

Long-wavelength density fluctuations as nucleation precursorsJames F. Lutsko¹ and Julien Lam¹*Center for Nonlinear Phenomena and Complex Systems, Code Postal 231, Université Libre de Bruxelles, Blvd. du Triomphe, 1050 Brussels, Belgium* (Received 8 October 2019; accepted 22 April 2020; published 18 May 2020)

Recent theories of nucleation that go beyond classical nucleation theory predict that diffusion-limited nucleation of both liquid droplets and of crystals from a low-density vapor (or weak solution) begins with long-wavelength density fluctuations. This means that in the early stages of nucleation, “clusters” can have low density but large spatial extent, which is at odds with the classical picture of arbitrarily small clusters of the condensed phase. We present the results of kinetic Monte Carlo simulations using forward flux sampling to show that these predictions are confirmed, namely, that on average, nucleation begins in the presence of low-amplitude, but spatially extended density fluctuations thus confirming a significant prediction of the nonclassical theory.

DOI: [10.1103/PhysRevE.101.052122](https://doi.org/10.1103/PhysRevE.101.052122)**I. INTRODUCTION**

Nucleation is a widely studied phenomenon that plays an important role in such diverse subjects as the function of cells [1], stellar evolution [2], and superhydrophobicity [3]. The particular example of crystallization is of great importance in chemistry, materials science, and fundamental physics and has been found to be far more complex than once believed. For this reason, the study of nucleation pathways—the steps by which thermal fluctuations build up a critical cluster—has become a major focus of modern work by experimentalists, simulators, and theorists alike.

In classical nucleation theory [4] (CNT) the nucleation of new phases is assumed to begin with the formation of small clusters that then evolve via the stochastic addition and loss of material. For example, in the most naive version, the formation of a crystal is supposed to begin with small oligomers of a few growth units arranged in the final crystalline form. This crystallite then grows while retaining its ordered structure. Droplets nucleating from a vapor and the converse, bubbles nucleating from a liquid, are imagined to proceed in an analogous manner. The idea that some sort of precursor process could play a role is one that has been suggested in several contexts. For example, Shen and Debenedetti [5] reported on Monte Carlo simulations showing large ramified voids as precursors to bubble nucleation in superheated fluids. These results were challenged by Wang *et al.* [6] who used dynamical simulations and who reported compact subcritical structures. They attribute the difference in the results to the choice of order parameter (local versus global). Intriguingly, these authors did report on another type of precursor, namely, that bubble nucleation seemed to preferentially begin in “hot spots” of locally elevated temperature.

There has long been speculation that local concentration would play a role in the nucleation of condensed phases such as a droplet from a liquid. In the 1960s, Russell [7]

coupled the usual Becker-Doring dynamics of CNT to simple models of the surrounding fluid and concluded that critical nuclei were likely to occur in enriched areas of the system. As Russell notes, when material clumps together to form a dense cluster, the surrounding system is necessarily depleted unless new material flows in to replace it. In his pioneering work on field-theoretic approaches to nucleation, Langer [8,9] also discusses the importance of coupling cluster formation to the local concentration but his work was mostly concerned with determining the nucleation rate and so focused on the end point of nucleation—the critical cluster. More recently, Peters [10] discussed a more detailed model of the coupling between the local concentration and cluster dynamics and concluded that local concentration gradients play a key role in determining the fate of clusters.

One approach that fully couples the fluctuations leading to cluster formation and transport processes is mesoscopic nucleation theory (MeNT) which combines classical density functional theory methods with fluctuating hydrodynamics and techniques from stochastic process and large deviation theory to give a complete, unconstrained framework for studying nonclassical pathways [11–13]. Recent successes of this approach have been the establishment of the connection between it and CNT [14], demonstrating the latter as a particularly simple approximation to the former, and the unbiased description of pathways for crystallization [13].

