



**HAL**  
open science

## **In search of a precursor for crystal nucleation of hard and charged colloids**

Marjolein de Jager, Frank Smallenburg, Laura Filion

► **To cite this version:**

Marjolein de Jager, Frank Smallenburg, Laura Filion. In search of a precursor for crystal nucleation of hard and charged colloids. *The Journal of Chemical Physics*, 2023, 159 (13), pp.134902. <10.1063/5.0161356>. <hal-04309399>

**HAL Id: hal-04309399**

**<https://hal.science/hal-04309399v1>**

Submitted on 27 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# In search of a precursor for crystal nucleation of hard and charged colloids

Marjolein de Jager,<sup>1, a)</sup> Frank Smalenburg,<sup>2</sup> and Laura Filion<sup>1</sup>

<sup>1)</sup>*Soft Condensed Matter, Debye Institute of Nanomaterials Science, Utrecht University, Utrecht, Netherlands*

<sup>2)</sup>*Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405 Orsay, France*

(Dated: 12 June 2023)

The interplay between crystal nucleation and the structure of the metastable fluid has been a topic of significant debate over recent years. In particular, it has been suggested that even in simple model systems such as hard or charged colloids, crystal nucleation might be foreshadowed by significant fluctuations in local structure around the location where the first nucleus arises. We investigate this using computer simulations of spontaneous nucleation events in both hard and charged colloidal particles. To detect local structural variations, we use both standard and unsupervised machine learning methods capable of finding hidden structures in the metastable fluid phase. We track numerous nucleation events for the face-centered cubic and body-centered cubic crystal on a local level, and demonstrate that all signs of crystallinity emerge simultaneously from the very start of the nucleation process. We thus conclude that there is no precursor for the nucleation of charged colloids.

## I. INTRODUCTION

Crystal nucleation plays an important role in fields ranging from colloidal self-assembly, to protein crystallization, and even polymorph selection in pharmaceuticals<sup>1,2</sup>. However, despite its importance in a number of essential fields, the detailed mechanism of forming a crystal nucleus still remains a topic of continuous debate.

The simplest theory which addresses crystal nucleation is classical nucleation theory (CNT). In CNT, the metastable fluid is continuously undergoing thermal fluctuations, where small, solid clusters form and dissolve until one appears which is large enough (critically large) to grow out into a macroscopic crystal. The size of such a critical cluster is given simply by balancing the bulk free-energy gain associated with transitioning into the more stable solid phase, with the surface free-energy cost of having a finite crystal cluster immersed in the fluid. This picture, however, becomes significantly more complicated when one considers the possibility of multiple competing crystal structures, typically referred to as polymorphs. In systems with crystal polymorphs, the crystalline phase that first nucleates in the metastable fluid is not necessarily the stable phase. Theories to address such situations, such as the Ostwald step rule<sup>3</sup>, and the Alexander-McTague theory<sup>4</sup> have proven unreliable in explaining polymorph selection (see e.g. Refs. 5–7).

One complication when studying such questions is the close interplay between local structural motifs that occur naturally in the fluid, and the ones that might emerge when the crystal forms. It has been suggested that motifs hiding in the fluid are predictive of, or even responsible for, the location or polymorph of the nucleus that forms<sup>8,9</sup>. To investigate this possibility, one avenue for-

ward could be to explore just how much information the metastable liquid is hiding regarding the nucleation process. Over the last two decades a plethora of studies have appeared presenting contradictory observations<sup>8,10–19</sup>. In particular, in simple systems such as charged colloids which nucleate into either the face-centered cubic (FCC) or body-centered cubic (BCC) crystal, some studies have argued that local structural order develops before the local density increases<sup>8,15,16</sup>, while other authors found evidence that the two processes happen simultaneously<sup>17</sup>.

To address this issue, some recent, intriguing studies have explored how modifying (via biasing) the structure of the fluid – either enhancing or suppressing specific local motifs – affects the nucleation process<sup>6,15</sup>. In principle, such studies might be able to give one direct evidence that a specific local structure either enhances or suppresses the nucleation process. Unfortunately, however, biasing the structure of the fluid modifies not only its local structure but also its thermodynamics, meaning that comparisons with the unbiased case are inconclusive.

The more direct route to trying to explore how various kinds of local ordering interplay in crystal nucleation is simply to simulate the nucleation event, and follow the various structural and density features as nucleation happens. At first glance this would appear to be a straightforward approach. However, the challenge in this case lies in the difficulty in creating local order parameters that are unbiased. For example, order parameters that are tuned to recognize the crystalline regions from fluid might struggle at the boundary between the fluid and crystal – a highly important aspect at the beginning of nucleation. Similar issues exist for other order parameters, making it very difficult to pinpoint the start of the nucleation process and hence to determine whether structural order emerges before, during, or after densification. Hence, in some cases instead of capturing accurately whether local structure exists in the highly fluctuating metastable fluid, one ends up examining the properties of the order parameter instead of the properties of

<sup>a)</sup>Electronic mail: m.e.dejager@uu.nl

the fluid.

While this problem is never fully avoidable, one option to try and avoid accidental biases is to exploit multiple different measures for local order – for instance measures associated with symmetries like bond-order parameters and order associated with the topological connections between neighboring particles – such as topological cluster classification (TCC). Interestingly, new unsupervised machine learning (UML) algorithms also give new avenues to probe structure (see e.g. Refs. 20–29). Recent studies have even demonstrated that simple, UML-based approaches are able to extract variations in disorder in the structure of supercooled fluids from e.g. a simple vector of bond order parameters<sup>23,27,28</sup>. Intriguingly, this includes identifying variations in local structure that are not easily extracted by looking at each element of the vector individually.

In this paper, we attempt to take the utmost care in identifying local signatures of the fluid and revisit the question: are there hidden local structures present in metastable fluid that foreshadow the location of the imminent formation of a crystal nucleus? Specifically, we apply both classical and UML-based methods to the nucleation of hard and charged colloids in both the regime of strong screening and weak screening, for which respectively the FCC and BCC crystals nucleate. To this end, we simulate numerous spontaneous nucleation events, and closely follow all nucleation events as a function of time. In particular, similar to Ref. 17 we zoom in on the regions where the nuclei are born and analyze the local fluctuations in density and structure of the metastable fluid. By doing this we can locally track whether there is a delay between the increase in local structural ordering and local density prior to the start of nucleation. Such a delay would indicate the presence of a precursor. However, within the limits of this study, we find no evidence of such a precursor in the systems we studied.

## II. MODEL

We consider a system of  $N$  like-charged hard spheres of diameter  $\sigma$  suspended in a solvent containing salt. The effective interaction potential between these colloids is given by the repulsive hard-core Yukawa potential

$$\beta\phi(r) = \begin{cases} \beta\epsilon \frac{e^{-\kappa\sigma(r/\sigma-1)}}{r/\sigma} & \text{for } r \geq \sigma, \\ \infty & \text{for } r < \sigma, \end{cases} \quad (1)$$

with contact value  $\beta\epsilon = Z^2\lambda_B/\sigma(1 + \kappa\sigma/2)^2$ , where  $Z$  is the charge of the colloids in electron charge,  $\lambda_B$  is the Bjerrum length,  $\kappa$  is the inverse Debye screening length, and  $\beta = 1/k_B T$ , with  $k_B$  the Boltzmann constant and  $T$  the temperature. Note that in the limit of zero charge ( $Z \rightarrow 0$ ) or infinite screening ( $\kappa\sigma \rightarrow \infty$ ), this potential reduces to the hard-sphere potential. The interaction

TABLE I. For each system studied, the packing fraction of the supersaturated fluid  $\eta^*$  at which the brute force nucleation is performed together with the corresponding supersaturation  $\beta|\Delta\mu|$ . The last columns give the critical nucleus size  $n^*$  and barrier height  $\beta\Delta G^*$  obtained using umbrella sampling. The error in  $\beta\Delta G^*$  is no more than 1.

	$\beta\epsilon$	$1/\kappa\sigma$	$\eta^*$	$\beta \Delta\mu $	$n^*$	$\beta\Delta G^*$
	hard spheres		0.5385	0.585	75	16.5
FCC	81	0.01	0.4681	0.584	84	16.3
	8	0.04	0.4400	0.541	69	14.8
BCC	81	0.40	0.1305	0.321	122	18.0

potential was truncated and shifted such that the shift was never more than  $10^{-5}k_B T$ .

Nucleation of both the BCC and FCC phases in this system has been studied in the past (see e.g. Refs. 30–34). In a previous study<sup>34</sup>, we used umbrella sampling to calculate the nucleation barriers and rates of highly screened charged particles. In this paper, we will study the nucleation of some of these (nearly-)hard systems, as well as the nucleation of weakly screened charged particles. To be able to compare the nucleation processes of different systems, we select state points with approximately equal barrier heights. In particular, we will simulate brute-force nucleation events of systems with barrier heights around  $15\text{--}18k_B T$ . Information on the nucleation barriers of the systems studied is given in Tab. I. Note that systems with a Debye screening length of  $1/\kappa\sigma = 0.01$  were found to behave essentially as “hard” spheres when mapped with an effective hard-sphere diameter<sup>34</sup>. A brief explanation of the methods used for computing the nucleation barriers as well as some additional information on these systems can be found in the Supplemental Materials (SM).

## III. METHODS

To explain the methods we use for studying the nucleation events, we need to discuss two things: i) how we identify local structure, and ii) how we track nucleation events locally.

### A. Identifying local structure

We use three different methods to classify the local structure. The first method considers just the averaged bond-orientational order parameters (BOPs) of Lechner and Dellago<sup>35</sup>. For this, we first calculate for each particle  $i$  the complex quantities

$$q_{lm}(i) = \frac{1}{N_b(i)} \sum_{j \in \mathcal{N}_b(i)} Y_l^m(\theta_{ij}, \phi_{ij}), \quad (2)$$

where  $\mathcal{N}_b(i)$  is the set of the  $N_b(i)$  nearest neighbors of particle  $i$ ,  $Y_{lm}(\theta, \phi)$  are the spherical harmonics with

$m \in [-l, l]$ , and  $\theta_{ij}$  and  $\phi_{ij}$  are the polar and azimuthal angles of the vector  $\mathbf{r}_{ij} = \mathbf{r}(j) - \mathbf{r}(i)$  connecting particles  $i$  and  $j$ . We use the SANN algorithm<sup>36</sup> to determine the nearest neighbors. Next, we average these complex quantities over the set of nearest neighbors as well as the particle itself

$$\bar{q}_{lm}(i) = \frac{1}{N_b(i) + 1} \sum_{j \in \{i, \mathcal{N}_b(i)\}} q_{lm}(j). \quad (3)$$

Finally, we compute the rotationally invariant averaged BOPs

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

and

$$\bar{w}_l(i) = \frac{w_l(i)}{\left(\sum_{m=-l}^l |\bar{q}_{lm}(i)|^2\right)^{3/2}}, \quad (5)$$

with

$$w_l(i) = \sum_{\substack{m_1, m_2, m_3 \\ m_1 + m_2 + m_3 = 0}} \begin{pmatrix} l & l & l \\ m_1 & m_2 & m_3 \end{pmatrix} q_{lm_1}(i) q_{lm_2}(i) q_{lm_3}(i), \quad (6)$$

where the term in brackets is the Wigner  $3j$  symbol, which is only non-zero when  $m_1 + m_2 + m_3 = 0$ . Note that  $w_l = 0$  when  $l$  is odd. Depending on the choice of  $l$ , these BOPs are sensitive to different (crystal) symmetries. For example,  $\bar{q}_6$  is very helpful in distinguishing more fluid-like environments from more solid-like environments such as FCC and BCC<sup>35</sup>.

While the BOPs are extremely useful for detecting specific symmetries in the local structure of the fluid, they are not necessarily optimal for detecting the most important structural variations in disordered systems such as the metastable fluid. Recent work has shown that BOPs in combination with unsupervised machine learning algorithms is highly effective at autonomously detecting variations that might be difficult to see by studying the individual BOPs<sup>22,23</sup>. Hence, for the second method, we use an unsupervised algorithm to autonomously detect local structural fluctuations in the metastable fluid. In Ref. 22, Boattini *et al.* showed that a neural-network-based autoencoder, which is given  $\bar{q}_l$  with  $l \in [1, 8]$  as input, does a good job in distinguishing a whole range of different local structures. Similarly, in Ref. 25 van Damme *et al.* showed that using principal component analysis (PCA) as dimensionality reduction method also does a good job in distinguishing the sizable assortment of crystal structures formed by rounded tetrahedra. For our system we found that PCA and an autoencoder performed equally well in distinguishing order, and hence chose to use the simpler PCA algorithm in our analysis.

