ACSNANO

www.acsnano.org

ARTICLE

An Artificial Neural Network Reveals the Nucleation Mechanism of a Binary Colloidal AB_{13} Crystal

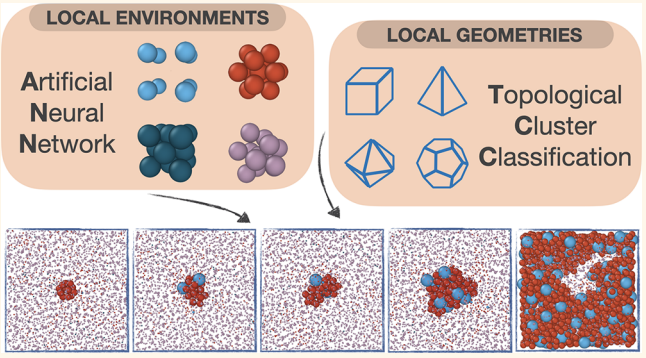
Gabriele M. Coli* and Marjolein Dijkstra*

Cite This: *ACS Nano* 2021, 15, 4335−4346

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: **Colloidal suspensions of two species have the ability to form binary crystals under certain conditions. The hunt for these functional materials and the countless investigations on their formation process are justified by the plethora of synergetic and collective properties these binary superlattices show. Among the many crystal structures observed over the past decades, the highly exotic colloidal *icosahedral* AB_{13} crystal was predicted to be stable in binary hard-sphere mixtures nearly 30 years ago, yet the kinetic pathway of how homogeneous nucleation occurs in this system is still unknown. Here we investigate binary nucleation of the AB_{13} crystal from a binary fluid phase of nearly hard spheres. We calculate the nucleation barrier and nucleation rate as a function of supersaturation and draw a comparison with nucleation of single-component and other binary crystals. To follow the nucleation process, we employ a neural network to identify the AB_{13} phase from the binary fluid phase and the competing fcc crystal with single-particle resolution and significant accuracy in the case of bulk phases. We show that AB_{13} crystal nucleation proceeds *via* a coassembly process where large spheres and icosahedral small-sphere clusters simultaneously attach to the nucleus. Our results lend strong support for a classical pathway that is well-described by classical nucleation theory, even though the binary fluid phase is highly structured and exhibits local regions of high bond orientational order.**



KEYWORDS: *colloidal particles, nanoparticles, nucleation, crystallization, neural network, machine learning, computer simulations*

INTRODUCTION

Understanding crystallization is important in many research fields such as protein crystallization for resolving the molecular structure, drug design in the pharmaceutical industry, ice crystal formation in clouds for weather forecasts, and crystallization of colloidal and nanoparticle suspensions with application perspectives in catalysis, optoelectronics, and plasmonics. Hence, it is not surprising that over the past decades many experimental and simulation studies have been devoted to studying crystal nucleation in a fluid of hard spheres, which is indisputably one of the simplest possible model systems to describe colloidal and nanoparticle systems and serves as a reference for systems with more complicated interactions, *e.g.*, depletion and electrostatic interactions.

Nucleation describes the process in which a crystal nucleus spontaneously forms due to a statistical fluctuation in the metastable fluid phase. Despite the significant amount of work spent on understanding crystal nucleation in hard spheres, the mechanism by which a hard-sphere fluid transforms into a crystal phase remains to be settled. Several scenarios such as a classical one-step crystallization process, a nonclassical two-step crystallization mechanism with precursors consisting of local regions with either high density, high bond-orientational order, or competing orders, or a spinodal-like process have been proposed, but all of these crystallization mechanisms are still heavily debated.[1−10]

Another unresolved issue is the huge discrepancy in crystal nucleation rate between simulations and experiments at low supersaturation. By employing a wide variety of simulation techniques, of which some of them even include the effect of hydrodynamics, a much lower nucleation rate was consistently found in simulations compared to experimental results.[11−18]

Received: September 7, 2020
Accepted: February 18, 2021
Published: February 23, 2021

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346

To enhance the structural diversity and functional composition of the self-assembled structures, one may resort to binary mixtures of large and small colloidal hard spheres with diameters σ_L and σ_S, respectively. Although the number of distinct binary crystal structures is relatively small, we wish to remark here that the structural diversity can be enhanced significantly by taking into account varying interaction potentials; for example, suspensions consisting of two types of particles with opposite charges can form a dazzling variety of binary superlattice structures.[19,20] In this work, we focus on binary mixtures of particles interacting with hard-sphere-like potentials as they serve as a reference for a wider class of soft repulsive interaction potentials, mimicking the interactions of many nanoparticle systems.

The phase behavior of binary hard-sphere mixtures is well studied by now and display a wide variety of behaviors ranging from a spindle-type to azeotropic and eutectic phase diagram, wide coexistence regions between phases with different compositions, pure single-component crystals, substitutionally disordered crystalline phases, interstitial solid solutions, and various binary crystal structures with different stoichiometries $x_L = N_L/(N_L + N_S)$, with N_L (N_S) denoting the number of large (small) particles.[21] Depending on the diameter ratio $q = \sigma_S/\sigma_L$, binary hard-sphere systems exhibit entropically stabilized binary superlattice structures analogous to their atomic counterparts NaCl ($0.2 \leq q \leq 0.42$), AlB_2 ($0.42 \leq q \leq 0.59$), $NaZn_{13}$ ($0.48 \leq q \leq 0.62$), and the Laves phases ($0.74 \leq q \leq 0.84$).[21]

The most intriguing structure of the above-mentioned binary crystals is without any doubt the $NaZn_{13}$, also termed the icosahedral AB_{13} structure in order to distinguish it from the cuboctahedral AB_{13} structure,[22] which has been found to be metastable due to a less efficient packing of the small spheres in the case of binary hard-sphere mixtures.[23,24] The stability of the AB_{13} structure has gained much attention because of its bizarre lattice. The large spheres are unusually distant from each other, as shown in Figure 1, and are arranged on a simple cubic lattice,

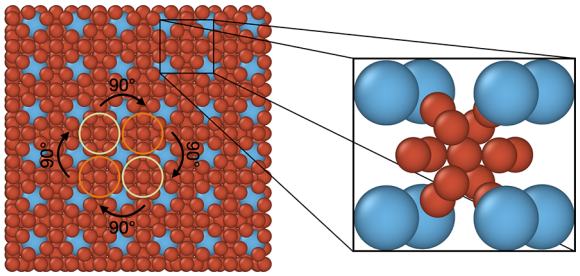


Figure 1. Icosahedral-AB_{13} structure consists of a simple cubic lattice of large (A) spheres and icosahedral clusters of 13 small (B) spheres denoted in blue and red, respectively. The icosahedral clusters of 13 small spheres are rotated by 90° with respect to their neighboring clusters as indicated by the differently colored circles. The zoom-in displays the icosahedral cluster formed by the small spheres inside a simple cubic subunit cell of large spheres.

which is a highly unusual crystal structure in the case of plain hard spheres. More intriguingly, each unit cell of this simple cubic lattice of large spheres contains an icosahedral cluster of 13 small spheres, which are all rotated by 90° with respect to their neighboring icosahedral clusters. Hence, the full unit cell of an icosahedral AB_{13} structure consists of 8 unit cells of this simple cubic lattice of large spheres with 8 icosahedral clusters of 13 small spheres in their centers, resulting in 112 particles in total.

The colloidal analogue of the $NaZn_{13}$ was for the first time observed by Sanders *et al.* in natural gem opals consisting of two sizes of silica spheres in 1987.[25,26] The same AB_{13} structure was later observed in systems of charged-stabilized colloids or PMMA particles[27−31] and in various nanoparticle systems, *e.g.*, mixtures of semiconductor, metal oxide, magnetic, silica, and polymer-grafted nanoparticles, as well as polyoxometalate clusters.[20,22,32−43]

In contrast to the considerable amount of work that has been devoted to studying crystal nucleation in single-component hard-sphere fluids, only a few studies have been focused on crystal nucleation in binary mixtures. Crystallization in fluid mixtures is generally much harder than in single-component systems. Spindle-, azeotropic-, eutectic-like phase transitions in binary systems usually involve fractionation, as the composition of the solid phase deviates from that of the supersaturated phase. Fractionation is known to slow down the rate of crystallization.[44,45] Additionally, the surface tension of the solid−fluid interface will increase when the compositions of the fluid and solid phase deviate substantially, leading to higher nucleation barriers and lower nucleation rates.[46,47] Moreover, nucleation of a binary (compound) crystal is believed to be orders of magnitude slower than that of pure crystals or substitutionally disordered crystalline phases due to a loss of mixing entropy, making binary crystal nucleation an extremely rare event.[44] This is one of the reasons that the number of simulation and experimental nucleation studies on binary colloidal crystals is very limited.