The main application of MeNT has so far been to diffusion-limited nucleation such as for liquid-liquid phase separation or crystallization of large molecules in solution. One key difference in the predicted pathways in this case, with respect to CNT, but consistent with the earlier work mentioned, is that they always begin with a long-wavelength, low-amplitude density fluctuation [11,12]. These can be thought of as clusters which are large in spatial extent but have a density only slightly greater than the background (see Fig. 1) or, alternatively, not as a single structure but simply a region of the system with a slightly elevated density. Such regions always exist due to thermal fluctuations: for example, given a snapshot of any simulation and dividing the volume in two,

*jlutsko@ulb.ac.be; <http://www.lutsko.com>

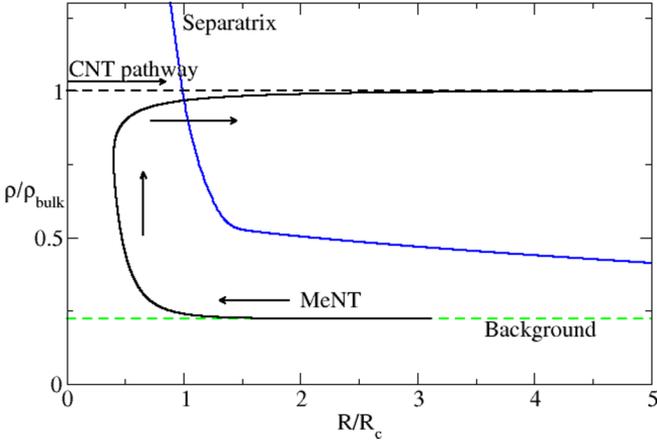


FIG. 1. Comparison of nucleation pathways predicted by MeNT (full black line) and CNT (broken black line) for the formation of a dense (concentrated) phase from a weak solution. The evolving nucleus is characterized by its interior concentration ρ (expressed as a ratio to the bulk density of the condensed phase) and by its radius R (expressed relative to the critical radius). The broken horizontal line at low concentration is the background, or initial, concentration of the weak solution. In CNT, the interior density is equal to that of the bulk and is constant and only the radius changes whereas in MeNT both variables change continuously and the path begins with a small density excess over a substantial area. Also shown are the separatrix which must be crossed to achieve nucleation: the critical cluster for each pathway is the intersection between the pathway and the separatrix.

one will always find the density of one half slightly above the system average and the other slightly below. The theory predicts that nucleation of a condensed phase is more likely to occur in the former than in the latter which makes intuitive sense. This result has proven very robust being found both for nucleation of liquid droplets from a vapor and for that of crystals from a weak solution and by means of simple, coarse-grained theories derived from MeNT [15] as well as in large-scale calculations based on the full theory [13]. It also persists for both the use of simple squared-gradient free-energy models as well as for sophisticated density functional theories. Ultimately, it is a result of the dynamics correctly conserving mass. In this work, we describe a validation of this prediction based on dynamical simulations.

In this work, the goal is to verify this element of the nucleation pathway via simulation. This goal is difficult as one cannot use the usual definitions of a “cluster.” For example, if one defines a cluster to be a collection of molecules that are each within some fixed distance of another member of the cluster, then one is biasing the definition to regions with a local density exceeding some threshold. This automatically rules out observation of a cluster consisting of a slight increase in density above the background. For this reason, we devote the next section to a detailed description of our simulation strategy. We describe a use of forward flux sampling to determine the nucleation pathway for an open system which allows us to identify a local increase in density. We present our simulation results in the following section where it is shown that nucleation is most likely to begin with a long-

wavelength, small-amplitude density fluctuation. We end with a short summary of our conclusions.

II. SIMULATION STRATEGY

Our strategy is inspired by recent calculations showing that the long-wavelength fluctuations manifest themselves as an overall increase in density in open, finite systems prior to cluster formation [13]. In order to simulate three-dimensional systems which are both open and yet with a mass-conserving dynamics, we use periodic boundaries in two directions and allow molecules to move out of the system in the natural course of the dynamics, with this being balanced by molecules randomly entering through the same boundaries. The number of molecules within the simulation cell therefore fluctuates while in the interior of the cell, mass is strictly, locally conserved. This model corresponds to a subvolume of a large system.