To specifically focus on finding local fluctuations or signatures in the metastable fluid, we train the PCA model on  $\bar{q}_l$  with  $l \in [1, 8]$  of configurations containing only fluid particles and no (significant) solid nuclei. We then use this trained PCA model to analyze the entire nucleation trajectory. Note that in this way, the first principal component corresponds to the largest BOPs-related structural variation in the metastable fluid.

For the third and last method, we use an altogether different approach for classifying local structure. In particular we use the topological cluster classification (TCC) algorithm developed by Malins *et al.*<sup>37</sup> to detect any local motifs that the BOPs might have overlooked. More specifically, we use TCC to calculate the population of certain types of clusters, as well as the number of clusters of a certain type a particle is involved in.

Note that for computing the nucleation barriers, we additionally need a binary classification method which labels a particle as either fluid or solid. For this we use the 6-fold Ten Wolde bonds<sup>38</sup>

$$d_6(i, j) = \frac{\sum_{m=-6}^{m=6} q_{6m}(i) q_{6m}^*(j)}{\sqrt{\left(\sum_{m=-6}^{m=6} |q_{6m}(i)|^2\right) \left(\sum_{m=-6}^{m=6} |q_{6m}(j)|^2\right)}}, \quad (7)$$

where  $*$  indicates the complex conjugate and  $q_{6m}$  are the bond-orientational order parameters given by Eq. (2). Particle  $i$  is classified as solid if it has 6 or more neighboring particles  $j$  with which it has a solid-like bond, i.e.  $d_6(i, j) > 0.7$ . We also use this fluid-solid classification to initially locate and follow the nucleation event. However, we want to point out that, although this provides a general overview of the nucleation event, it is not an ideal order parameter to study the onset of nucleation. Its binary nature with the thresholds for  $d_6(i, j)$  and the number of solid-like bonds causes it to overlook subtle increases in the local structural ordering of the fluid and hence reacts more slowly to the nucleation than other continuous order parameters.

## B. Simulating and tracking nucleation events

To obtain the spontaneous nucleation events we use brute force MC, KMC, and MD simulations in the  $NVT$ -ensemble, where we simulate  $N = 10976$  and  $N = 11664$  particles for the systems forming the FCC and BCC phase, respectively. The difference between the MC and KMC simulations is the acceptance ratio of the trial particle moves. For the MC simulations this acceptance ratio is around 30%, whilst for the KMC simulations it is around 85% resulting in dynamics that mimic Brownian motion<sup>39</sup>. The MD simulations are performed using LAMMPS with a Nose-Hoover thermostat<sup>40</sup>. We only perform MD simulations for the soft hard-core Yukawa system, i.e. the one with  $1/\kappa\sigma = 0.40$ . As this system does not feel its hard core, it can simply be ignored. Once we have obtained the numerous nucleation events,

we want to track the local density and structure to determine if there is a difference between the increase in local structural ordering and local density at the start of nucleation. To this end, for each nucleation event we find the position  $\mathbf{r}_0$  that best captures the center of the nucleus at the start of nucleation. For this we use the average center-of-mass of the precritical nucleus as a starting point and, if needed, by eye adjust it to best capture the birthplace of the crystal nucleus. Next, for each snapshot of the nucleation trajectory, starting well before the start of nucleation, we determine all particles inside a sphere of radius  $R$  around  $\mathbf{r}_0$ , and take the average of the local properties of these particles. This is similar to what Berryman *et al.* did in Ref. 17. The local structural properties that we consider are explained in the previous subsection. Additionally, we define for each particle a local packing fraction measured via the volume of its Voronoi cell. The volumes of the Voronoi cells were obtained using `voropp`<sup>41</sup>. As we are searching for local precursors, we choose  $R$  such that the selected region contains around 30-40 particles. This size provides a good balance between being large enough to obtain relatively stable averages of the local properties, and being small enough to ensure that the averaged properties still represent the local situation.

## IV. RESULTS

### A. Structure of the metastable fluid

Before we look into the actual crystal nucleation, we first characterize the structural properties of the metastable fluid.

First, we examine the globally averaged values of the local BOPs, and plot the results in Fig. 1 for the metastable fluids of essentially hard spheres and of soft spheres as a function of the supersaturation. We see that  $\bar{q}_6$ ,  $\bar{q}_8$ , and  $\bar{w}_8$  are most prominent in both metastable fluids, and that all BOPs are only marginally affected by the increase in supersaturation. Furthermore, notice that the values in both systems are surprisingly similar, even though the metastable fluid of essentially hard spheres later forms an FCC crystal, whereas the fluid of the soft spheres will form a BCC crystal. The most prevalent difference between the two systems can be found in  $\bar{w}_6$ , which is smaller for the nearly-hard spheres than for soft spheres, and for high supersaturation even becomes on average negative for the nearly-hard spheres whereas it stays positive for the soft spheres. See the SM for more analysis on the  $\bar{w}_i$ 's. We, thus, conclude that the fluid's "knowledge" about which crystal phase it should nucleate into is difficult to distinguish from the global values of the BOPs.

In addition to the BOPs, we take a look at the presence of the different TCC clusters in the metastable fluids. Figure 2 shows the population of various TCC clusters in the metastable fluid of hard spheres, and soft spheres.

Even though we see some small deviations in the populations of the different metastable fluids – e.g. clusters 6A, 8A, 8K, 9K, and BCC\_9 have a slightly higher population in the fluid of soft spheres and 9B, 10B, 11C, 11E, and 12D have a slightly higher population in the fluid of nearly-hard spheres – the values are again surprisingly similar. This indicates once more that it is difficult to determine which crystal phase will nucleate from the metastable fluid for the systems studied here.

Next, we characterize the local ordering of the metastable fluid on the single-particle level. As explained in the Methods, we train a PCA model using only configurations of the metastable fluid. In all cases the first principal component (PC1) explains around 70% of the total variance of the input. To illustrate what kind of fluctuations PCA picks up in the fluid, Figs. 3a,c) show for the metastable fluids of hard spheres and of soft spheres the weight of each BOP in the first and second principal component. We see that PC1 is mostly made up of  $\bar{q}_6$  and  $\bar{q}_8$ . Furthermore, Figs. 3b,d) show the distribution of these metastable fluid particles in the PC1-PC2 plane, as well as the distributions of the corresponding FCC and BCC phase. (Recall that the data for the crystalline particles was not used in training the PCA models.) In this scatter plot we can neatly see that the crystal phases lie in the region of large PC1. To get a better understanding of the real-space distribution of these particles with above or below average PC1, we take a look at a single snapshot of the metastable fluid of hard spheres and color the particles according to their local packing fraction, PC1, and the number of 9B and 11F clusters a particles is involved in, see Fig. 4. Even though the spatial correlations in the local packing fraction are not clearly visible, we can clearly distinguish by eye large spatial regions of above or below average PC1. The autocorrelation functions of these spatial correlations can be found in the SM. Notice that the regions with above average PC1 correspond to an absence of 9B clusters and a high presence of 11F clusters, while regions with below average PC1 correspond to a high presence of 9B clusters and an absence of 11F clusters. Thus, there is a negative correlation between PC1 and 9B clusters and positive correlation between PC1 and 11F. The precise correlations of these two and other TCC clusters with PC1 are shown in Fig. 3e). Analogous to what was found in Ref. 23, the TCC clusters can be roughly divided into two groups: those with a negative correlation, which essentially are all clusters consisting of one or more tetrahedral subclusters, and those with a positive correlation, which contain the clusters consisting of one or more square pyramidal subclusters.

Combining the observation of large spatial regions of above average PC1 with the proximity of these particles to the crystal phases in the PC1-PC2 scatter plot (Fig. 3), we can conclude that regions of above average PC1 form a good candidate for harboring a precursor for crystal nucleation. In the next section, we will investigate these regions whilst tracking the nucleation events. However, before we turn our attention to that,

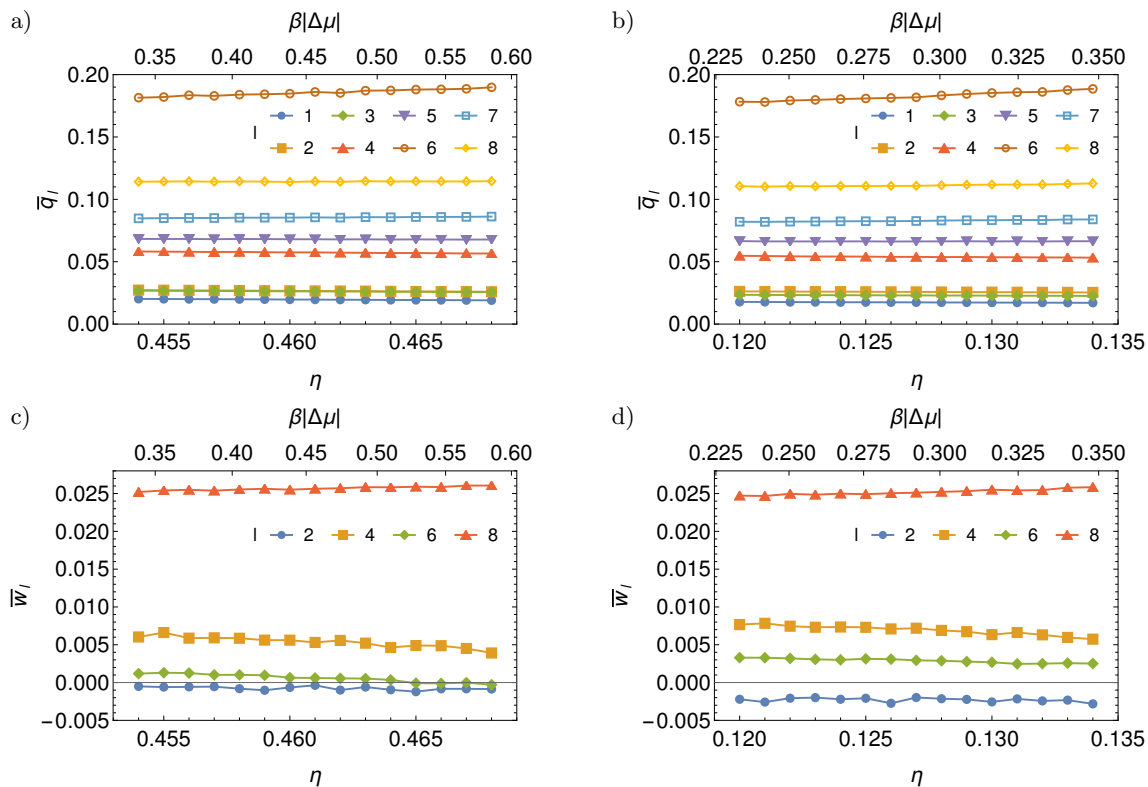


FIG. 1. The mean of a,b) the first eight  $\bar{q}_i$ 's and c,d) the first four even  $\bar{w}_l$ 's as a function of the supersaturation for the metastable fluids of hard-core Yukawa with a,c)  $\beta\epsilon = 81$  and  $1/\kappa\sigma = 0.01$ , and b,d)  $\beta\epsilon = 81$  and  $1/\kappa\sigma = 0.40$ .

