# Nested sampling for materials: the case of hard spheres

Lívia B. Pártay

*University Chemical Laboratory, University of Cambridge,*

*Lensfield Road, CB2 1EW Cambridge, United Kingdom*[*]

Albert P. Bartók and Gábor Csányi

*Engineering Laboratory, University of Cambridge,*

*Trumpington Street, CB2 1PZ Cambridge, United Kingdom*

(Dated: November 12, 2014)

## Abstract

The recently introduced nested sampling algorithm allows the direct and efficient calculation of the partition function of atomistic systems. We demonstrate its applicability to condensed phase systems with periodic boundary conditions by studying the three dimensional hard sphere model. Having obtained the partition function, we show how easy it is to calculate the compressibility and the free energy as functions of the packing fraction and local order, verifying that the transition to crystallinity has a very small barrier, and that the entropic contribution of jammed states to the free energy is negligible for packing fractions above the phase transition. We quantify the previously proposed schematic phase diagram and estimate the extent of the region of jammed states. We find that within our samples, the maximally random jammed configuration is surprisingly disordered.

## I. INTRODUCTION

Nested sampling (NS) was introduced by Skilling[1,2] about a decade ago in the field of applied probability and inference to efficiently sample probability densities in high dimensional spaces where the regions contributing most of the probability-mass are exponentially localised. It is an annealing algorithm that creates a sequence of probability distributions, each more concentrated than the previous one near the high-likelihood region of the parameter space (or, in the context of materials, the low energy region of configuration space) but with the crucial property that the ratio of phase space volumes that support the successive distributions is known, and is approximately constant during the annealing process. This enables the direct evaluation of the density of states, the partition function, and all other thermodynamic observables.

Since its original inception, nested sampling has been taken up in astrophysics[3] where there is a long history of association with the inference community, and it has also been applied to atomistic systems. It was shown to gain orders of magnitude in efficiency over parallel tempering (also known as replica exchange sampling) for exploring the configurations and calculating heat capacity curves of Lennard-Jones clusters and permitted the construction of phase stability diagrams.[4] Burkoff *et al.* used it to explore the folding landscape of proteins,[5] and Do *et al.* studied liquid vapour coexistence,[6] binary liquid mixtures[7] and combined it with their "cage model" for solids.[8] Recently Nielsen carried out nested sampling using a sequence of NVT trajectories.[9]

We start with a brief account of the nested sampling algorithm, and to aid understanding in the materials context we loosely follow the exposition of Wheatley rather than the original formulation of Skilling. Consider the partition function $Z(V, \beta)$ of a system of $N$ particles at constant volume $V$ and inverse temperature $\beta$,

$$Z(V, \beta) = \frac{1}{N!} \left( \frac{2\pi m}{\beta h^2} \right)^{3N/2} \int d\mathbf{p} \, d\mathbf{q} \, e^{-\beta H(\mathbf{p}, \mathbf{q}, V)}, \tag{1}$$

where $\mathbf{q}$ and $\mathbf{p}$ are the configurational and the corresponding momentum variables, respectively, and $H$ is the Hamiltonian. Separating out the momentum-dependent part of the Hamiltonian in the usual way, we focus on the configurational partition function and using the fact that the dependence of the Hamiltonian on the coordinates $\mathbf{q}$ is solely through the

potential energy $U$, we can write the partition function in terms of the density of states, $\Omega$,

$$Z(V, \beta) = Z_\beta \int d\mathbf{q} \, e^{-\beta U(\mathbf{q}, V)} = Z_\beta \int du \, e^{-\beta u} \, \Omega_V(u), \tag{2}$$

$$Z_\beta = \frac{1}{N!} \left( \frac{2\pi m}{\beta h^2} \right)^{3N/2}, \tag{3}$$

$$\Omega_V(u) = \int d\mathbf{q} \, \delta(U(\mathbf{q}, V) - u), \tag{4}$$

where we have moved $V$ into the subscript of $\Omega$ to emphasise that it is a parameter rather than a variable in the present context. In order to calculate $\Omega$, let us consider its cumulative function, defined by

$$\gamma_V(u) = \int_{-\infty}^{u} du' \, \Omega_V(u'). \tag{5}$$

Since the potential energy function for a given material is obviously bounded from below by the binding energy of the most stable phase, there is no problem with taking the lower limit of the integral to be negative infinity. Computing $\gamma(u)$ directly for any particular $u$ is very difficult. However, since it is the cumulative of a distribution, it is monotonic in $u$, and therefore we are free to consider its inverse, $u(\gamma)$. Noting that all thermodynamic observables of interest depend only on the logarithm of $Z$ and therefore we are only interested in $\gamma$ and $\Omega$ up to a constant multiplicative factor, nested sampling proceeds to compute $u(\gamma)$ as follows.

1. Select a very high potential energy value $U = u_0$, such that the structure of the corresponding subset of configuration space is little different from that of the ideal gas. Fix the corresponding value of $\gamma_0$ so that $u(\gamma_0) = u_0$, thus absorbing the arbitrary multiplicative constant in $\gamma$. Set $i = 0$.

2. Obtain $K$ samples in configuration space from the *uniform* distribution over $\mathbf{q}$, subject to the constraint $U(\mathbf{q}) < u_i$. This is easily accomplished by Markov chain Monte Carlo and rejection sampling.

3. Compute the median of the potential energies of the $K$ samples, and let us call it $u_{i+1}$. The volume of the phase space covered by the samples with potential energy $U(\mathbf{q}) < u_{i+1}$ is given by the random variable $\gamma_{i+1}$ whose distribution is approximately $\texttt{Beta}(K/2, K/2)\gamma_i$, with a mean of $\gamma_i/2$ whose statistical error that scales with $K^{-1/2}$. Thus we obtain the estimate $u(\langle \gamma_{i+1} \rangle) = u_{i+1}$.