We have performed kinetic Monte Carlo simulations for molecules on a lattice with lattice constant a and with these semiopen boundary conditions. Our simulations include a single species of molecules that bonds to any nearest neighbors in the six Cartesian directions with an energy $-\epsilon$ for each bond. (We scale all the temperature and all energies to this quantity so its physical value is not needed.) Molecules jump from one lattice site to an unoccupied nearest-neighbor lattice site with a rate proportional to $\min[1, \exp(-\beta \Delta E)]$ where ΔE is the total change in the system energy between the final and initial states, $\beta = 1/k_B T$, k_B is Boltzmann’s constant, and T is the temperature. In the low-density limit, this reduces to a three-dimensional random walk and hence, macroscopically, to diffusion. Note that only monomers move; there is no provision for the collective movement of clusters. This is important because it means that only monomers can attach to clusters and the concentration of monomers is generally less than the total number of molecules in the cell.

The control parameters in the simulation are the temperature and the number of molecules entering per time step via the open surfaces which, in turn, controls the average number of molecules in the simulation cell and the number of monomers. In equilibrium, with an average density of monomers n_1 , the number of monomers exiting through the boundaries in each time step will be $2n_1 A a v_0 \delta t$ where A is the area of the open boundaries and v_0 is the “attempt frequency.” (The latter is the fundamental timescale of the simulations: all times are expressed in terms of it and so its actual value is arbitrary.) If the number of monomers entering per time step, δt , is $2N_B$, then the concentration in equilibrium is given by $n_1 A a v_0 \delta t = N_B$. Balancing the rates of attachment and detachment at a kink site similarly show that at two-phase coexistence the concentration of monomers is $n_1^{\text{coex}} = e^{-3\beta\epsilon} a^{-3}$ [16] so that $N_B^{\text{coex}} = e^{-3\beta\epsilon} A a^{-2} v_0 \delta t$. When presenting results below, we use $S \equiv \ln(N_B/N_B^{\text{coex}})$ as a measure of the supersaturation. The simulation scheme is illustrated in Fig. 2.

For the simple lattice-based kinetic Monte Carlo (kMC) dynamics there are only two phases: a low-density vapor and a crystalline state. Crystalline clusters are identified via the nearest-neighbor bonds giving the cluster size as the natural order parameter; specifically, for any configuration of positions, the order parameter is taken to be the size of the largest