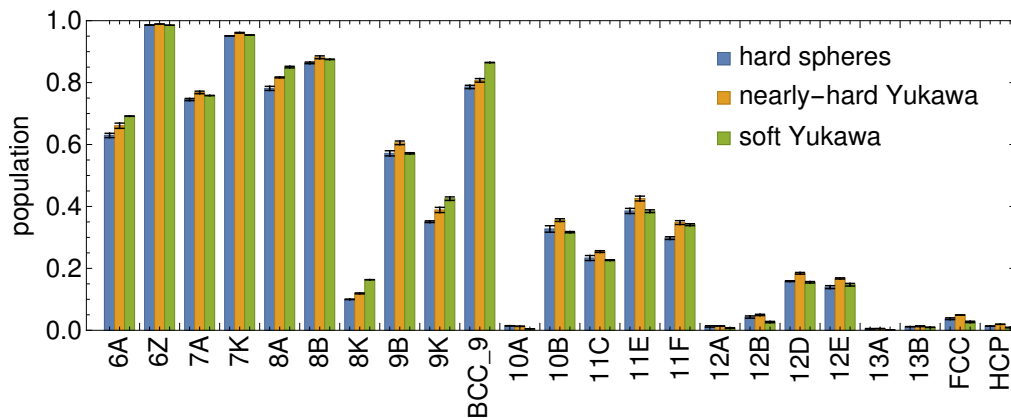


FIG. 2. The population of various TCC clusters in the metastable fluids of hard spheres ( $\eta = 0.5385$ ), nearly-hard hard-core Yukawa particles ( $\beta\epsilon = 8$ ,  $1/\kappa\sigma = 0.04$ ,  $\eta = 0.4400$ ), and soft hard-core Yukawa particles ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.40$ ,  $\eta = 0.1305$ ).

we need to determine the temporal correlations of the local structure such that we know the time window before the start of nucleation during which we can search for a precursor. Figure 5 shows, for multiple simulation methods, the autocorrelation functions (ACFs) of the first two principal components and the local packing fraction in the metastable fluids of hard spheres and of soft spheres. Here, we give the time in terms of the long-time diffusion time  $\tau_d = \sigma^2/6D_l$ , where  $D_l$  is the long-time diffusion coefficient obtained from the mean-squared displacement.

Notice that the ACFs are essentially independent of the simulation method, which confirms that the dynamics are also independent of the choice of simulation method. We see in both systems that PC2 and the local packing fraction decay extremely fast in time. Although PC1 decays more slowly, i.e. within half a diffusion time for the hard spheres and one diffusion time for the soft spheres, this is still relatively fast. The decay time of PC1 provides a good estimate for the time window before the start of nucleation in which we can search for a precursor.

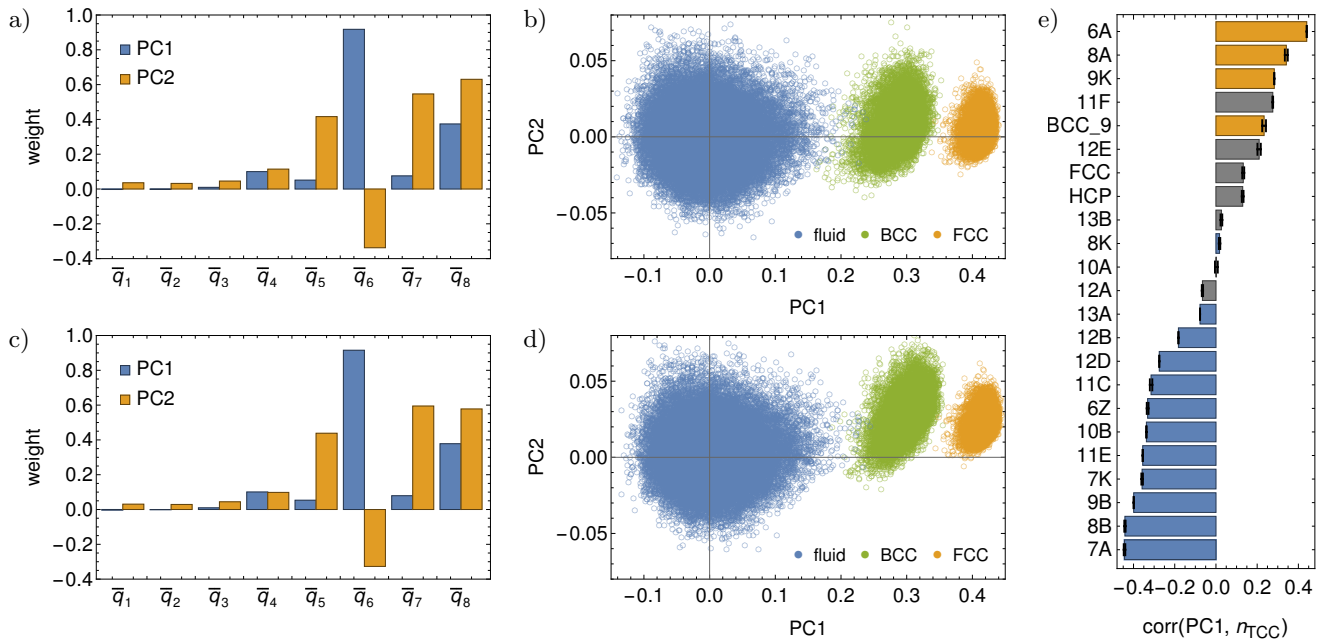


FIG. 3. PCA on the metastable fluids of a,b) hard spheres ( $\eta = 0.5385$ ) and c,d) soft hard-core Yukawa particles ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.40$ ,  $\eta = 0.1305$ ). a,c) give the weight of each BOP in the first and second principal component. b,d) show the distribution of the fluid particles in the PC1-PC2 plane, with the addition of the corresponding bulk FCC phase of hard spheres ( $\eta = 0.5981$ ) and bulk BCC phase of soft hard-core Yukawa ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.40$ ,  $\eta = 0.1311$ ). e) Pearson correlation between PC1 and the number of TCC clusters a particle is involved in for the hard spheres. The color of the bars indicate clusters that consist of one or more tetrahedral subclusters (blue), one or more square pyramidal subclusters (yellow), or both/neither (gray).

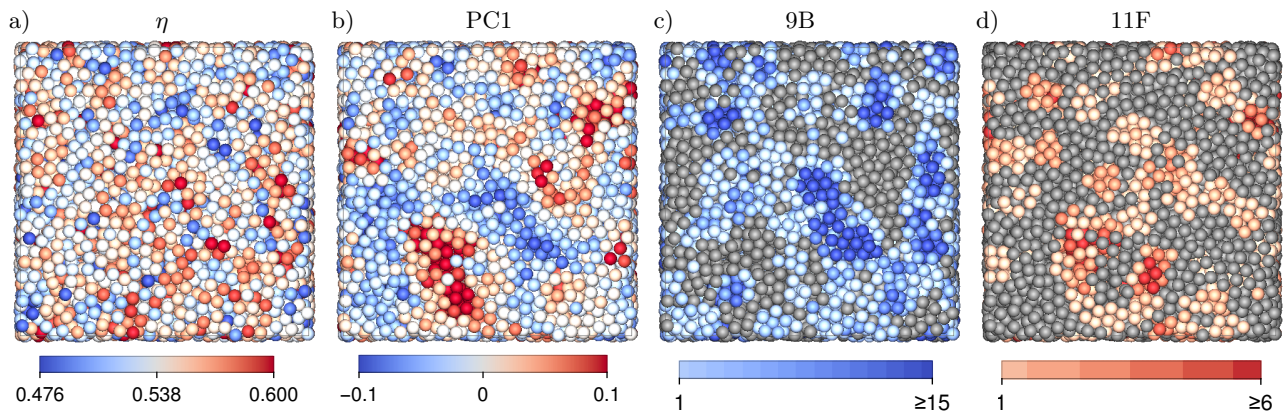


FIG. 4. Snapshot of the metastable fluid of hard spheres ( $\eta = 0.5385$ ) colored by a) the local packing fraction, b) the first principal component, c) the number of 9B clusters per particle, and d) the number of 11F clusters per particle.

## B. Nucleation study

We now turn our attention to crystal nucleation. As explained in the methods, we simulate numerous spontaneous nucleation events using (K)MC and MD simulations, and track for all these events the local properties of the region where nucleation starts. Here, we discuss our observations using two typical nucleation events: one of the hard-spheres system and one of soft spheres ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.40$ ). Both these nucleation events were ob-

tained using MC simulations. More nucleation events, where we either used other simulation methods or studied the other systems mentioned in Tab. I, can be found in the SM. To better illustrate which region we study while tracking a nucleation event, Fig. 6 shows a couple of snapshots of the nucleation event of hard spheres where the particles inside the studied region are colored red. Figure 7 shows for this event and the nucleation event of soft spheres the average properties of the particles in this studied region. Before we discuss what we see,

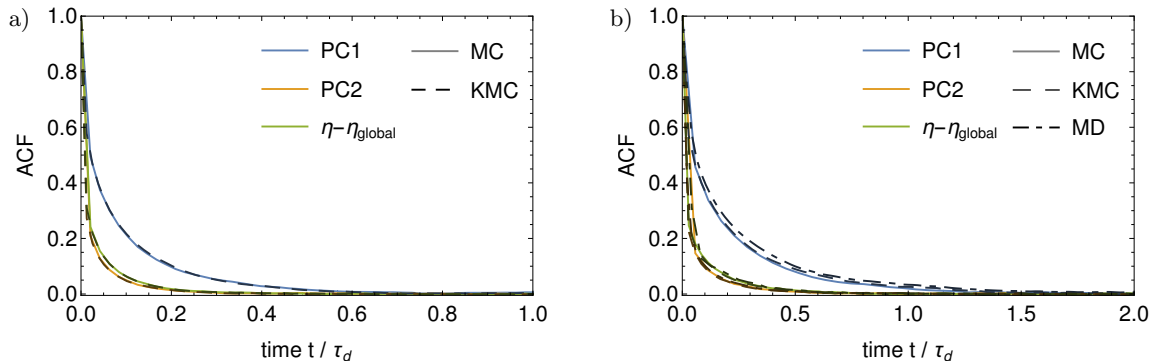


FIG. 5. Autocorrelation function of the first and second principal components and the local packing fraction in the metastable fluids of a) hard spheres ( $\eta = 0.5385$ ) and b) soft hard-core Yukawa particles ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.40$ ,  $\eta = 0.1305$ ). The different dashed and darkness of the color indicate the simulation method, and the time is in terms of the long-time diffusion time  $\tau_d$ .

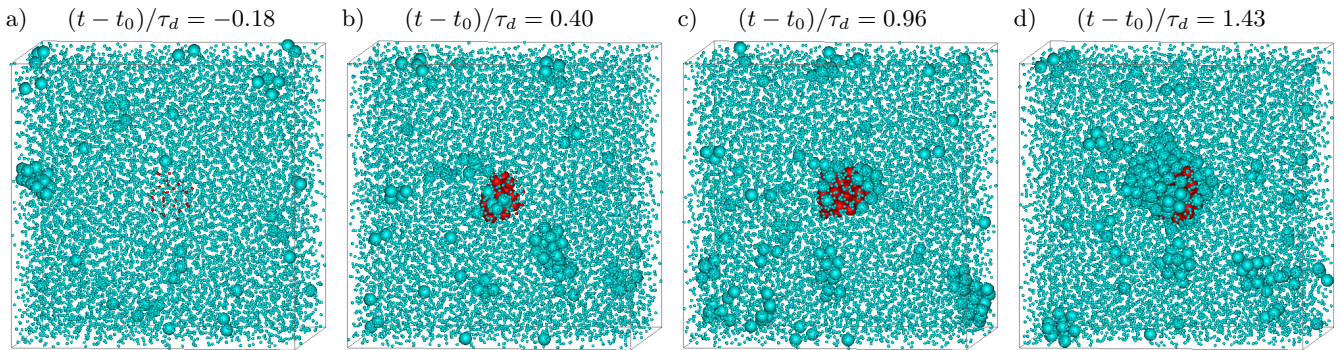


FIG. 6. Four snapshots of a typical nucleation event of hard spheres ( $\eta = 0.5385$ ). Here  $t_0$  indicates the start of nucleation. Fluid particles are displayed at a quarter of their actual size to make the nucleus visible, and red indicates the particles inside the studied region, i.e. those inside the sphere of radius  $R$  around the center of nucleation  $\mathbf{r}_0$ .

let us again point out that the nucleus size (black line) is not an ideal order parameter for tracking nucleation since its binary nature causes it to overlook subtle increases in the local structural ordering at the onset of nucleation. It does, however, provide a general overview of the nucleation event, such as when the nucleus reaches its critical size (see Tab. I). That being said, let us first discuss the BOPs of the studied region. We observe no notable change in the behavior of the BOPs before the start of nucleation, but, as soon as nucleation starts, we see a sharp increase in the values of  $\bar{q}_6$  and  $\bar{q}_8$  for both systems. Furthermore, for the hard spheres, we see that, once nucleation starts,  $\bar{q}_4$  increases,  $\bar{w}_6$  stays negative, and  $\bar{w}_4$  decreases. This all indicates that indeed the FCC phase nucleates. On the other hand, for the soft spheres, we see that as nucleation starts  $\bar{q}_4$  barely increases,  $\bar{w}_6$  stays positive, and  $\bar{w}_4$  keeps fluctuating around zero, which all indicates that indeed the BCC phase nucleates. Similar to the behavior of the BOPs, we observe no notable change in the behavior of PC1 before the start of nucleation, but see a sharp increase in its value once nucleation starts. Note that this increase in PC1 is visible before the number of particles classified as crystalline starts to rise (black line), demonstrating that PC1 is a better order pa-

rameter for tracking the start of nucleation than the nucleus size according to our definition. Lastly, for the hard spheres, we see that the local packing fraction increases simultaneously with PC1 as soon as nucleation starts, and that no notable behavior can be observed before the start of nucleation. This strongly indicates that increase in structural ordering and local density go together and, thus, that there is no apparent precursor. Unfortunately, as the difference between the packing fraction of the fluid and solid phases is extremely small for the soft spheres, i.e. less than 0.001, it is not possible to observe any increase in the local packing fraction on top of the normal fluctuations. Hence, we cannot draw any conclusions on the local packing fraction of the soft spheres.

