ARTICLE

Measuring the size of individual particles from three-dimensional imaging experiments

Rei Kurita[1], David B. Ruffner[2] & Eric R. Weeks[3]

Often experimentalists study particulate samples that are nominally monodisperse. In reality, many samples have a polydispersity of 4–10%. At the level of an individual particle, the consequences of this polydispersity are unknown as it is difficult to measure an individual particle size from images of a dense sample. Here we propose a method to estimate individual particle radii from three-dimensional data of the particle positions. We first validate our method with simulations. We then apply our method to experimental data of colloidal suspensions observed with confocal microscopy. We demonstrate that we can recover the full particle size distribution *in situ*. Finally, we use our method to study the relationship between homogeneous colloidal crystal nucleation and particle sizes. We show that nucleation occurs in regions that are more monodisperse than average.

[1] Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan. [2] Department of Physics and Center for Soft Matter Research, New York University, New York, New York 10003, USA. [3] Department of Physics, Emory University, Atlanta, Georgia 30322, USA. Correspondence and requests for materials should be addressed to R.K. (email: kurita0@iis.u-tokyo.ac.jp).

A wide variety of techniques exist for three-dimensional (3D) imaging of collections of particles, such as granular materials, soils and colloidal suspensions[1,2]. Our particular interest is in colloidal suspensions; these have been successfully used as model systems for understanding phase transitions for several decades[3,4], and are interesting in their own right owing to industrial relevance[5]. Confocal microscopy can be used to take 3D images of fluorescent colloidal particles deep within a sample[2,6,7]. When coupled with particle tracking techniques, the motion of thousands of individual colloidal particles can be followed over long periods of time[8–10]. This technique has been used to investigate the colloidal glass transition[7–9,11], crystallization[12–14], colloidal gels[15–17], capillary waves[18,19], sedimentation[12,20] and a variety of other questions (see ref. 2 for a review). One advantage of confocal microscopy of colloids is that particle tracking data are similar to simulations that also provide particle positions over long periods of time.

However, experimental samples are always polydisperse: even for a nominally single-component sample, the particles have a variety of sizes[21]. This is quantified by the polydispersity p, defined as the s.d. of particle radii divided by the mean radius. For many samples, $p \sim 0.04$–0.10 (ref. 21). Numerical simulations show that the effects of the particle size distribution are not negligible. For example, crystal nucleation is difficult or impossible for more polydisperse samples[22,23]. The crystal–liquid phase boundary depends on the polydispersity[24]. The sensitivity to volume fraction near the glass transition depends on the composition in nontrivial ways[25,26]. Experimentally, the influence of polydispersity on colloidal crystallization has been demonstrated[27,28], and there is also some understanding of how polydispersity influences the rheological behaviour of a colloidal sample[29]. However, these are limited to studies of the spatially averaged properties of the sample. Microscopy is useful for local properties, but particle size fluctuations of a few per cent are not easily detectable. It would be desirable to know particle sizes for more direct comparison with simulations. Furthermore, in some cases, neglecting these sizes in an experiment can lead to wrong conclusions. One example is that the pair correlation function $g(r)$ can show a qualitatively incorrect dependence on control parameters if the particle sizes are treated as all identical[30]. A second example is that the apparent compressibility of a random close-packed sample depends qualitatively on whether individual particle sizes are taken into account[31–33].

In this work, we introduce a general method for using 3D data to determine the size of individual particles in any moderately concentrated sample, in general with volume fractions $\phi \gtrsim 0.4$. We use simulation data to verify that our method works well in a variety of sample types. We then check our method with experimental data from confocal microscopy, which compares well with data on the same samples from holographic microscopy. Finally, we demonstrate the utility of our method using previously published experimental data from confocal microscopy of colloids. In particular, we show that colloidal crystal nucleation is sensitive to the local polydispersity: nucleation happens in locally monodisperse regions. Our method is not limited to confocal microscopy and colloidal samples, but rather works with any data of the 3D positions of a collection of particles.

Results

Size estimation algorithm. Owing to diffraction limits, it is difficult to directly determine the radii of individual particles from microscopy images to better than $\pm 0.1 \, \mu m$ (ref. 34). Defining the edge is somewhat arbitrary and varies depending on particle properties and the details of the illumination. Other 3D imaging techniques have similar issues[1]. In contrast, it is much easier to calculate the mean radius $\bar{a}$ of particles with a variety of techniques[21]. Likewise, from the centres of particles, the separations between neighbouring particles r_{ij} can be easily calculated. Our estimation method for
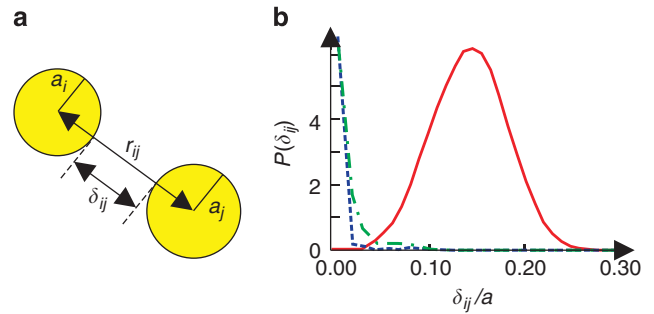

Figure 1 | Probability distribution function of δ_{ij}. (**a**) Sketch defining key distances. (**b**) Probability distribution function of δ_{ij} from simulation data for several systems. The systems and the number of neighbours Z in each case are: a dense liquid with $\phi = 0.51$ (solid red curve, $Z = 7$), a random close-packed sample (dashed blue curve, $Z = 5$) and a colloidal gel (dash-dotted green curve, $Z_i = 5c_i/13$). The details of the choices of Z are discussed in the text.

particle sizes uses only $\bar{a}$ and r_{ij}. The key idea of our method is that a large particle will be slightly farther from its neighbours and thus have larger values for its r_{ij}, and likewise a smaller particle will have smaller values of r_{ij}.