4. Let $i \leftarrow i + 1$ and go back to step 2.

Having obtained $u(\gamma)$ at a geometrically decreasing series of $\gamma$ values, we can invert it to get $\gamma(u)$ and obtain $\Omega_V(u)$ by finite difference, which can then be used to evaluate the partition function according to equation (2). The procedure hinges critically on our ability to obtain uniformly distributed samples in step 2 above. This can be aided by recycling the positions of the $K$ samplers from the previous iteration, so that only a modest number of MCMC steps are required to re-equilibrate them according to the new constraint. As was touched upon in previous works,[4,9] the number of samplers required for a given accuracy can be strongly affected by the basin-structure of the potential energy surface, because the simple MCMC rejection sampler can have difficulty in converging to the uniform distribution once the constraint results in the distribution having disconnected domains. As usual in any Monte Carlo procedure, this type of problem can be alleviated by choosing proposal distributions that are better suited to a given system. Furthermore, the risk that a disconnected domain is missed completely by the samplers is inversely proportional to the size of the domain (and also to the number of samplers of course), and therefore large domains are less likely to be missed. This tradeoff between the number of samples generated and the risk of missing relevant configurations exists in all stochastic simulation methods, and it is no better or worse in nested sampling.

It is helpful to compare the nested sampling with simulated annealing, which also creates a series of overlapping distributions, each one proportional to the Boltzmann measure with ever decreasing temperature parameters. The major difference is that the average phase space volume ratio of the successive distributions, which is needed to calculate the partition function, is directly available (indeed set by the user) in nested sampling, whereas it would need to be computed in simulated annealing, and if the successive distributions have little overlap, the estimate of this ratio would have an enormous error. This problem would be particularly severe near a first order phase transition, where a small change in the temperature corresponds to a very large change in the distribution, and therefore the overlap between the successive distributions across the phase transition vanishes in the thermodynamic limit. For this reason, the partition function is not considered accessible by simulated annealing. In contrast, nested sampling will maintain the distribution overlap by design, even through a phase transition.

Practitioners (and Skilling originally) often replace the median in step 3 with the $K/(K+1)$ fraction of the potential energy distribution of the $K$ samples, which does not change the

algorithm substantially and results in a finer sampling of the $u(\gamma)$ function. In this case, the next energy level $u_{i+1}$ is simply given by the potential energy of the sample with the highest value in the sample set, and the corresponding phase space volume is $\gamma_{i+1} = \text{Beta}(K, 1)\gamma_i$ (note that $\text{Beta}(K, 1)$ is equivalent to $\text{Uniform}(0, 1)^{1/K}$), thus $\langle \gamma_{i+1} \rangle = \langle \gamma_i \rangle K/(K + 1)$.

The errors of any quantity derived from the partition function are most easily computed by generating a large dataset using "bootstrapping", i.e. resampling the phase space volume ratios, $\gamma$, according to their theoretical distribution.[47] The mean and standard error of the thermodynamic quantities are then given by the distribution of the sample mean.

The nested sampling iterations can in principle be continued indefinitely, obtaining ever lower energy samples, however due to the bounded nature of the potential energy for atomistic systems, configurations whose energy is very close to the lower bound will only become relevant for computing thermodynamic observables at very low temperatures. Therefore a natural and very practical criterion for terminating the iterations is to aim for the convergence of desired observables for all temperature values above a preset value.

The purpose of the present work is to demonstrate the nested sampling technique for condensed phase systems. This is different from previous applications because the ordered/disordered phase transition is much sharper than it was in the case of isolated clusters, and therefore is particularly challenging to the annealing algorithm, while still having a macroscopic number of local minima in the configuration space. The hard sphere system presents these essential features of a condensed phase material. Much is known about the hard sphere system in the literature, and indeed this contributed to our choice to use it as an exemplar. The significance of present work lies in the *power and simplicity* of nested sampling algorithm, in that it yields the partition function (and thus all desirable thermodynamics observables) in a straightforward manner, thus unifying the methodology for calculating various quantities. This makes it eminently suitable for high throughput calculations in the future.

## A.  The hard sphere system

One of the first material systems ever studied using computer simulation was the hard sphere model and its simpler two-dimensional version consisting of hard disks.[10–12] Since then hard spheres have been widely used as model particles not just in the field of statistical
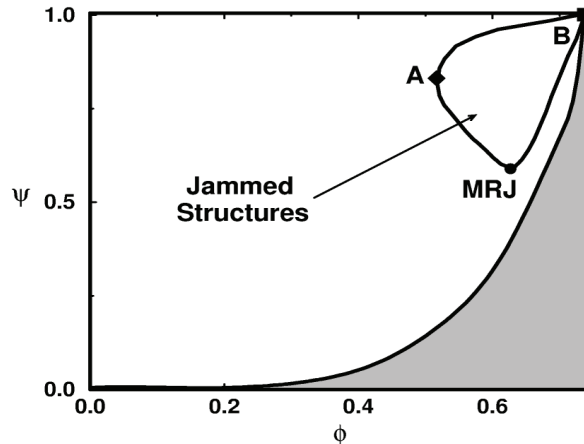
FIG. 1. Schematic plot of an arbitrary order parameter, $\Psi$, of hard sphere configurations as a function of the packing fraction, $\phi$. The grey area represents the inaccessible region, and the black squares show the maximally random jammed state (MRJ), the jammed state with the lowest possible packing fraction (A) and the perfect crystalline structure (B). (Reprint of Fig 1. of Ref. [39], with the permission of the authors. Copyright 2000 by The American Physical Society.)

mechanics,[13] but also in biology,[14,15] engineering and material science,[16,17] and consequently extensively studied by analytical means,[18] computer simulations,[19–29] and also experimental methods using colloids.[30–36]

The hard sphere system undergoes a first order solid-fluid phase transition[13] as the packing fraction $\phi$ is increased. The fluid phase is stable below the freezing point, $\phi \approx 0.494$, beyond which fluid and solid coexist up to the melting point at $\phi \approx 0.545$.[13,21] The densest possible packing fraction which can be achieved with identical hard spheres is $\pi/3\sqrt{2} \approx 0.74048$, corresponding to the fcc lattice and its stacking variants — this was known as Kepler's conjecture and has recently been proved.[37] However, metastable solid structures exist (both experimentally and in simulations) where the particles are disordered at relatively high density, in the so-called *random close packed* (RCP) state. Due to the proposed similarity of this disordered phase to glassy phases in more complex systems, these structures are widely studied, even though there is no glass transition in one-component monodisperse hard sphere systems.[22,24,38] It is known from experimental studies[31–33] and computer simulations[22,23] that the disordered phase is thermodynamically unstable with respect to the ordered crystal.