Next, to show that we have missed no subtle changes in the local structure and thus confirm that there is no precursor for nucleation, we further examine the local structure using TCC. In Fig. 8 we show for four of the most relevant TCC clusters the average number of clusters a particle is involved in and compare it with PC1. These four clusters are: i) 6A, which has the strongest positive correlation with PC1 and is present in both bulk FCC and bulk BCC, ii) 8A, which has the second strongest positive correlation with PC1 and is present in bulk BCC

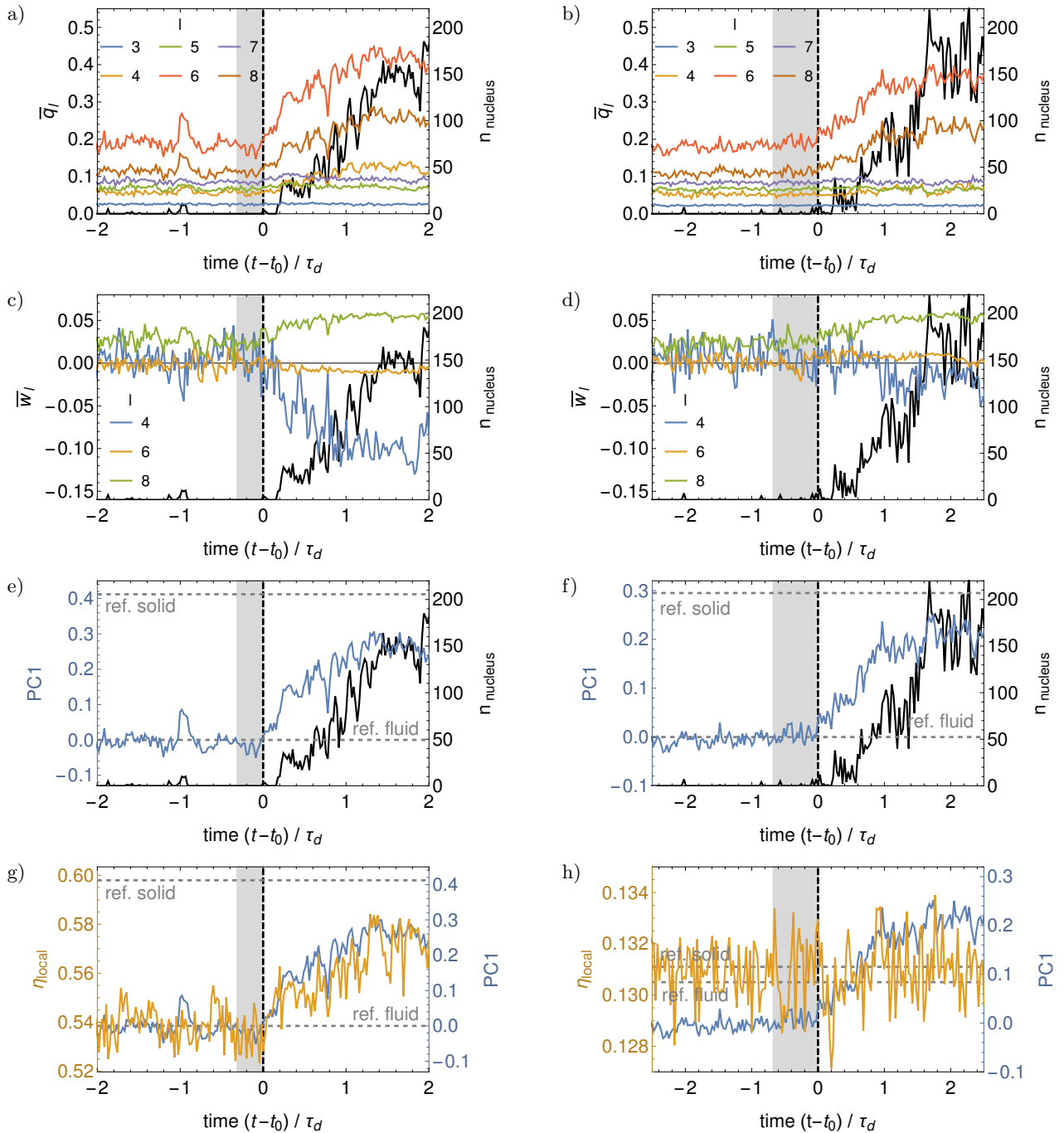


FIG. 7. Left: typical nucleation event of hard spheres ( $\eta = 0.5385$ ), same as in Fig. 6. Right: typical nucleation event of soft hard-core Yukawa particles ( $\beta\epsilon = 81, 1/\kappa\sigma = 0.40, \eta = 0.1305$ ). Both events were obtained using MC simulations. The vertical dashed line in each figure indicates the start of nucleation  $t_0$ , and the shaded area indicates the time window before  $t_0$  for which the ACF of PC1  $> 0.05$  (see Fig. 5). In a-f) the black line (right axis) gives the size of the biggest nucleus present in the studied region. The other lines give the average value of a-b)  $\bar{q}_l$  for  $l \in [3, 8]$ , c-d)  $\bar{w}_l$  for  $l \in [4, 6, 8]$ , e-f) PC1. In g-h) the blue line (right axis) gives PC1, while the yellow line (left axis) gives the local packing fraction  $\eta_{\text{local}}$ . Note that in e-g) the horizontal dashed lines give the reference value of PC1 in the fluid and solid phase. In g) the right axis is scaled in such a way that the reference values in the fluid and solid of PC1 and  $\eta_{\text{local}}$  lie on top of each other. In h) the horizontal dashed lines give the reference value of  $\eta_{\text{local}}$  in the fluid and solid phase.

but not bulk FCC, iii) FCC, which is present in bulk FCC but not bulk BCC, and iv) 7A, which has the strongest

negative correlation with PC1 and can neither be found in bulk FCC nor bulk BCC. Similar figures of other TCC

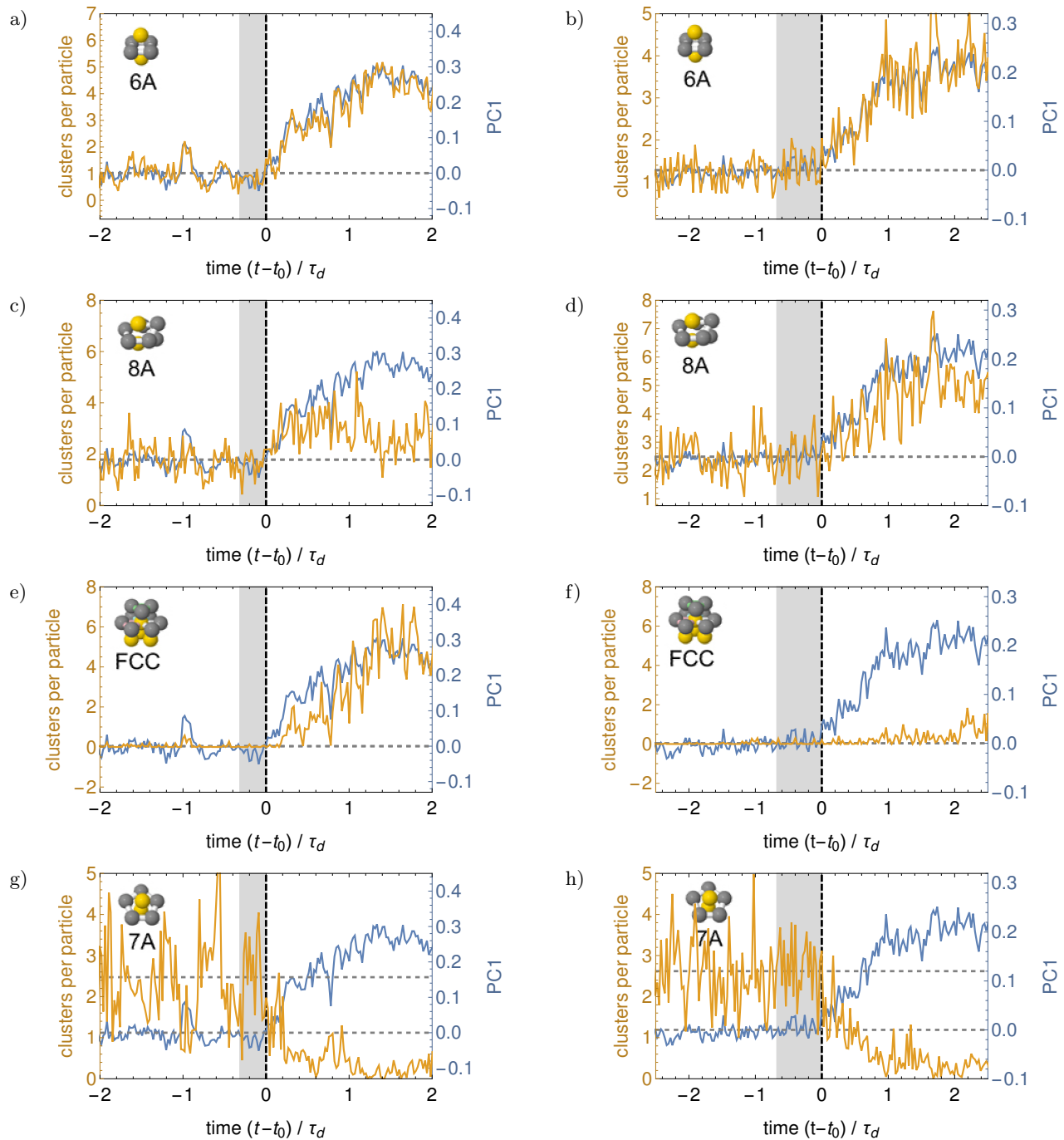


FIG. 8. For the same events as in Fig. 7, i.e. hard spheres (left) and soft hard-core Yukawa particles (right), the average number of clusters per particle (left axis, yellow) for a couple of TCC clusters together with the average value of PC1 (right axis, blue). The horizontal dashed lines indicate the reference values of the number of clusters per particle and PC1 in the fluid. In a-f) the left axis is scaled in such a way that these lines lie on top of each other. In g-h) this was not possible without inverting one of the axes.

clusters can be found in the SM. For all clusters we see that there is no significant change prior to the start of nucleation. Furthermore, we see that the trends of the 6A cluster coincide almost perfectly with those of PC1. Similarly, we see that the trends of the 8A cluster closely follow the trends of PC1. However, for hard spheres the

initial increase in 8A clusters is followed by a decrease. Since 8A is a cluster that is usually found in bulk BCC and not in bulk FCC, this initial increase might be surprising. This can be explained via the observation that 8A clusters are found in high concentrations near the surface of growing nuclei<sup>9</sup>. As a result, the number of

these clusters decreases once the nucleus grows beyond our averaging radius. For the FCC cluster, we observe a sharp increase during the nucleation of hard spheres. Notice, however, that this increase starts slightly later than the increase in PC1. This is not surprising as the FCC cluster is a relatively large cluster, i.e. it contains 13 particles, and consequently is not present in the first stages of nucleation. For the soft spheres there is no significant increase in FCC clusters, as expected. Lastly, we take a look at the 7A cluster. In contrast to the other three clusters, this five-fold symmetric cluster has a strong negative correlation with PC1. Moreover, it is strongly present in the metastable fluid phases, whereas its presence in the FCC and BCC phases is negligible. It is, therefore, not surprising that we observe an immediate and sharp decrease in 7A clusters as soon as nucleation starts.

