David M. Huang and Phil Attard

A random walk through the dynamics of homogeneous vapor-liquid nucleation

Journal of Chemical Physics, 2005; 122(17): 174503-1-174503-8

© 2005 American Institute of Physics

This article may be downloaded for personal use only. Any other use requires prior permission of the author and AIP Publishing. This article appeared in Journal of Chemical Physics, 2005; 122(17): 174503-1-174503-8 and may be found at http://dx.doi.org/10.1063/1.1884086

PERMISSIONS

https://publishing.aip.org/resources/researchers/rights-and-permissions/sharing-content-online/

For institutional or funder-designated repositories (e.g., DOE Pages)

- You may deposit the accepted manuscript immediately after acceptance, using the credit line formatting below
- You may deposit the VOR 12 months after publication, with the credit line and a link to the VOR on AIP Publishing’s site

Format for credit lines

- After publication please use: “This article may be downloaded for personal use only. Any other use requires prior permission of the author and AIP Publishing. This article appeared in (citation of published article) and may be found at (URL/link for published article abstract).
- Prior to publication please use: “The following article has been submitted to/accepted by [Name of Journal]. After it is published, it will be found at Link.”
- For Creative Commons licensed material, please use: “Copyright (year) Author(s). This article is distributed under a Creative Commons Attribution (CC BY) License.”

16 June 2021

http://hdl.handle.net/2440/64427
A random walk through the dynamics of homogeneous vapor-liquid nucleation

David M. Huang, and Phil Attard

ARTICLES YOU MAY BE INTERESTED IN

Optical description of solid-state dye-sensitized solar cells. II. Device optical modeling with implications for improving efficiency
Journal of Applied Physics 106, 073112 (2009); https://doi.org/10.1063/1.3204985

Phase transitions in systems small enough to be clusters
The Journal of Chemical Physics 118, 340 (2003); https://doi.org/10.1063/1.1524192

Perspective: Crossing the Widom line in no man’s land: Experiments, simulations, and the location of the liquid-liquid critical point in supercooled water
The Journal of Chemical Physics 149, 140901 (2018); https://doi.org/10.1063/1.5046687
A random walk through the dynamics of homogeneous vapor-liquid nucleation

David M. Huang and Phil Attard

School of Chemistry, University of Sydney, New South Wales 2006, Australia

(Received 11 January 2005; accepted 8 February 2005; published online 29 April 2005)

A method of calculating rates of homogeneous vapor-liquid nucleation based on Langevin dynamics of a few relevant degrees of freedom on a free-energy surface is proposed. The surface is obtained here from simulation and from a semi empirical expression. The mass and friction coefficients are derived from atomistic umbrella-sampling molecular-dynamics simulations. The calculated nucleation rate agrees with atomistic simulations for one particular state point of the Lennard-Jones fluid. The present method is about four orders of magnitude more computationally efficient than the direct atomistic simulation of the transmission coefficient. © 2005 American Institute of Physics. [DOI: 10.1063/1.1884086]

I. INTRODUCTION

Classical nucleation theory (CNT) provides a simple analytic expression for the rate of homogeneous vapor-liquid nucleation by relying on several assumptions. One such assumption is that the growing liquid droplet is spherical in shape and of uniform liquid density. The free energy of nucleation at a given temperature and supersaturation is a function of a single coordinate, such as the droplet radius or number of molecules in the droplet, and depends on the chemical potential difference between the coexistence and supersaturated vapor phases, and the interfacial tension between the two phases. Further simplifications are made for the nucleation rate, namely, that it depends upon the free-energy barrier height, a steady flux over the barrier, growth by monomer addition, ideal gas kinetics for the vapor, and complete removal of product.

Although CNT has been found to reproduce experimental estimates of the size of critical nuclei, for many of the same experiments it does not predict the nucleation rate as accurately.1 This result could be due to the neglect in CNT of many of the same experiments. The present method is about four orders of magnitude more computationally efficient than the direct atomistic simulation of the transmission coefficient. © 2005 American Institute of Physics.

