



Cite this: *Soft Matter*, 2025, 21, 8100

Received 6th August 2025,
Accepted 2nd October 2025

DOI: 10.1039/d5sm00776c

rsc.li/soft-matter-journal

The missing billions in hard sphere nucleation

Sahana Kale,^a Achim Lederer^b and Hans Joachim Schöpe^{id} *^a

The crystallisation of a metastable liquid is an everyday phenomenon, yet it still presents a number of puzzles. One such puzzle is the discrepancy between the crystallisation rate observed in experiments and that predicted by theory: the experimental and simulated rate densities for hard spheres – the “simplest” system showing a first-order freezing transition – disagree by up to 22 orders of magnitude. Nevertheless, it is precisely the utilisation of elementary model systems that facilitates the resolution of these enigmas. We present a comprehensive experimental investigation into the crystallisation of colloidal hard spheres at the particle level. Our ground breaking findings challenge the prevailing conceptualisation of crystal nucleation, elucidate the discrepancy between experiment and theory, and propose an alternative description.

Introduction

Crystallisation of a metastable melt is one of the most important non-equilibrium phenomena in statistical physics, underpinning numerous scientific phenomena. An overarching understanding of the crystallization process is not only of interest for fundamental research,^{1,2} but also plays a key role in the systematic development of new materials with tailored properties.^{3,4} Within the field, crystal nucleation is of particular importance, as it represents the “birth” of the new stable phase or new material. The prevalent picture is based on classical nucleation theory (CNT),^{5–7} whereby crystals form spontaneously in an undercooled but thermodynamic equilibrated liquid due to equilibrium-like fluctuations. If crystals exceed a critical size they grow and gain energy, if they are too small they dissolve. Nevertheless, CNT is seldom corroborated by experimental data, and this is the reason why it has been the subject of controversy for decades.^{8–12} Recent studies have revealed discrepancies of several orders of magnitude between theoretical and experimental results in elementary atomic liquids such as argon.¹³ Consequently, the precise mechanisms underlying the formation of crystals remain to be elucidated.

In this context a system of hard spheres (HS) plays a special role: it is the most basic system that exhibits a first order freezing transition.¹⁴ Due to its simplicity, it is regarded as “the working horse” for classical many body physics both in experiment and in theory. The phase behaviour of the monodisperse HS system is solely determined by the particle volume fraction

Φ . Freezing occurs at $\Phi_F = 0.494$ and melting at $\Phi_M = 0.545$. However, size polydispersity (spd) shifts the freezing transition to higher Φ and narrows the coexistence region.¹⁵ Above Φ_F , a shear-melted system is initially a fluid out of equilibrium, which subsequently crystallizes. Strictly speaking, this metastable fluid is over-packed – nevertheless, for the sake of simplicity, the term “metastable” is used in the following.

A series of experiments^{16–25} and simulations^{26–35} were conducted in order to determine the key parameters that characterize HS nucleation. The findings from these studies have both similarities and differences, and the results are not entirely consistent. In particular a direct comparison of the nucleation rate density (NRD) as function of metastability shows a highly unsatisfactory result: the shape of the curves from experiment and simulation are qualitatively different and the data diverge at $\Phi \approx 0.52$ by 22 orders of magnitude!

The fundamental reason for these significant discrepancies remains unknown. A summary of the possible reasons speculated thus far can be found in a recently published review article.³⁶ These include heterogeneous nucleation on aggregates or dust, deviation from the ideal HS interaction, sedimentation effects, hydrodynamic interaction, crystal twinning, and data analysis of light scattering (LS) methods.^{29,34,37–39} Nevertheless, the origin of the discrepancy remains elusive. In order to make a substantial contribution to this issue, we have conducted a comprehensive investigation of the nucleation process in colloidal HS in direct space using laser-scanning confocal microscopy (LSCM). A prerequisite for conducting such a study is the availability of a colloidal HS model system that can be measured using LSCM. This necessitates fluorescent particles exhibiting HS interaction, a dispersion medium with identical refractive index and mass density as the particles, and samples that are stable over time. Previous studies

^a Institute for Applied Physics, Eberhard Karls University Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany.
E-mail: hans-joachim.schoepe@uni-tuebingen.de

^b Retsch Technology GmbH, Retsch-Allee 1-5, 42781 Haan, Germany



investigating crystallisation in colloidal model systems using LSCM have unfortunately not been able to meet these requirements.^{40,41} We have recently presented a colloidal model system that fulfils all these criteria.⁴²

Materials and methods

For our study we use sterically stabilized and fluorescent PMMA particles dispersed in mixture of *cis*-decalin (CDL) and tetrachloroethylene (TCE). The gravity matched HS-system was characterized comprehensively to allow for meaningful comparison with other experiments and simulations.⁴² The particle size is $2R_{\text{EM}} = 1.388 \pm 0.002 \mu\text{m}$ in diameter with a size polydispersity of $\text{spd} = 5.75 \pm 0.06\%$ determined by electron microscopy. By analysing the radial distribution function in the equilibrium fluid state using a Verlet–Weiss–corrected Percus Yevick equation an effective HS diameter of $2R_{\text{VW}} = 1.452 \pm 0.010 \mu\text{m}$ was obtained. The polydispersity of the system is large enough to delay the onset of the crystallization process offering the possibility to perform measurements with good time resolution and it is small enough that no fractionation processes occur which may stabilize the fluid against crystallization. Fractionation typically occurs above $\text{spd} \approx 6\%$, whereby the exact threshold value depends on the shape of the particle size distribution (PSD).⁴³ The system under investigation exhibited only a minor skew in the PSD. The particles are sterically stabilized by chemically grafting poly-hydroxystearic-acid onto their surface and are homogeneously stained with fluorescent dye 1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate (DiIc18). They are dispersed in a mixture of CDL and TCE, which matches the mass density exactly and the refractive index n closely ($\Delta n = 0.01$).

In conducting LSCM measurements, we employed custom-built sample cells with screw caps that facilitate straightforward and highly accurate volume fraction adjustments, with a typical capacity of approximately 1 mL suspension. Heterogeneous nucleation on the container walls was eliminated by coating all walls with larger PMMA particles ($2R = 2.33 \mu\text{m}$).^{44,45} All samples were shear-molten by tumbling for several hours at a frequency of ≈ 1 Hz before studying the crystallization kinetics. The crystallization process was monitored using a white light LSCM (Leica TCS-SP8). Twenty-five volumes of $82 \times 82 \times 60 \mu\text{m}^3$ were observed in the lateral centre of the cell, situated approximately $50 \mu\text{m}$ from the bottom cell wall. The scanned volume contains $\sim 3 \times 10^6$ particles. The voxel size is $\sim 80 \times 80 \times 130 \text{ nm}^3$. The requisite time to scan such a volume is ~ 50 s. A minimum of two measurements were conducted for each packing fraction. The duration of these measurements varied depending on the volume fraction with the shortest period being a few hours ($\Phi \approx \Phi_{\text{M}}$) and the longest being three weeks ($\Phi \approx (\Phi_{\text{M}} + \Phi_{\text{F}})/2$). The observed freezing and melting volume fractions are $\Phi_{\text{F}} = 0.5075 \pm 0.0013$ and $\Phi_{\text{M}} = 0.5490 \pm 0.0017$, respectively. In the SI, the comprehensive methodology employed to determine the volume fraction is described in detail. It ensures minimal systematic and static errors.

