

# **Finding the arbitrary parameter $L$ in Renormalization Group Theory**

VIA fitting Monte Carlo simulations to Statistical  
Associating Fluid Theory

**Billy Edward Geerhart III**

A project presented for the degree of  
Masters of Science



Department of Physics  
Oregon State University  
August 25, 2016

# **Finding the arbitrary parameter $L$ in Renormalization Group Theory**

VIA fitting Monte Carlo simulations to SAFT

**Billy Edward Geerhart III**

## **Abstract**

I did as I was told.

Things happened as expected

Accept the flat landers universe.

# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>                                 | <b>5</b>  |
| <b>2</b> | <b>Methods</b>                                      | <b>7</b>  |
| 2.1      | Equation Summary: SAFT-VR . . . . .                 | 7         |
| 2.2      | Equation Summary: Monte Carlo Simulations . . . . . | 7         |
| 2.3      | Partition Function Seperation . . . . .             | 8         |
| 2.4      | Free Energy via Monte Carlo Simulations . . . . .   | 9         |
| 2.4.1    | MC Ideal gas free energy . . . . .                  | 9         |
| 2.4.2    | MC Hard sphere free energy . . . . .                | 10        |
| 2.4.3    | MC Dispersive free energy . . . . .                 | 12        |
| 2.5      | Finding Coexistence . . . . .                       | 14        |
| 2.6      | Computation . . . . .                               | 16        |
| 2.6.1    | Resolution reduction . . . . .                      | 16        |
| 2.6.2    | Coexistence points . . . . .                        | 17        |
| <b>3</b> | <b>Results</b>                                      | <b>18</b> |
| 3.1      | Trends in U,F . . . . .                             | 18        |
| 3.2      | Trends in S,Cv . . . . .                            | 19        |
| 3.3      | Using Cost to Estimate Best Box Size . . . . .      | 20        |
| 3.4      | Coexistence: T vs Density . . . . .                 | 20        |
| <b>4</b> | <b>Conclusion</b>                                   | <b>21</b> |
| 4.1      | Conclusion . . . . .                                | 21        |

# List of Figures

|     |   |    |
|-----|---|----|
| 2.1 | Here resolution reduction of the data is demonstrated. The data must be generated from a range of box sizes, and for each box size it is possible to run simulations From 2 atoms in a box all the way up to the desired filling fraction of 0.5; depending on the box size, a filling fraction of 0.5 can correspond to only 40 atoms or it could be as much as 100 atoms. Running simulations with every possible filling fraction will indeed produce a nice data set, but this will come at the cost of a lengthy computation time. Instead, pairs of simulations are run around filling fractions of [0.5,0.15,0.25,0.35,0.45]; the idea is each pair of simulations can be used to interpolate the thermodynamic observables to the selected filling fractions. . . . .   | 16 |
| 2.2 | Here the common tangent method is demonstrated. The free energy at a particular temperature is plotted as a function of density. The two points highlighted by red and blue X's on the curve share a common tangent; this means both points will have the same pressure and chemical potential. The left and right set of plots shows how the signal to noise ratio changes as the temperature increases. Far from the critical point the signal is relatively strong, while the near the critical point the noise becomes significant. To prevent the loss of the signal, first start at a low temperature where the signal is strong. Then gradually increase the temperature while finding common tangents. As the temperature and noise increases, the previous solutions at a slightly lower temperature can be used as a guess for the new higher temperature. . . . .  | 17 |
| 3.1 | These graphs show the Monte Carlo simulation results in terms of dispersive free energy and internal energy relative to SAFT at three different box sizes. Both sets of graphs show that a small box size has a higher energy than SAFT, while a larger box will have a lower energy than SAFT. This makes sense because as the box size increases longer wavelength fluctuations are allowed, and these longer wavelengths actually lower the allowed energy due to the decreased interactions in the low density regions being more than compensated by the increase in interactions in the high density regions. Given a constant box size, the high filling fraction simulations cannot fluctuate as much as the low filling fraction simulations, which explains the bulge in the two lower plots. . . . .   | 18 |
| 3.2 | These graphs show the Monte Carlo simulations results in terms of dispersive entropy and heat capacity per particle relative to SAFT at three different box sizes. It is interesting to see the entropy trends are nearly the inverse of the heat capacity trends. Both entropy and heat capacity require a derivative to find, it is just that entropy is the negative derivative of Free energy while heat capacity is the positive derivative of internal energy. Given fig. 3.1 showed the internal and free energy spread out in a similar fashion as the temperature was increased, it isn't surprising entropy and heat capacity have inverse trends. Regarding the trend itself, the graphs suggest entropy of a small box will have more entropy than SAFT, while a large box will have a lower entropy than SAFT. This trend can be explained by the particles binding together in ways that SAFT could not take into account, and this extra binding will decrease the entropy simply because some configurations are now favored over other configurations. . . . | 19 |
| 3.3 | These graphs show the cost function as a function of the box size for the dispersive free energy, internal energy, entropy, and heat capacity. The cost function is defined as the mean square difference between the Monte Carlo simulations and SAFT evaluated at a filling fraction of [0.15,0.25,0.35,0.45] and evaluated for $0.6 < T < 1.5$ . The dispersive free energy and internal energy show a minimum near $L=7.6$ , while any box size less than $L=8.0$ seems to be just fine for the dispersive entropy and heat capacity. . . . .   | 20 |
| 3.4 | Here the gas-liquid coexistence temperature as a function of filling fraction for three different Monte Carlo simulations are compared to SAFT. The box size $L=6.61$ clearly has a higher critical temperature than SAFT, while the box size $L=10.0$ clearly has a lower critical temperature than SAFT. . . . .  | 20 |

# Introduction

Outline: Introduction to liquid vapor coexistence Demonstrate the problem Fix the problem what I did

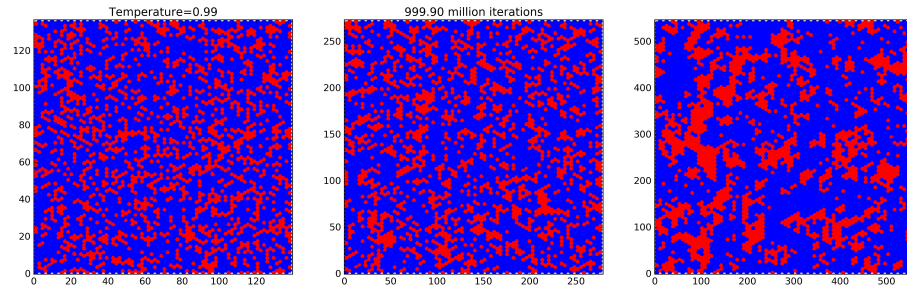
The liquid-vapor phase coexistence can be found by just placing some liquid into a sealed container. Some of the liquid will evaporate into the jar creating a vapor pressure. Provided we placed enough liquid into the jar, the system will now contain both a liquid and a vapor at the same time. By sealing the liquid into a jar, we have forced the pressure of the liquid to be the as the pressure of the gas. By allowing the atoms to swap between the liquid and gas phase, we have also allowed the chemical potential of the liquid to be the same as that of the gas.

Although we can find the phase coexistence by physically sealing a liquid in a jar, an equivalent computational method is to use an empirically determined fluid model which we can use to run simulations that also equalize the pressure and chemical potential. The problem with the computational method is the empirical model requires fitting a few parameters to data, data which probably already includes points on the phase coexistence curve. The main benefit to the computational approach is that once the empirical parameters are found, we can then create the phase coexistence of a mixture of fluids using the previously determined parameters.

