Classical nucleation theory from a dynamical approach to nucleation

James F. Lutsko and Miguel A. Durán-Olivencia

Additional information on J. Chem. Phys.

Classical nucleation theory from a dynamical approach to nucleation

James F. Lutsko[1,a)] and Miguel A. Durán-Olivencia[2,b)]

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

[2] *Laboratorio de Estudios Cristalográficos, Instituto Andaluz de Ciencias de la Tierra, CSIC-UGR, Avda. de las Palmeras, 18100 Granada, Spain*

It is shown that diffusion-limited classical nucleation theory (CNT) can be recovered as a simple limit of the recently proposed dynamical theory of nucleation based on fluctuating hydrodynamics [J. F. Lutsko, J. Chem. Phys. **136**, 034509 (2012)]. The same framework is also used to construct a more realistic theory in which clusters have finite interfacial width. When applied to the dilute solution/dense solution transition in globular proteins, it is found that the extension gives corrections to the nucleation rate even for the case of small supersaturations due to changes in the monomer distribution function and to the excess free energy. It is also found that the monomer attachment/detachment picture breaks down at high supersaturations corresponding to clusters smaller than about 100 molecules. The results also confirm the usual assumption that most important corrections to CNT can be achieved by means of improved estimates of the free energy barrier. The theory also illustrates two topics that have received considerable attention in the recent literature on nucleation: the importance sub-dominant corrections to the capillary model for the free energy and of the correct choice of the reaction coordinate. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4811490]

I. INTRODUCTION

The starting point for any discussion of nucleation is the set of ideas collectively known as Classical Nucleation Theory (CNT).[1] The aim of CNT is to describe the growth and dissipation of clusters of a new phase forming within a bath of the mother phase. (This paper will primarily address the problem of homogeneous nucleation although the ideas discussed here, and indeed those of CNT, are easily extended to the case of heterogeneous nucleation.) In CNT, the central quantity is the concentration, c_N, of clusters of the new phase of a given size, N. (Note that since we will limit the present discussion to phase transitions involving a single species, the number of molecules in a cluster and the mass of the cluster are equivalent and the terms will be used interchangeably along with the generic term "size".) The basic idea underlying CNT is to write equations for the time evolution of the cluster concentrations which are analogous in form to chemical rate equations (the Becker-Döring equation[1,2]). The physics of the problem enters via the rate coefficients which depend on the properties of the cluster (e.g., its size) as well as on those of the mother phase (e.g., its pressure). These equations are then used to extract a nucleation rate under the particular assumption of a quasi-stationary distribution in which there is a constant concentration of monomers and a steady production of clusters up to the of the critical cluster: critical clusters are removed and new monomers added thus allowing for the establishment of a steady state. This process is necessary since the system is, by definition, not in an equilibrium state and so the concept of an equilibrium distribution is not available.

One important assumption that is made in CNT, even at this level, is that the concept of a "cluster" is well defined. Of course, one can always devise a definition of what constitutes a cluster. For example, in their simulation studies of nucleation, ten Wolde and Frenkel have defined a liquid cluster forming in a vapor to be the collection of molecules for which the number of molecules within a given difference exceeds some threshold — i.e., that the local density in the neighborhood of a molecule is above some value.[3] The need for such a definition highlights the underlying fact that clusters are not physically distinct objects. Indeed, both simulation[3] and Density Functional Theory calculations[4–7] show, e.g., that a liquid cluster forming in a vapor has a broad and diffuse interface with the surrounding vapor. In fact, for small clusters, *all* molecules in the cluster could be classified as being part of the interface with no bulk material present at all. It is only when clusters are large, specifically when the radius of the cluster is large compared to the width of the interface, that the distinction between molecules being "within" a cluster and "outside" a cluster makes sense. However, this is not a particular problem in the context of CNT as the assumption of large clusters is usually made at some point in any case.

Another subtlety arises in the description of small clusters (i.e., clusters that result from density fluctuations). Let us consider for a moment an *under-saturated* solution for which there is no difficulty in assuming an equilibrium distribution of cluster sizes. In CNT, it is usually assumed that the probability density for observing a cluster of size N is proportional to the Boltzmann factor $P(N) \sim \exp(-\beta\Omega(N))$ where $\beta = 1/k_BT$, k_B is Boltzmann's constant, T is the temperature and $\Omega(N)$ is the free energy (e.g., in the capillary approximation) of a cluster of size (mass) N. On the other hand, in the capillary approximation, the free energy of a spherical cluster

a)Electronic mail: jlutsko@ulb.ac.be. URL: http://www.lutsko.com
b)Electronic mail: maduran@lec.csic.es

is often expressed in terms of its radius, R, $\widetilde{\Omega}(R) \equiv \Omega(N(R))$ and one could equally well suppose that the distribution of cluster radii is $\widetilde{P}(R) \sim \exp(-\beta\widetilde{\Omega}(R))$. However, these two expressions cannot both be true because, strictly speaking, there is the exact relation $P(N)dN = \widetilde{P}(R)dR$ from which one would deduce, e.g., $\widetilde{P}(R) = P(N(R))\frac{dN}{dR}$ or, more suggestively, $\widetilde{P}(R) \sim \exp(-\beta(\widetilde{\Omega}(R) + k_B T \ln \frac{dN(R)}{dR}))$. Alternatively, one could take $\widetilde{P}(R)$ as being more fundamental and deduce $P(N)$ from it with a corresponding shift of the effective free energy. Which of these is correct — or are they both wrong? In the CNT limit, these distinctions are unimportant since $\ln \frac{dN(R)}{dR} \sim \ln R$ and so the correction to the free energy is logarithmic and, hence, sub-dominant compared to the volume (R^D) and surface (R^{D-1}) contributions that determine $\widetilde{\Omega}(R)$ in D dimensions. However, in the description of small clusters, even at low supersaturations, these distinctions become important. One way to resolve this ambiguity would be to formulate a Fokker-Planck equation for the probability density and then to use this to determine the stationary (e.g., equilibrium) state. This is, in essence, the approach followed here.

Recently, there has been considerable interest in the possibility of extending CNT in two different but related directions. The first, discussed in detail by Prestipino et al.[8] concerns the correction of the capillary model for the free energy used in CNT whereas the second, discussed, e.g., by Lechner et al.,[9] concerns the choice of the reaction coordinate used to characterize the nucleation pathway. In both cases, of course, the quantity of most direct interest is the nucleation rate. For example, in the former work it was pointed out that sub-dominant terms in the expansion of the free energy as a function of cluster size make non-negligible contributions to the nucleation rate for small clusters. The goal of the present paper is to demonstrate a similar analysis which has the advantage that the entire nucleation processes can be described self-consistently. This is important because it turns out, as indicated above, that the choice of reaction coordinate also involves sub-dominant corrections to the nucleation rate so that the two avenues of investigation mentioned above are actually seen to come together.

The present analysis is based on a recently developed description of nucleation based on fluctuating hydrodynamics.[10–12] Its range of validity is restricted to systems governed by a diffusive dynamics typical of colloids or of macromolecules in solution. The new formulation was motivated by difficulties in finding a consistent way to use the tools of Density Functional Theory (an equilibrium theory) to describe the process of nucleation (a nonequilibrium process). In this approach, attention is focused on the formation of a single cluster and the fundamental quantity is the spatial density distribution that describes the cluster. One feature of the theory is that the density distribution can be parametrized in terms of a few physical quantities such as some measure of the size of a cluster, the width of the interface, etc. and a dynamical description of the evolution of these quantities results. When the cluster is parametrized by a single quantity, its "size", one makes contact with CNT and, indeed, one can reproduce CNT in the appropriate limits as described in Ref. 11

and below (subject to the same approximations such including that clusters are spherical). In this paper, the goal is to extend this approach to allow for an extension CNT that includes the finite width of the cluster interface in a self-consistent manner. It proves relatively straightforward to extract a nucleation rate from the formalism, essentially by following the development from CNT, and a comparison between the predictions of the extended theory and of CNT can be made (see, e.g., Eqs. (49), (50) and (73) and Table I below). The ambiguity discussed above concerning the equilibrium distribution is resolved and it is in fact found that neither the size nor the radius are the most natural reaction coordinate.

In Sec. II, the elements of Classical Nucleation Theory are reviewed with a particular focus on the assumptions that go into it. The formulation derived from fluctuating hydrodynamics is then described. It is noted that the formulation possesses a type of covariance with respect to the choice of reaction coordinate so that the ambiguities discussed above are resolved. An expression for the nucleation rate is also derived. Section III shows how the capillary model for a cluster can be used in conjunction with the dynamical theory to reproduce familiar results from CNT such as the expression for the monomer attachment rate, the nucleation rate, and the rate of growth of super-critical clusters. A model for clusters that allows for a finite interfacial width is described in Sec. IV. Section V presents comparisons between the models as well as tests of the assumptions underlying them. Finally, our results are summarized in Sec. VI.

II. THEORY

A. Classical nucleation theory

In CNT, one begins by assuming that clusters of a given size can grow by the addition of monomers, or can dissipate by spontaneously giving up a monomer to the surrounding solution. Processes involving the interaction of clusters which are both larger than monomers are ignored. Then, the concentration of clusters of size N, c_N, is determined by a set of rate equations of the form

$$\frac{dc_N}{dt} = f_{N-1}c_{N-1} - g_N c_N + g_{N+1}c_{N+1} - f_N c_N, \quad (1)$$

where f_N is the monomer attachment rate for a cluster of size N and g_N is the monomer detachment rate for a cluster of size N. This is known as the Becker-Döring model.[1,2] For condensation of a droplet from a diffuse solution, the monomer attachment rate is clearly going to be proportional to the rate at which monomers impinge on the cluster and so can be written as $f_N = \gamma_N j_N (4\pi R_N^2)$ where R_N is the radius of a cluster of size N, j_N is the rate per unit area at which molecules collide with the cluster and γ_N is a phenomenological factor describing the probability that a colliding molecule actually sticks to the cluster. The rate at which molecules in solution collide with the cluster can be determined by solving the diffusion equation with adsorbing boundary conditions with the result that in the long time (quasi-static approximation) $j_N = Dc_1/R_N$ where D is the monomer diffusion constant, so that $f_N = \gamma_N(4\pi DR_N)c_1$.[1] The monomer detachment rate is

much harder to estimate since it depends on the same physical details that are accounted for phenomenologically in the sticking constant (i.e., details concerning intermolecular interactions in the condensed phase). Kashchiev[1] gives a general argument for a relation between the monomer attachment frequency and the monomer detachment frequency for an *equilibrium* (i.e., under-saturated) system: namely, that in this case detailed balance demands that $f_{N-1}c_{N-1}^{(e)} = g_N c_N^{(e)}$ so that, if we can assume a Boltzmann distribution $c_N^{(e)} \propto \exp(-\beta\Delta\Omega_N)$ then it follows that $g_N = f_{N-1}\exp(\beta\Delta\Omega_N - \beta\Delta\Omega_{N-1})$. Note that this argument only holds for the under-saturated, equilibrium solution and that employing it for the supersaturated, *nonequilibrium* solution is a further assumption.