II. BEYOND CLASSICAL NUCLEATION THEORY

According to classical nucleation theory, the Gibbs free energy \( \Delta G_{\text{CNT}}(R) \) of a spherical droplet of liquid density \( \rho_l \) and radius \( R \) in a supersaturated vapor at a pressure \( p \) is

\[
\Delta G_{\text{CNT}}(R) = \Delta \mu \rho_l \frac{4 \pi}{3} R^3 + \gamma' 4 \pi R^2.
\]

This is measured relative to the uniform vapor phase, and \( \Delta \mu = \mu_l(p) - \mu_v(p) \) is the difference between the chemical potential of the liquid \( \mu_l \) and that of the vapor \( \mu_v \) at the pressure \( p \), and \( \gamma' \) is the liquid-vapor surface tension for a planar interface at liquid-vapor coexistence. The second term in Eq. (1), which is positive, dominates for small \( R \), while the first term, which is negative, dominates for large \( R \). Therefore, \( \Delta G_{\text{CNT}}(R) \) starts off positive, has a maximum, and eventually becomes negative for very large values of \( R \).

The CNT expression for the nucleation free energy, Eq. (1), can be generalized to include changes in the droplet density \( \rho \) (or some other degree of freedom) as well as the droplet radius \( R \).

In order to do this, we consider the constrained thermodynamic potential

\[
\mathcal{F}(N, A, V, p, \gamma, T) = F(N, A, V, T) + pV - \mu N + \gamma A,
\]

where \( \mu = \mu(\rho_v, T) \) and \( p = p(\rho_v, T) \) are, respectively, the chemical potential and pressure of a reservoir of density \( \rho_v \) at a temperature of \( T \), and \( \gamma = \gamma(\rho_v, T) \) is the surface tension between the reservoir and a subsystem of volume \( V \) containing \( N \) particles (\( \gamma \) is assumed to be independent of geometry). This thermodynamic potential is equal to difference in Helmholtz free energy between a total system of volume \( V_{\text{tot}} \) and particle number \( N_{\text{tot}} \) consisting of a subsystem of volume \( V \) (and surface area \( A \)) containing \( N \) particles and a reservoir of density \( \rho_v \) at temperature \( T \) and a total system of uniform density \( \rho_v \) at temperature \( T \).

We can show that

\[
\frac{\partial \mathcal{F}(N, A, V)}{\partial N} \bigg|_{N/V=\rho_v} = 0,
\]

and, therefore, that

\[
\mathcal{F}(N = \rho_v V, A, V = \text{const}) = \text{const}.
\]

But \( \gamma A \) is identically zero if both the subsystem and reservoir have density \( \rho_v \), and so

\[
\mathcal{F}(N = \rho_v V, A, V = \text{const}) = \text{const}
\]

is true for all \( A \). By setting \( N = V = 0 \), it is easy to show that this constant is zero. In other words, if \( N/V = \rho_v \), the total
system consists entirely of vapor and therefore this constrained thermodynamic potential, which is the total Helmholtz free-energy relative to a total system of uniform density \( \rho_c \) at temperature \( T \), must be zero.

To make it clear that we are considering changes in free energy in going from a uniform system of density \( \rho_c \) to a nonuniform system consisting of \( N \) particles in a volume \( V \) surrounded by a reservoir of density \( \rho_o \), we define the quantity

\[
\Delta F(N,V,A) = F(N,V,A) - F(N = \rho_c V, V,A)
\]

\[
= F(N,V,A). \tag{6}
\]

If we take \( \rho = \rho_i \), the liquid density, \( \Delta F(N = \rho_i V, V,A | \mu, p, \gamma, T) \) is the free-energy change to nucleate a droplet of volume \( V \) (and surface area \( A \)) in a vapor of density \( \rho_o \).

We now make a couple assumptions. Firstly, we assume that the free energy of the subsystem is equal to the free-energy density of a uniform system of density \( \rho \) at temperature \( T \) multiplied by the volume of the system,

\[
F(N = \rho V, V,A) = f(\rho, T)V. \tag{7}
\]

Therefore,

\[
\Delta F(N,V,A | \mu, p, \gamma, T) = \Delta F(\rho = N/V, V,A | \mu, p, \gamma, T)
\]

\[
= f(\rho, T)V + pV - \mu p V + \gamma(\rho, \rho_o)A. \tag{8}
\]