Yet, the few simulation studies on binary crystal nucleation revealed several interesting observations. For instance, simulations on homogeneous nucleation of a binary AB_2 crystal in a mixture of hard spheres revealed that in the case of multiple competing crystal structures the phase that nucleates is the one whose composition is closest to that of the fluid phase even when it is metastable.[44,47] In addition, it was found by simulations that kinetic barriers also play an important role in determining which crystal phase nucleates. In the case of oppositely charged colloids it was found that the disordered fcc phase that nucleates is metastable and has a higher free-energy barrier for nucleation than the thermodynamically stable binary CsCl crystal.[48] In this case the disordered fcc phase was favored by nonequilibrium nucleation. These results greatly challenge the commonly held assumption that subcritical clusters are always in quasi-equilibrium with the fluid phase.[49] Another simulation study showed that homogeneous nucleation of an interstitial solid solution in a binary mixture of hard spheres is driven by the nucleation of large spheres into an fcc crystal while maintaining chemical equilibrium of the small spheres throughout the system. More recent simulations showed that nucleation of Laves phases is severely suppressed by the presence of icosahedral clusters in a binary hard-sphere mixture, but that softness of the interaction potential reduces the degree of 5-fold symmetry in the binary fluid and enhances crystallization.[50] Finally, we also mention for completeness that spontaneous spinodal-like crystallization of structures isostructural to AlB_2, $NaZn_{13}$, and the Laves phases has been observed in simulations on highly supersaturated binary hard-sphere fluids with and without unphysical moves that swap the identities of large and small spheres.[51]

All nucleation studies share a common challenge: being able to recognize different phases starting from the raw particle coordinates of the system. In particular, one requires a criterion that is able to distinguish on a single-particle level particles

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346

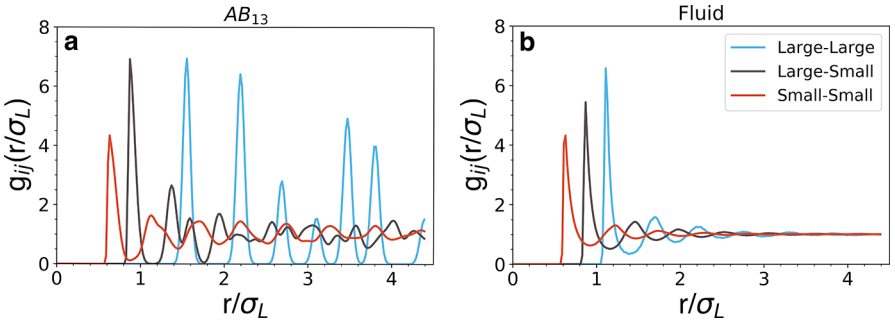


Figure 2. Pair correlation functions $g_{ij}(r/\sigma_L)$ of (a) the AB_{13} crystal phase and (b) the binary fluid phase at composition $x_L = 1/14$ with $i, j \in \{L, S\}$ denoting the large (L) and small (S) species, for a mixture of WCA spheres at $k_BT/\epsilon = 0.025$ to mimic hard spheres and at coexistence pressure $\beta P_{coex}\sigma_L^3 = 45.35$. The large−large pair correlation function $g_{LL}(r)$ of the AB_{13} shows that the large spheres are unusually distant from each other in comparison with the binary fluid phase. The small−small pair correlation function $g_{SS}(r)$ of the AB_{13} phase demonstrates that the small spheres exhibit fluid-like behavior, making it difficult to distinguish the binary fluid phase and the AB_{13} crystal on the basis of the small spheres.

belonging to the growing phase from those of the metastable parent phase. Most crystal nucleation studies are based on describing the local environment of each particle in terms of the so-called bond orientational order parameters, *i.e.*, rotational invariant combinations of the spherical harmonics of degree l, as introduced by Steinhardt *et al.* in 1983.[52] Specifically, the 4-fold and 6-fold bond order parameters, q_4 and q_6, suffice to distinguish the crystalline particles from the fluid-like particles, as most crystals exhibit either cubic and/or hexagonal symmetry. In the case of binary crystals, the local environment of each particle of each species can deviate substantially from cubic and hexagonal symmetry, and other symmetries should be taken into account in identifying the different crystal phases. Here, we describe the local environment of a particle by using a full expansion in spherical harmonics, and we train an artificial neural network to identify the different phases on a single-particle level using a set of bond order parameters up to degree $l = 12$ as input. We demonstrate the effectiveness of this method by studying nucleation of the AB_{13} crystal structure in binary mixtures of nearly hard spheres using simulations. We show that standard techniques fail in identifying the different phases and that machine learning is useful in achieving this goal.

Employing the trained neural network as an order parameter, we investigate how an AB_{13} crystal nucleates and grows and we shed light on the formation mechanism during the early stages. In addition, we study how icosahedral clusters of small spheres arrange themselves inside this simple cubic lattice and whether the growth of the binary nucleus proceeds *via* the attachment of individual small spheres, small clusters, or perfect or defective icosahedral clusters of small spheres. Specifically, the role of the icosahedral clusters on the AB_{13} nucleation is intriguing, as the presence of 5-fold clusters is often attributed to glassy dynamics and suppression of crystallization.[53] However, in ref 23, it was conjectured that the abundance of icosahedral clusters in both the fluid and AB_{13} crystal and thus the structural similarity of these two phases may result in an ultralow surface tension and hence a low nucleation barrier and high nucleation rate. To investigate this, we determine the nucleation barrier height and the nucleation rate using the seeding approach, and we compare our results with crystal nucleation in pure hard spheres and with nucleation of the Laves phases. Finally, we analyze the kinetic pathways of the spontaneous crystallization events of the AB_{13} phase in binary mixtures of hard spheres using brute force molecular dynamics (MD) simulations.

RESULTS AND DISCUSSION

The Model. We consider a binary mixture of N_L large (L) hard spheres with a diameter σ_L and N_S small (S) hard spheres with a diameter σ_S. For a binary hard-sphere (BHS) mixture, the AB_{13} phase is thermodynamically stable for a diameter ratio $q = \sigma_S/\sigma_L \in [0.54, 0.61]$.[21] In this work, we set $q = 0.55$. The particles interact *via* a Weeks−Chandler−Andersen (WCA) pair potential, which can straightforwardly be employed in MD simulations and which reduces to the hard-sphere potential in the limit that the temperature $T \to 0$. The WCA potential $u_{\alpha\beta}(r_{ij})$ between species $\alpha, \beta \in \{L, S\}$ reads[54]

$$u_{\alpha\beta}(r_{ij}) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma_{\alpha\beta}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r_{ij}}\right)^{6} + \frac{1}{4}\right] & r_{ij} < 2^{1/6}\sigma_{\alpha\beta} \\ 0 & r_{ij} \geq 2^{1/6}\sigma_{\alpha\beta} \end{cases} \quad (1)$$

with $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ the center-of-mass distance between particle i and j, $\mathbf{r}_i$ the position of particle i, ϵ the interaction strength, and $\sigma_{\alpha\beta} = (\sigma_\alpha + \sigma_\beta)/2$. The steepness of the repulsion between the particles can be tuned by the temperature k_BT/ϵ. We set $k_BT/\epsilon = 0.025$, which has been used extensively in previous simulation studies to mimic hard spheres.[3,13,16,55]

Using free-energy calculations in Monte Carlo (MC) simulations, we determine phase coexistence between the AB_{13} crystal and the binary fluid phase with the same composition as that of the AB_{13} crystal (see the Methods section). The pressure at which the crystal and fluid are at coexistence reads $\beta P_{coex}\sigma_L^3 = 45.35$. To study nucleation of the AB_{13} crystal, we perform simulations at pressures $P > P_{coex}$ in the regime where the fluid phase is metastable with respect to the crystal phase.

Local Structure Detection. *Bond Order Parameters.* In order to follow the nucleation process of the AB_{13} phase, we need to find a way to detect an embryo of the stable AB_{13} crystal structure in the supersaturated binary fluid phase with single-particle resolution. In many simulation studies, local bond orientational order parameters have been used to study crystal nucleation.[52,56,57] To calculate these local bond order parameters, we first have to define the local environment of particle i by determining a list of neighbors using, for instance, a distance criterion based on the first minimum of the pair

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346

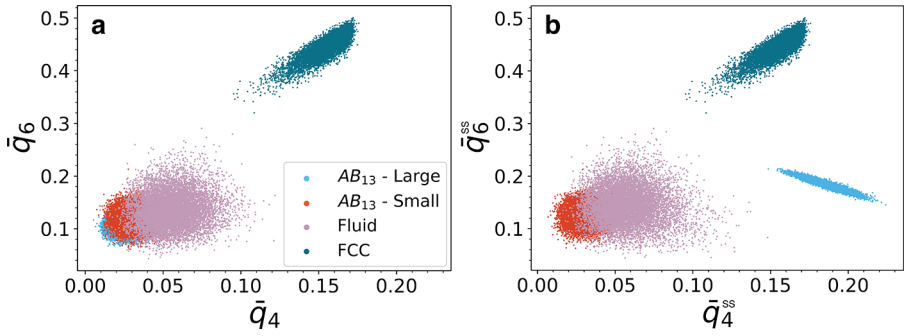


Figure 3. Scatter plot in the averaged bond order parameter $\bar{q}_4$−$\bar{q}_6$ plane for the four local particle environments we wish to distinguish: large species (light blue) and small species (red) of the AB_{13} phase, binary fluid phase (light purple), and the pure fcc phase (dark blue). (a) Averaged bond order parameters calculated using the solid-angle-based nearest-neighbor criterion irrespective of particle species, showing significant overlap of 3 of the 4 local structures, thereby making the classification impossible. (b) Averaged bond order parameters calculated by taking into account only the neighbors belonging to the same species. The distribution of the small species of the AB_{13} phase still overlaps with that of the binary fluid phase.

correlation function or by employing a Voronoi construction. The set of distance vectors between particle i and its neighbors is then expanded in spherical harmonics of order l. Finally, the quadratic and cubic rotationally invariant quantities, q_l and w_l, are defined to measure the local symmetry of bonds of particle i; see the Methods section. Due to the cubic or hexagonal symmetry of most crystals, the 4-fold and 6-fold bond order parameters, q_4 and q_6, have been extensively employed in the literature to study crystal nucleation.

We first study whether the 4-fold and 6-fold bond order parameters can be used to distinguish the AB_{13} crystal from the binary fluid phase with a composition $x_L = N_L/(N_L + N_S) = 1/14$ and from the pure fcc phase. For this purpose, we use the averaged bond order parameters $\bar{q}_l$, thereby taking into account also the second shell of neighbors of a particle.[58] We perform MC simulations in the isobaric−isothermal ensemble; that is, we fix the pressure P, the temperature T, and the number of particles $N = N_L + N_S$. We carry out bulk simulations of the AB_{13} crystal and the binary fluid at coexistence pressure $\beta P_{coex}\sigma_L^3 = 45.35$ and of the pure fcc crystal at $\beta P_{coex}\sigma^3 = 8.87$ corresponding to the pressure at bulk coexistence with the single-component fluid phase.

