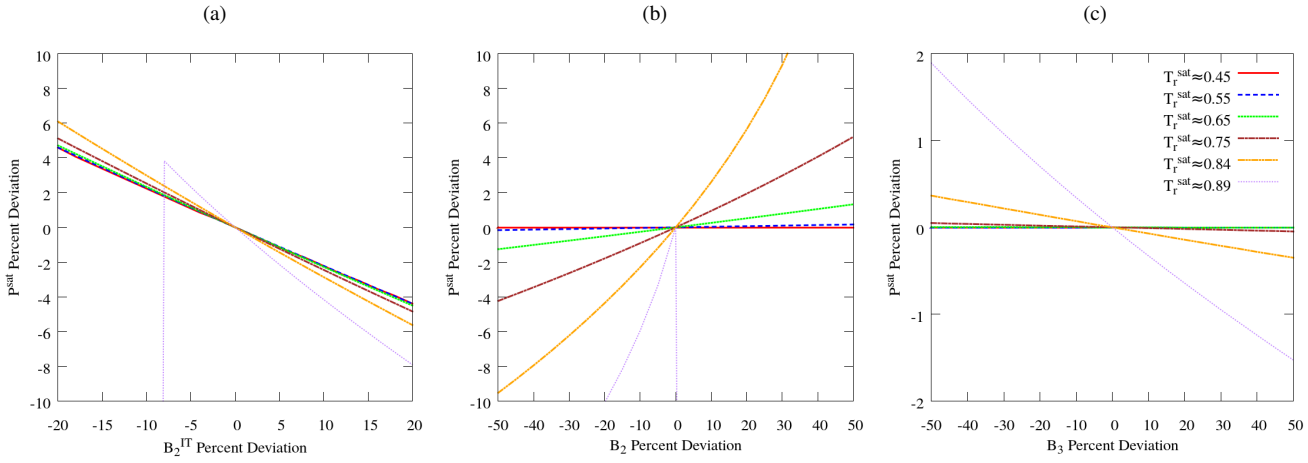


Figure 6: **a)** n-Dodecane P^{sat} sensitivity to isotherm B_2 ,**b)** P^{sat} sensitivity to second virial coefficient used in Eq. (15) , and **c)** P^{sat} sensitivity to third virial coefficient used in Eq. (15)