To start, we relate the pairwise separations r_{ij} as

$$r_{ij}(t) = a_i + a_j + \delta_{ij}(t), \tag{1}$$

where particle j is a nearest neighbour particle of particle i, r_{ij} is the measured distance between i and j, a_i and a_j are their radii, and $\delta_{ij}(t)$ is a surface-to-surface distance between their particles. These distances are sketched in Fig. 1a. When choosing neighbours, we use the particles j that are closest to i (smallest separations r_{ij}). Typically, we use $Z = 5$–7 of these closest neighbour particles; this choice is justified below. Often these data come from particle tracking, and so $r_{ij}(t)$ and $\delta_{ij}(t)$ depend on time t. Next, we take an average of r_{ij} with respect to the nearest neighbour particles j, and then $\langle r_{ij}(t) \rangle_j = a_i + \langle a_j \rangle_j + \langle \delta_{ij}(t) \rangle_j$, where $\langle \rangle_j$ means an average over particle j. Thus, we obtain

$$a_i = \langle r_{ij}(t) \rangle_j - \langle a_j \rangle_j - \langle \delta_{ij}(t) \rangle_j. \tag{2}$$

This is exact, but the quantities $\delta_{ij}(t)$ are unknown. We estimate this by replacing $\delta_{ij}(t)$ with its time- and particle-averaged value, the mean gap distance $\bar{\delta} \equiv \langle r_{ij}(t) \rangle_{i,j,t} - 2\bar{a}$, where the average is over all particle pairs and all times. Our algorithm is then:

$$a_i^{(0)}(t) = \bar{a}, \tag{3}$$

$$a_i^{(n)}(t) = \langle r_{ij}(t) \rangle_j - \langle a_j^{(n-1)}(t) \rangle_j - \bar{\delta}, \tag{4}$$

where the superscripts denote iteration. The more we iterate equation 4, the more information we obtain from particles far away from a given particle. In fact, $a_i^{(n)}(t)$ is unchanged for $n \geq 10$ as $a_i^{(10)}$ includes the information from several thousand particles, thus we fix $n = 10$ for the number of iterations in this paper. Of course, the particle radius does not depend on time, so after the 10th iteration, we time average $a_i^{(10)}(t)$ to obtain the estimated particle radius $a_i^{(10)}$. Time averaging after each iteration of equation 4 does not change the results.

Uncertainty of the estimation. There are several sources of uncertainty in this estimation. First, there is the uncertainty of each
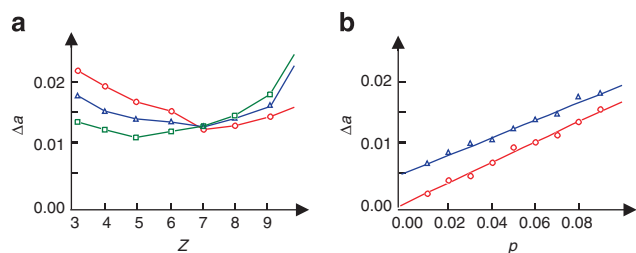
Figure 2 | Dependence of estimation uncertainty Δa on parameters.
(**a**) The uncertainty Δa as a function of the chosen number of neighbours Z used for the averaging. The circles, triangles and squares correspond to Δa for volume fractions $\phi = 0.51$, 0.56 and 0.64 (RCP), respectively. In each case, the sample polydispersity is $p = 0.07$. (**b**) Δa as a function of the bulk polydispersity p. The circles and the triangles correspond to Δa at RCP and $\phi = 0.51$, respectively. The solid lines are the fitting lines for Δa.
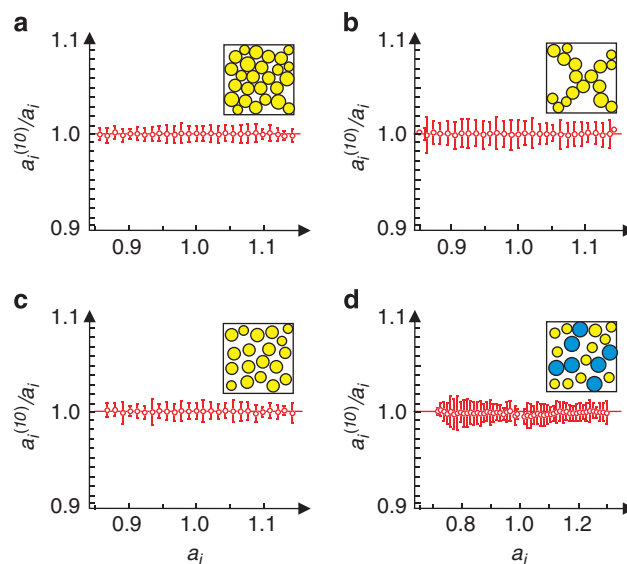

Figure 3 | The estimated radius $a_i^{(10)}/a_i$ as a function of the true radius a_i for four simulated systems. (**a**) RCP, (**b**) a colloidal gel at $\phi = 0.10$, (**c**) a nominally monodisperse suspension at $\phi = 0.51$ and (**d**) a nominally binary suspension at $\phi = 0.51$. For (**a**–**c**), the polydispersities are $p = 0.07$. For the binary sample, the size ratio is 1:1.3, the number ratio is 1:1, and each species has an individual polydispersity of $p = 0.04$. The error bars correspond to the s.d. of of $a_i^{(10)}/a_i$ between a_i and $a_i + 0.01$. The insets show sketches of each system.

particle position. Typically this is about 5–8% of the mean radius, leading to a 8–10% uncertainty of r_{ij} (refs 10,35). However, these errors are nearly time-independent, so those errors are greatly diminished by time averaging. Second, our approximation for $\bar{\delta}$ is weaker in the case that the distribution of $\delta_{ij}(t)$ is broad. Distributions of δ_{ij} are shown in Fig. 1b. As we choose neighbouring particles j to be those closest to particle i, choosing more neighbours results in a broader distribution of δ_{ij} because the additional neighbours have a larger δ_{ij} by construction. That is, $\bar{\delta}$ increases monotonically with increasing Z. However, using too few neighbours means that the average $\langle r_{ij}(t) \rangle$ in equation 4 is poor. Below, we use simulation data to determine that $Z = 5$–7 nearest neighbours is an optimal choice. Third, independent of a given choice of Z, some particles will simply be farther from their neighbours, and some will be closer. For example, in a dense suspension this relates to the size of the 'cage' formed by the nearest neighbour particles[9]. Again, time averaging helps. If particles can rearrange and find new neighbours, then $\bar{\delta}$ becomes a better approximation for $\langle \delta_{ij}(t) \rangle_t$. Fourth, an additional approximation is that we use $\langle \delta_{ij} \rangle \approx \bar{\delta}$ to go from equation 2 to equation 4, which assumes that δ_{ij} is uncorrelated with the particle size a_i. Our simulations suggest that this is reasonable, but one can imagine situations where it might not be (for example, with charged particles where the charge depended on particle size). In summary, the greatest strength of our algorithm is time averaging, and past that, a sensible choice for the number of nearest neighbours Z is useful. Our tests show that time averaging over ~20 different times is sufficient for reasonable results.