The normalised width of the coexistence region $\text{MS} = (\Phi - \Phi_{\text{F}})/(\Phi_{\text{M}} - \Phi_{\text{F}})$ is employed as an experimental measure of the chemical potential difference (metastability MS). This procedure enables the comparison of datasets with different Φ_{F} and Φ_{M} due to different size polydispersity or different experimental procedures in determining Φ . Given that the process of crystallisation can be observed in LSCM measurements at the level of individual particles, any instances of heterogeneous nucleation can be identified unambiguously. Our experiments have not revealed evidence of heterogeneous nucleation on dust, aggregates or walls.

The particle coordinates are determined using a self-written IDL routine, which is based on the main concepts of the algorithm developed by Jenkins.⁴⁶ The position uncertainty is about 5% of the particle diameter. To identify crystalline clusters and to analyse their structure we use local bond order parameters.⁴⁷ In our study a particle is identified as crystalline if it has at least 8 nearest neighbours in a distance smaller than $1.4 \times 2R_{\text{HS}}$ and the $q_{6\text{m}}(i)q_{6\text{m}}^*(j)$ scalar product is larger than 0.5. Four or more connected crystalline particles are identified as a crystalline cluster. Different structures (hcp, fcc, bcc, non-registered hexagonal layers, liquid) can be identified with the help of averaged bond order parameters aq_4 , aq_6 , aw_4 and aw_6 . More detailed information on the experimental procedure can be found in the SI.

Results and discussion

Ensemble averaged crystallisation kinetics

In order to gain insight into the crystallisation process, it is beneficial to initially ascertain ensemble-averaged variables that are analogous to the LS-experiments conducted in the past. For this purpose, all crystalline clusters are identified using the $q_{6\text{m}}(i)q_{6\text{m}}^*(j)$ scalar product. The crystallinity X is defined by the relative amount of crystalline particles $X = N_{\text{C}}/N_{\text{all}}$, where N_{C} denotes all particles in crystalline clusters and N_{all} all particles in the observed volume. The averaged crystal size is obtained by calculating the mean of all cluster sizes. The crystal number density ρ_{C} is calculated by normalizing the crystallinity with the averaged crystal volume $\rho_{\text{C}} = X/\langle V_{\text{C}} \rangle$. Fig. 1 shows the temporal evolution of the crystallinity, the crystal size and the crystal number density for a metastability of $\text{MS} = 0.75$. Global induction times t_{ind} , growth rates, nucleation rate density $J(t) = \text{d}/\text{d}t \rho_{\text{C}}/(1 - X(t))$ can be determined and compared with direct measurements. In the example shown the induction time equates $t_{\text{ind}} = 395 \pm 25$ min and the time averaged nucleation rate density equates $\langle J \rangle = (6 \pm 3) \times 10^9 \text{ m}^{-3} \text{ s}^{-1}$. The bracket denotes the time average in the nucleation and growth state ($t > t_{\text{ind}}$).

Crystallisation kinetics by tracing overcritical nuclei

In order to ascertain the NRD with the greatest possible degree of accuracy, we monitor the formation of each crystal individually and determine the number of supercritical nuclei N_{grcr} that are growing to extended crystals as a function of time.



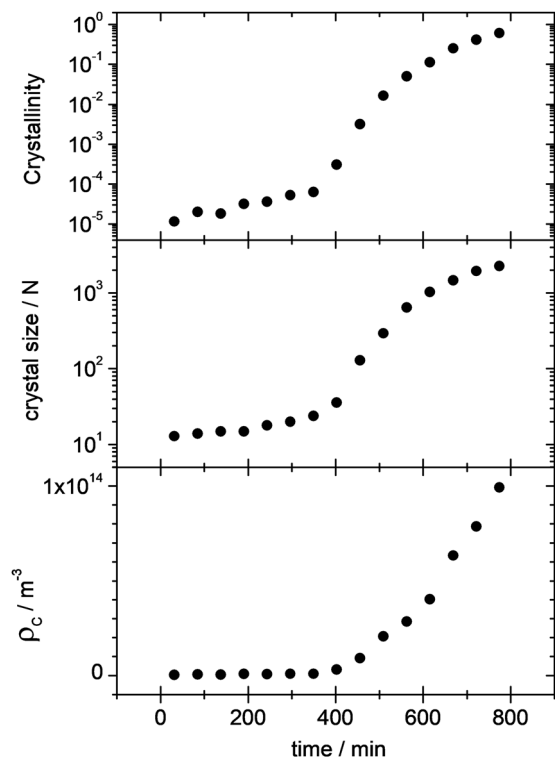


Fig. 1 Ensemble averaged crystallisation kinetics – from top to bottom: crystallinity, crystal size and crystal number density as function of elapsed time. Crystallisation sets in after ≈ 400 min. The metastability is $MS = 0.75$.

The tracing of the formation and growth of individual crystals provides direct access to a number of key parameters, including individual induction times, growth rates, critical radii and the nucleation rate. Fig. 2 presents ten out of 124 trajectories of growing crystals, selected at random. The individual critical size, induction times and growth rates display a wide range of values. By counting the number of growing crystalline clusters N_{grcr} as function of time the nucleation rate dN_{grcr}/dt can directly be measured (Fig. 2). The time averaged NRD is then given by $J = \langle 1/V(t) dN_{\text{grcr}}/dt \rangle$, with $V(t)$ the observation volume not occupied by crystals. The bracket denotes the time average in the nucleation and growth state. In the example the time

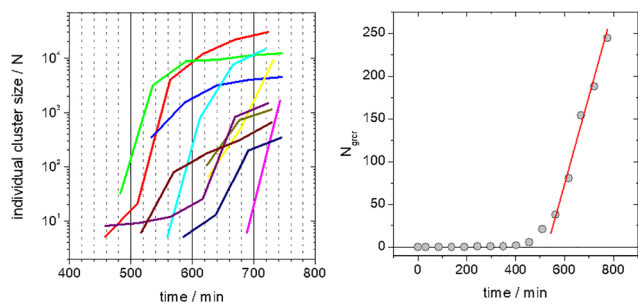


Fig. 2 Left: Ten randomly selected trajectories of growing crystals. Right: Number of overcritical growing crystals as function of elapsed time. The red line is a linear fit to the data. The slope represents the nucleation rate. The metastability is $MS = 0.75$.

averaged nucleation rate density equates $\langle J \rangle = (1.7 \pm 0.2) \times 10^9 \text{ m}^{-3} \text{ s}^{-1}$.

As previously stated in the introduction, the prevailing concept of crystal nucleation is founded upon CNT. Moreover, the majority of simulations conducted at low metastabilities are based on this theoretical framework. Consequently, we will initially present a concise overview of the fundamental principles of CNT.

CNT – a short review

CNT assumes that the crystal phase forms *via* the formation and subsequent growth of a crystal nucleus of spherical shape. The total Gibbs free energy cost to form a nucleus is

$$\Delta G = \frac{4}{3}\pi r_N^3 \rho_s \Delta\mu + 4\pi r_N^2 \gamma = \Delta\mu N + (6\sqrt{\pi}/\rho_s)^{2/3} N^{2/3} \gamma \quad (1)$$

where $\Delta\mu$ is the difference in the chemical potential between the metastable liquid and the thermodynamically stable crystal phase, γ is the surface free energy of the interface between solid and liquid, r_N the radius of the spherical nucleus which consists of N particles, and ρ_S is the number density of the solid. By assuming that the Boltzmann statistics is still valid out of equilibrium the probability that crystal nuclei appear in the metastable liquid is described by

$$P_{\text{CSD}}(N) \propto \exp(-\Delta G(N)/k_{\text{B}}T) \quad (2)$$

Only crystals larger than the critical size $r_{\text{crit}} = 2\gamma/(\rho_{\text{S}}|\Delta\mu|)$ overcome the nucleation barrier

$$\Delta G_{\text{crit}} = \frac{16\pi}{3} \frac{\gamma^3}{(\rho_s \Delta \mu)^2} \quad (3)$$

and grow continuously gaining energy. The crystal NRD is given by the product of the probability that a critical nucleus is formed and the kinetic prefactor J_0

$$J = J_0 \exp(-\Delta G_{\text{crit}}/k_{\text{B}}T) \quad (4)$$

J_0 is derived *via* rate equations, assuming single particle condensation and evaporation. It is finally expressed by the particle attachment rate K^+ ; the Zeldovich factor Z and the number density of the fluid

$$J_0 = K^+ \rho_{\text{liq}} Z = A \times \sqrt{\gamma} \times D_S^L \times \frac{1}{R^4} \times \Phi_s^{-1/3} \times \Phi_{\text{fl}}^{5/3}, \quad (5)$$

and can be rewritten for HS colloids²⁴ with an unknown scaling constant A , the particle radius R , the volume fraction of the solid and the fluid phase Φ_s and Φ_f .