In the limit of large ($N \gg 1$) clusters, Eq. (1) can be approximated by treating $c_N(t) \to C(N;t)$ as a continuous function and expanding to get

$$
\begin{aligned}
\frac{dC(N,t)}{dt} = &\left(f(N)C(N;t) - \epsilon\frac{\partial f(N)C(N;t)}{\partial N} \right. \\
&\left. + \frac{1}{2}\epsilon^2\frac{\partial^2 f(N)C(N;t)}{\partial N^2} + ... \right) - g(N)C(N;t) \\
&+ \left(g(N)C(N;t) + \epsilon\frac{\partial g(N)C(N;t)}{\partial N} \right. \\
&\left. + \frac{1}{2}\epsilon^2\frac{\partial^2 g(N)C(N;t)}{\partial N^2} + ... \right) - f(N)C(N;t),
\end{aligned}
\tag{2}
$$

where a formal ordering parameter, ϵ, has been introduced: it will, at the end of the calculation, be set to one. Simplifying gives the Tunitskii equation,[1]

$$
\begin{aligned}
\frac{dC(N,t)}{dt} = \frac{\partial}{\partial N}&\left(\epsilon(g(N) - f(N))C(N;t) \right. \\
&\left. + \frac{1}{2}\epsilon^2\frac{\partial}{\partial N}(f(N) + g(N))C(N;t) \right) + O(\epsilon^3),
\end{aligned}
\tag{3}
$$

once we set $\epsilon = 1$ and dropping the third and higher order contributions. Similarly, expanding the approximation for the detachment rate gives

$$
\begin{aligned}
g(N) &= f(N-1)\exp(\beta\Delta\Omega(N) - \beta\Delta\Omega(N-1)) \\
&= f(N) - \epsilon\frac{\partial f(N)}{\partial N} + \epsilon f(N)\frac{\partial\beta\Delta\Omega(N)}{\partial N} + O(\epsilon^2)
\end{aligned}
\tag{4}
$$

and combing with the Tunitskii equation results in

$$
\begin{aligned}
\frac{dC(N,t)}{dt} &= \epsilon^2\frac{\partial}{\partial N}\left(\left(-\frac{\partial f(N)}{\partial N} + f(N)\frac{\partial\beta\Delta\Omega(N)}{\partial N} \right)C(N;t) \right. \\
&\qquad\left. + \frac{\partial}{\partial N}f(N)C(N;t) \right) + O(\epsilon^3) \\
&= \epsilon^2\frac{\partial}{\partial N}\left(f(N)\frac{\partial\beta\Delta\Omega(N)}{\partial N}C(N;t) \right. \\
&\qquad\left. + f(N)\frac{\partial}{\partial N}C(N;t) \right) + O(\epsilon^3),
\end{aligned}
\tag{5}
$$

which is the well-known Zeldovich equation.[1]

Although these manipulations are formally correct, there is in fact an additional assumption being made. Although the derivative $\frac{\partial f(N)}{\partial N} \sim \frac{f}{N}$ can be considered to be small for large N, it is not clear that the same is true of the free energy since in fact $\Delta\Omega(N) - \beta\Delta\Omega(N-1) \sim O(1)$. Taking only the first term in the expansion of the exponential in Eq. (4) is therefore only justified if it can be assumed that $\frac{\partial\beta\Delta\Omega(N)}{\partial N} \ll 1$: *in other words, in the vicinity of the critical cluster*. It is therefore clear that the Zeldovich equation is, strictly speaking, only applicable for large N and for N near the critical size.

Noting that the probability to observe a cluster of size N, $P(N) \equiv \frac{n(N)}{\sum_{N'} n(N')}$ where $n(N)$ is the number of clusters of size N, it is easily seen that $C(N;t) = C_{tot}(t)P(N;t)$ where $C_{tot}(t) = \sum_N C(N;t)$ is the total number of molecules per unit volume, i.e., the average concentration. So, the Zeldovich equation implies a Fokker-Planck equation for $P(N;t)$

$$
\begin{aligned}
\frac{dP(N,t)}{dt} = \frac{\partial}{\partial N}&\left(f(N)\frac{\partial\beta\Delta\Omega(N)}{\partial N}P(N;t) \right. \\
&\left. + f(N)\frac{\partial}{\partial N}P(N;t) \right) - P(N;t)\frac{\partial\ln C_{tot}(t)}{\partial t},
\end{aligned}
\tag{6}
$$

where the source term vanishes if the total number of molecules is constant.

B. Generalization based on fluctuating hydrodynamics

In general, nucleation is a fluctuation-driven phenomena limited by the rate of transport of mass, heat, etc. within the system. In fluid systems, it is therefore naturally described by Landau's fluctuating hydrodynamics which includes both transport and thermal fluctuations.[13,14] The description of nucleation via fluctuating hydrodynamics has been described in Refs. 10–12. A particularly simple description is possible when the system is diffusion limited and over-damped, as is an appropriate description for colloids and macromolecules in solution. The dynamics can be reduced to an equation for the time-dependent spatial density and a reduced description is possible when the density field is parametrized in the form

$$
\rho(r;t) \to \rho(r; X_1(t), X_2(t)...X_N(t)),
\tag{7}
$$

so that the time dependence occurs via a set of N parameters which could include, e.g., the size of the cluster, the width of the interface, the density at the center of the cluster, etc. Here, we will be interested in the case that there is a single parameter. Then, it was shown[10–12] that the evolution of the order parameter may, in the case of an infinite system, be approximated by a stochastic differential equation of the form

$$
\begin{aligned}
\frac{dX}{dt} = &-Dg^{-1}(X)\frac{\partial\beta\Omega}{\partial X} - D\frac{1}{2}g^{-2}(X)\frac{\partial g(X)}{\partial X} \\
&+ \sqrt{2Dg^{-1}(X)}\xi(t),
\end{aligned}
\tag{8}
$$

where D is the diffusion constant, $\Omega(X) = F(x) - \mu N(X)$, is the grand potential for chemical potential μ and where $N(X)$ is the total number of molecules. The last term on the right

is proportional to $\xi(t)$ which is a delta-function correlated fluctuating force with mean zero and variance equal to one. The quantity $g(X)$ that determines the kinetic coefficient of the SDE as well as the amplitude of the noise is given by

$$g(X) = \int_0^\infty \frac{1}{4\pi r^2 \rho(r;X)} \left(\frac{\partial m(r;X)}{\partial X} \right)^2 dr, \qquad (9)$$

where the cumulative mass is

$$m(r;X) = 4\pi \int_0^r \rho(r';X) r'^2 dr'. \qquad (10)$$

Since the amplitude of the noise in the SDE is state-dependent, it is important to note that the equation is given according to the Ito interpretation. The function $g(X)$ will be called "the metric" as it can be used to define a meaningful length in the general theory,[11] although it will not be used for that purpose here.

1. Fokker-Planck equation

It is useful to also consider the equivalent Fokker-Planck equation which is[15]

$$\frac{\partial P(X,t)}{\partial t} = D \frac{\partial}{\partial X} \left(g^{-1}(X) \frac{\partial \beta \Omega(X)}{\partial X} \right.$$
$$\left. + g^{-1/2}(X) \frac{\partial}{\partial X} g^{-1/2}(X) \right) P(X,t). \quad (11)$$

Note that this can be written as

$$\frac{\partial P(X,t)}{\partial t} = D \frac{\partial}{\partial X} \left(g^{-1}(X) \frac{\partial (\beta \Omega(X) - \ln g^{1/2}(X))}{\partial X} \right.$$
$$\left. + g^{-1}(X) \frac{\partial}{\partial X} \right) P(X,t), \quad (12)$$

so that comparison with Eq. (6) shows that these are formally the same with the monomer attachment frequency replaced by $g^{-1}(X)$ and with the free energy shifted by a term logarithmic in $g(X)$. From the expression for $g(X)$, Eq. (9), it is evident that this function varies as the volume of the cluster so that the shift to the free energy will go like the log of the radius of the cluster.

Let us look for a stationary solution determined, for some constant J_s, from

$$-D \left(g^{-1}(X) \frac{\partial \beta \Omega(X)}{\partial X} + g^{-1/2}(X) \frac{\partial}{\partial X} g^{-1/2}(X) \right) P_s(X) = J_s, \qquad (13)$$

giving

$$P_s(X) = A g^{1/2}(X) \exp(-\beta \Omega(X)) - D^{-1} J_s g^{1/2}(X)$$
$$\times \exp(-\beta \Omega(X)) \int^X g^{1/2}(X') \exp(\beta \Omega(X')) dX'$$
$$(14)$$

for some constant A. Now, if we note that for some change of variables $X \to Y(X)$, we have that

$$\widetilde{g}(Y) = g(X(Y)) \left(\frac{dX}{dY} \right)^2 \qquad (15)$$

and so

$$\widetilde{P}_s(Y) = A \widetilde{g}^{1/2}(Y) \exp(-\beta \Omega(Y)) - D^{-1} J_s \widetilde{g}^{1/2}(Y)$$
$$\times \exp(-\beta \Omega(Y)) \int^Y \widetilde{g}^{1/2}(Y') \exp(\beta \Omega(Y')) dY'$$
$$= A g^{1/2}(X) \frac{dX}{dY} \exp(-\beta \Omega(X))$$
$$- D^{-1} J_s g^{1/2}(X) \frac{dX}{dY} \exp(-\beta \Omega(X))$$
$$\times \int^X g^{1/2}(X') \frac{dX'}{dY'} \exp(\beta \Omega(X')) dY'$$
$$= P_s(X) \frac{dX}{dY}. \qquad (16)$$

Hence, this solution is completely covariant — there is not ambiguity as to which variable is used. If the system is stable (i.e., under-saturated), then it makes sense to seek an equilibrium ($J = 0$) solution which is

$$P_{eq}(X) = A g^{1/2}(X) \exp(-\beta \Omega(X))$$
$$= A \exp \left(-\beta \left(\Omega(X) - \frac{1}{2} k_B T \ln g(X) \right) \right), \quad (17)$$

where the constant A is fixed by normalization. Note that this has the form of a canonical distribution in terms of the shifted free energy.

