Crystallization and arrest mechanisms of model colloids

Thomas K. Haxton,[*a] **Lester O. Hedges,**[a,b] **and Stephen Whitelam**[*a]

We performed dynamic simulations of spheres with short-range attractive interactions for many values of interaction strength and range. Fast crystallization occurs in a localized region of this parameter space, but the character of crystallization pathways is not uniform within this region. Pathways range from one-step, in which a crystal nucleates directly from a gas, to two-step, in which substantial liquid-like clusters form and only subsequently become crystalline. Crystallization can fail because of slow nucleation from either gas or liquid, or because of dynamic arrest caused by strong interactions. Arrested states are characterized by the formation of networks of face-sharing tetrahedra that can be detected by a local common neighbor analysis.

Colloidal crystallization is of considerable interest because of the value of colloidal assemblies to technology[1,2] and the value of colloidal dispersions as model systems[3–5]. Colloidal particles can be made of controlled size, interaction strength, and interaction range, making them useful models for exploring the thermodynamic and kinetic factors that lead to the assembly of equilibrium and nonequilibrium condensed states of matter[6–8].

A defining feature of many colloidal suspensions is that their interactions can be short ranged compared to the nm to μm size of the colloidal particles. For example, van der Waals interaction, depletion interactions[9], and DNA base-pairing interactions[5,10–12] used to promote colloidal crystallization typically act over a range of distances small compared to the particle size. As a result, colloidal suspensions can exhibit phase behavior and assembly kinetics not typically seen in atomic or molecular systems[13–19].

A minimal and well-studied model of a colloidal dispersion is a collection of spheres of radius R_0 interacting via the 'square-well' potential[13,20–23]

$$U(r_{ij}) = \begin{cases} \infty & r_{ij} \leq 2R_0 \\ -\varepsilon & 2R_0 < r_{ij} \leq 2(1+\lambda)R_0 \\ 0 & r_{ij} > 2(1+\lambda)R_0, \end{cases} \quad (1)$$

where r_{ij} is the distance between the centers of particles i and j. In this model the solvent in which particles are dispersed is not represented explicitly. Square-well spheres display the well-known phase behavior of particles with isotropic attractions[14,15,24], summarized in Fig. 1. When the interaction range λ is large (Fig. 1 (a)), the square-well system exhibits sequential phase transitions from gas to liquid to crystal as temperature (i.e. the combination k_BT/ε) decreases. (Here "gas"

refers to a dilute suspension of colloids and "liquid" to a concentrated suspension.) As the interaction range decreases, the gas-liquid coexistence curve decreases in temperature (Fig. 1 (b)), eventually becoming metastable with respect to fluid-crystal coexistence (Fig. 1 (c)).

Many studies have shown that such metastable liquid-gas phase separation can play a crucial role in colloidal crystallization. Free-energy calculations[14,15], and dynamic simulations[16,17,19,24,25] of short-range attractive spheres show that the metastable liquid can promote a two-step crystallization pathway in which colloids coalesce into liquid droplets from which crystals nucleate. Two-step pathways have been observed experimentally in colloidal particles confined in two dimensions[26–28], in DNA-tethered nanoparticles[29], and in proteins[30,31], as well as in simulations of DNA-tethered nanoparticles[32].

Colloidal liquid-gas phase separation also plays an important role in the formation of (physical) gels at deep supercooling. Gels are nonequilibrium, disordered networks of particles with solid-like mechanical properties that result from their percolating structures. Gelation occurs because strong inter-particle bonding causes particles within aggregates to rearrange too slowly to allow equilibration on observed timescales. In the deeply supercooled spinodal regime, rapid liquid-gas phase separation can cause the formation of extended, non-compact colloidal aggregates which fail to relax into compact colloidal droplets[3,33]. Many experimental studies of polymeric colloidal particles possessing depletion attractions have found gelation to occur in preference to crystallization[34,35] (potentially exacerbated by effects of polydispersity[36,37]). Microscopic analysis of the colloid-colloid interaction networks formed during gelation shows gelation to be characterized by certain (overlapping) locally-favored motifs: gels in long-range repulsive colloids consist of networks of face-sharing tetrahedra (maximally bonded clus-

[a]Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States

[b]Department of Physics, University of Bath, Bath, BA2 7AY, United Kingdom
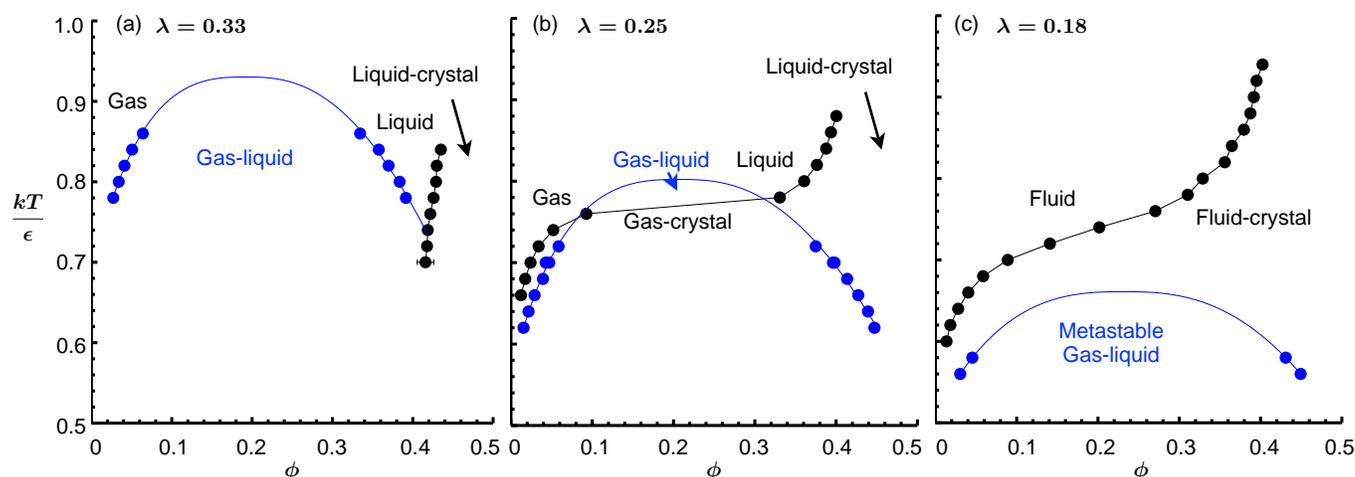
Fig. 1 Phase diagrams of hard spheres of radius R_0 with attractive square-well interactions of range $2(1+\lambda)R_0$, in the plane of density (hard-core packing fraction ϕ) and temperature (scaled interaction strength $k_\mathrm{B}T/\varepsilon$). Interaction ranges are (a) $\lambda = 0.33$, (b) $\lambda = 0.25$, and (c) $\lambda = 0.18$. Blue circles represent densities of coexisting fluids, calculated using Gibbs ensemble Monte Carlo simulations, and blue curves are fits to the coexistence curve for systems in the Ising universality class. Black circles represent densities of fluids coexisting with face-centered-cubic (fcc) crystals, calculated using direct coexistence simulations. Crystal densities, to the right of these plots, are not shown.

ters of four particles)[34,35], while gels of short-range-attractive spheres consist of networks of maximally bonded clusters of various sizes[38].
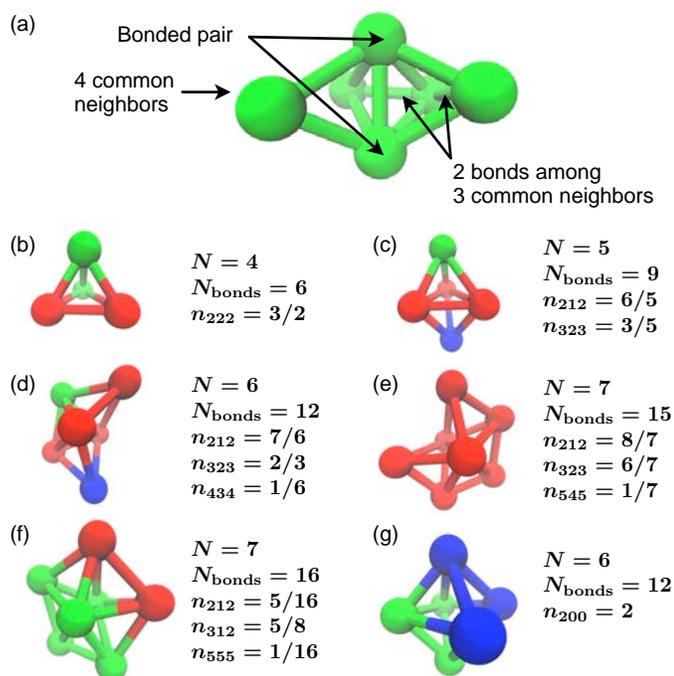
Recently, several authors have investigated the generality of crystallization and gelation mechanisms by characterizing colloid dynamics across broad sections of parameter space. Macfarlane *et al.* showed in experiments that DNA-linked nanoparticle crystallization occurs for each nanoparticle size only within a limited range of DNA lengths. Short lengths resulted in effective interaction ranges smaller than the nanoparticles' polydispersity, disfavoring the crystal thermodynamically, while large lengths inhibited kinetics[39].

Several authors have performed Monte Carlo, molecular dynamics, or Brownian dynamics simulations of spheres with short-range attractive interactions to investigate how assembly mechanism and product depend on interaction strength and/or concentration. Multiple studies have shown that crystallization occurs via two-step nucleation for a window of temperatures below the metastable liquid-gas transition, with deeper temperature quenches leading to gelation[16,17,24]. Extending these studies to multiple concentrations, Fortini *et al.* showed that crystallization coincides with the metastable liquid-gas transition temperature at concentrations below the liquid-gas critical concentration but occurs also at higher temperatures at supercritical concentrations[19]. Performing single-particle Monte Carlo simulations of square-well spheres at a single packing fraction of $\phi = 0.04$, Klotsa and Jack found an exception to the two-step rule: at a temperature near the metastable liquid-gas transition, they found that crystallization can proceed via a one-step pathway without significant formation of

amorphous clusters[25].

Complementing and extending these studies, we describe in this paper the self-assembly dynamics of square-well-attractive spheres over a broad spectrum of interaction strengths and ranges. We simulated sphere dynamics using the virtual-move Monte Carlo algorithm[40–43] parameterized so that colloidal clusters diffuse at rates agreeing with Stokes' law. Consistent with the previously mentioned studies at fixed interaction range, we find that efficient crystallization occurs in a localized region of parameter space, with a high-temperature boundary associated with the metastable liquid-gas transition. However, we find that the character of crystallization pathway varies within the region of efficient crystallization. Near the high-temperature boundary, crystallization proceeds along a one-step pathway via nucleation from the gas. Further below this boundary, crystallization proceeds along a two-step pathway via the formation of liquid-like clusters from which crystals subsequently nucleate. We find that poor crystallization at low temperature is characterized by the formation of networks of face-sharing tetrahedra that can be detected by a local common neighbor analysis[44].