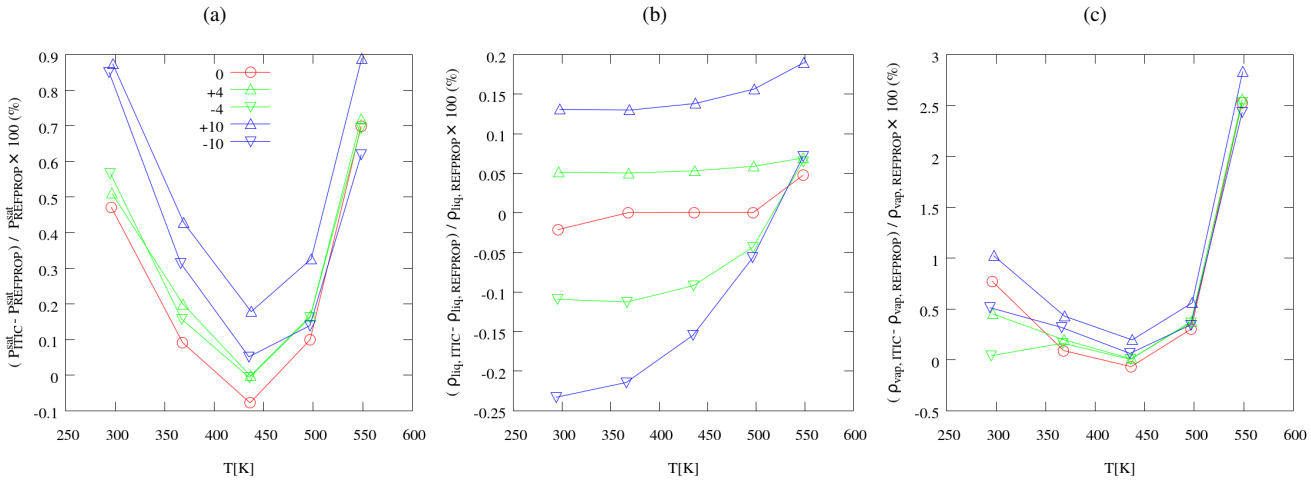


Figure 7: Sensitivity of the ITIC method to $T_{\text{est}}^{\text{sat}}$ for n-dodecane. The ITIC method is applied using REFPROP values of Z and U^{dep}. The y-axis represents percent deviation of the corresponding property calculated using ITIC compared with REFPROP saturation data. Triangles pointing up or down represent ITIC results when $T_{\text{est}}^{\text{sat}}$ is increased or decreased by the percentage shown in the legend, respectively. B_3 is not included in Eq. (14).

0.75, and 0.85. This temperature range is chosen because both GEMC and ITIC are reliable within $0.65 < T_{\text{r}} < 0.85$. 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^{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^{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. (10) requires estimating the ideal gas energy (U_{ig}) at each temperature simulated along a given isochore. The most rigorous approach to determine U_{ig} is by simulating a single molecule system at the corresponding temperature, such that

$$U^{\text{dep}} = \frac{E^{\text{tot}} - N \times E^{N=1}}{NRT} \tag{28}$$

where E^{tot} and $E^{N=1}$ are the total potential energy and the single molecule potential energy, respectively. We recommend Monte Carlo or stochastic dynamics to compute $E^{N=1}$. 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 [38].

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_{ig} is approximately equal to the intramolecular energy at the isochoric state point, such that

$$U^{\text{dep}} = \frac{E^{\text{tot}} - E^{\text{bonded}} - E^{\text{intra}}}{NRT} \tag{29}$$

where E^{bonded} and E^{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^{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^{intra} causes a significant error in vapor pressure.

We emphasize that Eq. (29) 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^{sat} are between 10% and 20% when applying Eq. (29) to larger molecules, where the deviation increases with decreasing T^{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^{dep} (i.e., Eq. (28)-(29)) is discussed in Supplementary Material.

4.2. Bootstrapping Method for Uncertainty Calculation

Bootstrapping is used to estimate the statistical uncertainties [39] in T^{sat}, P^{sat}, ρ_{vap}, and ΔH_{v}. Note that the ITIC method determines saturation conditions at a fixed value of ρ_{liq} (equal to

the isochoric density) and, therefore, the bootstrap uncertainty in ρ_{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 $T_{\text{r}} = 0.9$ 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 [40]. For example, Kofke and Schultz implement this approach for n-alkanes [41]. Efforts like those of Kofke and Schultz could make simulations on the lower end of supercritical isotherm unnecessary [42], 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 [43].

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 ρ. 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. (30)

$$\frac{Z-1}{\rho} = B_2 + B_3\rho \tag{30}$$

where the intercept (B_2) and slope (B_3) correspond to Schultz's values [41].

Figure 8(a) shows that for TraPPE-UA ethane at temperatures above $T_{\text{r}}^{\text{sat}} = 0.85$, B_2 values calculated using this method agree with values reported by Schultz to within 2 %. In Figure 8(a), Z at $T_{\text{r}} = 0.8$ 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. (14), it is important to have a correlation for B_2 and B_3 with respect to temperature, because T^{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 [45]. In this study, we obtain such a correlation from the formula used in the DIPPR [46] database, except the last term is omitted to decrease the number of parameters and avoid overfitting, as shown in Eq. (31)

$$B_2 = A + \frac{B}{T} + \frac{C}{T^3} \tag{31}$$

Eq. (32) and Eq. (33) are obtained by inserting B_2 values extrapolated using Eq. (30) and their corresponding temperatures into Eq. (31).