2. A canonical variable

In the present, single-variable, case there is always a special variable for which these expressions simplify. Given any variable, X, it is defined via

$$dY = \sqrt{g(X)} dX \qquad (18)$$

and an arbitrary boundary condition that will be taken to be $Y(0) = 0$. In terms of this variable, the Langevin equation becomes

$$\frac{dY}{dt} = -D \frac{\partial \beta \widetilde{\Omega}(Y)}{\partial Y} + \sqrt{2D} \xi(t), \qquad (19)$$

where $\widetilde{\Omega}(Y) = \Omega(X(Y))$. The corresponding Fokker-Planck equation is

$$\frac{\partial \widetilde{P}(Y,t)}{\partial t} = D \frac{\partial}{\partial Y} \left(\frac{\partial \beta \widetilde{\Omega}(Y)}{\partial Y} + \frac{\partial}{\partial Y} \right) \widetilde{P}(Y,t). \qquad (20)$$

These are equations that might have been written down on phenomenological grounds but they would have been ambiguous since there is no a priori reason to use the mass as the independent variable rather than, say, the equimolar radius. This illustrates the way in which grounding the theory on a more fundamental description serves to remove such ambiguities.

3. Nucleation rate

The standard argument to obtain the nucleation rate has a long history, as discussed by Hanggi et al.,[16] and is based on boundary conditions according to which (a) the distribution is stationary and (b) the distribution goes to zero at some point beyond the critical cluster, say at $X = X_+$ (see also Refs. 1

and [17]). The latter condition represents the physical fact that once a cluster is slightly larger than the critical cluster, it will almost certainly grow forever and so plays no role in the stochastic part of the process. One therefore imagines establishing a steady state by removing clusters once they reach size $X = X_+$ and simultaneously re-injecting the removed material in the form of monomers. Here, since we are only interested in the case of one-dimensional barrier-crossing, this rate is easily evaluated from the exact solution for the steady-state distribution. (For multivariate problems, further approximations are necessary as described in Refs. [16], [18], and [19].)

First, the nucleation rate, J, is the rate of production of super-critical nuclei per unit volume and is therefore

$$J = \frac{d}{dt} \int_{N^*}^{\infty} \frac{1}{V} n(N; t) dN = \frac{d}{dt} \int_{N^*}^{\infty} \frac{N_c(t)}{V} P(N; t) dN, \tag{21}$$

where $N_c(t) = \int_0^{\infty} n(N; t) dN$ is the total number of clusters. Making use of the Fokker-Planck equation, this becomes

$$J = \frac{1}{V} \frac{dN_c(t)}{dt} \int_{N^*}^{\infty} P(N, t) dN$$
$$- D \frac{N_c(t)}{V} \left(g^{-1}(N) \frac{\partial \beta \Omega(N)}{\partial N} P(N, t) \right.$$
$$\left. + g^{-1/2}(N) \frac{\partial}{\partial N} g^{-1/2}(N) P(N, t) \right)_{N=N^*}. \tag{22}$$

In the artificially imposed stationary state, the first term on the right does not contribute leaving

$$J = -D \frac{N_c}{V} \left(g^{-1}(N) \frac{\partial \beta \Omega(N)}{\partial N} P_s(N) \right.$$
$$\left. + g^{-1/2}(N) \frac{\partial}{\partial N} g^{-1/2}(N) P_s(N) \right)_{N=N^*}. \tag{23}$$

Since we remove clusters of size X_+ as they form, we need a steady-state distribution that satisfies $P(X_+) = 0$. When this is imposed, the general steady-state distribution, Eq. (14), becomes

$$P_s(X) = D^{-1} J_s g^{1/2}(X) \exp(-\beta \Omega(X))$$
$$\times \int_X^{X_+} g^{1/2}(X') \exp(\beta \Omega(X')) dX'. \tag{24}$$

Substitution into the expression for the nucleation rate gives

$$J = \frac{N_c}{V} J_s. \tag{25}$$

The remainder of the development concerns the relation between the imposed stationary flux, J_s, and real properties of

the system. Zeldovich, as described by Kashchiev[1] requires that the concentration of monomers be fixed at the "equilibrium" value

$$\rho_{\infty} = \frac{N_c}{V} P_s(X_1), \tag{26}$$

where X_1 is the value of the parameter X that fixes the number of molecules in the cluster to be 1. (The notation chosen here is motivated by the fact that one expects that ρ_{∞} is the density far from a cluster.) This then gives

$$J_{CNT}$$

$$= \frac{D \rho_{\infty}}{g^{1/2}(X_1) \exp(-\beta \Omega(X_1)) \int_{X_1}^{X_+} g^{1/2}(X') \exp(\beta \Omega(X')) dX'}$$

$$= \frac{D \rho_{\infty}}{\int_{X_1}^{X_+} \exp\left(\Delta \beta \Omega(X') - \frac{1}{2} \ln \frac{g(X')}{g(X_1)}\right) dX'}. \tag{27}$$

The second equality gives the result written in terms of the adjusted free energy. A saddle-point evaluation for the choice of number of molecules as the variable gives

$$J_{CNT} \simeq D \rho_{\infty} g^{-1}(N_{**}) \sqrt{\frac{1}{2\pi}} \sqrt{|\Delta \beta \Omega_{**}''|} \exp(-\Delta \beta \Omega_{**}), \tag{28}$$

where the adjusted critical size is determined from

$$\Delta \beta \Omega'(N_{**}) - \frac{1}{2} \frac{g'(N_{**})}{g(N_{**})} = 0 \tag{29}$$

and where

$$\Delta \beta \Omega_{**} \equiv \Delta \beta \Omega(N_{**}) - \frac{1}{2} \ln \frac{g(N_{**})}{g(1)}. \tag{30}$$

The problem with this expression is that it is not covariant and the result will depend on the choice of X. This is due to the fact that the condition, fixing the fraction of monomers, is itself not covariant since only the combination $P_s(X) dX$ is invariant under a change of variables. We stress that within the context of CNT this fact is irrelevant since the assumption of large N means that the lack of covariance is due to sub-dominant terms in the distribution. However, within the context of the general theory, this lack of covariance indicates unnecessary ambiguity since different results will be obtained for different choices of the variable X appearing in the boundary condition, Eq. (26).

The first step to solving this problem is to note that, as defined here, $P_s(X)$ is normalized so that we can immediately fix the unknown coefficient giving

$$P_s(X) = \frac{g^{1/2}(X) \exp(-\beta \Omega(X)) \int_X^{X_+} g^{1/2}(X') \exp(\beta \Omega(X')) dX'}{\int_0^{X_+} dX \, g^{1/2}(X) \exp(-\beta \Omega(X)) \int_X^{X_+} g^{1/2}(X') \exp(\beta \Omega(X')) dX'} \tag{31}$$

and

$$J = \frac{N_c}{V} \frac{D}{\int_0^{X_+} dX \, g^{1/2}(X) \exp(-\beta \Omega(X)) \int_X^{X_+} g^{1/2}(X') \exp(\beta \Omega(X')) dX'}. \tag{32}$$

This still leaves the question of determining the overall concentration of clusters. We do this following the general idea used in CNT but being careful to preserve covariance by imposing a condition on the total number of molecules, $\mathcal{N}(N_0)$ in clusters up

to size N_0,

$$\mathcal{N}(N_0) = \int_0^{N_0} N n(N) dN = N_c \int_0^{N_0} N P_s(N) dN, \tag{33}$$

so that

$$N_c = \mathcal{N}(N_0) \frac{\int_0^{X_+} dX g^{1/2}(X) \exp(-\beta\Omega(X)) \int_X^{X_+} g^{1/2}(X') \exp(\beta\Omega(X')) dX'}{\int_0^{X(N_0)} dX \, N(X) g^{1/2}(X) \exp(-\beta\Omega(X)) \int_X^{X_+} g^{1/2}(X') \exp(\beta\Omega(X')) dX'} \tag{34}$$

giving the nucleation rate as

$$J(N_0) = \frac{D\overline{\rho}(N_0)}{\int_0^{X(N_0)} dX \, N(X) g^{1/2}(X) \exp(-\beta\Omega(X)) \int_X^{X(N_+)} g^{1/2}(X') \exp(\beta\Omega(X')) dX'}, \tag{35}$$

where the average density is $\overline{\rho}(N_0) = \frac{\mathcal{N}(N_0)}{V}$ and where we have indicated that the boundary condition makes the nucleation rate a function of N_0. If N_0 is chosen to be large, say the size of the critical cluster N_*, then $\mathcal{N}(N_*)$ is the total population in the initial (i.e., precritical) state and this is essentially the standard expression for the reaction rate derived from one-dimensional barrier crossing (the "flux over population" expression, see, e.g., Hanggi et al.[16]) except that an extra factor of N occurs in the integral in the denominator, which is meant to count the population in the initial state, because each cluster contains many molecules. If N_0 is chosen to be on the order of one, then we are essentially counting the number of monomers. Importantly, if most of the material exists in the form of small clusters (monomers, dimers, etc.) then these estimates will be the same. If they are not the same, then a significant amount of mass is present in the form of clusters and the implicit assumption (made here as well as in CNT) that clusters do not interact is probably invalid. One of the goals below will therefore be to monitor the validity of this assumption.