The critical radius (or nucleation barrier) is the key parameter that exerts a decisive influence in this theory. Essentially the kinetic pre-factor serves to scale the absolute values of the data curve. In the following, we will determine the critical radius in four different ways and compare it with results obtained in the simulation. Firstly, we will do so within the framework of CNT. Secondly, we will present three additional approaches, each of which will determine r_{crit} directly from the experimental data.

Prior to presenting the procedure and the resulting data, we would like to make a remark. The critical nucleus size in CNT corresponds to a state in which crystals neither grow nor shrink. A direct data analysis, independent of theoretical models, is the most appropriate method of analysis for a crystallisation experiment, such as the one carried out here. This approach facilitates the determination of the critical radius. However, it should be noted that this critical radius can only be identified to a limited extent with the idealised definition. The data that were determined here in a direct approach therefore represent an upper estimate of the critical radius.

Temporal evolution of the crystal size distribution

In order to relate the experimental data to the central statements of CNT, it is necessary to determine and analyse the temporal evolution of the crystallite size distribution $f(N)$. Fig. 3 shows the crystallite size distribution (CSD) as function of elapsed time for three different metastabilities 0.62, 0.75 and 0.95. The upper row represents $f(N)$ in the induction phase prior to the onset of crystallisation in ($t < t_{\text{ind}}$) while the lower row shows $f(N)$ in the main crystallisation stage ($t > t_{\text{ind}}$) characterised by the nucleation and growth of individual crystals. The distributions measured in the induction phase

exhibits a strong decay commencing from several hundred crystalline clusters containing four particles. Clusters larger than ~ 30 particles are not observed. Additionally, a negligible waiting time dependence is observed. In the main crystallisation stage $f(N)$ shows a weaker decline that merges into a “tail” of larger crystals, which are overcritical and undergo growth over time. There is a significant temporal evolution of the CSD.

Upon examination of the temporal evolution of the CSD, it becomes immediately evident that it undergoes notable changes. The increase in the proportion of supercritical crystals is to be expected. However, the most striking observation is the substantial change in the distribution of subcritical nuclei over time. The data demonstrate the absence of a quasi-stationary state: the structural properties of the metastable melt undergo substantial changes following the initiation of crystallisation. The relative amount of larger subcritical clusters increases significantly with time (Fig. 4). Moreover, the transition from the induction period to the nucleation and growth regime can be identified by a marked increase in subcritical clusters. It is important to note that there is no change in the relevant parameters of the colloidal model system (*e.g.* number of particles in the observed volume, particle size, interaction, gravity match) during the measurements.

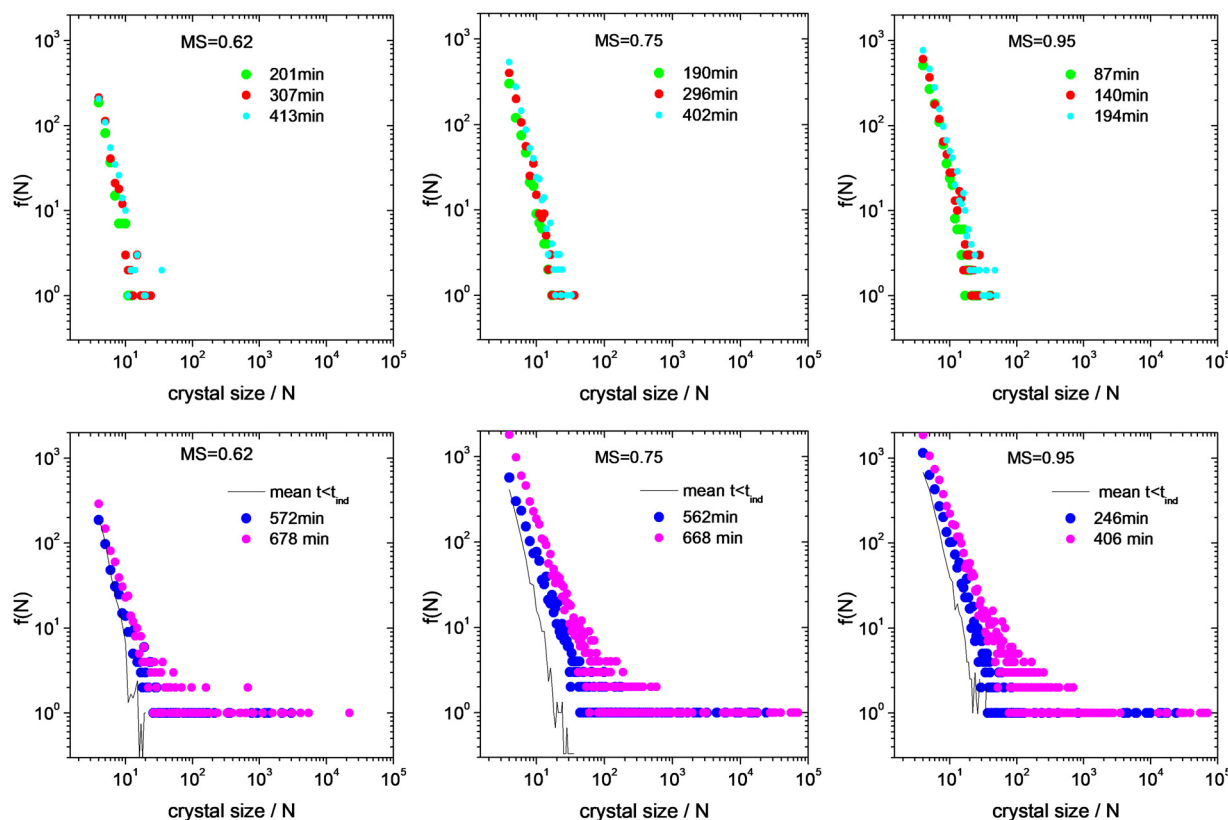


Fig. 3 Crystallite size distribution $f(N)$ for different times and metastabilities as indicated in the plots. The distributions measured in the induction phase (upper graphs) show a strong decay starting from several hundred crystalline clusters containing four particles. Clusters larger than about 30 particles are not observed. In the main crystallisation stage (lower row) $f(N)$ displays a weaker drop that merges into a “tail” of larger growing crystals. The thin line represents the mean of the data in the induction phase for comparison.

