

# The contact value approximation to the pair distribution function for an inhomogeneous hard sphere fluid.

Paho Lurie-Gregg

(Dated: October 3, 2014)

We construct the contact value approximation (CVA) for the pair distribution function,  $g^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ , for an inhomogeneous hard sphere fluid. The CVA is an average of two radial distribution functions, which each take as input the distance between the particles,  $|\mathbf{r}_2 - \mathbf{r}_1|$ , and the average value of the radial distribution function at contact,  $g_\sigma(\mathbf{r})$  at the locations of each of the particles. In a recently published paper, an accurate function for  $g_\sigma(\mathbf{r})$  was developed, and it is made use of here. We then make a separable approximation to the radial distribution function,  $g_S(r)$ , which we use to construct the separable contact value approximation (CVA-S) to the pair distribution function.

We compare the CVA and CVA-S to Monte Carlo simulations that we have developed and run as well as to two prior approximations to the pair distribution function. This comparison is done in three main cases: When one particle is near a hard wall; when there is an external particle the size of a sphere of the fluid; and for various integrals that illustrate typical use-cases of the pair distribution function. We show reasonable quantitative agreement between the CVA-S and simulation data, similar to that of the prior approximations. However, due to its separable nature, the CVA-S can be efficiently used in density functional theory, which is not the case of the prior approximations.

## CONTENTS

I. Introduction	2
II. Methods	3
A. Monte Carlo simulation	3
B. A separable fit to the radial distribution function for a hard sphere fluid	4
C. Construction of the CVA	6
III. Results	7
A. Pair distribution function	7
B. Triplet distribution function	9
C. Thermodynamic perturbation theory	10
IV. Discussion	12
V. Conclusion	13
References	13

## I. INTRODUCTION

In this project, we study hard sphere fluids at an interface with a hard wall. Hard sphere fluids are used as a reference fluid for virtually every current liquid theory. Hard spheres are defined as having no interactions except that they cannot overlap. One can imagine a hard sphere fluid as billiard balls bouncing around in a box in space. Various properties of this fluid are used in the foundation of the theory for real liquids. One such property is the pair distribution function, which we have set out to study. The pair distribution function takes the positions of two spheres, and gives the probability of finding a sphere at one position given that there is one at the other position, compared to the probability of finding it there otherwise. It is, in effect, a measure of how much a particle affects the distribution around it. The pair distribution function has been studied extensively in the homogeneous case. One can think of this case as the fluid extending infinitely in all directions. Then, the location of the first sphere does not matter, and the pair distribution function is a function purely of the distance between the particles, and is called the radial distribution function.

However, the pair distribution function is much less well understood when there is an interface, and many interesting things happen at interfaces. For example, a molecule in a fluid gives a small interface, acid-base interactions happen at the interface between the fluids, rust occurs at the interface between water and metal, and so on. The pair distribution function at an interface is very useful in many such calculations, but there are currently no approximations to it that can be used efficiently enough to be practical. The approximations that do exist tend to be made without a test of their accuracy[1–3]. In this paper, we develop a new approximation for the pair distribution function for a hard sphere fluid and compare it to the existing approximations as well as to simulation results, which gives a clear idea of its accuracy.

Simulation is carried out via the Monte Carlo method. The idea is to generate as many random states of the fluid as possible, and to average all of them to get a true picture of the behavior of the hard sphere fluid. Monte Carlo simulations give essentially exact results, with small and knowable error bars, and so provide an ideal reference for our theoretical results.

## II. METHODS

### A. Monte Carlo simulation

In order to have a reference for the CVA and CVA-S, we created and then ran Monte Carlo simulations. The idea behind a Monte Carlo simulation is to generate many states and average all of them to get a picture of the behavior of the hard sphere fluid. The simplest way to run such a simulation would be to construct a box and then place spheres at random locations within it. If any of the spheres overlap, then it is not a valid hard sphere fluid and must be discarded. If none of the spheres overlap, then it is a valid state for a hard sphere fluid, and this data is kept. Then, one would repeat this until there are enough randomly generated valid states to be able to average them and get an accurate representation of the fluid.