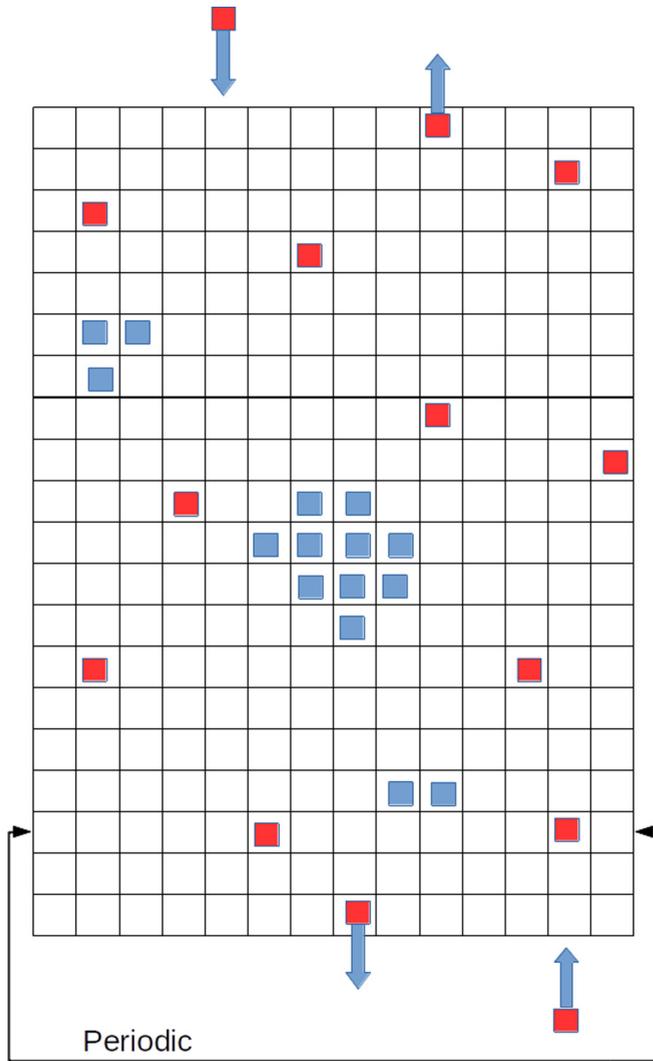


FIG. 2. Schematic illustration of the simulation scheme projected onto two dimensions. The simulations take place on a grid with molecules randomly hopping between nearest-neighbor sites. The boundaries are periodic in two dimensions and open in the third: molecules that jump out of the simulation via the top and bottom boundaries are lost to the simulation but balanced by molecules randomly entering at rates determined by the applied chemical potential. Clusters of molecules, i.e., those connected by nearest-neighbor bonds, are shown as blue.

cluster present in the system. For the conditions considered here, a system beginning in the vapor state will seldom spontaneously form clusters larger than a few molecules and those are, in general, short-lived. Thus we use a standard rare-event technique, forward flux sampling [17,18] (FFS), to drive the system in an unbiased manner from the initial, low-density state to crystallization. The key concept in FFS is a set of surfaces in phase space defined as being all configurations with a given value of the order parameter. In our implementation of FFS for this system, we begin with the vapor and allow the system to evolve. Each time a configuration having largest cluster size $N_0 - 1$ evolves into one having largest cluster size N_0 in the next time time step, we store the latter configuration thus creating a database, D_{N_0} , of configurations of systems

crossing the N_0 surface in the direction of increasing cluster size. Once this database has a given number of configurations, the second part of the simulation begins. This consists of randomly choosing one of the configurations from D_{N_0} and evolving the system forward in time. If it eventually returns to the initial metastable basin, i.e., the vapor state having order parameter less than a specified value N_{equil} , then the simulation is discarded and the process is repeated. If, before returning to the initial basin, the system evolves into one with order parameter $N_0 + 1$, the simulation is stopped, the configuration recorded in a new database, D_{N_0+1} , and the process is repeated until D_{N_0+1} has the required number of configurations. Once complete, this continues for order parameter $N_0 + 2$, etc., until some maximal value N_{max} is reached. In the initial stages, most trajectories fail to reach the next surface but as the cluster sizes increase, the success rate does as well until almost all trajectories successfully reach the next surface. Thus, the configurations in $D_{N_{\text{max}}}$ are effectively stable crystals that will only grow with time.

From the empirically determined probabilities of trajectories launched from one surface to reach the next, and from the empirically determined rate at which trajectories leave the initial basin, the transition rate can be determined. Here, however, the main interest is not the rate but the paths. For any given element of D_N one can trace back the parent configuration in D_{N-1} and the grandparent in D_{N-2} , etc., and in this way the entire trajectory going back to an initial configuration in D_{N_0} . If we denote by $\Gamma_N^{(j)}$ the j th configuration in database D_N , then a successful trajectory leading to a stable crystal will be a sequence of snapshots $(\Gamma_{N_0}^{(j_0)}, \Gamma_{N_0+1}^{(j_1)}, \dots, \Gamma_{N_{\text{max}}}^{(j_{N_{\text{max}}-N_0})})$. Any configuration $\Gamma_N^{(j)}$ will be a member of some number, $n^{(S)}(\Gamma_N^{(j)})$, of such successful trajectories. If we also record for each configuration the total number of times it was used as an initial value, $n(\Gamma_N^{(j)})$, then we also know the total number of unsuccessful trajectories launched from this configuration, $n^{(U)}(\Gamma_{N_0}^{(j)}) = n(\Gamma_N^{(j)}) - n^{(S)}(\Gamma_N^{(j)})$. This information is used to compute three types of average of the mass. Let $M(\Gamma_N^{(j)})$ be the mass, or number of molecules, in configuration $\Gamma_N^{(j)}$. Then,