Assuming that the bulk of the material is in the form of small clusters and that N_0 is chosen sufficiently large, an approximation to the exact expression, Eq. (35), can be developed as described in Appendix A. Assuming that the free energy and number of molecules as functions of the canonical variable Y have the expansions $\Delta\beta\widetilde{\Omega}(Y) = \widetilde{\Omega}_0 Y^\alpha + ...$ and $N(Y) = n_0 Y^\beta + ...$ for small Y, it turns out that the approximation implies that, for small X, the stationary distribution can be approximated by

$$P_s(X) \sim \frac{\alpha\widetilde{\Omega}_0^{1/\alpha}}{\Gamma\left(\frac{1}{\alpha}\right)} g^{1/2}(X) \exp(-\Delta\beta\Omega(X)) \tag{36}$$

and in which case the nucleation rate is approximated by

$$J \sim \frac{1}{\sqrt{2\pi}} \rho_\infty D \frac{\alpha\widetilde{\Omega}_0^{\frac{\beta+1}{\alpha}}}{\Gamma\left(\frac{\beta+1}{\alpha}\right) n_0}$$

$$\times \sqrt{\left|\beta\Omega''(X_*)g^{-1}(X_*)\right|} \exp(-\Delta\beta\Omega_*). \tag{37}$$

III. CAPILLARY MODEL: CLASSICAL NUCLEATION THEORY

The process we will describe is the nucleation of a liquid, with bulk density ρ_l, from a vapor with bulk density ρ_v at some temperature T and chemical potential μ. Let $\omega(\rho)$ be the free energy (grand potential) per unit volume, so that $\omega(\rho) = f(\rho) - \mu\rho$, where $f(\rho)$ is the Helmholtz free energy per unit volume. Then, the liquid and vapor densities are determined by the imposed chemical potential via $\omega'(\rho_l) = \omega'(\rho_v) = 0$ and since we choose thermodynamic conditions such that the vapor is metastable, $\omega(\rho_l) < \omega(\rho_v)$. To recover CNT it is only necessary to use the capillary model for the density

$$\rho(r) = \begin{cases} \rho_0, & r < R \\ \rho_\infty, & R < r \end{cases}, \tag{38}$$

where we take $\rho_0 = \rho_l$ for the density inside the cluster and $\rho_\infty = \rho_v$ for the density outside the cluster. The capillary-theory expression for the free energy of the cluster is

$$\Delta\beta\Omega(R) = \frac{4\pi}{3} R^3 \Delta\beta\omega(\rho_0) + 4\pi R^2 \gamma, \tag{39}$$

where the second term represents the effect of surface tension. Note that the only parameter that is allowed to vary is the radius of the cluster.

The model for the density gives the cumulative mass density

$$m(r) = \begin{cases} \frac{4\pi}{3} r^3 \rho_0, & r < R \\ \frac{4\pi}{3} R^3 \rho_0 + \frac{4\pi}{3}(r^3 - R^3)\rho_\infty, & R < r \end{cases} \tag{40}$$

and the metric

$$g(R) = \int_R^\infty \frac{1}{4\pi r^2 \rho_\infty}(4\pi R^2(\rho_0 - \rho_\infty))^2 dr$$

$$= \frac{(\rho_0 - \rho_\infty)^2}{\rho_\infty} 4\pi R^3. \tag{41}$$

The excess number of molecules in the cluster is

$$\Delta N = \frac{4\pi}{3} R^3(\rho_0 - \rho_\infty) \tag{42}$$

and the metric in terms of the number of molecules is

$$g(\Delta N) = g(R)\left(\frac{dR}{d\Delta N}\right)^2 = \frac{1}{4\pi\rho_\infty R(\Delta N)}. \quad (43)$$

The canonical variable is

$$Y = \int_0^R \sqrt{\frac{(\rho_0 - \rho_\infty)^2}{\rho_\infty}4\pi R'^3}\,dR' = \frac{2}{5}\sqrt{4\pi\frac{(\rho_0 - \rho_\infty)^2}{\rho_\infty}}R^{5/2}. \quad (44)$$

In comparing the Fokker-Planck equation for the generalized theory, Eq. (12), to that of Zeldovitch, Eq. (6), one finds that agreement provided that the logarithmic corrections to the free energy are neglected and the effective monomer attachment frequency is identified as

$$f(N) = g^{-1}(\Delta N) = 4\pi\rho_\infty R(\Delta N), \quad (45)$$

which is the usual result for diffusion-limited nucleation in the case that the phenomenological "sticking constant" is equal to one.[1] The stochastic differential equation for the radius is

$$\frac{dR}{dt} = -D\frac{\rho_\infty}{(\rho_0 - \rho_\infty)^2 4\pi R^3}\frac{\partial(\beta\Omega + \ln g^{1/2}(R))}{\partial R}$$
$$+ \sqrt{2Dg^{-1}(R)}\xi(t). \quad (46)$$

When the cluster is large (i.e., when it is super-critical), the noise becomes unimportant and the radius grows as

$$\frac{dR}{dt} = \frac{D\rho_\infty|\Delta\beta\omega(\rho_0)|}{(\rho_0 - \rho_\infty)^2}R^{-1} + O(R^{-2}), \quad (47)$$

which gives the classical result $R \sim t^{1/2}$ when the higher order terms are neglected.[20] In the weak liquid limit, $\rho_\infty \ll \rho_0$, the coefficient of R^{-1} agrees with that given by Lifshitz et al.[17]

The (non-covariant) CNT-like nucleation rate, from Eq. (28), is

$$J_{CNT} \sim 2\sqrt{2\pi}D\rho_\infty^2 R(\Delta N_{**})\sqrt{|\Delta\beta\Omega''(\Delta N_{**})|}$$
$$\times \exp(-\Delta\beta\Omega_{**}) \quad (48)$$

and in the limit that the logarithmic corrections to the free energy are negligible, this agrees with the result from CNT. In the following, we will use as a reference the usual CNT result that is obtained by ignoring the logarithmic shift in the free energy,

$$J_{CNT} \sim 2\sqrt{2\pi}D\rho_\infty^2 R(\Delta N_*)\sqrt{|\Delta\beta\Omega''(\Delta N_*)|}$$
$$\times \exp\left(-\Delta\beta\Omega_*^{CNT}\right)$$
$$= \frac{D\rho_\infty^2}{\rho_0}\frac{|\Delta\beta\omega|}{\sqrt{\gamma}}\exp\left(-\Delta\beta\Omega_*^{CNT}\right), \quad (49)$$

$$\Delta\beta\Omega_*^{CNT} = \frac{16\pi}{3}\frac{\gamma^3}{(\Delta\beta\omega)^2}.$$

The approximate nucleation rate based on a condition of fixed mass, Eq. (37), is

$$\frac{J}{J_{CNT}} \sim \frac{3\rho_0\sqrt{2}(4\pi)^{\frac{3}{4}}}{\Gamma\left(\frac{11}{4}\right)\sqrt{\rho_\infty}|\Delta\beta\omega|(\rho_0 - \rho_\infty)^2}$$
$$\times \gamma^{\frac{13}{4}}|\beta\Omega''(R_*)g^{-1}(R_*)|^{1/2}$$
$$= \frac{3\sqrt{2}(4\pi)^{\frac{3}{4}}}{\Gamma\left(\frac{11}{4}\right)}\frac{\rho_0}{(\rho_0 - \rho_\infty)^3}\gamma^{\frac{9}{4}}|\Delta\beta\omega(\rho_0)|^{1/2}. \quad (50)$$

Note that in all cases, we are assuming that the average concentration of material in the (hypothetical) steady state, $\overline{\rho}$, is the same as the background density, ρ_∞.

IV. EXTENDED MODEL: FINITE CLUSTER WIDTH

A. The cluster structure and the metric

A significant short coming of the capillary model is that the width of the cluster's interface is zero. A more realistic model will have a finite width which, from various simulations and Density Functional Theory calculations, might be expected to be two or three molecular diameters in width. A simple extension of the capillary model to take account of a finite width is the piecewise-linear profile,

$$\rho(r) = \begin{cases} \rho_0, & r < R - w \\ \rho_0 - (\rho_0 - \rho_\infty)\frac{r-(R-w)}{w}, & R - w < r < R \\ \rho_\infty, & R < r \end{cases}. \quad (51)$$

The corresponding cumulative mass distribution for $0 \leq R \leq w$ is

$$m(r) = \Theta(R - r)\frac{\pi}{3w}(\rho_0 - \rho_\infty)r^3(4R - 3r)$$
$$+ \Theta(r - R)\frac{\pi}{3w}(\rho_0 - \rho_\infty)R^4 + V(r)\rho_\infty, \quad (52)$$

while for $R > w$ it becomes

$$m(r) = (\rho_0 - \rho_\infty)V(r)\Theta(R - w - r)$$
$$+ \Theta(r - (R - w))\Theta(R - r)$$
$$\times \frac{\pi}{3w}(\rho_0 - \rho_\infty)(r^3(4R - 3r) - (R - w)^4)$$
$$+ \Theta(r - R)\frac{\pi}{3w}(\rho_0 - \rho_\infty)(R^4 - (R - w)^4)$$
$$+ V(r)\rho_\infty. \quad (53)$$

Calculation of the metric is then straightforward with the result that for $0 \leq R \leq w$,

$$g(r) = 4\pi\left(\frac{\rho_0 - \rho_\infty}{3w}\right)^2\left[\frac{wR^4}{(\rho_0 - \rho_\infty)}\left[a^4\ln\left(\frac{a}{a-1}\right)\right.\right.$$
$$\left.\left. - \left(a^3 + \frac{1}{2}a^2 + \frac{1}{3}a + \frac{1}{4}\right)\right] + \frac{1}{\rho_\infty}R^5\right] \quad (54)$$

with

$$a = 1 + \frac{w\rho_\infty}{R(\rho_0 - \rho_\infty)}. \quad (55)$$

Note that for small clusters, $R \ll w$, this gives

$$g(r) = \frac{8\pi}{15}\left(\frac{\rho_0 - \rho_\infty}{w}\right)^2 \frac{R^5}{\rho_\infty}\left[1 - \frac{(\rho_0 - \rho_\infty)}{36\rho_\infty}\frac{R}{w} + O\left(\left(\frac{R}{w}\right)^2\right)\right]. \tag{56}$$

