A dynamical theory of nucleation for colloids and macromolecules

James F. Lutsko

Related Articles

Evaporation rate of nucleating clusters
J. Chem. Phys. 135, 194504 (2011)

Reaction coordinates for the crystal nucleation of colloidal suspensions extracted from the reweighted path ensemble
J. Chem. Phys. 135, 154110 (2011)

Doubled heterogeneous crystal nucleation in sediments of hard sphere binary-mass mixtures
J. Chem. Phys. 135, 134115 (2011)

On the coupling between slow diffusion transport and barrier crossing in nucleation
J. Chem. Phys. 135, 044107 (2011)

A multigrid method for N-component nucleation
J. Chem. Phys. 135, 014114 (2011)

Additional information on J. Chem. Phys.

A dynamical theory of nucleation for colloids and macromolecules

James F. Lutsko[a)]
*Center for Nonlinear Phenomena and Complex Systems, Code Postal 231, Université Libre de Bruxelles,
Blvd. du Triomphe, 1050 Brussels, Belgium*

A general theory of nucleation for colloids and macromolecules in solution is formulated within the context of fluctuating hydrodynamics. A formalism for the determination of nucleation pathways is developed and stochastic differential equations for the evolution of order parameters are given. The conditions under which the elements of classical nucleation theory are recovered are determined. The theory provides a justification and extension of more heuristic equilibrium approaches based solely on the free energy. It is illustrated by application to the low-concentration/high-concentration transition in globular proteins, where a novel two-step mechanism is identified, where the first step involves the formation of long-wavelength density fluctuations, and the second step is the actual nucleation event occurring within the fluctuation. © *2012 American Institute of Physics.* [doi:10.1063/1.3677191]

I. INTRODUCTION

Nucleation is a fundamental physical process of importance in fields as diverse as chemistry, materials science, biology, and cosmology. Our basic understanding of it goes back to Gibbs' discussion of the nucleation of a liquid droplet in a metastable vapor, see, e.g., discussion in Ref. 1. The physics is governed by the fact that the excess free energy of a cluster of new phase relative to the mother phase has a negative contribution that scales as the volume and a positive contribution that scales as the surface area. The former is due to the fact that the bulk new phase has lower free energy than the bulk mother phase, while the latter arises from the free-energy cost of the interfacial region. When the cluster is small, the surface term dominates and the excess free energy of the cluster is positive making the cluster thermodynamically disfavored. When the cluster is large, the volume term dominates, the cluster has lower free energy than the mother phase and so cluster growth is favored. Separating the two regimes is the metastable critical cluster. In classical nucleation theory (CNT), it is assumed that the cluster is spherical, that its interior is in the bulk new-phase state and that the surface tension is the same as for a planar interface of the coexisting new and mother phases so that the free energy of the cluster as a function of its radius can be calculated giving a quantitative picture of homogeneous nucleation.[1]

The dynamics of nucleation – the nucleation rate, the determination of the nucleation pathway, and the nature of the post-critical growth of clusters – is a more involved subject. In CNT, the starting point for understanding nucleation rates is the monomer attachment/detachment picture that leads to the Becker-Döring equations.[1,2] These are simply rate equations for number of clusters of size N based on the processes of monomer attachment and detachment. Following Frenkel, Zeldovich, and co-workers,[1] these equations can be converted into a Fokker-Planck equation for the evolution of the distribution of clusters. All of these formalisms can be used to estimate nucleation rates. The physics enters via the monomer attachment and monomer detachment rates which in general depend on the thermodynamic driving force (the derivative of the CNT free energy with respect to cluster size) and the dynamics of the surrounding mother phase. In many cases, adequate approximations are available that express the nucleation rate in terms of $e^{-\beta \Delta \Omega_c}$ where $\beta = 1/k_B T$, k_B is Boltzmann's constant, T is the temperature, and $\Delta \Omega_c$ is the excess free energy of the critical cluster.

Although physically appealing, this formalism suffers from a number of deficiencies. First and foremost is the well-known fact that the nucleation rates estimated from CNT are often many of orders of magnitude too large or too small[3] (see also Fig. 13.8 of Ref. 1). This is largely attributed to a poor estimate of the free-energy barrier: more sophisticated methods allow for better estimates of the barriers. A common way to get improved estimates of the barriers is via classical density functional theory (DFT; Refs. 4, 5, and 7) (for early examples of this approach see Refs. 8 and 9). However, the use of DFT raises new conceptual questions. The monomer attachment/detachment picture requires estimates of the rates for all cluster sizes which in turn means estimating "free energies" for all clusters. But, DFT is an equilibrium theory that can only describe – at best – stationary states on the free-energy surface. This fact follows trivially since the procedure to determine free energies via DFT begins with the extremization of the free energy to get the structure of the cluster and naturally must result in a minimum, maximum, or saddle point but it is also connected to the theorems that underlay DFT, which only attach significance to minima of the free-energy functional.[5,10] This, therefore, raises the question as to whether DFT can be used to determine the nucleation pathway, i.e., the sequence of necessarily nonequilibrium states leading from the pure unstable phase to the critical cluster.

a)Electronic mail: jlutsko@ulb.ac.be; URL: http://www.lutsko.com.

Putting aside the conceptual questions, there are currently two primary approaches to determining nucleation pathways using DFT. The first, and oldest, is to extremize the free energy under some sort of constraint that stabilizes non-critical clusters: in effect, imposing a stabilizing external field.[8,9,11] The problem with this approach is that there are numerous possible ways to implement it. To take a simple example, one could minimize the free energy at fixed radius so that the pathway would be mapped out by varying the radius which acts as the reaction coordinate (or "order parameter"). However, one could equally well minimize at fixed number of molecules in the cluster, N, so that this becomes the order parameter. These do not necessarily give the same result since, e.g., the number of molecules can increase with the radius actually decreasing provided the interfacial width or interior density of the cluster increases enough to compensate; indeed, this actually occurs in some calculations.[12] Notwithstanding this problem, the constraint approach continues to be widely used.[13–16] A more recent alternative is to search for steepest-descent paths on the free-energy surface.[11,17–20] However, here there is also an ambiguity as steepest-descent requires a measure of distance in parameter space and there is *a priori* no unique prescription for this measure.[17]

It is reasonable to suppose that all of the ambiguity underlying the DFT approaches is due to the fact that DFT is an equilibrium theory, whereas nucleation is fundamentally a nonequilibrium, fluctuation-driven process. Viewed in this way, one can question other aspects of CNT as well, such as whether it is really necessary that the system even passes through the critical cluster since the free energy may not be the only factor controlling the nonequilibrium evolution. Similarly, one can also ask why the pre-critical process does not seem to depend on fluid-kinetics, while these are crucial in driving post-critical growth. (This question has also been addressed recently by Peters[6] using CNT free-energy models). All of this suggests that one would like a unified description with the following properties: it is inherently nonequilibrium in nature, it is dynamical; it allows for the determination of the nucleation pathway from the dynamics, it allows for the determination of nucleation rates and it gives a description of post-critical growth. Of course, one would expect that it involves free energies in a natural way without assuming equilibrium and that under some limit the CNT would result thus also helping to clarify its applicability. Additionally, one would hope that such a formalism would be sufficiently flexible to be applied to more complex problems involving multiple order parameters and to determine the relative probability of completely different nucleation pathways (e.g., whether crystals form directly from solution or whether they form in two steps, passing first through a dense amorphous phase.[21])

The aim of this paper is to illustrate an approach that can satisfy these requirements. Attention will primarily be focused on nucleation from a vapor or liquid phase, but the basic ideas could be extended to other systems in obvious ways. The development is based on Brownian dynamics where molecules move according to Newton's laws while being subject to a frictional force as well as fluctuating forces. This is a simple model for colloids and the important case of macromolecules in solution in which cases the friction and

the fluctuations come from the bath/solvent and has the virtue of allowing for useful simplifications. As described in Sec. II, the fundamental level of description adopted for the description of nucleation is fluctuating hydrodynamics, in which the free energy enters the dynamics via a local equilibrium expression for the pressure, resulting in a set of stochastic differential equations (SDEs). The most likely path (MLP) is introduced as a means of characterizing the nucleation pathway via a *deterministic* equation derived from the stochastic model. In Sec. III, a reduced description in terms of order parameters is developed. It is shown that in the weak-noise limit, the MLP is simply gradient descent on the free-energy surface with a metric uniquely determined by the dynamics and that it necessarily passes through the critical cluster thus making contact with the CNT picture. Finally, in Sec. IV three applications are discussed: nucleation of high-concentration droplets from a low-concentration solution of globular proteins; the reverse transition of high-concentration solution to low-concentration bubbles in an under-saturated solution; and finally the specialization to a single order parameter so as to recover well-known results from CNT. For the first application – droplet nucleation – calculations lead to an unexpected but physically appealing result, wherein it is found that the coupling of fluctuations, fluid dynamics, and thermodynamics implies that nucleation is most likely to occur within regions of enhanced density due to long-wavelength density fluctuations. Section V summarizes the results and discusses possibilities for further developments. A brief summary of some of these results has appeared previously.[22]

II. THE MODEL

A. Dynamics

The system consists of a collection of particles with positions and momenta $\mathbf{q}_i$, $\mathbf{p}_i$ interacting via a potential V. Additionally, the particles interact with a bath/solvent of small particles and this is described via a frictional drag and a fluctuating force, both characterized by the constant γ,

$$\dot{\mathbf{q}}_i = \mathbf{p}_i/m, \quad \dot{\mathbf{p}}_i = -\frac{\partial V}{\partial \mathbf{q}_i} - \gamma \mathbf{p}_i + \varepsilon \sqrt{2\gamma m k_B T} \mathbf{f}_i(t), \quad (1)$$

where all the components of the force are independent,

$$\langle \mathbf{f}_i(t)\mathbf{f}_j(t') \rangle = \mathbf{1}\delta_{ij}\delta(t - t'). \quad (2)$$

The coefficient of the noise term is determined by demanding that the system admits of an equilibrium Boltzmann distribution thus forcing a fluctuation-dissipation relation. Here and in the following, an additional parameter, ε, is included to allow for the discussion of a weak-noise limit. The physicality of such a limit will be discussed below.

Defining the local density and momentum density, respectively, as

$$\hat{\rho}(\mathbf{r};t) = \sum_i \delta(\mathbf{r} - \mathbf{q}_i),$$

$$\hat{\mathbf{j}}(\mathbf{r};t) = \sum_i \mathbf{p}_i \delta(\mathbf{r} - \mathbf{q}_i), \quad (3)$$

one sees that these satisfy the exact equations

$$\frac{\partial \widehat{\rho}(\mathbf{r};t)}{\partial t} = -\frac{1}{m}\nabla \cdot \widehat{\mathbf{j}}(\mathbf{r};t),$$

$$\frac{\partial \widehat{\mathbf{j}}(\mathbf{r};t)}{\partial t} = -\nabla \cdot \sum_i \frac{\mathbf{p}_i \mathbf{p}_i}{m} \delta(\mathbf{r} - \mathbf{q}_i) - \sum_i \frac{\partial V}{\partial \mathbf{q}_i}\delta(\mathbf{r} - \mathbf{q}_i)$$
$$- \gamma \widehat{\mathbf{j}}(\mathbf{r};t) + \varepsilon\sqrt{2\gamma m k_B T \widehat{\rho}(\mathbf{r};t)}\mathbf{F}(\mathbf{r};t) \qquad (4)$$

with

$$\langle \mathbf{F}(\mathbf{r};t)\mathbf{F}(\mathbf{r}';t')\rangle = \delta(t - t')\delta(\mathbf{r} - \mathbf{r}')\mathbf{1}. \qquad (5)$$

Note that the actual form of the noise term that occurs in the momentum equation is

$$\varepsilon\sqrt{2\gamma m k_B T}\sum_i \delta(\mathbf{r} - \mathbf{q}_i)\mathbf{f}_i(t). \qquad (6)$$

However, it is well known that a noise term of the form $B_{ij}\xi_j(t)$ (with summation over repeated indices) can be replaced by another of the form $C_{ij}\xi_j(t)$ without changing the statistical properties of the dependent variables provided that coefficients are related by $B_{il}B_{jl} = C_{il}C_{jl}$, i.e., the autocorrelation of the noise is unchanged.[23, 24]

There are two sources of stochasticity in this problem. The obvious source is the fluctuating force that represents the bath or solvent. The second arises because one is not interested in following a single trajectory of the system but, as usual, want to consider an average over an ensemble of systems. In equilibrium, the ensemble is drawn from a known distribution (the canonical distribution). The distribution for systems out of equilibrium is generally time-dependent and one must specify the ensemble for a particular point in time, e.g., a distribution of initial conditions. In the present case, because of the two sources of noise, one must carry out this procedure in two steps: first, the set of values of the fluctuating forces acting on the colloids at each instant in time is specified. Then, this fixed set of values for the noise simply acts as an external field coupled to the colloid dynamics and nonequilibrium statistical mechanics can be developed in the usual way as an average over an ensemble of initial conditions of the particles. Then, having developed whatever level of description is of interest, a second average is performed over the noise. As a result, one can go from the exact balance equations, Eq. (4), to a description of ensemble-averaged quantities.

To fix the notation, let $\Gamma(t) = \{\mathbf{q}_i, \mathbf{p}_i\}_{i=1}^N$ represent the collection of positions and momenta of all N large particles at time t and to be more explicit one can write $\Gamma(t; \Gamma(0))$, since the phase variables at time t are completely determined by their values at some earlier time. Similarly, one can write for the density, for example, $\widehat{\rho}(\mathbf{r};t) \equiv \widehat{\rho}(\mathbf{r}; \Gamma(t; \Gamma(0)))$. The probability for the phase variables to have particular initial values is denoted $f(\Gamma(0))$ and the ensemble-averaged density is then

$$\rho(\mathbf{r}, t) = \int \widehat{\rho}(\mathbf{r}; \Gamma(t; \Gamma(0)))\, f(\Gamma(0))\, d\Gamma(0),$$

$$= \int \widehat{\rho}(\mathbf{r}; \Gamma)\, f(\Gamma; t)\, d\Gamma, \qquad (7)$$

where in the second line the time-dependence has been moved to the distribution function.[25] Notice that the ensemble-averaged density and momentum density are all conditional on the fluctuating forces which for now just act as external forces. At this point, the usual methods of nonequilibrium statistical mechanics can be used to develop a hydrodynamic description – either a mesoscopic fluctuating hydrodynamics or a macroscopic, deterministic hydrodynamics, the former having the form

$$\frac{\partial \rho(\mathbf{r})}{\partial t} + \frac{1}{m}\nabla \cdot j(\mathbf{r}) = 0,$$

$$\frac{\partial \mathbf{j}(\mathbf{r})}{\partial t} + \nabla \cdot \mathbf{j}(\mathbf{r})\mathbf{j}(\mathbf{r})/\rho(\mathbf{r}) + \nabla p(\mathbf{r}, t) + \nabla \cdot \mathbf{\Pi}(\mathbf{r})$$
$$= -\gamma \mathbf{j}(\mathbf{r}) + \varepsilon\sqrt{2m\gamma k_B T\rho(\mathbf{r})}\xi(\mathbf{r};t), \qquad (8)$$

where $\rho(\mathbf{r})$ and $\mathbf{j}(\mathbf{r})$ are the ensemble-averaged local density and momentum density, and $\mathbf{\Pi}(\mathbf{r}, t)$ is the dissipative stress tensor which for the case of fluctuating hydrodynamics will have both a deterministic part and a fluctuating part. The other new term that appears here is $p(\mathbf{r}, t)$ which is the local pressure. Note that in the present case, the bath acts as a thermostat and so temperature fluctuations and heat transport are neglected, although in principle the model could be extended to include them.

The final approximation involves the pressure and can be motivated in different ways but always amounts to an assumption of local equilibrium. At its most heuristic, one starts with the Gibbs-Duhem equation for a single-component system, $Nd\mu = -SdT + Vdp$. For an isothermal system, this gives $dp = \rho d\mu$ where ρ is the density. Assuming local-equilibrium this is generalized to

$$dp(\mathbf{r}, t) = \rho(\mathbf{r}, t)d\mu(\mathbf{r}, t) = \rho(\mathbf{r}, t)d\frac{\delta F[\rho]}{\delta\rho(\mathbf{r})}, \qquad (9)$$

which is a widely used approximation (see, e.g., Sec. 8.4 of Ref. 1). Substituting into the equations for the density and momentum gives

$$\frac{\partial \rho(\mathbf{r})}{\partial t} + \frac{1}{m}\nabla \cdot \mathbf{j}(\mathbf{r}) = 0,$$

$$\frac{\partial \mathbf{j}(\mathbf{r})}{\partial t} + \nabla \cdot \mathbf{j}(\mathbf{r})\mathbf{j}(\mathbf{r})/\rho(\mathbf{r}) + \rho(\mathbf{r})\nabla\frac{\delta F[\rho]}{\delta\rho(\mathbf{r})} + \nabla \cdot \mathbf{\Pi}(\mathbf{r})$$
$$= -\gamma \mathbf{j}(\mathbf{r}) + \varepsilon\sqrt{2m\gamma k_B T\rho(\mathbf{r})}\xi(\mathbf{r};t). \qquad (10)$$

An alternative way to motivate this is to evaluate the pressure under the assumption of a local equilibrium (i.e., Maxwellian) velocity distribution and using the local-equilibrium version of the first Yvon-Born-Green equation[26] to rearrange the resulting virial expression for the pressure. The same system of equations has been developed starting with the Boltzmann equation.[27] Chavanis[28] has discussed the derivation of these equations for coarse-grained quantities (i.e., averaged over small volumes and time intervals) as well as their derivation based on the general Landau-Lifshitz fluctuation formalism.[28] Chavanis also discusses the difference between this model and the Cahn-Hillard dynamics, the main point being that the latter does not include the density in the noise amplitude and the free-energy term. However it is found, the most important

assumption is clearly the introduction of the local equilibrium expression for the pressure since this is what brings the free energy into the picture. This assumption of local equilibrium is, on the one hand, heuristic and represents an uncontrolled approximation. On the other hand, it is widely used in nonequilibrium statistical mechanics (e.g., a similar assumption underlies the generalized Enskog equation[25]) and similar approximations are even used for dissipative granular systems. Thus, while it cannot be rigorously justified it does represent the "state of the art."