In practice, this is impossible. Using such a method, there would be virtually zero states that do not have at least one overlap for any reasonable density, and so no data would ever be gathered. Instead, our simulations start by placing the spheres randomly in the box, and then moving them random but short distances, only keeping moves that reduce the total degree of overlap, until none of the spheres overlap. This gives a valid state to start with. Then, the simulation attempts to move one sphere a short distance in a random direction. If, after moving, it overlaps with another sphere, then it is an invalid move and the sphere is moved back. This procedure is then repeated

for each sphere in the fluid. Once each sphere has had an attempted move, the state of the fluid is recorded and the process is repeated over many iterations.

The state of the fluid is recorded by dividing space into regular bins and counting spheres observed in each bin. How exactly this is done varies based on the data being recorded. To get density, for example, it is sufficient to count the number of spheres seen at each location. To then obtain the density of a given bin, one can use the relation

$$n = \frac{N_{\text{bin}}}{N_{\text{total}}} \frac{N}{V_{\text{bin}}} \quad (1)$$

where  $n$  is the number density,  $N_{\text{bin}}$  is the number of times a sphere has been counted in the given bin,  $N_{\text{total}}$  is the number of times spheres have been counted across all of the bins,  $N$  is the number of spheres in the simulation, and  $V_{\text{bin}}$  is the volume of the bin in question. Note that, of these quantities,  $N_{\text{bin}}$  and  $N_{\text{total}}$  increase with simulation time, whereas  $N$  and  $V_{\text{bin}}$  are fixed. For a hard sphere fluid near a hard wall, the density is particularly nice because the system has cylindrical symmetry, so the density really is just a function of one variable (the distance from the wall), and each bin is a slice of the full cell parallel to the wall.

To get the pair distribution function is similar, but slightly more complicated as creating bins and counting the spheres observed in them requires looking at every pair of spheres in relation to each other. Due to the cylindrical symmetry of the system, we can think of the pair distribution function as a function of 3 scalar variables. For the first particle, only the distance from the wall is important. For the second particle, its distance from the wall and radial distance from the first particle are needed. In cylindrical coordinates, these are  $z_1$ ,  $z_2$ , and  $r_2$ . So, the pair distribution function can be thought of as a function of these three variables, which causes the binning to be a bit more involved than for the density. The pair distribution function for a given bin is defined as

$$g^{(2)} = \frac{N_{\text{bin}}}{N_{\text{total}}} \frac{1}{n_1 n_2 V_{\text{shell1}} V_2} \quad (2)$$

where  $N_{\text{bin}}$  and  $N_{\text{total}}$  are as before,  $n_1$  is the density for the bin that the first sphere is in,  $n_2$  is the density for the bin that the second sphere is in,  $V_{\text{shell1}}$  is the volume of the *shell* that the first particle is in (this is the bin size that was used for the density), and  $V_2$  is the volume of the *bin* that the second particle is in.

## B. A separable fit to the radial distribution function for a hard sphere fluid

The radial distribution function is equal to the pair distribution function for a homogeneous hard sphere fluid. For a given fluid, it is a function of how fixing particle at one position affects the



probability to find a particle at a second position. As it is a homogeneous fluid, the only thing that matters in terms of these two positions is the distance between them, and so the radial distribution function is often expressed as a function of one variable,  $g(r)$ . However, it really also depends on the density of the fluid. Rather than a straight density, we use the pair distribution function averaged over positions  $\mathbf{r}_2$  that are in contact with  $\mathbf{r}_1$ , called  $g_\sigma$  because  $\sigma$  is the diameter of a hard sphere and so  $g_\sigma(\mathbf{r}_1)$  is the average value of  $g^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$  when  $|\mathbf{r}_2 - \mathbf{r}_1| = \sigma$ . Instead of  $g_\sigma(\mathbf{r})$ , other works[2, 3] use the density of the fluid averaged over some volume. We use  $g_\sigma(\mathbf{r})$  because it holds more information about an inhomogeneous fluid and because a very good approximation for  $g_\sigma(\mathbf{r})$  was found in a previous paper[4]. For a homogeneous fluid,  $g_\sigma$  is constant in space, and so we can express the radial distribution function as  $g(r, g_\sigma)$ .