For larger radii, $R > w$, the result is

$$g(r) = \frac{4\pi}{9}\frac{(\rho_0 - \rho_\infty)^2}{\rho_\infty}\frac{(R^3 - (R-w)^3)^2}{w^2 R}$$
$$+ \frac{4\pi}{9}\left(\frac{\rho_0 - \rho_\infty}{w}\right)\left(\frac{(R-w)^6}{R^2 a^2}\ln\frac{R}{R-w} - \frac{((R-w)^3 - R^3 a^3)^2}{R^2 a^2}\ln\left(\frac{\rho_\infty}{\rho_0}\right)\right)$$
$$- \frac{4\pi}{9}\left(\frac{\rho_0 - \rho_\infty}{w}\right)\left(\begin{array}{c} -w\frac{(R-w)^5}{R^2 a} - \left(2(R-w)^3 - R^3 a^3\right)w \\ + \frac{1}{2}R^2 a^2(R^2 - (R-w)^2) + \frac{1}{3}Ra(R^3 - (R-w)^3) + \frac{1}{4}(R^4 - (R-w)^4) \end{array}\right). \tag{57}$$

One can again define a canonical variable using

$$\frac{dY}{dR} = \sqrt{g(r)} \tag{58}$$

and for small clusters one finds

$$Y = \frac{2}{7}\left(\frac{\rho_0 - \rho_\infty}{w}\right)\sqrt{\frac{8\pi}{15\rho_\infty}}R^{\frac{7}{2}}\left(1 + O(r)\right). \tag{59}$$

B. Free energy model

It would be somewhat inconsistent to use the capillary approximation for the free energy given that the assumed profile now has finite width. We therefore consider a simple, but more fundamental free energy model based on the squared-gradient approximation,

$$\Omega[\rho] = \int\left(\omega(\rho(\mathbf{r})) + \frac{1}{2}K(\nabla\rho(\mathbf{r}))^2\right)d\mathbf{r}, \tag{60}$$

where $\omega(\rho) = f(\rho) - \mu\rho$ is the grand potential per unit volume, $f(\rho)$ is the Helmholtz free energy per unit volume for a bulk system with uniform density ρ which can be determined based on a given pair potential using thermodynamic perturbation theory or liquid state integral equation methods. The squared-gradient coefficient, K, can be estimated from a model interaction potential using the results of Ref. 21. For the assumed density profile, this becomes

$$\Omega(R; w) - \Omega_\infty$$
$$= \frac{4\pi}{3}(\max(R - w, 0))^3 \Delta\omega(\rho_0)$$
$$+ 4\pi\int_{\max(0, R-w)}^{R}\Delta\omega\left(\rho_0 - (\rho_0 - \rho_\infty)\frac{r - (R-w)}{w}\right)r^2 dr$$
$$+ \frac{1}{2}K\frac{4\pi}{3}\left(R^3 - \max(R - w, 0)^3\right)\left(\frac{\rho_0 - \rho_\infty}{w}\right)^2. \tag{61}$$

The result for $R > w$, can also be written as

$$\Omega(R; w) - \Omega_\infty = \frac{4\pi}{3}(R - w)^3\Delta\omega(\rho_0)$$
$$+ 4\pi\left(\overline{\omega}_0 w + K\frac{(\rho_0 - \rho_\infty)^2}{2w}\right)R^2$$
$$- 4\pi\left(2\overline{\omega}_1 w + K\frac{(\rho_0 - \rho_\infty)^2}{2w}\right)Rw$$
$$+ 4\pi\left(\overline{\omega}_2 w + K\frac{(\rho_0 - \rho_\infty)^2}{6w}\right)w^2, \tag{62}$$

where the density moments of the excess free energy per unit volume are

$$\overline{\omega}_n = \frac{1}{(\rho_0 - \rho_\infty)^{n+1}}\int_{\rho_\infty}^{\rho_0}(\omega(x) - \omega(\rho_\infty))(x - \rho_\infty)^n dx, \tag{63}$$

thus showing how, for large clusters (in the $\frac{w}{R} \to 0$ limit), one recovers something like the capillary approximation but with a variable width. Minimizing with respect to the width at constant radius and solving as an expansion in the radius (i.e., assuming $R \gg w$) gives (see Appendix B)

$$w_{\min} = w_0\left(1 + \frac{\Delta\omega(\rho_0) - 2\overline{\omega}_1}{\Delta\omega(\rho_0) - \overline{\omega}_0}\frac{w_0}{R} + ...\right) \tag{64}$$

with

$$w_0 = \sqrt{\frac{K(\rho_0 - \rho_\infty)^2}{2(\overline{\omega}_0 - \Delta\omega(\rho_0))}} \tag{65}$$

and the free energy becomes

$$\Omega(R; w) - \Omega_\infty$$
$$= \left(\frac{4}{3}\pi R^3\right)\Delta\omega(\rho_0)$$
$$+ (4\pi R^2)\left(\frac{1}{2}\Delta\omega(\rho_0)w_0 + \frac{K(\rho_0 - \rho_\infty)^2}{w_0}\right) + O(R^1). \tag{66}$$

The higher order terms are a simple illustration of the post-CNT corrections to the free energy barrier recently discussed by Prestipino *et al.*[8] At lowest order, the implied capillary-like model is

$$\Omega(R; w) - \Omega_\infty = \left(\frac{4\pi}{3} R^3\right) \Delta\omega(\rho_0) + (4\pi R^2)\gamma \quad (67)$$

with

$$\gamma = \frac{1}{2}\Delta\omega(\rho_0)w_0 + \frac{K(\rho_0 - \rho_\infty)^2}{w_0}$$

$$= \left(1 + \frac{\Delta\omega(\rho_0)}{4(\overline{\omega}_0 - \Delta\omega(\rho_0))}\right)$$

$$\times \sqrt{2K(\rho_0 - \rho_\infty)^2(\overline{\omega}_0 - \Delta\omega(\rho_0))}. \quad (68)$$

Note that the coefficient γ depends on the supersaturation implicitly via its dependence on the densities. In CNT, it is more common to ignore the state dependence of this coefficient and to fix its value to that which gives the correct value of the planar surface tension at coexistence. Here, "correct" will be taken to mean that it gives the same value as the full model with finite width. Since at coexistence the free energies of the two phases are equal, this is

$$\gamma_{CNT} = \left(\rho_0^{coex} - \rho_\infty^{coex}\right)\sqrt{2K\overline{\omega}_0^{coex}}, \quad (69)$$

where the superscripts indicate that the densities are those at coexistence.

One peculiarity of this model is that the value of the width that minimizes the free energy undergoes a bifurcation as the radius increases. This is due to the fact that for $w > R$, the free energy is

$$\Omega(R; w) - \Omega_\infty$$

$$= 4\pi \int_0^R \Delta\omega\left(\rho_0 - (\rho_0 - \rho_\infty)\frac{r - (R - w)}{w}\right)r^2 dr$$

$$+ \frac{2\pi}{3} K R^3 \left(\frac{\rho_0 - \rho_\infty}{w}\right)^2 \quad (70)$$

and it is clear that the free energy difference decreases monotonically to zero as the width increases.

Expanding in R one has that

$$\Omega(R; w) - \Omega_\infty$$

$$= \frac{2\pi}{3} K \left(\frac{\rho_0 - \rho_\infty}{w}\right)^2 R^3$$

$$+ \frac{8\pi}{5!} \Delta\omega''(\rho_\infty) \left(\frac{\rho_0 - \rho_\infty}{w}\right)^2 R^5 + O(R^6) \quad (71)$$

and in terms of the canonical variable

$$\Omega(Y; w) - \Omega_\infty$$

$$= \frac{2\pi}{3} K \left(\frac{\rho_0 - \rho_\infty}{w}\right)^2$$

$$\times \left(\frac{7}{2}\left(\frac{w}{\rho_0 - \rho_\infty}\right)\sqrt{\frac{15\rho_\infty}{8\pi}}\right)^{\frac{6}{7}} Y^{\frac{6}{7}} + O\left(Y^{\frac{10}{7}}\right), \quad (72)$$

which can be used to evaluate the approximate nucleation rate

$$J_{ext} \sim Z_{ext}|\beta\Omega''(R_*)|^{1/2} \exp(-\beta\Omega_*) \quad (73)$$

with the prefactor

$$Z_{ext} = D\frac{\sqrt{5}}{\Gamma\left(\frac{13}{6}\right)}\left(\frac{9\pi}{2}\right)^{\frac{1}{6}}\sqrt{K\rho_\infty}(\overline{\omega}_0 - \Delta\omega(\rho_0))^{\frac{5}{3}} g^{-1/2}(R_*). \quad (74)$$

V. RESULTS AND COMPARISONS

In order to illustrate the theory developed above, we have performed detailed calculations for a model globular protein. We assume that the solvent can be approximated, crudely, by assuming Brownian dynamics of the (large) solute molecules which also experience an effective pair interaction for which we use the ten Wolde-Frenkel interaction potential

$$v(r) = \begin{cases} \infty, & r \leq \sigma \\ \frac{4\epsilon}{\alpha^2}\left(\left(\frac{1}{(\frac{r}{\sigma})^2 - 1}\right)^6 - \alpha\left(\frac{1}{(\frac{r}{\sigma})^2 - 1}\right)^3\right), & r \geq \sigma \end{cases} \quad (75)$$

with $\alpha = 50$ which is then cutoff at $r_c = 2.5\sigma$ and shifted so that $v(r_c) = 0$. The temperature is fixed at $k_B T = 0.375\epsilon$ and the equation of state computed using thermodynamic perturbation theory. The transition we study is that between the dilute phase and the dense protein phase which, in the present simplified picture, is completely analogous to the vapor-liquid transition for particles interacting under the given pair potential. Throughout this section, the supersaturation will be defined as the ratio of the density of the vapor phase to that of the vapor at coexistence, $S \equiv \rho_v/\rho_{vc}$. The gradient coefficient, K, is calculated from the pair potential using the approximation given in Ref. 21,

$$\beta K \simeq -\frac{2\pi}{45}d^5\beta v(d) + \frac{2\pi}{15}\int_d^\infty (2d^2 - 5r^2)\beta v(r)r^2 dr, \quad (76)$$

where d is the effective hard-sphere diameter for which we use the Barker-Henderson approximation. For the temperature used here we get $\beta K = 1.80322\sigma^5$.