One of the difficulties with the computer simulations is computation time. A simulation with a handful of atoms can be finished within 20 minutes, yet a simulation with a 100 atoms can take as long as 4 days. To recreate the phase coexistence graph requires setting the pressure and chemical potential of both phases equal to each other, but this is equivalent to plotting the free energy as a function of density and searching for a common cotangent. For even a simulation that uses a small box, the problem is plotting free energy as a function of density requires many simulations at various densities; some simulations containing only a handful of atoms while other simulations containing as much as a couple hundred atoms. Total computation time for 10 computer cores to create the free energy as a function of density graph can take as long as three weeks.

Naturally there is a tradeoff between a realistic computation time for a project and how realistic the results are going to be. The simulations use periodic boundary conditions which is just the thermodynamic way of saying the boundary should not affect the bulk properties. The key part being whatever is in the simulation is considered the bulk, which is a problem since a small simulation can hardly be considered the bulk. To get around this problem, the simulation size is usually increased until the simulation does represent the bulk.

Normally increasing the simulation size is just fine, but there is a regime where the bulk properties depend on length scales of the order of a millimeter. This regime occurs near the critical point where the liquid density approaches that of the vapor. Normally the liquid and vapor are entirely separated due to the surface tension. However; as the system approaches the critical point, the surface tension approaches zero which means the liquid and vapor will start to mix. Although figure 1 shows an extreme example of a liquid completely mixing with vapor, even under normal circumstances there exist natural fluctuations in the density of either phase. It just so happens that these fluctuations are significant as the density of the liquid approaches the density of the gas at the critical point. Basically the fluctuations within each phase would normally be rather small, but near the critical point the fluctuations are relatively big since the gas and liquid phases are morphing into each other. This is a problem since it requires the simulations to be about the same size as that of the largest fluctuation. For example, a 1mm length box would contain roughly  $10^{19}$  atoms, which would require an unfeasible computation time.



Renormalization group theory(RGT) can be applied to get around this large fluctuation problem. Simply put, RGT is a method in which fine details are omitted from calculations while still maintaining accuracy. This is done by transforming the original partition function into a smaller partition function, but the smaller system is renormalized to the original partition function. [add details...]

In particular, RGT is applied through an interaction wave length. The zeroth order is just the ideal gas term in which all atoms only interact with the walls of the enclosure. The next order is applied via SAFT-VR in which a mean interaction and an approximate wavelength interaction is applied. Longer wavelength interactions can then be applied using RGT via a iterative method that steadily increases the size of the box.

The main problem this project addresses is the transition from SAFT-VR to the application of RGT. Within RGT there is an arbitrary base wavelength  $L$  which is usually fit to the data. However; this arbitrary wavelength shouldn't really be an arbitrary parameter as SAFT-VR already includes interactions at some unknown wavelength. This project attempts to find the base wavelength that SAFT-VR implements, and tests how the results of RGT change when using this base wavelength.

The base wavelength is found by using Monte Carlo Markov Chain (MCMC) simulations at various box sizes. The idea is that a MCMC simulation using a fixed box size of  $L_0$  will incorporate interactions at this box size. The box size is then varied until the results from the MCMC simulations fit the results from SAFT-VR. Once the optimal box size is found, RGT is implemented using this optimal box size and the results are compared to previous results.

# Methods

The methods section is partitioned into theory, which define the equations used, and the actual implementation through computational techniques. The main point of the theory is to estimate the free energy  $F$  from which any of the thermodynamic variables can be found. For example, the entropy, chemical potential, or pressure can be found by the differential  $dF = -S \cdot dT - P \cdot dV + \mu \cdot dN$ . Further, the internal energy can be found by  $U = F + T \cdot S$ . Once the free energy is found, the next goal is to fit the Monte Carlo simulations to SAFT by using a cost function that compares each thermodynamic variable from the Monte Carlo simulations to SAFT. The comparisons are constrained to a temperature range of  $0.45 < T/T_c < 1.13$  where  $T_c$  is the critical point using SAFT; the lower bound is chosen mostly to decrease the computation time of the simulations, while the upper limit is chosen to stay away from the well known hard sphere behavior. To decrease computation times even further, the comparisons are again constrained to filling fractions of  $[0.15, 0.25, 0.35, 0.45]$ . Once the cost can be determined for a variety of box sizes, a 'best' box size is chosen that decreases the overall cost. The final step is to compare this best box size to slightly larger and smaller box sizes by looking at the coexistence densities of the liquid and gas phases.

## Equation Summary: SAFT-VR

$$\begin{aligned} \text{let } n &= \frac{N}{V} \\ \eta &= \text{packing Fraction} \\ \Lambda &= \text{de Broglie wavelength} \end{aligned} \quad g^{HS}(1; \eta_{eff}) = \frac{1 - \eta_{eff}/2}{(1 - \eta_{eff})^3} \quad (2.6)$$

$$\frac{F}{N \cdot k \cdot T} = a^{IDEAL} + a^{HS} + \beta \cdot a_1^{SW} + \beta^2 \cdot a_2^{SW} \quad n_{eff} = c_1 \cdot \eta + c_2 \cdot \eta^2 + c_3 \cdot \eta^3 \quad (2.1) \quad (2.7)$$

$$\begin{aligned} a^{IDEAL} &= \ln(n \cdot \Lambda^3) - 1 \quad (2.2) \\ a^{HS} &= -\ln(1 - 4 \cdot \eta) \quad (2.3) \end{aligned} \quad \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 2.25855 & -1.50349 & 0.249434 \\ -0.669270 & 1.40049 & -0.827739 \\ 10.1576 & -15.0427 & 5.30827 \end{pmatrix} \times \begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \end{pmatrix} \quad (2.8)$$

$$a_1^{SW} = a_1^{VDW} \cdot g^{HS}(1; \eta_{eff}) \quad (2.4) \quad a_2^{SW} = \frac{1}{2} \cdot \epsilon K^{HS} \eta \cdot \frac{\partial a_1^{SW}}{\partial \eta} \quad (2.9)$$

$$a_1^{VDW} = -4 \cdot \eta \cdot \epsilon \cdot (\lambda^3 - 1) \quad (2.5) \quad K^{HS} = \frac{(1 - \eta)^4}{1 + 4 \cdot \eta + 4 \cdot \eta^2} \quad (2.10)$$

## Equation Summary: Monte Carlo Simulations

$$\begin{aligned} Z &= Z_{ideal} \cdot Z_{HS} \cdot Z_{disp} \\ Z_{ideal} &= Z_{ideal}^{SAFTVR} \\ Z_{HS} &= e^{-S_{HS}/k} \end{aligned}$$

$$Z_{disp} = \frac{Z_{interaction}}{Z_{HS}} = \frac{\sum_i \alpha \cdot D(E_i) \cdot e^{-\beta \cdot E_i}}{\lim_{T \rightarrow \infty} \sum_i \alpha \cdot D(E_i) \cdot e^{-\beta \cdot E_i}} = \frac{\sum_i D(E_i) \cdot e^{-\beta \cdot E_i}}{\sum_i D(E_i)}$$

Note that  $S_{HS}$  is the hard sphere entropy. Past work by Brenden has shown the hard sphere entropy of the simulations can be modeled by the Carnahan-Starling equation of state, so all simulations used in this project use the entropy found via the Carnahan-Starling equation of state. Given this, if one wanted to find the hard sphere entropy using computaiton we can use a known starting point of a very large box at  $T = \infty$ . This very large box acts like an ideal gas, which means the hard sphere entropy is zero. Next this very large box is randomly scaled down to a smaller box. The scaling maps actually maps every atom within the larger box down into the smaller box while preserving angular orientation. This means a random squeeze can actually fail due hard sphere collisions between atoms when the box everything is scaled to the smaller box. Given the probability to succeed in scaling the box down as  $P_{success}$ , the change in hard sphere entropy is just  $\Delta S_{HS} = k \cdot \ln(P_{success})$ . This scaling can be done in as many small steps as necessary to scale the very large box down to the actual box size; and the net sum in the change in entropy is then  $S_{HS}$ .