Although the term "random close packed" is widely used in the literature, Torquato and

his co-workers pointed out that the RCP state is very ill defined,[39] and the corresponding packing fraction that is reported varies in the range of $\phi \approx 0.60 - 0.68$ depending on the protocol used to produce the configurations.[40–44] The problem lies in the fact that it is not well defined how random the "random", or how close packed the "close packed" really is. They suggest that the randomness should be quantified by an order parameter, and the term "close packed" should be replaced by the term *jammed*. A system is jammed, if all the particles are jammed in the system, and a particle is jammed if it cannot be translated while the rest of the system is kept fixed. The various states can be sketched on a two dimensional diagram using the packing fraction $\phi$ and an (as yet undefined, "ideal") order parameter $\Psi$ as axes, as shown in Figure 1, reproduced from[39]. The bottom left corner represents the completely disordered spheres of zero size, while the top right corner ($\phi \approx 0.74048$) corresponds to the perfectly ordered and close packed fcc crystal. The grey area represents the inaccessible region where the hard spheres would overlap and all the jammed configurations are located within the loop. Two specific points of the loop boundary can be identified: the point marked with A corresponds to the jammed structure with the lowest possible packing fraction, and the maximally random jammed (MRJ) state is the jammed structure having the lowest possible order parameter. More recently, Torquato et al. refined the definition of jammed states and distinguished three sub-categories (locally, collectively and strictly jammed),[45] and discussed the above figure in more details.[46]

In what follows, nested sampling enables us to quantitatively substantiate the sketch of Figure 1, estimate the parameters of the extremal points of the jammed phase and also calculate the free energies of the various states as a function of the packing fraction.

## II.  THERMODYNAMICS OF HARD SPHERES

We now briefly recall the fundamental thermodynamics of hard spheres. The interesting intensive variable is the pressure $p$, rather than the inverse temperature $\beta$, and so we consider the isotherm-isobar partition function, $\Delta$, that is given by

$$\Delta(\beta, p, N) = \frac{1}{h^{3N} N!} \beta p \int \exp[-\beta(H(\mathbf{p}, \mathbf{q}, V) + pV] \, d\mathbf{p} \, d\mathbf{q} \, dV. \tag{6}$$

After again taking out the momentum dependent part, $Z_p$, we change to fractional coordinates $\mathbf{s}$:

$$\Delta(\beta, p, N) = \beta p \, Z_p(\beta) \int_{(0,1)^{3N}} d\mathbf{s} \int_0^\infty dV \, V^N \exp[-\beta(U(\mathbf{s}, V) + pV]. \tag{7}$$

For hard spheres of a given radius, the potential energy is either zero or infinity that, for a given set of fractional coordinates, depends on the volume,

$$U(\mathbf{s}, V) = \begin{cases} 0 & \text{if } V \geq V_c(\mathbf{s}) \\ \infty & \text{if } V < V_c(\mathbf{s}) \end{cases}, \tag{8}$$

where $V_c(\mathbf{s})$ is the function that gives the lowest permissible volume for the configuration represented by the fractional coordinates $\mathbf{s}$. Substituting into eq. 7,

$$\Delta(\beta, p, N) = \beta p \, Z_p(\beta) \int_{(0,1)^{3N}} d\mathbf{s} \int_{V_c(\mathbf{s})}^\infty dV \, V^N \exp[-\beta pV], \tag{9}$$

and now the last integral can be transformed to a standard exponential integral,

$$\int_{V_c(\mathbf{s})}^\infty dV \, V^N \exp[-\beta pV] = V_c(\mathbf{s})^{N+1} \, E_{-N}(\beta V_c(\mathbf{s})p), \tag{10}$$

$$E_N(x) = \int_1^\infty dt \frac{e^{-xt}}{t^N} \tag{11}$$

The partition function thus becomes

$$\Delta(\beta, p, N) = \beta p \, Z_p(\beta) \int_{(0,1)^{3N}} d\mathbf{s} \, V_c(\mathbf{s})^{N+1} \, E_{-N}(\beta V_c(\mathbf{s})p) \tag{12}$$

$$= \frac{1}{\beta^N p^N} Z_p(\beta) \int_{(0,1)^{3N}} d\mathbf{s} \, \Gamma(N+1, \beta V_c(\mathbf{s})p), \tag{13}$$

where $\Gamma$ is the upper incomplete gamma function,

$$\Gamma(N, x) = \int_x^\infty t^{N-1} e^{-t} \, dt = x^N E_{1-N}(x). \tag{14}$$

For completeness, recall that given the partition function, thermodynamic observables can be computed easily. For example, the Gibbs function is given by

$$G(T, p, N) = -kT \ln \Delta, \tag{15}$$

and so the expected volume and the compressibility are,

$$V = \frac{\partial G}{\partial p} = -kT \frac{1}{\Delta} \frac{\partial \Delta}{\partial p} \tag{16}$$

$$\kappa = -\frac{1}{V} \frac{\partial V}{\partial p}. \tag{17}$$

8

We compute the final partition function by combining the density of states samples, $\Omega_V(u)$, and associated weights, $\gamma(u)$, (see eqs 5 and 2) from a number of independent simulations.

## III. NESTED SAMPLING OF HARD SPHERES

When applying nested sampling to the hard sphere problem, the role of the potential energy $U$ in the algorithm in the Introduction is taken by $V_c(\mathbf{s})$, and the density of states that is obtained is used to calculate the partition function according to equation (13). Notice that the temperature $\beta$ and pressure $p$ occur only as a product in this equation, and therefore a single nested sampling annealing run enables us to calculate observables at all pressures and temperatures.

The size of the sample set, $K$, and the number of attempted moves, $M$, during the MCMC rejection sampling in step 2 of the algorithm are convergence parameters. In the present work, $K$ was between 500 and 1500, and $M$ was between 80 and 320. The proposal distribution for the MCMC sampler was a Gaussian in the position of a single particle. The variance of the Gaussian was periodically adjusted to maintain a constant acceptance rate. In order to quantify finite size effects, systems of 32, 64 and 128 hard spheres were used. For each system size we performed four independent calculations starting from different initial conditions. The total number of MC steps needed to converge the partition function and to reliably reach the crystalline structure with the highest possible packing fraction are shown in Table I. We emphasise again, that after performing the above number of MC steps, arbitrary thermodynamic variables can be calculated at any given packing fraction, without further computational cost. In comparison, Noya et al.[21] performed $4 \times 10^9 - 5 \times 10^{10}$ MC steps to converge a solid/liquid coexistence simulation at a *single* value of the packing fraction, using a unit cell of 5184 particles.