In the following, we will characterize the size of a cluster by either the total excess number of molecules in the cluster,

$$\Delta N \equiv \int (\rho(\mathbf{r}) - \rho_\infty)d\mathbf{r}, \quad (77)$$

or by its equimolar radius, R_E, which is related to ΔN by

$$\Delta N = \frac{4\pi}{3}R_E^3(\rho_0 - \rho_\infty). \quad (78)$$

For the simple capillary model, one has that $\Delta N = \frac{4\pi}{3}R^3(\rho_0 - \rho_\infty)$ and $R_E = R$. For the extended model, a simple calculation gives

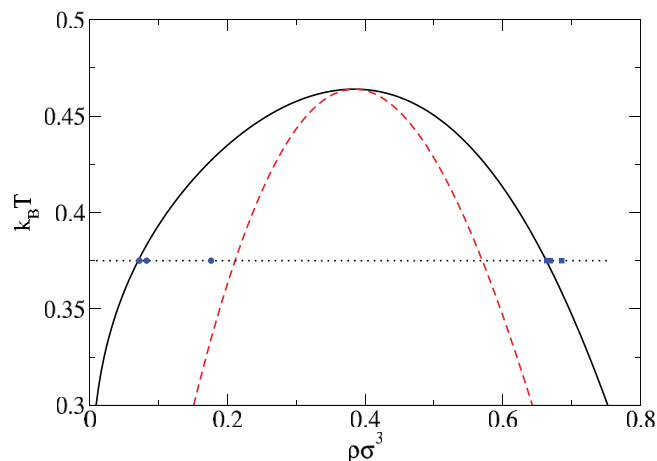$$\Delta N = \frac{\pi}{3w}(R^4 - (\max(R - w, 0))^4)(\rho_0 - \rho_\infty). \quad (79)$$

FIG. 1. The phase diagram for the model globular protein. The full lines show the dilute/dense solution coexistence, the broken lines are the spinodal and the horizontal line shows the temperature used in the calculations.

A. Energy of formation of a cluster

The dilute-solution/dense-solution phase diagram is shown in Fig. 1. By definition, the coexistence concentrations correspond to saturation $S = 1$ and at the spinodal the supersaturation is found to be $S = 2.99$. We will therefore illustrate the results of the models for supersaturations from $S = 1.025$ to $S = 2.5$ corresponding, as will be seen, to quite large and very small critical clusters, respectively.

Figure 2 shows the energy of formation of a cluster as a function of its size for different values of supersaturation.

The figures show the energy as determined using CNT (capillary model with the value of $\gamma = \gamma_{CNT}$ fixed to give the correct surface tension at coexistence ($S = 1$), Eq. (69)), the same model but with a supersaturation-dependent value of γ (Eq. (68)), the extended model with fixed width and the extended model minimized with respect to the width. It is clear in all cases that the capillary model with fixed γ gives lower free energies than when γ is allowed to vary and that it is also in closer agreement with the finite-width model. This is somewhat counter intuitive. On the other hand, the extended model with fixed width gives virtually the same results as when the energy is minimized with respect to the width, showing that the simple fixed-width model is adequate. For these reasons, only the capillary model with fixed $\gamma = \gamma_{CNT}$ and the extended model with fixed width will be used below.

B. The stationary distribution

The exact stationary distribution is given in Eq. (24). Although both the capillary model and the extended model depend on a single variable, a "radius" denoted R in both cases, the meaning of the parameter is not the same. In order to make a meaningful comparison, we therefore give the stationary distribution using the equimolar radius as the independent variable and noting that

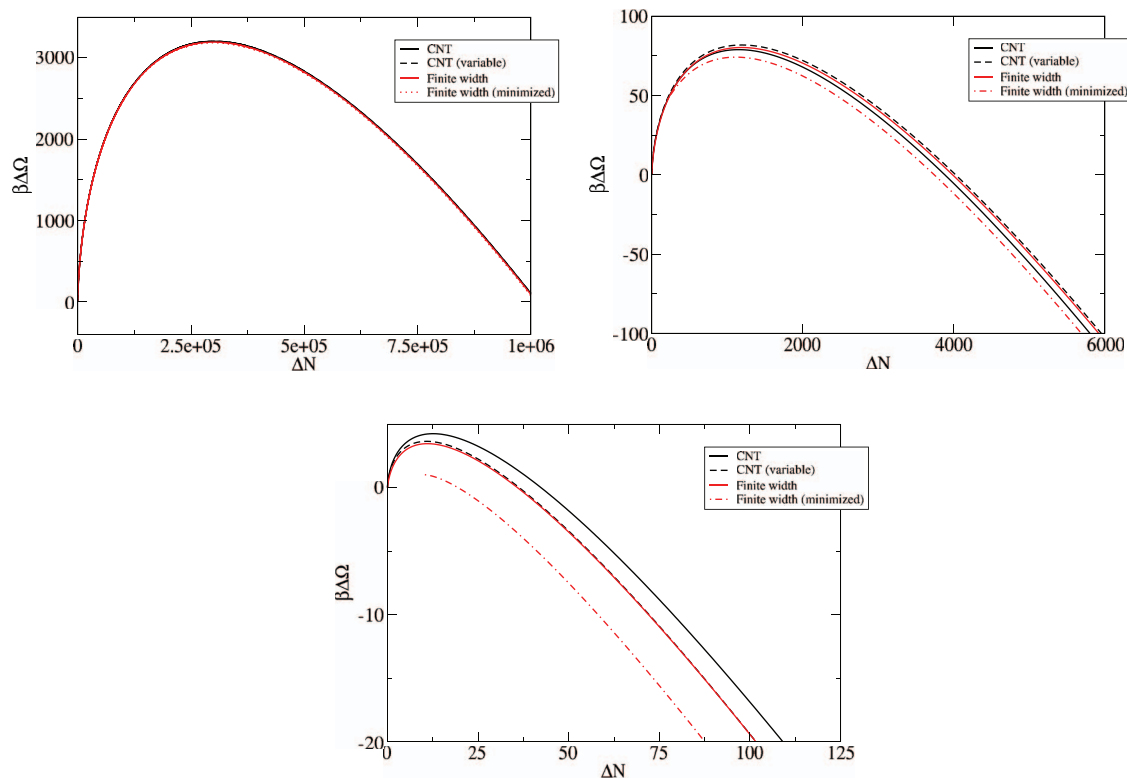$$\bar{P}(R_E) = P(R)\frac{dR}{dR_E}. \tag{80}$$

FIG. 2. The free energy as a function of cluster size (excess number of molecules in cluster) at $S = 1.025$, 1.175, and 2.5 as calculated using the capillary model with fixed γ, Eqs. (39) and (69), the capillary model with variable γ (Eq. (68)), the extended model with fixed width, Eqs. (61) and (65), and the extended model minimized with respect to the width.
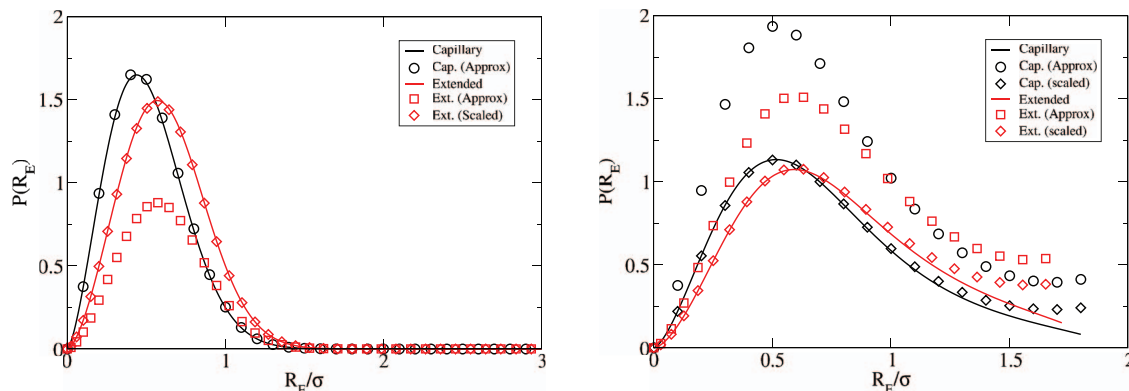
FIG. 3. The stationary distribution for $S = 1.025$ (left panel) and $S = 2.5$ (right panel). Exact results, i.e., evaluations of Eq. (24) are shown for the capillary and extended models in units such that $\sigma = 1$. Also shown is the approximation, Eq. (36), and the approximate distribution scaled so as to have the same maximum as the exact result.

For the capillary model $R = R_E$, so that $\bar{P}(R_E) = P(R)$. For the extended model, we find that

$$\frac{dR_E}{dR} = \frac{R^3 - (max(R - w, 0))^3}{3wR_E^2}. \tag{81}$$

Figure 3 shows the stationary distributions as calculated in CNT and using the extended model for different values of the supersaturation. Also shown are the approximate distribution used to evaluate the nucleation rate. It is apparent that the approximate distribution has the right shape so that it is indeed the case that the stationary distribution is well approximated by the "equilibrium" distribution, $g(X)e^{-\beta\Delta\Omega}$. However, in the case of the extended model, the approximate evaluation of the normalization of the distribution is poor due to the rapidly changing analytic structure of the free energy as a function of cluster radius for small clusters. A surprising result is that the distributions for the capillary model and for the extended model differ even for small supersaturation. This is simply a reflection of the fact that the differences induced by the two models are most pronounced for small clusters regardless of the supersaturation.