It is important to note that the square-well model we have studied neglects features of real colloidal particles that may lead to complexity beyond that discussed here. The pairwise nature of the square-well interaction cannot capture collective properties of counterions, depletants, or polymer coats that mediate multi-body interactions between colloidal particles. For example, counterion entropy can favor gelation over crystallization in a way that cannot be modeled at a pairwise level[45], and the entropy of mobile linkers in the dilute-linker

(a)

Bonded pair

4 common
neighbors

2 bonds among
3 common neighbors

(b)　$N = 4$
$N_{bonds} = 6$
$n_{222} = 3/2$

(c)　$N = 5$
$N_{bonds} = 9$
$n_{212} = 6/5$
$n_{323} = 3/5$

(d)　$N = 6$
$N_{bonds} = 12$
$n_{212} = 7/6$
$n_{323} = 2/3$
$n_{434} = 1/6$

(e)　$N = 7$
$N_{bonds} = 15$
$n_{212} = 8/7$
$n_{323} = 6/7$
$n_{545} = 1/7$

(f)　$N = 7$
$N_{bonds} = 16$
$n_{212} = 5/16$
$n_{312} = 5/8$
$n_{555} = 1/16$

(g)　$N = 6$
$N_{bonds} = 12$
$n_{200} = 2$

Fig. 2 (a) Illustration of the common neighbor analysis for a bonded pair of spheres with a 423 common neighbor environment: the bonded pair shares four common neighbors, there are two bonds among those common neighbors, and three common neighbors participate in those two bonds. (b-g) Low-energy bond topologies for clusters of size $N \leq 7$. Linear polytetrahedral networks (b-e) and a closed 5-loop of face sharing tetrahedral (f) exhibit nonzero values of $n_{212}, n_{323}, n_{434}, n_{545}$, and/or n_{555}. The octahedron (g) is a maximally bonded cluster that is not composed of face-sharing tetrahedra. It has the same number of spheres and bonds as the 3-tetrahedron (d) and a different common neighbor signature, $n_{200} = 2$.

limit can favor the liquid state to the point of removing the triple point from Fig. 1 (a)[46]. The square-well model also treats solvent in an implicit manner; explicitly accounting for solvent and the long-ranged hydrodynamics it mediates may be important for colloidal crystallization under certain conditions[47,48].

1 Methods

1.1 Structure characterization

We characterized the dynamics by performing a common neighbor analysis of the network of spheres linked by favorable square-well interactions, similar to the analysis developed by Honeycutt and Andersen[44]. At regular time intervals, we recorded the number of bonded pairs of particles N_{abc} with a common neighbors, b bonds among those

common neighbors, and c common neighbors participating in those bonds (see Fig. 2 (a)). We defined the relative number $n_{abc} \equiv N_{abc}/N$, where N is the number of particles. This analysis identifies gaseous, liquid, crystalline, and polytetrahedral configurations. Perfect face-centered cubic and hexagonally close packed crystals exhibit nonzero values only of n_{423} (bulk hexagonal close-packed (hcp)), n_{424} (bulk hcp and fcc), n_{212}, and n_{312} (boundaries). Weakly interacting gases exhibit few bonds and uniformly low values of all common neighbor metrics, while weakly structured liquids exhibit large values of n_{200}. Networks of face-sharing tetrahedra exhibit large values of $n_{323}, n_{434}, n_{545}$, and/or n_{555} (see Fig. 2 (b-f)). We characterize the crystallinity by the fraction f_c of particles participating in at least one 423 or 424 bond.

To quantify the difference in crystallinity between dynamic simulations and simulations begun from fcc or bcc crystals (see below), we use the 'distance-to-equilibrium' parameter

$$\Delta_{eq} \equiv \max \left(\frac{n_{42x}^{fcc} - n_{42x}^{random}}{n_{42x}^{perfect}}, \frac{n_{666}^{bcc} - n_{666}^{random}}{n_{666}^{perfect}} \right). \quad (2)$$

Here the superscript 'random' describes yield of common neighbor types from dynamic (randomly-initialized) simulations, the superscripts 'fcc' and 'bcc' describe yields obtained from fcc- and bcc-initialized simulations, respectively, and the superscript 'perfect' describes the yield of a given environment in a bulk (perfect) fcc or bcc crystal, i.e. $n_{42x}^{perfect} = 6$ and $n_{666}^{perfect} = 4$.

In Figs. 6, 7, 8, 10, and 11 we use the following color code to denote particle environments. Particles participating in crystalline common neighbor environments (423 or 424) are colored green. Particles that do not, and that participate in polytetrahedral common neighbor environments (323, 434, 545, or 555) are colored red. The remaining particles that participate in in liquid common neighbor environments (200) are colored blue. The remaining particles that participate in other common neighbor environments abc with $a \geq 2$ are colored magenta. The remaining (gas) particles are colored gray.

1.2 Thermodynamics

We tested the thermodynamic stability of finite-size crystals by performing virtual-move Monte Carlo simulations (see below) initialized from compact crystals containing approximately 1000 spheres in a simulation box with an overall hardcore packing fraction of $\phi = 0.1$. We used a 923-particle cuboctahedron for the fcc crystal[49] and a 1001-particle cuboctahedron for the bcc crystal[50]. We initialized these crystals with nearest-neighbor distances $d = 2(1 + \lambda/2)R_0$, placing nearest neighbors in the middle of their interaction range. We defined the boundary of finite-size fcc stability (Fig. 4 (a)) by the contour where $n_{42x} = n_{423} + n_{424} = 1$. (As shown by the
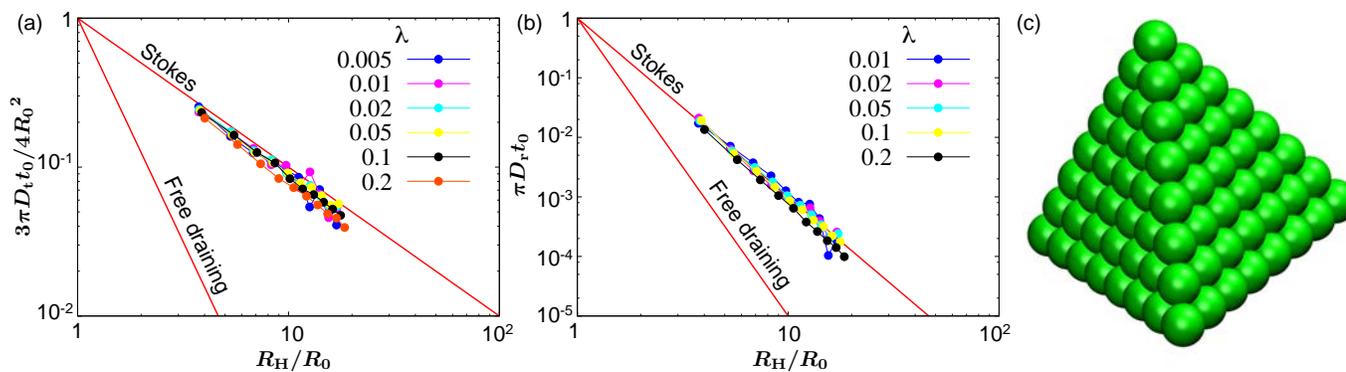
Fig. 3 (a) Translational and (b) rotational cluster diffusion constants, as a function of hydrodynamic radius R_H, for virtual-move Monte Carlo simulations of tetrahedral clusters composed of infinitely attractive square-well spheres. We show results for various interaction ranges λ (legend). The clusters used in these calculations were tetrahedral, and ranged in size from 1 to 120 particles; panel (c) shows a cluster of 120 particles. Algorithm parameters were $\Delta_t = 4\lambda R_0$, $\Delta_r = 1$, and $p_t = 0.4(R_0\Delta_r/\Delta_t)^2$ (see text). Cluster diffusion constants approximate the Stokes solutions ($D_t \propto R_H^{-1}$ and $D_r \propto R_H^{-3}$). For comparison, we show also the free-draining solutions ($D_t \propto R_H^{-3}$ and $D_r \propto R_H^{-5}$), which describe substantially slower collective motion.

relatively sharp decay of n_{42x} in Fig. 5 (a), the location of the boundary is relatively insensitive to choice of threshold.)

We calculated the boundary of stability of bulk fcc crystals by performing single-particle Monte Carlo (SPMC) direct co-existence simulations of 1000 spheres in a slab geometry at various temperatures and interaction ranges, choosing over-all packing fractions that allowed sufficient sampling of both fluid and crystal phases. We initialized the crystal slabs with nearest-neighbor distances $d = 2(1 + \lambda/2)R_0$, and we initialized the fluid phases with random configurations without hard-core overlaps. As shown for example by the black points in Fig. 1, these simulations allowed us to determine the coexistence concentrations for the fluid phase (gas or liquid, depending on T and λ) and the crystal phase (not shown). We defined the boundary of bulk crystal stability at $\phi = 0.1$ for each λ as the temperature at where the interpolated fluid coexistence concentration (black curves in Fig. 1) intersects $\phi = 0.1$.

We calculated the boundary of stability of the bulk liquid by performing SPMC Gibbs ensemble Monte Carlo simulations of 1000 spheres separated in two boxes that exchange spheres and volume[51]. Analogous to the direct coexistence simulations, we performed the Gibbs ensemble simulations at a range of temperatures for each interaction range, with overall packing fractions chosen to allow sufficient sampling of both phases, and we determined the boundary of liquid (meta)stability at $\phi = 0.1$ by interpolating the fluid coexistence curves (see blue points and curves in Fig. 1). We initialized both the gas and liquid box with random configurations without hard-core overlaps.

1.3 Dynamics

To approximate the overdamped dynamics of strongly-associating particles in solution we used the virtual-move Monte Carlo algorithm[40] (specifically, the version of the algorithm described in the appendix of Ref.[41]). Under this algorithm, which satisfies detailed balance, particles move locally according to the gradients of potential energy they experience, and collectively with a rate that can be controlled to a degree by the user. We parameterized the algorithm in a manner similar to that described in Ref.[43], in order to ensure that tightly-bound clusters of particles of hydrodynamic radius R_H diffused with rates close to those predicted by the Stokes' law,

$$D_t = \frac{k_B T}{6\pi\eta R_H},$$
$$D_r = \frac{k_B T}{8\pi\eta R_H^3}. \qquad (3)$$

The natural time unit of this motion is then the Brownian time scale

$$t_0 = \frac{\eta(2R_0)^3}{k_B T}, \qquad (4)$$

where η is the (implicit) solvent viscosity and $k_B T$ is the thermal energy. Simulations performed at particular values of $k_B T/\varepsilon$ and λ can therefore be considered to apply to a wide range of absolute particle sizes R_0, the latter determining only the value of t_0. For example, for spherical colloidal particles with radius $R_0 = 50$ nm at room temperature ($T = 293$ K) in water ($\eta = 1.00 \times 10^{-3}$ Pa s), the Brownian timescale is $t_0 = 2.5 \times 10^{-4}$ s.

In the Appendix we describe in detail the procedure we used. Briefly, the virtual-move algorithm generates collective

Monte Carlo moves by proposing trial 'virtual' particle translations or rotations, and probabilistically recruiting neighboring particles to join this motion, in an iterative fashion. The resulting trial move is accepted with a probability ensuring detailed balance. In addition, one is free to attenuate the rate at which collective motion is accepted, by imposing what are effectively kinetic constraints. We chose these constraints in order to enforce Eq. (3).