In Figure 2, we plot the pair correlation functions $g_{ij}(r)$ of the AB_{13} crystal and the binary fluid phase with $i, j \in \{L, S\}$ denoting the large (L) and small (S) species. We make the following noteworthy observations. We first observe from the small−small pair correlation function $g_{SS}(r)$ of the AB_{13} phase that the small spheres exhibit fluid-like behavior even though they are in a solid state. The structural similarity of the small spheres in the binary fluid and the AB_{13} phase makes it difficult to distinguish the two phases on the basis of the local symmetry of the small spheres. Additionally, we observe that the main peak of the large−large pair correlation function $g_{LL}(r)$ of the AB_{13} crystal is at a unusually large distance in comparison with that of the binary fluid phase. In order to nucleate the AB_{13} phase in the binary fluid phase, the large spheres have to be pushed away from each other to much larger distances to make room for the icosahedral clusters of small spheres. In addition, the huge difference in the position of the main peak of the $g_{LL}(r)$ of the binary fluid and the AB_{13} phase complicates the identification of neighboring particles on the basis of a simple cutoff distance.

To circumvent this problem, we use the parameter-free solid-angle-based nearest-neighbor (SANN) algorithm to identify the neighbors of each particle.[59] Using the SANN algorithm, we measure the 4-fold and 6-fold averaged bond order parameters, $\bar{q}_4$ and $\bar{q}_6$, and we show scatter plots in the $\bar{q}_4$−$\bar{q}_6$ plane for the large and small species of the AB_{13} phase, the binary fluid phase, and the fcc phase in Figure 3. We observe from Figure 3a that the distributions for the large and small species of the AB_{13} phase and the binary fluid phase overlap, making the distinction between the different phases very hard. To improve the separation of the bond order parameter distributions, we calculate $\bar{q}_4$ and $\bar{q}_6$ by taking into account only the neighbors of the same species as the particle of interest in the SANN algorithm. We plot the results in Figure 3b and observe that the distribution of the large species in the AB_{13} is well separated from the other structures. However, the distribution of the small species of the AB_{13} phase still overlaps with that of the binary fluid phase. We thus find that it remains a major challenge to correctly classify the small species of the AB_{13} phase, which can be misidentified as fluid particles due to their icosahedral arrangement.

Feed forward Neural Network. In order to overcome this problem, we describe the local environment of a particle using a full expansion in spherical harmonics, and we train an artificial neural network (ANN) to identify distinct structures on a single-particle level, thereby extending the approach of ref 60. The main goal of the ANN here is to detect the birth of a crystal nucleus of the AB_{13} structure in a binary fluid phase, which presents additional difficulties with respect to the identification of bulk phases[60] due to the solid−fluid interfaces of the crystal nuclei. Moreover, the identification of crystalline particles and estimating the number of crystalline particles are crucial for determining the barrier height and the nucleation rate using the seeding approach.[14,15] To minimize the effect of interfaces, we employ the standard, instead of the averaged coarse-grained, bond order parameters, thereby increasing the spatial resolution at the expense of the accuracy in the local structure detection. The idea behind the choice of nonaveraged bond order parameters is as follows: given the great predictive capacity of neural networks, due to the extremely effective way of combining information from many features and using nonlinear functions, we can use *less precise* but *more local* descriptors. A high accuracy can be reached thanks to the versatility of the neural network, and in this way it is possible to obtain, for each particle, an estimate of the class based only on the *first shell* of neighbors.

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346

We employ the bulk simulations of the AB_{13} phase, the pure fcc phase, and the binary fluid as described in the section Bond Order Parameters, and build a training set of 10^5 training samples for each of the local particle environments we wish to distinguish: large particles of the AB_{13} phase, small species in the AB_{13} phase, particles in a pure fcc phase, and particles irrespective of species in a binary fluid. We describe the local environment of each particle i by a 36-dimensional input vector of bond order parameters:

$$I(i) = (\{q_l(i)\}, \{w_{l'}(i)\}, \{q_l^{ss}(i)\}, \{w_{l'}^{ss}(i)\}) \quad (2)$$

where $l \in [1, 12]$ and l' varies in the same range but only assumes even values. The superscript ss means that bond order parameters are calculated by considering only particles of the same species as particle i.

In ref 60, a single-layer ANN, *i.e.*, only an input and output layer and no hidden layers, was employed to successfully classify the AB_{13} phase from a binary fluid phase with a composition $x_L = 1/3$ using the *averaged* bond order parameters as input. However, this network architecture with the *averaged* bond order parameters as input vector is not accurate enough to distinguish the AB_{13} phase from the binary fluid with a composition $x_L = 1/14$ equal to the stoichiometry of the AB_{13} phase, which is mostly due to the structural similarity of the small spheres in the AB_{13} and the fluid phase, both exhibiting an abundance of (defective) icosahedral clusters. In addition, the standard *nonaveraged* bond order parameters that we employ offer a poorer characterization of the bulk phases with respect to their averaged counterparts. In order to improve the accuracy of the classification, we add hidden layers to our neural network.

To be more specific, we employ a fully connected neural network with two hidden layers consisting of 72 neurons. Each neuron uses a rectified linear unit (ReLU) activation function to guarantee fast convergence and good generalization. The output layer has four neurons, corresponding to the four distinct local particle environments (classes) we wish to distinguish, and is activated with a Softmax function (Figure 4). The network was trained using stochastic gradient descent and L_2 regularization.[61] We employ 20% of the samples as validation data to predict the accuracies for each output node corresponding to the four different particle environments (see the Methods section). The accuracies related to each specific class are shown in Table 1.

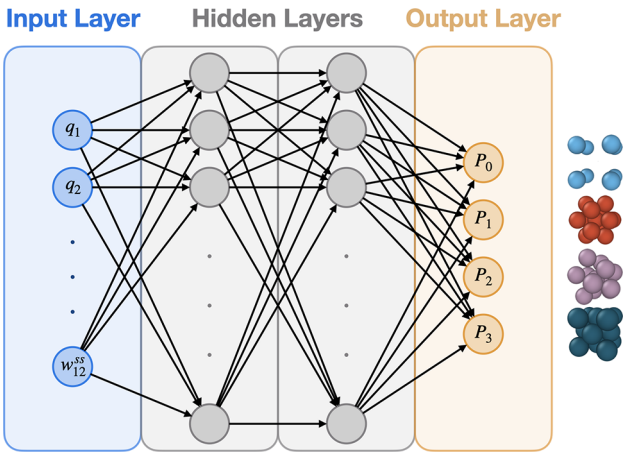


Figure 4. Architecture of our fully connected artificial neural network (ANN). The input layer has 36 units, as described by eq 2, while both hidden layers contain 72 neurons. The output layer consist of four neurons, yielding the probability that a particle corresponds to a certain class.

Table 1. Accuracies of the ANNs on the Validation Set Calculated for All Four Classes

class	accuracy
AB_{13}, large	100.0%
AB_{13}, small	98.1%
fluid	97.8%
fcc	99.4%

Seeding Approach. Numerical simulations have helped in elucidating nucleation for a plethora of model systems, but despite the possibility of following each single particle during the crystallization process, they suffer from an important drawback. In fact, the accessible time scales in MC or MD simulations are typically much shorter than in experiments. This is particularly disadvantageous for nucleation studies, as the birth of a crystalline nucleus in a metastable fluid is a rare event.

For this reason, it is often necessary to use special sampling schemes in simulations like umbrella sampling (US),[12,62−64] forward flux sampling (FFS),[12,65] metadynamics,[66,67] or transition path sampling.[68−70] These techniques are mainly employed to enhance or bias the sampling of the system in order to observe rare events such as nucleation. However, these simulation techniques are extremely expensive from a computational point of view, restricting nucleation studies to highly metastable conditions.

Recently, another technique has been proposed to study nucleation, known as the seeding approach.[14,15] The great merit of this technique is that it enables the determination of all relevant physical quantities to describe nucleation, *e.g.*, barrier height and nucleation rate, and that it divides the simulation study into short simulations with a standard and unbiased sampling of phase space. Moreover, the computational cost of these simulations is moderate, which allows studying nucleation under weakly metastable conditions where the critical nucleus consists of several thousands of particles. The seeding technique involves the following steps: (1) inserting a seed of the crystal structure of interest in a metastable fluid and running simulations to equilibrate the interface while keeping the crystalline particles fixed, (2) releasing this constraint and equilibrating the full system at a sufficiently high pressure that the seed does not melt, and finally (3) running simulations of this carefully equilibrated system for a wide range of pressures in order to determine the critical pressure P_c at which the probability that the seed will grow or melt will be equal, while for $P < P_c$ the seeds will predominately melt, or grow for $P > P_c$. In order to avoid finite-size effects, we perform simulations in the *NPT* ensemble. An illustration of the last step is shown in Figure 5, where we plot the size of the largest cluster N_{CL} as a function of time t/τ_{MD}, for 10 independent simulations, at pressure $\beta P\sigma_L^3 =$ 51.1, and 51.3. Here, $\tau_{MD} = \sigma_L\sqrt{m/k_BT}$ denotes the MD time unit and m the mass of the particles. In Figure 5b (5d) the majority of the simulations show a growing (melting) cluster, which means that the pressure $\beta P\sigma_L^3 = 51.3(50.9)$ is higher (lower) than the critical one. In Figure 5a, we observe that at $\beta P\sigma_L^3 = 51.1$ the cluster melts and grows with equal probability, indicating that this value corresponds to the critical pressure for a critical cluster size $N_c = 770$. The number of particles belonging to the main cluster is determined using the neural-