## Partition Function Seperation

The partition function can be partitioned into high and low temperature regimes. At low temperatures the interactions dominate, while at high temperatures the system acts like a hard sphere system in which the entropy dominates. Both regimes can be partitioned yet again by seperating the kinetic energies into an ideal gas term. In terms of equations, we start with the partition function and sepearate the summation into the three parts:

$$\begin{aligned} Z &= \frac{1}{N!} \sum_i e^{-\beta \cdot E_i} \\ &= \frac{1}{N!} \int \dots \int e^{-\beta \cdot [\{P_1^2/2m + P_2^2/2m + \dots\} + V(\vec{R}_1, \vec{R}_2, \dots)]} d\vec{P}_1 \cdot d\vec{P}_2 \dots d\vec{P}_N \cdot d\vec{R}_1 \cdot d\vec{R}_2 \dots d\vec{R}_N \end{aligned} \quad (2.11)$$

The momentum integrals are independent of the position integrals, so we sepearate the momentum integrals which we then incorporate into an ideal gas term by multiplying by  $V^N/V^N$ .

$$Z = Z_{ideal} \cdot \frac{1}{V^N} \cdot \int \dots \int e^{-\beta \cdot V(\vec{R}_1, \vec{R}_2, \dots)} d\vec{R}_1 \cdot d\vec{R}_2 \dots d\vec{R}_N \quad (2.12)$$

$$\text{where } Z_{ideal} = V^N \cdot \frac{1}{N!} \int \dots \int e^{-\beta \cdot [P_1^2/2m + P_2^2/2m + \dots]} d\vec{P}_1 \cdot d\vec{P}_2 \dots d\vec{P}_N \quad (2.13)$$

After taking into account the ideal gas term, the remaining potential energy term is seperated into the low temperature and high temperature regime. At high temperatures, the atom energies are much higher than the potential well; this means the finite potential well can be considered zero. However; even at infinite temperature the atoms are assumed to be a hard sphere. The expected hard sphere behavior is seperated by multiplying by  $Z_{HS}/Z_{HS}$ .

$$Z = Z_{ideal} \cdot Z_{HS} \cdot \frac{1}{Z_{HS}} \cdot \frac{1}{V^N} \cdot \int \dots \int e^{-\beta \cdot V(\vec{R}_1, \vec{R}_2, \dots)} d\vec{R}_1 \cdot d\vec{R}_2 \dots d\vec{R}_N \quad (2.14)$$



$$\text{where } Z_{HS} = \frac{1}{V^N} \int \cdots \int e^{-\beta \cdot V_{HS}(\vec{R}_1, \vec{R}_2, \dots)} d\vec{R}_1 \cdot d\vec{R}_2 \cdots d\vec{R}_N \quad (2.15)$$

After taking into account the expected high temperature hard sphere behavior, the remaining term is just the ratio between the interaction and hard sphere partition functions; this term is called the dispersive partition function. I.e.

$$Z = Z_{ideal} \cdot Z_{HS} \cdot \frac{Z_{interaction}}{Z_{HS}} = Z_{ideal} \cdot Z_{HS} \cdot Z_{disp} \quad (2.16)$$

$$\text{where } Z_{interaction} = \frac{1}{V^N} \cdot \int \cdots \int e^{-\beta \cdot V(\vec{R}_1, \vec{R}_2, \dots)} d\vec{R}_1 \cdot d\vec{R}_2 \cdots d\vec{R}_N \quad (2.17)$$

$$\text{and } Z_{disp} = \frac{Z_{interaction}}{Z_{HS}} \quad (2.18)$$

## Free Energy via Monte Carlo Simulations

Given the separated partition function, the free energy of the Monte Carlo simulations are just the sum of the ideal term, the hard sphere term, and the dispersive term. The ideal term is ignored because it stays the same between theory and the Monte Carlo simulations. Although the hard sphere energy can be determined by squeezing an infinitely large box down to the required box size, previous work by Brenden has already shown the hard sphere entropy follows the Carnahan-Starling entropy. This is fortunate since the hard sphere entropy computations take a relatively long time. Finally, the dispersion term is estimated using a \*method\* that finds the relative density of states; this density of states can then be used to find the dispersive free energy at temperatures higher than some constant  $T_0$ . The computation time increases as  $T_0$  is decreased, so  $T_0$  is set to roughly a half of the critical temperature of SAFT.

### MC Ideal gas free energy

To find the ideal gas free energy we can use previous results based on knowing the entropy of the system which we can then use to find the free energy by  $F_{ideal} = U_{ideal} - T \cdot S_{ideal}$ .

$$U_{ideal} = 3/2 \cdot N \cdot k \cdot T \quad (2.19)$$

$$S_{ideal*} \approx k \cdot N \left( \ln \left[ \frac{V}{N} \left( \frac{4 \cdot \pi \cdot m}{3 \cdot h^2} \cdot \frac{U}{N} \right)^{3/2} \right] + \frac{5}{2} \right) \quad (2.20)$$

The Sackur-Tetrode equation with the Stirling approximation actually does not make physical sense as the entropy per particle at constant temperature only depends on density. A simple example can show why this doesn't make sense. Consider two atoms in boxes separated by a partition, and compare the entropy per particle to when the partition in the box has been removed. Once the partition has been removed, the two atoms can explore many more configurations than previously, which means the entropy should have increased once the two atoms were allowed to mix. Unfortunately the Sackur-Tetrode equation with Stirlings approximation

would predict the entropy per particle with or without the partition should be the same. This is a bit of a problem since every box in a Monte Carlo simulation will contain exactly  $N$  particles, which means the entropy is actually lower than a real ideal gas.

The entropy problem can be fixed by using the Sackur-Tetrode equation without Stirling's approximation (eq 2.21). All we have to do is use a constant density and compare the entropy per particle of a small box to the entropy per particle of a very large box. The difference in entropy should show how the ideal gas entropy within a small box will deviate from the expected ideal gas entropy.

$$S_{ideal} = k \cdot N \left( \ln \left[ V \cdot N!^{-1/N} \left( \frac{4 \cdot \pi \cdot m}{3 \cdot h^2} \cdot \frac{U}{N} \right)^{3/2} \right] + \frac{3}{2} \right) \quad (2.21)$$

The entropy problem was not taken into account since the simulations should really just be a tool to explore the region that we don't know. At high temperatures we already know the system will act like an ideal gas in terms of the momentum integrals and a hard sphere system in terms of collisions. Basically keeping track of every entropy change will affect the boundary conditions at high temperature; for this reason the entropy problem is ignored for this project.