Verification of the estimation. To verify our radius estimation method, we simulate a variety of systems and compare the estimated radius of each particle with its true radius. The error is given by $\delta a_i = \left| a_i^{(10)} - a_i \right|$, where $a_i^{(10)}$ is the estimated value and a_i is the true value. $\Delta a \equiv \langle \delta a_i \rangle_i / \bar{a}$ is the mean fractional error in the estimated particle radius. Also relevant is the polydispersity p of the simulated sample, defined as $p = \sqrt{\langle (a_i - \bar{a})^2 \rangle} / \bar{a}$, where the averages are over all particles i. A priori, the best guess for each particle size is $\bar{a}$ with a fractional uncertainty p. If the mean estimation error Δa is less than p, the estimation method improves our knowledge of the particle sizes; we will show this is true for the simulated data.

Our first test case is a random close-packed sample. In such a sample particles do not move, and so time averaging cannot be used. However, particles are packed so that they can contact each other, that is, $\bar{\delta}$ is close to zero. The number of contacting neighbours varies from particle to particle, so it is not clear how many neighbours should be considered. Accordingly, we plot Δa as a function of Z in Fig. 2a. We find that Δa is a minimum at $Z = 5$, and is much smaller than p (0.01 versus 0.07 in this case). In practice we

find that the $Z = 5$ closest neighbours are nearly always in contact, and $\bar{\delta} \approx 10^{-4} a$. Figure 2a shows that increasing Z to include some non-contacting neighbours only increases the uncertainty a small amount.

It is possible that while Δa is small, that there are systematic errors depending on the real particle size a_i. To test this, in Fig. 3a we show the ratio between the estimated radius and the given radius $a_i^{(10)}/a_i$ as a function of a_i for a polydispersity $p = 0.07$ random close packing (RCP) sample. The symbols and the error bars correspond to the mean and s.d. of the distribution of $a_i^{(10)}/a_i$ between $[a_i, a_i + 0.01]$, respectively. $a_i^{(10)}/a_i$ should be 1 if our estimation is perfect and indeed we find $a_i^{(10)}/a_i = 1.000 \pm 0.013$. The quality of the results is nearly uniform as a function of particle size. To check the validity of our method for RCP samples with different polydispersity, we plot the uncertainty Δa as a function of sample polydispersity p in Fig. 2b. We find $\Delta a \approx p/6$ (ref. 33).

A colloidal gel shares a similarity to a RCP sample (touching or nearly touching particles), and has a significant difference (much lower volume fraction). As an example of how to adapt our algorithm, we consider a simulated depletion gel[36]. In a colloidal gel, particles are stuck to their neighbours and form a large network. Often the attractive interactions have a range, which is true for depletion gels (see discussion in Methods). Thus the distribution of δ_{ij} for gels can be slightly broader than that for RCP, as shown in Fig. 1b, though the mean average of δ_{ij} is close to 0. Some time averaging is possible, although such samples are frequently nonergodic.

The contacting particles make gels similar to RCP samples locally. However, the contact number can fluctuate greatly in a colloidal gel. Therefore, rather than fixing the number of neighbours Z to average over, we use Z_i neighbours for the average (equation 4). To determine Z_i, we define the coordination number c_i as the number of particles within a distance $2.8a$ of particle i, where $2.8a$ is the first minimum of the pair correlation function. We find the average coordination number $\bar{c} \approx 13.1$ for a RCP sample, but this can be smaller for some gels[15]. Thus for every particle we select the

number of neighbours using $Z_i = 5c_i/13$ where we round Z_i to the nearest integer, and as always we select the Z_i particles closest to particle i to be the neighbours. For tenous gels at dilute concentrations, for many particles Z_i is fairly small; also, δ_{ij} has a broader distribution, and so Δa will be worse than the RCP case. However, Δa is improved by time averaging. Figure 3b shows the ratio between the time-averaged estimated radius and the given radius $a_i^{(10)}/a_i$ as a function of the true radius a_i for the colloidal gel. We find that $a_i^{(10)}/a_i = 1.000 \pm 0.018$. $\Delta a = 0.018$ is much smaller than the polydispersity $p = 0.070$.

Moving from gels, we next consider a dense suspension of purely repulsive (hard-sphere) particles. Here, no particles are in contact, so δ_{ij} has a much broader distribution (Fig. 1b); however, time averaging is even more powerful. We show $a_i^{(10)}/a_i$ as a function of a_i at $\phi = 0.51$ in Fig. 3c, finding $a_i^{(10)}/a_i = 1.000 \pm 0.014$. Yet again $\Delta a = 0.014$ is much smaller than the polydispersity $p = 0.070$.

For a dense suspension, it is not obvious how many nearest neighbours should be used in the average (equation 4), so we plot Δa as a function of Z in Fig. 2a for two different volume fractions. Δa is minimized at $Z = 7$ for the non-RCP samples (circles and triangles in the figure), so we fix our choice $Z = 7$ for all our $\phi < 0.6$ experimental data (discussed below). Figure 2a demonstrates that Δa does not depend too sensitively on this choice. However, it should be expected that for a more dilute system, the importance of caging decreases, and the number of neighbours a particle has will fluctuate significantly. For fixed polydispersity $p = 0.070$, we find $\Delta a = 0.023$ for $\phi = 0.45$ and $\Delta a = 0.060$ for $\phi = 0.40$. This suggests that for $\phi \lesssim 0.40$, the estimation method may not be useful without further modifications.

To check the influence of the sample polydispersity at fixed $\phi = 0.51$, we vary p with results shown in Fig. 2b (triangles). We find $\Delta a \approx 0.005 + p/7$, suggesting that the estimation is useful for samples with $p > 0.01$, that is, any realistic sample. Δa is non-zero when $p = 0$, in contrast to the RCP case. This is due to the distribution of δ_{ij} in a dense but non-contacting sample.

The last case we examine with simulation data is a nominally binary sample. We simulate a dense suspension composed of particles with a size ratio 1:1.3 (mean sizes 0.877 and 1.14) and number ratio 1:1. For both 'small' and 'large' particles, there is a polydispersity $p = 0.04$. The results are shown in Fig. 3d, and we find $a_i^{(10)}/a_i = 1.000 \pm 0.024$ at $\phi = 0.51$. (Here we have fixed $Z = 7$.) Again, there is no strong dependence on the true particle size a_i. However, the uncertainty Δa for the binary mixture is larger than what is found for the nominally monodisperse distribution. This is consistent with the overall polydispersity of the sample being larger, $p = 0.14$.