Having developed the general framework, it is now specialized to the case of strong dissipation due to the bath. This assumption greatly simplifies the analysis as it means that (a) the dissipative stresses can be neglected since they will act more slowly than the friction due to the bath and (b) the momentum current will always be small so that the quadratic convective term can be neglected. The result is

$$\frac{\partial \rho(\mathbf{r})}{\partial t} + \frac{1}{m}\nabla \cdot \mathbf{j}(\mathbf{r}) = 0,$$

$$\frac{\partial \mathbf{j}(\mathbf{r})}{\partial t} + \rho(\mathbf{r})\nabla\frac{\delta F[\rho]}{\delta\rho(\mathbf{r})} = -\gamma\mathbf{j}(\mathbf{r}) + \varepsilon\sqrt{2\gamma m k_B T \rho(\mathbf{r})}\xi(\mathbf{r};t).$$

$$(11)$$

The linearity allows us to eliminate the momentum giving

$$\frac{1}{\gamma}\frac{\partial^2 \rho(\mathbf{r})}{\partial t^2} + \frac{\partial \rho(\mathbf{r})}{\partial t} = \frac{k_B T}{\gamma m}\nabla \cdot \left(\rho(\mathbf{r})\nabla\frac{\delta\beta F[\rho]}{\delta\rho(\mathbf{r})}\right)$$
$$-\nabla \cdot \varepsilon\sqrt{\frac{2k_B T}{\gamma m}\rho(\mathbf{r})}\xi(\mathbf{r};t). \quad (12)$$

Scaling time by $1/\gamma$ shows that the second-order time derivative is of higher order in $1/\gamma$ than the other terms so that it can be neglected in the strong damping approximation leaving

$$\frac{\partial \rho(\mathbf{r})}{\partial t} = \frac{k_B T}{\gamma m}\nabla \cdot \left(\rho(\mathbf{r})\nabla\frac{\delta\beta F[\rho]}{\delta\rho(\mathbf{r})}\right)$$
$$-\nabla \cdot \varepsilon\sqrt{\frac{2k_B T}{\gamma m}\rho(\mathbf{r})}\xi(\mathbf{r};t). \quad (13)$$

(Note that this result can be obtained directly from the full hydrodynamic equations by introducing a scaling of the time and momentum current of $t = \gamma t^*$ and $\mathbf{j} = (k_B T/\gamma l)\bar{\rho}\mathbf{j}^*$, with $\bar{\rho}$ a typical density and l a molecular length scale, and noting that the dissipative stress tensor is of at least first-order in the momentum current.) Although the coefficient of the noise term is state-dependent, it is nevertheless the case that the Ito and Stratonovich interpretations of this stochastic differential equation are the same as shown in Appendix A. This model is well known in the literature as the Dean-Kawasaki model[24, 29] and it has been widely used, e.g., in the study of the glass transition. Chavanis[28] discusses various approaches to its interpretation. The derivation sketched here emphasizes the relation between this model and the more general theory of fluctuating hydrodynamics because the latter can be used, in conjunction with the approach described below, to address related problems to which Brownian dynamics is not applicable such as nucleation in simple fluids, multicomponent reacting fluids, etc. However, for the specific problem of

nucleation of over-damped colloids, it is worth noting that the Dean-Kawasaki model can be derived by more rigorous methods as discussed, e.g., by Kawasaki[29] and Chavanis.[30]

Without the noise term, this is often referred to as dynamical density functional theory. Another way of thinking about this is to say that the usual DDFT results from averaging this result over the noise. The question then arises as to whether the "free energy" entering here is the thermodynamic free energy or some other quantity (such as a coarse-grained free energy) that becomes the free energy after averaging over the noise. From the sketch given above, it would appear consistent to interpret the free-energy functional in Eq. (13) as being the free-energy functional for the one-component system of large molecules or colloids since it arises from a local-equilibrium approximation for the pressure which is in turn entirely due to interactions between the large-particles.

Note that at low density one has[4]

$$\beta F[\rho] \simeq \int (\rho(\mathbf{r})\ln\rho(\mathbf{r}) - \rho(\mathbf{r}))\,d\mathbf{r}, \quad (14)$$

so that the dynamical equation becomes

$$\frac{\partial \rho(\mathbf{r})}{\partial t} \simeq \frac{k_B T}{\gamma m}\nabla^2\rho(\mathbf{r}) - \nabla \cdot \varepsilon\sqrt{\frac{2k_B T}{\gamma m}\rho(\mathbf{r})}\xi(\mathbf{r};t), \quad (15)$$

which is simple diffusion with noise. One can, therefore, identify $D = (k_B T)/\gamma m$ as the diffusion constant for the large particles in the bath so that it, and therefore the coefficient γ, can be determined from the Einstein-Stokes equation based solely on the size of the particles and the properties of the bath.

B. Spherical symmetry

In principle, this model could be studied numerically. Here, however, the goal is to make contact with the standard approach to nucleation and one of the key assumptions usually made is that of spherical symmetry of the clusters. One cannot, however, simply assume a spherically symmetric density profile in Eq. (13) as the noise term breaks spherical symmetry. One solution would be to average Eq. (13) over a spherical shell but it proves more convenient and natural to integrate over a spherical volume.

The total (or cumulative) mass inside a spherical shell of radius r, $m(r)$, is simply the integral of the density

$$m(r) = \int_{r'<r} \rho(\mathbf{r}')d\mathbf{r}'. \quad (16)$$

Integrating the dynamical equation over a spherical shell gives an evolution equation for the cumulative mass

$$\frac{\partial m(r)}{\partial t} = D\int_{r'<r}\nabla'\cdot\left(\rho(\mathbf{r}')\nabla'\frac{\delta\beta F[\rho]}{\delta\rho(\mathbf{r}')}\right)d\mathbf{r}'$$
$$-\varepsilon\int_{r'<r}\nabla'\cdot\sqrt{2D\rho(\mathbf{r}')}\xi(\mathbf{r}';t)\,d\mathbf{r}',$$
$$= D\int\left(\rho(\mathbf{r})\nabla\frac{\delta\beta F[\rho]}{\delta\rho(\mathbf{r})}\right)\cdot\widehat{\mathbf{r}}d\widehat{\mathbf{r}}$$
$$-\varepsilon\int_{r'<r}\nabla\cdot\sqrt{2D\rho(\mathbf{r}')}\xi(\mathbf{r}';t)\,d\mathbf{r}', \quad (17)$$

where Gauss' theorem has been used to convert the first integral on the right to a surface integral. It is straightforward to show that the autocorrelation of the noise term is

$$\delta(r - r')\delta(t - t')2D\varepsilon^2 4\pi r^2$$
$$\times \left(\frac{1}{4\pi r^2}\int \delta(r'' - r)\rho(\mathbf{r}'')d\mathbf{r}''\right), \qquad (18)$$

where the last term is the density averaged over a spherical shell of radius r,

$$\langle\rho; R\rangle_S \equiv \frac{1}{4\pi R^2}\int \delta(r - R)\rho(\mathbf{r})d\mathbf{r}. \qquad (19)$$

Hence, the noise term can be replaced by $\varepsilon\sqrt{8\pi D r^2 \langle\rho; r\rangle_S}\xi(r; t)$, where $\xi(r; t)\xi(r'; t') = \delta(r - r')\delta(t - t')$. At this point, since there is no more coupling to the non-radial noise terms, one can consistently seek a solution in terms of spherically symmetric density distributions, $\rho(r)$, thus giving

$$\frac{\partial m(r)}{\partial t} = D4\pi r^2 \rho(r)\frac{\partial}{\partial r}\left.\frac{\delta\beta F[\rho]}{\delta\rho(\mathbf{r})}\right|_{\rho(r)}$$
$$- \varepsilon\sqrt{8\pi r^2 D\rho(r)}\xi(r; t). \qquad (20)$$

It is again easy to verify the equivalence of the Ito and Stratonovich interpretations of this equation (see Appendix A).

This result can be made somewhat more transparent by noting that for a spherically symmetric function $\rho(\mathbf{r}) \equiv \rho(r)$, one has

$$\rho(r) = \frac{1}{4\pi r^2}\int \delta(r - r')\rho(\mathbf{r}')d\mathbf{r}', \qquad (21)$$

so that for any functional F of $\rho(r)$,

$$\frac{\delta F[\rho]}{\delta\rho(\mathbf{r})} = \int_0^\infty \frac{\delta F[\rho]}{\delta\rho(r')}\frac{\delta\rho(r')}{\delta\rho(\mathbf{r})}dr' = \frac{\delta F[\rho]}{\delta\rho(r)}\frac{1}{4\pi r^2} \qquad (22)$$

giving

$$\frac{\partial m(r)}{\partial t} = D4\pi r^2 \rho(r)\frac{\partial}{\partial r}\left(\frac{1}{4\pi r^2}\frac{\delta\beta F[\rho]}{\delta\rho(r)}\right)$$
$$- \varepsilon\sqrt{8\pi r^2 D\rho(r)}\xi(r; t). \qquad (23)$$

Next, using the functional chain rule gives

$$\frac{\partial}{\partial r}\left(\frac{1}{4\pi r^2}\frac{\delta\beta F[\rho]}{\delta\rho(r)}\right) = \frac{\partial}{\partial r}\frac{1}{4\pi r^2}\int_0^\infty \frac{\delta m(r')}{\delta\rho(r)}\frac{\delta\beta F[\rho]}{\delta m(r')}dr',$$
$$= \frac{\partial}{\partial r}\frac{1}{4\pi r^2}\int_0^\infty \Theta(r' - r)$$
$$\times 4\pi r^2 \frac{\delta\beta F[\rho]}{\delta m(r')}dr',$$
$$= -\frac{\delta\beta F[\rho]}{\delta m(r)}, \qquad (24)$$

where $\Theta(x)$ is the Heaviside step function, so that the SDE can be written as

$$\frac{\partial m(r)}{\partial t} = -D\frac{\partial m(r)}{\partial r}\frac{\delta\beta F[\rho]}{\delta m(r)} - \varepsilon\sqrt{2D\frac{\partial m(r)}{\partial r}}\xi(r; t). \qquad (25)$$

This shows very clearly that the spherically symmetric dynamics is gradient-driven with kinetic coefficients $D(\partial m(r)/\partial r)$ and obeys a fluctuation-dissipation relation. It also emphasizes the central role played by the *mass* in this approach to nucleation – as opposed to density or any order parameter – which is a consequence of the fact that the theory is founded in hydrodynamics where conservation of mass is always true.

C. Nucleation pathway

The stochastic evolution equation, Eq. (25), could be used as the basis of numerical simulations. However, the goal here is to provide a description of the nucleation pathway and to make connection with more phenomenological descriptions of the nucleation processes. It is therefore necessary to characterize the typical or expected pathway. One natural way to do so is to determine the most likely path. When the noise is weak, it can be expected that typical paths will be close to the MLP, so that the latter can be viewed as characterizing the process. If the noise is strong, the status of the MLP is less clear and it may only be one of many alternative paths that could be explored by the system.

1. The path probability

The question of the determination of the most likely path for a stochastic process dates back to Onsager and Machlup.[31] For Gaussian processes with constant diffusion matrix the result is relatively easy to derive, whereas for the most general cases there are several subtleties and the most general result seems to have first been derived by Graham.[32] Here, and in the following sub-subsection the theory of the MLP is reviewed and then its application to the problem of nucleation is discussed.

Consider the general stochastic dynamics given by

$$\frac{dx_i}{dt} = b_i(\mathbf{x}(t)) + \epsilon Q_{ij}(\mathbf{x}(t))\xi_j(t), \qquad (26)$$

where the noise terms, $\xi_j(t)$, are Gaussian, white and uncorrelated. Note that here, and in the following, the Einstein summation convention is used. Interpreted within the Ito calculus, Graham[32] shows that the probability to observe a given continuous path is given by

$$P(\mathbf{x}(t)) \sim \exp\left(-\int_{t_1}^{t_N}\mathcal{L}(\mathbf{x}(t), \dot{\mathbf{x}}(t))dt\right), \qquad (27)$$

where the Lagrangian is

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2}\epsilon^{-2}(\dot{x}_i - c_i(\mathbf{x}))D_{il}^{-1}(\mathbf{x})(\dot{x}_l - c_l(\mathbf{x}))$$
$$+ \frac{1}{2}\det(Q(\mathbf{x}))\frac{\partial}{\partial x_i}\frac{c_i(\mathbf{x})}{\det(Q(\mathbf{x}))} + \epsilon^2\frac{R}{12} \qquad (28)$$

with

$$D_{il} \equiv Q_{il}Q_{jl},$$
$$c_i \equiv b_i - \frac{1}{2}\epsilon^2 \det(Q)\frac{\partial}{\partial x_j}\frac{D_{ji}}{\det(Q)}, \qquad (29)$$

the Ricci curvature scalar formed using $D_{il}^{-1}(\mathbf{x}(t))$ as a metric in x−space is

$$R = D_{ab}D_{cd}\frac{1}{2}\left(\frac{\partial^2 D_{ab}^{-1}}{\partial x_c \partial x_d} + \frac{\partial^2 D_{cd}^{-1}}{\partial x_a \partial x_b} - \frac{\partial^2 D_{bc}^{-1}}{\partial x_a \partial x_d} - \frac{\partial^2 D_{ad}^{-1}}{\partial x_b \partial x_c}\right)$$

$$+ D_{ab}D_{cd}D_{ef}^{-1}\left(\Gamma_{ab}^e \Gamma_{cd}^f - \Gamma_{db}^e \Gamma_{ca}^f\right)$$

and the Christoffel symbols of the second kind are

$$\Gamma_{bc}^a = \frac{1}{2}D_{ad}\left(\frac{\partial D_{db}^{-1}}{\partial x_c} + \frac{\partial D_{dc}^{-1}}{\partial x_b} - \frac{\partial D_{bc}^{-1}}{\partial x_d}\right). \quad (30)$$

The probability of observing a path is simply the probability of observing the required noise needed to generate the path plus additional terms related to the change in variable from the noise to the path. The first term on the right in the Lagrangian is due to the fact that the probability to observe a sequence of noises is by definition $P(\xi(t)) = \exp(-1/2 \int_{t_1}^{t_N} \xi^2(t)dt)$ together with a replacement of $\xi_j(t)$ based on rearranging the stochastic differential equation. The second and third terms are due to the Jacobean in transforming variables from $\xi(t)$ to $\mathbf{x}(t)$. Note that the first term is of order ϵ^{-2}, the second of order ϵ^0, and the third of order ϵ^2 so that in the weak noise limit, only the first term contributes. In this case, another interpretation of the expression for the path probability is that it is the Freidlin-Wetzel action functional from the theory of large deviations.[33,34] The fact that these results can be written in the fully covariant form given above is, as noted by Graham,[32] necessary since the path probability must be independent of the choice of variables used to express it. Thus, while the original stochastic model has no explicit geometrical interpretation, the focus of *paths*, which are geometrical objects, induces a natural geometrical perspective in which the (inverse of the) covariance of the noise plays the role of a metric in the space of stochastic variables.

Given fixed endpoints, the MLP is determined by maximizing the probability with respect to variations in the path resulting in the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial}{\partial \dot{x}_i}\mathcal{L}(\mathbf{x}(t),\dot{\mathbf{x}}(t)) - \frac{\partial}{\partial x_i}\mathcal{L}(\mathbf{x}(t),\dot{\mathbf{x}}(t)) = 0 \quad (31)$$

or, after some manipulations,

$$\frac{d^2}{dt^2}x_r + \frac{1}{2}\Gamma_{lc}^r\frac{dx_c}{dt}\frac{dx_l}{dt} + D_{ri}\left(\frac{\partial D_{cl}^{-1}c_l}{\partial x_i} - \frac{\partial D_{il}^{-1}c_l}{\partial x_c}\right)\frac{dx_c}{dt}$$

$$= D_{ri}\frac{\partial}{\partial x_i}\left[\frac{1}{2}c_m D_{ml}^{-1}c_l + \epsilon^2\frac{1}{2}\det(Q)\frac{\partial}{\partial x_m}\frac{c_m}{\det(Q)} + \epsilon^4\frac{R}{12}\right]. \quad (32)$$

This can be interpreted as describing a particle with curvilinear coordinates $\mathbf{x}$ moving in response to a force given by the gradient of a potential under the metric $\mathbf{D}^{-1}$. (Note that here and below, it is assumed that one selects a solution that actually corresponds to a maximization of the probability, as opposed to a minimum or saddle point extremization.)

Equation (32) was derived under the condition that the end points of the path are fixed, so it defines a two-point boundary value problem with fixed initial point, $\mathbf{x}(0) = \mathbf{x}_0$, and end point, $\mathbf{x}(t_f) = \mathbf{x}_f$, and the time, t_f, is determined by

minimizing the action. In some cases, the endpoints are not both known *a priori* and one wishes instead to determine the MLP from any state satisfying some set of constraints to any other satisfying a different constraints. (For example, in the nucleation problem, this might be the path connecting a cluster of total excess mass M_1 to one with total excess mass M_2.) If these constraints involve only the coordinates and are represented, in general, as $K_a(\mathbf{x}(t_0)) = 0$ and $K_a(\mathbf{x}(t_f)) = 0$, then the boundary conditions become

$$\dot{x}_i(0) = c_i(\mathbf{x}(0)) + \sum_a \lambda_a^{(0)}D_{ij}^{-1}(x(0))\left(\frac{\partial K_a(x)}{\partial x_i}\right)_{\mathbf{x}(0)},$$

$$\dot{x}_i(t_f) = c_i(\mathbf{x}(t_f)) + \sum_a \lambda_a^{(f)}D_{ij}^{-1}(x(t_f))\left(\frac{\partial L_a(x)}{\partial x_i}\right)_{\mathbf{x}(t_f)},$$

$$(33)$$

where the λ's are Lagrange multipliers which must be determined. To illustrate, suppose the desired path started at point $\mathbf{x}_0$ but that the final point was determined as having a fixed mass, $M(\mathbf{x}(t_f)) = M_f$. If there were n parameters, there would be n constraint equations, Eq. (33) at the end point x_f together with the condition on the final mass. These $n+1$ constraints would take the place of the n conditions otherwise determined by specifying the final coordinate as well as providing one additional constraint for the Lagrange multiplier.

2. The most likely path when there is a fluctuation-dissipation relation

In the limit of weak noise, i.e., for $\epsilon \to 0$, it can be shown (see Appendix B) that for the case of a generalized potential driven dynamics of the form

$$b_i(\mathbf{x}) = -L_{ij}(\mathbf{x})\frac{\partial V(\mathbf{x})}{\partial x_j} \quad (34)$$

that obeys a fluctuation-dissipation relation,

$$L_{ij}(\mathbf{x}) \propto D_{ij}(\mathbf{x}) \quad (35)$$

the MLP between two metastable states, i.e., points satisfying $b_i(\mathbf{x}) = 0$, *must* pass through a saddle point and follow the deterministic path connecting the saddle point to the metastable states

$$\frac{dx_i}{dt} = \pm L_{ij}(\mathbf{x})\frac{\partial V(\mathbf{x})}{\partial x_j}. \quad (36)$$

It is easily verified by substitution that this is an exact solution to the general expression, Eq. (32), when Eqs. (34) and (35) hold and the initial and final states are metastable points. It can be interpreted as saying that, starting in the metastable state, one follows the *time-reversed dynamics* against the free-energy gradient up to the critical point; beyond the critical point, one follows the forward-time dynamics down to the stable state. Note that this is only a mathematical prescription that serves to determine the MLP: there is no suggestion that the highly dissipative physical system is in any sense time-reversal invariant. Similarly, the MLP is not the same as the deterministic (i.e., $\epsilon \to 0$) limit of the stochastic process: it is a separate concept that concerns the most likely path to move both up and down free-energy barriers, whereas the deterministic path can only describe movement down a free-energy

gradient and is *a priori* incapable of describing movement *up* an energy barrier. While it so happens that for a part of the path between two points that involves moving down a free-energy gradient, the MLP and the deterministic limit coincide, this is not a general identification of the two: thus, while the deterministic path is indeed given by Eq. (36) with the minus sign, the more general – and most certainly fluctuation-driven – MLP is given by Eq. (36) with the appropriate sign for the part of the path being determined.