$$B_2(T_{\text{IT}}) = A + \frac{B}{T_{\text{IT}}} + \frac{C}{T_{\text{IT}}^3} \tag{32}$$

$$B_2(T_{0.9}) = A + \frac{B}{T_{0.9}} + \frac{C}{T_{0.9}^3} \tag{33}$$

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

Subtracting Eq. (33) from Eq. (32) gives

$$B_2(T_{\text{IT}}) - B_2(T_{0.9}) = B\left(\frac{1}{T_{\text{IT}}} - \frac{1}{T_{0.9}}\right) + C\left(\frac{1}{T_{\text{IT}}^3} - \frac{1}{T_{0.9}^3}\right) \tag{34}$$

Taking the derivative of B_2 with respect to β, as shown in Eq. (24) leads to the internal energy departure function

$$\rho\frac{\beta\partial B_2}{\partial\beta} = \rho\left(\frac{B}{T} + \frac{3C}{T^3}\right) = \frac{U - U_{\text{ig}}}{RT} \tag{35}$$

Therefore, the intercept of $\frac{U-U_{\text{ig}}}{\rho RT}$ with respect to ρ gives the value of $\beta\frac{\partial B_2}{\partial\beta}$. Having this value can lead to Eq. (36) with two unknowns (B and C)

$$\beta\frac{\partial B_2}{\partial\beta} = \frac{B}{T_{0.9}} + \frac{3C}{T_{0.9}^3} \tag{36}$$

Solving three equations (Eq. (32)/(33), (34), and (36)) 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 [41].

B_3 calculation using Eq. (30) 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^{eff} from Eq. (30) 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.

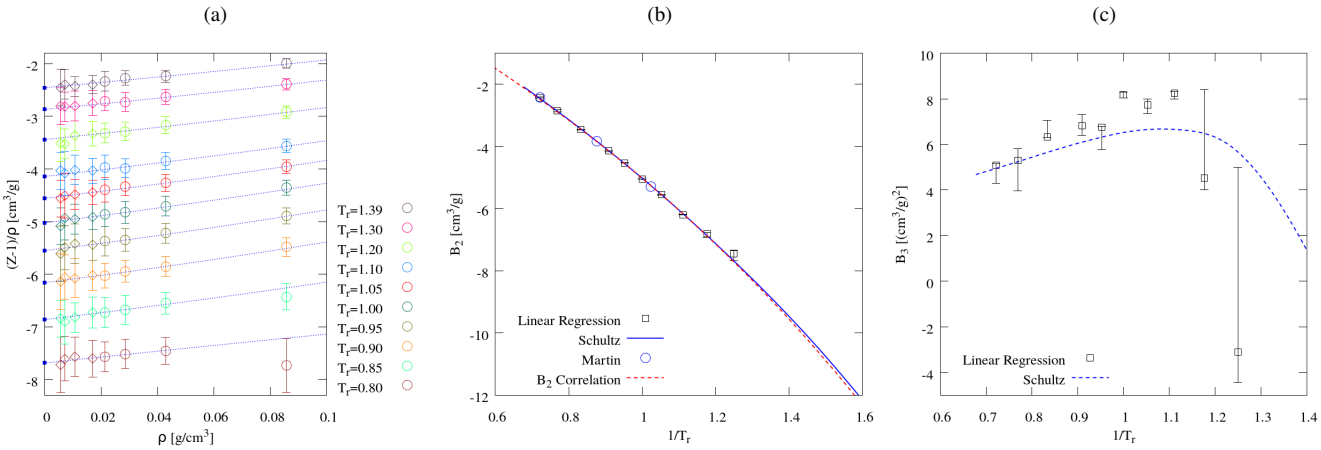


Figure 8: Panel **a)** shows the plot of $(Z-1)/\rho$ with respect to ρ for TraPPE-UA ethane. Blue solid squares represent Schultz's B_2 values. [41]. Blue dashed lines represent Schultz values of B_2 (intercept) and B_3 (slope). Circles and diamonds are NVT state points simulated with the GOMC package [32]. Using circles in a) suffices to obtain accurate virial coefficients, therefore diamond points are 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.(30). 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 [44].