Each Monte Carlo move begins with either a trial translation or a trial rotation, chosen with probability p_t and $p_r = 1 - p_t$, respectively. For translations, we randomly selected a particle and translated it randomly within a ball of radius Δ_t. For rotations, we randomly selected a particle, randomly selected a second particle within the interaction range of the first, and rotated the second particle by an angle, chosen uniformly from the range $(-\Delta_r, \Delta_r)$, around a randomly-oriented axis $\hat{n}$ passing through the center of the first particle. We chose $\Delta_t = 4\lambda R_0$, $\Delta_r = 1$, and $p_t = 0.4(R_0 \Delta_r / \Delta_t)^2$. As discussed in the Appendix, we found that with this choice of parameters we could enforce Eq. (3) by suppressing the acceptance rate for translation and rotation of a cluster of N particles of hydrodynamic radius R_H by factors $N^{-1} R_H^{-1}$ and $N^{-1} R_H^{-3}$, respectively. We chose to maximize the ratio of rates of internal cluster relaxation to whole-cluster diffusion, by working with the smallest trial displacement Δ_t that is large enough to induce substantial collective motion (i.e. is large enough to ensure that Stokes' law could be maintained). This choice is somewhat arbitrary, and whether it is physically appropriate will likely depend on details of the experimental system one wishes to model, but we note that one has some freedom to influence this ratio if necessary.

In Fig. 3 we show measured diffusion constants for tetrahedral clusters of between 1 and 120 square-well spheres, in the $k_B T / \varepsilon \to 0$ limit. D_t and D_r approximate the Stokes solutions (Eq. 3) over a broad range of cluster sizes and interaction ranges. In this respect our procedure therefore captures an important aspect of solvent-mediated diffusion, without representing solvent explicitly. Note that 'long-ranged' hydrodynamic coupling[47] is not taken into account by this procedure; to do so, one should represent solvent more explicitly[48,52–54]. Simple implementations of Brownian (Langevin) dynamics integrators, and single-particle Monte Carlo simulations in the limit of zero trial displacement[55], result instead in the 'free-draining' behavior $D_t \propto R_H^{-3}$ and $D_r \propto R_H^{-5}$. As shown in Fig. 3 (note the logarithmic scale), such diffusion is significantly slower than Stokes' diffusion, even for relatively modest cluster sizes.

As discussed in the Appendix, our procedure yields a time per Monte Carlo cycle of

$$t_{\text{cycle}} = \frac{6}{5}\pi p_t \lambda^2 t_0, \tag{5}$$

where t_0 is the Brownian time scale (Eq. 4). We present results relative to the physical time unit t_0.

We carried out simulations of 1000 square-well spheres, in periodically-replicated cubic simulation boxes, at a hard-core packing fraction of $\phi = 0.1$. We carried out independent simulations for interaction ranges between (and including) the values $\lambda = 0.005$ and 1.35, and for temperatures ranging from $k_B T / \varepsilon = 0.06$ to 0.86. We initialized dynamic simulations with random configurations, under the constraint that the particle hard cores could not overlap (equivalent to equilibrium configurations in the $k_B T / \varepsilon \to \infty$ limit).

An open-source C++ library for implementing the virtual-move Monte Carlo algorithm is available at http://vmmc.xyz[56].

## 2 Results

2.1 Dynamic and thermodynamic phase diagrams

In Fig. 4(a) we show in the temperature-range plane the fcc crystal yield f_c seen in dynamic simulations at time $t = 10^5 t_0$; panel (b) shows yield also at times $10^3 t_0$ and $10^4 t_0$. High yield (green) is found in a localized region of parameter space. Note that the phase diagrams of Fig. 1 intersect the diagram of Fig. 4 via three vertical lines corresponding to particular values $\lambda = 0.18$, $\lambda = 0.25$, and $\lambda = 0.33$.

A necessary condition for high crystal yield is that the fcc crystal is stable thermodynamically; this condition holds for our compact fcc crystals below the solid green curve, marked "g-c (finite)" in Fig. 4. Note that the dashed green bulk gas-crystal coexistence curve (labeled "g-c") derived from direct coexistence simulations in a slab geometry lies above the finite-size curve. This difference simply reflects the fact that a finite crystalline cluster with free boundaries can melt within the regime of bulk crystal stability, i.e. can be smaller than the critical cluster size. Note also that the solid green curve bends toward small λ and small $k_B T / \varepsilon$ at around $\lambda = 0.32$, $k_B T / \varepsilon = 0.44$. This bend occurs because the fcc crystal becomes unstable with respect to to a body-centered cubic (bcc) crystal at large λ and small $k_B T / \varepsilon$ (see below).

A second necessary condition for high crystal yield expected from previous work[14–19] is that a state point must lie within the regime of metastable liquid-gas phase coexistence. For large interaction ranges λ we determined the boundary of liquid stability in bulk from Gibbs ensemble simulations (dashed gray curve in Fig. 4), a technique that eliminates interfacial effects by putting gas and liquid phases in separate boxes that interchange both volume and particles[51]. For interaction ranges smaller than $\lambda = 0.15$, we could not accurately determine gas-liquid coexistence because crystallization rapidly occurred within the liquid box. Instead, we estimated the onset of transient liquid-like structure as the curve
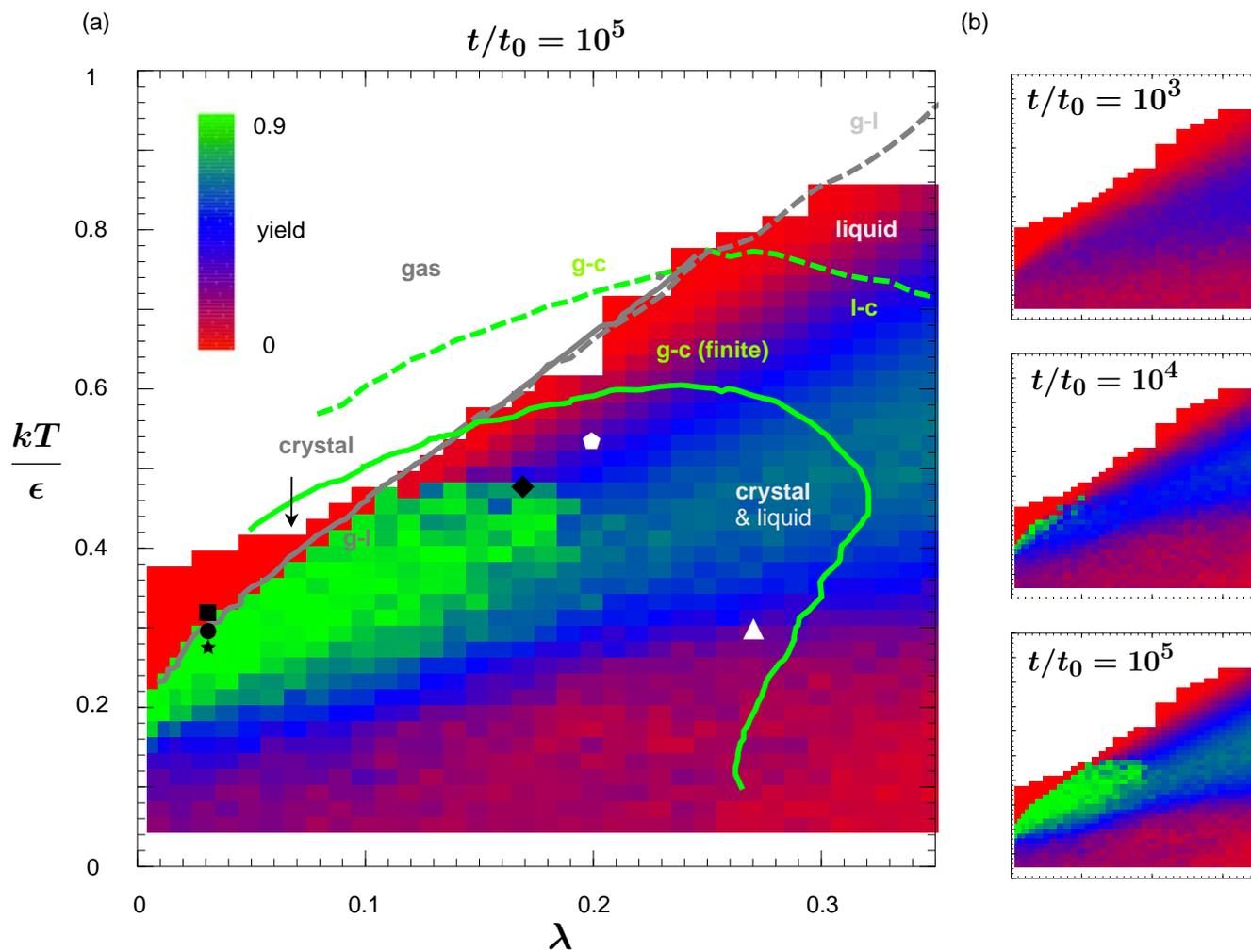
Fig. 4 (a) Crystal yield f_c from dynamic simulations after $t = 10^5 t_0$ as a function of interaction range λ and temperature $k_B T/\varepsilon$, for systems of 1000 square-well spheres at hard-core packing fraction $\phi = 0.1$. The dashed (solid) green curve indicates the boundary of stability of bulk (finite-size) fcc crystals, and the dashed (solid) gray curve indicates the boundary of stability of the bulk (finite-size) liquid (see text for details). Symbols indicate representative state points for the various dynamic regimes shown in Figs. 6, 7, 8, 10, and 11. (b) Yield at three times.
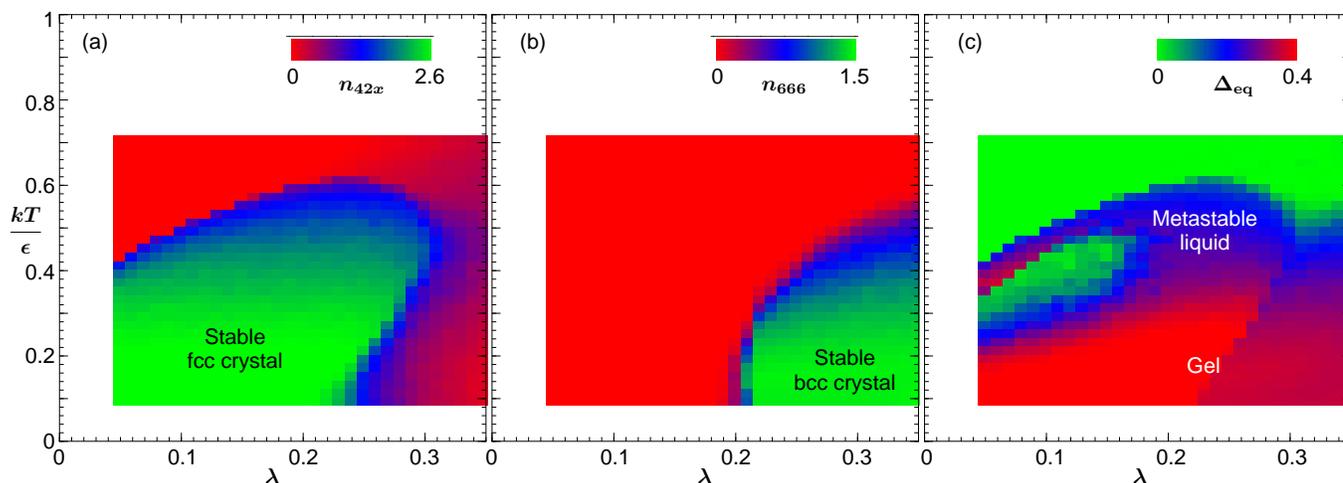
**Fig. 5** (a) Common neighbor metric n_{42x} indicating fcc crystallinity after simulations of length $t = 10^5 t_0$ initiated from perfect fcc crystals. (b) Common neighbor metric n_{666} indicating bcc crystallinity after simulations of length $t = 10^5 t_0$ initiated from perfect bcc crystals. (c) Distance-to-equilibrium Δ_{eq} (defined in Section 1.1) indicating how close the common neighbor metrics in the dynamic simulations (Fig. 4) get to the closest of the two crystal-initiated simulations.

below which our dynamic simulations attained a relative number of liquid-like bonds $n_{200} \geq 0.1$ at some point during our dynamic simulations (solid gray curve in Fig. 4; see also Fig. 12 (g)). We find that this curve coincides with the bulk liquid curve over the interval of interaction ranges for which both could be calculated, indicating that metastability of the liquid is not strongly influenced by the existence of free boundaries for system sizes on the order of 1000 particles. Although it is not clear to what extent the liquids at small interaction range can be considered metastable, comparison of the onset of high crystal yield (green pixels) with the liquid boundary (dashed and solid curves) is consistent with crystallization coinciding with the onset of transient liquid order and/or an extrapolation of the metastable liquid curve.