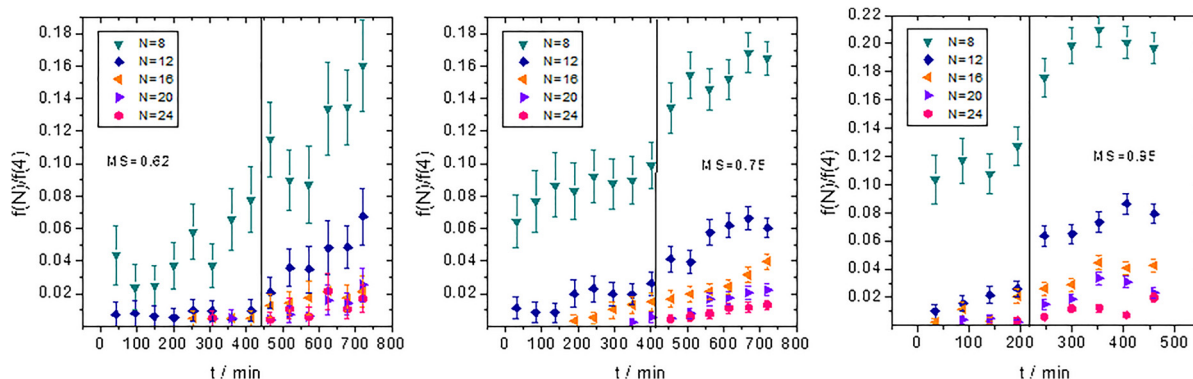


Fig. 4 Relative amount of subcritical clusters as function of elapsed time for different metastabilities as indicated in the plots. The vertical line marks the transition from the induction period ($t < t_{\text{ind}}$) to the nucleation and growth stage ($t > t_{\text{ind}}$).

CNT describes the nucleation process under the assumption of a stationary CSD and a constant surface tension. However, our experimental data demonstrates that this condition is not fulfilled *per se*. Especially in the nucleation and growth phase the CSD changes significantly. Strictly speaking, the assumption of a Boltzmann distribution to describe the CSD is not permissible. In the context of our experiment, quasi-stationarity it is most likely that to be fulfilled during the induction phase. Consequently, the ensuing analysis of the CSD within the framework of the CNT is based on data from this specific time regime.

Critical radius

In the following, we will determine the critical radius by employing a variety of methodologies. We will analyze the temporal evolution of the averaged crystal size, determine the frequency distribution of the critical radii from the growth trajectories of the overcritical crystals, determine the transition from subcritical to overcritical nuclei directly from the temporal development of the CSD, and determine the critical nucleus size in the context of the CNT analysing the CSD.

Analysing averaged crystal growth. The critical nucleus size can be estimated analysing the time trace of the averaged crystal size (Fig. 1). The crystal size at which crystal growth becomes apparent ($N \approx 25$ at $t \approx 400$ min in the shown example, $MS = 0.75$) represents a reasonable measure of the critical size.

Analysing individual trajectories of overcritical nuclei. As already mentioned, tracing the formation and growth of individual crystals (Fig. 2) gives direct access to r_{crit} . This is achieved by identifying the size at which continuous growth of the individual cluster commences. It has been frequently observed that this is the initial value of the trajectory. The determination of the critical size of individual overcritical growing crystals in a given sample enables the calculation of the critical size distribution. As illustrated in Fig. 5, the corresponding critical size distributions are characterised by a pronounced skewness. The majority of crystals starts their life with quite small size ($N < 10$), however, there are also critical nuclei containing more than a hundred particles. The mean

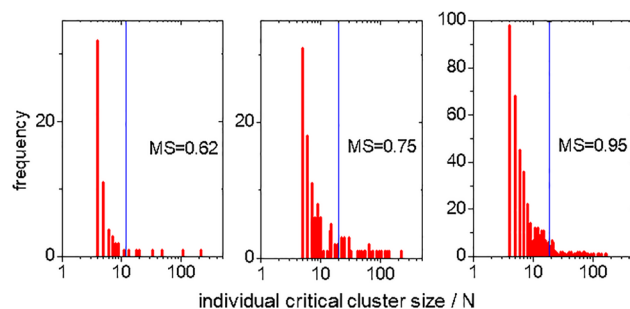


Fig. 5 Frequency distribution of the critical size for different metastabilities as indicated in the plots determined by analysing the trajectories of growing crystals; the blue line represents the mean value.

value, which is represented by the blue line, aligns closely with the other non-CNT based methods.

Direct analysis of the CSD. The time dependent crystal size statistics allows another determination of r_{crit} which is independent of CNT. A comparison of the CSD for times larger than the induction time with those less than the induction time (Fig. 3, lower row) reveals the transition from subcritical to supercritical nuclei in a direct manner. The size at which the continuous distribution of subcritical nuclei ($t < t_{\text{ind}}$) transverse to the growing ones ($t > t_{\text{ind}}$) can be identified with r_{crit} , which is about $N \approx 30$ in the shown examples.

Analysing the CSD in the framework of CNT. Here we will use the same approach to determine $\Delta G(N)$ as in the corresponding simulations.^{26,27,30,31,33} Within the framework of CNT the statistics of the crystalline clusters $f(N)$ gives access to ΔG_{crit} , γ and r_{crit} (eqn (1)–(3)). Experimentally, the Gibbs free energy can be determined by fitting²⁶

$$B - \ln(f(N)) = \Delta G(N) \quad (6)$$

to the data with B and γ as free fit parameter, where $f(N)$ is the absolute frequency of the crystals of size N . $\Delta\mu$ is calculated from the equation of state⁴⁸ and ρ_s is measured *via* Voronoi construction analysing the crystalline clusters.

ΔG of the crystals close to the induction time and the corresponding fit (eqn (6)) is shown in Fig. 6 for three



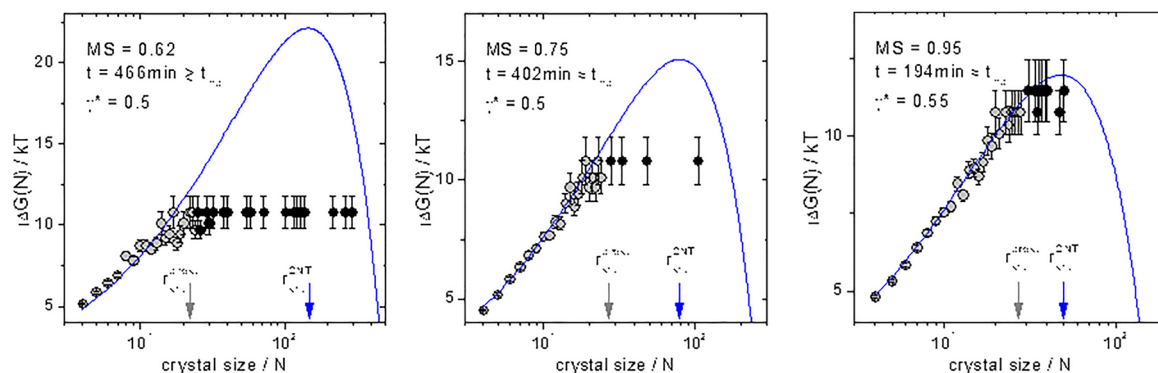


Fig. 6 Gibbs free energy at three metastabilities 0.62, 0.75 and 0.95. The time is close to the induction time as indicated in the plots. Grey points are subcritical cluster, black points represent overcritical and continuously growing crystals. From the fit (blue line, eqn (6)) the surface tension is determined as indicated in the plots. The grey and blue arrow mark the critical crystal size determined directly or by the CNT-fit.

metastabilities. The reduced crystal surface tension $\gamma^* = \gamma(2R_{HS})^2/k_B T \approx 0.5$ is in good agreement with theoretical studies.^{49–53}

The curve fitting according to CNT describes the data progression of the subcritical nuclei (grey dots) with high accuracy; however, the description of the supercritical nuclei (black dots) fails. The transition from subcritical to overcritical nuclei occurs much earlier in the experiment than predicted by the classical theory. This discrepancy becomes progressively smaller as one approaches the melting point ($MS = 1$). The nucleation barrier observed in the experiment is evidently smaller than that proclaimed by CNT. The underlying reason for this lower effective nucleation barrier will be identified in the chapters below.