For computational efficiency in future use in density function theory, we desire a separable form of the contact value approximation (CVA-S), which means that we need a separable form of the radial distribution function. While there is no known analytic form for the radial distribution function, there do exist some good approximations, but none are separable. We start by defining our separable approximation to  $g$ ,  $g_S$  in the general form for a separable function

$$g_S(r, g_\sigma) = \sum_i a_i(r) b_i(g_\sigma) \quad (3)$$

Then, we apply several constraints. First, we want to match the value at contact

$$g_S(\sigma, g_\sigma) = g_\sigma \quad (4)$$

Second, we found a very good approximation to the slope at contact. We were unable to find a theoretical justification for this approximation, but found that it is a good approximation by comparison with simulation results. We match it as follows:

$$g'_S(\sigma, g_\sigma) = g_\sigma - g_\sigma^2 \quad (5)$$

A final constraint fixes the integral over all space,

$$1 + n \int_{\mathbb{R}^3} (g(r) - 1) d\mathbf{r} = n\chi \quad (6)$$

where  $n$  is the number density, and  $\chi$  is the dimensionless form of the isothermal compressibility of a hard sphere fluid, which is given in Equation 9. By using this constraint, we lose a lot of the oscillatory behavior of the radial distribution function, but deemed that an acceptable concession for a general purpose approximation.

The simplest function that we found that fits these constraints that is reasonably accurate is

$$g_S(r, g_\sigma) = 1 + (g_\sigma - 1)e^{-c_0(\frac{r}{\sigma}-1)} + (g_\sigma - 1)(c_0 - 2g_\sigma) \left(\frac{r}{\sigma} - 1\right) e^{-c_1(\frac{r}{\sigma}-1)} + I(g_\sigma) \left(\frac{r}{\sigma} - 1\right)^2 e^{-c_2(\frac{r}{\sigma}-1)} \quad (7)$$

where

$$I(g_\sigma) = c_2^5 \frac{\frac{\chi-1}{24\eta} + (1-g_\sigma) \frac{(c_0-2g_\sigma)(c_1^2+4c_1+6)}{c_1^4} + (1-g_\sigma) \frac{c_0^2+2c_0+2}{c_0}}{2c_2^2 + 12c_2 + 24} \quad (8)$$

and

$$\chi = \frac{(1-\eta)^4}{\eta^4 - 4\eta^3 + 4\eta^2 + 4\eta + 1} \quad (9)$$

Finally, we can convert between the filling fraction  $\eta$  (which is the fraction of space that is filled by hard spheres) and the radial distribution function at contact  $g_\sigma$  using the Carnahan-Starling approximation

$$g_\sigma = \frac{1 - \frac{\eta}{2}}{(1 - \eta)^3} \quad (10)$$

and its inverse which we calculated using Mathematica

$$\eta = 1 - \frac{1}{\sqrt[3]{54g_\sigma^2 - 6\sqrt{81g_\sigma^4 - 6g_\sigma^3}}} - \frac{\sqrt[3]{54g_\sigma^2 - 6\sqrt{81g_\sigma^4 - 6g_\sigma^3}}}{6g_\sigma}. \quad (11)$$

The remaining three parameters,  $c_0$ ,  $c_1$ , and  $c_2$  were found by using a least squares fit to Monte Carlo simulation data. The values used are  $c_0 = 3.68$ ,  $c_1 = 2.16$ ,  $c_2 = 2.79$ . The final fit can be seen in Figure 1. While  $g_S$  under-approximates both the number and amplitude of oscillations, the fit was done such that it gets the value at and near contact correct as well as the integral over all space, making it a viable approximation for near and far interactions.

### C. Construction of the CVA

The goals for our approximation to the pair distribution function for a hard sphere fluid are as follows. First, it should be based on the radial distribution function as we have good insight into it and it contains similar information. All prior approximations that we have found also make use of the radial distribution function. Second, given a separable radial distribution function, the pair distribution function should also be separable. A geometric mean, for example, would not meet this goal. Third, it should be as simple as possible.