While our results show that the stability of the crystal and the onset of transient liquid structure are necessary conditions for rapid crystallization, they are clearly not sufficient: large regions of parameter space below the crystal and liquid boundaries appear blue or red in Fig. 4, indicating low crystal yield after $t = 10^5 t_0$, despite the fact that the systems must eventually assemble into a thermodynamically favored fcc crystal. The eventual (infinite-time) fate of the system can be inferred from Fig. 5, in which we plot in panels (a) and (b) the crystal yield that results from simulations initiated from a single fcc or bcc crystal, respectively. Note that the fcc crystal becomes unstable with respect to the body-centered cubic (bcc) crystal at large λ and low $k_B T / \varepsilon$, because bcc spheres can accommodate 8 nearest neighbors and 8 second-nearest neighbors at these values of λ, while fcc spheres can only accommodate 12 nearest neighbors. In this region of parameter space the bcc crystal is therefore lower in energy than the fcc crystal.
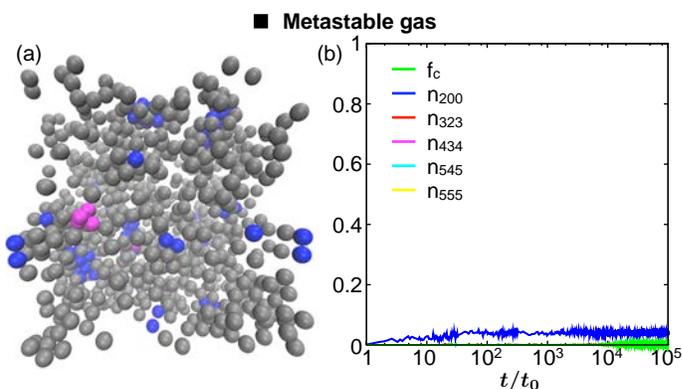
Fig. 6 (a) Snapshot from a simulation with $\lambda = 0.03$ and $k_B T / \varepsilon = 0.32$. The system remains in the metastable gas phase up to time $t = 10^5 t_0$. Particle color code is described in Section 1.1. (d) Time series of the crystal yield f_c (green), liquid (blue) and polytetrahedral common neighbor metrics.
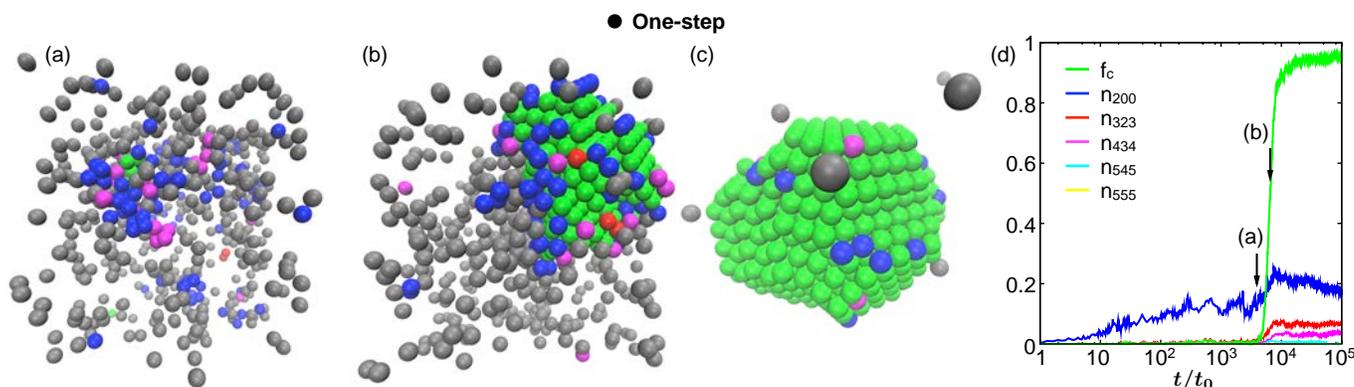
Fig. 7 (a-c) Snapshots from a simulation with $\lambda = 0.03$ and $k_B T / \varepsilon = 0.3$ that exhibits one-step crystallization (a) before nucleation, (b) after nucleation, and (c) at the end of the simulation ($t = 10^5 t_0$). (d) Time series of the crystal yield f_c and liquid and polytetrahedral common neighbor metrics. Liquid-like environments (blue curve and particles) are seen throughout crystallization, but the crystal that nucleates and grows does not have substantial liquid-like character. Arrows indicate the time points of the snapshots.

Panel (c) displays a 'distance-to-equilibrium' parameter (see Section 1.1) that summarizes how close dynamic simulations come to equilibrium: anything not shown green corresponds either to a metastable liquid or to an arrested gel.

As we will discuss below, low crystal yield within the thermodynamically stable crystal region is due to one of two kinetic effects, depending on the state point: either nucleation from the liquid is slow, or crystallization is arrested by gelation. Furthermore, we find that the kinetics of crystallization varies strongly even within the region of high yield: crystallization may proceed in a two-step pathway via a liquid-like intermediate, or it may proceed directly from a relatively homogeneous gas. The following five subsections discuss the five qualitatively different dynamic regimes encountered when broadly varying the interaction range and strength.

2.2 Metastable gas

The black square in Fig. 4 lies in the metastable gas regime. Here, as shown in Fig. 6, the system remains in a gas state, with low values of all common neighbor metrics. In this regime the face-centered cubic (fcc) crystal is the stable state, while the liquid is unstable with respect to the gas. The metastability of the gas for times up to $t = 10^5 t_0$ indicates the existence of large free-energy barriers for direct crystal nucleation from the gas.
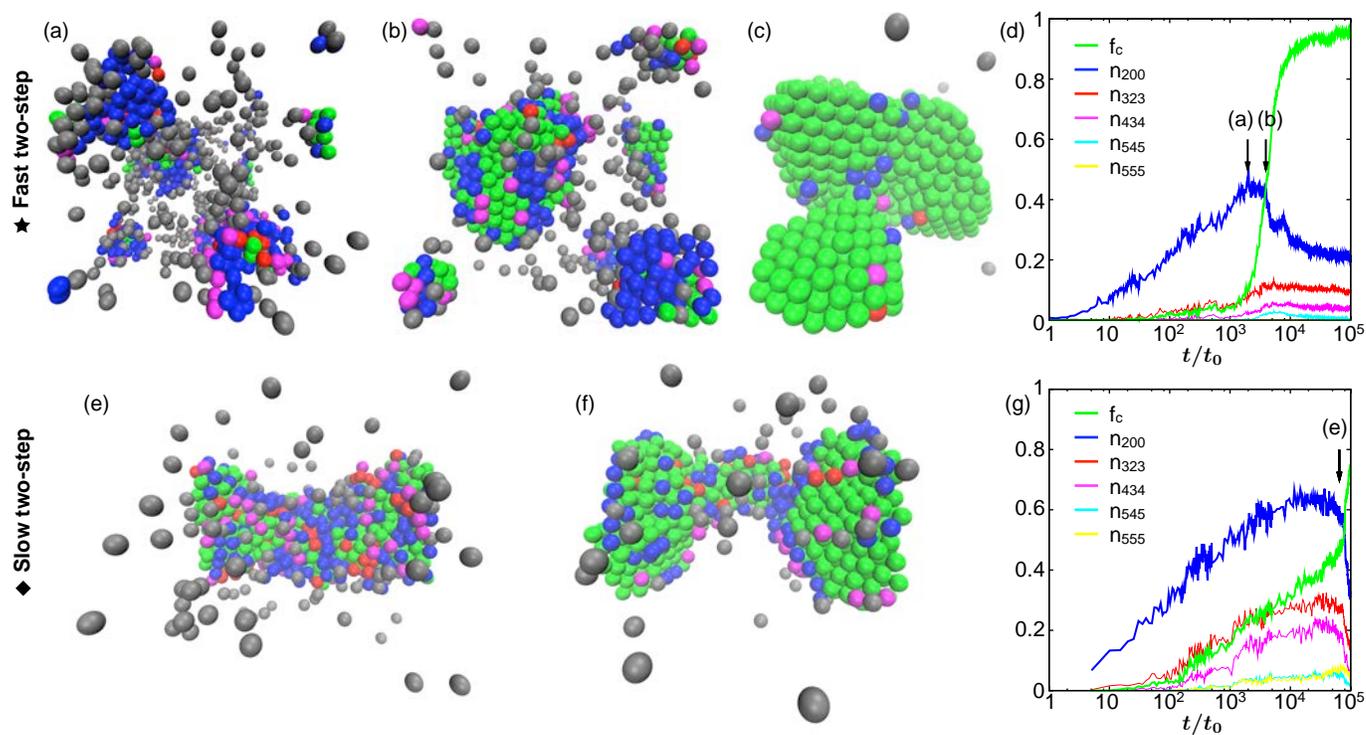
2.3 One-step crystallization

As the temperature decreases below the metastable liquid transition (dashed and solid gray curves in Fig. 4) the crystal yield increases, as indicated by the sharp change from red (low yield) to green (high yield) pixels in Fig. 4. As shown in Fig. 7 (black circle in Fig. 4), dynamic crystallization pathways near

the metastable liquid transition involve crystal nucleation from a fluid with substantial liquid-like fluctuations but no significant gas-liquid phase separation. Before the sharp increase in crystal fraction f_c at around $6 \times 10^3 t_0$ (Fig. 7 (b)), there are substantial fluctuations in the liquid common neighbor metric n_{200}, but the average value of n_{200} remains less than 0.2 (one liquid-like bond per five spheres). The nucleation event at around $6 \times 10^3 t_0$ does not occur at the expense of liquid-like structure, as it would if the crystal nucleated from within a liquid-like droplet. Instead, the value of n_{200} increases during the nucleation event. Thus, Fig. 7 illustrates a *one-step* pathway that appears to be facilitated by strong but non-critical density fluctuations. A similar pathway was identified in Ref. [25]. As discussed below, the region of parameter space in which we observe a one-step pathway is very narrow, consistent with the fact that it was not found in many other studies.

2.4 Two-step crystallization

Beyond this narrow region of one-step nucleation we find a broad range of parameters where crystallization occurs via a two-step pathway, illustrated in Fig. 8 (a-d) and (e-g) (black star and diamond, respectively, in Fig. 4). First, liquid-like clusters (Fig. 8 (a) and (e)) quickly nucleate, grow, and merge, resulting in an increase in the liquid common neighbor metric n_{200}. Later, crystals (Fig. 8 (b) and (f)) nucleate from within those droplets, resulting in a decrease in n_{200} and an increase in the crystallinity metric f_c. The nucleation time increases with increasing range λ, as can be seen by comparing the common neighbor time series of Fig. 8 (d) and (g). For $\lambda \geq 0.2$ the time for crystal nucleation from the liquid exceeds our simulation time $t = 10^5 t_0$.