## V. CONCLUSIONS

To conclude, we have characterized the local structure of various metastable fluids of charged colloids using multiple methods: bond-orientational order parameters (BOPs), principal component analysis (PCA) on the BOPs, and topological cluster classification (TCC). In doing this we have attempted to avoid artefacts due to biases in our chosen order parameters. For all systems we have found that any local structural ordering has a relatively short lifetime, resulting in a short time window prior to the start of nucleation in which a precursor could exist. By tracking the local structure of the spatial region coinciding with the birthplace of the crystal nucleus, we show that inside this time window no atypical behavior in the local structural order is observed using any of our structural order parameters. Furthermore, we demonstrate that all structural characteristics that differ significantly between the fluid and crystal phase start changing simultaneously as soon as nucleation starts. Specifically in the case of FCC, this includes the local density, which starts growing immediately as soon as structural order emerges. We, thus, conclude that we find no evidence for a precursor for the crystal nucleation of hard and charged colloids.

## VI. ACKNOWLEDGEMENTS

L.F. and M.d.J. acknowledge funding from the Vidi research program with project number VI.VIDI.192.102 which is financed by the Dutch Research Council (NWO).

<sup>1</sup>S. Karthika, T. Radhakrishnan, and P. Kalaichelvi, *Crystal Growth & Design* **16**, 6663 (2016).

- <sup>2</sup>G. C. Sosso, J. Chen, S. J. Cox, M. Fitzner, P. Pedevilla, A. Zen, and A. Michaelides, *Chemical reviews* **116**, 7078 (2016).
- <sup>3</sup>W. Ostwald, *Zeitschrift für physikalische Chemie* **22**, 289 (1897).
- <sup>4</sup>S. Alexander and J. McTague, *Phys. Rev. Lett.* **41**, 702 (1978).
- <sup>5</sup>J. Russo and H. Tanaka, *Soft Matter* **8**, 4206 (2012).
- <sup>6</sup>J. Taffs and C. Patrick Royall, *Nat. Commun.* **7**, 13225 (2016).
- <sup>7</sup>W. Ouyang, C. Fu, Z. Sun, and S. Xu, *Phys. Rev. E* **94**, 042805 (2016).
- <sup>8</sup>J. Russo and H. Tanaka, *Sci. Rep.* **2**, 1 (2012).
- <sup>9</sup>W. Gispen, G. M. Coli, R. van Damme, C. P. Royall, and M. Dijkstra, *ACS Nano* **17**, 8807–8814 (2023).
- <sup>10</sup>P. R. ten Wolde and D. Frenkel, *Physical Chemistry Chemical Physics* **1**, 2191 (1999).
- <sup>11</sup>T. Schilling, H. J. Schöpe, M. Oettel, G. Opletal, and I. Snook, *Phys. Rev. Lett.* **105**, 025701 (2010).
- <sup>12</sup>J. Russo and H. Tanaka, *J. Chem. Phys.* **145**, 211801 (2016).
- <sup>13</sup>P. Tan, N. Xu, and L. Xu, *Nat. Phys.* **10**, 73 (2014).
- <sup>14</sup>M. Li, Y. Chen, H. Tanaka, and P. Tan, *Science advances* **6**, eaaw8938 (2020).
- <sup>15</sup>Y.-C. Hu and H. Tanaka, *Nat. Commun.* **13**, 4519 (2022).
- <sup>16</sup>Y. Lu, X. Lu, Z. Qin, and J. Shen, *Solid State Communications* **217**, 13 (2015).
- <sup>17</sup>J. T. Berryman, M. Anwar, S. Dorosz, and T. Schilling, *J. Chem. Phys.* **145**, 211901 (2016).
- <sup>18</sup>H. Tanaka, H. Tong, R. Shi, and J. Russo, *Nature Reviews Physics* **1**, 333 (2019).
- <sup>19</sup>W. Lechner, C. Dellago, and P. G. Bolhuis, *Phys. Rev. Lett.* **106**, 085701 (2011).
- <sup>20</sup>S. Becker, E. Devijver, R. Molinier, and N. Jakse, *Phys. Rev. E* **105**, 045304 (2022).
- <sup>21</sup>W. F. Reinhart, A. W. Long, M. P. Howard, A. L. Ferguson, and A. Z. Panagiotopoulos, *Soft Matter* **13**, 4733 (2017).
- <sup>22</sup>E. Boattini, M. Dijkstra, and L. Filion, *J. Chem. Phys.* **151**, 154901 (2019).
- <sup>23</sup>E. Boattini, S. Marín-Aguilar, S. Mitra, G. Foffi, F. Smalenburg, and L. Filion, *Nat. Commun.* **11**, 1 (2020).
- <sup>24</sup>G. M. Coli and M. Dijkstra, *ACS nano* **15**, 4335 (2021).
- <sup>25</sup>R. van Damme, G. M. Coli, R. van Roij, and M. Dijkstra, *ACS nano* **14**, 15144 (2020).
- <sup>26</sup>A. Gardin, C. Perego, G. Doni, and G. M. Pavan, *arXiv* (2021).
- <sup>27</sup>D. Coslovich, R. L. Jack, and J. Paret, *J. Chem. Phys.* **157**, 204503 (2022).
- <sup>28</sup>J. Paret, R. L. Jack, and D. Coslovich, *J. Chem. Phys.* **152**, 144502 (2020).
- <sup>29</sup>C. S. Adorf, T. C. Moore, Y. J. Melle, and S. C. Glotzer, *J. Phys. Chem. B* **124**, 69 (2019).
- <sup>30</sup>S. Auer and D. Frenkel, *Adv. Comput. Simul.* **173**, 149 (2005).
- <sup>31</sup>C. Desgranges and J. Delhommelle, *The Journal of chemical physics* **126**, 054501 (2007).
- <sup>32</sup>A. R. Browning, M. F. Doherty, and G. H. Fredrickson, *Physical Review E* **77**, 041604 (2008).
- <sup>33</sup>W. Gispen and M. Dijkstra, *Physical Review Letters* **129**, 098002 (2022).
- <sup>34</sup>M. de Jager and L. Filion, *J. Chem. Phys.* **157**, 154905 (2022).
- <sup>35</sup>W. Lechner and C. Dellago, *J. Chem. Phys.* **129**, 114707 (2008).
- <sup>36</sup>J. A. van Meel, L. Filion, C. Valeriani, and D. Frenkel, *J. Chem. Phys.* **136**, 234107 (2012).
- <sup>37</sup>A. Malins, S. R. Williams, J. Eggers, and C. P. Royall, *J. Chem. Phys.* **139**, 234506 (2013).
- <sup>38</sup>P. R. ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, *Faraday Discuss.* **104**, 93 (1996).
- <sup>39</sup>E. Sanz and D. Marenduzzo, *J. Chem. Phys.* **132**, 194102 (2010).
- <sup>40</sup>S. Plimpton, *J. Comp. Phys.* **117**, 1 (1995).
- <sup>41</sup>C. Rycroft, *Chaos* **19**, 041111 (2009).

# Supplemental Material for “In search of a precursor for crystal nucleation of hard and charged colloids”

Marjolein de Jager,<sup>1</sup> Frank Smalenburg,<sup>2</sup> and Laura Filion<sup>1</sup>

<sup>1</sup>*Soft Condensed Matter, Debye Institute of Nanomaterials Science, Utrecht University, Utrecht, Netherlands*

<sup>2</sup>*Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405 Orsay, France*

## I. NUCLEATION BARRIERS

In the main paper, we give the barrier height and critical nucleus size of systems studied for their spontaneous nucleation. Here we provide some more background on these systems and their nucleation barriers. To compute the nucleation barriers, we use Monte Carlo (MC) simulations in the  $NPT$ -ensemble combined with umbrella sampling. We bias the simulations using the number of particles in the nucleus measured via the Ten Wolde order parameter<sup>7</sup> with cutoff values  $d_c = 0.7$  and  $\xi_c = 6$ . The nearest neighbors for this order parameter are determined using a radial cutoff  $r_c$ , which is chosen to be approximately the position of the first minimum of the radial distribution function. The full methods are explained in our previous work<sup>7</sup>. For all nucleation barriers, we use a biasing strength of  $\lambda = 0.02$  and bias around the target nucleus size  $n_c$ , where we take  $n_c$  with an interval of 10. We simulate  $N = 10976$  and  $N = 11664$  particles for the systems forming the face-centered cubic (FCC) and body-centered cubic (BCC) phase, respectively. Furthermore, to obtain more accurate nucleation barriers, we perform 2 to 4 independent runs for each window. The resulting nucleation barriers are shown in Fig. S1 and Tab. S1 provides additional information on these barriers, such as the interfacial free energy obtained from fitting the nucleation barrier to classical nucleation theory (see Ref. ? ). Note that we use the radial cutoff  $r_c$  for determining nearest neighbors in the fluid-solid classification, which is done on the fly, whereas for Lechner and Dellago’s bond-orientational order parameters (BOPs)<sup>7</sup>, which are computed in the post processing, we use the SANN algorithm<sup>7</sup> to determine the nearest neighbors.

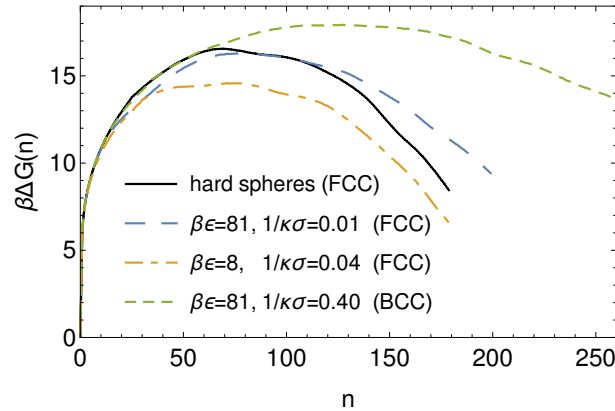


FIG. S1. Nucleation barriers of (nearly-)hard spheres forming an FCC crystal and soft spheres forming a BCC crystal. For more information see Tab. S1.

TABLE S1. For each system studied, the packing fraction of the supersaturated fluid  $\eta^*$  at which the brute force nucleation is performed together with the corresponding packing fraction of the solid phase  $\eta_s$ , pressure  $\beta P\sigma^3$ , and supersaturation  $\beta|\Delta\mu|$ . The last columns give cutoff radius  $r_c/\sigma$  as well as the reduced interfacial free energy  $\beta\gamma/\rho_s^{2/3}$ , critical nucleus size  $n^*$ , and barrier height  $\beta\Delta G^*$  obtained from fitting the nucleation barriers. The error in  $\beta\Delta G^*$  is no more than 1.

	$\beta\epsilon$	$1/\kappa\sigma$	$\eta^*$	$\eta_s$	$\beta P\sigma^3$	$\beta \Delta\mu $	$r_c/\sigma$	$\beta\gamma/\rho_s^{2/3}$	$n^*$	$\beta\Delta G^*$
		hard spheres	0.5385	0.5981	17.5	0.585	1.40	0.52	75	16.5
FCC	81	0.01	0.4681	0.5190	15.4	0.584	1.46	0.49	84	16.3
	8	0.04	0.4400	0.4711	16.7	0.541	1.49	0.44	69	14.8
BCC	81	0.40	0.1305	0.1311	21.3	0.321	2.32	0.35	122	18.0

## II. DISTRIBUTION OF BOPS IN THE METASTABLE FLUIDS

In Fig. 1 of the main text, we show the average values of Lechner and Dellago's BOPs, both  $\bar{q}_l$ 's and  $\bar{w}_l$ 's, in the metastable fluids of nearly hard spheres and of soft spheres as a function of the supersaturation. Remember that the metastable fluid of nearly hard spheres will later nucleate the FCC phase, while the one of soft spheres will nucleate the BCC phase. From this figure we observe that the most prevalent difference between the two systems can be found in  $\bar{w}_6$ . Here we focus more on this difference by performing a similar analysis to Russo and Tanaka's analysis in Refs. ???. This analysis is based on the observation that  $\bar{w}_6$  can be used to distinguish between the BCC crystal and close-packed crystals (FCC and/or HCP), since  $\bar{w}_6 > 0$  for the BCC crystal and  $\bar{w}_6 < 0$  for the FCC and HCP crystals. Furthermore,  $\bar{w}_4$  can be used to distinguish FCC from HCP, as  $\bar{w}_4 < 0$  for the FCC crystal and  $\bar{w}_4 > 0$  for the HCP crystal. Thus, by looking at the probability distributions of  $\bar{w}_6$  and  $\bar{w}_4$  in the metastable fluid for different threshold values of  $\bar{q}_6$ , Russo and Tanaka demonstrated that the more crystal-like regions in the metastable fluid of hard spheres tend more towards the FCC phase, than the HCP or BCC phase.