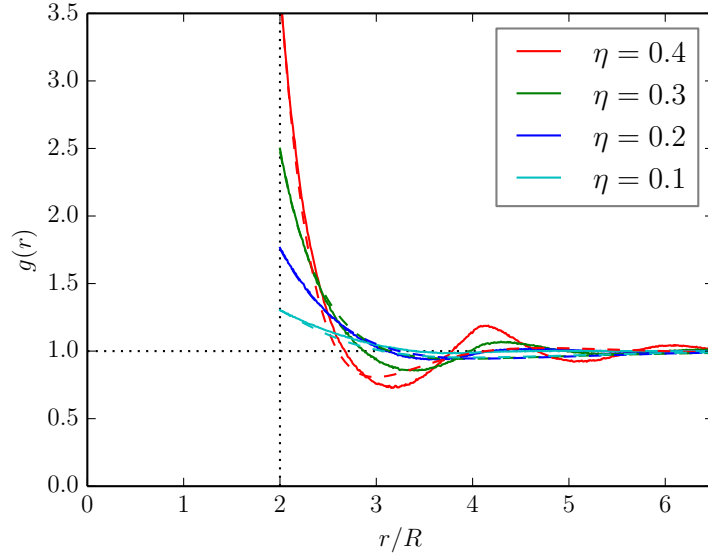


FIG. 1: Plot of the hard-sphere radial distribution function, with our separable fit. The solid lines show the true radial distribution function  $g(r)$ , and the dashed lines show our separable fit to the radial distribution functions  $g_S(r)$ .

With these goals in mind, we define the CVA as

$$g^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \frac{g(r_{12}, g_\sigma(\mathbf{r}_1)) + g(r_{12}, g_\sigma(\mathbf{r}_2))}{2} \quad (12)$$

which is just the averaged value of the radial distribution function at the two points, using in both cases the distance between the points.

### III. RESULTS

#### A. Pair distribution function

We first examine the pair distribution function when one sphere is in contact with a hard wall. In Figures 2a and 2c, this data is shown for fluids with filling fractions of  $\eta = 0.1$  and  $\eta = 0.3$ , respectively. The top halves of these figures are Monte Carlo simulation results and the bottom halves are the results from the CVA-S. The gray circle indicates the location of the sphere in contact with the wall, with the black semi-circle surrounding it indicating a volume where no other spheres can be (as any spheres in this region would collide with the fixed sphere), giving the pair distribution function a value of 0. The color scale is such that the plot is white when the pair distribution function has a value of 1, indicating that the gray sphere has no affect on the