The approximate forms for the nucleation rates given in Eqs. (50) and (73) are not calculated from the exact stationary distribution but, rather, from the approximation given in Eq. (A6). The figures show that for CNT this approximation is

quite good at low supersaturations but is in considerable error for small clusters. The reason for the increasing error is illustrated in Fig. 4 which shows the convergence of the "exact" expression, Eq. (35), as a function of the domain of integration. For large supersaturations, the convergence is slow indicating that larger clusters are contributing significantly to the evaluation of the nucleation rate. That is to say, that there is a substantial population of larger clusters in contradiction to the assumption that the population is dominated by monomers. For large supersaturations, this calls into question both the monomer attachment picture that underlies CNT and the assumption of non-interacting clusters that is tacitly used in both CNT and the dynamical theory thus pointing to the internal inconsistency of the theory. The main conclusion to be drawn from this result is that even if one could accurately calculate the free energy and distribution of clusters at large supersaturations, it is most likely the case that cluster-cluster interactions would invalidate the assumptions necessary to calculate the nucleation rate, both in CNT and in the general framework described here.

C. The nucleation rate

Table I gives the nucleation rate as calculated for the two models. For the lowest supersaturation, the extended model
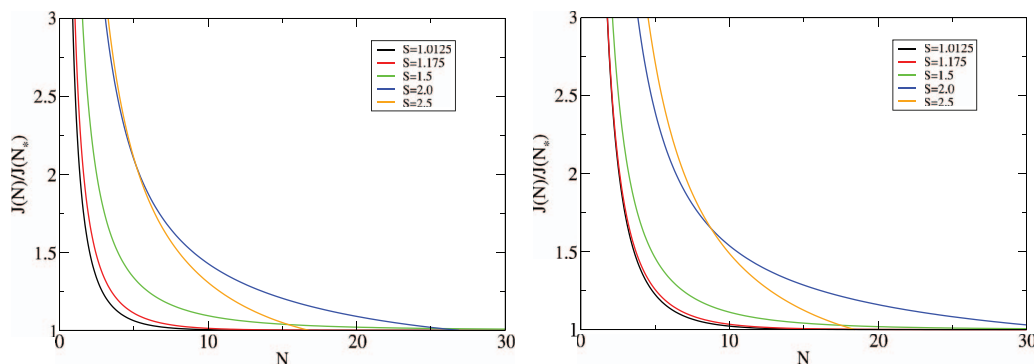

FIG. 4. The convergence of $J(N)$, Eq. (35), as a function of N for the capillary model (left panel) and the extended model (right panel) for different values of the supersaturation. Note that in the case of the highest supersaturations, the nucleation rate does not reach a stable value before the integral extends up to the size of the critical cluster, N_*.

TABLE I. Properties of the capillary and extended cluster models as functions of the supersaturation, S. The classical nucleation rate, J_{CNT} is given followed by the critical equimolar radius, R_{*E}, the excess mass, ΔN_*, and excess free energy, $\Delta \Omega_*$, of the critical cluster. Also given are the nucleation rates as calculated from the "exact" expression, Eq. (35), and the approximate expression, Eq. (37), J and J_{app}, respectively. Lengths are in units of σ and times are in units of σ^2/D.

		Capillary					Extended model				
S	$J_{CNT}e^{\beta\Delta\Omega}$	R_{*E}	ΔN_*	$\Delta\beta\Omega_*$	$\frac{J}{J_{CNT}}$	$\frac{J_{app}}{J_{CNT}}$	R_{*E}	ΔN_*	$\Delta\beta\Omega_*$	$\frac{J}{J_{CNT}}$	$\frac{J_{app}}{J_{CNT}}$
1.025	1.7×10^{-4}	49.5	300741	3205	0.4	0.37	49.5	300297	3195	6396	5564
1.175	1.5×10^{-3}	7.8	1149	79	0.84	0.93	7.9	1188	80	0.09	0.11
1.5	5.7×10^{-3}	3.27	83.6	14.0	1.03	1.45	3.39	92.9	14.9	0.22	0.40
2.0	1.5×10^{-2}	2.1	21.6	5.9	0.7	1.9	2.2	22.9	5.9	0.5	1.6
2.5	2.8×10^{-2}	1.8	12.6	4.3	0.9	2.2	1.7	11	3.5	1.1	3.8

gives a much higher nucleation rate. This is almost entirely attributable to the fact that the free energy of the critical cluster is about $10 k_B T$ lower in the extended model than in the capillary model (and, indeed, the log of the ratio of the nucleation rates is about 9.7). In all other cases, the nucleation rates of the two models are comparable and similar to that given by the classical CNT formula.

VI. CONCLUSIONS

In this paper, an alternative approach to the semi-phenomenological description of nucleation has been presented that is based on a systematic development of the nonequilibrium dynamics describing thermal fluctuations.[10–12] The theory admits of a reduced description whereby density fluctuations, which is to say unstable clusters, are characterized by one or more order paramters and one of the goals of the present work has been to demonstrate that in the simplest limit, when clusters are characterized by a single order parameter, one recovers Classical Nucleation Theory. Another goal has been to show that, because the theory is the result of a systematic development some seemingly arbitrary choices made in the heuristic development of CNT, such as the choice of reaction coordinate, are resolved. For the example of the reaction coordinate, the systematic theory is a priori covariant making the physical results independent of the chosen reaction coordinate.

Furthermore, because the theory is based on a more general framework, it is straightforward to go beyond the assumptions underlying CNT. One could, e.g., allow for multiple order parameters so that aside from cluster size, the interfacial width and the density within the cluster could be treated dynamically. We intend to discuss such generalizations at a later time. Here, working within the spirit of CNT, we have explored a simpler generalization whereby clusters are allowed to have a finite interfacial width. This gives rise to sub-dominant corrections to the free energy of the type described by Prestipino et al.[8] It was, nevertheless, not surprising that we found the dominant effect on the nucleation rate was simply due to the difference in the free energy of the critical cluster in the different models (i.e., with zero and finite interfacial width), see, e.g., Table I. Thus, despite the fact that the model for the structure of a cluster affects all aspects of

the dynamics in our general approach, it was found that, as is commonly assumed, only the correction to the free energy barrier is really important in determining the nucleation rate. This provides post hoc justification for the combination of the framework of CNT with the use, e.g., of Density Functional Theory to get better estimates for the free energy barrier.

The nucleation rates were evaluated under the usual scenario of a steady state whereby post-critical clusters above a certain size are removed and replaced my monomers. We explicitly calculated the stationary distribution of cluster sizes and it was shown that the stationary distribution is well approximated, for small clusters, by the "equilibrium" distribution as is usually assumed in the determination of nucleation rates. However, at high supersaturations, it was found that the stationary distribution becomes quite broad and that the nucleation rate becomes dependent on the domain of integration over the cluster sizes. This indicates that the assumption of cluster growth by monomer attachment and detachment is probably invalid as is the assumption of non-interacting clusters. It is therefore the case that quantitative prediction of nucleation rates is not really possible with the theory discussed here. For the conditions considered here, the breakdown occurs when the critical cluster contains between 20 and 100 molecules. A cluster of 100 molecules is still very small and our results illustrate, in some detail, the fact that Classical Nucleation Theory is remarkably robust given the crude assumptions that go into it.

There are, of course, many other important physical effects that we have ignored, as they are in Classical Nucleation Theory. These include the variation of the density within a cluster with cluster size, variation of surface tension with cluster size, the non-spherical shape of small clusters, capillary waves, etc. Most of these can in fact be addressed within the context of the more general field theory that underlies the reduced description in terms of order parameters used here (see, e.g., Ref. 22). The point of the present study has been a critical and detailed examination of the limits placed by the requirement of internal consistency of CNT and of CNT-like theories. If these break down beyond a certain point — e.g., for sufficiently high supersaturations — then it is not necessary to invoke additional physics — such as nonspherical clusters — to say that these theories are inadequate. Instead, the present results provide an additional criterion to be considered when arguing for or against the applicability of CNT to a particular system.

ACKNOWLEDGMENTS

APPENDIX A: APPROXIMATE NUCLEATION RATE

The nucleation rate is

$$J^{-1} = \rho_\infty^{-1} D^{-1} \int_0^{X_*} N(X) g^{1/2}(X) \exp(-\beta\Omega(X))$$
$$\times \left(\int_X^{X_+} g^{1/2}(X') \exp(\beta\Omega(X')) dX' \right) dX \quad \text{(A1)}$$

and is the exact expression for the nucleation rate given the assumptions of the theory. In order to develop an approximation it is convenient to switch to the canonical variable so that the nucleation rate becomes

$$J^{-1} = \rho_\infty^{-1} D^{-1} \int_0^{Y_*} N(Y) \exp(-\beta\widetilde{\Omega}(Y))$$
$$\times \left(\int_Y^{Y_+} \exp(\beta\widetilde{\Omega}(Y')) dY' \right) dY. \quad \text{(A2)}$$

The first step is to assume that the free energy has a maximum at Y_* which is defined by

$$\left. \frac{\partial\widetilde{\Omega}(Y)}{\partial Y} \right|_{Y_*} = 0, \quad \text{(A3)}$$

so that we can evaluate the inner integral by expanding

$$\exp(\beta\widetilde{\Omega}(Y')) \sim \exp\left(\beta\widetilde{\Omega}(Y_*) + \frac{1}{2}\beta\widetilde{\Omega}''(Y_*)(Y' - Y_*)^2 \right) \quad \text{(A4)}$$

giving

$$J^{-1} \sim \rho_\infty^{-1} D^{-1} \exp(\beta\widetilde{\Omega}(Y_*)) \sqrt{\frac{2\pi}{|\beta\widetilde{\Omega}''(Y_*)|}} \int_0^{Y_*} N(Y)$$
$$\times \exp(-\beta\widetilde{\Omega}(Y)) dY. \quad \text{(A5)}$$

Before proceeding, note that this is really an approximation to the stationary distribution of

$$\widetilde{P}(Y) \sim J D^{-1} \exp(\beta\widetilde{\Omega}(Y_*)) \sqrt{\frac{2\pi}{|\beta\widetilde{\Omega}''(Y_*)|}} \exp(-\beta\widetilde{\Omega}(Y)), \quad \text{(A6)}$$

or, in general,

$$P(X) \sim J D^{-1} \exp(\beta\Omega(X_*)) \sqrt{\frac{2\pi g(X) g(X_*)}{|\beta\Omega''(X_*)|}} \exp(-\beta\Omega(X)). \quad \text{(A7)}$$