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346

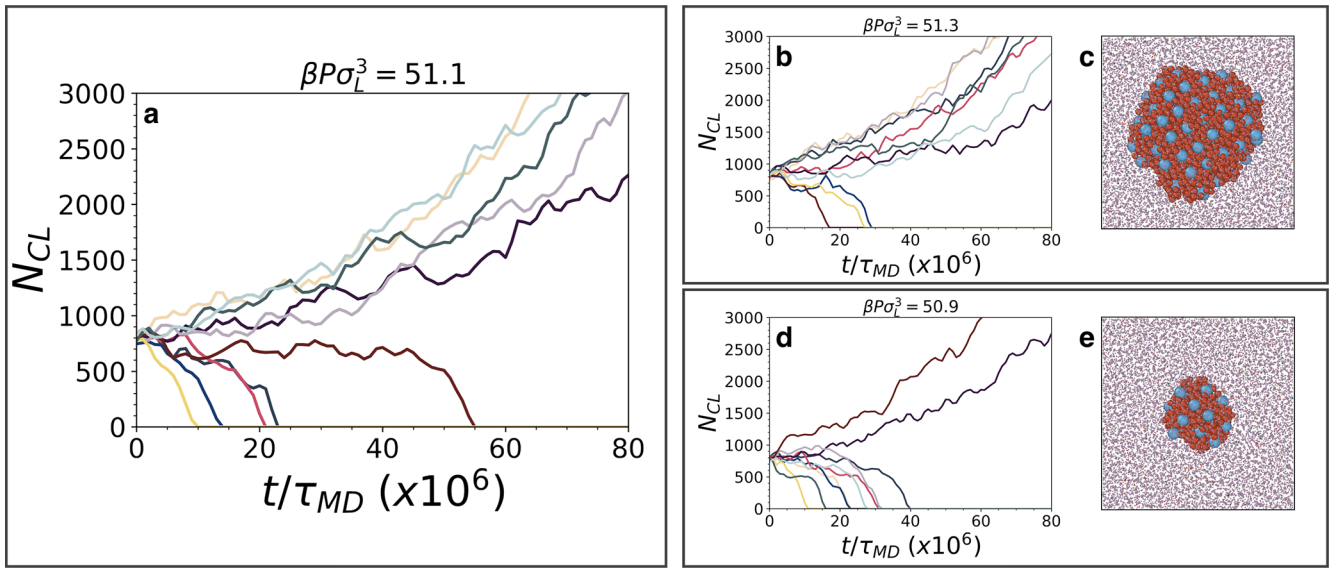


Figure 5. (a) Largest cluster size N_{CL} of the AB_{13} phase as recognized by the ANN as a function of time t/τ_{MD} using the seeding approach in 10 independent MD simulations of a binary mixture of WCA spheres at temperature $k_BT/\epsilon = 0.025$ to mimic hard spheres with a diameter ratio $q = 0.55$ in the *NPT* ensemble, composition $x_L = 1/14$, and at a pressure (a) $\beta P\sigma_L^3 = 51.1$, where the cluster melts and grows with equal probability, indicating that this pressure value corresponds to the critical pressure for this cluster size, (b) $\beta P\sigma_L^3 = 51.3$, where the cluster grows in the majority of the simulations, and (d) $\beta P\sigma_L^3 = 50.9$, at which the cluster melts in most of the simulations. Typical configurations of the growth and melting of the cluster are shown in (c) and (e), respectively. Note that the size of fluid-like and fcc-like particles is reduced for visual clarity.

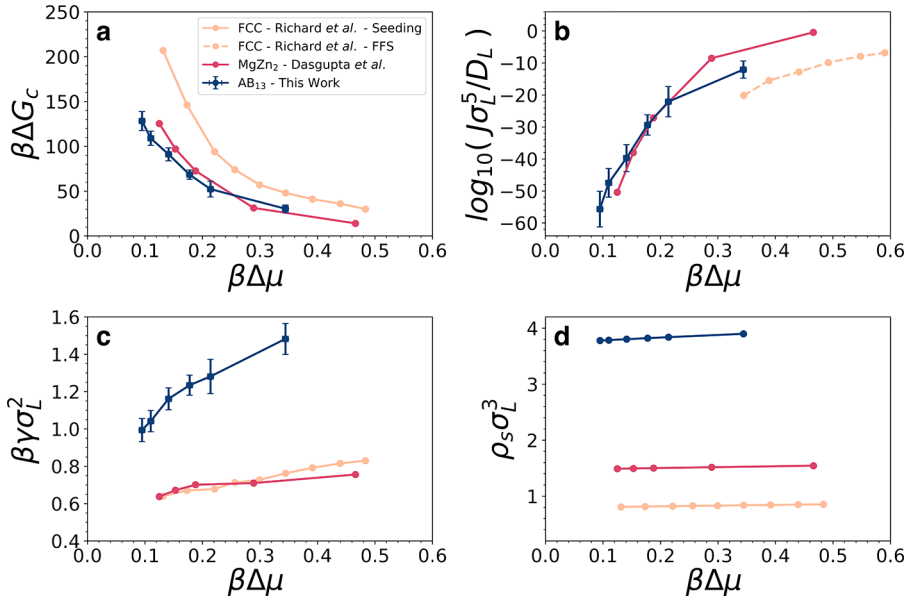


Figure 6. (a) Height of the Gibbs free-energy barrier $\beta\Delta G_c$, (b) nucleation rate $J\sigma_L^5/D_L$, (c) interfacial free energy $\beta\gamma\sigma_L^2$, and (d) the crystal number density $\rho_s\sigma_L^3$ as a function of the chemical potential difference $\beta\Delta\mu$ between the fluid and the AB_{13} phase of a binary WCA mixture at $k_BT/\epsilon = 0.025$ to mimic hard spheres, with a diameter ratio $q = 0.55$ and composition $x_L = 1/14$ as obtained from the seeding approach. For comparison, we also plot the results on nucleation of the Laves phase in a binary hard-sphere mixture from ref 50 and of the fcc phase in a fluid of pure hard spheres from refs 16 and 55.

network-based order parameter as described in the section Feed forward Neural Network together with a clustering algorithm to identify clusters of mutually bonded solid particles.

Subsequently, several physical quantities can be calculated using classical nucleation theory (CNT), such as the height of the Gibbs free-energy barrier ΔG_c using

$$\Delta G_c(N_c) = N_c\Delta\mu/2 \tag{3}$$

and the nucleation rate J,

$$\frac{J\sigma_L^5}{D_L} = \sqrt{\frac{\beta\Delta\mu}{6\pi N_c}\frac{f^+\sigma_L^2}{D_L}\rho_f\sigma_L^3\exp(-\beta\Delta G_c)} \tag{4}$$

where N_c denotes the critical nucleus size, $\beta\Delta\mu$ is the supersaturation, *i.e.*, the difference in chemical potential between the supersaturated fluid and the stable crystal phase, $f^+ = \langle(N(t) - N_c)^2\rangle/t$ is the attachment rate of particles to the critical cluster, t is the time, $\rho_f(\beta P_c\sigma_L^3)$ is the critical density of the fluid at the critical pressure P_c, and D_L is the long-time diffusion coefficient

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346

Table 2. Values of the Most Significant Variables Involved in the Seeding Approach Calculations[a]

N_c	N	$\beta\Delta\mu$	$\beta P_c\sigma_L^3$	$\rho_f\sigma_L^3$	$\rho_s\sigma_L^3$	$\beta\Delta G_c$	$\beta\gamma\sigma_L^2$	$\log_{10}(J\sigma_L^5/D_L)$
2706	40334	0.095	48.40	3.383	3.778	128.3	0.994	−55.66
1977	29204	0.110	48.90	3.390	3.786	108.9	1.042	−47.47
1290	19376	0.141	49.90	3.405	3.802	91.01	1.161	−39.72
770	13692	0.178	51.10	3.423	3.820	68.38	1.234	−29.37
488	8988	0.214	52.30	3.440	3.838	52.22	1.281	−22.10
176	4746	0.344	56.70	3.499	3.897	30.29	1.482	−12.08

[a]Each row corresponds to a different critical nucleus size N_c. See the main text for the meaning of each variable.

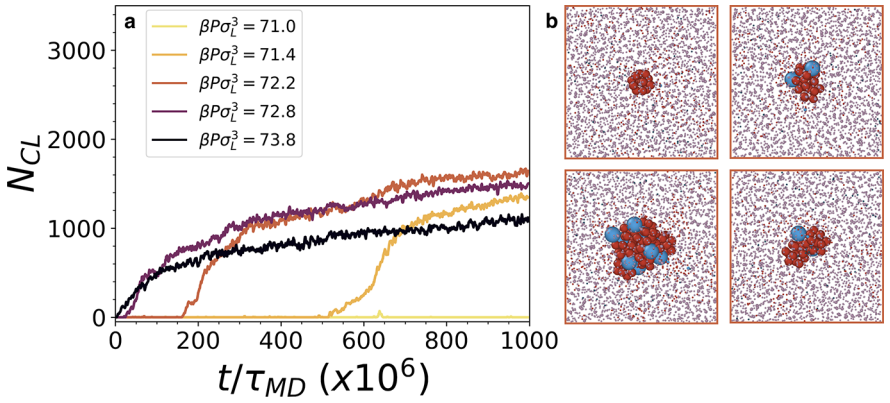


Figure 7. (a) Size of the largest AB_{13} cluster N_{CL} as recognized by the ANN for a binary mixture of WCA spheres at $k_BT/\epsilon = 0.025$ to mimic hard spheres, with a diameter ratio $q = 0.55$ and composition $x_L = 1/14$ as a function of time t/τ_{MD} for five different pressures $\beta P\sigma_L^3$ using MD simulations in the *NPT* ensemble. (b) Four configurations of a spontaneous nucleation event at $\beta P\sigma_L^3 = 72.2$, showing a time sequence of the early stages of AB_{13} nucleation with time increasing from the upper-left corner and then proceeding clockwise. Particles that do not belong to the main crystalline cluster have been reduced in size for visual clarity.

at the same ρ_f.[14,15] The attachment rate f^+ is measured from 10 independent simulation trajectories at density ρ_f. Assuming, on average, a spherical cluster shape, the crystal-fluid interfacial free energy γ can be calculated from

$$N_c = \frac{32\pi\gamma^3}{3\rho_s^2\Delta\mu^3} \tag{5}$$

with ρ_s the density of the solid phase. We note that these equations rely on the validity of CNT, which will be checked and proven in the section Spontaneous Nucleation. We emphasize that all variables computed through the seeding approach using eqs 3, 4, and 5 are sensitive to the numerical value of N_c and thus to an estimate of the nucleus interface. This particular estimate is problematic for all classification algorithms and can, in principle, lead to systematic errors when evaluating the aforementioned variables. In the Supporting Information we show a detailed analysis of the performance of the ANN with respect to the interface detection.