In Fig. S2, we show the probability distribution functions of  $\bar{w}_6$  and  $\bar{w}_4$  for the metastable fluid of hard spheres, nearly hard spheres ( $\beta\epsilon = 8$ ,  $1/\kappa\sigma = 0.04$ ), and soft spheres ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.40$ ). The black line in each

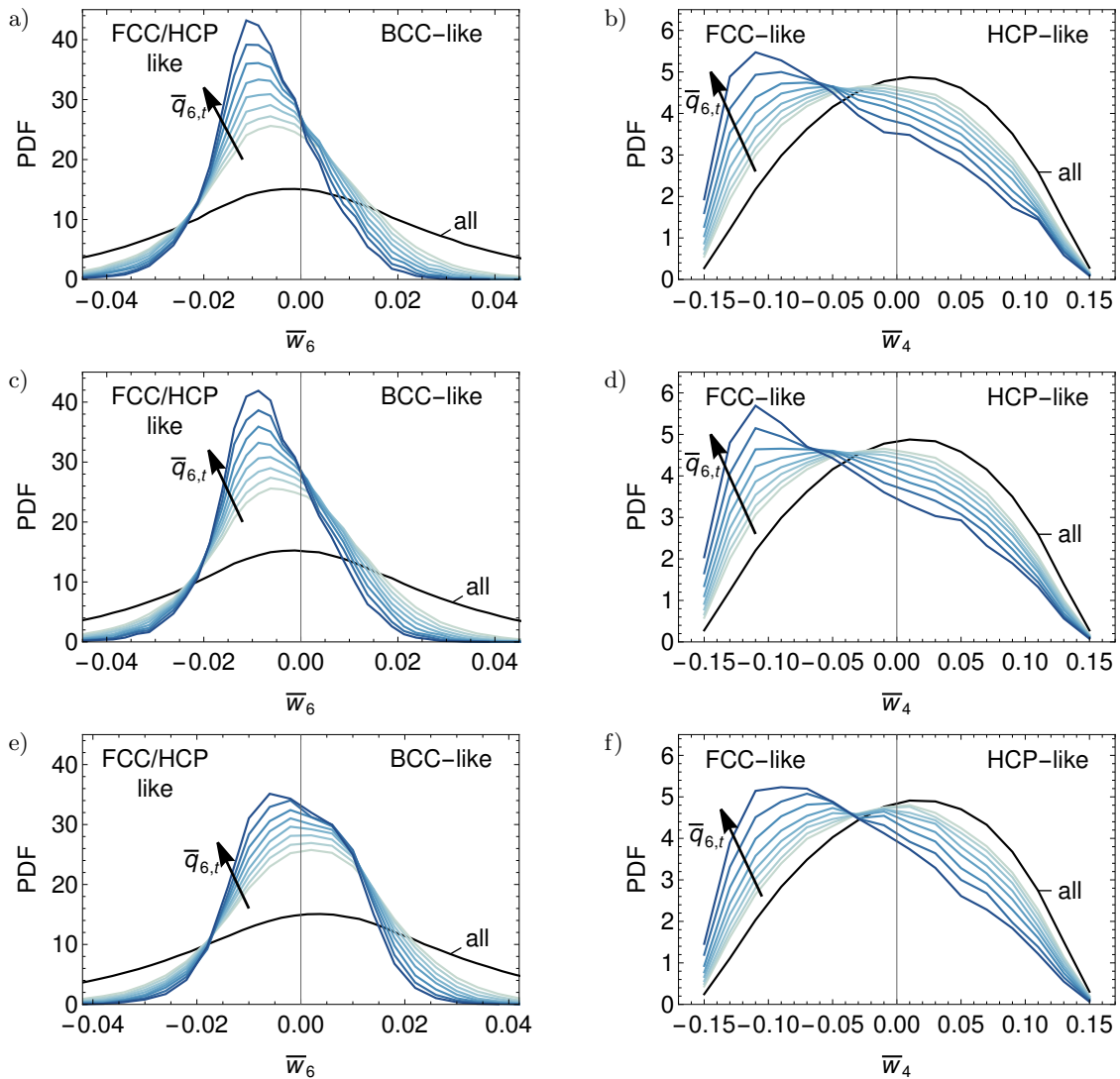


FIG. S2. The probability distribution of a,c,e)  $\bar{w}_6$  and b,d,f)  $\bar{w}_4$  in the metastable fluids of a,b) hard spheres ( $\eta = 0.5385$ ) soft hard-core Yukawa particles with c,d)  $\beta\epsilon = 8$  and  $1/\kappa\sigma = 0.04$  ( $\eta = 0.1305$ ), and e,f)  $\beta\epsilon = 81$  and  $1/\kappa\sigma = 0.40$  ( $\eta = 0.1305$ ). In each figure, the black line gives the total distribution and the eight lines with the blue color gradient give the distribution for  $\bar{q}_6 > \bar{q}_{6,t}$ , where the arrow indicates the increase in the threshold value from  $\bar{q}_{6,t} = 0.25$  to  $\bar{q}_{6,t} = 0.32$ .

figure shows the probability distribution for the entire system, while the lines with the blue color gradient show the probability distributions for particles with  $\bar{q}_6 > \bar{q}_{6,t}$  for increasing value of the threshold  $\bar{q}_{6,t}$ . The probability distributions for hard spheres and nearly hard spheres are practically identical to the ones shown by Russo and Tanaka in Ref. ? . From these figures one can draw the conclusion that high  $\bar{q}_6$  regions show a predominant preference for negative  $\bar{w}_6$  and negative  $\bar{w}_4$ , and thus correspond to FCC-like symmetries. Furthermore, this preference increases with increasing  $\bar{q}_{6,t}$ . The probability distributions of the soft spheres, on the other hand, show a slight preference for positive  $\bar{w}_6$ , which corresponds to BCC-like symmetries. However, notice that the distribution shifts towards FCC-like for increasing  $\bar{q}_{6,t}$ .

### III. RADIAL CORRELATIONS

In the main text, we qualitative discuss the spatial correlations in the local packing fraction and first principal component (PC1). Here, we quantify the spatial correlations. Figure S3 shows the radial distribution function together with the radial correlations of PC1 and the local packing fraction for the metastable fluids of hard spheres and of soft spheres. We clearly see that PC1 has a measurable correlation even 4 or 5 shells removed, whereas the correlation of the local packing fraction decays to zero within the first two shells.

### IV. MORE FIGURES AND MORE NUCLEATION EVENTS

In the main paper, we discuss our observations using two typical nucleation events: one of the hard-spheres system and one of the soft spheres ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.40$ ). Here, we show more tracking results on these two events, as well as the results on some other events. As in the main paper, all figures show the average values inside the studied region, i.e. the small spatial region containing the birthplace of the crystal nucleus. Moreover, the layout of the figures is the same as in the main paper. The vertical dashed line in each figure indicates the start of nucleation  $t_0$ , the shaded area indicates the time window before  $t_0$  for which the ACF of the first principal component (PC1) is greater than 0.05, and, if present, the horizontal dashed lines give the reference value of a parameter in the fluid and sometimes in the solid phase as well.

Figure S4 shows for the same events as in the main paper the average number of a certain TCC cluster a particle is involved in. It is an addition to Fig. 8 of the main paper, which already gives these results for the 6A, 8A, FCC, and 7A clusters. Again, we see that there is no significant change prior to the start of nucleation. The BCC\_9, 9K, and 11F clusters all increase as soon as nucleation starts, while the 7K and 8B clusters both decrease as soon as nucleation starts. This is to be expected as the former group have a positive correlation with PC1 and the latter a negative. However, notice that BCC\_9 is more present in the nucleation of the FCC phase than of the BCC phase. This is not entirely surprising as bulk FCC contains more BCC\_9 clusters than bulk BCC. In contrast, the 9K and 11F clusters are found in bulk BCC and not in bulk FCC. Hence, similar to the 8A cluster, we first observe an initial increase in 9K and 11F clusters for the nucleation of the FCC phase, later this increase is followed by a decrease.

The two events discussed in the main paper were obtained using MC simulations. In order to demonstrate that our

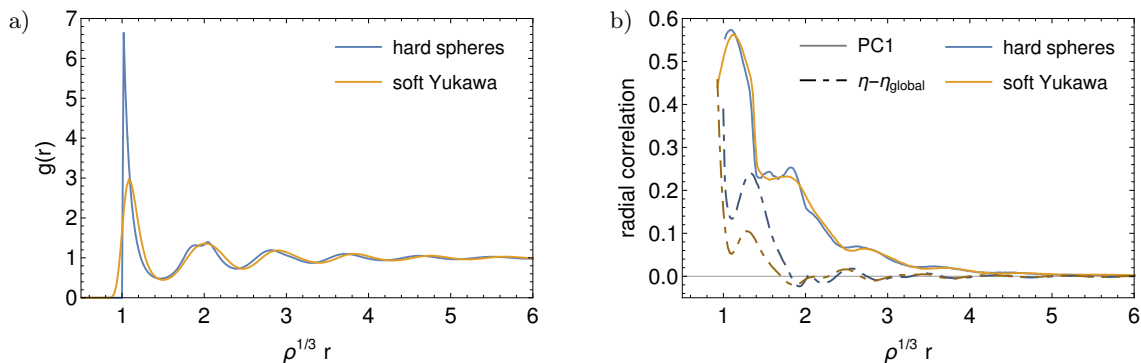


FIG. S3. a) radial distribution function and b) radial correlation of the first principal component (solid) and local packing fraction (dashed) in the metastable fluids of hard spheres ( $\eta = 0.5385$ ) and soft hard-core Yukawa particles ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.40$ ,  $\eta = 0.1305$ ). Note that the horizontal axes are scaled by the cubic root of the density such that the first peaks of the  $g(r)$  are located at the same position.

observations and conclusions do not depend on the specific dynamics of the simulations method, we also show the results on three more nucleation events of the same systems but then obtained using a different simulation method. In Fig. S5 we show the tracking results for a typical nucleation event of the hard-spheres system obtained using kinetic MC (KMC), and in Figs. S6 and S7 two of the soft spheres obtained using, respectively, KMC and molecular dynamics (MD). The results are very similar to those of the nucleation events obtained using MC. Again we observe no anomalous behavior in the metastable fluid prior to the start of nucleation, and all local order parameters abruptly change as soon as nucleation starts. We, thus, conclude that our results and conclusions do not depend on the specific simulation method used.

Lastly, we show a typical nucleation event of a system of essentially-hard spheres ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.01$ ), see Fig. S8, and of a system of nearly-hard spheres ( $\beta\epsilon = 8$ ,  $1/\kappa\sigma = 0.04$ ), see Fig. S9. Again we observe no anomalous behavior in the metastable fluid prior to the start of nucleation, and all local order parameters abruptly change as soon as nucleation starts.