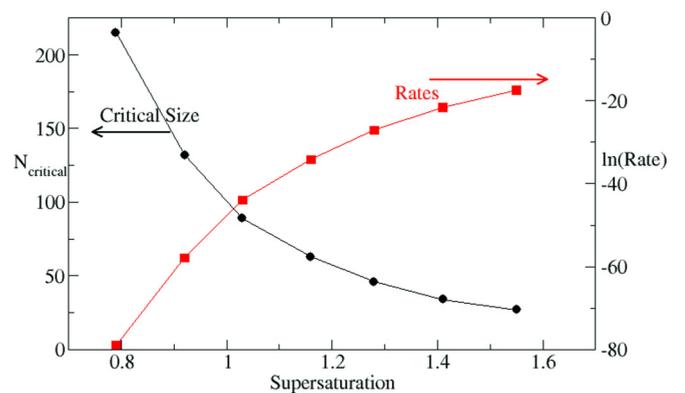


FIG. 3. The number of molecules in the critical cluster (circles, left-hand scale) and the log of the measured nucleation rates (squares, right-hand scale) as functions of the supersaturation at $k_B T = 0.4\epsilon$. The nucleation rate is for the entire volume of the simulation cell and is expressed in terms of the fundamental KMC timescale defined by ν_0 .

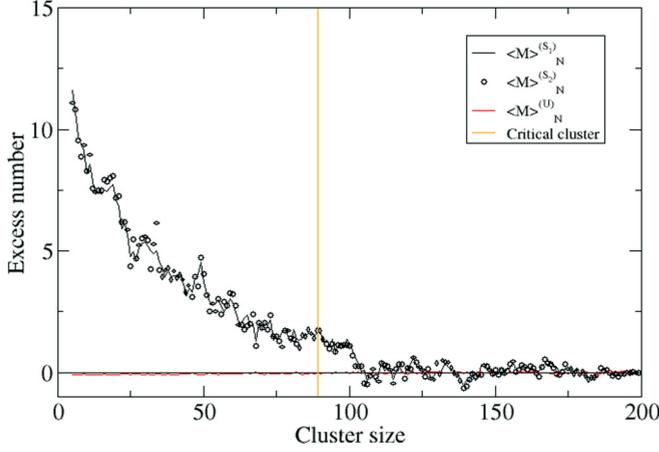


FIG. 4. Difference between the average number of molecules as computed by using the various weightings and the reference, unweighted average, as a function of cluster size. The vertical line indicates the critical cluster which contains approximately 90 molecules (supersaturation = 1.03), and the average number of monomers in the simulation cell in the initial metastable state was approximately 220.

first, we define the simple average mass of a database, that is, the average mass of all configurations possessing a cluster of given size N ,

$$\langle M \rangle_N \equiv \frac{\sum_{\Gamma_N^{(j)} \in D_N} M(\Gamma_N^{(j)})}{\sum_{\Gamma_N^{(j)} \in D_N} 1}. \quad (1)$$

Next is a *weighted-average* number of molecules in each database where the weight for each configuration $\Gamma_N^{(j)} \in D_N$

is the number of successful trajectories that include it,

$$\langle M \rangle_N^{(S_1)} \equiv \frac{\sum_{\Gamma_N^{(j)} \in D_N} n^{(S)}(\Gamma_N^{(j)}) M(\Gamma_N^{(j)})}{\sum_{\Gamma_N^{(j)} \in D_N} n^{(S)}(\Gamma_N^{(j)})}. \quad (2)$$

We have also defined a variant called $\langle M \rangle_N^{(S_2)}$ by replacing $n^S(\Gamma_N^{(j)})$ as the weight by $\frac{n^S(\Gamma_N^{(j)})}{n(\Gamma_N^{(j)})}$ [for the cases that $n(\Gamma_N^{(j)}) > 0$] to thus correct for the possibility that a configuration is chosen many times and only yields a small fraction of successful trajectories. Finally, we define the average mass of completely unsuccessful configurations, that is, those configurations that were used as initial values but did not lead to any successful trajectories:

$$\langle M \rangle_N^{(U)} \equiv \frac{\sum_{\substack{\Gamma_N^{(j)} \in D_N \\ n(\Gamma_N^{(j)}) > 0}} \delta_{n^{(S)}(\Gamma_N^{(j)})=0} M(\Gamma_N^{(j)})}{\sum_{\substack{\Gamma_N^{(j)} \in D_N \\ n(\Gamma_N^{(j)}) > 0}} \delta_{n^{(S)}(\Gamma_N^{(j)})=0}}. \quad (3)$$