Fig. 8 (a-c) Snapshots from a simulation with $\lambda = 0.03$ and $k_B T/\varepsilon = 0.28$ exhibiting fast two-step crystallization: (a) before crystal nucleation, (b) after crystal nucleation, and (c) at the end of the simulation ($t = 10^5 t_0$). (d) Time series of the crystal yield f_c and liquid and polytetrahedral common neighbor metrics. Arrows indicate the time points of the snapshots. (e-f) Snapshots from a simulation with $\lambda = 0.17$ and $k_B T/\varepsilon = 0.48$ exhibiting slow two-step crystallization: (e) before crystal nucleation and (f) after crystal nucleation, at the end of the simulation ($t = 10^5 t_0$). (g) Time series of the crystal yield f_c and common neighbor metrics.
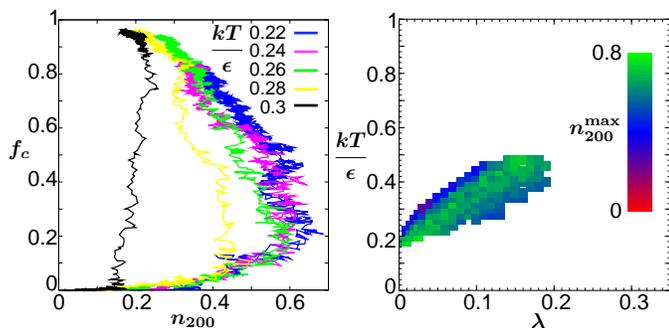
Fig. 9 Parametric pathway diagram illustrating the evolution of liquid structure (n_{200}) on the horizontal axis and crystalline structure (f_c) on the vertical axis, for a slice of state points with $\lambda = 0.03$ and $0.22 \leq k_B T/\varepsilon \leq 0.3$. All trajectories show high crystal yield $f_c > 0.7$ after $t = 10^5 t_0$. For most state points ($k_B T/\varepsilon \leq 0.28$) crystallization follows a two-step pathway, where first n_{200} increases and then n_{200} decreases while f_c increases. For $k_B T/\varepsilon = 0.3$ crystallization proceeds via largely a one-step mechanism, with n_{200} remaining low throughout assembly. (b) Maximum value of n_{200} during assembly as a function of λ and $k_B T/\varepsilon$, restricted to those state points for which $f_c > 0.7$ after $t = 10^5 t_0$. Most points are green, indicating a two-step pathway with large intermediate values of n_{200}. A narrow strip of state points near the extrapolated location of the gas-liquid curve are blue, indicating one-step assembly for which n_{200} is low throughout.

To illustrate the crossover from one-step to two-step pathway more generally, we show in Fig. 9(a) a parametric plot of the liquid common neighbor metric n_{200} versus the fractional crystal yield f_c, for a slice of state points ($\lambda = 0.03$ and $0.22 \leq k_B T/\varepsilon \leq 0.3$) having crystal yields $f_c > 0.7$ at $t = 10^5 t_0$. Most of these systems ($k_B T/\varepsilon \leq 0.28$, including the example $k_B T/\varepsilon = 0.28$ from Fig. 8) follow a pronounced two-step pathway. First, n_{200} increases with little increase in f_c, corresponding to liquid droplet nucleation, growth, and coalescence. Subsequently, n_{200} decreases and f_c increases, corresponding to crystal nucleation (fast or slow) from within the liquid droplet. In contrast, the example system with $k_B T/\varepsilon = 0.3$ (Fig. 7) does not exhibit pronounced two-step nucleation; instead, a crystal nucleates directly from the gas.

In Fig. 9 (b) we show the maximum value of n_{200} along each pathway for which $f_c \geq 0.7$ after $t = 10^5 t_0$. The narrow strip of values along the top of this region, near the metastable liquid transition, exhibit one-step behavior, with correspondingly low (blue) values of n_{200}^{max}.

2.5 Metastable liquid

At larger interaction ranges ($\lambda \geq 0.2$) and for temperatures close to the gas-liquid curve we find that the liquid remains metastable up to times $t = 10^5 t_0$. The liquid metric n_{200} is
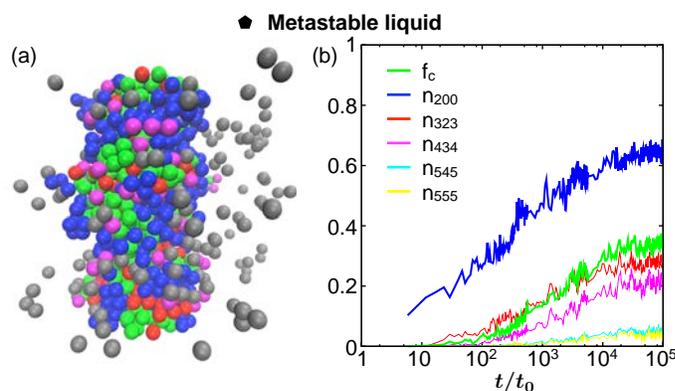

Fig. 10 (a) Snapshot from a system with $\lambda = 0.2$ and $k_B T/\varepsilon = 0.54$ which remains as a metastable liquid up to time $t = 10^5 t_0$. (b) Time series of the crystal yield f_c and liquid and polytetrahedral common neighbor metrics.

large, and the crystal metric f_c increases until it reaches a plateau that persists until the end of the simulation. An example trajectory is shown in Fig. 10 (white pentagon in Fig. 4). For larger systems and times that are longer (but still accessible to the corresponding experiments), this region of parameter space may give rise to good crystals.

2.6 Gelation

At low temperature, our simulated systems display the fast formation and persistence of polytetrahedral gels[34,35], which are bonded networks of face-sharing tetrahedra. As illustrated for small networks in Fig. 2 (b-f), polytetrahedral networks exhibit nonzero values of the common neighbor metrics $n_{212}, n_{323}, n_{434}, n_{545}$, and n_{555}. All but the n_{212} metric do not appear in perfect close-packed crystals; 212 environments appear on the 100 and 110 surfaces of fcc crystals. As shown for example in Fig. 11 (white triangle in Fig. 4) the metrics $n_{323}, n_{434}, n_{545}$, and/or n_{555} increase quickly and remain large up to times $t = 10^5 t_0$, while metrics characterizing crystals (f_c) and mobile liquids (n_{200}) remain low.

We find that the decrease in crystal yield at low temperature is accompanied by proliferation of these polytetrahedral networks. This correspondence can be seen across parameter space by referring to Fig. 12 (a-e), in which we compare (a) the crystal yield f_c after $t = 10^5 t_0$ with (b-e) the maximum value of the polytetrahedral common neighbor metrics $n_{323}, n_{434}, n_{545}$, and n_{555} achieved during the simulations. Although each common neighbor metric shows somewhat different dependence on λ and $k_B T/\varepsilon$, comparison of Fig. 12 (a) with Fig. 12 (b-e) shows that pathways involving large maximum values of the polytetrahedral common neighbor metrics account for most of the region in Fig. 12 (a) where the close-packed crystals are stable but yield is low.
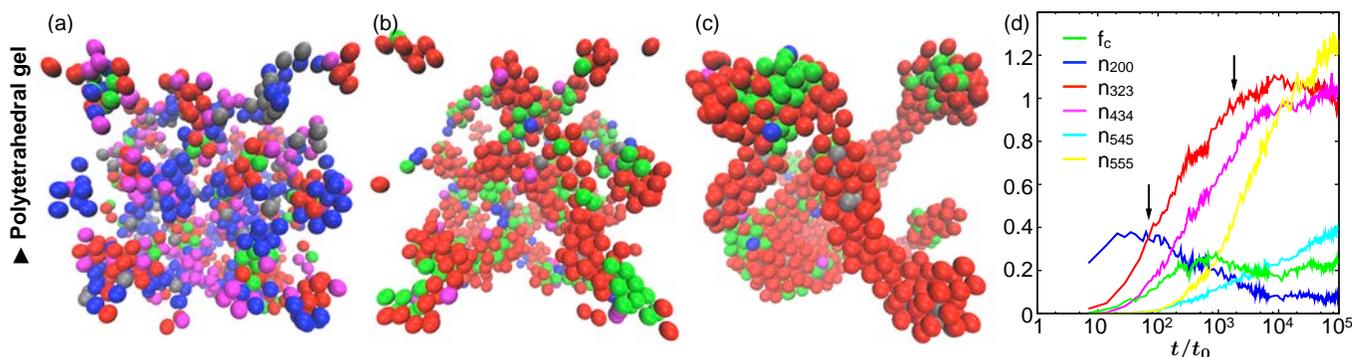
Fig. 11 Snapshots from a simulation with $\lambda = 0.27$ and $k_{\rm B}T/\varepsilon = 0.3$ that forms a polytetrahedral gel (a) before gelation, (b) after gelation, and (c) at the end of the simulation ($t = 10^5 t_0$). (d) Time series of the crystal yield $f_{\rm c}$ and liquid and polytetrahedral common neighbor metrics. Arrows indicate the time points of the snapshots.
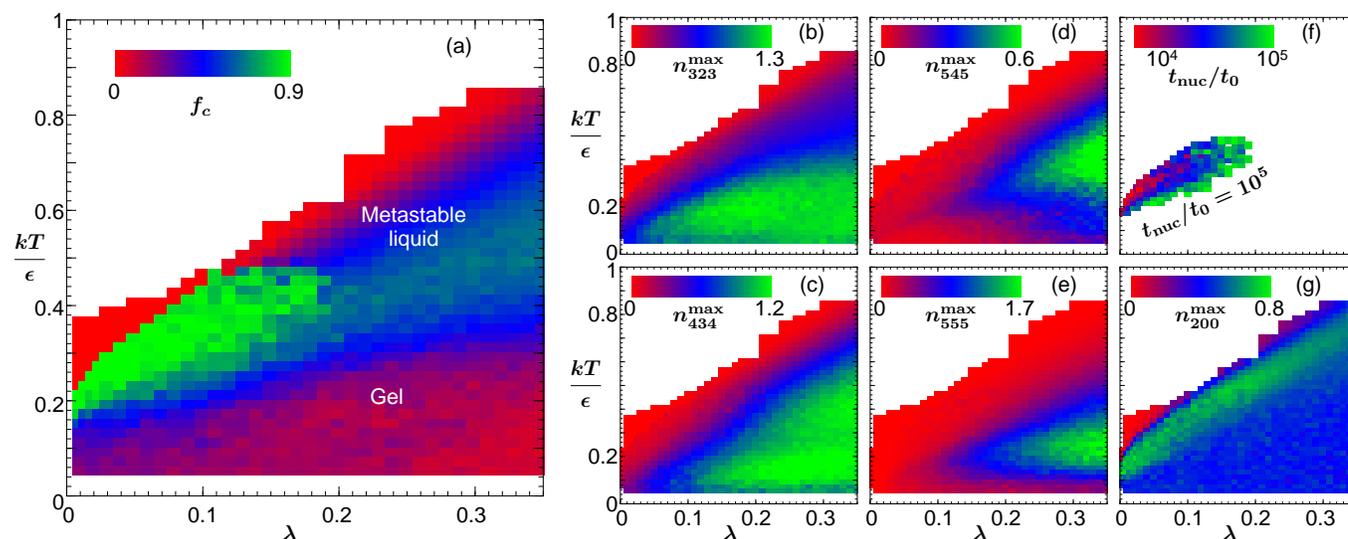
Fig. 12 (a) Relative crystal yield $f_{\rm c}$ after $t = 10^5 t_0$ as a function of interaction range λ and temperature $k_{\rm B}T/\varepsilon$. (b-e) Maximum values of the polytetrahedral common neighbor metrics (b) n_{323}, (c) n_{434}, (d) n_{545}, and (e) n_{555} over the course of the simulations, as a function of interaction range λ and temperature $k_{\rm B}T/\varepsilon$. The region of parameter space in the lower right of (a) ("gel") where $f_{\rm c} < 0.7$ is mostly associated with large values of $n_{323}^{\max}, n_{434}^{\max}, n_{545}^{\max}$, and/or $n_{555}^{\max}$ (b-e), except for the region labeled "metastable liquid". (f) Nucleation time $t_{\rm nuc}$ (time to first achieve $f_{\rm c} = 0.7$) on a logarithmic scale. (g) Maximum values of the liquid common neighbor metric n_{200}.