6. Accurate Low Density Simulations

According to Figure 6(a), P^{sat} accuracy is sensitive to accuracy of B_2 values at the isothermal temperature (B_2^{IT}). This sensitivity is due to accumulation of errors when integrating along isotherm, such that an error in B_2^{IT} affects the A^{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 ρ shown in Figure 9(a) demonstrates the effect of varying number of molecules in MD simulation of ethane with rigid bonds. As shown in Figure 9(a), in order to match the $B_2 + B_3\rho + B_4\rho^2$ line which represents the rigorous values of $(Z-1)/\rho$, we need 3200 ethane molecules to achieve low enough uncertainties for accurate extrapolation of B_2 for the simulation times chosen. In the MD simulations shown in Figure 9(a) C-C bonds are held constant using the SHAKE algorithm [47].

The effect of using flexible bonds is shown in Figure 9(b). The systematic discrepancy from rigorous values (solid black line) as well as large uncertainties suggests not using flexible bonds at very low densities. The divergence of pressure at low densities is presumably due to the small number of intermolecular collisions at low density relative to the large number of intramolecular collisions, inhibiting the equilibration of the various components of momentum. In Supplementary Materials, we illustrate one manner of correcting for this deficiency by modifying the intramolecular virial when different virial contributions are separately accessible.

Figure 9(c) shows the low density NVT state points simulated using GOMC [32]. 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, 55, 49], Mie-UA [5, 51, 56, 57], and TIP4P/2005 [58] models were chosen for the purpose of testing 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 ρliq and within a few percent for P^{sat} and ρliq.

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. (14) and Eq. (15), while B_3 is set to zero for simplicity. Initial T^{sat} values for all the compounds shown in Figures 10 and 11 were obtained from the DIPPR [46] 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 [59]. Figure 10, 11(a), and 11(b) compare the ITIC and GEMC results for pure ethane,