## IV. COMPRESSIBILITY, EQUATION OF STATE AND FREE ENERGY

Given the partition function, it is now easy to evaluate thermodynamic observables, such as the compressibility, $\kappa$, and the compressibility factor, $z = pV/Nk_BT$ (i.e. the equation of state, EoS) as a function of packing fraction, which are shown in Figure 2. There is no exact

TABLE I. Parameters used to produce the converged nested sampling simulations for different number of particles, $N$. $K$ is the number of walkers and $M$ is the number of random walk steps in each iteration. The total number of MC steps is equivalent to the number of evaluations of the $V_c(\mathbf{s})$ function during the whole simulation.

| $N$ | $K$ | $M$ | MC steps |
|---|---|---|---|
| 32 | 500 | 80 | $5.3 \times 10^8$ |
| 64 | 1000 | 80 | $3.9 \times 10^9$ |
| 128 | 1500 | 320 | $8.0 \times 10^{10}$ |

theoretical solution for the EoS of the hard sphere system, and a huge variety of fits can be found in the literature, separately for the fluid and solid phases.[27,48] Nested sampling, on the other hand, gives the equation of state naturally for the whole range of packing densities. The liquid branch is compared to the simple but successful Carnahan-Stirling EoS[49], to the EoS proposed by Kolafa[50], and to one of the most recently published closed-virial EoS, the work of Bannerman et al.[28]. The solid branch is compared to the EoS of Speedy[51] and Alder et al.[25], both developed to predict the behaviour of the fcc crystal. The compressibility shows the expected peak at $\phi = 0.49 - 0.55$, in the coexistence region.[13,21] Both the compressibility and the compressibility factor clearly have a significant finite size error for 32 particles, but it is remarkable that the finite size error of the nested sampling simulation is already very small for $N = 64$, in contrast to what is expected in the case of explicit coexistence simulations.

Access to the partition function enables the calculation of absolute free energies as a post-processing step after the nested sampling simulation is complete, therefore we are able to compute the Gibbs free energy of the hard sphere system. The natural variable of the Gibbs free energy is the pressure, and to shed light on the phase transition, we consider the well known bond order parameter $Q_6$[52] as a further thermodynamic variable. Its value is zero for completely disordered systems and increases with degrees of crystallisation. The $Q_6$ value differs for the perfect fcc and hcp crystals and their stacking variants, $Q_6(\mathrm{fcc}) = 0.57452$ and $Q_6(\mathrm{hcp}) = 0.48476$. Thus the Gibbs free energy becomes, up to an additive constant,

$$G(p, Q_6) = -\beta^{-1} \ln(\Delta(\beta, p, Q_6)) \approx -\beta^{-1} \ln \left[ \int d\mathbf{s} \, \Gamma(N + 1, \beta p V_c(\mathbf{s}) \delta(Q_6(\mathbf{s}) - Q_6)) \right]. \quad (18)$$
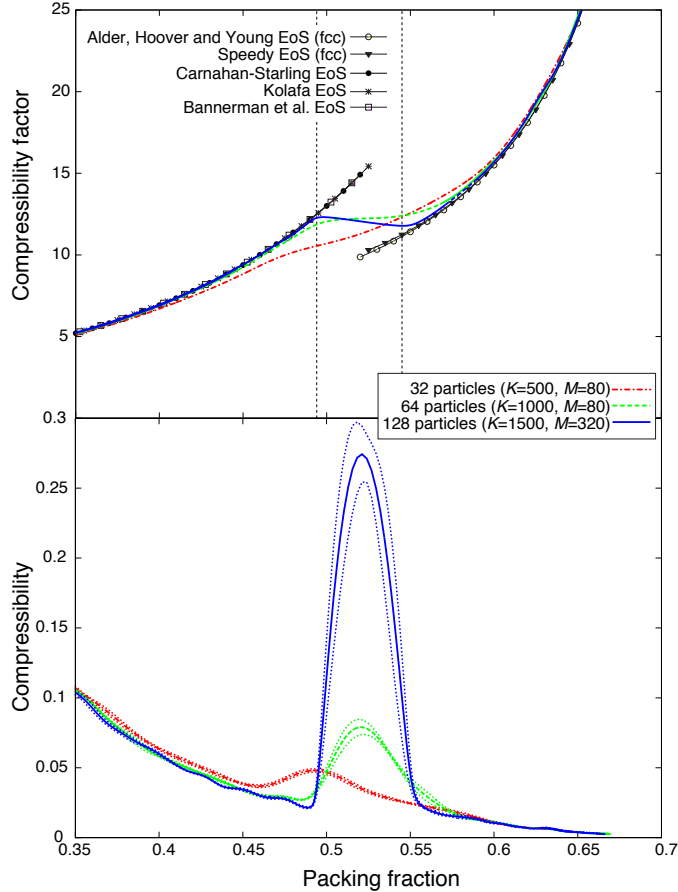
FIG. 2. (Color online) Compressibility factor (top panel) and the compressibility itself (in units of $\beta\sigma^3$, bottom panel) as a function of packing fraction from our nested sampling runs, for 32 (red, dash-dot), 64 (green, dash), 128 (blue, solid) hard sphere particles. Five well known equations of state for hard spheres are shown for comparison with black curves and different symbols. The vertical dashed lines indicate the packing fraction values of freezing and melting. The error of the compressibility factor is on the order of the line thickness, and for the compressibility the standard errors are shown by coloured dotted lines.

The calculated Gibbs free energies are shown in Figure 3 as a function of $Q_6$ for a series of fixed pressure values (with the corresponding average packing fraction in parentheses) in case of the 128 particle system. The free energy barrier corresponding to the phase transition at $p \approx 12.0, \phi \approx 0.513$ is seen to be small, $0.04\,\beta^{-1}$ per particle.