The predominance of polytetrahedral common neighbor metrics at low temperature suggests an explanation for the dynamic inaccessibility of the bcc crystal. Since the region of parameter space where the bcc crystal is stable (Fig. 5 (b)) is contained within the region where the polytetrahedral common neighbor metrics are large (Fig. 12 (a-e)), there is no region of parameter space where bcc crystallization proceeds efficiently. Rather, polytetrahedral gelation forestalls bcc crystallization wherever the bcc crystal is thermodynamically stable.

The only large region below the gas-liquid and liquid-crystal curves where neither f_c nor the polytetrahedral metrics are large is the region marked "metastable liquid" in Fig. 12 (a). As discussed above, crystal yield is low in this regime because the nucleation time is longer than $10^5 t_0$, not because there is substantial polytetrahedral gelation. To see that nucleation times extrapolate beyond $10^5 t_0$ in the metastable liquid regime, in Fig. 12 (f) we plot the time (on a log scale) at which each system first achieves a crystal yield of $f_c = 0.7$. Nucleation times increase to the right and top within the high-yield region until they pass beyond the $t = 10^5 t_0$ window. Fig. 12 (g) shows that beyond the high-yield region, the maximum value of n_{200} remains high, indicating that these systems achieve similar levels of liquid structure as in the high-yield region, the main difference being that crystals have not yet nucleated from the liquid after $10^5 t_0$.

3 Conclusions

Our results confirm that colloidal crystallization displays features common to many examples of self-assembly, in that it happens efficiently in only a small regime or 'sweet spot' of parameter space [25,57–59]. We confirm that efficient crystal nucleation of spherical particles with short-range attractions happens via a two-step pathway throughout most of parameter space. However, by performing a systematic investigation of colloidal crystallization as a function of interaction strength and range, we have found that largely one-step crystallization from the gas can occur for certain combinations of parameters. Specifically, over a narrow range of interaction strengths and ranges near the extrapolated location of the metastable liquid-gas boundary, we find that crystal nucleation is not preceded by substantial liquid-vapor phase separation and does not occur at the expense of liquid-like order. It is possible that small liquid-like fluctuations associated with the liquid-vapor phase boundary serve to speed nucleation, perhaps by lowering the interfacial tension between crystal and gas [14]. Given that we see the onset of liquid-like fluctuations in the simulation boxes we used for our dynamic simulations exactly where Gibbs-ensemble simulations suggest that the bulk liquid-vapor phase boundary lies, we do not expect the location of the one-step nucleation regime to change strongly with system size.

Our simulations show that low crystal yield at low temperatures or large interaction strengths is accompanied by particular local bond geometries. Crystal yield is low whenever there are many common neighbor configurations associated with face-sharing tetrahedra. Local bond geographies have been associated with arrested dynamics in related systems. For example, Karayiannis *et al.* showed that crystallization of hard spheres from concentrated amorphous fluids is limited by the persistence of five-fold coordinated local configurations [60,61].

Our result that gelation in short-range attractive spheres is associated with polytetrahedral order is largely consistent with a recent analysis of gelation in an experimental colloid-depletant mixture [38]. In Ref. [38] Royall *et al.* found that gelation can be explained by the formation of overlapping networks of locally favored states, defined for each cluster size $m \leq 13$ as the structure (set of bonds) that minimizes the potential energy for a reference potential. Royall *et al.* used a Morse potential for their calculations, which fit the potential of mean force for their system and led to locally favored states with predominantly polytetrahedral order [62]: all but one of the locally favored states in Ref. [38] are sections of a 13-particle icosahedron consisting of 20 face-sharing tetrahedra, and are thus themselves clusters of face-sharing tetrahedra. The one exception is the octahedron (Fig. 2 (g)), which has the same number of spheres (6) and bonds (12) as three face-sharing tetrahedra (Fig. 2 (d)), but is not composed of tetrahedra. The octahedron thus shows a distinct common neighbor signature: instead of exhibiting nonzero values of the polytetrahedral common neighbor metrics, the octahedron exhibits only a nonzero value of the common neighbor metric n_{200} that is prevalent at temperatures above gelation (Fig. 12 (g)).

We note also that within the square-well system, gelation can be detected by the study of local bond environments and does not require identifying minimum-energy clusters. Indeed, our results suggest that gels contain large numbers of some local configurations that do not minimize the potential energy: while 323, 434, and 555 bonds associated with polytetrahedral gelation are found in clusters that maximize the number of bonds, 545 bonds are not. Instead, 545 bonds are found in curved, linear polytetrahedral motifs as shown in Fig. 2 (e). As seen by comparing Fig. 2 (e) to Fig. 2 (f), these motifs do not maximize the number of bonds because they do not close into complete loops of face-sharing tetrahedra. The usefulness of local bond environments vs minimum-energy clusters is encouraging because the identity of maximally bonded clusters is known to vary strongly with potential. For example, in contrast to the Morse potential, square-well spheres with vanishingly small interaction ranges can have many distinct maximally bonded clusters for each cluster size [63–68].

Finally, our dynamic protocol suggests which combinations of temperature and colloid interaction range will yield best crystallization on the time scale $10^5 t_0$, where $t_0 =$

$\eta (2R_0)^3/k_B T$. For example, for spherical colloidal particles with radius $R_0 = 50$ nm, at room temperature ($T = 293$ K) in water ($\eta = 1.00 \times 10^{-3}$ Pa s), our results predict that crystallization after $t = 10^5 t_0 = 25$ s will be best for an attractive interaction of range $0.04R_0 = 2$ nm and strength $0.32 k_B T = 0.19$ kcal/mol.

4 Acknowledgements

5 Appendix: Virtual-move Monte Carlo algorithm

To efficiently approximate the overdamped and hydrodynamically coupled dynamics of strongly associating particles in solution, we used the virtual-move Monte Carlo (VMMC) algorithm[40,42], specifically the variant described in the appendix of Ref.[41]. We parameterized the algorithm to satisfy the Stokes solutions for the translational and rotational diffusion of clusters of hydrodynamic radius R_H,

$$
\begin{aligned}
D_t &= \frac{k_B T}{6\pi \eta R_H}, \\
D_r &= \frac{k_B T}{8\pi \eta R_H{}^3},
\end{aligned} \tag{6}
$$

while allowing as much internal relaxation of each cluster as possible. Our parameterization allows us to present results across interaction ranges and particles sizes relative to the natural Brownian time unit

$$
t_0 = \eta (2R_0)^3/k_B T, \tag{7}
$$

where η is the solvent viscosity, R_0 is the hard-core radius of the spherical particles, and $k_B T$ is the thermal energy.

The VMMC algorithm[40–42] moves individual particles and groups of particles with attempt and success frequencies designed to (1) preserve the correct equilibrium distribution, (2) ensure that particles move according to gradients in the potential energy, and (3) allow the dependence of diffusion coefficients on cluster size (e.g. Eq. (6)) to be controlled. The algorithm achieves this by proposing individual Monte Carlo moves, self-consistently generating individual or collective moves from the proposed moves, and accepting those moves in a way that satisfies the above three conditions.

In our implementation, trial individual translations are attempted with probability p_t by randomly selecting a particle and then attempting to translate it randomly within a ball of radius Δ_t. Because our particles are spherically symmetric, collective rotations cannot be generated from trial rotations about the center of mass of a single particle. Our trial rotations are attempted with probability $p_r = 1 - p_t$ by randomly selecting a particle, randomly selecting a second particle within the interaction range of the first, and then attempting to rotate the second particles by a randomly-chosen angle in the range $(-\Delta_r, \Delta_r)$ around a randomly oriented axis $\hat{n}$ centered at the first particle.

Acceptance rates in the VMMC algorithm consist of three factors: (1) a factor built on the Metropolis criterion ensuring that the system relaxes toward equilibrium and particles move according to gradients in the potential energy (2) a factor ensuring that motions of clusters are not oversampled with respect to the motion of isolated particles, and (3) a factor enforcing a prescribed dependence of diffusion coefficients on cluster size. The first factor is generic for any application of the VMMC algorithm. The second factor is usually used to produce realistic dynamics, but can be omitted when the algorithm is used only to sample an equilibrium distribution. The third factor has a general form that depends on the scaling of diffusion coefficients D_t and D_r on hydrodynamic radius R_H (e.g. $D_t \propto R_H{}^{-1}$ and $D_r \propto R_H{}^{-3}$ in Eq. 6), but the parameters of the algorithm (p_t, Δ_t, and Δ_r) must be tuned to ensure that the prefactor of the diffusion laws match the prescribed values[43]. Below, we review the VMMC algorithm and discuss our parameter optimization.

5.1 Review of the VMMC algorithm

We followed the 'symmetrized' version of the VMMC algorithm discussed in Refs.[41,42]. Trial 'virtual' single-particle translations and two-particle rotations are generated as discussed above, with probability p_t and $p_r = p_t - 1$, respectively. Trial collective moves are generated by iteratively testing each link between particles inside the moving group and particles outside the moving group, starting with an initial translation or rotation, until no links remain to be tested. When each link ij is tested, the particle outside the moving group (j) is *pre-linked* to the moving group with probability

$$
p_{ij}^{\text{link}} = \mathscr{I}_{ij} \max \left(0, 1 - \exp \left(\beta (U_{ij} - U_{i'j}) \right) \right), \tag{8}
$$

where $\mathscr{I}_{ij} = 1$ if particles i and j are within their mutual interaction range and $\mathscr{I}_{ij} = 0$ otherwise, U_{ij} is the initial interaction energy, and $U_{i'j}$ is the interaction energy when particle i executes a virtual move but particle j does not. After making a virtual move particle i is returned to its initial coordinates. If particle j is pre-linked to the moving group, then the probability of the reverse move (translation or rotation in the opposite

direction) is calculated,

$$p_{ij}^{\text{reverse}} = \max\left(0, 1 - \exp\left(\beta(U_{ij} - U_{i''j})\right)\right), \quad (9)$$

where $U_{i''j}$ is the interaction energy when particle i does the reverse virtual move and particle j stays still. After the reverse virtual move i is restored to its initial position. Then, the pre-linked particle j is *linked* (admitted) to the moving group with probability

$$p_{ij}^{\text{link}} = \min\left(1, \frac{p_{ij}^{\text{reverse}}}{p_{ij}^{\text{link}}}\right); \quad (10)$$

otherwise, the link is marked as *frustrated*. This procedure is continued until it converges on a final moving group, with all links having been tested between particles inside and outside the group.