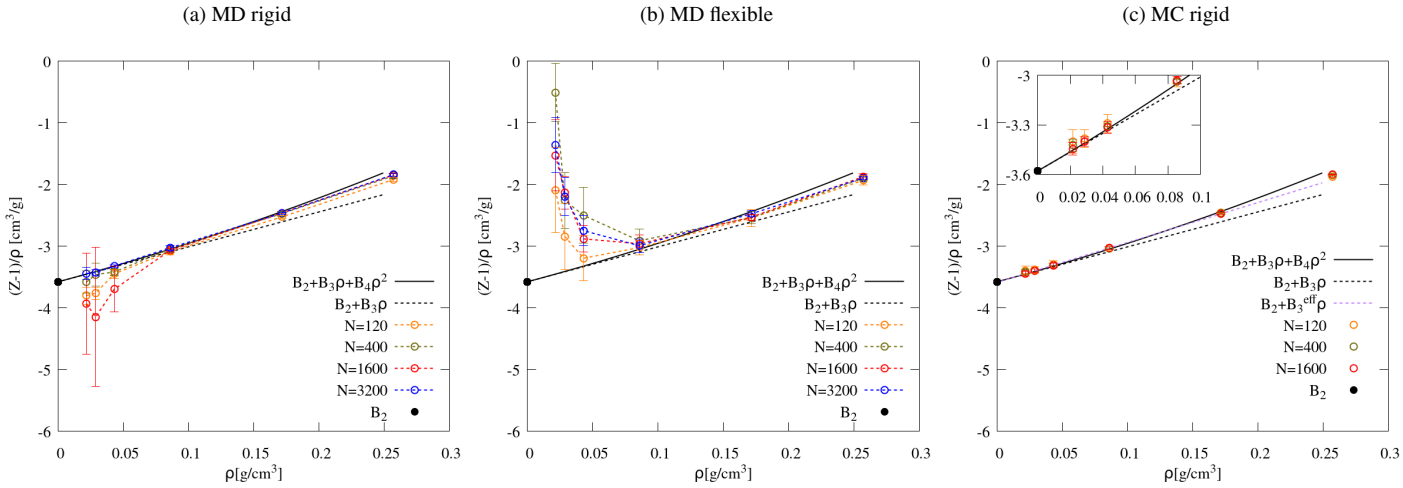


Figure 9: Effect of number of ethane molecules on compressibility factor at low densities at $T^{\mathrm{IT}} = 360$. Panel **a)** is from MD with rigid bonds (using SHAKE algorithm in LAMMPS), panel **b)** is from MD with flexible bonds (harmonic potential using multiple-time-step algorithm RESPA [48] 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 [41]. Dashed black lines represent $B_2 + B_3\rho$ line where B_{2-3} are obtained from Schultz's work. Solid black circle shows the Schultz's value of B_2. Dashed purple lines represent $B_2 + B_3^{\mathrm{eff}}\rho$ where B_2 and B_3^{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.

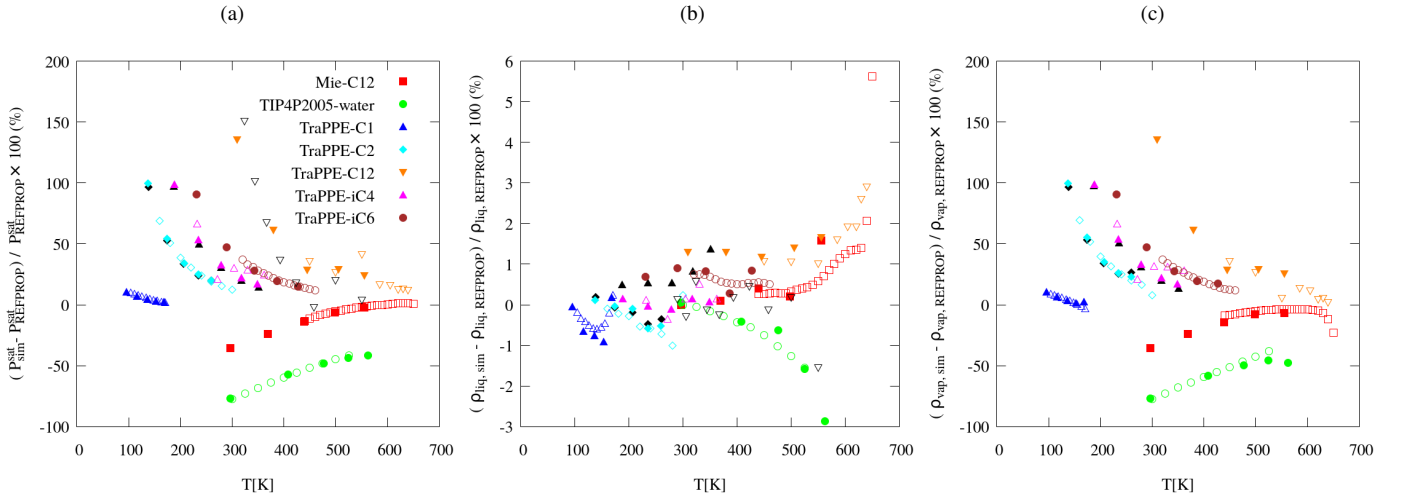


Figure 10: Comparison between ITIC (filled) with respect to GEMC and GCMC methods (unfilled). The y-axis represents deviation from REFPROP data. GEMC: TraPPE-UA *n*-dodecane (orange) [4], TraPPE-UA ethane (cyan) [4], TraPPE-UA isobutane (purple) [49]; GCMC: TraPPE methane[50] (blue), Mie-UA *n*-dodecane (red) [5], TIP4P/2005 water (green) [50], and TraPPE-UA isohexane (brown) [51]. Black filled symbols represent ITIC results simulated with GROMACS [36]. Black unfilled triangles represent TraPPE-UA *n*-dodecane from Ref. [52]. U^{dep} is computed with Eq. (28) for *n*-dodecane and isohexane by performing single molecule simulations, while U^{dep} is computed with Eq. (29) for other compounds.