III. RESULTS

We have performed simulations on a system with a simulation cell of size $50 \times 50 \times 50$, and a temperature of $k_B T = 0.4\epsilon$. The FFS calculations were performed using approximately 160 000 configurations in each database. (This number was chosen simply empirically as one in which the noise in the final results was sufficiently small to reveal the trends.) At these conditions, the equilibrium system has a concentration of approximately $e^{-3/k_B T} = 5.5 \times 10^{-4}$ monomers per unit cell or 69 monomers on average in the whole simulation cell. Our nonequilibrium simulations spanned the range of supersaturations from $S = 0.79$ (about 170 molecules) to $S = 1.55$

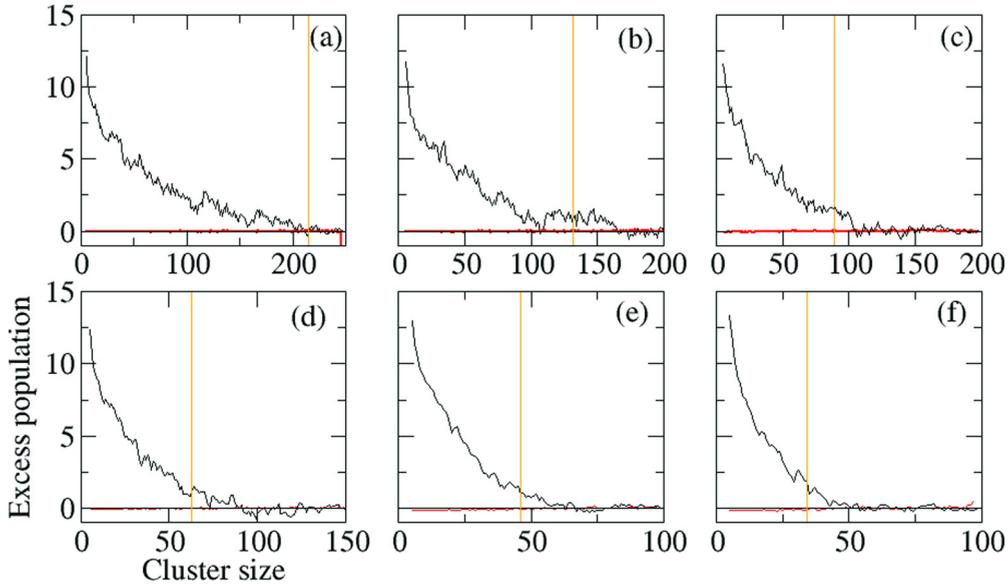


FIG. 5. Difference between the average number of molecules in the simulation cell calculated by weighting each configuration by the number of successful trajectories it participates in and the average overall of all configurations as a function of the size of the clusters. The different panels are for different supersaturations and in each case the critical cluster is indicated by a vertical line. The temperature was $k_B T = 0.4\epsilon$ and the simulation cell was a cube with 50 sites per side. Defining the supersaturation as $S = \ln \langle N \rangle / \langle N \rangle_{\text{coex}}$ where $\langle N \rangle$ is the average number of molecules in the initial state and $\langle N \rangle_{\text{coex}}$ is that for coexistence, panels (a)–(f) correspond to supersaturations $S = 0.79, 0.92, 1.03, 1.16, 1.28, \text{ and } 1.41$, respectively.

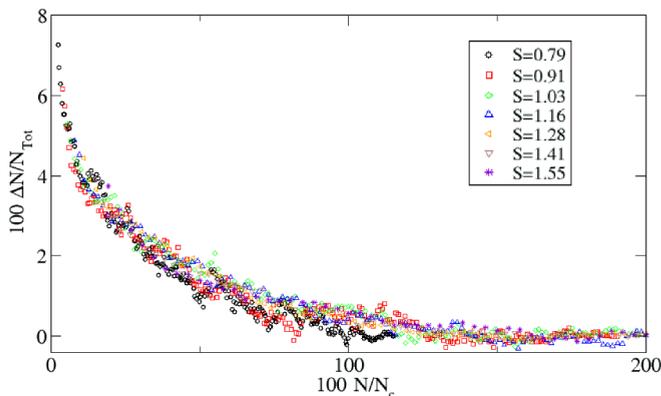


FIG. 6. The data from Fig. 5 but with the number of molecules in the clusters scaled by the critical size and the excess number of molecules scaled by the total number in the simulation cell in equilibrium.