## MC Hard sphere free energy

As been said previously, the Carnahan-Starling equation of state was used to find the hard sphere partition function. Although if one wanted to calculate the partition function via simulations, the hard sphere partition function can be found by calculating the excess entropy at  $T \rightarrow \infty$ . This can be achieved by the following:

$$Z_{exc} = Z_{HS} \cdot Z_{disp} \Rightarrow \quad (2.22)$$

$$F_{exc} = -k \cdot T \cdot \ln(Z_{exc}) = -k \cdot T \cdot \ln(Z_{HS}) - k \cdot T \cdot \ln(Z_{disp}) \Rightarrow \quad (2.23)$$

$$S_{exc.\infty} = \lim_{T \rightarrow \infty} - \left( \frac{\partial F_{exc}}{\partial T} \right)_{V,N} \quad \text{but} \quad \lim_{T \rightarrow \infty} Z_{exc} = \text{constant} \Rightarrow \quad (2.24)$$

$$S_{exc.\infty} = -k \cdot \ln(Z_{HS.\infty}) - k \cdot \ln(Z_{disp.\infty}) \quad (2.25)$$

$$\text{but } Z_{HS} \text{ is also constant and } Z_{disp.\infty} = \frac{Z_{interaction.\infty}}{Z_{HS}} = Z_{HS}/Z_{HS} = 1 \Rightarrow \quad (2.26)$$

$$S_{exc.\infty} = -k \cdot \ln(Z_{HS}) - k \cdot \ln(1) = -k \cdot \ln(Z_{HS}) \Rightarrow \quad (2.27)$$

$$Z_{HS} = e^{-S_{exc.\infty}/k} \quad (2.28)$$

To find the hard sphere partition function, we need to find the maximum excess entropy as  $T \rightarrow \infty$ . The method used to find the maximum excess entropy starts by considering  $N$  atoms in an extremely large box at a very large temperature. The

atoms in the large box should behave like an ideal gas, so the entropy of this large box is a known starting point. Then all atom positions are scaled relative to a corner as the origin. Although each individual atom can now exist within a smaller box, it is not guaranteed an atom will not overlap with another atom after all positions are scaled to the smaller box.

What we can observe is the probability that a particular scaling fails, and from this probability the change in entropy from the bigger box down to the smaller box can be calculated. The details are as follows:

$$F = U - T \cdot S \Rightarrow S = (U - F)/T \approx -F/T = k \cdot \ln(Z) \quad (2.29)$$

$$\text{let } Z_{small} = \text{partition function of the smaller box} \quad (2.30)$$

$$\text{let } Z_{big} = \text{partition function of the bigger box} \quad (2.31)$$

$$\Delta S = S_{small} - S_{big} = k \cdot \ln(Z_{small}) - k \cdot \ln(Z_{big}) = k \cdot \ln(Z_{small}/Z_{big}) \quad (2.32)$$

From here note that although every configuration in the bigger box is not guaranteed to scale down into the smaller box, it is guaranteed that the smaller box can be expanded into the bigger box. This means there is a mapping from the partition function of the smaller box to the bigger box; as such we can break  $Z_{big}$  into valid and invalid configurations where the valid configurations are guaranteed to shrink down into the smaller box while the invalid configurations will not. The  $\Delta S$  can then be simplified by making a connection between  $Z_{small}$  and  $Z_{valid}$ .

$$\text{let } Z_{big} = Z_{valid} + Z_{invalid} \quad (2.33)$$

$$\text{note } Z_{small} \text{ maps into } Z_{valid} \Rightarrow Z_{small} = \alpha \cdot z_{valid} \text{ for some constant } \alpha \quad (2.34)$$

This constant  $\alpha$  can be found directly by looking at the partition function of both the big and small box. All configurations in the small box can be mapped into the valid regions of the large box which means to find  $\alpha$  all we have to do is rewrite  $Z_{small}$  in terms of  $Z_{big}$ .

$$Z_{small} = Z_{ideal} \frac{1}{V^N} \int \dots \int e^{-\beta \cdot V(\vec{R}_1, \vec{R}_2, \dots)} d\vec{R}_1^{small} \cdot d\vec{R}_2^{small} \dots d\vec{R}_N^{small} \quad (2.35)$$

From here make the change of variables for each spatial coordinate to remap the integral from the small partition function into the big partition function.

$$\text{note } d\vec{R}_1^{small} = d\vec{R}_1^{big} \cdot \frac{L_{small}^3}{L_{big}^3} \Rightarrow \quad (2.36)$$

$$Z_{small} = Z_{ideal} \frac{1}{V^N} \int \dots \int e^{-\beta \cdot V(\vec{R}_1, \vec{R}_2, \dots)} d\vec{R}_1^{small} \cdot \frac{L_{small}^3}{L_{big}^3} \cdot d\vec{R}_2^{small} \cdot \frac{L_{small}^3}{L_{big}^3} \dots d\vec{R}_N^{small} \cdot \frac{L_{small}^3}{L_{big}^3} \Rightarrow \quad (2.37)$$

$$Z_{small} = \left( \frac{L_{small}}{L_{big}} \right)^{3N} \cdot Z_{big} = \left( \frac{V_{small}}{V_{big}} \right)^N \cdot Z_{big} \doteq \alpha \cdot Z_{big} \Rightarrow \quad (2.38)$$

$$\alpha = (V_{small}/V_{big})^N \quad (2.39)$$

Now that  $\alpha$  is known, we can get back to finding the entropy.

$$\Delta S = k \cdot \ln(Z_{small}/Z_{big}) = k \cdot \ln(\alpha \cdot Z_{valid}/Z_{big}) \quad (2.40)$$

$$\Delta S = k \cdot \ln\left(\frac{V_{small}^N}{V_{big}^N} \cdot Z_{valid}/Z_{big}\right) \quad (2.41)$$

$$\text{but } Z_{valid}/Z_{big} = \text{probability of success when scaling} = P_{success} \Rightarrow \quad (2.42)$$

$$\Delta S = k \cdot \ln\left(\frac{V_{small}^N}{V_{big}^N} \cdot P_{success}\right) \quad (2.43)$$

At this point this  $\Delta S$  represents the total change in entropy from squeezing the box. However; only the change in excess free entropy was wanted. To compensate, simply subtract the change in entropy due to the ideal gas using constant temperature and constant number of atoms.

$$\Delta S_{exc.\infty} = \Delta S_{\infty} - \Delta S_{ideal.\infty} \quad (2.44)$$

$$\text{but from eq 2.20 } \Delta S_{ideal.\infty} = k \cdot N \cdot \ln(V_{small}/V_{big}) \Rightarrow \quad (2.45)$$

$$\Delta S_{exc.\infty} = k \cdot \ln\left(\frac{V_{small}^N}{V_{big}^N} \cdot P_{success}\right) - k \cdot N \cdot \ln(V_{small}/V_{big}) \quad (2.46)$$

$$\Delta S_{exc.\infty} = k \cdot \ln(P_{success}) \quad (2.47)$$

## MC Dispersive free energy

The dispersive term is found by estimating the density of states using a \* method. The main problem is the density of states is found by sampling the energy levels, but the relative distribution of energy levels is not unique to a single density of states. In other words, the partition function can be changed by an arbitrary factor without actually changing the total energy. This can be seen as follows:

$$\text{Suppose } Z = \sum_i D(E_i) \cdot e^{-\beta \cdot E_i} \quad (2.48)$$

$$\text{then } U = \sum_i E_i \cdot D(E_i) \cdot e^{-\beta \cdot E_i} / Z \quad (2.49)$$

$$\text{Now suppose we let } D_2(E_i) = \alpha \cdot D(E_i) \quad (2.50)$$

$$\text{then } Z_2 = \sum_i \alpha \cdot D(E_i) \cdot e^{-\beta \cdot E_i} = \alpha \cdot Z \Rightarrow \quad (2.51)$$

$$U_2 = \sum_i E_i \cdot D_2(E_i) \cdot e^{-\beta \cdot E_i} / Z_2 \quad (2.52)$$

$$U_2 = \sum_i E_i \cdot \alpha \cdot D(E_i) \cdot e^{-\beta \cdot E_i} / (\alpha \cdot Z) \quad (2.53)$$

$$U_2 = \sum_i E_i \cdot D(E_i) \cdot e^{-\beta \cdot E_i} / Z = U \quad (2.54)$$

Basically the partition function will define the total energy at all temperatures, but knowing the total energy at all temperatures will not define a unique partition function. Specifically, the partition function up to some arbitrary factor will give the same total energy. This is a bit problematic considering the method to find the density of states makes observations of the total energy to estimate the density of states, which then results in a density of states that is only defined up to some unknown factor.