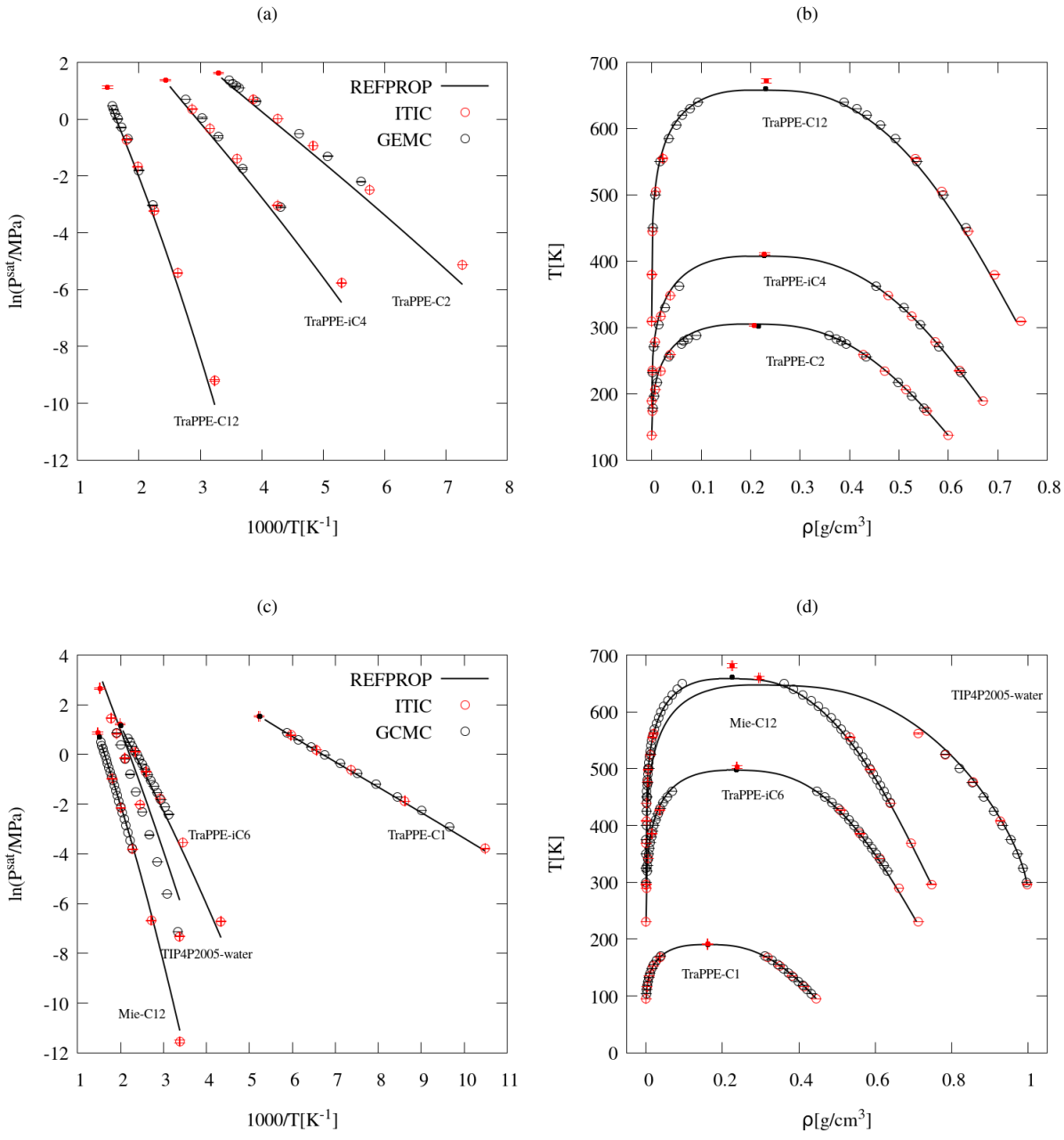


Figure 11: Comparison between ITIC method and Monte Carlo methods: GEMC (TraPPE-UA *n*-dodecane [4], TraPPE-UA ethane [4], and TraPPE-UA isobutane [49]), and GCMC (TraPPE-UA methane [50], Mie-UA *n*-dodecane [5], TIP4P/2005 water [50], and TraPPE-UA isohexane [51]). B_2 values at saturation temperatures were obtained using low density simulations described in Section 5, except for TIP4P/2005 simulation where B_2 correlation is obtained from Ref. [53, 54]. Black and red filled symbols show the critical points from literature and ITIC, respectively.

n-dodecane, and isobutane systems. *NVT* simulations at ITIC state points are performed using the Cassandra Monte Carlo [31] and GROMACS molecular dynamics [36] packages.

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^{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 [50] 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. (30) 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 $T_r^{sat} \approx 0.85$, 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. [60]. Accurate estimation of critical points by molecular simulation requires a careful and deliberate effort that includes accounting for system size effects [61]. 