At low pressures (below $p \approx 8.05, \phi \approx 0.45$) there is only one free energy minimum present, corresponding to the fluid phase. As the pressure and packing fraction increases,

11

two distinct free energy minima appear, corresponding to the hcp and fcc stacking orders. Although the two structures have the same density at maximum packing, their thermodynamic properties differ due to the difference in their second and subsequent shell neighbours.[53] The relative stability of the two stacking variants had remained a long-standing question, until the fcc structure was found to be more stable than hcp,[54] however, the Helmholtz free energy difference is only $9 \times 10^{-4} \, \beta^{-1}$ per particle at the melting point, $\phi = 0.545$.[29] At this packing fraction we found that the free energy difference is within the error of our calculations (see second panel on Fig. 3), and the difference is the same order of magnitude as the literature value. More interestingly, at lower packing fractions, $\phi \approx 0.513$ and $\phi \approx 0.50$, we found larger free energy differences, and in favour of the hcp structure. Although the difference is still small, it might explain the experimental observation that hard-sphere colloids rarely crystallise directly into the fcc structure, but initially form a randomly stacked hexagonal close-packed (rhcp) crystal[35,36], which then slowly transforms to the stable fcc phase.[20,55]

## V.   ORDER PARAMETER VS PACKING FRACTION

To quantitatively substantiate the diagram of Figure 1, we again use the bond order parameter $Q_6$ as the order parameter for the vertical axis. The calculated $Q_6$ parameters of the nested sampling configurations as a function of their packing fraction are shown with black dots in Figure 4 for all system sizes. The phase transition corresponding to the compressibility peak near $\phi \approx 0.49$ is revealed as a sharp increase in crystallinity.

Nested sampling allocates samples to various regions of phase space in proportion to the region's phase space volume. Thus white areas in Figure 4 (i.e. regions with no samples in our simulation) correspond to regions with very small phase space volume (approximately $< 1/K$ for a given packing fraction). Accordingly, at low packing fraction, the crystalline states while clearly accessible, occupy a negligible phase space volume. Conversely, at high packing fraction, the disordered states have negligible volume, which manifests itself in the well known result that the disordered phases are thermodynamically unstable here. The samples with the highest possible packing fraction fall into distinct groups with different order parameter values, because the nested sampling simulation found both the fcc and the hcp phases. However, it is interesting to see that the branches leading to either stacking variants of the crystal separate at relatively small $\phi$ values, already around the melting
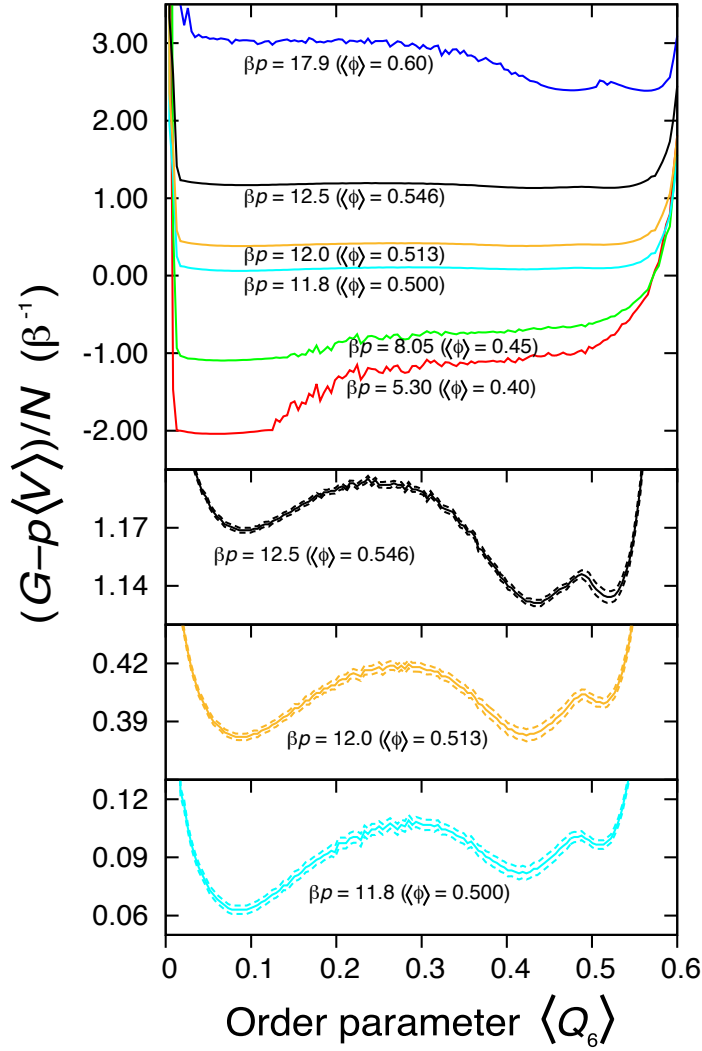
FIG. 3. (Color online) Gibbs free energy per particle of a periodic lattice of hard spheres with 128 particles in the unit cell as a function of the $Q_6$ order parameter at different pressure values ($\beta p$ in $\sigma^{-3}$ units) calculated by eq. 18. In order to show the curves for different pressure values on the same plot, $-p\langle V \rangle$ was added to the free energy. The corresponding packing fraction ($\phi$) values are also indicated. The bottom three panels show the free energy curves in a larger scale around the phase transition, and the error is shown by dashed lines.

point.