Since the algorithm will move the entire group by the initial translation or rotation, the acceptance of collective moves must be scaled by a factor of $1/N$ to prevent over-sampling the motion of particles that are in large clusters. This is achieved by rejecting moves *in situ* when the number of particles in a moving group grows larger than $1/x$ (translations) or $2/x$ (rotations), where x is a random number on the interval $(0, 1]$ chosen at the beginning of each Monte Carlo step. The factor 2 appears because only rotations of groups of at least two particles are explicitly simulated for spherical particles; rotations of individual particles can be assumed to occur at any arbitrary rate while having no effect on the center of mass motion of the particles.

An additional scaling of acceptance probabilities can be applied to control the dependence of diffusion on cluster size. The Stokes scalings (Eq. 6) can be perfectly enforced in the limit where particles are tightly bound to each other within well-defined clusters and do not interact outside of these clusters. In this limit, any trial Monte Carlo move results in the recruitment of the entire cluster into a moving group, followed by the translation or rotation of the entire cluster. The VMMC algorithm can enforce Eq. 6 in this limit by rejecting moves *in situ* when the hydrodynamic radius R_{H} of the moving group exceeds R_{min}/y (translations) or R_{min}/y^3 (rotations), where R_{min} is the minimum possible hydrodynamic radius (see below) and y is another random number on the interval $(0, 1]$. Following previous implementations of the VMMC algorithm [40] we estimate the hydrodynamic radius R_{H} of a group $\mathscr{G}$ as a generalization of the radius of gyration,

$$R_{\text{H}}^2 \equiv 10\langle|(\vec{r} - \vec{r}_{\text{center}}) \times \hat{n}|^2\rangle_{\vec{r} \in \mathscr{G}}, \quad (11)$$

where $\vec{r}_{\text{center}}$ is the group's center of mass (center of rotation) and $\hat{n}$ is the direction of the translation (axis of rotation) for translations (rotations). This factor is the same for for translations that occur in opposite directions, and for rotations that occur with opposite sense, as is required for detailed balance.

We take $\vec{r} \in \mathscr{G}$ to include all points within the hard cores of the particles. The minimum hydrodynamic radius for both translations and rotations (single-sphere translations or effective two-sphere rotations about an axis $\hat{n}$ parallel to the separation vector between the particles) is the physical sphere radius R_0 (it would be $R_0/\sqrt{10}$ without the factor of 10 that appears in Eq. (11)).

Once a moving group has been generated that is not rejected *in situ* due to its number of particles or hydrodynamic radius, two additional factors contribute to its acceptance probability. First, the move is rejected if there are any frustrated links between particles inside and outside the moving group. Second, moves that remain valid are accepted with probability

$$W_{\text{acc}} = \min\left(1, \prod_{\langle ij\rangle_{0\leftrightarrow\text{p}}} \exp\left(-\beta\left(U_{i'j} - U_{ij}\right)\right)\right), \quad (12)$$

where the product runs over all pairs of particles (i in the moving group and j outside it) that are non-interacting before the move and have positive pair energy after the move, or vice versa. Together, these factors ensure that the system satisfies superdetailed balance, a condition that implies detailed balance [69]. For square well spheres that have only zero, negative, or infinite positive interaction, Eq. 12 reduces to a rejection if the move results in any hard-core overlap,

$$W_{\text{acc}} = \prod_{i\in\mathscr{G}, j\notin\mathscr{G}} \theta(r_{i'j} - 2R_0), \quad (13)$$

where θ is the heaviside step function.

For rotations of moving groups large enough to interact with their periodic images we imposed an additional rejection if a move resulted in a hard core overlap with a periodic image. Such overlaps occurred only for gels.

5.2 Parameter optimization

The prescribed dependence of diffusion coefficients with cluster size (Eq. 6) is derived for the VMMC algorithm in the limit of vanishingly narrow potential energy wells. In this limit, trial moves always take particles out these wells. If the wells are deep relative to $k_{\text{B}}T$ and the clusters are isolated, this causes the algorithm to recruit the entire cluster into the moving group, always accept the move, and thus generate diffusion coefficients dictated by the *in situ* rejection of collective moves as a function of hydrodynamic radius.

In practice, potential energy wells are not vanishingly narrow. Properly modeling the motion of such systems requires (1) allowing degrees of freedom internal to the clusters to relax and (2) ensuring that the prescribed diffusion laws are obeyed even when whole-cluster moves are not always generated. We sought to satisfy these conditions by choosing algorithm parameters that allow clusters to internally relax as much as possible without violating the Stokes scaling. We achieved this

by selecting the smallest values of Δ_t and Δ_r that resulted in the Stokes solutions for a test set of tetrahedral clusters of size 1 to 120 (8 spheres along each edge) for a range of interaction ranges λ and $T \to 0$ (infinite square well attractive interaction). We determined the optimal algorithm parameters in four steps.

First, we recorded the translational diffusion coefficient D_t from translation-only VMMC simulations of isolated tetrahedral clusters with range $\lambda = 0.11$ and compared it to the predicted value from Eq. 6,

$$D_t^{\text{Stokes}} = D_t^0 \frac{R_0}{R_H}, \tag{14}$$

where

$$D_t^0 = \frac{\Delta_t^2}{10 t_{\text{cycle}}} \tag{15}$$

is the translational diffusion coefficient for a single sphere translating with Monte Carlo dynamics. As shown in Fig. 13 (a), we found that for a range of tetrahedral cluster sizes $D_t \ll D_t^{\text{Stokes}}$ for $\Delta_t \leq 0.2 R_0$, $D_t \simeq 0.7 D_t^{\text{Stokes}}$ for $\Delta_t = 0.4 R_0$, and $D_t \simeq D_t^{\text{Stokes}}$ for $\Delta_t \geq R_0$. Although the translational diffusion coefficient is not quite saturated at the Stokes limit at $\Delta_t = 0.4 R_0$, we chose to parameterize Δ_t near this value because this balances a diffusion coefficient near the Stokes limit with ample internal relaxation: as shown in Fig. 13 (b), more single-particle than whole-tetrahedron moves are accepted for $\Delta_t = 0.4 R_0$, while the reverse is true for $\Delta_t = R_0$.

Second, we recorded the rotational diffusion coefficient D_r from rotation-only VMMC simulations of the same tetrahedral clusters and compared it to the predicted value from Eq. 6,

$$D_r^{\text{Stokes}} = D_r^0 \left(\frac{R_0}{R_H} \right)^3, \tag{16}$$

where

$$D_r^0 = \frac{\theta_{\text{ms}}(\Delta_r) \Delta_r^2}{6 t_{\text{cycle}}} \tag{17}$$

is the rotational diffusion coefficient for a single sphere rotating with Monte Carlo dynamics. In Eq. 17 $\theta_{\text{ms}}(\Delta_r)$ is the mean-squared rotation angle (about a fixed arbitrary axis centered at the center of mass) that a sphere would experience if we applied our Stokes-scaled VMMC algorithm to volume elements within the sphere. This factor is necessary for rotations because, unlike for translations, the hydrodynamic radius (Eq. 11) depends on the center and axis of rotation. Defining $\theta(\Delta_r, \hat{n}, \vec{r})$ as the center-of-mass rotation angle for a rotation by Δ_r around $\hat{n}$ centered at $\vec{r}$, we find

$$\theta_{\text{ms}}(\Delta_r) = \frac{1}{\Delta_r^2} \left\langle \left(\theta(\Delta_r, \hat{n}, \vec{r}_{\text{center}}) \right)^2 \left(\frac{R_0}{R_H(\vec{r}_c, \hat{n})} \right)^3 \right\rangle_{\hat{n}, |\vec{r}_c| < R_0}, \tag{18}$$
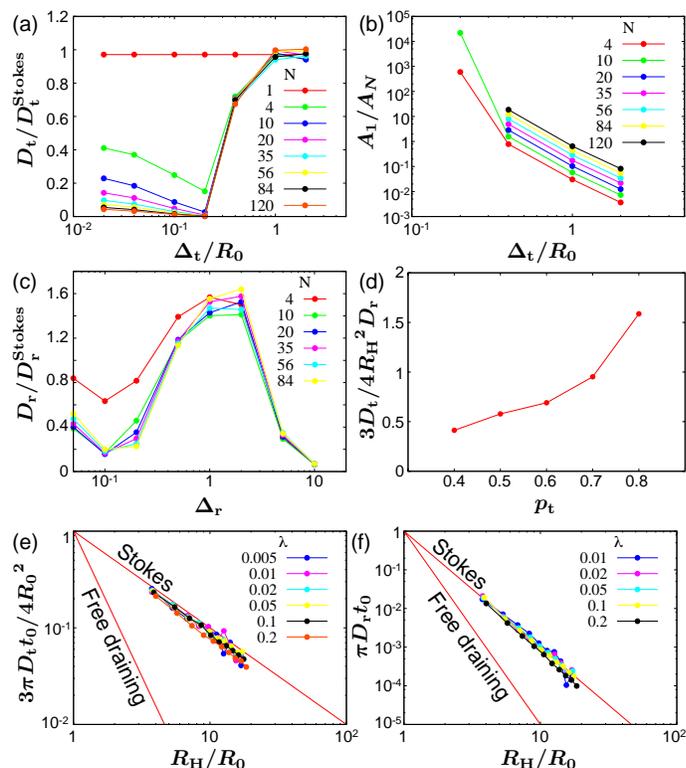
Fig. 13 (a) Ratio of translational diffusion coefficient to the value predicted by the Stokes solution (Eq. 14) for translation-only VMMC simulations of tetrahedral clusters composed of spheres with infinite square well attractive interactions of range $\lambda = 0.11$. (b) Ratio of accepted single-sphere moves to whole-cluster moves for the same set of simulations. (c) Ratio of rotational diffusion coefficient to the value predicted by the Stokes solution (Eq. 16) for rotation-only VMMC simulations of the same tetrahedral clusters. (d) Ratio of translational to rotational diffusion coefficients for a four-sphere tetrahedron with infinite square well attractive interactions as a function of the attempt probability for translations. The ratio is normalized such that the Stokes solution corresponds to a value of 1 (see Eq. 20). (e) Translational and (f) rotational diffusion coefficients vs hydrodynamic radius for simulations of tetrahedral clusters of various sizes composed of infinite square well attractive spheres with various interaction ranges λ (legend), using the full translating and rotating VMMC algorithm with parameters $\Delta_t = 4\lambda R_0$, $\Delta_r = 1$, and $p_t = 0.4(R_0 \Delta_r / \Delta_t)^2$. For comparison, the Stokes solutions ($D_t \propto R_H^{-1}$ and $D_r \propto R_H^{-3}$) and free draining solutions ($D_t \propto R_H^{-3}$ and $D_r \propto R_H^{-5}$) are shown as lines.

where

$$R_{\mathrm{H}}(\vec{r}_{\mathrm{c}}, \hat{n}) = \left(\frac{15}{2\pi R_0^3} \int_{\vec{r} < R_0} |(\vec{r} - \vec{r}_{\mathrm{c}}) \times \hat{n}|^2 \right)^{1/2} \qquad (19)$$

is the center- and axis-dependent hydrodynamic radius. In the limit $\Delta_r \to 0$, we numerically calculated $\theta_{\mathrm{ms}}(\Delta_r) \to \theta_{\mathrm{ms}}^0 \simeq 0.14$. As shown in Fig. 13 (c), we find that for a range of tetrahedral cluster sizes D_r saturates from below near a value $1.5 D_r^{\mathrm{Stokes}}$ for $\Delta_r \geq 1$. (D_r drops again for $\Delta_r \geq 5$ due to the inability to resolve rotations when individual rotations exceed π radians.) We chose to parameterize $\Delta_r = 1$ to ensure that rotations are saturated at the large-Δ_r limit, allowing translations to accommodate internal relaxation.