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_{est}^{sat}. The ITIC calculations shown so far are done for well-known molecules for which extensive experimental data are available. All force fields considered (TraPPE-UA, Mie-UA, TIP4P/2005) provide accurate representation of the T^{sat} vs. ρ_{liq} curve compared to experimental data. Therefore, the T_{est}^{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_{est}^{sat}, i.e. the temperature at which $Z \approx 0$. Once T_{est}^{sat} is calculated, the procedure in Figure 2 is followed. Figure 12(b-c) shows a good agreement between the ITIC method with GEMC results.

Alternatively, one can improve ITIC results when a reasonable estimate for T_{est}^{sat} is not available, by running the ITIC procedure iteratively, i.e. using T^{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, 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 with fixed bond lengths when feasible.

It is important for engineering applications to be able to simulate systems at temperatures as low as $T_r = 0.45$. 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 The ITIC method, hence, outperforms GEMC and GCMC when T_r is less than 0.6. This 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.

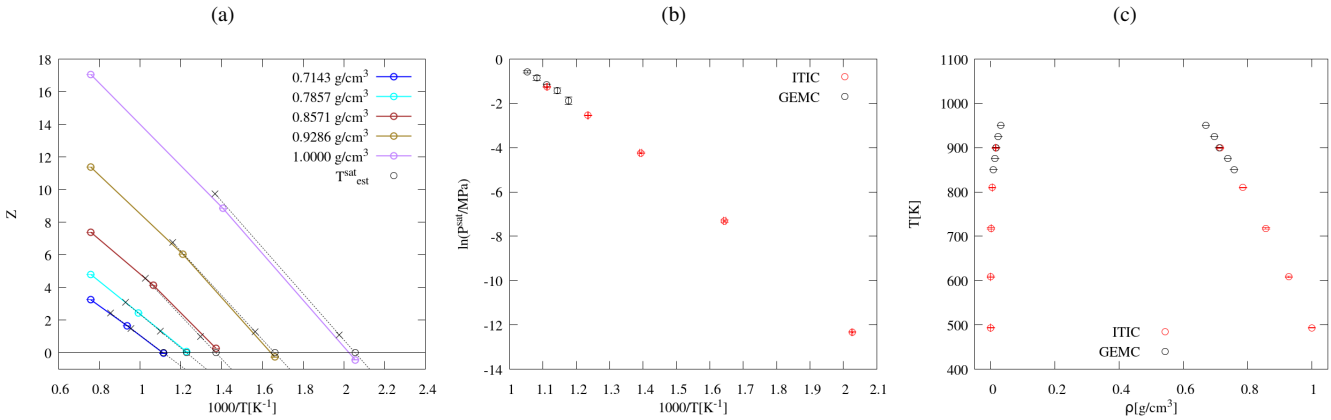


Figure 12: Simulation of 1-naphthalenyl,4-phenanthrenyl butane based on an extended TraPPE-UA model [62]. Both ITIC and GEMC results are obtained using GOMC. Black (x) symbols are simulated at two arbitrarily chosen temperatures to obtain $T_{\text{est}}^{\text{sat}}$ (black circles) for each isochore.