Recognizing that Eq. (36) is equivalent to steepest-descent in curvilinear coordinates with metric $g_{ij} \equiv L_{ij}^{-1}$ means that the MLP can be determined using the standard gradient-descent algorithms.[35] This involves first locating the saddle point, $\mathbf{x}_s$, using, e.g., eigenvector-following techniques,[35] and then perturbing slightly in the direction of the unstable eigenvector as determined from the generalized eigenvalue problem

$$H_{ij} v_j = \lambda L_{ij}^{-1} v_j, \tag{37}$$

where the Hessian at the saddle point is

$$H_{ij} = \left. \frac{\partial^2 V(\mathbf{x})}{\partial x_i \partial x_j} \right|_{\mathbf{x}_s}. \tag{38}$$

In principle, the perturbation should be infinitesimal but in practice, some small, finite perturbation must be used with the approximation becoming exact as the size of the perturbation goes to zero. From this initial point, the steepest-descent path is determined using Eq. (36) with the positive sign (i.e., the forward-time, deterministic dynamics).

3. Application to nucleation

For the nucleation problem, instead of a collection of variables, x_i, as considered above, one has the density $\rho(\mathbf{r})$ or the cumulative mass function, $m(r)$, which are both fields. Generally, stochastic differential equations involving fields are understood to be defined by some sort of spatial discretization scheme so that the field, say $m(r)$, is actually a collection of values, $m(r_i) \equiv m_i$, defined on a lattice thus allowing the use of the results developed for discrete variables.

With this understanding, it is easy to verify that the general model, Eq. (13), does indeed posses a fluctuation-dissipation relation between the deterministic and fluctuating parts of the dynamics so that one can immediately infer that the MLP between the initial and final phases will pass through a saddle point determined from

$$\frac{\delta F}{\delta \rho(\mathbf{r})} = \mu, \tag{39}$$

where μ is the chemical potential of the mother phase and will be determined by moving away from this state in the forward and backward directions either by following the time-reversed dynamics from the metastable state to the critical state and then the forward-time dynamics thereafter,

$$\frac{\partial \rho(\mathbf{r})}{\partial t} = \pm D \nabla \cdot \rho(\mathbf{r}) \nabla \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})}, \tag{40}$$

or, equivalently, by perturbing slightly away from the critical cluster in the direction (or anti-direction) of the most unstable

eigenvector and following the forward-time dynamics to the respective minima. This is the first significant result of this analysis and represents a generalization of the usual dynamic density functional theory to barrier crossing problems. Lacking this result, applications of DDFT have until now had to rely on the artificial introduction of noise via the initial conditions of the density in order to induce nucleation (for a recent example, see Ref. 36). Such a procedure is in fact inconsistent in so far as the density appearing in DDFT is already a noise- and initial-condition-averaged quantity and so, for example, for mother phases which are fluids it should simply be a constant.

In the case of spherical symmetry, the dynamical model given in Eq. (25) is obviously of the same form as Eq. (26) and obeys the fluctuation-dissipation constraints of Eqs. (34) and (35). The MLP in the limit of weak noise, therefore, passes through the saddle point determined from

$$0 = \frac{\delta \beta F[\rho]}{\delta m(r)} = \left. \frac{\partial}{\partial r} \frac{\delta \beta F[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho(r)} \tag{41}$$

and is given by

$$\frac{\partial m(r)}{\partial t} = \pm D \frac{\partial m(r)}{\partial r} \frac{\delta \beta F[\rho]}{\delta m(r)}$$

$$= \pm D \frac{\partial m(r)}{\partial r} \left. \frac{\partial}{\partial r} \frac{\delta \beta F[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho(r)}, \tag{42}$$

which can be given a geometric interpretation as steepest-descent on the free-energy surface in mass-space with a metric of $(\partial m(r)/\partial r)^{-1}$. Note that this is exactly equivalent to Eqs. (39) and (40) with the assumption of spherical symmetry, thus showing the equivalence of the approaches. It also allows one to define the distance between two mass distributions, $m_0(r)$ and $m_1(r)$, as

$$s[m_0, m_1]$$
$$= \min_T \min_{\text{paths}} \int_0^T \sqrt{\int_0^\infty \frac{\partial m(r,t)}{\partial t} \left(\frac{\partial m(r,t)}{\partial r} \right)^{-1} \frac{\partial m(r,t)}{\partial t} dr} \, dt, \tag{43}$$

where the right-hand side is minimized over all "acceptable" paths connecting the two mass distributions with $m(r, 0) = m_0(r)$ and $m(r, T) = m_1(r)$. An "acceptable" path is one obeying the basic constraints on the cumulative mass distribution that it be monotonically increasing as a function of r and that $m(0) = 0$. These properties mean that the position r may be replaced by $m(r)$ as the variable of integration yielding

$$s[m_0, m_1]$$
$$= \min_T \min_{\text{paths}} \int_0^T \sqrt{\int_0^\infty \left(\frac{\partial m(r,t)}{\partial t} \middle/ \frac{\partial m(r,t)}{\partial r} \right)^2 \frac{\partial m(r,t)}{\partial r} dr} \, dt$$

$$= \min_T \min_{\text{paths}} \int_0^T \sqrt{\int_0^\infty \left(\frac{\partial r(m,t)}{\partial t} \right)^2 dm} \, dt$$

$$= \sqrt{\int_0^\infty \left(m_0^{-1}(x) - m_1^{-1}(x) \right)^2 dx}, \tag{44}$$

where the last line follows from recognizing that the second line now defines a Euclidean metric. Since $m_0^{-1}(x)$ is the location at which distribution $m_0(r)$ has value x and $m_1^{-1}(x)$ is the analogous quantity for the second distribution, this expresses the distance between the distributions in terms of the sum over Euclidean distances between equal-mass points. This structure is entirely due to the presence of the factor $(\partial m(r,t)/\partial r) \sim \rho(r)$ in the dynamical equations and, via the fluctuation-dissipation relation, in the noise. Without it, Eq. (43) would already be Euclidean and the distance function would simply depend on the integral over $(m_0(r) - m_1(r))^2$.

The fact that the exact distance function can be easily calculated means that powerful approximate methods for determining the steepest-descent path, such as the string method[37] and the nudged elastic band,[38,39] can be used to determine the steepest-descent paths defined by Eq. (42) which is to say, the MLP. Another implication concerns the stochastic evolution equations themselves. The fact that the induced distance function is Euclidean in the variable $r(m)$, immediately implies that if the SDE is written in terms of this variable, the noise will be white (i.e., the coefficient of the noise variable $\xi(r;t)$ will be constant rather than state-dependent). That this is the case is easily verified by first observing that

$$\frac{\partial r(m)}{\partial t} = \frac{\partial r}{\partial m}\frac{\partial m}{\partial t} = \left(\frac{\partial m}{\partial r}\right)^{-1}\frac{\partial m}{\partial t}. \quad (45)$$

Using the evolution equation, Eq. (25), gives

$$\frac{\partial r(m)}{\partial t} = \left(\frac{\partial m}{\partial r}\right)^{-1}\frac{\partial m(r)}{\partial t}$$

$$= -D\frac{\delta \beta F[\rho]}{\delta m(r)} - \epsilon \sqrt{2D\left(\frac{\partial m(r)}{\partial r}\right)^{-1}}\xi(r;t), \quad (46)$$

where the term $\delta F[\rho]/\delta m(r)$ must be expressed in terms of $r(m)$. At the moment, it does not seem as if the goal of getting white noise is going to be achieved. However, note that one must express the noise in terms of m as the dependent variable rather than r. Since

$$\langle \xi(r;t)\xi(r';t')\rangle = \delta(r - r')\delta(t - t'),$$

$$= \delta(r(m) - r(m'))\delta(t - t'),$$

$$= \left(\frac{\partial r(m)}{\partial m}\right)^{-1}\delta(m - m')\delta(t - t'), \quad (47)$$

it is clear that one can make the substitution $\xi(r;t) \longmapsto \sqrt{(\partial r(m)/\partial m)^{-1}}\xi(m;t)$ giving

$$\frac{\partial r(m)}{\partial t} = -D\frac{\delta \beta F[\rho]}{\delta m(r)}\bigg|_{r(m)} - \epsilon\sqrt{2D}\xi(m;t) \quad (48)$$

as expected. It is not clear whether this formulation is of any practical significance.

III. ORDER PARAMETERS

So far, the central quantity governing the description of nucleation has been the cumulative mass, $m(r;t)$. However, the classical description of nucleation is typically formulated in terms of a small number, often just one, order

parameter. For example, in classical nucleation theory clusters of the new phase are assumed to be spherical, with a radius R, to have the bulk new phase in the interior of the cluster and to have an interface of vanishing width. The only quantity that can vary is, therefore, the radius and this is the order parameter in terms of which the dynamics are formulated.

The point of view adopted here in order to make a connection with the concept of order parameters is that the parameters must determine an approximation to the cumulative mass distribution or, equivalently, to the density. For example, in the CNT picture described above, a cluster has interior density ρ_0 which is just the bulk density of the new phase; exterior density ρ_∞ which is the density of the metastable phase, a radius R and an interface of negligible width. Hence, it implicitly specifies a density model

$$\rho(r;t) \longmapsto \rho(r;R(t)) \equiv \rho_0\Theta(R(t) - r) + \rho_\infty\Theta(r - R(t)) \quad (49)$$

and a cumulative mass of

$$m(r;t) \longmapsto m(r;R(t))$$

$$\equiv \frac{4\pi}{3}r^3\rho_0\Theta(R(t) - r)$$

$$+ \frac{4\pi}{3}(R^3(t)\rho_0 + (r^3 - R^3(t))\rho_\infty)\Theta(r - R(t)). \quad (50)$$

In general, the approximating function can depend on many parameters so that the general case is $\rho(r;t) \longmapsto \rho(r;\mathbf{x}(t))$ for a set of time-dependent parameters $\mathbf{x}(t)$. The goal in this section is to determine how the time-dependence of the restricted description in terms of the order parameters can be developed from the general theory. The reason for investigating this question is that it has been shown that one can derive the CNT description of the critical cluster as the lowest order result in a systematic expansion of the DFT free energy based on these types of parameterizations of the density[12,17] so that this gives a path for making contact between the general theory and CNT. Finally, it is also relevant to note that any practical scheme for the integration of Eq. (25) will involve some sort of representation of the field in terms of a finite number of parameters, so the results obtained here will be of equal use in numerical calculations as discussed below.

In the following, two different approaches to the use of order parameters will be described. The first is solely concerned with the determination of the most likely path for nucleation and involves relatively few assumptions. The second is aimed at determining approximate stochastic equations for the order parameters which, in turn, can be used to determine the MLP. The approaches have different strengths and weaknesses: the former involves fewer assumptions but is more restricted in its use, while the latter gives broader contact with earlier work but is somewhat heuristic. Both methods become exact in the limit of a complete set of order parameters and of course agree with one another in that limit as demonstrated below.

A. Direct determination of pathway using order parameters

In the weak noise limit, the probability for a path based on the spherically-symmetric dynamics, Eq. (23), is

$$P \sim \exp\left(-\frac{1}{8\pi D\varepsilon}\int_0^T \mathcal{L}\,dt\right) \quad (51)$$

with Lagrangian

$$\mathcal{L} = \frac{1}{2}\int_0^\infty \frac{1}{r^2\rho(r)}\left(\frac{\partial m(r)}{\partial t} - Dr^2\rho(r)\frac{\partial}{\partial r}\frac{1}{r^2}\frac{\delta\beta F[\rho]}{\delta\rho(r)}\right)^2 dr. \quad (52)$$

It is shown in Appendix C that if the density is parametrized as $\rho(r;t) = \rho(r;\mathbf{x}(t))$, then the weak-noise Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}g_{ab}(\mathbf{x})\frac{dx_a}{dt}\frac{dx_b}{dt} + D\frac{dx_a}{dt}\frac{\partial\beta\Omega[\rho]}{\partial x_a} + V(\mathbf{x}), \quad (53)$$

where $\Omega[\rho] = F[\rho] - \mu\rho$ and

$$V(\mathbf{x}) = \frac{1}{2}D^2\int_0^\infty r^2\rho(r)\left(\frac{\partial}{\partial r}\frac{1}{r^2}\frac{\delta\beta F[\rho]}{\delta\rho(r)}\right)^2 dr \quad (54)$$

and where the metric $g_{ab}(\mathbf{x})$ is

$$g_{ab} = \int_0^\infty \frac{1}{4\pi r^2\rho(r)}\frac{\partial m(r)}{\partial x_a}\frac{\partial m(r)}{\partial x_b}dr. \quad (55)$$

Maximizing the path probability gives the Euler-Lagrange equations

$$g_{il}\frac{d^2x_l}{dt^2} + \frac{1}{2}\frac{dx_l}{dt}\frac{dx_j}{dt}\left(\frac{\partial g_{il}}{\partial x_j} + \frac{\partial g_{ij}}{\partial x_l} - \frac{\partial g_{lj}}{\partial x_i}\right) = \frac{\partial}{\partial x_i}V(\mathbf{x}) \quad (56)$$

or

$$\frac{d^2x_i}{dt^2} + \Gamma_{lj}^i\frac{dx_l}{dt}\frac{dx_j}{dt} = g_{ir}^{-1}\frac{\partial}{\partial x_r}V(\mathbf{x}), \quad (57)$$

where the Christoffel symbol of the second kind is

$$\Gamma_{kl}^i = \frac{1}{2}g_{im}^{-1}\left(\frac{\partial g_{mk}}{\partial x_l} + \frac{\partial g_{ml}}{\partial x_k} - \frac{\partial g_{kl}}{\partial x_m}\right). \quad (58)$$

Equation (57) is recognized as the equation of motion of a particle expressed in curvilinear coordinates with a force derived from the potential $V(\mathbf{x})$. If the end points are determined by the constraints,

$$J_l(\mathbf{x}(0)) = 0, \quad l = 1, \ldots, n,$$
$$K_l(\mathbf{x}(T)) = 0, \quad l = 1, \ldots, n, \quad (59)$$

then the boundary conditions are

$$\frac{dx_i}{dt}\bigg|_{t=0} + \frac{D}{4\pi}g_{il}^{-1}(\mathbf{x})\frac{\partial\beta F[\rho]}{\partial x_l}\bigg|_{t=0} = \mu_r g_{il}^{-1}(\mathbf{x})\frac{\partial J_r}{\partial x_l}\bigg|_{t=0},$$

$$\frac{dx_i}{dt}\bigg|_{t=T} + \frac{D}{4\pi}g_{il}^{-1}(\mathbf{x})\frac{\partial\beta F[\rho]}{\partial x_l}\bigg|_{t=T} = \lambda_r g_{il}^{-1}(\mathbf{x})\frac{\partial K_r}{\partial x_l}\bigg|_{t=T}, \quad (60)$$

where μ_r and λ_r are Lagrange multipliers introduced in the course of minimizing the action under the constraints. Note that if, e.g., the initial point is fixed, then $J_r = x_r(T) - x_{0r}$ and the constraint just serves to evaluate the (uninteresting)

Lagrange multiplier: in this case the constraint equations play no role. On the other hand, if a nontrivial constraint is applied, then instead of fixing the value of $x_r(T)$ a condition is imposed on the velocities at time T. For example, for nucleation one might know the initial state, $\mathbf{x}(0)$, and might want to know the most likely final state with a given excess mass

$$M = 4\pi\int_0^\infty (\rho(r;\mathbf{x}(T)) - \rho_\infty)r^2 dr. \quad (61)$$

Then, the boundary conditions at the end point would be

$$\frac{dx_i}{dt}\bigg|_{t=T} + \frac{D}{4\pi}g_{il}^{-1}(\mathbf{x})\frac{\partial\beta F[\rho]}{\partial x_l}\bigg|_{t=T}$$
$$= \lambda g_{il}^{-1}(\mathbf{x}(T))4\pi\int_0^\infty \frac{\partial\rho(r;\mathbf{x}(T))}{\partial x_l(T)}r^2 dr. \quad (62)$$

B. An order-parameter dynamics

The spatial and temporal components of the density can be separated by writing it in some parametrized form as $\rho(r;t) = \rho(r;\mathbf{x}(t))$, where $\mathbf{x}(t)$ is a collection of scalar parameters. One common method of doing so would be to expand the density in a complete set of basis functions. Alternatively, the parameters could represent a discretization of the density, $x_n = \rho(n\Delta)$ for some small distance Δ. As these examples indicate, there will be in general an arbitrary number of parameters. Their time evolution can be developed by noting that the $m(r)$ will also be a function of the parameters so that from Eq. (20), one can write

$$\frac{\partial m(r;\mathbf{x}(t))}{\partial x_i}\frac{dx_i}{dt} = D4\pi r^2\rho(r;\mathbf{x}(t))\frac{\partial}{\partial r}\frac{\delta\beta F[\rho]}{\delta\rho(\mathbf{r})}\bigg|_{\rho(r;\mathbf{x}(t))}$$
$$- \epsilon\sqrt{8\pi r^2 D\rho(r;\mathbf{x}(t))}\xi(r;t). \quad (63)$$

Suppose that this is multiplied by some function $W_j(r;\mathbf{x}(t))$ and integrated giving

$$g_{ij}(\mathbf{x})\frac{dx_i}{dt} = D\int_0^\infty W_j(r;\mathbf{x}(t))4\pi r^2\rho(r;\mathbf{x}(t))$$
$$\times \frac{\partial}{\partial r}\left(\frac{\delta\beta F[\rho]}{\delta\rho(\mathbf{r})}\right)_{\rho(r;\mathbf{x}(t))}dr$$
$$- \int W_j(r;\mathbf{x}(t))\epsilon\sqrt{8\pi r^2 D\rho(r;\mathbf{x}(t))}\xi(r;t)dr, \quad (64)$$

where

$$g_{ij}(\mathbf{x}) = \int_0^\infty W_i(r;\mathbf{x}(t))\frac{\partial m(r;\mathbf{x}(t))}{\partial x_j}dr. \quad (65)$$