Using different seed sizes in the seeding approach, we determine ΔG_c, J, γ, and ρ_s for different critical pressures corresponding to different supersaturations $\Delta\mu$. We plot our results as a function of $\Delta\mu$ in Figure 6 and present the numerical data in Table 2. For comparison, we also plot the results of previous simulation studies on the nucleation of the Laves phase in a binary hard-sphere mixture[50] and of the fcc phase in a fluid of pure hard spheres.[16,55]

Comparing the nucleation barrier ΔG_c for the three different phases at the same thermodynamic driving force $\Delta\mu$, we clearly observe from Figure 6a that ΔG_c is consistently higher for the fcc phase than for the AB_{13} crystal and the Laves phases. This is in contradiction with the assumption that the nucleation barrier for binary nucleation should be higher due to a loss of mixing entropy. As the nucleation rate J is predominantly determined by ΔG_c, we find a similar behavior for J, where J of the fcc phase seems to be smaller compared to that of the binary crystals. However, a direct comparison of the nucleation rates J for the three different systems is difficult, as J is only measured for relatively high supersaturations $\beta\Delta\mu > 0.34$ in the case of hard spheres,[55] whereas J is determined for $\Delta\mu < 0.4$ for the binary crystals. In addition, we plot the interfacial free energy γ for the three phases in Figure 6c. We wish to remark here that we express the surface tensions in units of k_BT/σ_L^2, which is an arbitrary choice, and hence a direct comparison of the three systems cannot be made as the dimensions of the spheres and the compositions are very different for the fcc, Laves, and AB_{13} phase. Hence, the conjecture of ref 23 that the surface tension of the AB_{13}−fluid interface may be low due to the structural similarity of these phases is difficult to verify. Moreover, one might expect that the much higher dimensionless interfacial tension $\beta\gamma\sigma_L^2$ of the AB_{13} phase may give rise to a much higher ΔG_c, but this is counterbalanced by a higher reduced crystal density $\rho_s\sigma_L^3$ in eq 5. On the other hand, by comparing nucleation of the fcc phase with that of the Laves phase, we find that although the dimensionless interfacial free energies $\beta\gamma\sigma_L^2$ are comparable, the difference in crystal density $\rho_s\sigma_L^3$ can yield a difference in ΔG_c. Thus, in order to compare the effect of interfacial tension and crystal density on the nucleation of different crystal structures, one should compare the ratio $\gamma^3/\rho_s\rho_s^2$ for the various systems, as this ratio is directly related to the barrier height and critical nucleus size *via* eqs 3 and 5 and is independent of an arbitrary choice of length scale.

Finally, we observe not only much higher reduced surface tensions $\beta\gamma\sigma_L^2$ for the AB_{13} phase with respect to the other

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346

examined crystals but also a much stronger increase of γ with supersaturation. This steep rise in surface tension with $\Delta\mu$ is responsible for the flattening of the nucleation barrier and nucleation rate at high supersaturation in Figure 6a and b, indicating that spontaneous nucleation, *i.e.*, where the nucleation barrier is sufficiently low, may be at surprisingly high driving forces $\Delta\mu$.

Spontaneous Nucleation. Figure 6a shows that the nucleation barrier $\beta\Delta G_c$ decreases with supersaturation $\beta\Delta\mu$. When $\beta\Delta G_c$ is sufficiently low, we expect to observe spontaneous nucleation of the AB_{13} phase using brute force MD simulations, *i.e.*, without any nonphysical biasing of the sampling of phase space.

In order to observe spontaneous AB_{13} nucleation, we initialize the system in a highly supersaturated binary fluid phase and perform MD simulations for a wide range of pressures in the *NPT* ensemble. Using our trained ANN to identify the AB_{13} particles, we monitor and plot the size of the largest cluster N_{CL} of the AB_{13} phase as a function of time t/τ_{MD} in Figure 7a. We distinguish three different regimes. At pressure $\beta P\sigma_L^3 = 71.0$, the metastable fluid does not show any sign of crystallization within our simulation times, and hence, the supersaturation $\beta\Delta\mu \simeq 0.74$ is too low to observe spontaneous nucleation. At a slightly higher pressure $\beta P\sigma^3 = 71.4$ ($\beta\Delta\mu \simeq 0.75$), we find a critical nucleus appearing after some waiting time, which subsequently grows out in time, showing a spontaneous crystallization event of the AB_{13} structure proceeding *via* nucleation. Increasing the pressure even further, we enter the third regime, where as soon as the simulation is started, multiple nuclei form immediately throughout the fluid. In this regime, the supersaturated fluid phase is mechanically unstable, and hence, crystallization sets in immediately and exhibits spinodal-like behavior. We thus confirm that the instability regime of the binary fluid phase with respect to AB_{13} crystallization is at much higher driving forces $\beta\Delta\mu > 0.75$ than in the case of the Laves phases, for which we found $\beta\Delta\mu > 0.53$ in ref 50 due to a much stronger increase of γ with supersaturation of the AB_{13} phase. More specifically, the significant increase of interfacial tension with supersaturation of the AB_{13} phase implies a relatively slow decrease (increase) of the nucleation barrier $\beta\Delta G_c$ (nucleation rate J). Hence, spontaneous nucleation of the AB_{13} is found at surprisingly high $\beta\Delta\mu$ with respect to the Laves phases or the fcc phase. Moreover, the validation of our estimate of $\beta\Delta\mu$ where we should expect spontaneous nucleation based on the results from the seeding approach supports our assumption that AB_{13} nucleation is well-described by classical nucleation theory.

In Figure 7b we show a time sequence of the early stages of the spontaneous AB_{13} nucleation from the metastable binary fluid phase at $\beta P\sigma_L^3 = 72.2$ with time increasing from the upper-left corner and then proceeding clockwise. We make two remarks here. First, we observe that the ANN classification, combined with the clustering algorithm, is capable of following the nucleation process from the early stages, thereby revealing the kinetic pathways toward the formation of an embryo. Second, we find that the nucleation starts with a (defective) icosahedral cluster of small spheres around which large spheres start to order themselves on a simple cubic lattice. We thus show that the local bond orientational order of small spheres into clusters with icosahedral symmetry plays a crucial role in the kinetic pathway of the fluid-to-solid transition. This immediately begs the question whether nucleation of the AB_{13} phase proceeds *via* a classical pathway or a nonclassical two-step crystallization scenario where relatively dense or bond orientational ordered structures in the fluid phase act as precursors for nucleation. The nucleation kinetics is of paramount importance, as the seeding approach for estimating the Gibbs free-energy barrier heights and nucleation rates is only valid in the case that nucleation proceeds *via* a classical nucleation pathway. To this end, we analyze particle configurations of spontaneous nucleation events in time. We observe that the system remains in the metastable binary fluid phase in which small crystalline nuclei appear and dissolve until a crystal nucleus of the AB_{13} phase exceeds its critical size at intermediate times and grows out. The induction time and the growth of a crystal nucleus when its size is larger than the critical nucleus size demonstrate that binary nucleation of the AB_{13} phase proceeds *via* a classical nucleation pathway. In Figure 7b, we observe the spontaneous formation of a crystalline nucleus in the metastable fluid phase, which grows further when its size is larger than the critical size. This observation, together with the video that we include in the Supporting Information, shows that AB_{13} formation occurs *via* classical nucleation.

Local Motifs Analysis. In order to shed light on the early stages of the AB_{13} crystal nucleation, we employ a recently developed method called topological cluster classification (TCC) algorithm[71] to detect predetermined particle arrangements in particle configurations. In particular, we focus on two topological clusters, the square shortest-path four-membered ring (sp4a) and the regular icosahedral cluster of 13 particles (13a), which are relevant particle clusters of the large and small species, respectively, in the AB_{13} crystal (see Figure 8b). The fraction of particles belonging to these two clusters has on average a nonzero value in the fluid phase and reaches one as crystallization proceeds.

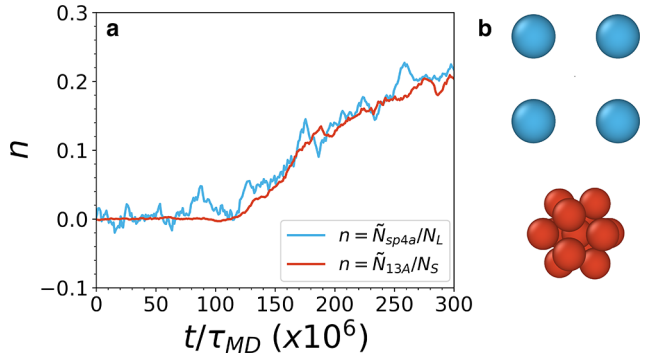


Figure 8. (a) Normalized number of large $\tilde{N}_{sp4a}/N_L$ (small $\tilde{N}_{13a}/N_S$) particles belonging to a square sp4a (icosahedral 13a) cluster as a function of time t/τ_{MD} in a spontaneous nucleation event, corrected for the averaged number observed in the fluid phase. The large statistical fluctuations in $\tilde{N}_{sp4a}/N_L$ is due to a much lower number of large species in the system. (b) Sketch of the square shortest-path four-membered ring sp4a (blue) and the 13a (red) icosahedral cluster.

In Figure 8a we plot the evolution of the fraction of large and small particles belonging to the square and the icosahedral clusters, respectively, during a spontaneous nucleation event. In order to facilitate the comparison between the two clusters, we subtract the corresponding averaged particle fraction observed in the fluid phase. Hence, the curves fluctuate around zero until nucleation occurs. Interestingly, the fraction of square and icosahedral clusters both increase at the same time when nucleation occurs, and the growth behavior of both particle clusters is similar. Hence, we conclude that the nucleation of the AB_{13} phase proceeds *via* a coassembly process, in which the large

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346

spheres form a simple cubic lattice and the small species form the body-centered icosahedral clusters.