Before we compare the nucleation rate densities and the different critical radii with the simulation data and previous experiments, it is essential to examine the structural evolution of the crystals. This is a critical component in analysing the data and a prerequisite for understanding the nucleation scenario.

Structural evolution of overcritical nuclei

In order to obtain information on the structural evolution of the supercritical nuclei, a detailed analysis of their structure was conducted using local bond order parameters,⁴⁷ and their development over time was monitored. Details of the procedure are given in the SI. As illustrated in Fig. 7, the temporal evolution of a crystallizing cluster exhibits a sequential progression. In the nascent stage of nucleation, the cluster is predominantly composed of non-registered hexagonal layers, depicted as green particles in the figure. As time progresses, the cluster undergoes an initial growth phase, followed by a subsequent registration of the non-registered layers, a process referred to as “locking in.” This registration event leads to the formation of a predominantly hcp structure, indicated by red particles. This sequence of events is a hallmark of the growth process of virtually all crystallizing clusters. The final crystal structure is achieved *via* the mediation of a metastable intermediate phase.

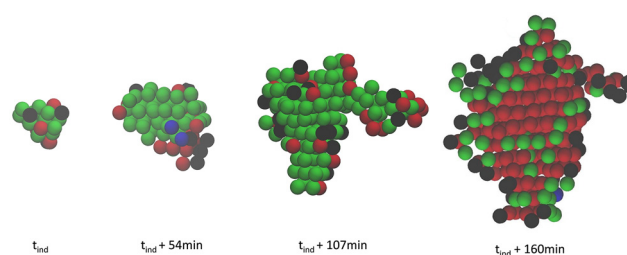


Fig. 7 Time series of a nucleation event. Color code: liquid (blue), non-registered hexagonal layers (green), hcp (red), fcc (black).

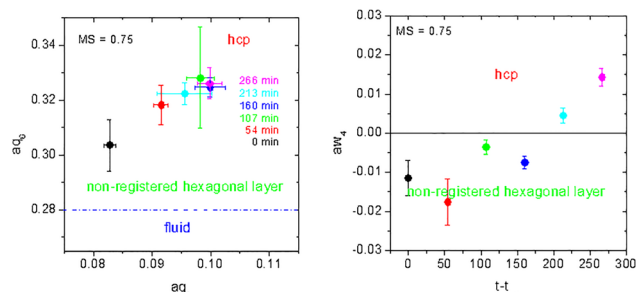


Fig. 8 Temporal evolution of the crystalline structure within the $aq_4(i)$ – $aq_6(i)$ -plane (left) and $aw_4(i)$ as function of elapsed time (right) for 124 crystallising clusters ($MS = 0.75$). The temporal evolution is shown in respect to the individual induction time $t - t_{ind}$ of each crystal.

This scenario can also be identified in the averaged data of all growing nuclei. As illustrated in Fig. 8, the mean LBOP values of all crystallizing clusters at $MS = 0.75$ have been calculated. It should be noted that the temporal evolution of a single cluster was corrected with the individual induction time in order to ensure the accuracy of the mean values. Shortly after cluster formation, the combination of aq_4 , aq_6 and aw_4 values shows clusters made of non-registered hexagonal layers. Later on, hcp crystals are formed.

The results of our analysis unequivocally demonstrate that the crystalline state of colloidal HS is reached through metastable intermediate states. This is in stark contrast to the CNT,



in which it is assumed that the nucleus has a 'perfect' fcc or hcp structure. Metastable intermediate states are easier to form than perfect crystals and obviously have a lower effective nucleation barrier.

Nucleation rate density and critical radius in comparison with previous data

In the following section, we will be conducting a comparative analysis of our novel results with existing data. The NRDs determined by LSCM are presented in Fig. 9 (normalized with the long time self diffusion coefficient D_s^L) in comparison with data from LS experiments and simulations. In instances where the long-time diffusion coefficient has not been explicitly measured, we have used the expression $D_s^L = D_0(\Phi/\Phi_{GT})^{2.58}$ to normalise the data.⁵⁴ Φ_{GT} denotes the position of the glass transition. If the Φ_{GT} was not determined we used a value of 0.58. It should be noted that the NRDs are plotted *versus* the position in the normalized coexistence width $MS = (\Phi - \Phi_F)/(\Phi_M - \Phi_F)$. This is an experimental measure of the difference in chemical potential (metastability, MS). This plot allows for a direct comparison of the data, as any artefacts caused by polydispersity and experimental concentration determination – systematic errors in Φ – are eliminated. Furthermore it offers the possibility to use the particle number density $\rho^\# = N/V_{\text{sample}}$ to calculate MS, which is directly accessible in confocal microscopy with very high precision – please see SI for more details. N denotes the number of particles in the observed volume V_{sample} and V_{particle} is the averaged volume of a single particle.

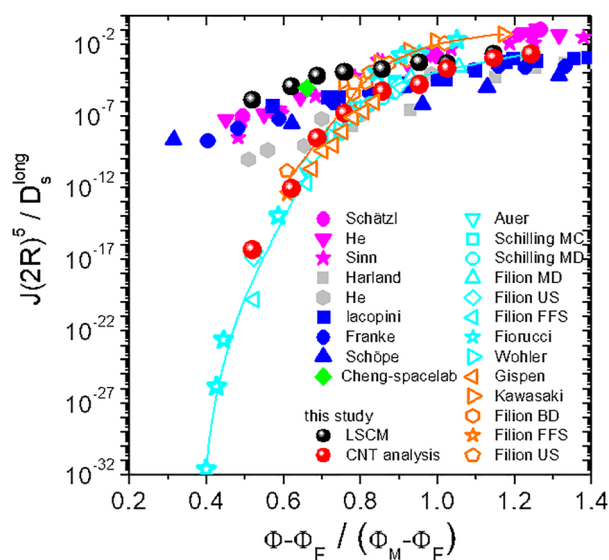


Fig. 9 Normalized crystal nucleation rate densities as function of metastability. Data from previous experiment^{16–19,21–25} (closed symbols in magenta, blue, grey, green) and simulations^{26,28–35} (open symbols in orange (WCA) and cyan (HS)) are shown. The lines are guides to the eye to highlight the numerical data. NRDs determined in this study by counting growing crystals (black circles) and within the CNT framework using the experimental barrier height (red circles). The experimental uncertainties are about one, in the CNT analysis two orders of magnitude. For details see text and SI.

The new experimental dataset reproduces the pre-existing experimental data. The absolute values and the slope below metastability $MS = 0.8$ cannot be harmonized with the simulations. Furthermore, it is noticeable that – within the experimental uncertainties – all but one experimental data overlap around the middle of the coexistence region ($MS = 0.5–0.6$). The only exception here is one of the datasets measured by He.¹⁸ We speculate that this may be due to a very asymmetric particle size distribution.⁵⁵

As mentioned above, several mechanisms have been proposed to explain the discrepancy between experiment and simulation: heterogeneous nucleation, deviation from the ideal HS interaction, sedimentation, hydrodynamic interaction, formation of twins, the data analysis in LS experiments, and systematic errors determining the hard sphere volume fraction. As listed in Table 1 in the experimental studies four different particle species with different chemical composition, five different solvent combinations with different viscosities, nine different particle sizes with different sedimentation rates and different hydrodynamic interactions have been used. One sample was measured under μ -gravity. In our experiment, no pronounced twinning was observed: on average, we detected around two twins per crystal at the end of the measurement for metastability, $MS = 0.75$. As all experimental data do overlap around the middle of the fluid crystal coexistence region it is clear that none of the speculated mechanisms is viable – please see Table 2.