To get around this problem of an unknown factor in the density of states, note that the dispersive partition function will go towards one as T approaches infinity. This happens because  $Z_{disp}$  is the ratio of the interaction term to the hard sphere term, and the interaction term approaches the hard sphere term as the temperature increases. Essentially we can leave the unknown  $\alpha$  term alone, and it will show up in both the interaction and hard sphere partition function.

That is:

$$Z_{disp} = \frac{Z_{interaction}}{Z_{HS}} = \frac{\sum_i \alpha \cdot D(E_i) \cdot e^{-\beta \cdot E_i}}{\lim_{T \rightarrow \infty} \sum_i \alpha \cdot D(E_i) \cdot e^{-\beta \cdot E_i}} = \frac{\sum_i D(E_i) \cdot e^{-\beta \cdot E_i}}{\sum_i D(E_i)} \quad (2.55)$$

## Finding Coexistence

At phase coexistence we want  $\mu_{liquid} = \mu_{gas}$  and  $P_{liquid} = P_{gas}$ , so provided we know free energy we can simply find coexistence by the following constraints:

$$\mu = \left( \frac{\partial F(T, V1, N)}{\partial N} \right)_{T, V1} \Big|_{N1} = \left( \frac{\partial F(T, V1, N)}{\partial N} \right)_{T, V1} \Big|_{N2} \quad (2.56)$$

$$P = \left( \frac{\partial F(T, V, N1)}{\partial V} \right)_{T, N1} \Big|_{V1} = \left( \frac{\partial F(T, V, N2)}{\partial V} \right)_{T, N2} \Big|_{V2} \quad (2.57)$$

Notice that one set of derivatives are at constant V while the other set of derivatives are at constant N. Luckily this means we can simplify the derivatives by instead considering the free energy per volume; the reason this helps is both the free energy and volume are extensive quantities, which means the ratio  $F/V$  must be an intensive quantity which further means  $F/V = f$  must only depend on intensive parameters. I.e.  $F(T, N, V)/V = f(T, N/V) = f(T, n)$ . First start simplifying eq. 2.56 by:

$$\mu = \left( \frac{\partial F(T, V, N)}{\partial N} \right)_{T, V} \cdot \frac{V}{V} = \left( \frac{\partial (F(T, V, N)/V)}{\partial (N/V)} \right)_{T, V} = \left( \frac{\partial f(T, n)}{\partial n} \right)_{T, V} \quad (2.58)$$

At this point we know  $\mu_1 = \mu_2$  which combined with eq. 2.58 just indicates that if we plotted the free energy as a function density, then coexistence must occur when the instantaneous slope at one density matches the instantaneous slope at another density. This is useful to know, but it isn't enough to pin down a unique coexistence point. As such we must move on to simplifying eq. 2.57:

$$P = \left( \frac{\partial F(T, V, N) \cdot V/V}{\partial V} \right)_{T, N} = \left( \frac{\partial f(T, n) \cdot V}{\partial V} \right)_{T, N} = f + V \cdot \left( \frac{\partial f(T, n)}{\partial V} \right)_{T, N} \quad (2.59)$$

To simplify the last partial, use the cyclic chain rule by holding T constant throughout the process while considering the set of variables  $(f, N, V)$ . Proceed as follows:

$$\begin{aligned} \left( \frac{\partial X}{\partial Y} \right)_Z \cdot \left( \frac{\partial Y}{\partial Z} \right)_X \cdot \left( \frac{\partial Z}{\partial X} \right)_Y &= -1 \rightarrow \\ \left( \frac{\partial f(T, n)}{\partial V} \right)_{N, T} \cdot \left( \frac{\partial V}{\partial N} \right)_{f(T, n), T} \cdot \left( \frac{\partial N}{\partial f(T, n)} \right)_{V, T} &= -1 \end{aligned} \quad (2.60)$$

But we can simplify the last partial since from eq. 2.58 we already know that  $\left( \frac{\partial N}{\partial f(T, n)} \right)_{V, T} = \frac{V}{\mu}$ . Further, we can simplify the middle partial by noting  $f(T, n) = \text{constant} \Rightarrow f(T, N/V) = \text{constant} \Rightarrow N/V = n = \text{constant} \Rightarrow \left( \frac{\partial V}{\partial N} \right)_{f(T, n), T} = 1/n$ . Putting this all together into 2.60 we get:

$$\left( \frac{\partial f(T, n)}{\partial V} \right)_{N, T} \cdot \frac{1}{n} \cdot \frac{V}{\mu} = -1 \quad (2.61)$$

Finally we can use eq. 2.61 inside eq. 2.59 to achieve the following:

$$P = f + V \cdot \left( \frac{n \cdot \mu}{V} \right) = f - n \cdot \mu \quad (2.62)$$

Having a simplified result for pressure, we can now apply the coexistence constraint that  $P_1 = P_2 \Rightarrow f1 + n1 \cdot \mu = f2 + n2 \cdot \mu$ . To compare this result with eq. 2.58, solve for  $\mu$  to get the following results:

$$\mu = \frac{f2 - f1}{n2 - n1} \quad (2.63)$$

$$\text{from eq. 2.58: } \mu = \left( \frac{\partial f(T, n)}{\partial n} \right)_{T, V1} \Big|_{n1} = \left( \frac{\partial f(T, n)}{\partial n} \right)_{T, V} \Big|_{n2} \quad (2.64)$$

The above constraints indicate that coexistence only occurs when the average slope must be equal to the instantaneous slope. This is equivalent to plotting the free energy as a function of density and finding a common tangent.