CONCLUSIONS

In conclusion, we have investigated homogeneous nucleation of an AB_{13} crystal in a binary fluid of hard spheres with a size ratio of $q = 0.55$ and a composition corresponding to the stoichiometry of the AB_{13} phase. To achieve this, we have trained a neural network using a large set of nonaveraged bond order parameters as input to distinguish the AB_{13} phase from all competing phases, *i.e.*, the binary fluid and the fcc phase, with significant accuracy in the case of bulk phases. We showed that using two (averaged) bond order parameters is not sufficient to identify the different phases of interest in a single-particle level, while an artificial neural network with two hidden layers provides an elegant and powerful way of combining a high number of bond order parameters and to successfully distinguish the different phases with high accuracy.

Using the trained neural network as an order parameter in our nucleation study, we were able to follow crystal nucleation of the AB_{13} phase in a supersaturated binary fluid phase. We used the seeding approach to calculate Gibbs free-energy barriers and nucleation rates without prior knowledge about the system. Subsequently, we made a comparison of the nucleation of the AB_{13} phase with another binary hard-sphere crystal, the Laves phase, and with a single-component fcc phase of pure hard spheres. Our key findings are that (1) the assumption that the nucleation barrier for binary nucleation is higher due to a loss of mixing entropy is incorrect, *e.g.*, the barrier for the pure fcc phase is higher than for the AB_{13} and the Laves phases in the case of hard spheres at the same thermodynamic driving force $\Delta\mu$, and that (2) the assumption that the nucleation barrier is high due to a high interfacial free energy is not always valid, as it also depends on the number density of the solid phase, *e.g.*, the reduced surface tensions $\beta\gamma\sigma_L^2$ for the fcc and the Laves phase are very similar, but fcc has a higher barrier height due to a lower reduced solid density $\rho_s\sigma_L^3$. Hence, in order to compare the effect of interfacial tension and crystal density on the nucleation of different crystal structures, one should compare the ratio γ^3/ρ_s^2 for the various systems, as this ratio is directly related to the barrier height and critical nucleus size *via* eqs 3 and 5 and is independent of an arbitrary choice of length scale.

Subsequently, we used the dependence of the nucleation barrier height $\beta\Delta G_c$ on supersaturation $\Delta\mu$ to obtain an estimate of $\Delta\mu$ where spontaneous nucleation should occur. In this way, we were able to observe spontaneous nucleation of the AB_{13} phase using brute force MD simulations. To shed light on the nucleation mechanism, we analyzed the spontaneous nucleation events by measuring the fraction of large particles belonging to a square shortest-path four-membered ring (sp4a) and the fraction of small particles belonging to an icosahedral (13a) cluster. We observed a similar growth behavior of both clusters, demonstrating that the AB_{13} nucleation proceeds *via* a coassembly process. This finding is corroborated by our analysis of the early stages of nucleation when the first embryo forms using the trained neural network as an order parameter. Figure 7b shows clearly that the embryo grows by the attachment of both large particles and icosahedral clusters of small particles. Our results show that AB_{13} crystal nucleation proceeds *via* a classical pathway that can be well-described by CNT, even though the binary fluid is highly structured. Due to thermal fluctuations, the local regions of high bond orientational order and many-body correlations appear and disappear in the metastable fluid phase. The crystal only forms when also the large species are coordinated in the right way around the icosahedral clusters, thereby making AB_{13} nucleation classical.

Finally, our method for the identification of local structures using a neural network can straightforwardly be extended to other crystal structures, liquid crystal phases, and glasses and can be employed for future nucleation studies, analyzing not only numerical but also experimental data stacks.

METHODS

We determine bulk coexistence of the binary fluid phase with a composition $x_L = 1/14$ and the AB_{13} crystal of a binary mixture of nearly hard spheres (see the section The Model) with a diameter ratio $q = \sigma_S/\sigma_L = 0.55$ by computing the Helmholtz free energy per particle $f = F/N$ as a function of density $\rho = N/V$ for both phases with N the number of particles and V the volume of the system. We calculate f using thermodynamic integration of the equations of state

$$\beta f(\rho) = \beta f(\rho_0) + \int_{\rho_0}^{\rho} d\rho' \frac{\beta P(\rho')}{\rho'^2} \tag{6}$$

where $f(\rho_0)$ denotes the Helmholtz free energy per particle for a reference density ρ_0, $\beta = 1/k_BT$ is the inverse temperature, and P is the pressure. We use the ideal gas as a reference state for the binary fluid phase, and we employ the Frenkel−Ladd method to calculate the Helmholtz free energy at a reference density ρ_0 using MC simulations in the NVT ensemble.[72]

Subsequently, we calculate, for both the AB_{13} crystal and the binary fluid phase, the chemical potential $\beta\mu$ at pressure P:

$$\beta\mu = \frac{\beta G}{N} = \beta f + \frac{\beta P}{\rho} \tag{7}$$

with βG the dimensionless Gibbs free energy. The chemical potential difference or supersaturation is obtained *via* $\beta\Delta\mu = \beta\mu_{fluid} - \beta\mu_{AB_{13}}$, whereas two-phase coexistence between the AB_{13} and fluid phase is determined by imposing $\beta\Delta\mu = 0$, resulting in

$$\frac{\beta\Delta f}{\Delta(1/\rho)} = -\beta P \tag{8}$$

which is equivalent to the common tangent construction on the free-energy curves in the $\beta f - 1/\rho$ plane with $\Delta f = f_{fluid} - f_{AB_{13}}$ and $\Delta(1/\rho) = (1/\rho_{fluid}) - (1/\rho_{AB_{13}})$.

In order to train the ANN, we first build a training set of the different local particle environments (classes) we wish to distinguish. To this end, we perform MC simulations in the NPT ensemble of the AB_{13} crystal and the binary fluid phase with a composition $x_L = 1/14$, both at coexistence pressure $\beta P_{coex}\sigma_L^3 = 45.35$, and of the pure fcc phase at bulk coexistence with the fluid phase at pressure $\beta P_{coex}\sigma^3 = 8.87$. In total we collect 100 000 configurations of each local particle environment that we wish to classify. We describe each local particle environment with a 36-dimensional input vector of nonaveraged bond order parameters; see eq 2. The training of the network was done using the Keras package, enabled by Tensorflow backend. Specifically, we trained the network minimizing the categorical cross-entropy loss function with the addition of an L_2 regularization term using a weight decay prefactor of 10^{-4}. The minimization was carried out using minibatch stochastic gradient descent with momentum,[61,73,74] and we set the learning rate to 10^{-2}.

Both equilibration parts of the seeding approach have been carried out using MC simulations in the NPT ensemble, using pressure $\beta P\sigma_L^3 = 56.0$ and a total number of MC cycles equal to 10^3. For our investigations on the seeded growth and the spontaneous nucleation, we perform MD simulations using HOOMD-blue (highly optimized object-oriented many-particle dynamics)[75,76] in the NPT ensemble. We use varying system sizes for the seeding approach (see Table 2) and employ a total number of 3024 particles to study spontaneous nucleation. The temperature T and pressure P are kept constant *via* the Martyna−Tobias−Klein (MTK) integrator,[77] with the thermostat and

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346

barostat coupling constants $\tau_T = 1.0\tau_{MD}$ and $\tau_P = 1.0\tau_{MD}$, respectively, and $\tau_{MD} = \sigma_L\sqrt{m/\epsilon}$ is the MD time unit. The time step is set to $\Delta t = 0.004\tau_{MD}$, which is small enough to ensure stability of the simulations. We ran the simulations for $10^9\tau_{MD}$ time steps, unless specified otherwise. The simulation box is cubic, and periodic boundary conditions are applied in all directions.

To calculate the nucleation free-energy barrier heights and nucleation rates, we used the seeding approach. To this end, we initialized the system with a crystal seed of the AB_{13} phase surrounded by fluid particles at an overall composition of $x_L = 1/14$. We first equilibrated the system using two steps as described in the main text, *via* MC simulations in the *NPT* ensemble involving trial moves to displace particles and to isotropically scale the volume of the system. We employed six different seed sizes to determine the free-energy barrier height $\beta\Delta G_c$ and nucleation rate $J\sigma_L^5/D_L$ as a function of supersaturation $\beta\Delta\mu$ as shown in Figure 6. Finally, we employ the TCC algorithm[71] to analyze spontaneous nucleation events. The algorithm is used separately on both species; that is, we take into account one species at a time. Bonds between particles are detected using a modified Voronoi construction method.[71] The free parameter f_c, controlling the amount of asymmetry that a four-membered ring can show before being identified as two three-membered rings, is set to 0.82.[71]

All simulation images are realized using the OVITO software.[78]

Bond Orientational Order Parameters. To describe the local environment of a particle, we employ the standard bond orientational order parameters introduced by Steinhardt *et al.*[52] We first define the complex vector $q_{lm}(i)$ for each particle i

$$q_{lm}(i) = \frac{1}{N_b(i)}\sum_{j=1}^{N_b(i)} Y_{lm}(\theta(\mathbf{r}_{ij}), \phi(\mathbf{r}_{ij})) \tag{9}$$

where $N_b(i)$ is the number of neighbors of particle i, $Y_{lm}(\theta(\mathbf{r}_{ij}), \phi(\mathbf{r}_{ij}))$ denotes the spherical harmonics, $m \in [-l, l]$, $\theta(\mathbf{r}_{ij})$ and $\phi(\mathbf{r}_{ij})$ are the polar and azimuthal angles of the distance vector $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$, and $\mathbf{r}_i$ denotes the position of particle i. Subsequently, we define rotationally invariant quadratic and cubic order parameters as

$$q_l(i) = \sqrt{\frac{4\pi}{2l+1}\sum_{m=-l}^{l} |q_{lm}(i)|^2} \tag{10}$$

and

$$w_l(i) = \frac{\sum_{m_1+m_2+m_3}\begin{pmatrix} l & l & l \\ m_1 & m_2 & m_3 \end{pmatrix} q_{lm_1}(i) q_{lm_2}(i) q_{lm_3}(i)}{\left(\sum_{m=-l}^{l} |q_{lm}(i)|\right)^{3/2}} \tag{11}$$

Additionally, we also use the averaged bond orientational order parameters. The averaged $\bar{q}_{lm}(i)$ is defined as

$$\bar{q}_{lm}(i) = \frac{1}{\tilde{N}_b(i)}\sum_{j=1}^{\tilde{N}_b(i)} q_{lm}(j) \tag{12}$$

where $\tilde{N}_b(i)$ is the number of neighbors including particle i itself. The rotationally invariant quadratic and cubic averaged bond order parameters are defined as

$$\bar{q}_l(i) = \sqrt{\frac{4\pi}{2l+1}\sum_{m=-l}^{l} |\bar{q}_{lm}(i)|^2} \tag{13}$$

and

$$\bar{w}_l(i) = \frac{\sum_{m_1+m_2+m_3}\begin{pmatrix} l & l & l \\ m_1 & m_2 & m_3 \end{pmatrix} \bar{q}_{lm_1}(i) \bar{q}_{lm_2}(i) \bar{q}_{lm_3}(i)}{\left(\sum_{m=-l}^{l} |\bar{q}_{lm}(i)|\right)^{3/2}} \tag{14}$$

To identify all the neighbors of particle i, we employ the parameter-free solid-angle-based nearest-neighbor algorithm of Van Meel.[59] This algorithm assigns a solid angle to every potential neighbor j of i and defines the neighborhood of particle i to consist of the $N_b(i)$ particles nearest to i for which the sum of solid angles equals 4π. We note that in SANN the identification of neighbors is not necessarily symmetric; that is, it is not ensured that if particle A is a neighbor of particle B, particle B is a neighbor of particle A.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsnano.0c07541.