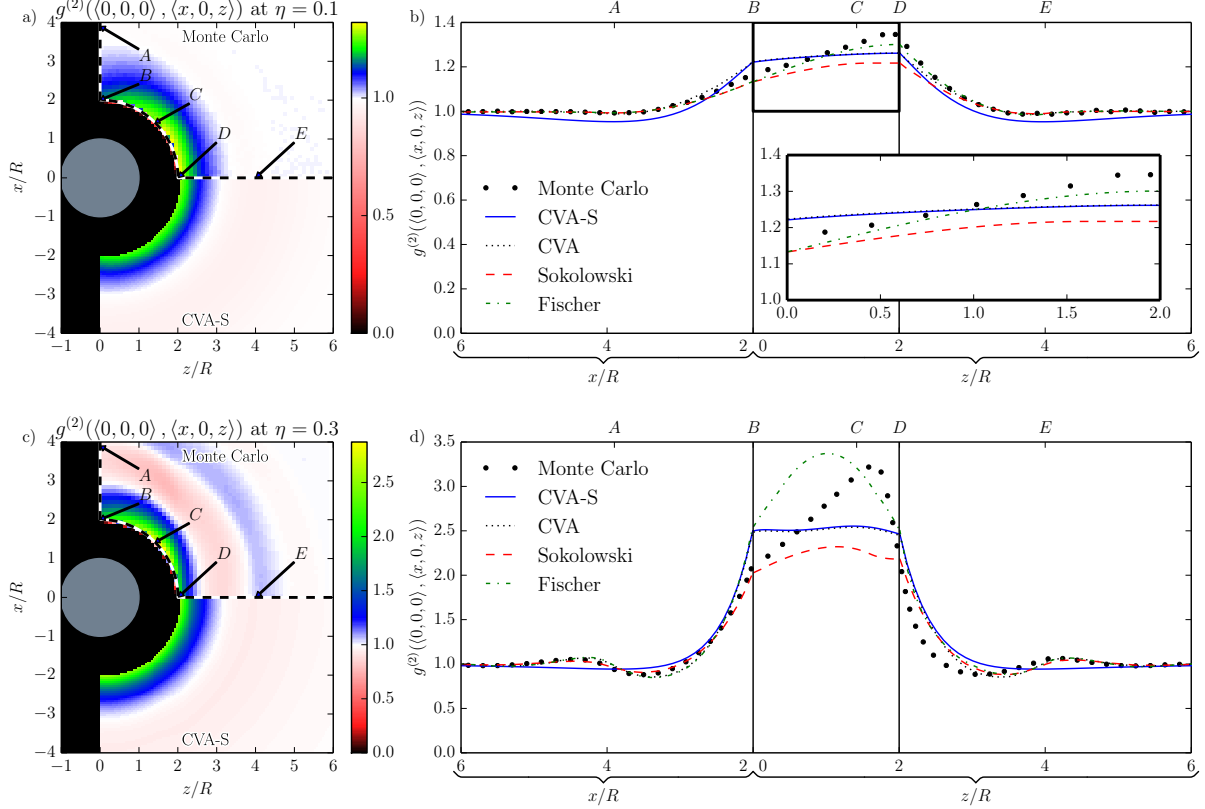


FIG. 2: The pair distribution function near a hard wall, with packing fractions of 0.1 and 0.3 and  $\mathbf{r}_1$  in contact with the hard wall. On the left are 2D plots of  $g^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$  as  $\mathbf{r}_2$  varies. The top halves of these figures show the results of Monte Carlo simulations, while the bottom halves show the CVA-S. On the right are plots of  $g^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$  on the paths illustrated in the figures to the left. These plots compare the CVA-S (blue solid line) and CVA (cyan dotted line) with Monte Carlo results (black circle) and the results of Sokolowski and Fischer (red dashed line) [2], and those of Fischer and Methfessel (green dot-dashed line) [3]. The latter is only plotted at contact, where it is defined.

likelihood of finding a sphere at that location. In both of these figures, there is a dashed line that comes down along the wall, around the gray sphere, and to the right, perpendicular to the wall. In figures 2b and 2d, the value of the pair distribution is plotted along the path indicated by the dashed line. Both the CVA and CVA-S are plotted in these figures, as well as the Monte Carlo results and the results of two other methods for comparison; those of Sokolowski and Fischer[2] and of Fischer and Methfessel[3]. The method of Fischer and Methfessel is only defined when the two spheres are in contact, so their results are only plotted over the central region.