The autocorrelation of the noise term is easily calculated giving

$$\epsilon^2 2D\int_0^\infty 4\pi r^2 W_i(r;\mathbf{x}(t))W_j(r;\mathbf{x}(t))\rho(r;\mathbf{x}(t))dr. \quad (66)$$

While these equations are exact regardless of the choice of $W_i(r;\mathbf{x}(t))$, it will soon be apparent the result is particularly simple if the diffusion matrix and the matrix $g_{ij}(\mathbf{x})$ are required to be proportional which is only possible if, modulo a

multiplicative constant,

$$W_i(r; \mathbf{x}(t)) = \frac{1}{4\pi r^2 \rho(r; \mathbf{x}(t))} \frac{\partial m(r; \mathbf{x}(t))}{\partial x_i}, \quad (67)$$

so that $\mathcal{D}_{ij}(\mathbf{x}) = 2Dg_{ij}(\mathbf{x})$ with

$$g_{ij}(\mathbf{x}) = \int_0^\infty \frac{1}{4\pi r^2 \rho(r; \mathbf{x}(t))} \frac{\partial m(r; \mathbf{x}(t))}{\partial x_i} \frac{\partial m(r; \mathbf{x}(t))}{\partial x_j} dr, \quad (68)$$

i.e., the same metric as found above (see Eq. (55)). Another result of this choice is that the thermodynamic driving force becomes

$$\int_0^\infty W_j(r; \mathbf{x}(t)) 4\pi r^2 \rho(r; \mathbf{x}(t)) \frac{\partial}{\partial r} \left.\frac{\delta \beta F[\rho]}{\delta \rho(\mathbf{r})}\right|_{\rho(r;\mathbf{x}(t))} dr$$

$$= \int_0^\infty \frac{\partial m(r; \mathbf{x}(t))}{\partial x_i} \frac{\partial}{\partial r} \left.\frac{\delta \beta F[\rho]}{\delta \rho(\mathbf{r})}\right|_{\rho(r;\mathbf{x}(t))} dr$$

$$= \left[\frac{\partial m(r; \mathbf{x}(t))}{\partial x_i} \left.\frac{\delta \beta F[\rho]}{\delta \rho(\mathbf{r})}\right|_{\rho(r;\mathbf{x}(t))} \right]_0^\infty$$

$$- \int_0^\infty \frac{\partial^2 m(r; \mathbf{x}(t))}{\partial x_i \partial r} \left.\frac{\delta \beta F[\rho]}{\delta \rho(\mathbf{r})}\right|_{\rho(r;\mathbf{x}(t))} dr$$

$$= \left[\frac{\partial m(r; \mathbf{x}(t))}{\partial x_i} \left.\frac{\delta \beta F[\rho]}{\delta \rho(\mathbf{r})}\right|_{\rho(r;\mathbf{x}(t))} \right]_0^\infty$$

$$- \int \frac{\partial \rho(r; \mathbf{x}(t))}{\partial x_i} \left.\frac{\delta \beta F[\rho]}{\delta \rho(\mathbf{r})}\right|_{\rho(r;\mathbf{x}(t))} d\mathbf{r}. \quad (69)$$

For the problem of nucleation, the first term gives no contribution at $r = 0$, while for very large r, one expects the system to have the bulk properties giving

$$\lim_{r \to \infty} \frac{\partial m(r; \mathbf{x}(t))}{\partial x_i} \left.\frac{\delta \beta F[\rho]}{\delta \rho(\mathbf{r})}\right|_{\rho(r;\mathbf{x}(t))} = \frac{\partial N(\mathbf{x}(t))}{\partial x_i} \mu, \quad (70)$$

where N is the total number of particles and μ is the chemical potential in the bulk. Using the functional chain rule, the second term is recognized as

$$\int \frac{\partial \rho(r; \mathbf{x}(t))}{\partial x_i} \left.\frac{\delta \beta F[\rho]}{\delta \rho(\mathbf{r})}\right|_{\rho(r;\mathbf{x}(t))} d\mathbf{r} = \frac{\partial \beta F(\mathbf{x})}{\partial x_i}, \quad (71)$$

where $F(\mathbf{x}) = F[f]$. (Note that this is only exact if the parametrization is complete.) Combining these results gives the stochastic model

$$g_{ij}(\mathbf{x}) \frac{dx_j}{dt} = -D \frac{\partial \beta \Omega}{\partial x_i} - \epsilon \int \sqrt{\frac{2D}{4\pi r^2 \rho(r; \mathbf{x}(t))}}$$

$$\times \frac{\partial m(r; \mathbf{x}(t))}{\partial x_i} \xi(r; t) dr, \quad (72)$$

where again $\Omega = F - \mu N$. Ideally, one would like to replace the noise by a simpler form giving the same autocorrelation function; however, there is a complication. The use of the usual chain rule for derivatives means that the resulting stochastic differential equation must be understood in the Stratonovich interpretation[23] and, unfortunately, in the present case the resulting equation is not Ito-Stratonovich equivalent. It is, therefore, not the case that one can substitute one noise term for another with the same autocorrela-

tion matrix as this only holds in the Ito form. As described in Appendix F, the spurious drift arising from Ito-Stratonovich inequivalence gives rise to an additional contribution to the deterministic part of the equation with the final result

$$\frac{dx_j}{dt} = -Dg_{ij}^{-1}(\mathbf{x}) \frac{\partial \beta \Omega}{\partial x_i} + 2D\epsilon^2 A_i(\mathbf{x}) - \epsilon \sqrt{2D} q_{ji}^{-1}(\mathbf{x}) \xi_j(t) \quad (73)$$

with $g_{ij} = q_{il} q_{jl}$ (note that g is positive semi-definite so this decomposition is always possible) and what will be termed the "anomalous force" (since it does not arise from the thermodynamic driving force) is

$$A_i(\mathbf{x}) = q_{ik}^{-1}(\mathbf{x}) \frac{\partial q_{jk}^{-1}(\mathbf{x})}{\partial x_j} - \frac{1}{2} g_{il}^{-1}(\mathbf{x}) \frac{\partial g_{jm}^{-1}(\mathbf{x})}{\partial x_l} g_{mj}(\mathbf{x})$$

$$+ \frac{1}{2} \left(g_{il}^{-1}(\mathbf{x}) g_{jm}^{-1}(\mathbf{x}) - g_{ij}^{-1}(\mathbf{x}) g_{lm}^{-1}(\mathbf{x}) \right)$$

$$\times \int_0^\infty \frac{1}{4\pi r^2 \rho^2(r; \mathbf{x})} \frac{\partial \rho(r; \mathbf{x})}{\partial x_l} \frac{\partial m(r; \mathbf{x})}{\partial x_j} \frac{\partial m(r; \mathbf{x})}{\partial x_m} dr. \quad (74)$$

The presence of the anomalous force means that there is no exact fluctuation-dissipation relation. However, since it has its origin in the noise, and is therefore of order ϵ^2, it will not affect the weak-noise limit and the MLP will be determined as usual from the forward- and backward-time dynamics starting at the critical point. Again, it is important to remember that in the strong-noise regime Eq. (73) must be understood in the Stratonovich interpretation: the equivalent Ito equation will have a modified form of the anomalous force (see Appendix F for details).

The advantages of the choice for $W_i(r; \mathbf{x}(t))$ are now evident. First, the particular choice used here results in the free-energy term having the form of a simple gradient of the free energy due to the ability to integrate by parts in Eq. (69). Second, the fact that the fluctuation-dissipation relation is preserved means that in the weak-noise limit one recovers the classical behavior with the MLP passing through the critical cluster.

Equation (73) is exact when the parameters are determined from a complete representation of the density. For a *finite* collection of parameters, it seems reasonable to continue to use Eq. (73) as an approximation and numerical evidence in support of this will be given below. Appendix E gives an alternative derivation of this model showing that it is reasonable even for an incomplete (i.e., approximate) representation of the density provided that the difference between $\rho(r; \mathbf{x}(t))$ and the actual density $\rho(r; t)$ is in some sense small.

C. Most likely path

A heuristic stochastic dynamics for a set of order parameters is given in Eq. (73) above. In the weak noise limit, in which the anomalous force can be neglected, this dynamics is gradient-driven and satisfies a fluctuation-dissipation relation, so that it immediately follows that the MLP connecting the two metastable phases passes through the saddle point

defined by

$$0 = \frac{\partial \beta \Omega}{\partial x_i} \qquad (75)$$

and is determined by steepest-descent according to

$$\frac{dx_i}{dt} = -D g_{ij}^{-1}(\mathbf{x}) \frac{\partial \beta \Omega}{\partial x_j}. \qquad (76)$$

Following the same arguments given above, the general MLP connecting any two states is determined from the Euler-Lagrange equations

$$g_{ij} \frac{d^2 x_j}{dt^2} + \frac{1}{2} \left(\frac{\partial g_{is}}{\partial x_r} + \frac{\partial g_{ir}}{\partial x_s} - \frac{\partial g_{rs}}{\partial x_i} \right) \frac{dx_r}{dt} \frac{dx_s}{dt}$$
$$= \frac{\partial}{\partial x_i} \left(\frac{D^2}{2} \frac{\partial \beta \Omega}{\partial x_j} g_{jl}^{-1} \frac{\partial \beta \Omega}{\partial x_l} \right). \qquad (77)$$

It is easy to show by direct substitution that Eq. (76) is a solution to Eq. (77).

Two expressions for the evolution of order parameters have been derived: Eq. (76), which follows from the general equations for the MLP, Eq. (77), for the special case that the endpoints are metastable states, and Eq. (56), which follows from the minimization of the path probability after assuming a particular test function. In essence, the former results from introducing the approximate density after determining the equations for the MLP, while the latter results from introducing the approximate density before determining the MLP. Each formulation has its advantages: the gradient formulation shares certain properties with the exact MLP, such as the role of the critical cluster, whereas it is not even clear *a priori* that Eq. (56) will give a path passing through the critical point. On the other hand, Eq. (56) is exact, given the assumed form for the density, whereas the gradient equations are derived via a series of manipulations which are not unique: one could derive any number of equations of motion for the parameters from Eq. (63) by similar manipulations to those used above. It is therefore of some interest to examine the connection between these two formulations.

Comparing Eqs. (57) and (77), it is clear that the only difference is in the source terms and that equivalence would demand that

$$\int_0^\infty r^2 \rho(r) \left(\frac{\partial}{\partial r} \frac{1}{r^2} \frac{\delta \beta F[\rho]}{\delta \rho(r)} \right)^2 dr = \frac{\partial \beta \Omega}{\partial x_j} g_{jl}^{-1} \frac{\partial \beta \Omega}{\partial x_l}. \qquad (78)$$

Now, from the derivations above, it is already known that

$$\frac{\partial \beta \Omega}{\partial x_j} = -\int_0^\infty \frac{\partial m(r)}{\partial x_j} \frac{\partial}{\partial r} \frac{1}{r^2} \frac{\delta \beta F[\rho]}{\delta \rho(r)} dr, \qquad (79)$$

so

$$\frac{\partial \beta \Omega}{\partial x_j} g_{jl}^{-1} \frac{\partial \beta \Omega}{\partial x_l}$$
$$= \int_0^\infty dr' \int_0^\infty dr\, r^2 \rho(r) \left(\frac{\partial}{\partial r} \frac{1}{r^2} \frac{\delta \beta F[\rho]}{\delta \rho(r)} \right)$$
$$\times \left(\frac{\partial}{\partial r'} \frac{1}{r'^2} \frac{\delta \beta F[\rho]}{\delta \rho(r')} \right) \left(r^{-2} \rho^{-1}(r) \frac{\partial m(r)}{\partial x_j} g_{jl}^{-1} \frac{\partial m(r')}{\partial x_l} \right),$$

which is the same as Eq. (78) provided that

$$\frac{\partial m(r)}{\partial x_i} g_{ij}^{-1} \frac{\partial m(r')}{\partial x_j} = r^2 \rho(r) \delta(r - r'). \qquad (80)$$

As demonstrated in Appendix D, this is a completeness relation and is sufficient to give equality between the two descriptions of the MLP. It is satisfied when the parameters result from an expansion of the density in a complete set of basis functions. The use of the order-parameter dynamics for a *finite* collection of order parameters is, therefore, heuristic. However, as shown below, only a small number of parameters is necessary for the heuristic dynamics to show convergence to the exact result.

IV. APPLICATIONS

The classic problem of the nucleation of a liquid droplet in a supersaturated vapor will be used to illustrate the theory described above. Calculations were performed using the squared-gradient free-energy model

$$F[\rho] = \int \left(f(\rho(\mathbf{r})) + \frac{1}{2} K (\nabla \rho(\mathbf{r}))^2 \right) d\mathbf{r}, \qquad (81)$$

where $f(\rho)$ is the bulk free-energy per unit volume. The calculations reported here were performed for the ten Wolde-Frenkel model potential for globular proteins,[40] which consists of a hard core of radius σ and an effective pair potential outside the core

$$v(r > \sigma) = \frac{4\epsilon}{\alpha^2} \left(\left(\frac{1}{(\frac{r}{\sigma})^2 - 1} \right)^6 - \alpha \left(\frac{1}{(\frac{r}{\sigma})^2 - 1} \right)^3 \right), \qquad (82)$$

where ϵ is the energy scale of the potential and the parameter α controls its range. For large values of α, such as $\alpha = 50$ used here, this gives a phase diagram typical of proteins including a metastable high-concentration ("liquid") phase.[41,42] The results presented here pertain to the transition from the low-concentration ("gas") phase to a high-concentration ("liquid") phase which is known to sometimes play a role as the first step in the processes of protein crystallization.[41,43] The bulk Helmholtz free-energy function, $f(\rho)$, was approximated using first-order perturbation theory[5,44] as described in detail in Refs. 42 and 45. The coefficient of the squared-gradient term, K, was evaluated using a recently derived approximation based on the potential.[12]

The most direct approach towards the determination of the nucleation pathway based on the field equations, Eqs. (41) and (42), would be to discretize them by first discretizing the radial coordinate, $r \longmapsto r_i \equiv \delta$, for some small length δ, which then implies a discretization of the density, $\rho(r) \longmapsto \rho_i \equiv \rho(i\delta)$. Since the density is only expected to vary quickly near the interface, the efficiency of this scheme can be significantly improved by allowing for a dynamic discretization whereby the density and position of the lattice points are allowed to vary. Here, this is done by introducing a particular model for the density based on a continuous,

piecewise-linear approximation,

$$\rho(r)$$

$$= \begin{cases} \rho_0, & r < w_0 \\ \rho_0 + (\rho_1 - \rho_0)\frac{r-w_0}{w_1}, & w_0 < r < w_0 + w_1 \\ \rho_1 + (\rho_2 - \rho_1)\frac{r-w_0-w_1}{w_2}, & w_0 + w_1 < r < w_0 + w_1 + w_2 \\ \cdots \\ \rho_\infty, & w_0 + \cdots + w_{N-1} < r. \end{cases}$$

$$(83)$$

The parameters that are allowed to vary freely and dynamically are the combinations ρ_i, w_i (referred to as "links"), while the density ρ_∞ is always fixed at that of the surrounding bulk (e.g., the vapor density for nucleation of liquid droplets). Note that the simplest model consisting of just two links

$$\rho(r) = \begin{cases} \rho_0, & r < w_0 \\ \rho_0 + (\rho_\infty - \rho_0)\frac{r-w_0}{w_1}, & w_0 < r < w_0 + w_1 \\ \rho_\infty, & w_0 + w_1 < r \end{cases} \quad (84)$$

give the minimal description of a cluster in terms of a radius, say w_0, an interfacial width, w_1, and a central density, ρ_0. Similar parameterizations based on these three quantities are common in the literature and here I note in particular one introduced by da Gama and Evans[46] and used recently by Ghosh and Ghosh[13] to study both planar interfaces and spherical clusters

$$\rho(r) = \left[\rho_0 - \frac{\rho_0 - \rho_\infty}{2} \exp(a(r - R)) \right] \Theta(R - r)$$

$$+ \left[\rho_\infty + \frac{\rho_0 - \rho_\infty}{2} \exp(-a(r - R)) \right] \Theta(r - R),$$

$$(85)$$

where the radius is here denoted R. The parameters occurring in all of these approximations are "order parameters" in the sense discussed above and the most likely path of the density can be determined by means of the order-parameter equations, Eq. (76). These are of course only approximations but the piecewise-linear scheme becomes exact as the number of links is increased.

As discussed above, once a parametrization is chosen, there are three steps involved in determining the MLP. The first is to determine the critical cluster, i.e., the saddle point of the free-energy surface in parameter space. This was done using standard eigenvector-following methods[35,47] which result in a given set of parameters describing the approximation to the critical cluster, $\mathbf{x}_c$. At the saddle point, the Hessian matrix determined from the free energy has one negative eigenvalue and a corresponding eigenvector, $\mathbf{v}_{neg}$, which specifies the unstable direction. The second and third parts of the calculation are the numerical integration of the order-parameter equations, Eq. (76), with initial conditions consisting of a small perturbation away from the critical point in the unstable direction, $\mathbf{x}_c \pm \epsilon\mathbf{v}_{neg}$ where the parameter ϵ is chosen so that the change in free energy is some specified amount. The result of one of the perturbations will be that the system falls back to the metastable state, i.e., the cluster will diminish, while the other will send the system towards the final, stable state which means the cluster will grow.

TABLE I. Excess free energy, $\Delta\Omega$, and mass (number of molecules, ΔN) for different density models. The label "PWM" refers to piecewise-linear profiles with M links, Eq. (83).