The critical nucleus size r_{crit} obtained by analysing $f(N)$ in the CNT framework (Fig. 10, red diamonds) shows, as expected, a decreasing size with increasing metastability from $N = 280$ from the middle to $N = 40$ at the end of the coexistence region. The data agree very well with the simulation results^{26,27,30,31} (Fig. 10, stars). The same is true, of course, for ΔG_{crit} . On the other hand, the direct measurements indicate a constant r_{crit} that remains unaffected by metastability (Fig. 10, circles). This result is highly unexpected. Within the fluid crystal coexistence region the CNT based r_{crit} are incompatible with the real ones – particularly for $MS < 0.75$. Notably, this is also the MS-range where the NRDs exhibit the largest discrepancy.

This substantial discrepancy can be attributed to the formation of the crystalline structure *via* metastable intermediate states. Evidently, these intermediates possess a reduced critical radius.

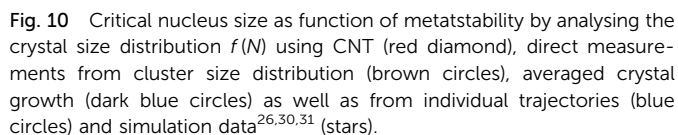
In order to complete the comparison with CNT, the NRD was calculated within the CNT framework (eqn (4) and (5)) using ΔG_{crit} obtained from the cluster size distribution by eqn (6). The scaling factor A is chosen in such a way that the data meet the experimental NRD at the highest metastability, $MS = 1.241$. As can be seen in Fig. 9, this data set (red circles) reproduces the simulation data fairly well.

All in all, the nucleation barriers, critical radii and NRDs determined from experimental data within the framework of CNT are consistent with those from simulations. However, they do not reflect the critical radii and NRDs determined directly from experiments without assuming any model.

In the following section, a summary of the findings will be provided and compared to CNT, followed by a discussion of the comparison with the simulation data.

Ref.	Core	s.l.	Dye	R [nm]	spd [%]	PSD	Sol	Pe	cw [%]	Method
Schätzel ¹⁶	PMMA	PHSA	—	500	≈ 5	?	D+T	0.192	5.1	SALS
He ¹⁸	PMMA	PHSA	—	495	≈ 5	?	D+T	0.192	5.1	SALS
Sinn ¹⁹	PMMA	PHSA	—	445	3.8	s.n.	D+T	0.121	5.1	SALS&BLS
He ¹⁸	PMMA	PHSA	—	215	7	h.n.	D+T	0.007	5.6	SALS
Harland ¹⁷	PMMA	PHSA	—	200	≈ 5	s.n.	D+CS ₂	0.005	5.1	BLS
Iacopini ²³	PS	PS	—	423	6.5	s.n.	2-EN	0.006	3.5	BLS
Franke ^{24,25}	PS	PS	—	410	5.5	sym	2-EN	0.005	4.4	BLS
Schöpe ²²	PMMA+TFEA	PHSA	—	320	4.8	sym	CD	0.036	3.5	BLS
Cheng ²¹	PMMA	PHSA	—	300	≈ 5	?	D+T	$< 10^{-6}$	4.6	BLS
Kale ⁴²	PMMA	PHSA	DilC18	726	5.75	s.n	CD+TCE	0.001	4.15	LSCM

Potential cause	Why it does not apply
Heterogeneous nucleation	Heterogeneous nucleation is not observed in LSCM It can be distinguished from homogenous nucleation in LS
Deviation from the ideal HS interaction (charge)	All data sets do overlap, although there are fundamental differences in chemical composition of particles and solvent. The system used in LSCM display HS interaction
Sedimentation	All data sets exhibit significant overlap in the fluid crystal coexistence region, despite substantial variations in sedimentation rates.
Hydrodynamic interaction	All data sets demonstrate substantial overlap in the fluid crystal coexistence region, despite considerable disparities in particle size and solvent composition.
Formation of twins	Excessive twinning is not observed in LSCM.
Data analysis of LS-experiment	Data from LSCM and LS overlap. In LSCM the ensemble averaged analysis agrees with the averaged analysis of individual crystal.
Systematic error in ϕ determination	By plotting the NRD data as function of metastability the systematic error in volume fraction determination becomes negligible. Furthermore, LSCM facilitates the determination of the particle concentration with a high degree of precision, as illustrated in the SI.



(i) The present study has demonstrated unequivocally that the metastable fluid's structural configuration undergoes alterations over time – we observe non-stationarity. In contrast, CNT assumes that a stationary Boltzmann distribution of subcritical nuclei is established immediately (or at least shortly) after the quench is completed. The nucleation rate is time-independent. The characteristics of the nucleation process are considered in terms of steady-state kinetics. The steady-state nucleation rate is calculated for the condition that the crystal size distribution remains constant over time. It is evident that these assumptions do not hold. In a recent article, Schilling *et al.* have presented a generally valid and exact equation of motion for the size distribution of the nuclei.¹² They finally come to the conclusion that stationarity is not given as a general rule, but may occur as a special rare case. With the obtained results our work supports this theoretical study.

(ii) The analysis of the structural evolution of the crystallising clusters reveals that crystallisation is mediated by intermediate states (precursors). These metastable intermediate states act as a bridge between the fluid and crystal. This phenomenon has been documented in a number of previous studies.^{22,24,28,32,56–59} As a consequence the nucleation barrier

This journal is © The Royal Society of Chemistry 2025

Table 3 Comparison of characteristic properties of MD³⁵ and experiment

	MD	Experiment
Sample size	$\sim 10^4$	$\sim 10^{12}$ (LSCM) $\sim 10^{13}$ (LS)
Analysed particles	$\sim 10^4$	$\sim 10^6$ (LSCM) $\sim 10^{13}$ (LS)
Boundaries	Periodic boundary conditions	Sample cell
Symmetry	Torus	Mirror planes
Solvent	No	Yes
Motion	Ballistic	Diffusive
Hydrodynamic interaction	No	Yes
Coupled systems	Spheres and bath via thermostat	Spheres, solvent and universe
Nucleation events per run	1	$\sim 10^2$ – 10^3 (LSCM) $\sim 10^5$ – 10^7 (LS)
Nucleation rate	N/t	dN/dt

Moreover, periodic boundary conditions are employed in the simulation to emulate the physics of an “infinitely” large sample. In the experiment, the sample is contained within a cuvette, with the particles and solvent molecules being reflected by the walls. The distinct boundary conditions give rise to different symmetries. The simulation exhibits torus-shaped symmetry, while the experiment features mirror planes. This results in fundamentally different conserved variables (Noether’s theorem), especially in the non-equilibrium fluctuations. Consequently, the mode spectrum (collective dynamics) in the metastable colloidal dispersion exhibits a different spectrum to that in the simulation, which influences the nucleation process.

It is important to note that the simulation does not take into account the dispersion medium, resulting in disparities in the dynamics of the particles, their interactions, and collective effects. In general, a minimum of two physical systems that are coupled with each other are aligned in their dynamics.^{68–70} That is to say, the systems synchronise. In the context of a colloidal suspension, this process entails the integration of three distinct systems: the colloidal particles, the dispersion medium (background fluid) and the laboratory (strictly speaking, the universe). Out of equilibrium, the non-equilibrium fluctuations of these three systems are coupled – the dynamics of the colloidal particles also includes the non-equilibrium fluctuations of the background fluid and the universe. In MD, the coupling of the ballistic HS-system to the bath (which is in TD equilibrium) is implemented by using a suitable thermostat. Gipsen *et al.* used a Nosé–Hoover thermostat (a deterministic one). The spectrum of non-equilibrium fluctuation in the experiment is, therefore, fundamentally different from that of the ballistic HS in MD. The experimental system is characterised by the presence of correlations – causality is not given, in contrast to the simulation, which exhibits a demonstrable causal chain. All said influences the nucleation kinetics in a fundamental way.