An important use of our technique is to measure a particle size distribution *in situ*; we wish to validate this idea with the simulation data. To do this, we compare the estimated radius distribution $P(a_i^{(10)})$ with the true radius distribution $P(a_i)$ in Fig. 4a,b. In both the nominally monodisperse sample and the nominally binary sample, the estimated distribution (symbols) is quite close to the true distribution (curves). We find that for a single-species sample with a Gaussian distribution of radii with polydispersity p, the apparent polydispersity is $\sqrt{p^2 + \Delta a^2}$. Given that for most situations we have shown $\Delta a \ll p$, our technique will only slightly increase the apparent polydispersity of a sample.

One key difference between simulations and experiments is the boundary condition. Our simulations have periodic boundaries. In an experiment, we cannot find all nearest neighbours of a particle when the particle is located at the edges of an image. This situation is similar to colloidal gels, where the number of nearest neighbours varies for each particle, and we adopt the same solution used there. For each particle, we average over a number of nearest neighbours given by $Z = 7c_i/13$, where c_i is the observed coordination number
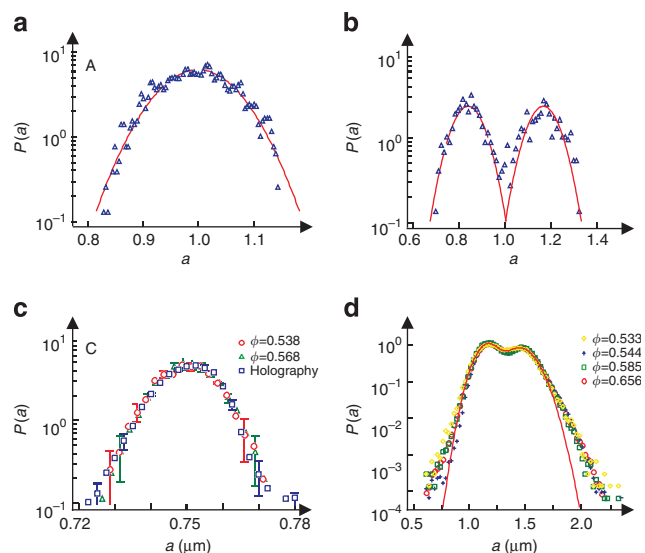
Figure 4 | Particle size distributions. (**a**) The distribution of the true radius (solid curve) and the estimated radius $a_i^{(10)}$ (triangles) for a simulated nominally monodisperse sample with $\phi = 0.51$ and polydispersity $p = 0.07$. (**b**) The distribution of the true radius (solid curve) and the estimated radius $a_i^{(10)}$ (triangles) for a simulated nominally binary sample with $\phi = 0.51$; see text for more details. (**c**) The radii distributions for a nominally monodisperse experimental suspension, for two volume fractions as indicated. The blue square symbols are a separate determination of the particle size distribution from holographic microscopy. Representative error bars are shown for $P(a)$; see text for discussion. (**d**) The radii distributions for the nominally binary experimental suspension from ref. 11, for four volume fractions as indicated. Here, the solid curve is a fit to the sum of two Gaussians. With the size ratio (1:1.3) and the polydispersity of each species ($p = 0.049$ for the small species and $p = 0.050$ for the large species), the two subdistributions have substantial overlap.

defined before, and we round Z_i to the nearest integer. The denominator 13 is chosen as the number of neighbours in a close-packed sample, and the numerator 7 is from the results of Fig. 2a.

Furthermore, we need one more improvement when we apply our method to a nominally binary sample. It usually happens that we know the mean radii a_1 and a_2 of the two subspecies, while the number ratio of two species is unknown, which means that $\bar{a}$ is unknown. In this situation, we start with a reasonable guess for $\bar{a}'$ to be used in equation 3. Then we compute the particle radii and obtain the double peak distribution, which depends on our guess $\bar{a}'$. The two peaks of the distribution will be at radii a_1 and a_2: these will differ from the true values. However, the difference will be the same for both, that is, $a_1' - a_1 = a_2' - a_2 = \bar{a}' - \bar{a}$. Adding this difference onto all of the data corrects the distributions and also identifies the true $\bar{a}$.

Application to experimental data. We now apply our method to experimental data taken on fluorescent silica particles. The particles have a mean radius $a = 0.752\,\mu m$ and polydispersity $p = 0.015$, both determined by holographic microscopy[37,38]. Holographic microscopy is a powerful technique, which can measure individual particle sizes to high accuracy (uncertainty $\Delta a/a \lesssim 0.005$). Using techniques similar to ref. 38, we determine the sizes of more than 8,000 particles, (see Methods). From these images, we determine the radius of each particle, and show the distribution of radii in Fig. 4c (blue squares). We additionally take 3D confocal microscopy movies of two high-volume fraction samples composed of the same silica particles and apply our method to determine the particle sizes

from more than 4,000 particles per sample. These results are plotted in Fig. 4c (circles and triangles). To estimate the uncertainty of $P(a)$, we take our list of a data and add additional Gaussian-distributed random noise to each value, with width corresponding to the uncertainty of a based on the technique (holography or our estimation technique). We then compute the new $P(a)$: repeating this process 10,000 times gives us fluctuations for $P(a)$ for each value of a, shown in the error bars for the three distributions plotted in Fig. 4c. The quantitative agreement between these results is strong evidence that our method works. Note that holographic microscopy works best with dilute samples, whereas our analysis works best with dense suspensions.

Separate evidence that our method works is shown in Fig. 4d. Here, we analyse previously published experimental data from ref. 11 (a nominally binary sample, composed of poly-methyl-methacrylate particles). Data from several different volume fractions are shown. Each different volume fraction was a sample taken from the same stock jar and therefore should have the same size distribution, as we observe, confirming the results are robust over a range of volume fractions.

A strength of our algorithm is that it only depends on data from the centres of particles. It therefore will work well with core-shell particles (dyed cores and transparent shells), for which the full size of the particle is not optically apparent. It would also work with particles with inhomogeneous dye distributions, so long as their centres could be identified. However, one limitation is that homogeneously dyed particles that are close together have overlapping images that pulls their apparent centres closer together[39]. This may somewhat average out for a random close-packed sample (with neighbours on all sides), but for a gel sample it would be more problematic. Measures can be taken to avoid such errors[39–41], and using core-shell particles minimizes the problems. In addition, time averaging allows particles to explore different environments and these errors should cancel out over time for samples other than gels and close-packed samples.