To distinguish the region of jammed states from the totally inaccessible region of high packing fraction and low order at the right hand side of the plot, we chose 10000 points from the nested sampling configurations in the packing fraction range of $0.40 - 0.73$ and rapidly
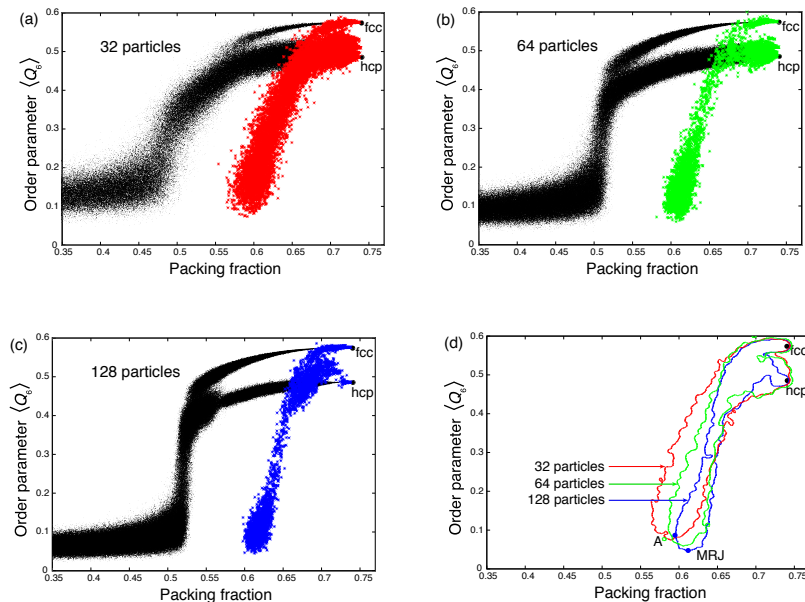
FIG. 4. (Color online) Order parameter $Q_6$ as a function of the packing fraction for $N = 32, 64$ and $128$ hard spheres. The black dots represent the configurations generated during the nested sampling runs and the coloured symbols represent the quenched jammed configurations (see text). Panel (d) shows the area occupied by the jammed states of 32 (red), 64 (green), 128 (blue) using a contour line after a little gaussian smearing. Perfect fcc and hcp structures are indicated by filled black circles, and the points MRJ (maximally random jammed state found) and A (jammed state with lowest possible packing fraction found) are also shown on the last panel for the 128 particle system.

quenched them by increasing the particle diameter (and hence the packing fraction) with minimal change in the location of particle centres until the entire system became jammed. Further details on this procedure are given in the Appendix.

The final jammed configurations are plotted by coloured stars in Figure 4. The size and shape of the region they occupy is converges with increasing particle number to the region of RCP states. It can be seen from these figures that jammed configurations exist in states corresponding to a wide variety of order and packing fraction. There are jammed states with packing fraction smaller than 0.6, and many others are as disordered as the fluid phase, or almost perfect crystals containing different defects.

To further characterise the structure of the extremal configurations "A" and "MRJ" of

TABLE II. Proportion of particles with given number of neighbours in the extremal configurations "A" and "MRJ" of Figure 4. The number of neighbours was counted with a tolerance of 1% on the inter-particle distance. Proportions less than 5% are likely unreliable and thus are omitted, and percentage values are rounded.

| Configuration | $\phi$ | $Q_6$ | Mean number of neighbours | Number of neighbours | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | 4 | 5 | 6 | 7 | 8 |
| MRJ | 0.612 | 0.046 | 5.3 | 14% | 21% | 22% | 23% | 9% |
| A | 0.598 | 0.083 | 5.6 | 13% | 32% | 23% | 23% | - |

the jammed region found in our simulations, we show the fractions of particles with varying numbers of neighbours in Table II. The distribution of the number of neighbours peaks near 5 and 6 neighbours. To put the neighbour count in context we note that in the closed packed crystalline phases particles have 12 neighbours each, and 8 in the bcc phase.

## VI. CONCLUSION

We have demonstrated the nested sampling algorithm in the condensed phase by applying to the three dimensional periodic hard sphere system, and studied the solid/liquid phase transition. Remarkably, the observed finite size effect is small so that both the location of the melting and freezing points and the compressibility and equation of state as a function of packing fraction are reasonably well converged for 128 particles. Coexistence simulations using a similar number of Monte Carlo moves and $\sim 50$ times more particles obtained just the location of the transition[21].

Using the samples collected, we were able to plot the Gibbs free energy as a function of order parameter and packing fraction, showing the phase transition explicitly and quantitatively, demonstrating that the barrier for transition to crystallinity is small, and that above the phase transition the entropic contribution of the jammed states to the free energy is negligible. The global unbiased nature of this sampling scheme allowed us to study the phase diagram using the packing fraction and the $Q_6$ order parameter as thermodynamic variables, identifying the extent of the jammed phase. The "maximally random jammed" state we found, i.e. the jammed state with the lowest $Q_6$ order parameter, and the least

dense jammed state, are both surprisingly disordered, having $Q_6$ values of less than 0.1. While we do not claim that there are no remaining finite size effects in the 128 particle system for any observable, the extent of the "forbidden region" in the phase diagram of Figure 4 shows little change in going from 32 to 128 particles.

The ease and simplicity of nested sampling combined with its power to calculate the partition function directly and all thermodynamic observables that follow make it an excellent tool to use in computational statistical mechanics of particle systems in the condensed phase.

## ACKNOWLEDGMENTS

*  lb415@cam.ac.uk

[1]  J. Skilling, in *AIP Conference Proceedings*, Vol. 735 (2004) p. 395.

[2]  J. Skilling, J. of Bayesian Analysis **1**, 833 (2006).

[3]  F. Feroz and M. P. Hobson, Mon. Not. R. Astron. Soc. **384**, 449 (2008).

[4]  L. B. Pártay, A. P. Bartók, and G. Csányi, J. Phys. Chem. B **114**, 10502 (2010).

[5]  N. S. Burkoff, C. Várnai, S. A. Wells, and D. L. Wild, Biophys. J. **102**, 878 (2012).

[6]  H. Do, J. D. Hirst, and R. J. Wheatley, J. Chem. Phys. **135**, 174105 (2011).

[7]  H. Do, J. D. Hirst, and R. J. Wheatley, J. Phys. Chem. **116**, 4535 (2012).

[8]  H. Do, , and R. J. Wheatley, J. Chem. Theory Comput. **9**, 165 (2013).

[9]  S. O. Nielsen, J. Chem. Phys. **139**, 124104 (2013).

[10]  M. N. Rosenbluth and A. W. Rosenbluth, J. Chem. Phys. **22**, 881 (1954).

[11]  W. W. Wood and J. D. Jacobson, J. Chem. Phys. **27**, 1207 (1957).

[12]  B. J. Alder and T. E. Wainwright, J. Chem. Phys. **27**, 1208 (1957).

[13]  W. G. Hoover and F. H. Ree, J. Chem. Phys. **49**, 3609 (1968).

[14]  Q. Wang and M. S. Cheung, Biophys. J. **102**, 2353 (2012).

[15] T. Ishikawa and T. J. Pedley, J. Fluid Mech. **588**, 399 (2007).

[16] A. Najafi and R. Golestanian, Phys. Rev. E **69**, 062901 (2004).

[17] S. Torquato and F. Lado, Phys. Rev. B **33**, 6428 (1986).