To further simplify, we assume that the free energy has a minimum at $Y = 0$ and that $\beta\Delta\widetilde{\Omega}(Y) = aY^\alpha + \dots$ and $N(Y)$

$= N_0 Y^\beta + \dots$ for some values of $\alpha, \beta > 0$ giving

$$J^{-1} \sim \rho_\infty^{-1} D^{-1} \exp(\beta\widetilde{\Omega}(Y_*)) \sqrt{\frac{2\pi}{|\beta\widetilde{\Omega}''(Y_*)|}}$$
$$\times \int_0^{Y_*} N_0 Y^\beta \exp(-aY^\alpha) dY$$
$$\sim \rho_\infty^{-1} D^{-1} \exp(\beta\Delta\widetilde{\Omega}(Y_*)) \sqrt{\frac{2\pi}{|\beta\widetilde{\Omega}''(Y_*)|}} N_0 \frac{1}{\alpha} a^{-\frac{\beta+1}{\alpha}}$$
$$\times \int_0^\infty Z^{\frac{\beta+1-\alpha}{\alpha}} \exp(-Z) dZ$$
$$= \rho_\infty^{-1} D^{-1} \exp(\beta\Delta\widetilde{\Omega}(Y_*)) \sqrt{\frac{2\pi}{|\beta\widetilde{\Omega}''(Y_*)|}}$$
$$\times N_0 \frac{1}{\alpha} a^{-\frac{\beta+1}{\alpha}} \Gamma\left(\frac{\beta+1}{\alpha}\right) \quad \text{(A8)}$$

or

$$J \sim D\rho_\infty \sqrt{\frac{1}{2\pi}} a^{\frac{1+\beta}{\alpha}} \frac{\alpha N_0^{-1}}{\Gamma\left(\frac{1+\beta}{\alpha}\right)} \left|\beta\widetilde{\Omega}''(Y_*)\right|^{1/2} \exp(-\beta\Delta\widetilde{\Omega}(Y_*)). \quad \text{(A9)}$$

It is easy to translate into the original variables to find

$$J \sim D\rho_\infty \sqrt{\frac{1}{2\pi}} a^{\frac{1+\beta}{\alpha}} \frac{\alpha N_0^{-1}}{\Gamma\left(\frac{1+\beta}{\alpha}\right)} \left|\beta\Omega''(R_*) g^{-1}(R_*)\right|^{1/2}$$
$$\times \exp(-\beta\Delta\widetilde{\Omega}(R_*)). \quad \text{(A10)}$$

APPENDIX B: EXPANSION OF THE FREE ENERGY

Beginning with the model expression for the free energy,

$$\Omega(R; w) - \Omega_\infty = \frac{4\pi}{3}(R - w)^3 \Delta\omega(\rho_0)$$
$$+ 2\pi \left(2\overline{\omega}_0 w + K\frac{(\rho_0 - \rho_\infty)^2}{w} \right) R^2$$
$$- 2\pi \left(4\overline{\omega}_1 w + K\frac{(\rho_0 - \rho_\infty)^2}{w} \right) Rw$$
$$+ 2\pi \left(2\overline{\omega}_2 w + \frac{1}{3}K\frac{(\rho_0 - \rho_\infty)^2}{w} \right) w^2, \quad \text{(B1)}$$

we minimize by setting the derivative with respect to the width, w, equal to zero giving

$$0 = -2(R - w)^2 \Delta\omega(\rho_0) + \left(2\overline{\omega}_0 - K\frac{(\rho_0 - \rho_\infty)^2}{w^2} \right)$$
$$\times R^2 - 8\overline{\omega}_1 wR + 6\overline{\omega}_2 w^2 + \frac{1}{3}K(\rho_0 - \rho_\infty)^2, \quad \text{(B2)}$$

or, after dividing by R^3 and rearranging,

$$0 = ((-2\Delta\omega(\rho_0) + 2\overline{\omega}_0)w^2 - K(\rho_0 - \rho_\infty)^2)$$
$$+ (4\Delta\omega(\rho_0) - 8\overline{\omega}_1)w^2 \frac{w}{R}$$
$$+ \left(-2\Delta\omega(\rho_0)w^2 + 6\overline{\omega}_2 w^2 + \frac{1}{3}K(\rho_0 - \rho_\infty)^2\right)\left(\frac{w}{R}\right)^2.$$
$$(B3)$$

We now expand in inverse powers of the radius,

$$w = w_0\left(1 + a_1\frac{w_0}{R} + a_2\left(\frac{w_0}{R}\right)^2 + ...\right) \qquad (B4)$$

and solve order by order in $\frac{w_0}{R}$ giving

$$w_0^2 = \frac{K(\rho_0 - \rho_\infty)^2}{2(-\Delta\omega(\rho_0) + \overline{\omega}_0)}$$
$$a_1 = \frac{\Delta\omega(\rho_0) - 2\overline{\omega}_1}{\Delta\omega(\rho_0) - \overline{\omega}_0} \qquad (B5)$$
$$a_2 = \frac{1}{6(\Delta\omega(\rho_0) - \overline{\omega}_0)^2}\big(11(\Delta\omega(\rho_0))^2$$
$$+ \Delta\omega(\rho_0)(5\overline{\omega}_0 - 60\overline{\omega}_1 + 9\overline{\omega}_2) - 9\overline{\omega}_0\overline{\omega}_2 - \overline{\omega}_0^2 + 60\overline{\omega}_1^2\big)$$

and

$$\Omega(R; w) - \Omega_\infty$$
$$= \left(\frac{4}{3}\pi R^3\right)\Delta\omega(\rho_0) + \left(4\pi R^2\right)\frac{K(\rho_0 - \rho_\infty)^2}{w_0}$$
$$+ (4\pi R)w_0^2(2\Delta\omega(\rho_0) - \overline{\omega}_0 - 2\overline{\omega}_1) + O(R^0). \quad (B6)$$

Of course, we could solve the equation for the width exactly since it is simply a fourth order polynomial in w,

$$0 = -K(\rho_0 - \rho_\infty)^2 + \left[\frac{K(\rho_0 - \rho_\infty)^2}{3R^2} + 2\overline{\omega}_0 - 2\Delta\omega(\rho_0)\right]w^2$$
$$+ \left[4\frac{\Delta\omega(\rho_0) - 2\overline{\omega}_1}{R}\right]w^3 + \left[\frac{6\overline{\omega}_2 - 2\Delta\omega(\rho_0)}{R^2}\right]w^4.$$
$$(B7)$$

We can also express the free energy in terms of the equimolar radius

$$R_E = \left(\frac{1}{4w}\left(R^4 - (R-w)^4\right)\right)^{1/3}$$
$$= R - \frac{1}{2}w_0 + \frac{w_0}{12}(1 - 6a_1)\frac{w_0}{R}$$
$$+ \frac{w_0}{24}(4a_1 - 12a_2 + 1)\left(\frac{w_0}{R}\right)^2 + O\left(\frac{w_0}{R}\right)^3, \quad (B8)$$

we find that

$$R = R_E + \frac{w_0}{2} + \frac{w_0}{12}(6a_1 - 1)\frac{w_0}{R_E}$$
$$+ \frac{w_0(6a_2 - 5a_1)}{12}\left(\frac{w_0}{R_E}\right)^2 + O\left(\left(\frac{w_0}{R_E}\right)^3\right) \quad (B9)$$

giving

$$\Omega(R_E; w) - \Omega_\infty$$
$$= \Delta\omega(\rho_0)\left(\frac{4\pi}{3}R_E^3\right)$$
$$+ w_0\left(\frac{K(\rho_0 - \rho_\infty)^2}{w_0^2} + \frac{\Delta\omega(\rho_0)}{2}\right)\left(4\pi R_E^2\right)$$
$$+ w_0^2\bigg(6\frac{K(\rho_0 - \rho_\infty)^2}{w_0^2} + (13 + 3a_1)\Delta\omega$$
$$- 6\overline{\omega}_0 - 12\overline{\omega}_1\bigg)\frac{2\pi}{3}R_E + O(R^0). \quad (B10)$$

[1] D. Kashchiev, *Nucleation: Basic Theory with Applications* (Butterworth-Heinemann, Oxford, 2000).
[2] R. Becker and W. Döring, Ann. Phys. **2416**, 719 (1935).
[3] P. Rien ten Wolde and D. Frenkel, J. Chem. Phys. **109**, 9901 (1998).
[4] D. W. Oxtoby and R. Evans, J. Chem. Phys. **89**, 7521 (1988).
[5] J. F. Lutsko, J. Chem. Phys. **129**, 244501 (2008).
[6] S. Ghosh and S. K. Ghosh, J. Chem. Phys. **134**, 024502 (2011).
[7] A. J. Archer and R. Evans, Mol. Phys. **109**, 2711 (2011).
[8] S. Prestipino, A. Laio, and E. Tosatti, Phys. Rev. Lett. **108**, 225701 (2012).
[9] W. Lechner, C. Dellago, and P. G. Bolhuis, Phys. Rev. Lett. **106**, 085701 (2011).
[10] J. F. Lutsko, J. Chem. Phys. **135**, 161101 (2011).
[11] J. F. Lutsko, J. Chem. Phys. **136**, 034509 (2012).
[12] J. F. Lutsko, J. Chem. Phys. **136**, 134502 (2012).
[13] L. Landau and E. Lifshitz, *Fluid Mechanics* (Pergamon, New York, 1975).
[14] P.-H. Chavanis, Physica A **387**, 5716 (2008).
[15] C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer, Berlin, 2004).
[16] P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).
[17] E. Lifshitz and L. Pitaevskii, *Physical Kinetics. Vol. 10: Course of Theoretical Physics* (Elsevier, Amsterdam, 2010).
[18] J. S. Langer, Ann. Phys. **54**, 258 (1969).
[19] J. S. Langer and L. A. Turski, Phys. Rev. A **8**, 3230 (1973).
[20] Y. Saito, *Statistical Physics of Crystal Growth* (World Scientific, Singapore, 1998).
[21] J. F. Lutsko, J. Chem. Phys. **134**, 164501 (2011).
[22] J. F. Lutsko, J. Chem. Phys. **137**, 154903 (2012).