A second limitation is that systematic and correlated errors can exist in the particle positions. For example, if the calibration (microns/pixel) is slightly incorrect for the z axis, then particles separated in z will have systematic errors in their separation. To mimic an incorrect z calibration, we reconsider the $p = 0.06$, $\phi = 0.51$ data and multiply all the z coordinates of the simulation data by 0.98 before we do the radius-estimation algorithm. We find that Δa only slightly increases (from 0.013 to 0.015). An additional concern is that owing to optical limitations, particle positions are inherently more uncertain in z than in x and y (refs 2,6,7,10). For time averaging to be helpful in solving these systematic problem, it would require particles to rearrange significantly during the observation period. Such time averaging is impossible or difficult for random close-packed samples, gels and glassy samples, and so such samples would be more prone to these types of systematic errors. On the other hand, having neighbours distributed in a spherically symmetric fashion around the central particle will help average out the errors to some extent.

A third limitation is that charged particles with a polydispersity of charge would have their apparent radii influenced by both their core size and their charge. Probably their estimated radii would depend on a combination of these effects. For example, the estimated radius might relate to the distance where the pairwise particle interaction energy is of magnitude k_BT (ref. 42).

Discussion

We now demonstrate the utility of our algorithm by studying colloidal crystal nucleation. The nucleation of crystals in a dense particle suspension depends sensitively on polydispersity[23,27,28]. We examine data of the $\phi = 0.46$ sample from ref. 9, analysed at longer times to examine the crystallization process that was discarded from the
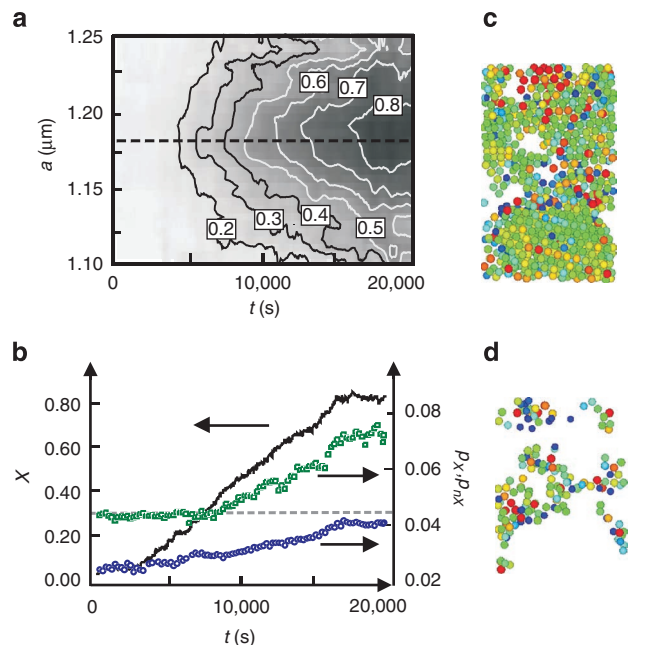
Figure 5 | Relationship between individual particle size a_i and crystallization. (**a**) The contour plot of the fraction of crystalline particles X as a function of a_i and time t. The contour lines are numbered by the value of X. The dashed line shows the mean radius $\bar{a}$. Particles with radii close to $\bar{a}$ crystallize faster. (**b**) The fraction of the sample X that is crystalline as a function of time (solid black curve), along with the mean polydispersity of all crystalline particles (blue circles) and all non-crystalline particles (green squares). The grey dashed line corresponds to the bulk polydispersity $p = 0.045$. (**c**) The crystalline particles are shown at $t = 20,000$ s. (**d**) The non-crystalline particles are shown at $t = 20,000$ s. For (**c,d**), the colour indicates a_i for each particle, where blue corresponds to smaller a_i, light green corresponds to a_i close to $\bar{a}$ and red corresponds to larger a_i.

analysis in ref. 9. These particles are slightly charged, shifting the freezing point to $\phi_{freeze} \approx 0.38$ and the melting point to $\phi_{melt} \approx 0.42$ (ref. 13). In this data, we confirm that the crystal nucleus appears at the centre of our microscopic image: this is homogeneous nucleation, not heterogeneous nucleation near the wall.

At each time step, we calculate the number of ordered neighbours N_o for each particle using standard techniques[13,43] (see Methods). By convention, a crystalline particle has $N_o \geq 8$. At each time, we compute the number fraction of the sample that is crystallized, $X(t)$. Figure 5a shows $X(t)$ as a function of both individual particle size and time, where darker colours correspond to larger values of $X(t)$. Below $t = 3,000$ s, $X < 0.2$ for all a, and essentially all crystal clusters are below the critical size (~ 100 particles)[13]. At $t = 3,000$ s, a sufficiently large crystalline region appears and begins to grow. X increases first for particles with a close to the mean radius, and these particles continue to be the subpopulation that is most crystallized at any given time. At longer times, the particles with a farther from $\bar{a}$ gradually begin to crystallize.

We next consider an alternate way of thinking about the same data. Figure 5b shows the relationship between the sample-averaged $X(t)$ (solid black curve), the polydispersity p_X for all crystalline particles (blue circles), and the polydispersity p_{nX} for all non-crystalline particles (green squares). X starts to increase at $t = 3,000$ s, and those particles that are crystalline at that time have $p_X \sim 0.03$, smaller than the bulk polydispersity $p = 0.045$. As the sample crystallizes, we observe that both p_X and p_{nX} increase. The growth of p_X indicates that the crystal, while nucleating in a fairly monodisperse

region, can grow by incorporating particles that are farther from the mean size. In the final state, the local polydispersity of the crystalline particles has nearly reached the mean polydispersity p. The growth of p_{nX} indicates that those particles that are still outside the crystal are more likely to be those with unusual sizes.

The spatial distribution of particles at the end of the experiment is shown in Fig. 5c,d. Figure 5c shows the locations of the crystalline particles, while D shows the locations of the non-crystalline particles. Green particles have a_i close to $\bar{a}$, and preferentially occupy the cores of the crystal regions.