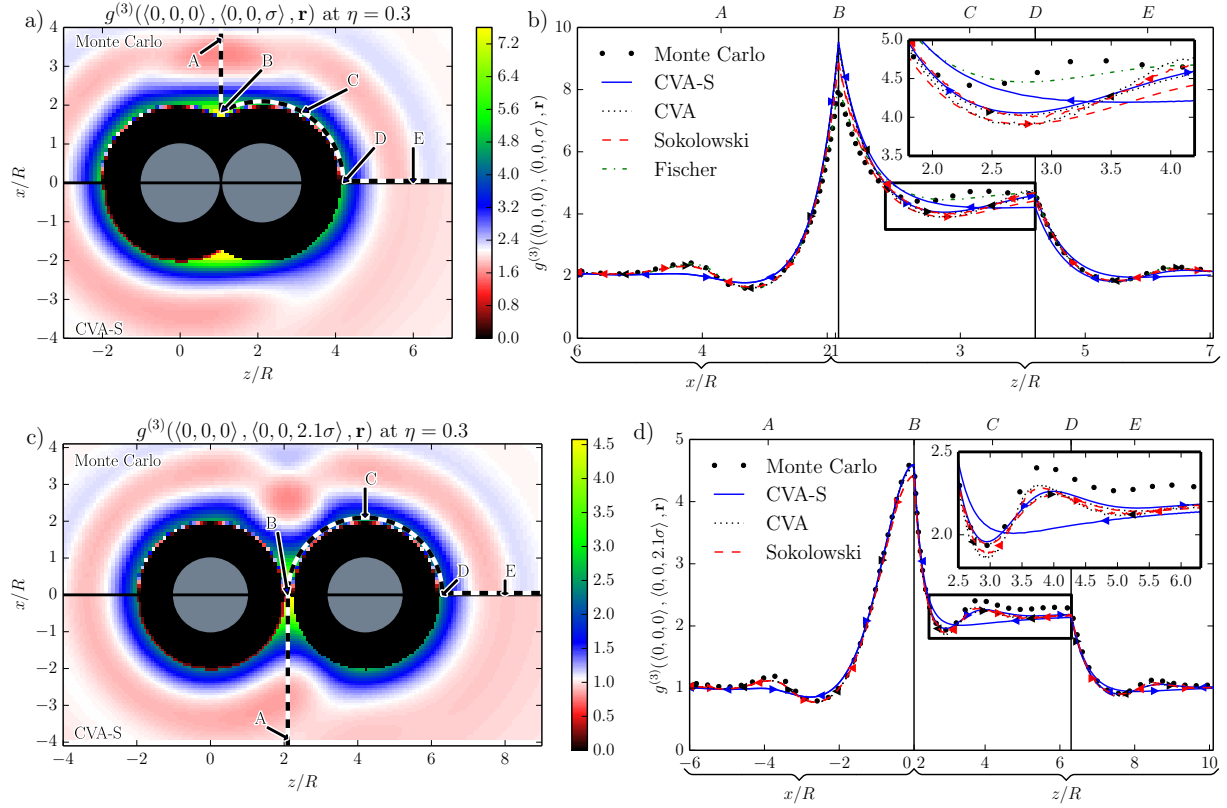


FIG. 3: The triplet distribution function  $g^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$  at packing fraction 0.3, plotted when  $\mathbf{r}_1$  and  $\mathbf{r}_2$  are in contact (a,b) and when  $\mathbf{r}_1$  and  $\mathbf{r}_2$  are separated by a distance  $2.1\sigma$  (c,d). On the left are 2D plots of  $g^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$  as  $\mathbf{r}_3$  varies. The top halves of these figures show the results of Monte Carlo simulations, while the bottom halves show the CVA-S. On the right are plots of  $g^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$  on the paths illustrated in the figures to the left. We also plot these curves along a left-right mirror image of this path. The data for the right-hand paths (as shown in the 2D images) are marked with right-pointing triangles, while the left-hand paths are marked with left-pointing triangles.

### B. Triplet distribution function

The triplet distribution can be thought of as the relative probability of finding a sphere in a position if we fix *two* spheres, instead of just one. If, instead of a flat wall, one were to make a small spherical wall, the exact size of a sphere in the fluid, and place it in an otherwise homogeneous hard sphere fluid, then the pair distribution function near this spherical wall would be identical to the triplet distribution function of a homogeneous fluid. We plot this function in figures 3a and 3c for a fluid with filling fraction  $\eta = 0.3$ . In figure 3a, the two fixed spheres are in contact,

and in figure 3c, they are just far enough apart to fit a third sphere between them. Again, the top half of these figures contains the results of Monte Carlo simulations, where we computed the triplet distribution function of a homogeneous hard sphere fluid. The bottom half contains CVA-S results, with the gray circle on the left indicating the external “wall” particle, and the gray circle on the right indicating the fixed particle, similar to figures 2a and 2c. While the pair distribution function has left-right symmetry, it can be seen in these figures that the CVA-S does not. This is a result of inaccuracies in the radial distribution function used, because the two spheres are treated computationally differently. Again, there are dashed lines indicating a path along which the triplet distribution function is plotted in figures 3b and 3d. Because of the asymmetry in the theoretical triplet distribution results, the CVA, CVA-S, and the method of Sokolowski and Fischer are each plotted twice, once along the path as shown, indicated with right-facing triangles, and again along a mirror image path, indicated with left-facing triangles.

### C. Thermodynamic perturbation theory

All approximations were tested for how well they predict the first term  $F_1$  of the interaction energy due to a pair potential. In order to observe how well each approximation does as a function of distance, the derivative of  $F_1$  with respect to  $z$ ,

$$\frac{dF_1}{dz} = \frac{1}{2} \int g^{(2)}(\mathbf{r}, \mathbf{r}') n(\mathbf{r}) n(\mathbf{r}') \Phi(|\mathbf{r} - \mathbf{r}'|) d\mathbf{r}' dx dy \quad (13)$$

is what we have computed. In this equation,  $n(\mathbf{r})$  is the number density of the hard sphere fluid at  $\mathbf{r}$  and  $\Phi(r)$  is the pair potential for spheres separated by a distance  $r$ . Three different pair potentials were studied.

First, the potential for a fluid of sticky hard-spheres was used. In this case, the potential is given by

$$\Phi(r) = \delta(\sigma - r) \quad (14)$$

The results are plotted in Figure 4a. In this case, the CVA and CVA-S do better than the prior approximations at contact. This is due to the use of  $g_\sigma$  in the CVA and CVA-S; the form of this integral at contact is essentially the same as that for computing  $g_\sigma$  and we have used a very accurate method for computing  $g_\sigma$ [4].