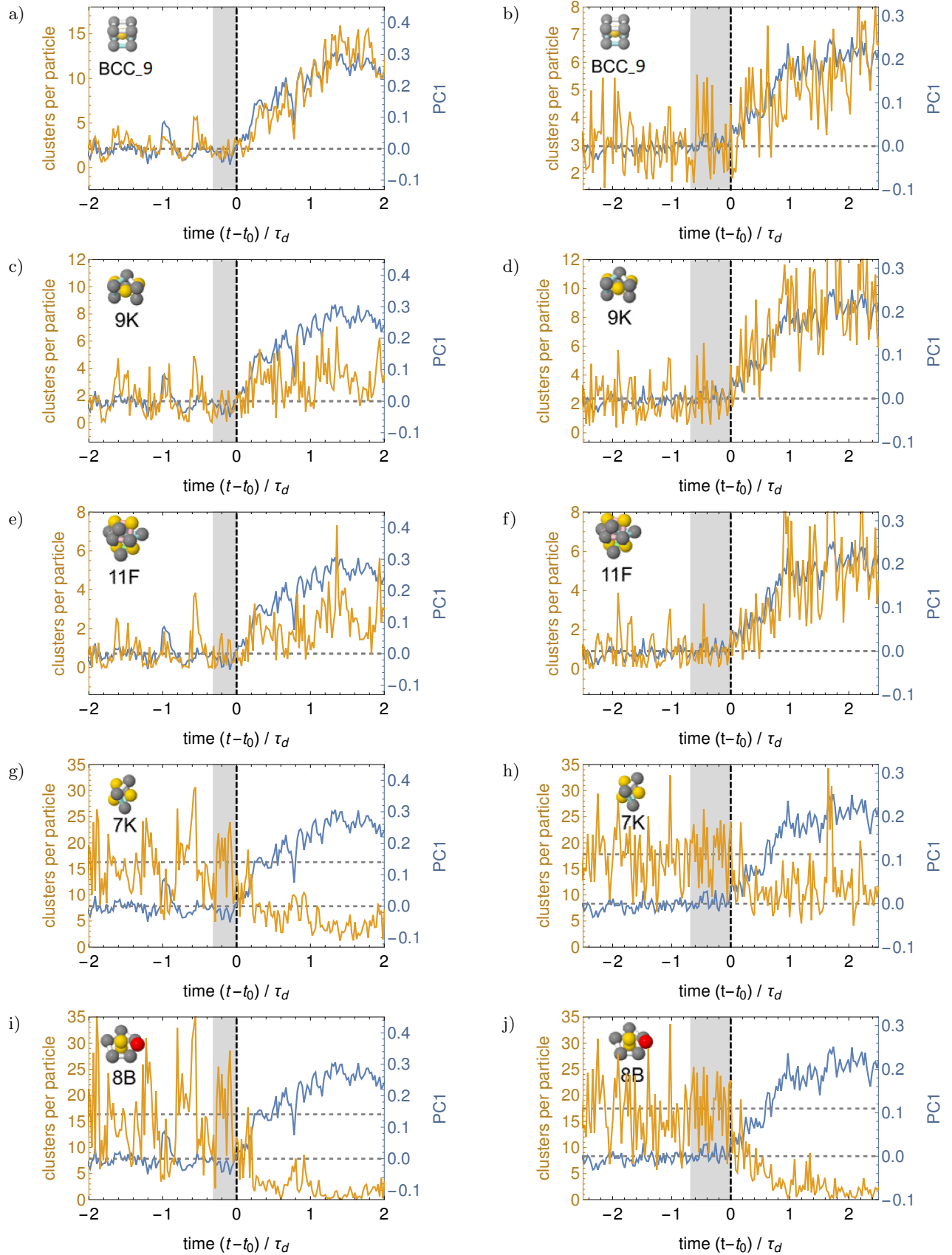


FIG. S4. For the same events as in Figs. 7 and 8 of the main paper, i.e. hard spheres (left) and soft hard-core Yukawa particles (right), the average number of clusters per particle (left axis, yellow) for a couple of TCC clusters together with the average value of PC1 (right axis, blue). The horizontal dashed lines indicate the reference values of the number of clusters per particle and PC1 in the fluid. In a-f) the left axis is scaled in such a way that these lines lie on top of each other. In g-j) this was not possible without inverting one of the axes.

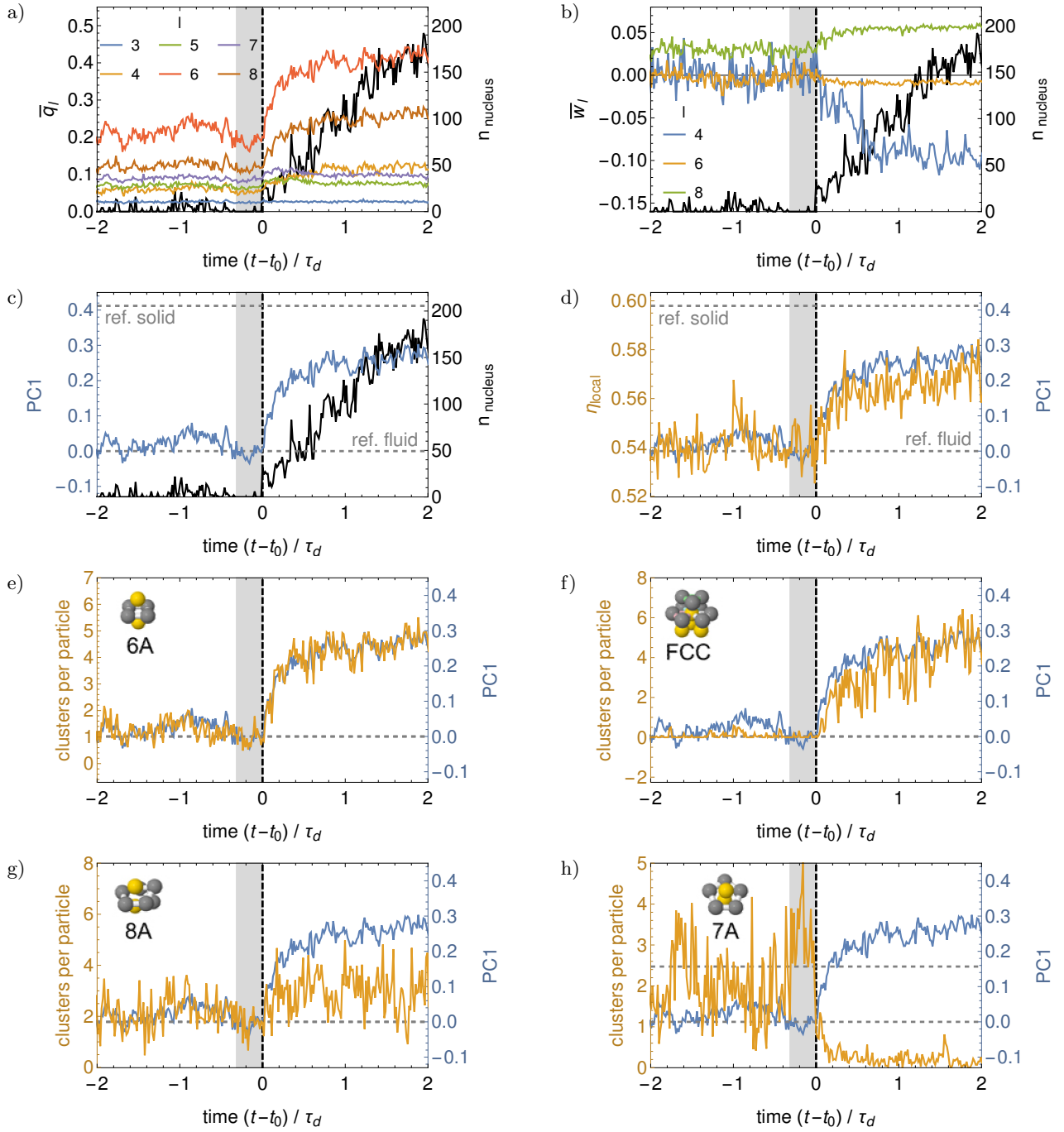


FIG. S5. Typical nucleation event of hard spheres ( $\eta = 0.5385$ ) obtained using KMC. In a-c) the black line (right axis) gives the size of the biggest nucleus present in the studied region, while the other lines (left axis) give the average value of a-b) the BOPs  $\bar{q}_l$  and  $\bar{w}_l$ , and c) PC1. In d-h) blue line (right axis) gives PC1, while the yellow line (left axis) gives d) the local packing fraction and e-h) the average number of clusters per particle for a couple of TCC clusters.

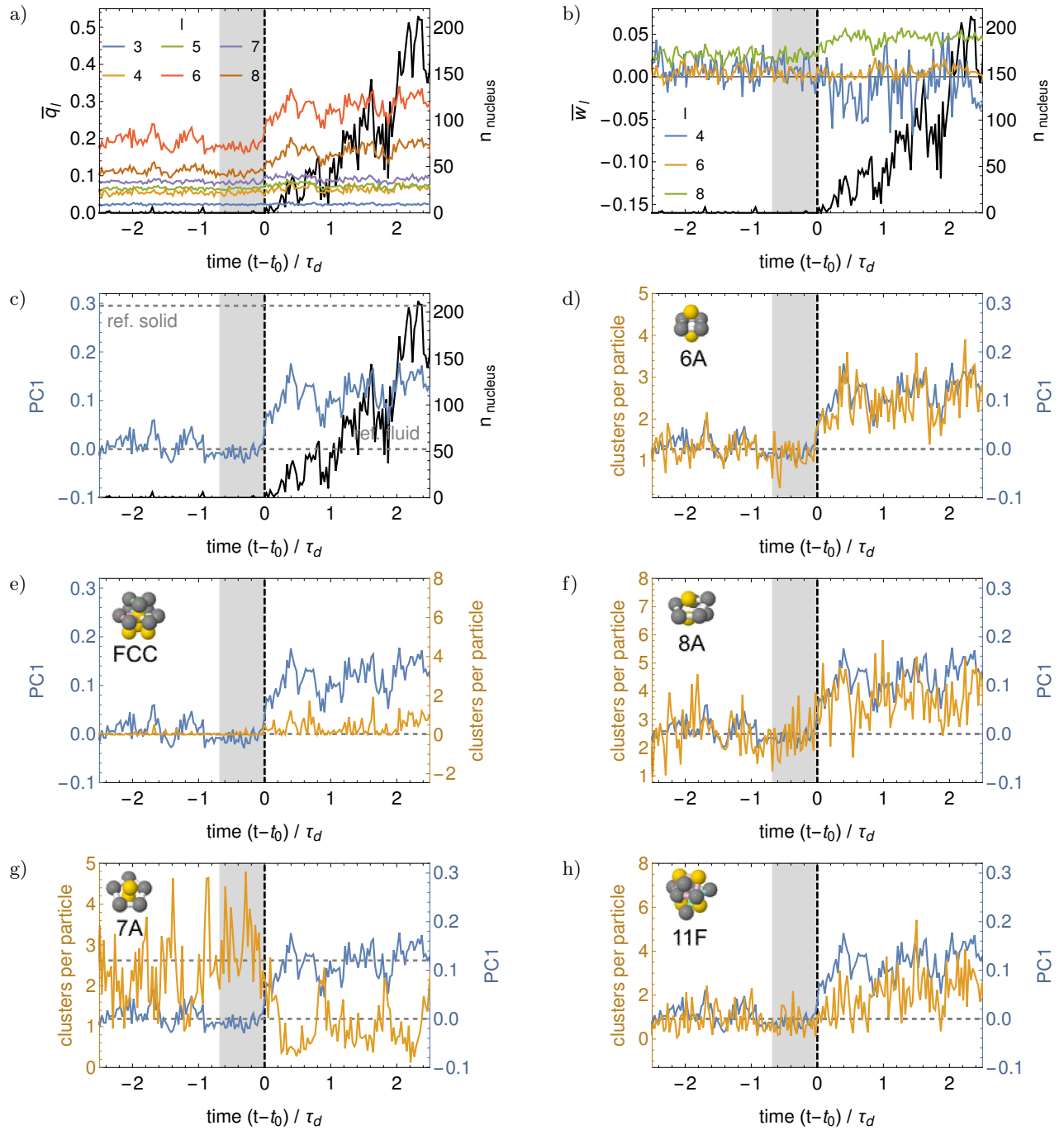


FIG. S6. Typical nucleation event of soft hard-core Yukawa particles ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.40$ ,  $\eta = 0.1305$ ) obtained using KMC. In a-c) the black line (right axis) gives the size of the biggest nucleus present in the studied region, while the other lines (left axis) give the average value of a-b) the BOPs  $\bar{q}_l$  and  $\bar{w}_l$ , and c) PC1. In d-h) blue line (right axis) gives PC1, while the yellow line (left axis) gives the average number of clusters per particle for a couple of TCC clusters.