Next, we examine the beginning of the crystal nucleation process. While many particles are close to the mean size, only a few end up being the nucleation site. To understand which ones nucleate, we now focus on the particles close to the mean size: radii $1.175\,\mu m < a_i < 1.185\,\mu m$. Among those particles, we define the nucleus particles as those that are crystalline particles at $t = 5,000\,s$, the remainder are non-nucleus particles. We next define the local polydispersity $p_i(r)$ of particle i as

$$p_i = \sqrt{\langle (a_j - a_i)^2 \rangle}/a_i \qquad (5)$$

where the angle brackets $\langle \rangle$ indicate an average over all particles j with centres within a distance r from particle i. Figure 6a,b show space-time plots of the mean value of $p_i(r)$ for the nucleus particles (Fig. 6a) and the non-nucleus particles (Fig. 6b). In all cases, $p_i(r)$ is lower close to the particles and increases with increasing r. However, notably the contours for low p_i are at larger values of r for the nucleus particles (Fig. 6a): the region of low polydispersity is larger for these particles. After $t \approx 3,000\,s$, the region of low p_i slightly spreads to even larger values of r for the nucleus particles (Fig. 6a), while little change is seen for the non-nucleus particles (Fig. 6b). Simultaneously, we show the temporal change of N_o in Fig. 6c for the nucleus particles (solid black curve) and the non-nucleus particles (dashed grey curve). This confirms that the onset of crystallization at $t \approx 3,000\,s$ coincides with the expansion of the low local polydispersity region seen in Fig. 6a. This is all evidence that crystal nuclei are formed from regions where the particles are all similar sizes. A conjecture is that nucleation rates are possibly sensitive to how well mixed the sample initially is: poorly mixed samples with regions of lower local polydispersity will crystallize faster.

The particles in this experiment are slightly charged and potentially polydisperse in their charge. We therefore note that the effects we see are based on the estimated radius, which likely includes both the core size and the effect of charge: our results show that this estimated radius has a role in nucleation. We cannot disentangle the two effects, but presumably our estimated radius—which relates to how closely particles approach each other—is similar to whatever effective radius is relevant for nucleation.

We have developed a general method to estimate the particle sizes in a dense particulate samples where the particle positions are known. This method can be applied to any cases where 3D particle positions can be found; while we have focused on colloidal samples, granular media are quite similar[1,44]. We have demonstrated the utility of our method by examining homogeneous colloidal crystal nucleation. While it has been known that nucleation is faster for more monodisperse samples, we find this is true on a quite local scale. Nucleation happens in regions that are locally more monodisperse, and crystal growth proceeds by preferentially incorporating particles close to the mean size.

Methods

Simulations. We simulate four particle suspension systems, which are RCP, colloidal gel, single-component suspension and a binary system. The polydisperse RCP sample is generated using the algorithm of ref. 45. For the three other cases, we perform 3D Monte Carlo simulations with hard spheres. Additionally for gels, we wish to model colloid-polymer mixtures and so we use the
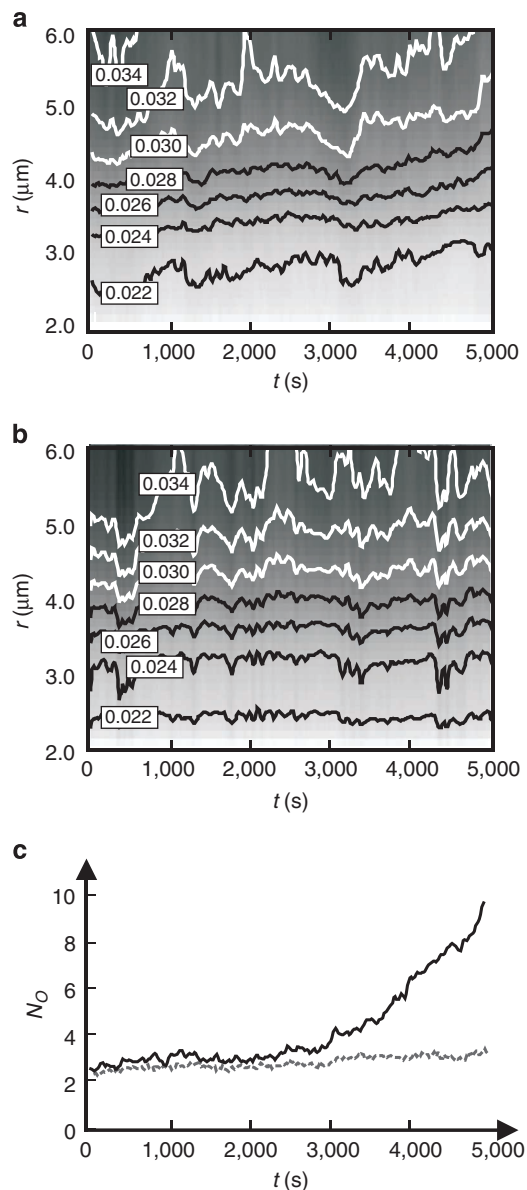
Figure 6 | Relationship between nucleation properties and the local polydispersity $p_i(r)$. (**a**) The contour plot of the mean $p_i(r)$ as a function of r and t, averaged over all particles that are crystalline at $t = 5,000\,s$. The numbers on the contour lines represent the value of $p_i(r)$. (**b**) The contour plot of the mean $p_i(r)$ as a function of r and t, averaged over all particles that are non-crystalline at $t = 5,000\,s$. The particles considered in (**a**) and (**b**) are only those with radii close to the mean radii; see text for details. (**c**) The number of ordered neighbours N_o as a function of time for those particles plotted in (**a**) (nucleus particles, solid curve) and (**b**) (non-nucleus particles, dashed curve).

Asakura and Oosawa model[36]. This model leads to a pair interaction between two hard colloidal spheres in a solution of ideal polymers as $U(r) = \infty$ for $r < \sigma_{ij}$, $U(r) = -(\pi/12)k_B T \rho_p [r^3 - 3(\sigma_{ij} + R_G)^2 r + 2(\sigma_{ij} + R_G)^3]$ for $\sigma_{ij} \leq r < \sigma_{ij} + 2R_G$, $U(r) = 0$ for $r \geq \sigma_{ij} + 2R_G$, where $\sigma_{ij} = (\sigma_i + \sigma_j)/2$, σ_i is the diameter of particle i, k_B is the Boltzmann constant, T is temperature, ρ_p is the number density of polymers and R_G is the polymer radius of gyration. We fix $R_G = 0.1\bar{\sigma}$ and $\phi_p = 4\pi/3 R_G^3 \rho_p = 0.1$ where $\bar{\sigma}$ is the mean diameter of the hard spheres. For our single-component and two-component hard-sphere suspensions, particles interact via $U(r) = \infty$ for $r < \sigma_{ij}$, otherwise $U(r) = 0$. We use 1,024 particles with the mean radius $\bar{a} = 1$ and variable polydispersity for all simulations.