Possibility of realizing an accurate classification either without using machine learning techniques or with different machine learning techniques, the selected ANN architecture (number of hidden layers and number of output classes) and its performance in detecting particles at the interface, and the input features it prioritizes during the training phase (PDF)

Video showing that AB_{13} formation occurs *via* classical nucleation (MP4)

AUTHOR INFORMATION

Corresponding Authors

Gabriele M. Coli − *Soft Condensed Matter, Debye Institute for Nanomaterials Science, Department of Physics, Utrecht University, 3584 CC Utrecht, The Netherlands;* orcid.org/0000-0002-5125-8007; Email: g.m.coli@uu.nl

Marjolein Dijkstra − *Soft Condensed Matter, Debye Institute for Nanomaterials Science, Department of Physics, Utrecht University, 3584 CC Utrecht, The Netherlands;* Email: m.dijkstra@uu.nl

Complete contact information is available at: https://pubs.acs.org/10.1021/acsnano.0c07541

Author Contributions

M.D. initiated the project and supervised G.M.C. during the work. G.M.C. performed the simulations and analysis. M.D. and G.M.C. discussed the results together and cowrote the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

G.M.C. and M.D. acknowledge financial support from NWO (grant no. 16DDS003). We thank Jayden Savage for a preliminary study of the AB_{13} crystal nucleation.

REFERENCES

(1) Schöpe, H. J.; Bryant, G.; van Megen, W. Two-Step Crystallization Kinetics in Colloidal Hard-Sphere Systems. *Phys. Rev. Lett.* **2006**, *96*, 175701.

(2) Erdemir, D.; Lee, A. Y.; Myerson, A. S. Nucleation of Crystals From Solution: Classical and Two-Step Models. *Acc. Chem. Res.* **2009**, *42*, 621−629.

(3) Kawasaki, T.; Tanaka, H. Formation of a Crystal Nucleus from Liquid. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 14036−14041.
(4) Schilling, T.; Schöpe, H. J.; Oettel, M.; Opletal, G.; Snook, I. Precursor-Mediated Crystallization Process in Suspensions of Hard Spheres. *Phys. Rev. Lett.* **2010**, *105*, 025701.
(5) Sear, R. P. The Non-Classical Nucleation of Crystals: Microscopic Mechanisms and Applications to Molecular Crystals, Ice and Calcium Carbonate. *Int. Mater. Rev.* **2012**, *57*, 328−356.
(6) Russo, J.; Tanaka, H. The Microscopic Pathway to Crystallization in Supercooled Liquids. *Sci. Rep.* **2012**, *2*, 505.
(7) Karthika, S.; Radhakrishnan, T.; Kalaichelvi, P. A Review of Classical and Nonclassical Nucleation Theories. *Cryst. Growth Des.* **2016**, *16*, 6663−6681.
(8) Russo, J.; Tanaka, H. Nonclassical Pathways of Crystallization in Colloidal Systems. *MRS Bull.* **2016**, *41*, 369−374.
(9) Berryman, J. T.; Anwar, M.; Dorosz, S.; Schilling, T. The Early Crystal Nucleation Process in Hard Spheres Shows Synchronised Ordering and Densification. *J. Chem. Phys.* **2016**, *145*, 211901.
(10) Li, M.; Chen, Y.; Tanaka, H.; Tan, P. Revealing Roles of Competing Local Structural Orderings in Crystallization of Polymorphic Systems. *Science advances* **2020**, *6*, No. eaaw8938.
(11) Auer, S.; Frenkel, D. Prediction of Absolute Crystal-Nucleation Rate in Hard-Sphere Colloids. *Nature* **2001**, *409*, 1020−1023.
(12) Filion, L.; Hermes, M.; Ni, R.; Dijkstra, M. Crystal Nucleation of Hard Spheres Using Molecular Dynamics, Umbrella Sampling, and Forward Flux Sampling: A Comparison of Simulation Techniques. *J. Chem. Phys.* **2010**, *133*, 244115.
(13) Filion, L.; Ni, R.; Frenkel, D.; Dijkstra, M. Simulation of Nucleation in Almost Hard-Sphere Colloids: The Discrepancy between Experiment and Simulation Persists. *J. Chem. Phys.* **2011**, *134*, 134901.
(14) Espinosa, J. R.; Vega, C.; Valeriani, C.; Sanz, E. Seeding Approach to Crystal Nucleation. *J. Chem. Phys.* **2016**, *144*, 034501.
(15) Espinosa, J. R.; Vega, C.; Valeriani, C.; Sanz, E. The Crystal-Fluid Interfacial Free Energy and Nucleation Rate of NaCl from Different Simulation Methods. *J. Chem. Phys.* **2015**, *142*, 194709.
(16) Richard, D.; Speck, T. Crystallization of Hard Spheres Revisited. I. Extracting Kinetics and Free Energy Landscape from Forward Flux Sampling. *J. Chem. Phys.* **2018**, *148*, 124110.
(17) Tateno, M.; Yanagishima, T.; Russo, J.; Tanaka, H. Influence of Hydrodynamic Interactions on Colloidal Crystallization. *Phys. Rev. Lett.* **2019**, *123*, 258002.
(18) Fiorucci, G.; Coli, G. M.; Padding, J. T.; Dijkstra, M. The Effect of Hydrodynamics on the Crystal Nucleation of Nearly Hard Spheres. *J. Chem. Phys.* **2020**, *152*, 064903.
(19) Leunissen, M. E.; Christova, C. G.; Hynninen, A.-P.; Royall, C. P.; Campbell, A. I.; Imhof, A.; Dijkstra, M.; Van Roij, R.; Van Blaaderen, A. Ionic Colloidal Crystals of Oppositely Charged Particles. *Nature* **2005**, *437*, 235−240.
(20) Shevchenko, E. V.; Talapin, D. V.; Kotov, N. A.; O'Brien, S.; Murray, C. B. Structural Diversity in Binary Nanoparticle Superlattices. *Nature* **2006**, *439*, 55−59.
(21) Dijkstra, M. Entropy-Driven Phase Transitions in Colloids: From Spheres to Anisotropic Particles. *Adv. Chem. Phys.* **2014**, *156*, 35−71.
(22) Shevchenko, E. V.; Talapin, D. V.; O'Brien, S.; Murray, C. B. Polymorphism in AB13 Nanoparticle Superlattices: An Example of Semiconductor- Metal Metamaterials. *J. Am. Chem. Soc.* **2005**, *127*, 8741−8747.
(23) Eldridge, M.; Madden, P.; Frenkel, D. Entropy-Driven Formation of a Superlattice in a Hard-Sphere Binary Mixture. *Nature* **1993**, *365*, 35−37.
(24) Filion, L.; Dijkstra, M. Prediction of Binary Hard-Sphere Crystal Structures. *Phys. Rev. E* **2009**, *79*, 046714.
(25) Sanders, J.; Murray, M. Ordered Arrangements of Spheres of Two Different Sizes in Opal. *Nature* **1978**, *275*, 201−203.
(26) Murray, M.; Sanders, J. Close-Packed Structures of Spheres of Two Different Sizes II. The Packing Densities of Likely Arrangements. *Philos. Mag. A* **1980**, *42*, 721−740.
(27) Hachisu, S.; Yoshimura, S. Optical Demonstration of Crystalline Superstructures in Binary Mixtures of Latex Globules. *Nature* **1980**, *283*, 188−189.
(28) Yoshimura, S.; Hachisu, S. *Frontiers in Colloid Science In Memoriam Professor Dr. Bun-ichi Tamamushi*; Springer, New York, NY, 1983 59−70.
(29) Bartlett, P.; Ottewill, R.; Pusey, P. Freezing of Binary Mixtures of Colloidal Hard Spheres. *J. Chem. Phys.* **1990**, *93*, 1299−1312.
(30) Bartlett, P.; Ottewill, R.; Pusey, P. Superlattice Formation in Binary Mixtures of Hard-Sphere Colloids. *Phys. Rev. Lett.* **1992**, *68*, 3801.
(31) Hunt, N.; Jardine, R.; Bartlett, P. Superlattice Formation in Mixtures of Hard-Sphere Colloids. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2000**, *62*, 900.
(32) Redl, F. X.; Cho, K.-S.; Murray, C. B.; O'Brien, S. Three-Dimensional Binary Superlattices of Magnetic Nanocrystals and Semiconductor Quantum Dots. *Nature* **2003**, *423*, 968−971.
(33) Shevchenko, E. V.; Talapin, D. V.; Murray, C. B.; O'Brien, S. Structural Characterization of Self-Assembled Multifunctional Binary Nanoparticle Superlattices. *J. Am. Chem. Soc.* **2006**, *128*, 3620−3637.
(34) Overgaag, K.; Evers, W.; de Nijs, B.; Koole, R.; Meeldijk, J.; Vanmaekelbergh, D. Binary Superlattices of PbSe and CdSe Nanocrystals. *J. Am. Chem. Soc.* **2008**, *130*, 7833−7835.
(35) Chen, Z.; O'Brien, S. Structure Direction of II- VI Semiconductor Quantum Dot Binary Nanoparticle Superlattices by Tuning Radius Ratio. *ACS Nano* **2008**, *2*, 1219−1229.
(36) Evers, W. H.; Nijs, B. D.; Filion, L.; Castillo, S.; Dijkstra, M.; Vanmaekelbergh, D. Entropy-Driven Formation of Binary Semiconductor-Nanocrystal Superlattices. *Nano Lett.* **2010**, *10*, 4235−4241.
(37) Bodnarchuk, M. I.; Kovalenko, M. V.; Heiss, W.; Talapin, D. V. Energetic and Entropic Contributions to Self-Assembly of Binary Nanocrystal Superlattices: Temperature as the Structure-Directing Factor. *J. Am. Chem. Soc.* **2010**, *132*, 11967−11977.
(38) Kang, Y.; Ye, X.; Chen, J.; Qi, L.; Diaz, R. E.; Doan-Nguyen, V.; Xing, G.; Kagan, C. R.; Li, J.; Gorte, R. J.; et al. Engineering Catalytic Contacts and Thermal Stability: Gold/Iron Oxide Binary Nanocrystal Superlattices for CO Oxidation. *J. Am. Chem. Soc.* **2013**, *135*, 1499−1505.
(39) Bodnarchuk, M. I.; Erni, R.; Krumeich, F.; Kovalenko, M. V. Binary Superlattices from Colloidal Nanocrystals and Giant Polyoxometalate Clusters. *Nano Lett.* **2013**, *13*, 1699−1705.
(40) Sakamoto, Y.; Kuroda, Y.; Toko, S.; Ikeda, T.; Matsui, T.; Kuroda, K. Electron Microscopy Study of Binary Nanocolloidal Crystals with ico-AB13 Structure Made of Monodisperse Silica Nanoparticles. *J. Phys. Chem. C* **2014**, *118*, 15004−15010.
(41) Boles, M. A.; Talapin, D. V. Many-Mody Effects in Nanocrystal Superlattices: Departure from Sphere Packing Explains Stability of Binary Phases. *J. Am. Chem. Soc.* **2015**, *137*, 4494−4502.
(42) Ye, X.; Zhu, C.; Ercius, P.; Raja, S. N.; He, B.; Jones, M. R.; Hauwiller, M. R.; Liu, Y.; Xu, T.; Alivisatos, A. P. Structural Diversity in Binary Superlattices Self-Assembled from Polymer-Grafted Nanocrystals. *Nat. Commun.* **2015**, *6*, 1−10.
(43) Wei, J.; Schaeffer, N.; Pileni, M.-P. Ligand Exchange Governs the Crystal Structures in Binary Nanocrystal Superlattices. *J. Am. Chem. Soc.* **2015**, *137*, 14773−14784.
(44) Punnathanam, S.; Monson, P. Crystal Nucleation in Binary Hard Sphere Mixtures: a Monte Carlo Simulation Study. *J. Chem. Phys.* **2006**, *125*, 024508.
(45) Jungblut, S.; Dellago, C. Crystallization of a Binary Lennard-Jones Mixture. *J. Chem. Phys.* **2011**, *134*, 104501.
(46) Ganagalla, S. R.; Punnathanam, S. N. Free Energy Barriers for Homogeneous Crystal Nucleation in a Eutectic System of Binary Hard Spheres. *J. Chem. Phys.* **2013**, *138*, 174503.
(47) Bommineni, P. K.; Punnathanam, S. N. Molecular Simulation of Homogeneous Crystal Nucleation of AB2 Solid Phase from a Binary Hard Sphere Mixture. *J. Chem. Phys.* **2017**, *147*, 064504.
(48) Sanz, E.; Valeriani, C.; Frenkel, D.; Dijkstra, M. Evidence for Out-of-Equilibrium Crystal Nucleation in Suspensions of Oppositely Charged Colloids. *Phys. Rev. Lett.* **2007**, *99*, 055501.