Approximation	$\beta\Delta\Omega$	ΔN
PW2	98.9	1471
PW4	82.9	1250
PW6	79.4	1211
PW8	77.8	1186
PW10	77.1	1175
Exponential, Eq. (85)	80.4	1246
Exact, Eq. (86)	75.8	1158

A. Nucleation of droplets

The critical cluster for temperature $k_B T/\epsilon = 0.375$ and supersaturation $S \equiv \rho_v/\rho_{vc} = 1.175$, where ρ_{vc} is the vapor density at coexistence, was determined for both the piecewise-linear profiles with varying numbers of links and the exponential profile. The "exact" critical cluster was determined by solving

$$0 = \frac{\delta\Omega[\rho]}{\delta\rho(\mathbf{r})} = K\nabla^2\rho(\mathbf{r}) - f'(\rho(\mathbf{r})) + \mu \quad (86)$$

numerically using a relaxation technique,[48] where the chemical potential is simply $\mu = f'(\rho_v)$. The excess free energy and excess mass of the critical cluster is given in Table I as calculated from several approximations. It demonstrates the gradual convergence of the properties of the critical cluster to the exact result as the number of links in the piecewise-linear profiles is increased. The exponential profile gives a relatively good description of the critical cluster based on only three parameters. Figure 1 shows the exact density profile of the critical cluster and the piecewise profile based on 2, 4, and 10 links. It is apparent that while the simple profile based on
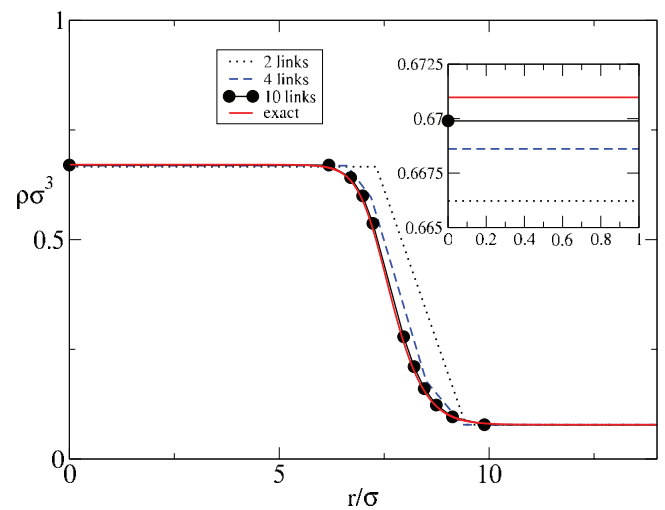

FIG. 1. The density profile of the critical cluster as determined by numerically finding a stationary solution of Eq. (86), the "exact" solution, and from the piecewise linear model with N links, Eq. (83). The inset shows the convergence of the central density towards the exact value as the number of links increases. For the model with 10 links, the figure shows the boundaries of the links as circles. Notice how the density of the discretization is automatically adjusted to be highest in the region of most rapid variation of the profile.

2 links is quite crude, the 10-link profile is already a good approximation to the continuous profile.

The evolution of the droplet was determined beginning with the profile for the critical nucleus. The Hessian for the free energy was evaluated and diagonalized giving in all cases a single negative eigenvalue. The profile was perturbed by adding a multiple of the corresponding eigenvector with the coefficient chosen so as to give a change in free energy of $1\,k_BT$. The equations for the MLP, expressed in terms of order parameters, Eq. (76), were then integrated using the Intel ODE solver library.[49] Depending on the sign of the coefficient used to perturb the state, the system would either fall backwards towards the metastable state (i.e., decreasing cluster size) or forwards towards the stable state (increasing cluster size). The integration was stopped in both cases when the excess free energy reached $1\,k_BT$. For the forward (growing) direction, this was simply a convenient place to stop but in the backward direction, some such cutoff is necessary since the weak-noise assumption must break down.

The independent variable in the integrations is time but care must be taken in interpreting the results. The "time" has the physical meaning of the time taken by the deterministic dynamics to drive the system in the desired direction. When moving forwards in time, this is therefore the physical time. However, when evaluating the MLP via the time-reversed dynamics, this "time" does not correspond to the physical time needed for fluctuations to drive the system up the free-energy gradient. Discussion as to what physical meaning *can* be attached to this time can be found in Bier *et al.*[50] In order to avoid any ambiguity, the results here are displayed in terms of the dimensionless distance along the path defined as

$$ s = \pm\sigma^{-2}\int_{t_1}^{t_2}\sqrt{g_{ab}\frac{dx_a}{dt}\frac{dx_b}{dt}}\,dt \qquad (87) $$

with $s = 0$ being taken to correspond to the smallest cluster and the sign being chosen so that s increases monotonically moving from the smallest cluster to the critical cluster and on to post-critical clusters.

Figure 2 shows the evolution of the central density, the excess mass and the excess free energy of the cluster as determined using the piecewise-linear parametrization with 10 links (19 parameters). At first, the central density remains very close to that of the vapor and the excess free energy is small – on the order of a few k_BT. Surprisingly, even at the very beginning of the process the excess mass is finite and a significant fraction of the mass of the critical cluster. After some time during which the density and mass increase very slowly the growth enters a new regime in which the density increases rapidly to nearly the bulk liquid density. From that point onward, the evolution is unremarkable as the cluster grows to criticality and beyond. Corresponding behavior is seen in the size of the cluster which is characterized in Fig. 3 in two ways: the equimolar radius, R_{eq} defined via $\frac{4\pi}{3}R_{eq}^3(\rho(0) - \rho_\infty) = \int(\rho(\mathbf{r}) - \rho_\infty)\,d\mathbf{r}$, and the total spatial extent, defined for the model given in Eq. (84), as $R_{\text{total}} \equiv \sum_{i=0}^{N-1} w_i$. (For a general representation of the density, this might be characterized as the distance at which the density reaches some small threshold above the background.)
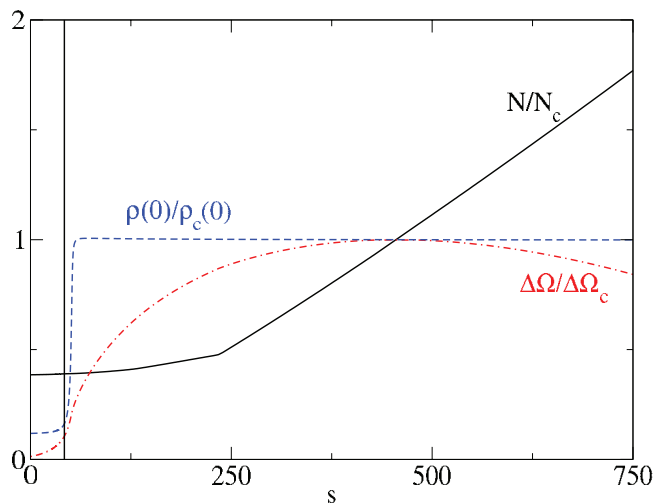

FIG. 2. The variation with path distance, s as defined in Eq. (87), of the excess number of molecules, N, total excess free energy, $\Delta\Omega$, and central density, $\rho(0)$, relative to the values for the critical cluster for a nucleating droplet. The calculations were performed using the piecewise-linear model with 10 links (19 parameters). For reference, the values for the critical cluster are $N = 1175$, $\Delta\beta\Omega = 77.1$, and $\rho(0)\sigma^3 = 0.68$. To aid in comparison, the vertical line picks out the same value of s as in Fig. (3).

Initially, the equimolar radius is nearly constant, while the cluster has very large spatial extent and, counter-intuitively, the cluster *shrinks*. At the same point as the density begins its rapid increase, the equimolar radius also begins to increase even while the total spatial extent continues to diminish. Eventually, the latter reaches a minimum and the cluster grows according to both measures.

These results are in stark contrast to the usual picture of droplet growth as assumed in the Becker-Döring picture (where droplets begin as dimers and grow by attachment and detachment of monomers) as well as the corresponding results from constrained DFT calculations (where droplets begin as zero mass objects that grow monotonically). In fact, how can a droplet "begin" with a finite mass? I shall attempt to answer this question from two different perspectives: first, that of the formalism and second from a physical point of view.
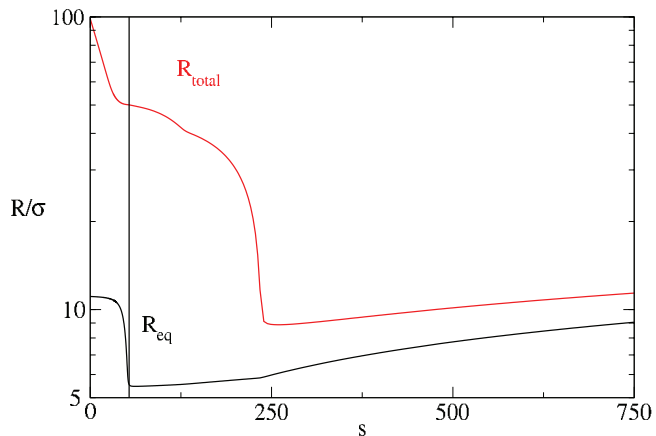

FIG. 3. The variation of the equimolar radius, R_{eq}, and the total radius, R_{total}, as functions of distance along the path, s as defined in Eq. (87), for a nucleating droplet. To aid in comparison, the vertical line picks out the same value of s as in Fig. (2).
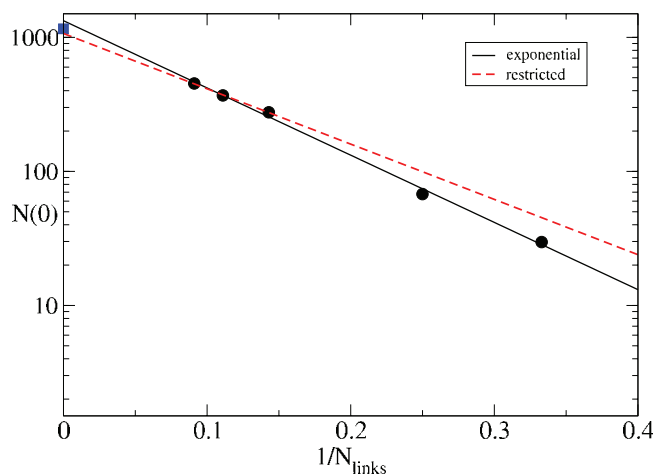
FIG. 4. The excess number of molecules in the cluster at the beginning of the nucleation process, $N(0)$, as a function of the inverse of the number of links in the profile, N_{links}. The log-linear plot shows a nearly linear variation indicating that $N(0) \sim A e^{b/N_{\text{links}}}$. Fitting all of the data to this functional form and extrapolating gives an asymptotic value of $N(0) = 1331$ and using only the three points with the largest value of N_{links} gives $N(0) = 1068$. These values are consistent with the expectation that the limiting value should be the mass of the critical cluster, $N_c = 1175$, shown as the square symbol.
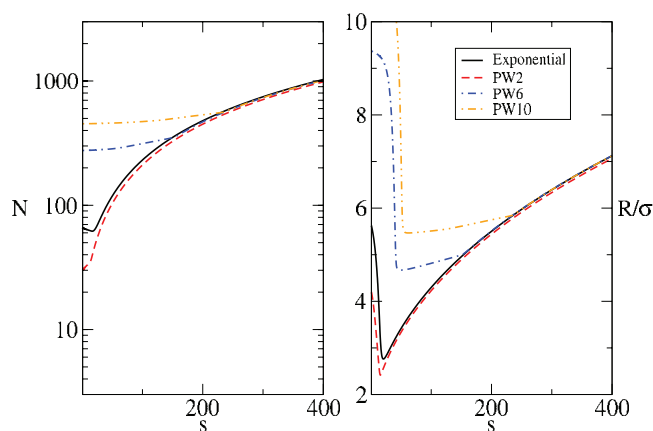

FIG. 5. The excess number of molecules, N, (left panel) and the total radius, R, (right panel) as functions of distance along the nucleation pathway using piecewise-linear models with N links, "PWN,' and the exponential profile. The figures demonstrate that the short-time behavior is sensitive to how well the density profile is represented, whereas all of the models converge at longer times. The exponential profile behaves very similar to the PW2 profile, which has the same number of free parameters, but gives a better estimate of the properties of the critical cluster (see Table I).

In terms of the formalism, the MLP going from "nothing" to the critical droplet is the time-reversed evolution of the dynamics with no noise: in other words, it is the time-reversal of the deterministic evolution starting with a droplet slightly smaller than the critical droplet. What happens in such a case is obvious: the droplet evaporates by shedding mass to the bulk. Since mass is conserved, the excess mass originally in the droplet cannot just disappear but must diffuse over ever larger volumes until it is lost in the (infinite) bulk. It is easy to understand, then, why the total mass is almost constant in the beginning of the process. Indeed, one would expect that the total mass should be the same as that of the critical cluster and should not change at all. That this is the actual result, if one could solve the model with the number of parameters tending to infinity, is indicated in Fig. 4 which shows the initial mass as a function of the number of links in the profile. It is clear that the mass increases with increasing refinement of the profile and that an extrapolation gives a mass close to that of the critical cluster. The fact that the mass is less than that of the critical cluster can be attributed to the difficulty of approximating the density distribution when the density becomes very dilute. Figure 5 illustrates this as well, showing a large variation of the various parameterizations at short times with close convergence at long times. This includes the case of the exponential profile which is also not well suited to representing the diffusive distribution at short times. Note that according to this picture, one would expect the excess density at early times to behave purely diffusively, giving, e.g.,

$$\Delta\rho(r, t) \sim \Delta N (2\pi D |t|)^{-3/2} \exp\left(-\frac{r^2}{2D|t|}\right) \qquad (88)$$

(where "early times" means that t is large and negative). Then, the total mass would be conserved, whereas the equimolar ra-

dius would be given by

$$\Delta N = \frac{4\pi}{3} R_{eq}^3(t) \Delta\rho(0, t)$$

$$\Rightarrow R_{eq}(t) = \left(\frac{3}{4\pi}\right)^{1/3} (2\pi D |t|)^{1/2}, \qquad (89)$$

thus explaining the slow change in the equimolar radius. The fact that the result follows from such simple considerations suggests that, formally, it is very robust being independent of the assumed free-energy model, potential, or any other details.

From a physical perspective, the explanation given so far is not very helpful. Instead, it is useful to divide the process into two parts as indicated in Figs. 2 and 3. During the first stage, matter comes together "from infinity" in the form of a diffuse structure. This can be understood as the formation of a long-wavelength density fluctuation. The presence of such fluctuations is expected and the fact that the structure formed only costs a few $k_B T$ in energy supports this expectation. What appears strange is the directed way that matter seems to come together but this is simply a result of two artifacts. The first is the assumption of spherical symmetry which means that the same thing must occur in all directions at the same time; in reality, such a structure is expected to be messy and ramified. Second is the nature of the MLP. It is well known that the MLP generally follows the shortest possible path up the gradient, with no backtracking or deviation.[50] This gives it an appearance of being directed towards reaching a final state (i.e., the critical state) like a hiker deliberately climbing a hill. Any individual realization of the stochastic process will, in contrast, consist of moves in both directions – up and down the gradient – and will only, on average, eventually trace out the MLP (or something like it). A final artifact is that this is the MLP for a particular process: in other words, it represents a conditional probability in which the condition is that one does actually reach the critical state. A real system will make many abortive attempts to climb the potential gradient

and only after many such attempts will it succeed. Here, in contrast, one studies the one successful attempt which again gives the appearance of being directed.

So far, only the first part of the process – the formation of a density fluctuation – has been discussed. The second part of the process – marked by the rapid increase in density in the nascent cluster – can be interpreted as the occurrence of an actual nucleation event. This part of the process essentially proceeds classically and can be visualized in terms of the Becker-Döring picture or of the typical DFT calculations.

Putting these two pictures together gives a plausible and physically appealing interpretation. Density fluctuations of all types occur in the fluid and it makes intuitive sense that droplet nucleation should be more common in regions of higher density since the free-energy barrier will be lower. Furthermore, the presence of a density fluctuation means that there is excess mass from which to build the cluster. Even in the classical picture (which does not assume any pre-nucleation enrichment), a cluster must somehow draw in material from the surrounding vapor but this part of the process is typically treated separately,[51] if at all. Considered in these terms, the classical process, with no enrichment, seems arguably more mysterious: a small, unstable cluster not only survives but manages to draw in material and, in the fluid, mass flows at such a rate that the cluster is always in contact with a vapor of exactly the bulk density. In the present picture, the needed mass is already locally present so that the formation of the cluster does not lead to the formation of a depletion zone or require the orchestrated flow of matter from far away. And finally, let us note that it is not the case that a fluctuation with exactly the right density appears in order to host the nucleation process: rather, fluctuations of all different sizes occur and if nucleation begins in one that is too small, then it may fail to reach the critical size. On the other hand, it is likely that the probability of a fluctuation decreases with its size so that the *most likely* scenario is of a successful nucleation event occurring in a fluctuation which is just large enough to supply the needed mass. This is the case that the MLP picks out.

B. Nucleation of bubbles

As a second application, the evolution of the reverse process – the formation of a low-concentration "bubble" within the high-concentration "liquid" – was determined. For these calculations, the temperature was the same as previously, $k_B T/\epsilon = 0.375$ but the supersaturation was taken to be $S \equiv \rho_v/\rho_{vc} = 0.9875$ so that the solution is under-saturated making the dense phase unstable towards nucleation of the less dense phase. The results are shown in Fig. 6. A bubble consists of a mass deficit relative to the bulk fluid and so, according to the deterministic dynamics, a sub-critical bubble is expected to vanish by gradually spreading the deficit over ever larger volumes. Therefore, based on the discussion of droplet nucleation, one would expect the process of bubble nucleation to begin with the formation of a long-wavelength, small amplitude deficit of density within which the bubble would nu-
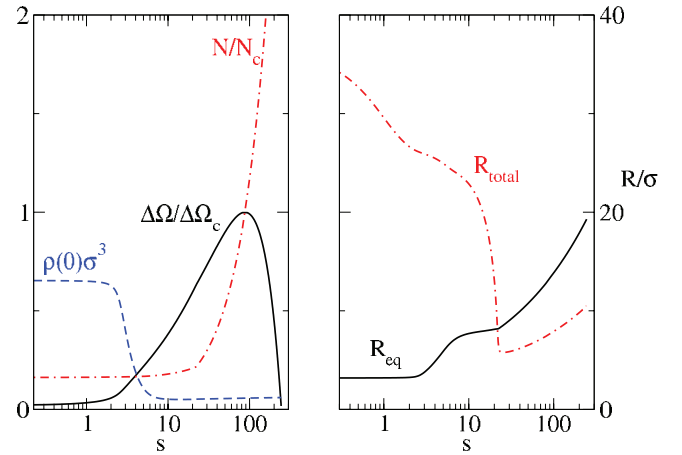

FIG. 6. The left panel shows the variation with path distance, s as defined in Eq. (87), of the excess number of molecules, N, the total excess free energy, $\Delta\Omega$ and the central density, $\rho(0)$. The right panel shows the evolution of the equimolar and total radii. The calculations were performed using the piecewise-linear model with 8 links (15 parameters). The excess number and free energies are scaled to the values for the critical cluster which are $N = -546$, $\Delta\Omega = 48.2\,k_B T$.

cleate. The results of the numerical calculation bear this out and all of the preceding discussion of droplet nucleation can be carried over to the dual process of bubble nucleation.