(about 394 molecules). To give some context to our analysis of the nucleation pathways, we show in Fig. 3 the critical clusters and nucleation rates as functions of the supersaturation. The critical cluster contains about 215 molecules at the lowest supersaturation down to only 27 at the highest concentrations. Note that the critical clusters are easily determined from the FFS simulations since they directly provide the probability of a cluster of size N growing to size $N + 1$ and that one of size $N + 1$ grows to size $N + 2$, etc.: multiplying these gives the probability that a cluster of size N grows indefinitely and the cluster for which this probability is 50% is the critical cluster.

Figure 4 shows the difference between the weighted-average number of molecules and the unweighted average for the three types of weighted average as functions of the cluster size for the case $S = 1.03$ as determined from the FFS simulations. It is clear that both the S_1 and S_2 weightings give virtually identical results and that both show a systematic excess relative to the unweighted average during the nucleation phase. The average over unsuccessful configurations, however, shows only a very slight decrease relative to the reference. These observations were the same in all other simulations reported below.

Figure 5 shows the excess number of molecules ΔN in the successful trajectories relative to the average over all configurations at each value of N for different supersaturations corresponding to critical clusters ranging from 35 to 220 molecules. It is clear that there is a systematic trend towards excess numbers of molecules in the successful trajectories thus confirming the theoretical prediction. For the smallest clusters, the excess amounts to approximately 5% of the average number of molecules in equilibrium.

We also show in Fig. 6 the same data with ΔN scaled to the average number of molecules in the simulation cell in equilibrium and with N scaled to the critical radius. The data collapse to a single curve, thus further supporting the systematic nature of the observed excess in the successful trajectories.

Our interpretation of these results is that in the early stages of nucleation, the total number of molecules in the simulations varies stochastically and that configurations with a higher than average density are more likely to participate in successful nucleation events. This is in agreement with the predictions of MeNT and serves to support its more complex description of nucleation, compared to CNT.

IV. CONCLUSIONS

We have used simple kMC simulations to confirm a robust prediction of MeNT, namely, that diffusion-limited nucleation is most likely to begin with a long-wavelength, small-amplitude density fluctuation. This supports the nuanced picture of the process that arises when coupling mass transport, mass conservation, and thermodynamics. The results do not themselves invalidate CNT—at least at low supersaturations, the difference between the nucleation pathways predicted by CNT and observed in the simulations is only significant in the early stages of nucleation—but they do serve to show that even in relatively simple cases, the intuitively appealing classical picture has limitations. One practical consequence is that studies that rely on local density to define clusters may miss the complexity of the early stages of nucleation completely. Indeed, referring back to Fig. 1, note that if one imposes a minimum density so that anything below that is not identified as a “cluster,” then observed clusters will have a minimum size that is larger than zero. This is consistent with previous work of Trudu *et al.* [19] who reported such an observation in simulations of crystal nucleation, namely, that precritical crystallites occur only above some finite size. It also suggests that the manipulation of density fluctuations, e.g., as reported in the GradFlex microgravity experiment [20], could be useful in controlling nucleation.

Finally, we speculate that while our present results are applicable to a particular form of nucleation, namely, that of a dense phase from a dilute phase, analogous phenomena could well occur in other systems. For example, density fluctuations can be both positive and negative: one might expect bubbles to preferentially nucleate in a liquid in regions of lower-density fluctuations. When liquids freeze into solids or, vice versa, solids melt to form liquids, the change in density is often relatively small but in these cases, heat transport is important due to the latent heat involved in the phase transitions and so long-wavelength, small-amplitude temperature fluctuations could be the analog to the precursors studied here, i.e., the “hot spots” and “cold spots” alluded to in the Introduction.