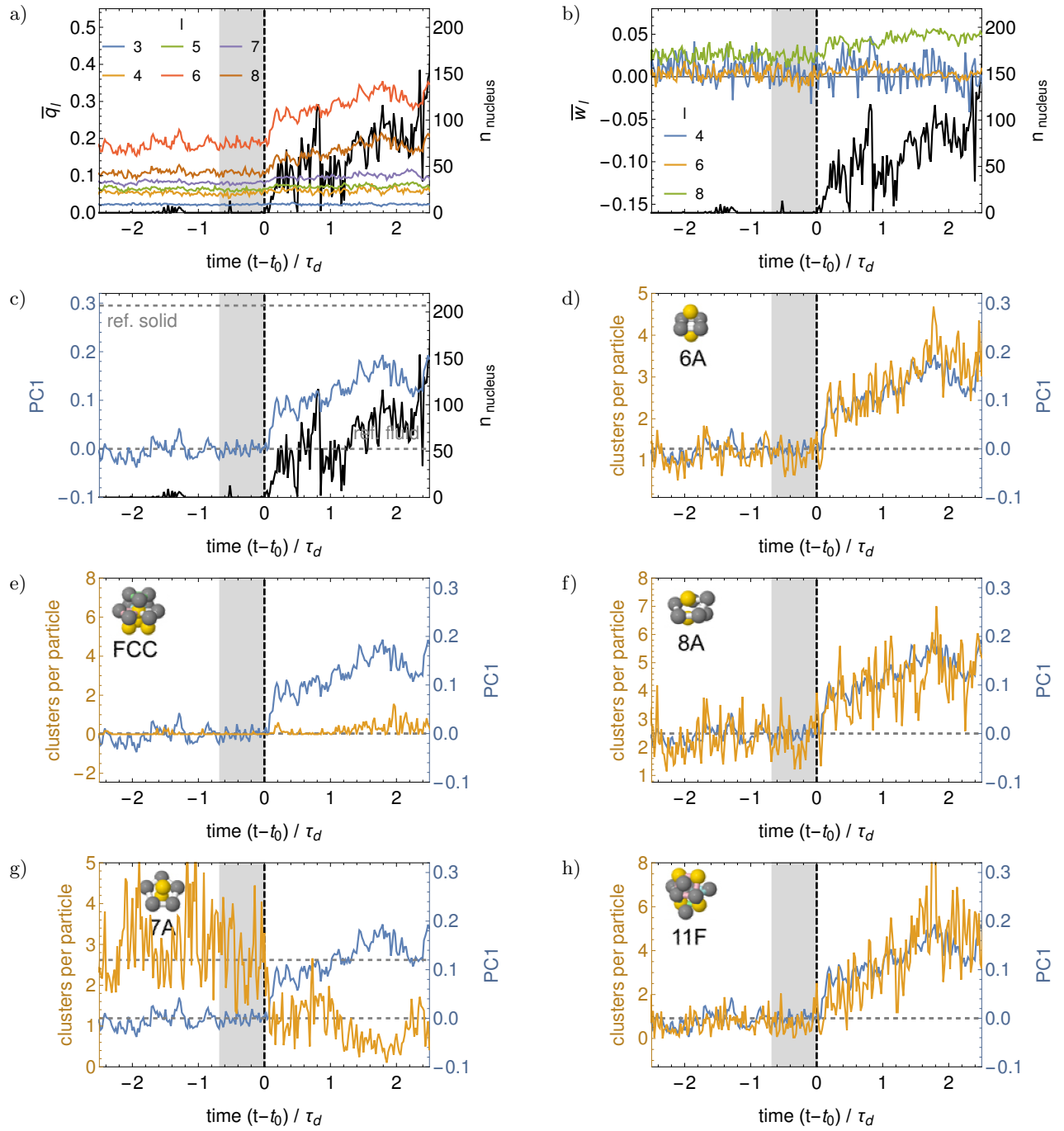


FIG. S7. Typical nucleation event of soft hard-core Yukawa particles ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.40$ ,  $\eta = 0.1305$ ) obtained using MD. In a-c) the black line (right axis) gives the size of the biggest nucleus present in the studied region, while the other lines (left axis) give the average value of a-b) the BOPs  $\bar{q}_i$  and  $\bar{w}_i$ , and c) PC1. In d-h) blue line (right axis) gives PC1, while the yellow line (left axis) gives the average number of clusters per particle for a couple of TCC clusters.

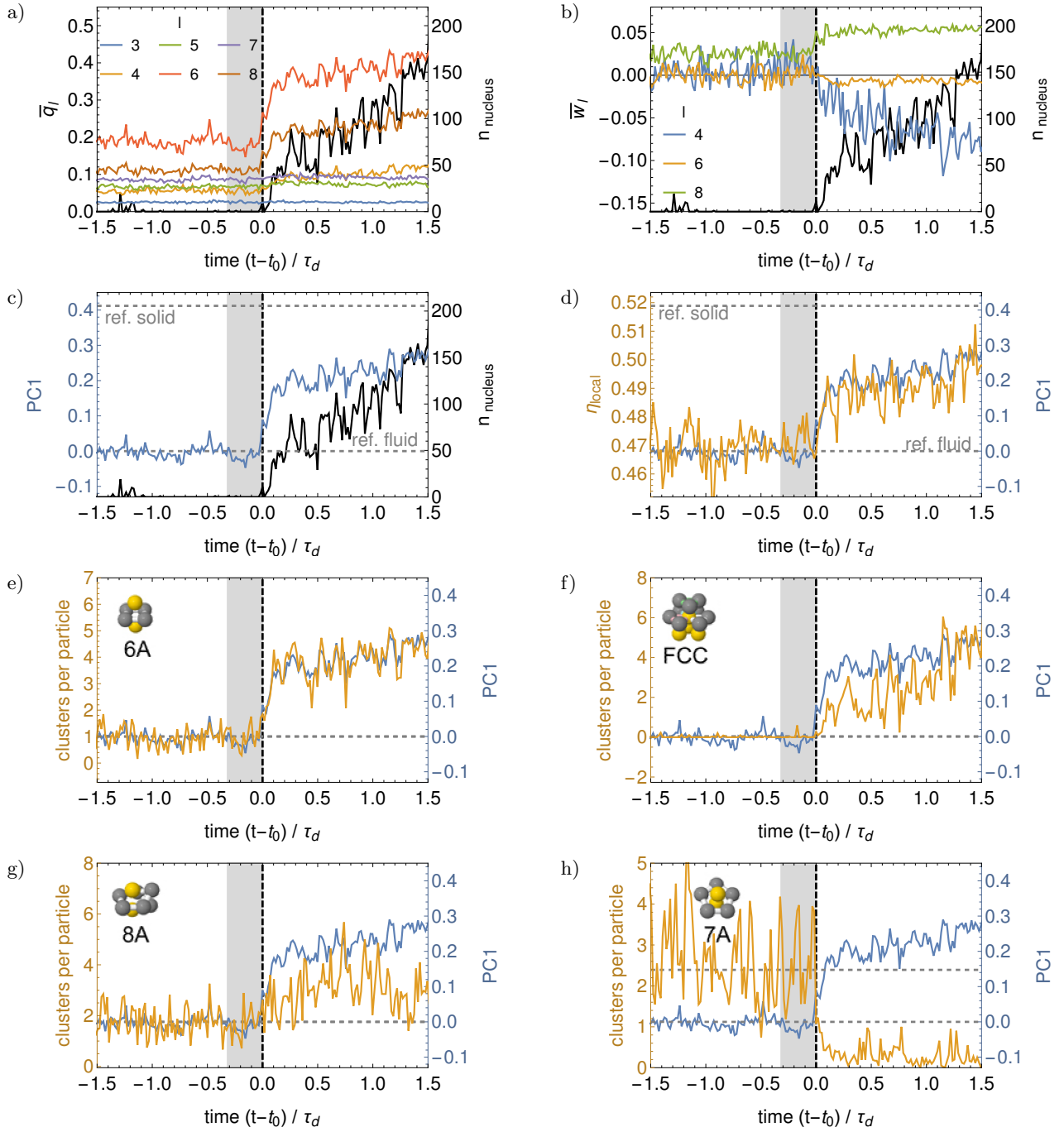


FIG. S8. Typical nucleation event of essentially-hard hard-core Yukawa particles ( $\beta\epsilon = 81$ ,  $1/\kappa\sigma = 0.01$ ,  $\eta = 0.4681$ ) obtained using MC. In a-c) the black line (right axis) gives the size of the biggest nucleus present in the studied region, while the other lines (left axis) give the average value of a-b) the BOPs  $\bar{q}_i$  and  $\bar{w}_i$ , and c) PC1. In d-h) blue line (right axis) gives PC1, while the yellow line (left axis) gives d) the local packing fraction and e-h) the average number of clusters per particle for a couple of TCC clusters.

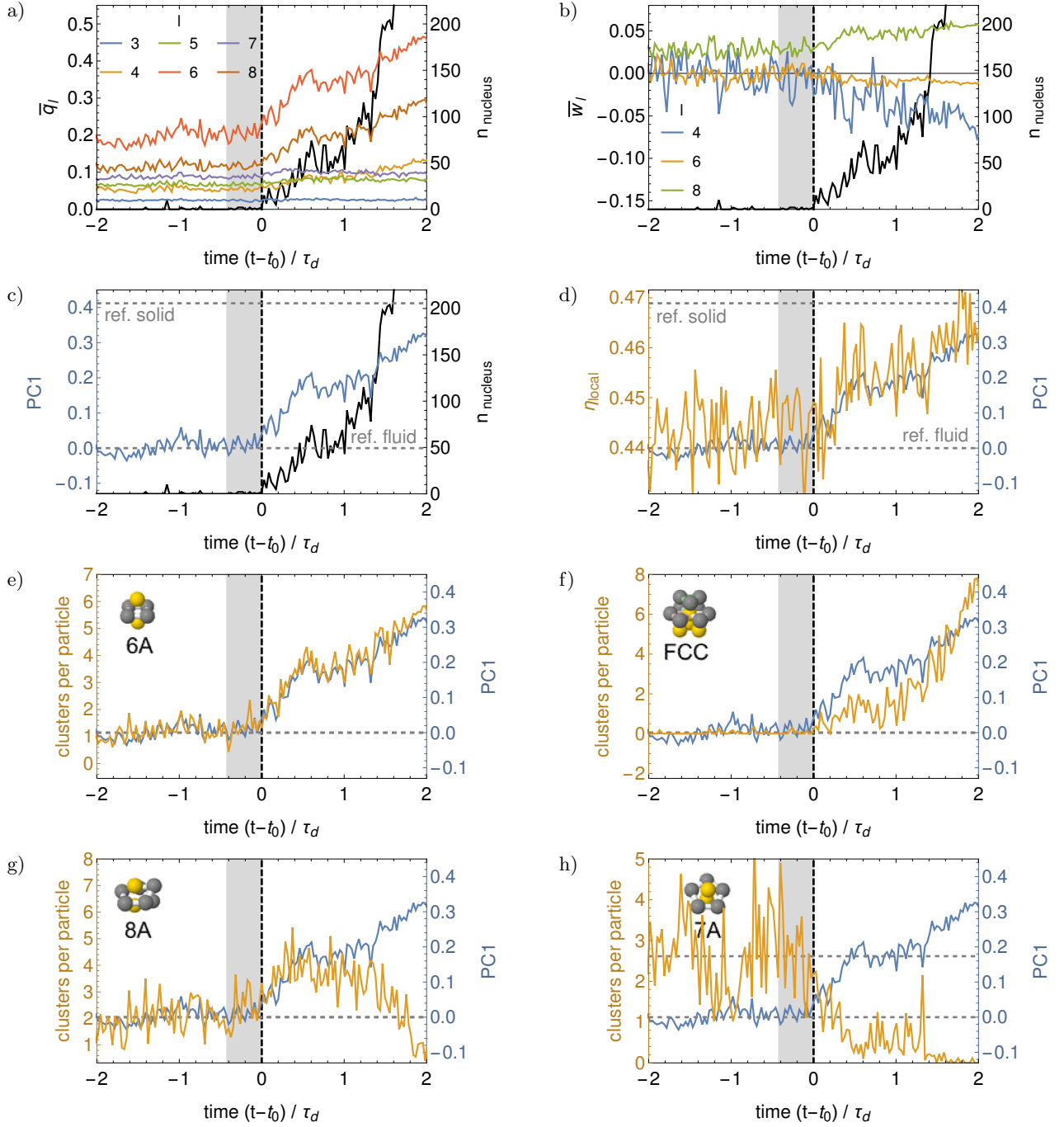


FIG. S9. Typical nucleation event of nearly-hard hard-core Yukawa particles ( $\beta\epsilon = 8$ ,  $1/\kappa\sigma = 0.04$ ,  $\eta = 0.4400$ ) obtained using MC. In a-c) the black line (right axis) gives the size of the biggest nucleus present in the studied region, while the other lines (left axis) give the average value of a-b) the BOPs  $\bar{q}_i$  and  $\bar{w}_i$ , and c) PC1. In d-h) blue line (right axis) gives PC1, while the yellow line (left axis) gives d) the local packing fraction and e-h) the average number of clusters per particle for a couple of TCC clusters.