C. Order-parameter dynamics

As a final illustration of the formalism, the general expressions for the order-parameter dynamics, Eq. (73), are specialized to give a CNT-level of description of the nucleation process for either droplet or bubble nucleation. By "CNT-level" it is meant that there will be only a single order parameter, the radius of the cluster R, and the calculations will be performed to lowest order in w/R where w is the width of the interface. The first step is the evaluation of the metric. This is most easily done for a concrete model and here the minimal piecewise linear model with 2 links will be used. The central density is fixed at the bulk density of the new phase, ρ_0, and the width w is held constant. Then, a straightforward calculation gives

$$g_{RR} = \int_0^\infty \frac{1}{4\pi r^2 \rho(r)} \left(\frac{\partial m(r)}{\partial R}\right)^2 dr$$

$$= \frac{4\pi(\rho_0 - \rho_\infty)^2}{\rho_\infty} R^3 (1 + O(w/R)), \qquad (90)$$

where ρ_∞ is the density of the initial metastable phase. Using this, the SDE for the order parameter becomes

$$\frac{dR}{dt} = -D\frac{\rho_\infty}{4\pi(\rho_0 - \rho_\infty)^2 R^3}\frac{\partial \beta\Omega}{\partial R}$$

$$+ \sqrt{2D\frac{\rho_\infty}{4\pi(\rho_0 - \rho_\infty)^2 R^3}}\,\xi(t) \qquad (91)$$

or, written in terms of the mass of a cluster, $N = \frac{4\pi}{3} R^3 \rho_0$,

$$\frac{dN}{dt} = -D \frac{4\pi \rho_0^2 \rho_\infty}{(\rho_0 - \rho_\infty)^2} R(N) \frac{\partial \beta \Omega}{\partial N}$$

$$+ \sqrt{2D \frac{4\pi \rho_0^2 \rho_\infty}{(\rho_0 - \rho_\infty)^2} R(N)} \xi(t). \quad (92)$$

Recalling that this equation must be interpreted in the Stratonovich sense, it is equivalent to a Fokker-Planck equation for the probability of observing a cluster of size N at time t, $p(N, t)$, of the form[23]

$$\frac{\partial p(N, t)}{\partial t} = \frac{\partial}{\partial N} \left[D \frac{4\pi \rho_0^2 \rho_\infty}{(\rho_0 - \rho_\infty)^2} R(N) \left(\frac{\partial \beta \Omega}{\partial N} + \frac{1}{6N} \right) \right.$$

$$+ \left. D \frac{4\pi \rho_0^2 \rho_\infty}{(\rho_0 - \rho_\infty)^2} R(N) \frac{\partial}{\partial N} \right] p(N, t). \quad (93)$$

This can be compared to the classical result of Zeldovich (see Eq. (9.27) of Ref. 1), which is derived using the Becker-Döring picture

$$\frac{\partial p(N, t)}{\partial t} = \frac{\partial}{\partial N} \left[f(N) \frac{\partial \beta \Omega}{\partial N} + f(N) \frac{\partial}{\partial N} \right] p(N, t), \quad (94)$$

where the monomer attachment frequency, $f(N)$, in the case of diffusion-limited kinetics is given by (see Eq. (10.18) of Ref. 1)

$$f(N) = \eta 4\pi R(N) D \rho_\infty \quad (95)$$

and where all unknown details of the monomer attachment process are contained in the sticking coefficient, η. In the CNT limit of large N, the factor of $1/(6N)$ in Eq. (93) can be ignored and it then agrees with the Zeldovich result with the sticking coefficient identified as

$$\eta = \frac{\rho_0^2}{(\rho_0 - \rho_\infty)^2}. \quad (96)$$

For the nucleation of droplets from vapor, one expects that $\rho_0 \gg \rho_\infty$ so that this gives $\eta \approx 1$. This shows that by specializing the general theory to the CNT regime of large clusters and a single order parameter, one is able to recover all of the elements of the classical formalism.

Similarly, the free energy in the squared-gradient approximation is[12]

$$\Omega[\rho] - \Omega_\infty = 4\pi \frac{R^3}{3} (\omega_0 - \omega_\infty) + 4\pi R^2 \gamma, \quad (97)$$

where the excess surface free energy is

$$\gamma = w\overline{\omega} + K \frac{(\rho_\infty - \rho_0)^2}{2w} \left(1 + O\left(\frac{w}{R} \right) \right) \quad (98)$$

and the density-averaged free-energy per unit volume is

$$\overline{\omega} = \frac{1}{(\rho_0 - \rho_\infty)} \int_{\rho_\infty}^{\rho_0} (\omega(x) - \omega_\infty) \, dx. \quad (99)$$

For a large interface, the width can be estimated by minimizing the free energy of the critical cluster giving

$$w = \sqrt{\frac{(\rho_0 - \rho_\infty)^2 K}{2(\overline{\omega} - \omega_{coex})}}, \quad (100)$$

where ω_{coex} is the free-energy per unit volume at coexistence.[12] All quantities here are, therefore, directly determined by the interaction potential.

Substituting these elements into Eq. (73) gives

$$\frac{dR}{dt} = -D \frac{\rho_\infty (\omega_0 - \omega_\infty)}{(\rho_0 - \rho_\infty)^2} R^{-1} \left(1 - \frac{R_c}{R} \right)$$

$$- \sqrt{D \frac{\rho_\infty (\omega_\infty - \omega_0)}{2\pi (\rho_0 - \rho_\infty)^2} R^{-3}} \xi(t), \quad (101)$$

where

$$R_c = \frac{2\gamma}{\omega_\infty - \omega_0} \quad (102)$$

is the usual CNT expression for the critical radius. Note that for a single order parameter, the anomalous force vanishes and so this model applies to the strong-noise regime. In principle, it can be used to determine a mean first-passage time and, hence, nucleation rate as will be discussed elsewhere. For now, it is interesting to make one further observation which is that for large, *post-critical* clusters, where the noise becomes unimportant, this gives the deterministic growth law

$$\frac{dR}{dt} \approx -D \frac{\rho_\infty (\omega_0 - \omega_\infty)}{(\rho_0 - \rho_\infty)^2} R^{-1} \quad (103)$$

implying that $R \sim t^{1/2}$ which is the well-known result for diffusion-limited cluster growth.[51,52]

V. CONCLUSIONS

The goal of this work is a description of nucleation based on a dynamical, non-equilibrium description of fluctuations with the additional aim of making contact with the usual ideas of the classical nucleation theory. The purpose is to develop a consistent and unique description of nucleation pathways, kinetics, and post-critical cluster growth. Starting with fluctuating hydrodynamics this goal is achieved for the particular case of Brownian particles where the limit of strong dissipation provides several simplifications. The nucleation pathway is characterized as the most likely path and a reduced description in terms of order parameters is developed. The order parameter description includes both a Langevin dynamics for the order parameters and a description of the most likely path. It is noted that in the weak noise limit, all levels of description lead to a unique description of the MLP that passes through the saddle point of the free-energy functional and that is determined by steepest-descent on the free-energy surface governed by a metric that is unambiguously specified. At this level, the goal of making contact with CNT is achieved with the role of the free energy emerging from the nonequilibrium description rather than being assumed *a priori*. Furthermore, as illustrated in the simple case of a single order parameter, one also has a Langevin description of the process which can be used to develop a rate theory that recovers the classical results and is consistent with the expected post-critical growth law. Thus, the goal of giving a unified description of these three elements – rates, pathways, and growth-laws – is also achieved.

One of the main theoretical results concerns the determination of the nucleation pathway as characterized by the most likely path. It is shown that, in the limit of weak noise, this is determined by the deterministic part of the dynamics moving either forward or backward in time away from the critical point. This result represents a generalization of dynamic density functional theory to problems involving barrier crossing. The same result holds when the dynamics are written in terms of order parameters. In all cases, the MLP can, alternatively, be viewed as being determined by steepest-descent on the free-energy surface under a prescribed metric. It is noted that the strong-noise limit may also be handled but investigation of the effect of strong noise is deferred to a later date. In general, it is seen that one recovers the elements of CNT in the weak-noise limit which, therefore, plays a role analogous to the quasi-classical limit in quantum theory.

It is interesting to compare this result with previous approaches. As discussed in the Introduction, there have been two primary approaches to the description of nucleation pathways using the DFT. The first is the minimization of the free energy under a constraint that stabilizes pre-critical and post-critical clusters. There is no direct analogy to this method in the present theory. The other approach that has more recently begun to be used is steepest-descent on the free-energy surface. The problem then is that it is necessary to define a distance in the parameter space (whether it be the space of density distributions or a space of parameters) and there has not been a compelling, much less unique, prescription for doing so. In Refs. 11 and 17, it was proposed that a natural choice was the Euclidean distance in density space,

$$d^2[\rho_1, \rho_2] = \int (\rho_1(\mathbf{r}) - \rho_2(\mathbf{r}))^2 d\mathbf{r} \qquad (104)$$

giving, for arbitrary parameters,

$$g_{ab}^\rho = \int \frac{\partial \rho(\mathbf{r})}{\partial x_a} \frac{\partial \rho(\mathbf{r})}{\partial x_b} d\mathbf{r} \qquad (105)$$

which should be contrasted with the present results, Eqs. (44) and (68). The differences between these metrics are not trivial: integration of the steepest-descent equations, Eq. (76), using the heuristic g_{ab}^ρ gives completely different results for droplet nucleation from those presented above. In fact, the result is the "classical" one whereby a droplet begins as a cluster of zero mass and the mass slowly increases as the density and radius increases during cluster formation.[11, 12] It is interesting to note that the difference in the results is not due to the factor of the inverse density in the expression for the correct metric, but rather is due to the fact that it is based on the cumulative mass distribution rather than the density. In retrospect, this could be argued to be quite physical: in moving from one density configuration to another, what actually required is the movement of mass from one place to another so that the "closeness" of one configuration to another is more naturally characterized by how different they are in mass distribution. Furthermore, it is to be expected to result from a formulation that respects the conservation of mass and in which, therefore, the rate of transport of mass is a limiting factor. Conversely, the incorrect, heuristic metric would result from

determining the MLP from a dynamics of the form

$$\frac{d\rho(\mathbf{r}, t)}{dt} = -D \frac{\delta\Omega[\rho]}{\delta\rho(\mathbf{r}, t)} + \sqrt{2Dk_BT}\xi(\mathbf{r}, t), \qquad (106)$$

which is to say, a non-conserved dynamics. (In this equation, $\Omega[\rho] = F[\rho] - \mu N[\rho]$ for fixed chemical potential, μ. In other words, this equation for a non-conserved total mass is to be understood in the grand-canonical ensemble.) While this could be appropriate for describing some types of simulations that do not conserve mass, it is clearly inappropriate for real physical systems. Finally, it goes without saying that doing steepest-descent using an *ad hoc* metric based on some parametrization of the free energy is the same thing as assuming an *ad hoc* dynamics which will not, in general, be physically relevant.

Another point that has not been emphasized is the ability of this formalism to distinguish between multiple pathways. Just as there may be multiple local minima, and only one true minimum, of a function, there can be multiple local MLPs and only one truly *most* likely path. The simplest and most relevant case occurs when the free-energy surface possesses multiple saddle points. For example, in the case of protein nucleation there is a question of whether the system proceeds directly from low-concentration solution to solid or whether it proceeds in two steps by passing through an intermediate, metastable high-concentration phase of the type studied above. Until now, the only question that was accessible theoretically was the height of the energy barriers and this alone was assumed to determine the relevant path. However, this ignores the role of dynamics (for a very interesting discussion of this point see Ref. 53). The present theory allows a way to address this problem. First, it is possible to determine local MLPs, e.g., in the weak-noise limit by means of steepest-descent from the various saddle points. Then, given these paths, their relative probabilities are easily determined by evaluating Eqs. (27)–(29). The absolute probability of the paths requires knowledge of possibly unknown normalization constants but the relative probabilities can be determined directly by taking the ratio the results. This will be elaborated at a later time.

When applied to globular proteins, namely, the case of nucleation of the high-concentration amorphous phase from the low-concentration solution, an unexpected result is found whereby the process involves two steps. In the first step, a long-wavelength, low-amplitude density fluctuation occurs. In the second step, a nucleation event takes place within this region of enhanced density. I argued that this is to be expected since, on the one hand, such density fluctuations *do* occur, and at little cost in free energy, and on the other hand, nucleation should be easier in such a region since the free-energy barriers should be lower and, equally importantly, since the excess mass required for cluster formation is present. This result is very robust in that it is independent of the interaction potential and of the free-energy functional used, while it is valid in the weak-noise limit and could be modified when the effects of strong noise are considered. It is interesting that some of the free-energy-with-constraint methods have given hints of a similar pathway, most notably as discussed in Refs. 12 and 16. It is unclear whether this result is of any practical

significance although one might speculate that the enhancement or suppression of long-wavelength density (concentration) fluctuations could have a strong effect on the nucleation rate.

The theory developed here is specific to the case of Brownian particles. In some sense, however, it can rather be viewed more as an illustration of how a theory can be constructed in other cases. For example, for pure fluids, one still has fluctuating hydrodynamics and the basic ideas can be developed in a similar way. In that case, temperature fluctuations and energy transport are important so that the fluctuating hydrodynamic theory will have to be extended to include the heat equation. The main complication might be that the hydrodynamic theory cannot be collapsed into a single equation for the density, but rather that one must deal will multiple quantities throughout. This may limit the practical utility of the theory but it could still yield useful insights into, e.g., the MLP in the weak noise limit and the role of the free energy.

Future directions for further work include the investigation of the role of strong noise. Also interesting would be the application to heterogeneous nucleation, nucleation in confined systems and the nucleation of solids.

ACKNOWLEDGMENTS

APPENDIX A: ITO AND STRATONOVICH INTERPRETATIONS ARE THE SAME

In general, using the Riemann-Stieltjes definition of integration, stochastic integrals differ from non-stochastic integrals in that their value depends on where the sampling point is chosen within each interval. This means that stochastic differential equations differ according to whether they are interpreted in terms of Ito or Stratonovich interpretations. However, in ordinary fluctuating hydrodynamics, the two interpretations turn out to be the same.[54] Here, it is shown that this is also true for the over-damped dynamical model in both the general and spherically symmetric cases. Note that this does *not* hold for the approximate order-parameter dynamics thus giving rise to the anomalous force discussed in Sec. III B and in Appendix F. Incidentally, it is worth noting that the fact that the two interpretations differ in the latter case is not directly attributable to any approximation: Ito-Stratonovich equivalence can be broken simply by a change in variables.

1. Equivalence for the general model

In general, the Stratonovich SDE

$$\frac{dx_i}{dt} = F_i(\mathbf{a}) + M_{ij}(\mathbf{a})\xi_j(t), \quad \langle \xi_j(t)\xi_l(t')\rangle = \delta_{jl}\delta(t-t')$$

(A1)

corresponds to an Ito SDE of the form

$$\frac{dx_i}{dt} = F_i(\mathbf{a}) + \frac{\partial M_{ij}(\mathbf{a})}{\partial a_k} M_{kj}(\mathbf{a}) + M_{ij}(\mathbf{a})\xi_j(t), \quad \langle \xi_j(t)\xi_l(t')\rangle$$
$$= \delta_{jl}\delta(t-t'),$$

(A2)

so the question comes down to an investigation of the second term on the right in the Ito SDE, the so-called spurious drift. To evaluate this for the hydrodynamic model, the procedure of van Saarloos[54] *et al.* is used whereby the stochastic differential equation is discretized in the spatial variable. Specifically, $\mathbf{r} \mapsto \mathbf{r_n} = \mathbf{n}\Delta$ where $\mathbf{n} = (n_1, n_2, \ldots, n_D)$ are the (integer) coordinates of a lattice point in D dimensions and Δ is the lattice spacing. Also writing $\rho_\mathbf{n} \equiv \rho(\mathbf{r_n})$, the noise term for the over-damped model becomes

$$\left[\nabla \cdot \sqrt{\rho(\mathbf{r})}\xi(\mathbf{r},t)\right]_{\mathbf{r_n}}$$
$$\rightarrow \frac{1}{2\Delta}\sum_{a=1}^{D}(\sqrt{\rho_{\mathbf{n}+\widehat{\mathbf{e}}^{(a)}}}\xi_{a,\mathbf{n}+\widehat{\mathbf{e}}^{(a)}} - \sqrt{\rho_{\mathbf{n}-\widehat{\mathbf{e}}^{(a)}}}\xi_{a,\mathbf{n}-\widehat{\mathbf{e}}^{(a)}}), \quad (A3)$$

where $\widehat{\mathbf{e}}^{(a)}$ is the unit vector in the a-direction (i.e., $\widehat{\mathbf{e}}_i^{(a)} = \delta_{ia}$) and the time-arguments have been suppressed. Introducing the notation $\delta_{\mathbf{n},\mathbf{n}'}$ meaning that all components of the vectors $\mathbf{n}$ and $\mathbf{n}'$ are the same the equivalent of M_{ij} can be identified as

$$M_{\mathbf{n},\mathbf{n}'a} = \frac{1}{2\Delta}(\sqrt{\rho_{\mathbf{n}+\widehat{\mathbf{e}}^{(a)}}}\delta_{\mathbf{n}+\widehat{\mathbf{e}}^{(a)},\mathbf{n}'} - \sqrt{\rho_{\mathbf{n}-\widehat{\mathbf{e}}^{(a)}}}\delta_{\mathbf{n}-\widehat{\mathbf{e}}^{(a)},\mathbf{n}'}),$$
$$= \frac{1}{2\Delta}\sqrt{\rho_{\mathbf{n}'}}(\delta_{\mathbf{n}+\widehat{\mathbf{e}}^{(a)},\mathbf{n}'} - \delta_{\mathbf{n}-\widehat{\mathbf{e}}^{(a)},\mathbf{n}'}). \quad (A4)$$

We, therefore, need

$$\sum_{\mathbf{k}\mathbf{n}'a}\frac{\partial M_{\mathbf{n},\mathbf{n}'a}}{\partial \rho_\mathbf{k}}M_{\mathbf{k},\mathbf{n}'a}$$
$$= \frac{1}{8\Delta^2}\sum_{\mathbf{k}\mathbf{n}'a}\frac{1}{\sqrt{\rho_\mathbf{k}}}\delta_{\mathbf{n}',\mathbf{k}}(\delta_{\mathbf{n}+\widehat{\mathbf{e}}^{(a)},\mathbf{n}'} - \delta_{\mathbf{n}-\widehat{\mathbf{e}}^{(a)},\mathbf{n}'})$$
$$\times \sqrt{\rho_{\mathbf{n}'}}(\delta_{\mathbf{k}+\widehat{\mathbf{e}}^{(a)},\mathbf{n}'} - \delta_{\mathbf{k}-\widehat{\mathbf{e}}^{(a)},\mathbf{n}'})$$
$$= \frac{1}{8\Delta^2}\sum_{\mathbf{k}a}(\delta_{\mathbf{n}+\widehat{\mathbf{e}}^{(a)},\mathbf{k}} - \delta_{\mathbf{n}-\widehat{\mathbf{e}}^{(a)},\mathbf{k}})(\delta_{\mathbf{k}+\widehat{\mathbf{e}}^{(a)},\mathbf{k}} - \delta_{\mathbf{k}-\widehat{\mathbf{e}}^{(a)},\mathbf{k}})$$
$$= 0 \quad (A5)$$

thus proving equivalence of the two interpretations.