Moreover, the definition of the nucleation rate density used in the simulation study is not consistent with the one employed in the experiments and various textbooks. Fig. 11 illustrates the discrepancy between the two definitions. Consequently, the

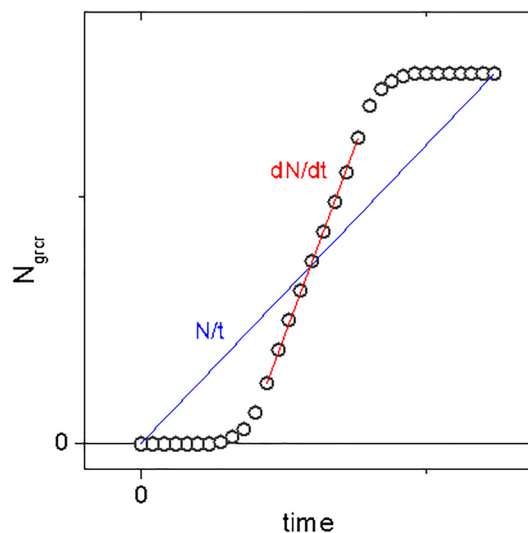


Fig. 11 Schematic of number of crystal's time trace of a large crystallising sample. The red line represents the steady state nucleation rate. The blue line is the expression used in MD simulation.²⁹

rate densities determined in the simulations are too low and subject to systematic errors. Based on the experimental data this effect can be as large as 4 orders of magnitude for $MS = 0.5$. This effect may offer a partial explanation for the significantly different slopes of the NRD. Nevertheless, it is unable to account for the missing orders of magnitude. The extent to which the mentioned discrepancies between experiment and simulation are responsible for the missing orders of magnitude must be clarified in future research. Subsequent publications on this subject are in preparation.

In summary, the inconsistency of the experimental data with those from US and FFS is mainly due to the presence of precursors and lack of stationarity in the experiment. The underlying reasons for the deviations from the MD data remain speculative: we hypothesise that the differences observed can be attributed to a number of factors, including variations in data analysis, different system size, the absence of correlation between nucleation events in MD, specific boundary condition leading to fundamentally different symmetries, the absence of background fluid in MD, and the type of coupling to the “bath”. Fundamentally different spectra of non-equilibrium fluctuations in MD and in the experiment result from the last three mentioned factors.

The experimental and simulation-based research provides a foundation for understanding crystal nucleation in HS systems, with all studies demonstrating validity under their respective boundary conditions. In comparison with the physics of the simulation, the experiment is characterised by a greater degree of complexity. However, a comprehensive cross-comparison of these studies is hindered by the significant variation in boundary conditions, which complicates the interpretation of results and the establishment of generalizable conclusions. It is therefore not surprising that the simulation data cannot be harmonized with the experimental data.



Conclusions

In conclusion we presented the first systematic study in which the nucleation process in gravity matched colloidal hard spheres was investigated by confocal microscopy over a wide volume fraction range. The nucleation rate densities, as determined by direct counting of growing supercritical nuclei, corroborates the outcomes of time-resolved light scattering experiments. These deviate by up to 22 orders of magnitude from simulated nucleation rate densities. The experimental data obtained in this study are comprehensive and detailed, and demonstrate clear deviations from existing simulations and classical theory.

The crystal size distribution as function of time demonstrates that stationarity – one of the most fundamental assumptions of CNT and transition state theory (TST) – is not fulfilled. Moreover, an approximately constant critical nucleus size is observed in the experiment, independent of the degree of metastability. In our opinion these findings are a surprising and significant discovery.

A detailed microscopic analysis of the nucleation process necessitates a re-evaluation of the conventional perspective on crystallisation. The crystal's genesis can be attributed to heterogeneities within metastable fluids. The scenario of “precursor mediated crystal nucleation” is confirmed: the crystalline state is reached *via* metastable intermediate states. The metastable intermediate states have a smaller nucleation barrier in comparison with perfect crystals. Consequently, the observed critical nucleus size appears to be a function of the transient metastable states that emerge, rather than the metastability itself. This ultimately results in a constant nucleus size. These significant and partly completely unexpected results call into question the classical picture of crystal nucleation and the applicability of transition state theory in general.

It is not expedient to engage in a discourse concerning whether the experiment or the simulation contains the correct physics of crystal nucleation in HS. All data obtained in the various studies are exact and correct in and of themselves. At present, experiments and simulations are inherently unequivalent, as they observe different physics under different boundary conditions. The commonality they share is their examination of the crystallisation process in HS systems albeit each study does so in its own way.

Author contributions

HJS designed and supervised the project; AL, SK and HJS characterized particles and designed sample cells; AL and SK wrote IDL codes; AL, SK and HJS prepared samples for LSCM and performed LSCM measurements; SK and HJS analysed experimental data; the manuscript was written by HJS. All authors have given approval to the final version of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Data availability

Data for this article are available on request – please contact the corresponding author.

Supplementary information (SI) is available. See DOI: <https://doi.org/10.1039/d5sm00776c>.

Acknowledgements

We gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (SCHO 1054/7-1). We thank Bill van Megen, Tanja Schilling and Martin Oettel for fruitful discussions, Andrew Schofield for synthesizing the particles and the German Center for Neurodegenerative Diseases (DZNE) within the Helmholtz Association, Tübingen for giving us the opportunity to perform LSCM measurements.

Notes and references

- 1 M. Polacci, F. Arzilli, G. La Spina, N. Le Gall, B. Cai, M. E. Hartley, D. Di Genova, N. T. Vo, S. Nonni, R. C. Atwood, E. W. Llewellyn, P. D. Lee and M. R. Burton, *Sci. Rep.*, 2018, **8**, 8377.
- 2 M. Jucker and L. C. Walker, *Nature*, 2013, **501**, 45–51.
- 3 M. D. Hollingsworth, *Science*, 2002, **295**, 2410–2413.
- 4 A. R. Tao, S. Habas and P. D. Yang, *Small*, 2008, **4**, 310–325.
- 5 M. Volmer and A. Weber, *Z. Phys. Chem.*, 1926, 277–301.
- 6 R. Becker and W. Döring, *Ann. Phys.*, 1935, **24**, 719.
- 7 S. P. Das, *Liquids at Freezing and Beyond*, Cambridge University Press, Cambridge, 2011.
- 8 D. W. Oxtoby, *Phys. Scr.*, 1993, **T49a**, 65–69.
- 9 P. G. Debenedetti, *Metastable Liquids: Concepts and Principles*, Princeton University Press, 1997.
- 10 D. Gebauer and H. Cölfen, *Nano Today*, 2011, **6**, 564–584.
- 11 S. Karthika, T. K. Radhakrishnan and P. Kalaichelvi, *Cryst. Growth Des.*, 2016, **16**, 6663–6681.
- 12 A. Kuhnhold, H. Meyer, G. Amati, P. Pelagejcev and T. Schilling, *Phys. Rev. E*, 2019, **100**, 052140.
- 13 J. Möller, A. Schottelius, M. Caresana, U. Boesenberg, C. Kim, F. Dallari, T. A. Ezquerro, J. M. Fernández, L. Gelisio, A. Glaesener, C. Goy, J. Hallmann, A. Kalinin, R. P. Kurta, D. Lapkin, F. Lehmkuhler, F. Mambretti, M. Scholz, R. Shayduk, F. Trinter, I. A. Vartanians, A. Zozulya, D. E. Galli, G. Gruebel, A. Madsen, F. Caupin and R. E. Grisenti, *Phys. Rev. Lett.*, 2024, **132**, 206102.
- 14 P. N. Pusey and W. van Megen, *Nature*, 1986, **320**, 340–342.
- 15 M. Fasolo and P. Sollich, *Phys. Rev. Lett.*, 2003, **91**, 068301.
- 16 K. Schätzel and B. J. Ackerson, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1993, **48**, 3766–3777.
- 17 J. L. Harland, S. I. Henderson, S. M. Underwood and W. van Megen, *Phys. Rev. Lett.*, 1995, **75**, 3572–3575.
- 18 Y. He, B. J. Ackerson, W. van Megen, S. M. Underwood and K. Schätzel, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1996, **54**, 5286–5297.
- 19 C. Sinn, A. Heymann, A. Stipp and T. Palberg, *Prog. Colloid Polym. Sci.*, 2001, **118**, 266–275.