## Computation

### Resolution reduction

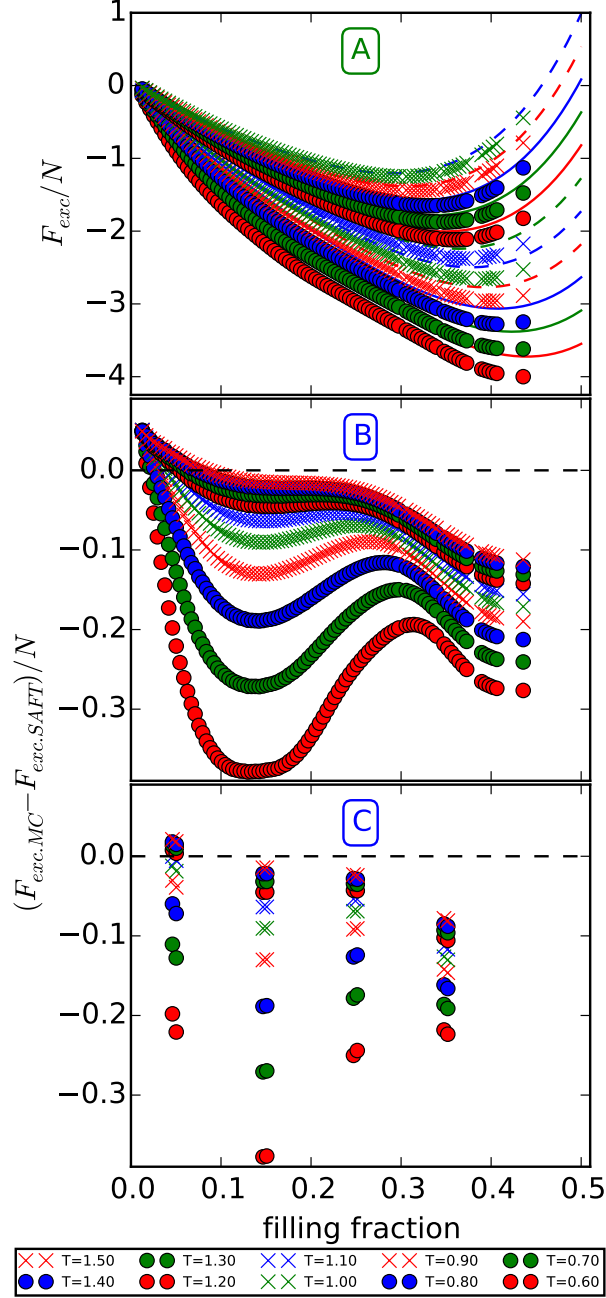


Figure 2.1: Here resolution reduction of the data is demonstrated. The data must be generated from a range of box sizes, and for each box size it is possible to run simulations From 2 atoms in a box all the way up to the desired filling fraction of 0.5; depending on the box size, a filling fraction of 0.5 can correspond to only 40 atoms or it could be as much as 100 atoms. Running simulations with every possible filling fraction will indeed produce a nice data set, but this will come at the cost of a lengthy computation time. Instead, pairs of simulations are run around filling fractions of [0.5,0.15,0.25,0.35,0.45]; the idea is each pair of simulations can be used to interpolate the thermodynamic observables to the selected filling fractions.



### Coexistence points

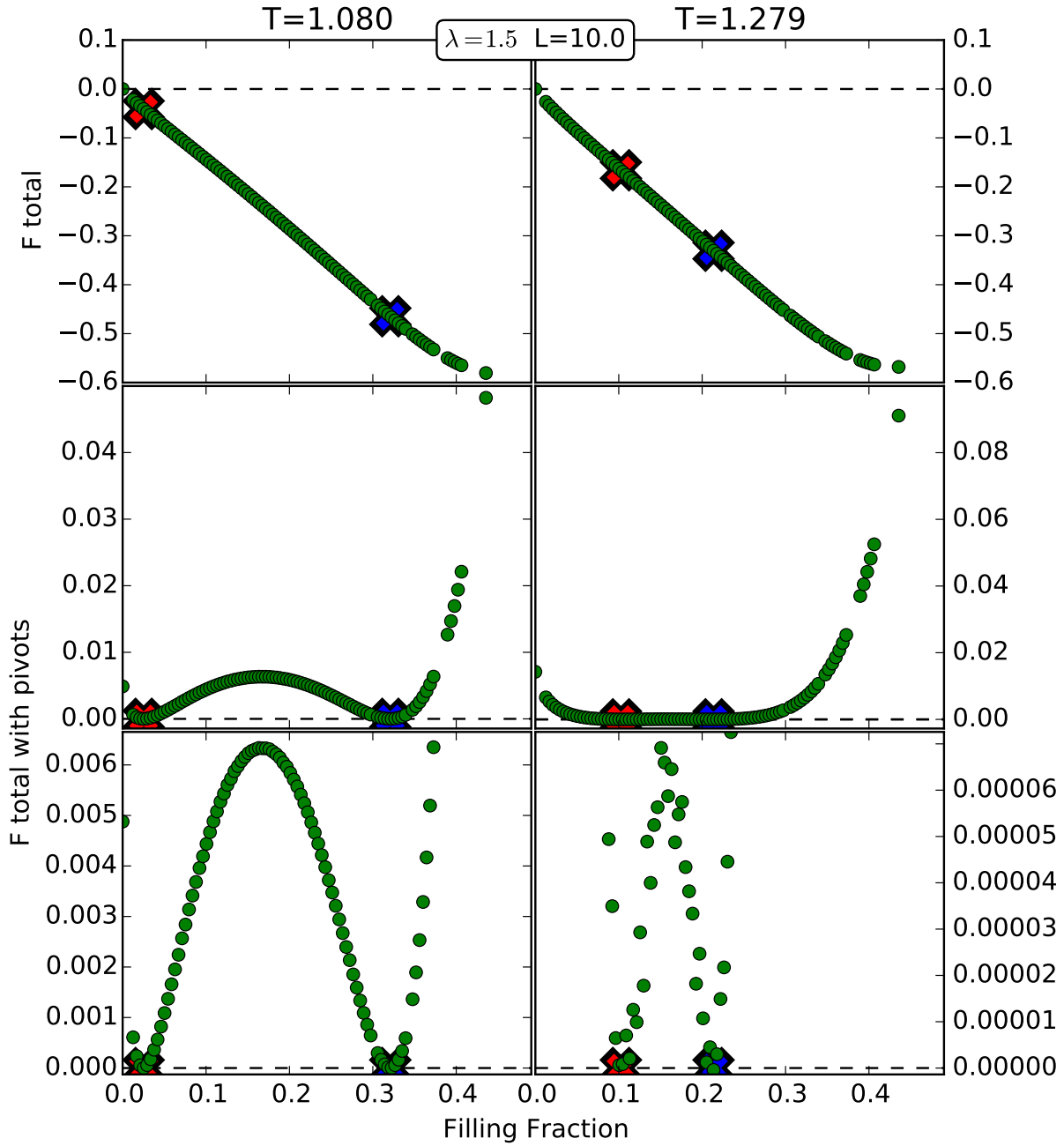


Figure 2.2: Here the common tangent method is demonstrated. The free energy at a particular temperature is plotted as a function of density. The two points highlighted by red and blue X's on the curve share a common tangent; this means both points will have the same pressure and chemical potential. The left and right set of plots shows how the signal to noise ratio changes as the temperature increases. Far from the critical point the signal is relatively strong, while near the critical point the noise becomes significant. To prevent the loss of the signal, first start at a low temperature where the signal is strong. Then gradually increase the temperature while finding common tangents. As the temperature and noise increases, the previous solutions at a slightly lower temperature can be used as a guess for the new higher temperature.