2. Equivalence for the spherically symmetric model

In the case of spherical symmetry, one has that

$$\frac{dm(r)}{dt} = 4\pi r^2 D\rho(r)\frac{\partial}{\partial r}\frac{1}{4\pi r^2}\frac{\delta\beta F}{\delta\rho(r)} + \sqrt{D4\pi r^2\rho(r)}\xi(r).$$

(A6)

We discretize as in the previous case but now only require a one-dimensional lattice. Writing $r \mapsto r_n = n\Delta$ and keeping in mind that

$$4\pi r^2\rho(r) = \frac{\partial m(r)}{\partial r}, \quad (A7)$$

one finds the noise term

$$\sqrt{D \frac{m_{n+1} - m_{n-1}}{2\Delta}} \xi_n, \tag{A8}$$

where $m_n = m(r_n; t)$, etc. Thus

$$M_{ij} = \sqrt{D \frac{m_{i+1} - m_{i-1}}{2\Delta}} \delta_{ij} \tag{A9}$$

and so

$$\frac{\partial M_{ij}(\mathbf{a})}{\partial a_k} M_{kj}(\mathbf{a}) = \delta_{ik} \frac{\partial \sqrt{D \frac{m_{i+1}-m_{i-1}}{2\Delta}}}{\partial m_k} \sqrt{D \frac{m_{k+1}-m_{k-1}}{2\Delta}} = 0. \tag{A10}$$

APPENDIX B: MLP FOR POTENTIAL-DRIVEN DYNAMICS WITH FLUCTUATION-DISSIPATION RELATION

The purpose of this Appendix is to sketch a straightforward extension of the results of Vanden-Eijnden and Heymann[55] whereby the assumption of a constant Onsager matrix and white noise is lifted while still assuming a fluctuation-dissipation relation.

In the following, I consider a set of N stochastic variables, $x_i(t)$, governed by a diffusive, gradient-driven stochastic dynamics with multiplicative noise

$$\frac{d\mathbf{x}}{dt} = -\mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x}) + \sqrt{2\epsilon} \sigma(\mathbf{x}) \cdot \xi(t), \tag{B1}$$

where $\mathbf{L}(\mathbf{x})$ is a state-dependent matrix of kinetic coefficients, where the scalar constant ϵ and matrix $\sigma(\mathbf{x})$ determine the noise amplitude and the noise itself is Gaussian, white and diagonally correlated

$$\langle \xi_i(t) \xi_j(t') \rangle = \delta_{ij} \delta(t - t'). \tag{B2}$$

Note that the probability density for $\xi(t)$ to assume some value, say $\mathbf{z}$, is

$$P(\xi(t) = \mathbf{z}) = \left(\frac{1}{2\pi}\right)^{N/2} \exp(-z^2/2). \tag{B3}$$

The key assumption in the following is that a fluctuation-dissipation relation (FDR) holds, namely,

$$\mathbf{L}(\mathbf{x}) = \sigma(\mathbf{x}) \cdot \sigma^{\mathrm{T}}(\mathbf{x}) \equiv \mathbf{D}(\mathbf{x}), \tag{B4}$$

where the second equality reminds that the middle quantity is the diffusion matrix occurring in the Fokker-Planck equation.

Following Refs. 56 and 57 the probability density for a given path taking place from time $t = 0$ to $t = T$ is $P = \exp(-\frac{1}{4\epsilon} S_T[x])$, where the action is

$$S_T[x] = \int_0^T \left(\frac{d\mathbf{x}}{dt} + \mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x})\right) \cdot \mathbf{D}^{-1}(\mathbf{x})$$
$$\cdot \left(\frac{d\mathbf{x}}{dt} + \mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x})\right) dt \tag{B5}$$

and it must be remembered that $\mathbf{x}$ depends on time in this and all following expression. A path between two points $\mathbf{x}_1$ and $\mathbf{x}_2$ is therefore a curve $\mathbf{x}(t)$ such that $\mathbf{x}(0) = \mathbf{x}_1$ and $\mathbf{x}(T) = \mathbf{x}_2$.

The MLP is determined by minimizing $S_T[x]$ over both $\mathbf{x}(t)$, subject to the constraints on the end points, and time, T.

In the following, we specialize to the situation that $\mathbf{x}_1$ and $\mathbf{x}_2$ are metastable points and in fact attractors. We assume that x-space is divided by a separatrix into two regions: region I in which points are attracted to $\mathbf{x}_1$ and region II in which points are attracted to $\mathbf{x}_2$. The separatrix is a curve which will be called $\mathcal{S}$. Any path from $\mathbf{x}_1$ to $\mathbf{x}_2$ must cross $\mathcal{S}$ at least once. For the moment, it will be assumed that only one such crossing occurs and the possibility of multiple crossings will be discussed below. Any such path can, therefore, be separated into two pieces: one running from $\mathbf{x}_1$ to some point $\mathbf{x}_s \in \mathcal{S}$ and a second part from $\mathbf{x}_s$ to $\mathbf{x}_2$,

$$S_T[x] = S^I[x] + S^{II}[x],$$
$$S^I[x] = \int_0^{T_s} \left(\frac{d\mathbf{x}}{dt} + \mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x})\right) \cdot \mathbf{D}^{-1}(\mathbf{x})$$
$$\cdot \left(\frac{d\mathbf{x}}{dt} + \mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x})\right) dt,$$
$$S^{II}[x] = \int_{T_s}^T \left(\frac{d\mathbf{x}}{dt} + \mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x})\right) \cdot \mathbf{D}^{-1}(\mathbf{x})$$
$$\cdot \left(\frac{d\mathbf{x}}{dt} + \mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x})\right) dt \tag{B6}$$

with $\mathbf{x}(T_s) = \mathbf{x}_s$. Clearly, the MLP is determined by minimizing over $\mathbf{x}_s \in \mathcal{S}$ and $0 \leq T_s \leq T$ as well. Consider the second term first. Once the separatrix is crossed all points are attracted to $\mathbf{x}_2$ by hypothesis, so the path

$$\frac{d\mathbf{x}}{dt} = -\mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x}),$$
$$\mathbf{x}(T_s) = \mathbf{x}_s \tag{B7}$$

will eventually reach $\mathbf{x}_2$: the time required determines T given T_s. (Note that one does not really start on the separatrix but rather at a point infinitesimally near it on the region II side.) However, this path has the property that $S^{II}[x] = 0$ and this is minimal since the integrand is positive definite. This is just the trivial result that the deterministic path is the MLP, if it passes through the desired points.

It is not possible to take the deterministic path in region I since the system must go from $\mathbf{x}_1$ to $\mathbf{x}_s$ and the deterministic dynamics is assumed to always take points in region I towards $\mathbf{x}_1$. So, upon noting that expanding the action gives

$$S^I[x] = \int_0^{T_s} \left(\frac{d\mathbf{x}}{dt} \cdot \mathbf{D}^{-1}(\mathbf{x}) \cdot \frac{d\mathbf{x}}{dt} + \left(\mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x})\right)\right.$$
$$\left. \cdot \mathbf{D}^{-1}(\mathbf{x}) \cdot \left(\mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x})\right)\right) dt$$
$$+ 2 \int_0^{T_s} \frac{d\mathbf{x}}{dt} \cdot \mathbf{D}^{-1}(\mathbf{x}) \cdot \mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x}) dt. \tag{B8}$$

Invoking the FDR, the last term is

$$\int_0^{T_s} \frac{d\mathbf{x}}{dt} \cdot \mathbf{D}^{-1}(\mathbf{x}) \cdot \mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x}) \, dt = \int_0^{T_s} \frac{d\mathbf{x}}{dt} \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x}) \, dt,$$

$$= \int_0^{T_s} \frac{d}{dt} V(\mathbf{x}) \, dt,$$

$$= V(\mathbf{x}_s) - V(\mathbf{x}_1) \quad \text{(B9)}$$

One can, therefore, write

$$S^I[x] = \int_0^{T_s} \left(\frac{d\mathbf{x}}{dt} \cdot \mathbf{D}^{-1}(\mathbf{x}) \cdot \frac{d\mathbf{x}}{dt} + \left(\mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x}) \right) \right.$$

$$\left. \cdot \mathbf{D}^{-1}(\mathbf{x}) \cdot \left(\mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x}) \right) \right) dt$$

$$+ 2V(\mathbf{x}_s) - V(\mathbf{x}_1),$$

$$= \int_0^{T_s} \left(\frac{d\mathbf{x}}{dt} + \mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x}) \right)$$

$$\cdot \mathbf{D}^{-1}(\mathbf{x}) \cdot \left(\frac{d\mathbf{x}}{dt} + \mathbf{L}(\mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} V(\mathbf{x}) \right) dt$$

$$+ 4\left(V(\mathbf{x}_s) - V(\mathbf{x}_1)\right). \quad \text{(B10)}$$

Notice the change in sign of the gradient term. Reversing the sign of the integration variable and introducing $\mathbf{y}(t) = \mathbf{x}(-t)$, so that $\mathbf{y}(0) = \mathbf{x}_0$ and $\mathbf{y}(-T_s) = \mathbf{x}_1$ the action can be written as

$$S^I[x] = \int_{-T_s}^0 \left(\frac{d\mathbf{y}}{dt'} - \mathbf{L}(\mathbf{y}) \cdot \frac{\partial}{\partial \mathbf{y}} V(\mathbf{y}) \right) \cdot \mathbf{D}^{-1}(\mathbf{y})$$

$$\cdot \left(\frac{d\mathbf{y}}{dt'} - \mathbf{L}(\mathbf{y}) \cdot \frac{\partial}{\partial \mathbf{y}} V(\mathbf{y}) \right) dt' + 4\left(V(\mathbf{x}_s) - V(\mathbf{x}_1)\right).$$

$$\text{(B11)}$$

The integral is now the action for a path going from $\mathbf{x}_1$ at $t' = -T_s$ to $\mathbf{x}_0$ at $t' = 0$ so that the deterministic path can again be used to set the integral to zero and this will also determine T_s. One, therefore, gets that

$$S \geq 4\left(V(\mathbf{x}_s) - V(\mathbf{x}_1)\right) \quad \text{(B12)}$$

with equality if the system follows the deterministic paths connecting $\mathbf{x}_s$ to the end points $\mathbf{x}_1$ and $\mathbf{x}_2$. Finally, the action is minimized by choosing $\mathbf{x}_s$ to be the minimal value on the separatrix which is just the saddle point.

Note that recrossing the separatrix will involve a deviation from the deterministic path on both sides of the separatrix and so will not minimize the action. For this reason, one need only consider a single crossing of the separatrix.

This serves to establish the claim that for the dynamics given by Eq. (B1) with the fluctuation-dissipation relation, Eq. (B4), the MLP crosses the separatrix at the critical point and follows the deterministic path away from the critical point. Note the key role played by the FDR in this process: the same result will not necessarily hold in the general case.

APPENDIX C: EXACT ACTION WITH ORDER PARAMETERS

Expanding the Lagrangian in Eq. (52) gives

$$\mathcal{L} = \frac{1}{2} \int_0^\infty \frac{1}{r^2 \rho(r)} \left(\frac{\partial m(r)}{\partial t} \right)^2 dr$$

$$- D \int_0^\infty \left(\frac{\partial m(r)}{\partial t} \frac{\partial}{\partial r} \frac{1}{r^2} \frac{\delta F[\rho]}{\delta \rho(r)} \right) dr$$

$$+ \frac{1}{2} D^2 \int_0^\infty r^2 \rho(r) \left(\frac{\partial}{\partial r} \frac{1}{r^2} \frac{\delta F[\rho]}{\delta \rho(r)} \right)^2 dr. \quad \text{(C1)}$$

If the density is parametrized as $\rho(r; t) = \rho(r; \mathbf{x}(t))$, then the first term becomes

$$\frac{1}{2} \int_0^\infty \frac{1}{r^2 \rho(r)} \left(\frac{\partial m(r)}{\partial t} \right)^2 dr$$

$$= \left[\frac{1}{2} \int_0^\infty \frac{1}{r^2 \rho(r;x)} \frac{\partial m(r)}{\partial x_a} \frac{\partial m(r)}{\partial x_b} dr \right] \frac{dx_a}{dt} \frac{dx_b}{dt}$$

$$= 2\pi g_{ab}(\mathbf{x}) \frac{dx_a}{dt} \frac{dx_b}{dt} \quad \text{(C2)}$$

and the second is

$$- D \int_0^\infty \left(\frac{\partial m(r)}{\partial t} \frac{\partial}{\partial r} \frac{1}{r^2} \frac{\delta F[\rho]}{\delta \rho(r)} \right) dr$$

$$= - D \int_0^\infty \frac{\partial}{\partial r} \left(\frac{\partial m(r)}{\partial t} 4\pi \left. \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho(r)} \right) dr$$

$$+ D \int_0^\infty \left(\frac{\partial}{\partial r} \frac{\partial m(r)}{\partial t} \right) 4\pi \left. \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho(r)} dr$$

$$= -4\pi D \frac{dx_a}{dt} \lim_{r \to \infty} \left(\frac{\partial m(r)}{\partial x_a} \left. \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho(r)} \right)$$

$$+ 4\pi D \frac{dx_a}{dt} \int_0^\infty 4\pi r^2 \frac{\partial \rho(r)}{\partial x_a} \left. \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho(r)} dr$$

$$= -4\pi D \frac{dx_a}{dt} \frac{\partial N}{\partial x_a} \mu + 4\pi D \frac{dx_a}{dt} \frac{\partial F[\rho]}{\partial x_a}. \quad \text{(C3)}$$

To evaluate this, note that $\lim_{r \to \infty} m(r) = N$, the total number of particles in the system. Assuming that the boundary condition is that the density assume the bulk value, ρ_∞, far from the interface gives

$$\lim_{r \to \infty} \left(\frac{\partial m(r)}{\partial x_a} \left. \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho(r)} \right) = \frac{\partial N}{\partial x_a} \left. \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho_\infty} \equiv \frac{\partial N}{\partial x_a} \mu,$$

$$\text{(C4)}$$

where the last equality defines the chemical potential. Combining, the final result is

$$- D \int_0^\infty \left(\frac{\partial m(r)}{\partial t} \frac{\partial}{\partial r} \frac{\delta F[\rho]}{\delta \rho(r)} \right) dr = 4\pi D \frac{dx_a}{dt} \frac{\partial \Omega[\rho]}{\partial x_a},$$

$$\text{(C5)}$$

where $\Omega[\rho] = F[\rho] - \mu N$ is the grand potential. Thus, the Lagrangian is

$$\frac{1}{4\pi} \mathcal{L} = \frac{1}{2} g_{ab}(\mathbf{x}) \frac{dx_a}{dt} \frac{dx_b}{dt} + D \frac{dx_a}{dt} \frac{\partial \beta \Omega[\rho]}{\partial x_a} + V(\mathbf{x}) \quad \text{(C6)}$$

with

$$V(\mathbf{x}) = \frac{1}{2} D^2 \int_0^\infty r^2 \rho(r) \left(\frac{\partial}{\partial r} \frac{1}{r^2} \frac{\delta F[\rho]}{\delta \rho(r)} \right)^2 dr. \qquad (C7)$$

APPENDIX D: EQ. (80) AS A COMPLETENESS RELATION

Assume that $\{p_i(r)\}_{i=1}^\infty$ is a complete set of basis functions so that one can write

$$m(r) = \sum_{i=1}^\infty x_i p_i(r),$$
$$(D1)$$
$$x_i = \int_0^\infty m(r) q_i(r) \, dr,$$

where $\{q_i(r)\}_{i=1}^\infty$ is the bi-orthogonal set satisfying

$$\int_0^\infty p_i(r) q_j(r) \, dr = \delta_{ij}. \qquad (D2)$$

Then, since $m(r)$ is arbitrary, one has the completeness relation

$$\sum_i p_i(r) q_i(r') = \delta(r - r'). \qquad (D3)$$

First note that the metric is

$$g_{ij} = \int_0^\infty r^{-2} \rho^{-1}(r) \frac{\partial m(r)}{\partial x_i} \frac{\partial m(r)}{\partial x_j} dr$$
$$= \int_0^\infty r^{-2} \rho^{-1}(r) p_i(r) p_j(r) \, dr. \qquad (D4)$$

The inverse metric is then

$$g_{jl}^{-1} = \int_0^\infty r'^2 \rho(r') \frac{\delta x_j}{\delta m(r')} \frac{\delta x_l}{\delta m(r')} dr'$$
$$= \int_0^\infty r'^2 \rho(r') q_j(r') q_l(r') \, dr' \qquad (D5)$$

as can be verified by direct evaluation of

$$\sum_j g_{ij} \int_0^\infty r'^2 \rho(r') q_j(r') q_l(r') dr'$$
$$= \sum_j \int_0^\infty dr \int_0^\infty dr' \, (r^{-2} \rho^{-1}(r) p_i(r) p_j(r))$$
$$\times (r'^2 \rho(r') q_j(r') q_l(r'))$$
$$= \int_0^\infty dr \int_0^\infty dr' \, (r^{-2} \rho^{-1}(r) p_i(r))$$
$$\times \delta(r - r')(r'^2 \rho(r') q_l(r'))$$
$$= \int_0^\infty dr \, p_i(r) q_l(r')$$
$$= \delta_{il}, \qquad (D6)$$

where the second line follows from the completeness relation and the last line by bi-orthogonality. Then, one easily verifies

that

$$\sum_{ij} \frac{\partial m(r)}{\partial x_i} g_{ij}^{-1} \frac{\partial m(r')}{\partial x_j}$$
$$= \sum_{ij} p_i(r) \left(\int_0^\infty r''^2 \rho(r'') q_i(r'') q_j(r'') dr'' \right) p_j(r')$$
$$= \int_0^\infty r''^2 \rho(r'') \delta(r - r'') \delta(r' - r'') dr''$$
$$= r^2 \rho(r) \delta(r - r') \qquad (D7)$$

as claimed.