Experiments. The silica particles are fluorescently dyed and are in a refractive index-matched mixture of water and DMSO (dimethyl sulphoxide, $(CH_3)_2SO$).

The silica particles are imaged with a Leica confocal microscope, and their positions are tracked using standard techniques[10,35]. For the holographic microscopy of the silica particles, we illuminate the sample with a 445-nm Coherent Cube diode laser mounted above a Nikon Eclipse TE2000-U inverted microscope. The scattered light is collected by a ×1001.4 numerical aperture oil immersion objective (Nikon Plan Apo) and is relayed to a grayscale video camera (NEC TI-324A II) that records at 30 frames per second. Diluted samples are flowed through a glass microfluidic cell and each particle is imaged at least $N = 10$ times. The s.d. of an individual particle's measured radius a from different images is 0.012 μm. The particle's radius is estimated from the average a over the N images that the particle was seen in, and the uncertainty of that estimation is the s.d. divided by $\sqrt{N}$. The resulting uncertainty of the particle radius for an individual particle, for $N \geq 10$, is $\Delta a/a \lesssim 0.005$.

The poly(methyl methacrylate) (PMMA) data come from a prior experiment[11]. This experiment used sterically stabilized PMMA particles that were imaged with a VisiTech confocal microscope and tracked using standard methods[10,35]. Each data set has more than 2,800 particles. A more detailed experimental discussion is in ref. 11.

The volume fraction numbers reported in our paper are given to two or three significant digits. The volume fractions are measured by direct counting of particles in a given volume[21], and converting from the measured number density n to volume fraction ϕ using $\phi = nv_P$ using the particle volume $v_P = (4/3)\pi a^3$. The particle radii were determined by holographic microscopy (for the silica particles) or by measuring the hydrodynamic radii from dilute suspensions (for the prior PMMA data). The uncertainty of ϕ with these methods is about 5%, that is, a number such as $\phi = 0.40$ is uncertain by about $0.40 \times 0.05 = \pm 0.02$ (ref. 21). Note that this uncertainty is systematic, so for example the data Fig. 4d are correct relative to each other to within the third decimal place, but overall have a systematic uncertainty of ±5%. For example, they may all be 5% larger than the reported values. The same is true for the data of Fig. 4c, which come from different experiments and so have a different systematic error from those of Fig. 4d.

Order parameters for crystalline particles. We use previously developed order parameters to look for crystalline particles and ordered structure[13,43,46]. For each particle i, we find its nearest neighbours j and identify unit vectors $\hat{r}_{ij}$ pointing to the neighbours. We then define a complex order parameter $\hat{q}_{lm}$ using $q_{lm}(i) = \sum_{j=1}^{c_i} Y_{lm}(\hat{r}_{ij})$ where c_i is the number of nearest neighbours of particle i and Y_{lm} is a spherical harmonic function; we normalize this as $\hat{q}_{lm} = q_{lm}/N$ where N is a normalization factor such that $\sum_m \hat{q}_{lm}(i)\hat{q}_{lm}^*(i) = 1$ (ref. 13). We use $l = 6$. For each particle pair, we compute the complex inner product $d_6 = \sum_m \hat{q}_{lm}(i)\hat{q}_{lm}^*(j)$. Two neighbouring particles are termed 'ordered neighbors' if d_6 exceeds a threshold value of 0.5. For each particle, we focus on N_o, the number of ordered neighbours it has at a given time. N_o^i measures the amount of similarity of structure around neighbouring particles. $N_o^i = 0$ corresponds to random structure around particle i, while a large value of N_o^i means that particle i and its neighbour particles have similar surroundings[43].