# Results

## Trends in U,F

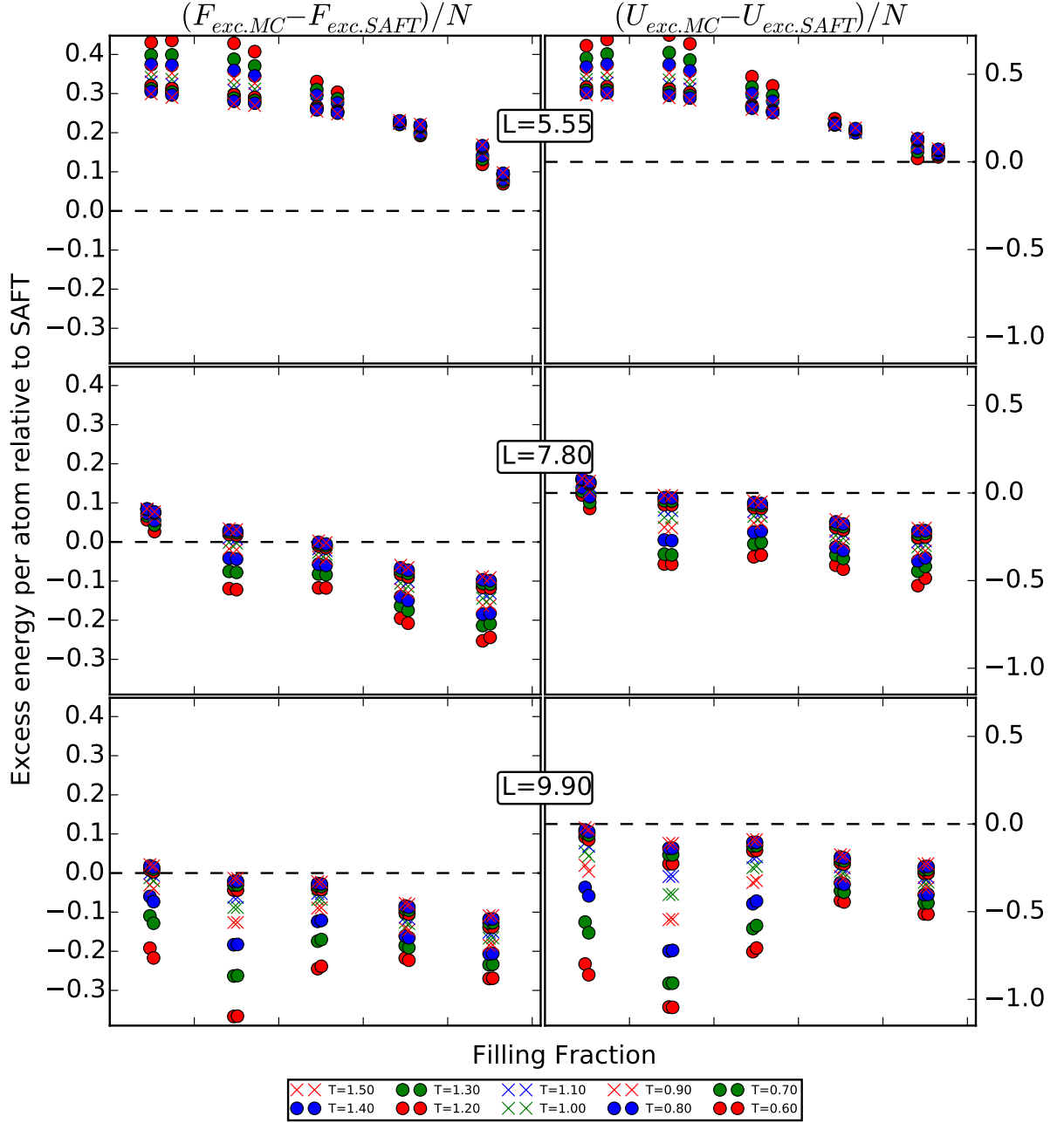


Figure 3.1: These graphs show the Monte Carlo simulation results in terms of dispersive free energy and internal energy relative to SAFT at three different box sizes. Both sets of graphs show that a small box size has a higher energy than SAFT, while a larger box will have a lower energy than SAFT. This makes sense because as the box size increases longer wavelength fluctuations are allowed, and these longer wavelengths actually lower the allowed energy due to the decreased interactions in the low density regions being more than compensated by the increase in interactions in the high density regions. Given a constant box size, the high filling fraction simulations cannot fluctuate as much as the low filling fraction simulations, which explains the bulge in the two lower plots.

## Trends in S,Cv

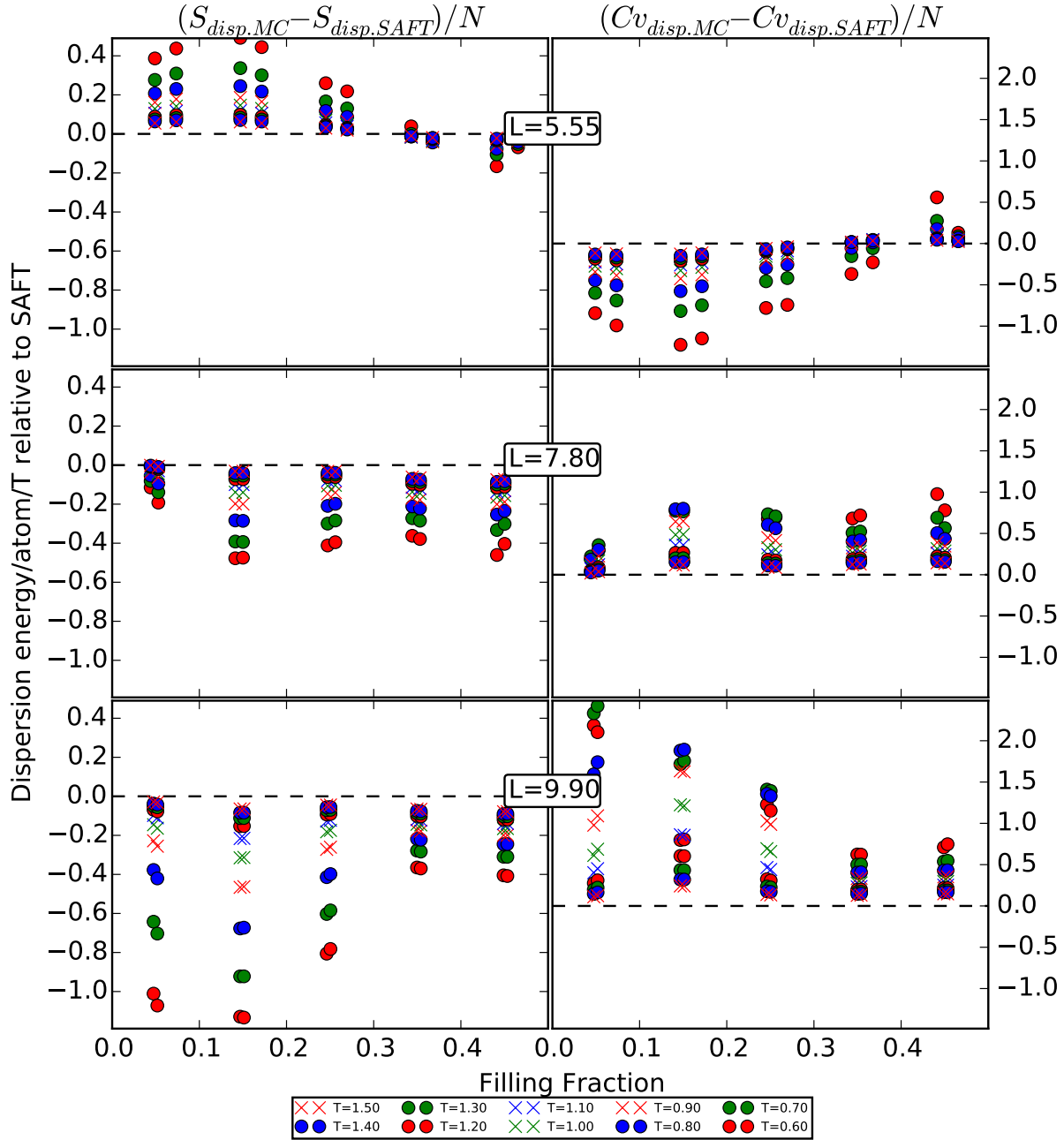


Figure 3.2: These graphs show the Monte Carlo simulations results in terms of dispersive entropy and heat capacity per particle relative to SAFT at three different box sizes. It is interesting to see the entropy trends are nearly the inverse of the heat capacity trends. Both entropy and heat capacity require a derivative to find, it is just that entropy is the negative derivative of Free energy while heat capacity is the positive derivative of internal energy. Given fig. 3.1 showed the internal and free energy spread out in a similar fashion as the temperature was increased, it isn't surprising entropy and heat capacity have inverse trends. Regarding the trend itself, the graphs suggest entropy of a small box will have more entropy than SAFT, while a large box will have a lower entropy than SAFT. This trend can be explained by the particles binding together in ways that SAFT could not take into account, and this extra binding will decrease the entropy simply because some configurations are now favored over other configurations.