Next, a potential for a square well fluid was used. In this case, the potential is given by

$$\Phi(r) = \Theta(1.79\sigma - r) \quad (15)$$

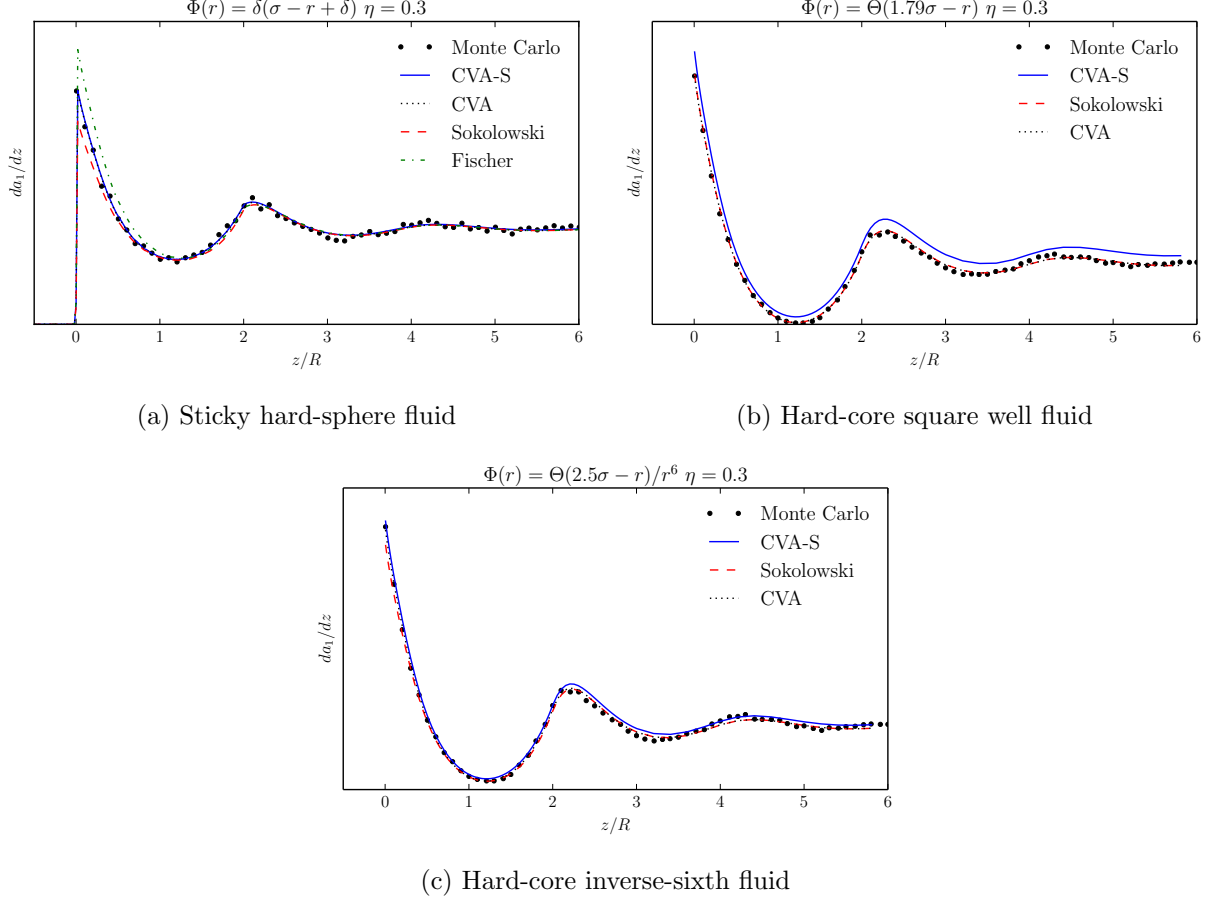


FIG. 4: Plot of  $\frac{dF_1}{dz}$  near a hard wall. (a) shows a sticky hard-sphere fluid defined by a pair potential  $\delta(\sigma - r + \delta)$  where  $\sigma$  is the hard-sphere diameter, and  $\delta$  is an infinitesimal distance; (b) shows a square well fluid defined by a pair potential  $\Theta(1.79\sigma - r)$ ; and (c) shows a hard-core inverse-sixth potential fluid with an attractive pair potential proportional to  $r^{-6}$ .