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346

(49) Peters, B. Competing Nucleation Pathways in a Mixture of Oppositely Charged Colloids: Out-of-Equilibrium Nucleation Revisited. *J. Chem. Phys.* **2009**, *131*, 244103.
(50) Dasgupta, T.; Coli, G. M.; Dijkstra, M. Tuning the Glass Transition: Enhanced Crystallization of the Laves Phases in Nearly Hard Spheres. *ACS Nano* **2020**, *14*, 3957−3968.
(51) Bommineni, P. K.; Klement, M.; Engel, M. Spontaneous Crystallization in Systems of Binary Hard Sphere Colloids. *Phys. Rev. Lett.* **2020**, *124*, 218003.
(52) Steinhardt, P. J.; Nelson, D. R.; Ronchetti, M. Bond-Orientational Order in Liquids and Glasses. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1983**, *28*, 784.
(53) Frank, F. C. Supercooling of Liquids. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1952**, *215*, 43−46.
(54) Weeks, J. D.; Chandler, D.; Andersen, H. C. Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids. *J. Chem. Phys.* **1971**, *54*, 5237−5247.
(55) Richard, D.; Speck, T. Crystallization of Hard Spheres Revisited. II. Thermodynamic Modeling, Nucleation Work, and the Surface of Tension. *J. Chem. Phys.* **2018**, *148*, 224102.
(56) Ten Wolde, P. R.; Ruiz-Montero, M. J.; Frenkel, D. Numerical Evidence for bcc Ordering at the Surface of a Critical fcc Nucleus. *Phys. Rev. Lett.* **1995**, *75*, 2714.
(57) Ten Wolde, P. R.; Ruiz-Montero, M. J.; Frenkel, D. Numerical Calculation of the Rate of Crystal Nucleation in a Lennard-Jones System at Moderate Undercooling. *J. Chem. Phys.* **1996**, *104*, 9932−9947.
(58) Lechner, W.; Dellago, C. Accurate Determination of Crystal Structures Based on Averaged Local Bond Order Parameters. *J. Chem. Phys.* **2008**, *129*, 114707.
(59) van Meel, J. A.; Filion, L.; Valeriani, C.; Frenkel, D. A Parameter-Free, Solid-Angle Based, Nearest-Neighbor Algorithm. *J. Chem. Phys.* **2012**, *136*, 234107.
(60) Boattini, E.; Ram, M.; Smallenburg, F.; Filion, L. Neural-Network-Based Order Parameters for Classification of Binary Hard-Sphere Crystal Structures. *Mol. Phys.* **2018**, *116*, 3066−3075.
(61) Bishop, C. M. *Pattern Recognition and Machine Learning*; Springer Science+Business Media: New York, NY, 2006.
(62) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.
(63) Warmflash, A.; Bhimalapuram, P.; Dinner, A. R. Umbrella Sampling for Nonequilibrium Processes. *J. Chem. Phys.* **2007**, *127*, 114109.
(64) Cuetos, A.; Dijkstra, M. Kinetic Pathways for the Isotropic-Nematic Phase Transition in a System of Colloidal Hard Rods: A Simulation Study. *Phys. Rev. Lett.* **2007**, *98*, 095701.
(65) Allen, R. J.; Warren, P. B.; Ten Wolde, P. R. Sampling Rare Switching Events in Biochemical Networks. *Phys. Rev. Lett.* **2005**, *94*, 018104.
(66) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.
(67) Trudu, F.; Donadio, D.; Parrinello, M. Freezing of a Lennard-Jones Fluid: From Nucleation to Spinodal Regime. *Phys. Rev. Lett.* **2006**, *97*, 105701.
(68) Moroni, D.; Ten Wolde, P. R.; Bolhuis, P. G. Interplay between Structure and Size in a Critical Crystal Nucleus. *Phys. Rev. Lett.* **2005**, *94*, 235703.
(69) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. Transition Path Sampling: Throwing Ropes over Rough Mountain Passes, in the Dark. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.
(70) Lechner, W.; Dellago, C.; Bolhuis, P. G. Role of the Prestructured Surface Cloud in Crystal Nucleation. *Phys. Rev. Lett.* **2011**, *106*, 085701.
(71) Malins, A.; Williams, S. R.; Eggers, J.; Royall, C. P. Identification of Structure in Condensed Matter with the Topological Cluster Classification. *J. Chem. Phys.* **2013**, *139*, 234506.
(72) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Computational Science Series; Elsevier: Cambridge, MA, 2002; Vol. *1*, pp 241−245.
(73) Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323*, 533−536.
(74) Sutskever, I.; Martens, J.; Dahl, G.; Hinton, G. On the Importance of Initialization and Momentum in Deep Learning. *Proceedings of the 30th International Conference on Machine Learning*, Atlanta, GA, USA, 2013; pp 1139−1147.
(75) Anderson, J. A.; Lorenz, C. D.; Travesset, A. General Purpose Molecular Dynamics Simulations Fully Implemented on Graphics Processing Units. *J. Comput. Phys.* **2008**, *227*, 5342−5359.
(76) Glaser, J.; Dac Nguyen, T.; Anderson, J. A.; Lui, P.; Spiga, F.; Millan, J. A.; Morse, D. C.; Glotzer, S. C. Strong Scaling of General-Purpose Molecular Dynamics Simulations on GPUs. *Comput. Phys. Commun.* **2015**, *192*, 97−107.
(77) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101*, 4177−4189.
(78) Stukowski, A. Visualization and Analysis of Atomistic Simulation Data with OVITO−the Open Visualization Tool. *Modell. Simul. Mater. Sci. Eng.* **2010**, *18*, 015012.

NOTE ADDED AFTER ASAP PUBLICATION

Originally published ASAP on February 23, 2021; Equation 1 updated February 24, 2021.

https://dx.doi.org/10.1021/acsnano.0c07541
ACS Nano 2021, 15, 4335−4346