### Using Cost to Estimate Best Box Size

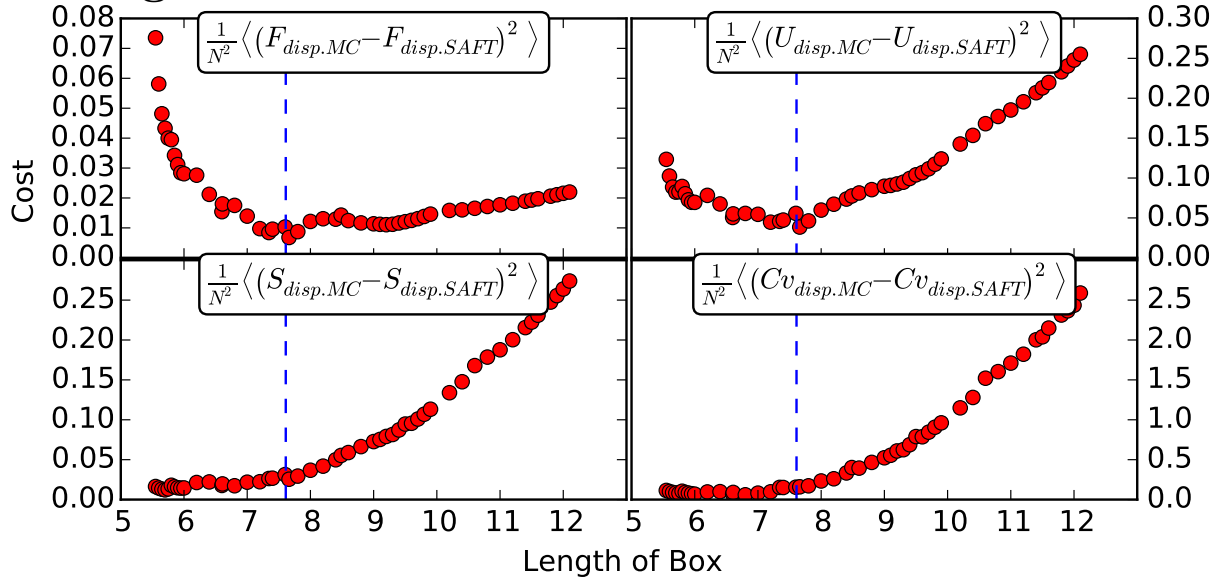


Figure 3.3: These graphs show the cost function as a function of the box size for the dispersive free energy, internal energy, entropy, and heat capacity. The cost function is defined as the mean square difference between the Monte Carlo simulations and SAFT evaluated at a filling fraction of [0.15,0.25,0.35,0.45] and evaluated for  $0.6 < T < 1.5$ . The dispersive free energy and internal energy show a minimum near  $L=7.6$ , while any box size less than  $L=8.0$  seems to be just fine for the dispersive entropy and heat capacity.

### Coexistence: T vs Density

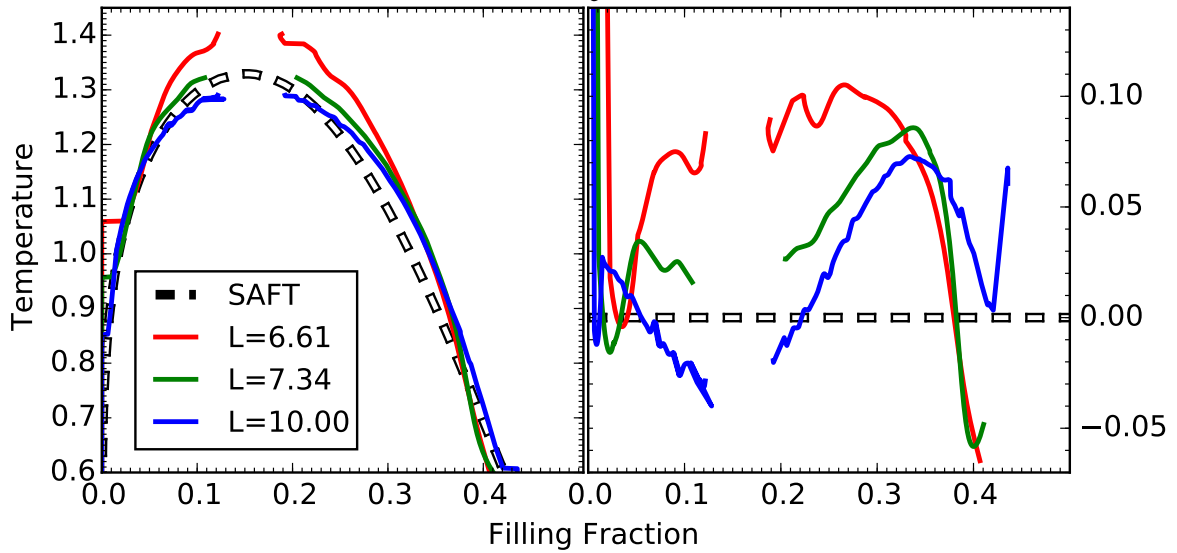


Figure 3.4: Here the gas-liquid coexistence temperature as a function of filling fraction for three different Monte Carlo simulations are compared to SAFT. The box size  $L=6.61$  clearly has a higher critical temperature than SAFT, while the box size  $L=10.00$  clearly has a lower critical temperature than SAFT.

# Conclusion

## Conclusion

SAFT-VR uses perturbation to incorporate interactions up to some unknown length scale  $L$ . GRG builds upon SAFT-VR by incorporating even longer wavelengths in an iterative manner. Unfortunately the unknown length scale is a free parameter. Free parameters are usually varied to find the best fit with experiment, but extra free parameters that aren't needed usually saps a model of its predictive abilities. In the extreme case, a bunch of experiments can be interpolated to 'predict' what will happen in the valid range of the data, yet the interpolation will fail far from the actual data. This failure to make predictions can also happen to models with too many free parameters which is undesirable for everyone involved.

This project explored how to fix the arbitrary parameter  $L$  in GRG by instead comparing the results of Monte Carlo simulations to SAFT-VR. Given that SAFT-VR only incorporates interactions up to some length scale  $L$ , the Monte Carlo simulations with a box size of length  $L$  should be able to reproduce the results of SAFT-VR. This project did vary the box size of the Monte Carlo simulations and found a 'best' fit length of roughly  $L=(7.6)$ . To test this best fit, the phase coexistence was found for three box sizes and compared to SAFT-VR. Boxes much smaller than 7.6 had a critical point higher than SAFT-VR, while box sizes much larger than 7.6 had a critical point below SAFT-VR.

While working on this project, I honestly thought the noise of the simulations would drown out the coexistence points. It was interesting to see the coexistence plots actually see a signal as low as (0.00006) (see fig 2.2). Work for the future would include running multiple simulations to increase the signal to noise ratio, auto detection of outliers, and an squeue interface that can fairly balance the workload between multiple users.