The results of the square well fluid are plotted in Figure 4b. In this case, the CVA and the two prior approximations all perform very well. The CVA-S has a systematic error. This is due to the square well width  $1.79\sigma$  occurring at a point where  $g_S(r, g\sigma)$  does not accurately reflect the radial distribution function. For other square well widths, the CVA-S may be more accurate or have systematic error in the other direction.

Lastly, the potential of a fluid comprised of hard spheres with an inverse sixth attractive potential. In this case, the potential is given by

$$\Phi(r) = \Theta(2.5\sigma - r)r^{-6} \quad (16)$$

In this case, all approximations perform very similarly.

#### IV. DISCUSSION

For the pair distribution function, all of the approximations perform their worst for the angular dependence at contact. This is a fundamental flaw in doing this kind of approximation; the angular information is largely what has been averaged out. By design, the CVA and CVA-S perform identically over this region. Due to its use of the averaged value of the pair distribution function at contact,  $g_\sigma$ , the CVA will obtain the correct result when an integral over the region at contact is taken. While the two prior approximations show a stronger angular dependence than the CVA, they each contain a degree of systematic error, either too high or too low, and will not perform as well for such an integral, as seen in Figure 4a. Away from contact, the CVA performs similarly to the prior approximations, while the CVA-S lacks some of the oscillatory behavior; a result of  $g_S$  doing the same.

For the triplet distribution function, all of the approximations perform very similarly, with a few exceptions. When all three spheres are in contact, the method of Fischer and Methfessel captures the angular dependence better (Figure 3b). When the two initial spheres are placed apart just enough for the third sphere to fit between, the CVA-S does not reflect the symmetry of the system accurately, with its values along the “right-traveling” path similar to those of the other approximations, but those along the “left-traveling” path being a good deal less accurate. The CVA and two prior approximations do very well away from contact, with the CVA-S again somewhat undervaluing oscillations.

The inaccuracies in the CVA-S that are not present in the CVA are due exclusively to the approximation that we have used for the radial distribution function. This is by no means the only approximation appropriate to use. So long as it is separable in terms of  $r$  and  $g_\sigma$ , any approximation for  $g(r, g_\sigma)$  can be used and still allow for FFTs to be used when evaluating integrals. We have also developed a short range approximation, defined only for  $r \leq 4R$  where  $R$  is the hard sphere radius. This approximation is quite accurate over the range it is defined; for  $\eta \leq 0.3$  it is essentially exact. The approximation presented in this paper serves as a decent general purpose approximation, and is used largely as an example in order to get results for the CVA-S. For the interactions presented in Section IIIC, the short range version of  $g_S$  would be appropriate for both the sticky hard sphere fluid and the square well fluid, but not for the inverse-sixth potential as that has a small dependence long range interactions. For the final version of the publication of this work, the short range approximation was used in lieu of the  $g_S$  presented here[5].



## V. CONCLUSION

The contact value approximation to the pair distribution function for a hard sphere fluid is presented in this paper, and then tested against two prior approximations and to simulation results. In both separable and non-separable forms, it has been tested both near a hard wall and near a hard spherical wall the size of a particle in the hard sphere fluid, as well as in some integrals typical in thermodynamic perturbation theory. It has been found to have generally comparable accuracy to the prior approximations, performing better in some cases, while the separable form is lacking in other cases. The separable form presented here is a general case, and could be made more accurate for specific uses. The separable form has the distinction that integrals involving it will be able to be computed using fast Fourier transforms, allowing them to be computed efficiently, and so it is suitable for use in density functional theory.

- 
- [1] G. J. Gloor, G. Jackson, F. Blas, E. M. Del Rio, and E. De Miguel, *The Journal of Physical Chemistry C* **111**, 15513 (2007).
  - [2] S. Sokolowski and J. Fischer, *The Journal of chemical physics* **96**, 5441 (1992).
  - [3] J. Fischer and M. Methfessel, *Physical Review A* **22**, 2836 (1980).
  - [4] J. B. Schulte, P. A. Kreitzberg, C. V. Haglund, and D. Roundy, *Physical Review E* **86**, 061201 (2012).
  - [5] P. Lurie-Gregg, J. B. Schulte, and D. Roundy, *arXiv preprint arXiv:1403.6893* (2014).