[18] W. R. Smith and D. Henderson, Mol. Phys. **19**, 411 (1970).

[19] R. L. Davidchack and B. B. Laird, J. Chem. Phys. **108**, 9452 (1998).

[20] S. Pronk and D. Frenkel, J. Chem. Phys. **110**, 4589 (1999).

[21] E. G. Noya, C. Vega,  and E. de Miguel, J. Chem. Phys. **128**, 154507 (2008).

[22] E. Sanz, C. Valeriani, E. Zaccarelli, W. C. K. Poon, P. N. Pusey,  and M. E. Cates, Phys. Rev. Lett. **106**, 215701 (2011).

[23] T. Schilling, H. J. Schöpe, M. Oettel, G. Opletal,  and I. Snook, Phys. Rev. Lett. **105**, 025701 (2010).

[24] S. R. Williams, I. K. Snook,  and W. van Megen, Phys. Rev. E **64**, 021506 (2001).

[25] B. J. Alder, W. G. Hoover,  and D. A. Young, The Journal of chemical physics **49**, 3688 (1968).

[26] A. Santos and M. López de Haro, The Journal of chemical physics **130**, 214104 (2009).

[27] W. Wang, M. K. Khoshkbarchi,  and J. H. Vera, Fluid phase equilibria **115**, 25 (1996).

[28] M. N. Bannerman, L. Lue,  and L. V. Woodcock, The Journal of chemical physics **132**, 084507 (2010).

[29] P. G. Bolhuis, D. Frenkel, S.-C. Mau,  and D. A. Huse, Nature **388**, 235 (1997).

[30] P. N. Pusey and W. van Megen, Nature **320**, 340 (1986).

[31] H. J. Schöpe, G. Bryant,  and W. van Megen, Phys. Rev. Lett. **96**, 175701 (2006).

[32] K. Schätzel and B. J. Ackerson, Phys. Rev. E **48**, 3766 (1993).

[33] J. L. Harland and W. van Megen, Phys. Rev. E **55**, 3054 (1997).

[34] Z. Cheng, P. Chaikin, W. Russel, W. Meyer, J. Zhu, R. Rogers,  and R. Ottewill, Materials & Design **22**, 529 (2001).

[35] J. Zhu, M. Li, R. Rogers, W. Meyer, R. H. Ottewill, W. B. Russel,  and P. M. Chaikin, Nature **387**, 883 (1997).

[36] A. Petukhov, I. Dolbnya, D. Aarts, G. Vroege,  and H. Lekkerkerker, Physical Review Letters **90**, 028304 (2003).

[37] T. C. Hales, Ann. Math. **162**, 1065 (2005).

[38] M. D. Rintoul and S. Torquato, Phys. Rev. Lett. **77**, 4198 (1996).

[39] S. Torquato, T. M. Truskett,  and P. G. Debenedetti, Phys. Rev. Lett. **84**, 2064 (2000).

[40] G. Scott and D. Kilgour, Br. J. Appl. Phys. **2**, 863 (1969).

[41] O. Pouliquen, M. Nicolas, and P. D. Weidman, Phys. Rev. Lett. **79**, 3640 (1997).

[42] W. S. Jodrey and E. M. Tory, Phys. Rev. A **32**, 2347 (1985).

[43] J. Tobochnik and P. Chapin, J. Chem. Phys. **88**, 5824 (1988).

[44] W. M. Visscher and M. Bolsterli, Nature **239**, 504 (1972).

[45] S. Torquato and F. H. Stillinger, J. Phys. Chem. B **105**, 11849 (2001).

[46] S. Torquato and F. H. Stillinger, Rev. Mod. Phys. **82**, 2633 (2010).

[47] J. Skilling, in *AIP Conference Proceedings*, Vol. 1193 (2009) p. 277.

[48] A. Mulero, C. A. Galan, M. I. Parra, and F. Cuadros, *Theory and Simulation of Hard-Sphere Fluids and Related Systems*, Lecture Notes in Physics, Vol. 753 (Springer-Verlag, Berlin, 2008).

[49] N. F. Carnahan and K. E. Starling, J. Chem. Phys. **51**, 635 (1969).

[50] T. Boublik, Mol. Phys. **59**, 371 (1986).

[51] R. J. Speedy, Journal Of Physics-Condensed Matter **10**, 4387 (1998).

[52] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, Phys. Rev. B **28**, 784 (1983).

[53] S. Pronk and D. Frenkel, Physical Review Letters **90**, 255501 (2003).

[54] L. V. Woodcock, Nature **385**, 141 (1997).

[55] Z. Cheng, P. M. Chaikin, J. Zhu, W. B. Russel, and W. V. Meyer, Phys. Rev. Lett. **88**, 015501 (2001).

## VII. APPENDIX

### A. Computational details

During the Monte Carlo process, both the particle positions and the shape of the simulation box are changed. In each Monte Carlo step either the movement of a single particle or a lattice shape change is proposed (the probability of the latter being $1/(N+1)$). In case of a shape change, the three lattice vectors are perturbed by small random vectors and, in order to avoid very flattened box shapes for which finite size effects would be very large, we applied the constraint that the smallest height of the box may not become less than 55% of the height of the cubic box at the same volume. To keep the acceptance ratio between 10% and 20% for both the particle position and lattice shape changes, the step sizes of each are readjusted at every 1000 NS iteration.

When evaluating sums in eq. 13 we work in quadruple precision. Alternatively, if using the usual double precision floating point arithmetic one has to be very careful to avoid under- or overflow.

### B. Quenching

Here we give details on the procedure that we used to quench configurations in order to obtain noncrystalline jammed configurations (note that configurations that include "rattle" particles are not excluded). During the quenching process the particle size was increased after every MC step until the whole configuration was jammed. We considered a configuration jammed when further relaxation could only increase the packing fraction by less then 0.3% in $10^6$ trial movements of the particles. Figure 5 shows some evidence that this procedure really does result in configurations that are very close to being jammed: the largest of the nearest-neighbour distances as a function of iteration number approaches the particle diameter approximately according to a power law with an exponent of $-1/2$, as expected for a Monte Carlo process. We have to note that this quenching process produces locally jammed configurations which are not necessarily always collectively jammed too.