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.

10. Supplementary Material

See supplementary material for the isothermal/isochoric plots of A^{dep}, U^{dep}, Z, and plots of ΔH_{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.

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.

[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] 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.

[10] 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.

[11] J. R. Elliott, L. Hu, Vapor-liquid equilibria of square-well spheres, The Journal of Chemical Physics 110 (6) (1999) 3043–3048.

[12] N. F. Carnahan, Equation of State for Nonattracting Rigid Spheres, The Journal of Chemical Physics 51 (2) (1969) 635. doi:10.1063/1.1672048.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] G. Rutkai, J. Vrabec, Empirical fundamental equation of state for phosgene 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.

[20] 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.
[21] 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.
[22] 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.
[23] J. R. Elliott, C. T. Lira, Introductory chemical engineering thermodynamics, Prentice Hall PTR Upper Saddle River, NJ, 1999.
[24] R. L. Burden, J. Douglas Faires, Numerical analysis, Prindle, Weber & Schmidt, 1985.
[25] K. E. Atkinson, An introduction to numerical analysis, John Wiley & Sons, 2008.
[26] J. S. Rowlinson, Liquids and Liquid Mixtures (1982).
[27] J. S. Rowlinson, B. Widom, Molecular theory of capillarity, Courier Corporation, 2013.
[28] J. S. Rowlinson, F. Swinton, Liquids and liquid mixtures: Butterworths monographs in chemistry, Butterworth-Heinemann, 2013.
[29] 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.
[30] 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.
[31] 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.
[32] 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.
[33] 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).
[34] M. P. Allen, D. J. Tildesley, Computer simulation of liquids, Oxford university press, 2017.
[35] S. Plimpton, P. Crozier, A. Thompson, LAMMPS-large-scale atomic/molecular massively parallel simulator, Sandia National Laboratories 18.
[36] 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.
[37] 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.
[38] 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.
[39] B. Efron, Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods, Biometrika 68 (3) (1981) 589–599.
[40] J. K. Singh, D. A. Kofke, Mayer sampling: Calculation of cluster integrals using free-energy perturbation methods, Physical review letters 92 (22) (2004) 220601.
[41] A. J. Schultz, D. A. Kofke, Virial coefficients of model alkanes, The Journal of Chemical Physics 133 (10). doi:10.1063/1.3486085.
[42] 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).
[43] 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.
[44] V. I. Harismiadis, I. Szleifer, 2nd Virial-Coefficients Of Chain Molecules - A Monte-Carlo Study, Molecular Physics 81 (4) (1994) 851–866.
[45] 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.
[46] DIPPR, Design Institute for Physical Property Data (2004).
[47] 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.
[48] M. Tuckerman, B. J. Berne, G. J. Martyna, Reversible multiple time scale molecular dynamics, The Journal of chemical physics 97 (3) (1992) 1990–2001.
[49] 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.
[50] V. Shen, D. Siderius, W. Krekelberg, H. E. Hatch, NIST Standard Reference Simulation Website (2008). doi:http://doi.org/10.18434/T4M88Q.
[51] 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.
[52] 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.
[53] 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.
[54] 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.
[55] 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.
[56] 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.
[57] 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.
[58] 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.
[59] 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.
[60] 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.
[61] G. Orkoulas, M. E. Fisher, A. Panagiotopoulos, Precise simulation of criticality in asymmetric fluids, Physical Review E 63 (5) (2001) 051507.
[62] 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 Equilibriadoi:10.1016/j.fluid.2018.07.001.

Supplementary Material

Click here to download Supplementary Material: ITIC-supp.pdf