APPENDIX E: ALTERNATIVE DERIVATION OF ORDER PARAMETER EQUATIONS

One way to relate a general functional form, $\widetilde{\rho}(r; \mathbf{x}(t))$, to the exact density, $\rho(r; t)$, is by defining a distance in density space and then minimizing the difference between the two functions. It is shown in the main text that the dynamics imposes a unique distance functional which is expressed in terms of the corresponding mass functionals as

$$s^2 = \int_0^\infty \left(\widetilde{m}^{-1}(u; \mathbf{x}(t)) - m^{-1}(u; t) \right)^2 du. \qquad (E1)$$

Determining the fitting parameters at each moment in time by minimizing the difference between the actual mass distribution and the approximating distribution gives

$$0 = \frac{\partial}{\partial x_i(t)} \int_0^\infty \left(\widetilde{m}^{-1}(u; \mathbf{x}(t)) - m^{-1}(u; t) \right)^2 du,$$
$$= \int_0^\infty \left(\widetilde{m}^{-1}(u; \mathbf{x}(t)) - m^{-1}(u; t) \right) \frac{\partial \widetilde{m}^{-1}(u; \mathbf{x}(t))}{\partial x_i(t)} du,$$
$$= \int_0^\infty \left(\widetilde{m}^{-1}(m(r; t); \mathbf{x}(t)) - r \right)$$
$$\times \frac{\partial \widetilde{m}^{-1}(u; \mathbf{x}(t))}{\partial x_i(t)} \frac{\partial m(r; t)}{\partial r} dr. \qquad (E2)$$

In the following, in order to simplify the notation, the dependence on time will not be explicitly indicated.

Noting that if

$$y(m; \mathbf{x}) = \widetilde{m}^{-1}(m; \mathbf{x}), \qquad (E3)$$

then

$$m = \widetilde{m}(y(m; \mathbf{x}); \mathbf{x}) \qquad (E4)$$

so

$$0 = \widetilde{m}_r(y(m, \mathbf{x}); \mathbf{x}) y_i(m; \mathbf{x}) + \widetilde{m}_i(y(m; \mathbf{x}); \mathbf{x}), \qquad (E5)$$

where the short-hand notation

$$\widetilde{m}_r(y; \mathbf{x}) \equiv \frac{\partial \widetilde{m}(r; \mathbf{x})}{\partial r} \bigg|_{r=y}$$
$$(E6)$$
$$\widetilde{m}_i(y; \mathbf{x}) \equiv \frac{\partial \widetilde{m}(y; \mathbf{x})}{\partial x_i}$$

is used. Thus, Eq. (E2) becomes

$$0 = \int_0^\infty (\tilde{m}^{-1}(m(r);\mathbf{x}) - r)$$
$$\times \frac{\tilde{m}_i(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x})}{\tilde{m}_r(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x})} \frac{\partial m(r)}{\partial r} dr. \qquad \text{(E7)}$$

Another derivative gives

$$0 = -\frac{dx_j}{dt} \int_0^\infty \frac{\tilde{m}_j\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)}{\tilde{m}_r\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)}$$
$$\times \frac{\tilde{m}_i\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)}{\tilde{m}_r\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)} \frac{\partial m(r)}{\partial r} dr$$
$$+ \int_0^\infty \tilde{m}_r^{-1}(m(r);\mathbf{x}) \frac{dm(r)}{dt}$$
$$\times \frac{\tilde{m}_i\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)}{\tilde{m}_r\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)} \frac{\partial m(r)}{\partial r} dr$$
$$+ \frac{dx_j}{dt} \int_0^\infty (\tilde{m}^{-1}(m(r);\mathbf{x}) - r) \frac{d}{dt}$$
$$\times \left[\frac{\tilde{m}_i\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)}{\tilde{m}_r\left(\tilde{m}^{-1}(m(r),\mathbf{x});\mathbf{x}\right)} \frac{\partial m(r)}{\partial r} \right] dr \qquad \text{(E8)}$$

or

$$0 = -\frac{dx_j}{dt} \int_0^\infty \frac{\tilde{m}_j\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)}{\tilde{m}_r\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)}$$
$$\times \frac{\tilde{m}_i\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)}{\tilde{m}_r\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)} \frac{\partial m(r)}{\partial r} dr$$
$$+ \int_0^\infty \frac{1}{\tilde{m}_r\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)} \frac{dm(r)}{dt}$$
$$\times \frac{\tilde{m}_i\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)}{\tilde{m}_r\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)} \frac{\partial m(r)}{\partial r} dr$$
$$+ \frac{dx_j}{dt} \int_0^\infty (\tilde{m}^{-1}(m(r);\mathbf{x}) - r) \frac{d}{dt}$$
$$\times \left[\frac{\tilde{m}_i\left(\tilde{m}^{-1}(m(r);\mathbf{x});\mathbf{x}\right)}{\tilde{m}_r\left(\tilde{m}^{-1}(m(r),\mathbf{x});\mathbf{x}\right)} \frac{\partial m(r)}{\partial r} \right] dr. \qquad \text{(E9)}$$

Expanding in the difference $m(r) - \tilde{m}(r;\mathbf{x})$ gives

$$g_{ij}(\mathbf{x}(t)) \frac{dx_i(t)}{dt} = \int_0^\infty \frac{\partial \tilde{m}(r;\mathbf{x}(t))}{\partial x_i(t)} \left(\frac{\partial m(r;t)}{\partial r} \right)^{-1}$$
$$\times \frac{dm(r;t)}{dt} dr + O(m(r;t) - \tilde{m}(r;\mathbf{x}(t))), \qquad \text{(E10)}$$

which gives the result, Eq. (73).

APPENDIX F: THE SPURIOUS DRIFT

The stochastic model is

$$\frac{dx_i}{dt} = -Dg_{ij}^{-1}(\mathbf{x}) \frac{\partial \beta \Omega}{\partial x_j} - \epsilon g_{ij}^{-1}(\mathbf{x})$$
$$\times \int_0^\infty \sqrt{\frac{2D}{4\pi r^2 \rho(r;\mathbf{x})}} \frac{\partial m(r;\mathbf{x})}{\partial x_j} \xi(r;t) dr, \qquad \text{(F1)}$$

which must be understood in the Stratonovich interpretation. In order to replace the noise term by a simpler form, this must first be written as an Ito SDE. Then, the change of the noise can be made and the result transformed back to the Stratonovich interpretation. The first step is accomplished using the standard transformation rule[23] giving the equivalent Ito form of the SDE,

$$\frac{dx_i}{dt} = -Dg_{ij}^{-1}(\mathbf{x}) \frac{\partial \beta \Omega}{\partial x_j}$$
$$+ \frac{1}{2}\epsilon^2 \int_0^\infty \left(\frac{\partial}{\partial x_l} g_{ij}^{-1}(\mathbf{x}) \sqrt{\frac{2D}{4\pi r^2 \rho(r;\mathbf{x})}} \frac{\partial m(r;\mathbf{x})}{\partial x_j} \right)$$
$$\times \left(g_{lk}^{-1}(\mathbf{x}) \sqrt{\frac{2D}{4\pi r^2 \rho(r;\mathbf{x})}} \frac{\partial m(r;\mathbf{x})}{\partial x_k} \right) dr$$
$$- \epsilon g_{ij}^{-1}(\mathbf{x}) \int_0^\infty \sqrt{\frac{2D}{4\pi r^2 \rho(r;\mathbf{x})}} \frac{\partial m(r;\mathbf{x})}{\partial x_j} \xi(r;t) dr. \qquad \text{(F2)}$$

The autocorrelation of the noise is

$$D_{il} \equiv \epsilon^2 g_{ij}^{-1} g_{lm}^{-1} \int_0^\infty \frac{2D}{4\pi r^2 \rho(r;\mathbf{x})} \frac{\partial m(r;\mathbf{x})}{\partial x_j} \frac{\partial m(r;\mathbf{x})}{\partial x_m} dr \qquad \text{(F3)}$$

$$= 2D\epsilon^2 g_{il}^{-1}, \qquad \text{(F4)}$$

so that the noise term in Eq. (F2) can be replaced by the equivalent form $\epsilon\sqrt{2D}q_{ij}^{-1}(\mathbf{x})\xi_j(t)$ with $q_{il}(\mathbf{x})q_{jl}(\mathbf{x}) = g_{ij}(\mathbf{x})$ and $\langle \xi_i(t)\xi_j(t') \rangle = \delta_{ij}\delta(t-t')$.

Up to a factor of D, the second term on the right is

$$A_i^{(I)} = \frac{\partial g_{ij}^{-1}(\mathbf{x})}{\partial x_j}$$
$$+ g_{ij}^{-1}(\mathbf{x}) g_{lk}^{-1}(\mathbf{x}) \int_0^\infty \frac{1}{4\pi r^2 \rho(r;\mathbf{x})} \left(\frac{\partial^2 m(r;\mathbf{x})}{\partial x_l \partial x_i} \right.$$
$$\left. - \frac{1}{2} \frac{1}{\rho(r;\mathbf{x})} \frac{\partial \rho(r;\mathbf{x})}{\partial x_l} \frac{\partial m(r;\mathbf{x})}{\partial x_i} \right) \frac{\partial m(r;\mathbf{x})}{\partial x_k} dr. \qquad \text{(F5)}$$

Using the symmetry of g_{ij}^{-1}, one has that

$$g_{ij}^{-1}(\mathbf{x})\, g_{lk}^{-1}(\mathbf{x}) \int_0^\infty \frac{1}{4\pi r^2 \rho(r;\mathbf{x})} \frac{\partial^2 m(r;\mathbf{x})}{\partial x_l \partial x_i} \frac{\partial m(r;\mathbf{x})}{\partial x_k} dr$$

$$= \frac{1}{2} g_{ij}^{-1}(\mathbf{x})\, g_{lk}^{-1}(\mathbf{x}) \int_0^\infty \frac{1}{4\pi r^2 \rho(r;\mathbf{x})} \left(\frac{\partial^2 m(r;\mathbf{x})}{\partial x_l \partial x_i} \right.$$

$$\times \frac{\partial m(r;\mathbf{x})}{\partial x_k} + \frac{\partial^2 m(r;\mathbf{x})}{\partial x_k \partial x_i} \frac{\partial m(r;\mathbf{x})}{\partial x_l} \bigg) dr$$

$$= \frac{1}{2} g_{ij}^{-1}(\mathbf{x})\, g_{lk}^{-1}(\mathbf{x}) \int_0^\infty \frac{1}{4\pi r^2 \rho(r;\mathbf{x})} \frac{\partial}{\partial x_i}$$

$$\times \left(\frac{\partial m(r;\mathbf{x})}{\partial x_l} \frac{\partial m(r;\mathbf{x})}{\partial x_k} \right) dr$$

$$= \frac{1}{2} g_{ij}^{-1}(\mathbf{x})\, g_{lk}^{-1}(\mathbf{x}) \left[\frac{\partial g_{lk}(\mathbf{x})}{\partial x_i} \right.$$

$$+ \int_0^\infty \frac{1}{4\pi r^2 \rho^2(r;\mathbf{x})} \frac{\partial \rho(r;\mathbf{x})}{\partial x_i} \frac{\partial m(r;\mathbf{x})}{\partial x_l} \frac{\partial m(r;\mathbf{x})}{\partial x_k} dr \bigg]$$

$$\tag{F6}$$

so the extra term becomes

$$A_i^{(I)} = \frac{\partial g_{ij}^{-1}(\mathbf{x})}{\partial x_j} + \frac{1}{2} g_{ij}^{-1}(\mathbf{x})\, g_{lk}^{-1}(\mathbf{x}) \frac{\partial g_{lk}(\mathbf{x})}{\partial x_i}$$

$$+ \frac{1}{2} g_{ij}^{-1}(\mathbf{x})\, g_{lk}^{-1}(\mathbf{x}) \int_0^\infty \frac{1}{4\pi r^2 f^2(r;\mathbf{x})}$$

$$\times \left(\frac{\partial \rho(r;\mathbf{x})}{\partial x_i} \frac{\partial m(r;\mathbf{x})}{\partial x_l} - \frac{\partial \rho(r;\mathbf{x})}{\partial x_l} \frac{\partial m(r;\mathbf{x})}{\partial x_i} \right)$$

$$\times \frac{\partial m(r;\mathbf{x})}{\partial x_k} dr$$

$$= \frac{\partial g_{ij}^{-1}(\mathbf{x})}{\partial x_j} + \frac{1}{2} g_{ij}^{-1}(\mathbf{x})\, g_{lk}^{-1}(\mathbf{x}) \frac{\partial g_{lk}(\mathbf{x})}{\partial x_i}$$

$$+ \frac{1}{2} \left(g_{ij}^{-1}(\mathbf{x})\, g_{lk}^{-1}(\mathbf{x}) - g_{lj}^{-1}(\mathbf{x})\, g_{ik}^{-1}(\mathbf{x}) \right)$$

$$\times \int_0^\infty \frac{1}{4\pi r^2 \rho^2(r;\mathbf{x})} \frac{\partial \rho(r;\mathbf{x})}{\partial x_i} \frac{\partial m(r;\mathbf{x})}{\partial x_l} \frac{\partial m(r;\mathbf{x})}{\partial x_k} dr.$$

$$\tag{F7}$$

Combining these results gives the equivalent Ito form of the SDE with simplified noise term

$$\frac{dx_i}{dt} = -2D g_{ij}^{-1}(\mathbf{x}) \frac{\partial \beta\Omega}{\partial x_j} - D\epsilon^2 A_i^{(I)} - \epsilon\sqrt{2D}\, q_{lj}^{-1}(\mathbf{x})\, \xi_l(t).$$

$$\tag{F8}$$

To get the Stratonovich form, one must transform back giving the same equation but with spurious drift

$$A_i^{(S)} = A_i^{(I)} - q_{kj}^{-1} \frac{\partial q_{ij}^{-1}}{\partial x_k}.$$

$$\tag{F9}$$

[1]D. Kashchiev, *Nucleation: Basic Theory With Applications* (Butterworths, Oxford, 2000).
[2]R. Becker and W. Döring, Ann. Phys. **2416**, 719 (1935).
[3]L. Gránásy, AIP Conf. Proc. **534**, 209 (2000).
[4]R. Evans, Adv. Phys. **28**, 143 (1979).
[5]J.-P. Hansen and I. McDonald, *Theory of Simple Liquids* (Academic Press, San Diego, 1986).
[6]B. Peters, J. Chem. Phys. **135**, 044107 (2011).
[7]J. F. Lutsko, Adv. Chem. Phys. **144**, 1 (2010).
[8]D. W. Oxtoby and R. Evans, J. Chem. Phys. **89**, 7521 (1988).
[9]V. Talanquer and D. W. Oxtoby, J. Chem. Phys. **100**, 5190 (1994).
[10]N. D. Mermin, Phys. Rev. **137**, A1441 (1965).
[11]J. F. Lutsko, J. Chem. Phys. **129**, 244501 (2008).
[12]J. F. Lutsko, J. Chem. Phys. **134**, 164501 (2011).
[13]S. Ghosh and S. K. Ghosh, J. Chem. Phys. **134**, 024502 (2011).
[14]M. J. Uline and D. S. Corti, Phys. Rev. Lett. **99**, 076102 (2007).
[15]M. J. Uline, K. Torabi, and D. S. Corti, J. Chem. Phys. **133**, 174511 (2010); M. J. Uline, K. Torabi, and D. S. Corti, **133**, 174512 (2010).
[16]A. J. Archer and R. Evans, Mol. Phys. **00**, 1 (2011).
[17]J. F. Lutsko, Europhys. Lett. **83**, 46007 (2008).
[18]C. Qiu and T. Qian, J. Chem. Phys. **131**, 124708+ (2009).
[19]M. Iwamatsu, J. Chem. Phys. **130**, 244507 (2009).
[20]T. Philippe and D. Blavette, J. Chem. Phys. **135**, 134508 (2011).
[21]J. F. Lutsko, e-print arXiv:cond-mat/1104.3413.
[22]J. F. Lutsko, J. Chem. Phys. **135**, 161101 (2011).
[23]C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer, Berlin, 2004).
[24]D. S. Dean, J. Phys. A **29**, L613 (1996).
[25]J. A. McLennan, *Introduction to Nonequilibrium Statistical Mechanics* (Prentice-Hall, Englewood Cliffs, NJ, 1989).
[26]A. J. Archer and R. Evans, J. Chem. Phys. **121**, 4246 (2004).
[27]A. J. Archer, J. Chem. Phys. **130**, 014509 (2009).
[28]P.-H. Chavanis, Physica A **387**, 5716 (2008).
[29]K. Kawasaki, J. Stat. Phys. **93**, 527 (1998).
[30]P.-H. Chavanis, Physica A **390**, 1546 (2011).
[31]L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).
[32]R. Graham, Z. Phys. B **26**, 281 (1977).
[33]M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, New York, 1998).
[34]S. R.S. Varadhan, Ann. Probab. **36**, 397 (2008).
[35]D. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, England, 2003).
[36]M. J. Robbins, A. J. Archer, and U. Thiele, J. Phys.: Condens. Matter **23**, 415102 (2011).
[37]E. Weinan, W. Ren, and E. Vanden-Eijnden, J. Chem. Phys. **126**, 164103 (2007).
[38]G. Henkelman and H. Jónsson, J. Chem. Phys. **113**, 9978 (2000).
[39]H. Jonsson, G. Mills, and G. K. Schenter, Surf. Sci. **324**, 305 (1995).
[40]P. R. ten Wolde and D. Frenkel, Science **77**, 1975 (1997).
[41]J. D. Gunton, A. Shiryayev, and D. L. Pagan, *Protein Condensation: Kinetic Pathways to Crystallization and Disease* (Cambridge University Press, Cambridge, England, 2007).
[42]J. F. Lutsko and G. Nicolis, J. Chem. Phys. **122**, 244907 (2005).
[43]P. G. Vekilov, Cryst. Growth Des. **4**, 671 (2004).
[44]J. A. Barker and D. Henderson, J. Chem. Phys. **47**, 4714 (1967).
[45]P. Grosfils and J. F. Lutsko, Langmuir **26**, 8510 (2010).
[46]M. M.T. da Gama and R. Evans, Mol. Phys. **38**, 367 (1979).
[47]J. C. Mauro, R. J. Loucks, and J. Balakrishnan, J. Phys. Chem. A **109**, 9578 (2005).
[48]W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Clarendon, Oxford, 1993).
[49]See http://software.intel.com/en-us/articles/intel-ordinary-differential-equ-http://ations-solver-library/ for use of the Intel ODE solver.
[50]M. Bier, I. Derényi, M. Kostur, and R. D. Astumian, Phys. Rev. E **59**, 6422 (1999).
[51]E. Lifshitz and L. Pitaevskii, *Physical Kinetics: Course of Theoretical Physics* (Elsevier, Amsterdam, 2010), Vol. 10.
[52]Y. Saito, *Statistical Physics of Crystal Growth* (World Scientific, Singapore, 1998).
[53]L. O. Hedges and S. Whitelam, e-print arXiv:cond-mat/1103.0334.
[54]W. van Saarloos, D. Bedeaux, and P. Mazur, Physica A **110**, 147 (1982).
[55]E. Vanden-Eijnden and M. Heymann, J. Chem. Phys. **128**, 061103 (2008).
[56]M. Heymann and E. Vanden-Eijnden, Phys. Rev. Lett. **100**, 140601 (2008).
[57]M. Heymann and E. Vanden-Eijnden, Commun. Pure Appl. Math. **61**, 1052 (2008).