The question arises how the location of a given jammed state on the phase diagram depends on the location of the initial configuration it was quenched from. We chose several packing fraction values and took the corresponding sample set of the nested sampling run
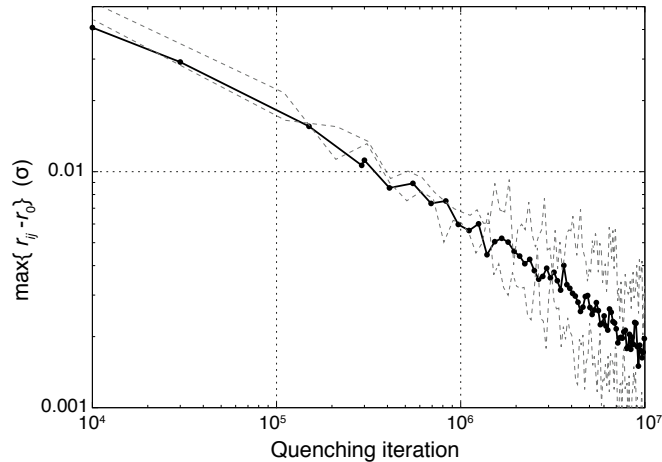
FIG. 5. Convergence towards jammed states during the quenching process. The difference between the furthest and closest first neighbour distances (the latter is, by definition, the particle diameter) is calculated within every configuration, and plotted against the number of MC steps. The black line is the average of several quenching processes. The grey dashed curves represents the slowest and the fastest process from our sample. Both axes are in the logarithmic scale.

in order to have sets of configurations uniformly distributed in configuration space at given packing fraction levels and we quenched those with the same method as previously described. The chosen packing fraction levels and the quenched jammed states are shown with matching colours in Figure 6 for an individual run with 64 and two runs of the 128 hard sphere systems. The figures show the same general picture for both system sizes. Different initial packing fraction levels result in different groups of jammed states. Starting from the fluid phase ($\phi \approx 0.45$) the jammed states have relatively small packing fraction and low order, while in case of quenching configurations taken from above the freezing point the jammed states end up in the region of crystalline structures.

This suggests that around the freezing point the relative phase space volume of the crystalline structures is small, approximately only 0.2% of the points quenched from the set with $\phi \approx 0.485$ ended up with a high order parameter. Thus, a large part of the phase space here leads to different disordered local minima. However, above the freezing point, the relative phase space volume of ordered structures becomes dominant, regardless of the

20

fact that an immense number of distinct basins exist at the same packing fraction level, corresponding to jammed states with disordered structures.
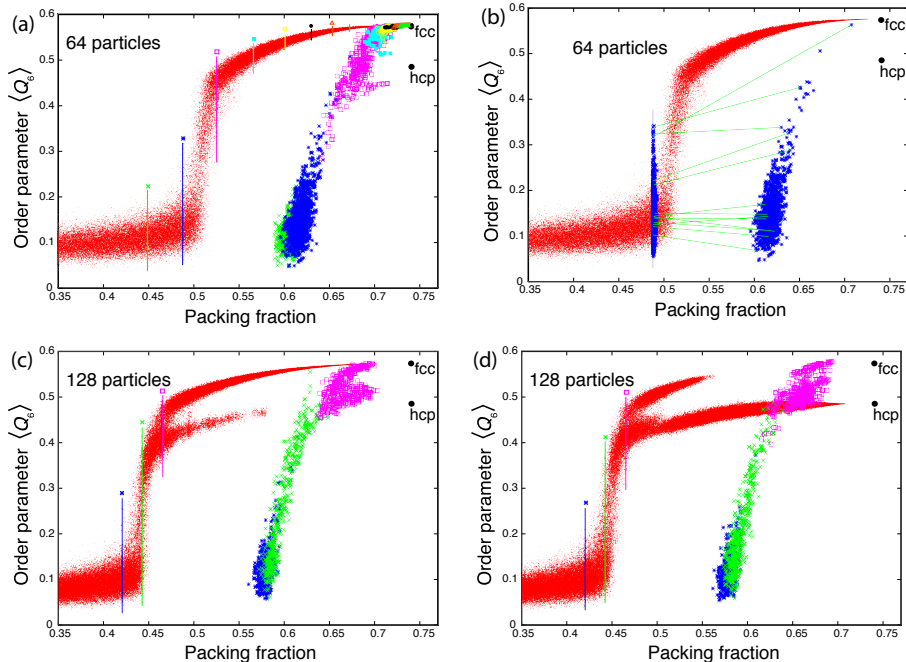


FIG. 6. (Color online) The $Q_6$ order parameter as a function of the packing fraction in case of an individual run with 64 hard spheres (top panels) and two runs with 128 hard spheres (bottom panels). The red dots represent the configurations generated during the nested sampling runs. The vertical lines indicate the packing fraction levels at which the sample sets were saved and subsequently quenched in order to obtain the jammed configurations. The quenched jammed configurations are shown by symbols with matching colours. For easier identification, at the top of the vertical lines, the corresponding symbols are also shown. In the top right panel some of the initial configurations and their corresponding jammed structures are connected by green lines to show how the quenching changes the order parameter. The points corresponding to perfect fcc and hcp structures are shown by black circles.

The quenching process allows the characterisation of the phase transition without the use of an explicit and somewhat arbitrarily chosen order parameter. In Figure 7 we plot a random subset of our samples with the packing fraction before and after quenching on the horizontal and vertical axes. The phase transition is again visible, showing that configurations just below and above the critical packing fraction are close to jammed states with relatively low and high packing fraction, respectively.
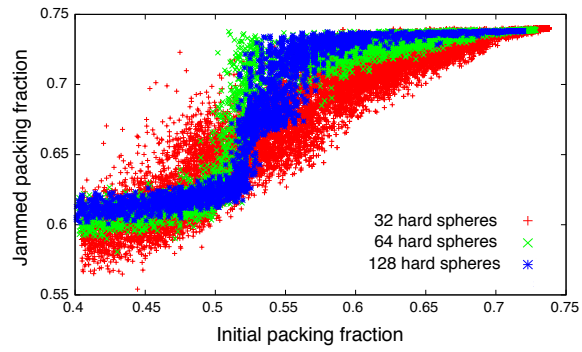
FIG. 7. (Color online) Packing fraction of the quenched jammed configurations as a function of their initial packing fraction for the 32, 64 and 128 hard spheres systems.