- 20 U. Gasser, E. W. Weeks, A. B. Schofield, P. N. Pusey and D. A. Weitz, *Science*, 2001, **292**, 258–262.
- 21 Z. D. Cheng, P. M. Chaikin, J. X. Zhu, W. B. Russel and W. V. Meyer, *Phys. Rev. Lett.*, 2002, **88**, 015501.
- 22 H. J. Schöpe, G. Bryant and W. van Megen, *Phys. Rev. Lett.*, 2006, **96**, 175701.
- 23 S. Iacopini, T. Palberg and H. J. Schöpe, *J. Chem. Phys.*, 2009, **130**, 084502.
- 24 M. Franke, A. Lederer and H. J. Schöpe, *Soft Matter*, 2011, **7**, 11267–11274.
- 25 M. Franke, S. Golde and H. J. Schöpe, *Soft Matter*, 2014, **10**, 5380–5389.
- 26 S. Auer and D. Frenkel, *Nature*, 2001, **409**, 1020–1023.
- 27 S. Auer and D. Frenkel, *J. Chem. Phys.*, 2004, **120**, 3015–3029.
- 28 T. Schilling, H. J. Schöpe, M. Oettel, G. Opletal and I. Snook, *Phys. Rev. Lett.*, 2010, **105**, 025701.
- 29 T. Schilling, S. Dorosz, H. J. Schöpe and G. Opletal, *J. Phys.: Condens. Matter*, 2011, **23**, 194120.
- 30 L. Filion, M. Hermes, R. Ni and M. Dijkstra, *J. Chem. Phys.*, 2010, **133**, 244115.
- 31 L. Filion, R. Ni, D. Frenkel and M. Dijkstra, *J. Chem. Phys.*, 2011, **134**, 134901.
- 32 T. Kawasaki and H. Tanaka, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 14036–14041.
- 33 G. Fiorucci, G. M. Coli, J. T. Padding and M. Dijkstra, *J. Chem. Phys.*, 2020, **152**, 064903.
- 34 W. Wohler and T. Schilling, *Phys. Rev. Lett.*, 2022, **128**, 238001.
- 35 W. Gispen and M. Dijkstra, *J. Chem. Phys.*, 2023, **159**, 086101.
- 36 C. P. Royall, P. Charbonneau, M. Dijkstra, J. Russo, F. Smalenburg, T. Speck and C. Valeriani, *Rev. Mod. Phys.*, 2024, **96**, 045003.
- 37 S. Auer, W. C. K. Poon and D. Frenkel, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2003, **67**, 039904.
- 38 J. Russo, A. C. Maggs, D. Bonn and H. Tanaka, *Soft Matter*, 2013, **9**, 7369–7383.
- 39 M. Radu and T. Schilling, *EPL*, 2014, **105**, 26001.
- 40 J. Taffs, S. R. Williams, H. Tanaka and C. P. Royall, *Soft Matter*, 2013, **9**, 297–305.
- 41 S. Ketsetzi, J. Russo and D. Bonn, *J. Chem. Phys.*, 2018, **148**, 064901.
- 42 S. Kale, A. Lederer, M. Oettel and H. J. Schöpe, *Soft Matter*, 2023, **19**, 2146–2157.
- 43 M. Fasolo and P. Sollich, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2004, **70**, 041410.
- 44 S. Dorosz and T. Schilling, *J. Chem. Phys.*, 2012, **136**, 044702.
- 45 A. Lederer, M. Franke and H. J. Schöpe, *Eur. Phys. J.: Spec. Top.*, 2014, **223**, 389–407.
- 46 M. C. Jenkins and S. U. Egelhaaf, *Adv. Colloid Interface Sci.*, 2008, **136**, 65–92.
- 47 W. Lechner and C. Dellago, *J. Chem. Phys.*, 2008, **129**, 114707.
- 48 B. J. Ackerson and K. Schätzel, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1995, **52**, 6448–6460.
- 49 R. L. Davidchack, *J. Chem. Phys.*, 2010, **133**, 234701.
- 50 A. Härtel, M. Oettel, R. E. Rozas, S. U. Egelhaaf, J. Horbach and H. Löwen, *Phys. Rev. Lett.*, 2012, **108**, 226101.
- 51 R. Benjamin and J. Horbach, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2015, **91**, 032410.
- 52 M. Bültmann and T. Schilling, *Phys. Rev. E*, 2020, **102**, 042123.
- 53 C. Schoonen and J. F. Lutsko, *Phys. Rev. E*, 2022, **106**, 064110.
- 54 W. van Megen, T. C. Mortensen and G. Bryant, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2005, **72**, 031402.
- 55 Y. He, B. Oliver and B. J. Ackerson, *Langmuir*, 1997, **13**, 1408–1412.
- 56 M. Leocmach and H. Tanaka, *Nat. Commun.*, 2012, **3**, 974.
- 57 T. Kawasaki and H. Tanaka, *J. Phys.: Condens. Matter*, 2010, **22**, 232102.
- 58 J. T. Berryman, M. Anwar, S. Dorosz and T. Schilling, *J. Chem. Phys.*, 2016, **145**, 211901.
- 59 D. Gebauer, M. Kellermeier, J. D. Gale, L. Bergstrom and H. Colfen, *Chem. Soc. Rev.*, 2014, **43**, 2348–2371.
- 60 S. Golde, T. Palberg and H. J. Schöpe, *Nat. Phys.*, 2016, **12**, 712.
- 61 M. D. Ediger, *Annu. Rev. Phys. Chem.*, 2000, **51**, 99–128.
- 62 R. Richert, *J. Phys.: Condens. Matter*, 2002, **14**, R703–R738.
- 63 E. R. Weeks and D. A. Weitz, *Phys. Rev. Lett.*, 2002, **89**, 095704.
- 64 A. Widmer-Cooper and P. Harrowell, *J. Phys.: Condens. Matter*, 2005, **17**, S4025–S4034.
- 65 J. C. Conrad, P. P. Dhillon, E. R. Weeks, D. R. Reichman and D. A. Weitz, *Phys. Rev. Lett.*, 2006, **97**, 265701.
- 66 G. S. Matharoo, M. S. G. Razul and P. H. Poole, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2006, **74**, 050502.
- 67 P. K. Gupta, D. R. Cassar and E. D. Zanotto, *J. Chem. Phys.*, 2016, **145**, 211920.
- 68 L. M. Pecora and T. L. Carroll, *Phys. Rev. Lett.*, 1990, **64**, 821–824.
- 69 L. Kocarev and U. Parlitz, *Phys. Rev. Lett.*, 1996, **76**, 1816–1819.
- 70 S. Boccaletti, J. Kurths, G. Osipov, D. L. Valladares and C. S. Zhou, *Phys. Rep.*, 2002, **366**, 1–101.