ACKNOWLEDGMENTS

J.F.L. was supported by the European Space Agency (ESA) and the Belgian Federal Science Policy Office (BELSPO) in the framework of the PRODEX Programme, Contract No. ESA AO-2004-070. J.L. acknowledges financial support from the Fonds de la Recherche Scientifique–FNRS. Computational resources have been provided by the Shared ICT Services Centre, Université Libre de Bruxelles.

- [1] A. Ablasser, Phase separation focuses DNA sensing, *Science* **361**, 646 (2018).
- [2] P.-E. Tremblay, G. Fontaine, N. P. G. Fusillo, B. H. Dunlap, B. T. Gänsicke, M. A. Hollands, J. J. Hermes, T. R. Marsh, E. Cukanovaite, and T. Cunningham, Core crystallization and pile-up in the cooling sequence of evolving white dwarfs, *Nature (London)* **565**, 202 (2019).
- [3] A. Giacomello, S. Meloni, M. Müller, and C. M. Casciola, Mechanism of the Cassie-Wenzel transition via the atomistic and continuum string methods, *J. Chem. Phys.* **142**, 104701 (2015).
- [4] D. Kashchiev, *Nucleation: Basic Theory with Applications* (Butterworth-Heinemann, Oxford, 2000).
- [5] V. K. Shen and P. G. Debenedetti, A computational study of homogeneous liquid-vapor nucleation in the Lennard-Jones fluid, *J. Chem. Phys.* **111**, 3581 (1999).
- [6] Z.-J. Wang, C. Valeriani, and D. Frenkel, Homogeneous bubble nucleation driven by local hot spots: A molecular dynamics study, *J. Phys. Chem. B* **113**, 3776 (2009).
- [7] K. C. Russell, Linked flux analysis of nucleation in condensed phases, *Acta Metall.* **16**, 761 (1968).
- [8] J. S. Langer, Kinetics of metastable states, in *Systems Far from Equilibrium*, edited by Luis Garrido (Springer, Berlin/Heidelberg, 1980), pp. 12–47.
- [9] J. S. Langer, Metastable states, *Physica* **73**, 61 (1974).
- [10] B. Peters, On the coupling between slow diffusion transport and barrier crossing in nucleation, *J. Chem. Phys.* **135**, 044107 (2011).
- [11] J. F. Lutsko, A dynamical theory of homogeneous nucleation for colloids and macromolecules, *J. Chem. Phys.* **135**, 161101 (2011).
- [12] J. F. Lutsko, A dynamical theory of nucleation for colloids and macromolecules, *J. Chem. Phys.* **136**, 034509 (2012).
- [13] J. F. Lutsko, How crystals form: A theory of nucleation pathways, *Sci. Adv.* **5**, eaav7399 (2019).
- [14] J. F. Lutsko, Systematically extending classical nucleation theory, *New J. Phys.* **20**, 103015 (2018).
- [15] J. F. Lutsko and M. A. Durán-Olivencia, A two-parameter extension of classical nucleation theory, *J. Phys.: Condens. Matter* **27**, 235101 (2015).
- [16] D. Maes and J. F. Lutsko, Molecular viewpoint on the crystal growth dynamics driven by solution flow, *Cryst. Growth Des.* **20**, 2294 (2020).
- [17] R. J. Allen, D. Frenkel, and P. R. ten Wolde, Simulating rare events in equilibrium or nonequilibrium stochastic systems, *J. Chem. Phys.* **124**, 024102 (2006).
- [18] R. J. Allen, C. Valeriani, and P. R. ten Wolde, Forward flux sampling for rare event simulations, *J. Phys.: Condens. Matter* **21**, 463102 (2009).
- [19] F. Trudu, D. Donadio, and M. Parrinello, Freezing of a Lennard-Jones Fluid: From Nucleation to Spinodal Regime, *Phys. Rev. Lett.* **97**, 105701 (2006).
- [20] A. Vailati, R. Cerbino, S. Mazzoni, C. J. Takacs, D. S. Cannell, and M. Giglio, Fractal fronts of diffusion in microgravity, *Nat. Commun.* **2**, 290 (2011).