Third, we adjusted the attempt probability for translation and rotation, p_t and $p_r = 1 - p_t$, to correct for the numerical discrepancies between the predicted and measured diffusion coefficients for tetrahedral clusters (the factors 0.6 for translations and 1.5 for rotations discussed above). We achieved this by performing VMMC simulations of a tetrahedra of four spheres with various translation attempt probabilities p_t and comparing the relative diffusion coefficients to the expected relationship (see Eq. 6)

$$\frac{D_t}{D_r} = \frac{4}{3} R_{\mathrm{H}}^2. \qquad (20)$$

If the prefactors for the diffusion coefficients followed Eq. 15 and Eq. 17, Eq. 20 would be satisfied for

$$\frac{p_t}{p_r} = \frac{20}{9} \theta_{\mathrm{ms}}(\Delta_r) \left(\frac{R_0 \Delta_r}{\Delta_t} \right)^2 = c \left(\frac{R_0 \Delta_r}{\Delta_t} \right)^2. \qquad (21)$$

Inserting the $\Delta_r \to 0$ limit, $\theta_{\mathrm{ms}}(\Delta_r) \simeq 0.14$, into Eq. 21, we find $c \simeq 0.31$ and $p_t \simeq 0.6$. Instead, Fig. 13 (d) shows that $p_t = 0.7$ (corresponding to $c = 0.4$) results in better agreement with Eq. 20. We therefore fix p_t via Eq. 21 with $c = 0.4$.

Finally, we showed that fixing the parameters as above results in diffusion coefficients agreeing with the Stokes solutions (Eq. 6) for tetrahedra of various sizes and with various interaction ranges. We fixed the translational step size $\Delta_t = 4\lambda R_0$ to be proportional to the interaction range to ensure that the ratio of single-particle to whole-cluster moves be consistent across interaction ranges. We fixed $\Delta_r = 1$ and $p_t = 0.4(R_0\Delta_r/\Delta_t)^2$ as described above. Together, this parameterization fixes the time per Monte Carlo cycle,

$$t_{\mathrm{cycle}} = \frac{6}{5} \pi p_t \lambda^2 t_0, \qquad (22)$$

where t_0 is the natural Brownian time scale (Eq. 7). Fig. 13 (e) and (f) show that D_t and D_r follow the Stokes solutions (Eq. 6) over a broad range of cluster sizes and interaction ranges, in stark contrast to single-particle Monte Carlo simulations, which follow the much more strongly size-dependent free draining solutions $D_t \propto R_{\mathrm{H}}^{-3}$ and $D_r \propto R_{\mathrm{H}}^{-5}$ for small step sizes.

References

1 A. D. Dinsmore, J. C. Crocker and A. G. Yodh, *Current opinion in colloid & interface science*, 1998, **3**, 5–11.
2 D. V. Talapin, J.-S. Lee, M. V. Kovalenko and E. V. Shevchenko, *Chem. Rev.*, 2010, **110**, 389.
3 V. J. Anderson and H. N. W. Lekkerkerker, *Nature*, 2002, **416**, 811.
4 P. J. Lu and D. A. Weitz, *Annu. Rev. Condens. Matt. Phys.*, 2013, **4**, 217.
5 Y. Gao and Z. Tang, *Small*, 2011, **7**, 2133.
6 M. E. Leunissen, C. G. Christova, A.-P. Hynninen, C. P. Royall, A. I. Campbell, A. Imhof, M. Dijkstra, R. Van Roij and A. Van Blaaderen, *Nature*, 2005, **437**, 235–240.
7 A. Ivlev, H. Löwen, G. Morfill and C. P. Royall, *Complex plasmas and colloidal dispersions*, World Scientific, 2012.
8 Y. Mao, M. E. Cates and H. N. W. Lekkerkerker, *Physica A: Statistical Mechanics and its Applications*, 1995, **222**, 10–24.
9 R. Tuinier, J. Rieger and C. G. de Kruif, *Adv. Colloid Interface Sci.*, 2003, **103**, 1.
10 P. L. Biancaniello, A. J. Kim and J. C. Crocker, *Phys. Rev. Lett.*, 2005, **94**, 058302.
11 S. Y. Park, A. K. R. Lytton-Jean, B. Lee, S. Weigard, G. C. Schatz and C. A. Mirkin, *Nature*, 2008, **451**, 553.
12 D. Nykypanchuk, M. M. Maye, D. van der Lelie and O. Gang, *Nature*, 2008, **451**, 549.
13 N. Asherie, A. Lomakin and G. B. Benedek, *Physical Review Letters*, 1996, **77**, 4832–4835.
14 P. R. ten Wolde and D. Frenkel, *Science*, 1997, **277**, 1975.
15 J. F. Lutsko and G. Nicolis, *Physical Review Letters*, 2006, **96**, 046102.
16 K. G. Soga, J. R. Melrose and R. C. Ball, *J. Chem. Phys.*, 1999, **110**, 2280.
17 D. Costa, P. Ballone and C. Caccamo, *J. Chem. Phys.*, 2002, **116**, 3327.
18 A. Lomakin, N. Asherie and G. B. Benedek, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 10254.
19 A. Fortini, E. Sanz and M. Dijkstra, *Phys. Rev. E*, 2008, **78**, 041402.
20 B. J. Alder, D. A. Young and M. A. Mark, *J. Chem. Phys.*, 1972, **56**, 3013.
21 D. A. Young and B. J. Alder, *J. Chem. Phys.*, 1980, **73**, 2430.
22 H. Liu, S. Garde and S. Kumar, *J. Chem. Phys.*, 2005, **123**, 174505.
23 D. L. Pagan and J. D. Gunton, *J. Chem. Phys.*, 2005, **122**, 184515.
24 P. Charbonneau and D. R. Reichman, *Physical Review E*, 2007, **75**, 011507.
25 D. Klotsa and R. L. Jack, *Soft Matter*, 2011, **7**, 6294–6303.
26 E. K. Hobbie, *Phys. Rev. Lett.*, 1998, **81**, 3996.
27 T. H. Zhang and X. Y. Liu, *J. Am. Chem. Soc.*, 2007, **129**, 13520.
28 J. R. Savage and A. D. Dinsmore, *Phys. Rev. Lett.*, 2009, **102**, 198302.
29 R. J. Macfarlane, B. Lee, H. D. Hill, A. J. Senesi, S. Seifert and C. A. Mirkin, *Proc. Natl. Acad. Sci. U.S.A.*, 2009, **109**, 10493.
30 O. Galkin and P. G. Vekilov, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 6277.
31 O. Galkin and P. G. Vekilov, *J. Crystal Growth*, 2001, **232**, 63.
32 W. Dai, S. K. Kumar and F. W. Starr, *Soft Matter*, 2010, **6**, 6130.
33 E. Zaccarelli, *J. Phys. Condens. Matter*, 2007, **19**, 323101.
34 A. I. Campbell, V. J. Anderson, J. S. van Duijneveldt and P. Bartlett, *Phys. Rev. Lett.*, 2005, **94**, 208301.
35 F. Sciortino, P. Tartaglia and E. Zaccarelli, *J. Phys. Chem. B*, 2005, **109**, 21942.
36 P. Sollich and N. B. Wilding, *Phys. Rev. Lett.*, 2010, **104**, 118302.
37 I. Zhang, C. P. Royall, M. A. Faers and P. Bartlett, *Soft Matter*, 2013, **9**, 2076.
38 C. P. Royall, S. R. Williams, T. Ohtsuka and H. Tanaka, *Nature Mater.*, 2008, **7**, 556.
39 R. J. Macfarlane, M. R. Jones, A. J. Senesi, K. L. Young, B. Lee, J. Wu and C. A. Mirkin, *Angew. Chem. Int. Ed.*, 2010, **49**, 4589.

40 S. Whitelam and P. L. Geissler, *J. Chem. Phys.*, 2007, **127**, 154101.

41 S. Whitelam, E. H. Feng, M. F. Hagan and P. L. Geissler, *Soft Matter*, 2009, **5**, 1251–1262.

42 S. Whitelam, *Molecular Simulation*, 2011, **37**, 606.

43 T. K. Haxton and S. Whitelam, *Soft Matter*, 2012, **8**, 3558.

44 J. D. Honeycutt and H. C. Andersen, *J. Phys. Chem.*, 1987, **91**, 4950.

45 J. D. Schmit, S. Whitelam and K. Dill, *J. Chem. Phys.*, 2011, **135**, 085103.

46 F. J. Martinez-Veracoechea, B. M. Mladek, A. V. Tkachenko and D. Frenkel, *Phys. Rev. Lett.*, 2011, **107**, 045902.

47 H. Tanaka and T. Araki, *Physical review letters*, 2000, **85**, 1338.

48 M. E. Cates, K. Stratford, R. Adhikari, P. Stansell, J. C. Desplat, I. Pagonabarraga and A. J. Wagner, *Journal of Physics: Condensed Matter*, 2004, **16**, S3903.

49 http://uk.mathworks.com/matlabcentral/fileexchange/30371-fcc-cluster.

50 http://uk.mathworks.com/matlabcentral/fileexchange/30372-bcc-cluster.

51 A. Z. Panagiotopoulos and M. R. Stapleton, *Fluid Phase Equilibr.*, 1989, **53**, 133.

52 T. Ihle and D. M. Kroll, *Physical Review E*, 2001, **63**, 020201.

53 C. M. Pooley and J. M. Yeomans, *The Journal of Physical Chemistry B*, 2005, **109**, 6505–6513.

54 J. Sané, J. T. Padding and A. A. Louis, *Physical Review E*, 2009, **79**, 051402.

55 K. Kikuchi, M. Yoshida, T. Maekawa and H. Watanabe, *Chemical Physics Letters*, 1991, **185**, 335–338.

56 L. O. Hedges, `http://vmmc.xyz`.

57 A. W. Wilber, J. P. K. Doye, A. A. Louis, E. G. Noya, M. A. Miller and P. Wong, *J. Chem. Phys.*, 2007, **127**, 085106.

58 M. F. Hagan and D. Chandler, *Biophys. J.*, 2006, **91**, 42.

59 S. Whitelam and R. L. Jack, *Annual Review of Physical Chemistry*, 2015, **66**, null.

60 N. C. Karayiannis, R. Malshe, J. J. de Pablo and M. Laso, *Phys. Rev. E*, 2011, **83**, 061505.

61 N. C. Karayiannis, R. Malshe, M. Kröger, J. J. de Pablo and M. Laso, *Soft Matter*, 2012, **8**, 844.

62 J. P. K. Doye, D. J. Wales and R. S. Berry, *J. Chem. Phys.*, 1995, **103**, 4234.

63 N. Arkus, V. N. Manoharan and M. P. Brenner, *Phys. Rev. Lett.*, 2009, **103**, 118303.

64 N. Arkus, V. N. Manoharan and M. P. Brenner, *SIAM J. Discrete Math.*, 2011, **4**, 1860.

65 R. S. Hoy and C. S. O'Hern, *Phys. Rev. Lett.*, 2010, **105**, 068001.

66 R. S. Hoy, J. Harwayne-Gidansky and C. S. O'Hern, *Phys. Rev. E*, 2012, **85**, 051403.

67 R. S. Hoy, *Phys. Rev. E*, 2015, **91**, 012303.

68 M. C. Holmes-Cerfon, submitted (available online at arxiv.org/abs/1407.3285).

69 D. Frenkel, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, **101**, 17571.

A thorough computational investigation of colloidal crystallization across parameter space illustrates several distinct mechanisms for crystallization and dynamical arrest.