References

1. Dijksman, J. A., Rietz, F., Lörincz, K. A., van Hecke, M. & Losert, W. Invited article: refractive index matched scanning of dense granular materials. *Rev. Sci. Inst.* **83,** 011301 (2012).
2. Prasad, V., Semwogerere, D. & Weeks, E. R. Confocal microscopy of colloids. *J. Phys.: Cond. Matt.* **19,** 113102 (2007).
3. Pusey, P. N. & van Megen, W. Phase behaviour of concentrated suspensions of nearly hard colloidal spheres. *Nature* **320,** 340–342 (1986).
4. Anderson, V. J. & Lekkerkerker, H. N. Insights into phase transition kinetics from colloid science. *Nature* **416,** 811–815 (2002).
5. Vincent, B. *Introduction to Colloidal Dispersions* (Blackwell Publishing, 2005).
6. van Blaaderen, A., Imhof, A., Hage, W. & Vrij, A. Three-dimensional imaging of submicrometer colloidal particles in concentrated suspensions using confocal scanning laser microscopy. *Langmuir* **8,** 1514–1517 (1992).
7. van Blaaderen, A. & Wiltzius, P. Real-space structure of colloidal hard-sphere glasses. *Science* **270,** 1177–1179 (1995).
8. Kegel, W. K. & van Blaaderen, A. Direct observation of dynamical heterogeneities in colloidal hard-sphere suspensions. *Science* **287,** 290–293 (2000).
9. Weeks, E. R. & Weitz, D. A. Properties of cage rearrangements observed near the colloidal glass transition. *Phys. Rev. Lett.* **89,** 095704 (2002).
10. Dinsmore, A. D., Weeks, E. R., Prasad, V., Levitt, A. C. & Weitz, D. A. Three-dimensional confocal microscopy of colloids. *App. Optics* **40,** 4152–4159 (2001).
11. Narumi, T., Franklin, S. V., Desmond, K. W., Tokuyama, M. & Weeks, E. R. Spatial and temporal dynamical heterogeneities approaching the binary colloidal glass transition. *Soft Matter* **7,** 1472–1482 (2011).
12. van Blaaderen, A., Ruel, R. & Wiltzius, P. Template-directed colloidal crystallization. *Nature* **385,** 321–324 (1997).
13. Gasser, U., Weeks, E. R., Schofield, A., Pusey, P. N. & Weitz, D. A. Real-space imaging of nucleation and growth in colloidal crystallization. *Science* **292,** 258–262 (2001).
14. Dullens, R. P. A., Aarts, D.G.A.L. & Kegel,, W. K. Dynamic broadening of the crystal-fluid interface of colloidal hard spheres. *Phys. Rev. Lett.* **97,** 228301 (2006).
15. Dinsmore, A. D. & Weitz, D. A. Direct imaging of three-dimensional structure and topology of colloidal gels. *J. Phys.: Cond. Matt.* **14,** 7581–7597 (2002).
16. Dibble, C. J., Kogan, M. & Solomon, M. J. Structure and dynamics of colloidal depletion gels: coincidence of transitions and heterogeneity. *Phys. Rev. E* **74,** 041403 (2006).
17. Gao, Y. & Kilfoil, M. L. Direct imaging of dynamical heterogeneities near the colloid-gel transition. *Phys. Rev. Lett.* **99,** 078301 (2007).
18. Aarts, D.G.A.L., Schmidt, M. & Lekkerkerker, H. N. W. Direct visual observation of thermal capillary waves. *Science* **304,** 847–850 (2004).
19. Hernández-Guzmán, J. & Weeks, E. R. The equilibrium intrinsic crystal-liquid interface of colloids. *Proc. Natl Acad. Sci. USA* **106,** 15198–15202 (2009).
20. Royall, C. P., Vermolen, E. C. M., van Blaaderen, A. & Tanaka, H. Controlling competition between crystallization and glass formation in binary colloids with an external field. *J. Phys.: Cond. Matt.* **20,** 404225 (2008).
21. Poon, W. C. K., Weeks, E. R. & Royall, C. P. On measuring colloidal volume fractions. *Soft Matter* **8,** 21–30 (2012).
22. Auer, S. & Frenkel, D. Prediction of absolute crystal-nucleation rate in hard-sphere colloids. *Nature* **409,** 1020–1023 (2001).
23. Pusey, P. N. *et al.* Hard spheres: crystallization and glass formation. *Phil. Trans. Roy. Soc. A* **367,** 4993–5011 (2009).
24. Sollich, P. & Wilding, N. B. Crystalline phases of polydisperse spheres. *Phys. Rev. Lett.* **104,** 118302 (2010).
25. Kawasaki, T., Araki, T. & Tanaka, H. Correlation between dynamic heterogeneity and medium-range order in two-dimensional glass-forming liquids. *Phys. Rev. Lett.* **99,** 215701 (2007).
26. Kurita, R. & Weeks, E. R. Glass transition of two-dimensional binary soft-disk mixtures with large size ratios. *Phys. Rev. E* **82,** 041402 (2010).
27. Schöpe, H. J., Bryant, G. & van Megen, W. Effect of polydispersity on the crystallization kinetics of suspensions of colloidal hard spheres when approaching the glass transition. *J. Chem. Phys.* **127,** 084505 (2007).
28. Henderson, S. I. & van Megen, W. Metastability and crystallization in suspensions of mixtures of hard spheres. *Phys. Rev. Lett.* **80,** 877–880 (1998).
29. D'Haene, P. & Mewis, J. Rheological characterization of bimodal colloidal dispersions. *Rheologica Acta* **33,** 165–174 (1994).
30. Pond, M. J., Errington, J. R. & Truskett, T. M. Implications of the effective one-component analysis of pair correlations in colloidal fluids with polydispersity. *J. Chem. Phys.* **135,** 124513 (2011).
31. Berthier, L., Chaudhuri, P., Coulais, C., Dauchot, O. & Sollich, P. Suppressed compressibility at large scale in jammed packings of size-disperse spheres. *Phys. Rev. Lett.* **106,** 120601 (2011).
32. Zachary, C. E., Jiao, Y. & Torquato, S. Hyperuniform long-range correlations are a signature of disordered jammed hard-particle packings. *Phys. Rev. Lett.* **106,** 178001 (2011).
33. Kurita, R. & Weeks, E. R. Incompressibility of polydisperse random-close-packed colloidal particles. *Phys. Rev. E* **84,** 030401(R) (2011).
34. Brujić, J. *et al.* 3D bulk measurements of the force distribution in a compressed emulsion system. *Faraday Disc.* **123,** 207–220 (2003).
35. Crocker, J. C. & Grier, D. G. Methods of digital video microscopy for colloidal studies. *J. Colloid Interf. Sci.* **179,** 298–310 (1996).
36. Asakura, S. & Oosawa, F. Surface tension of high-polymer solutions. *J. Chem. Phys.* **22,** 1255–1256 (1954).
37. Lee, S.- H. *et al.* Characterizing and tracking single colloidal particles with video holographic microscopy. *Opt. Express* **15,** 18275–18282 (2007).
38. Cheong, F. C. *et al.* Flow visualization and flow cytometry with holographic video microscopy. *Opt. Express* **17,** 13071–13079 (2009).
39. Baumgartl, J. & Bechinger, C. On the limits of digital video microscopy. *Europhys. Lett.* **71,** 487–493 (2005).
40. Jenkins, M. C. & Egelhaaf, S. U. Confocal microscopy of colloidal particles: towards reliable, optimum coordinates. *Adv. Coll. Int. Sci.* **136,** 65–92 (2008).
41. Ueberschär, O., Wagner, C., Stangner, T., Gutsche, C. & Kremer, F. A novel video-based microsphere localization algorithm for low contrast silica particles under white light illumination. *Opt. Lasers Eng.* **50,** 423–439 (2012).
42. Royall, C. P., Poon, W. C. K. & Weeks, E. R. In search of colloidal hard spheres Preprint at arXiv: 1205.6665 (2012).
43. Rein ten Wolde, P., Ruiz-Montero, M. J. & Frenkel, D. Numerical calculation of the rate of crystal nucleation in a Lennard-Jones system at moderate undercooling. *J. Chem. Phys.* **104,** 9932–9947 (1996).
44. Slotterback, S., Toiya, M., Goff, L., Douglas, J. F. & Losert, W. Correlation between particle motion and voronoi-cell-shape fluctuations during the compaction of granular matter. *Phys. Rev. Lett.* **101,** 258001 (2008).
45. Xu, N., Blawzdziewicz, J. & O'Hern, C. S. Random close packing revisited: ways to pack frictionless disks. *Phys. Rev. E* **71,** 061306 (2005).
46. Steinhardt, P. J., Nelson, D. R. & Ronchetti, M. Bond-orientational order in liquids and glasses. *Phys. Rev. B* **28,** 784–805 (1983).

7

Authors contributions

R.K. conceived the idea, performed numerical simulations and analysed the data for crystallization. D.B.R. performed the holography experiment. E.R.W. took the new confocal microscopy data and supervised the work. R.K. and E.R.W. wrote the paper.

Additional information

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Kurita R. *et al.* Measuring the size of individual particles from three-dimensional imaging experiments. *Nat. Commun.* 3:1127 doi: 10.1038/ncomms2114 (2012).