Secondly, we assume that the surface tension is given by the interpolation formula,\(^3\)

\[
\gamma(\rho, \rho_o) = \gamma^i \left( \frac{S^0 - S}{S^0 - 1} \right) \left( \frac{\rho - \rho_o}{\rho^i - \rho_o} \right), \tag{9}
\]

where \( S^0 = \rho^i / \rho^o \) and \( S = \rho_c / \rho^i \), and the superscript “\( i \)” denotes the spinodal while the superscript “\( o \)” denotes liquid-vapor coexistence. This equation gives \( \gamma(\rho_c, \rho_o) = 0 \), \( \gamma(\rho^i, \rho^o) = \gamma^i \), and \( \gamma(\rho, \rho_c) = 0 \). We also specialize to a spherical geometry; i.e., \( V = 4\pi R^3 / 3 \) and \( A = 4\pi R^2 \), which leads to

\[
\Delta F(\rho, R) = \left[ f(\rho, T) + \mu p \right] \frac{4\pi R^3}{3} + \gamma(\rho, \rho_o)4\pi R^2. \tag{10}
\]

With \( f(\rho, T) \) from an equation of state for the fluid and \( \gamma^i \) from simulation or experiment, \( \Delta F(\rho, R) \) can be calculated as a function of \( \rho \) and \( R \). For the Lennard-Jones (LJ) fluid, for example, accurate equations of state are available\(^3\) and so is simulation data for the liquid-vapor surface tension. The pathway to nucleation in classical nucleation theory (CNT) is the slice through this surface at \( \rho = \rho_i \), although the CNT free-energy values use \( \gamma^i \) instead of \( \gamma(\rho, \rho_c) \).

The free energy \( \Delta F(\rho, R) \) is plotted for the LJ fluid under two different sets of thermodynamic conditions in Fig. 1. Allowing fluctuations in the spherical droplet’s density \( \rho \) as well as its radius \( R \), it can be seen from Fig. 1 that close to the critical point—the same occurs at high supersaturation—that changes in \( \rho \) constant \( R \) cost little energy near the barrier to nucleation. This result indicates the potential importance of \( \rho \) in determining the nucleation rate.

Accurate empirical free-energy surfaces \( \Delta F(\mathbf{q}) \), in terms of a few relevant degrees of freedom, \( \mathbf{q} = (q_1, q_2, \ldots, q_N) \), such as those in Fig. 1, may provide a more accurate means than CNT for calculating nucleation rates while avoiding costly atomic simulations, provided a way of carrying out dynamics on the surface is available.

One shortcoming with this method is that it does not provide a way of determining a priori which degrees of freedom are relevant to the reaction. However, comparisons with the experimental or simulation results may indicate whether a particular choice of coordinates is suitable.

## III. Langevin Dynamics

One way of carrying out dynamics on the free-energy surface \( \Delta F(\mathbf{q}) \) is to use the Langevin equation,\(^6,7\)

\[
m_i \dot{q}_i(t) = -\sum_j \xi_{ij} [\mathbf{q}(t) - \mathbf{q}_i(t)] + f_i(\mathbf{q}(t)) + \sum_j \alpha_{ij} \dot{q}_j(t), \tag{11}
\]

where \( m_i \) is the mass corresponding to the coordinate \( q_i \), and \( \mathbf{q} = \mathbf{q}_i \). There are two contributions to the force on the coordinates \( \mathbf{q} \). The first is a force due to variations in the free-energy surface \( \Delta F \),

\[
f_i(\mathbf{q}) = -\frac{\partial \Delta F(\mathbf{q})}{\partial q_i}. \tag{12}
\]

The second is a random force which has a mean of zero and a delta-function correlation in time,
\( \langle \Gamma_j(t) \rangle = 0, \quad \langle \Gamma_j(t) \Gamma_j(t') \rangle = 2 \delta_{jj} \delta(t-t') \). \hspace{1cm} (13)

The friction coefficients \( \zeta_{ij} \) can be related to the strength of the random force via

\[
\zeta_{ij} = \frac{1}{k_B T} \sum_i \alpha_{ij},
\]

and to the diffusion coefficients \( D_{ij} \) via

\[
\sum_j \zeta_{ij} D_{ji} = \sum_j D_{ij} \zeta_{ji} = k_B T \delta_{ij},
\]

In principle, \( m_i \) and \( \zeta_{ij} \) can vary with \( q \). It should also be noted that \( m_i \) and \( \zeta_{ij} \) are masses and friction coefficients associated with generalized coordinates and do not have a simple interpretation in terms of atomic masses or the friction acting on individual particles.

The parameters \( m_i \) and \( \zeta_{ij} \) can, however, be estimated from molecular dynamics (MD) in which the time dependence of the coordinates \( q \) are measured. In order to obtain accurate statistics for \( m_i \) and \( \zeta_{ij} \) as functions of \( q \), it is also necessary to define the system to a narrow range of values of \( q \), particularly regions of configuration space that the system rarely explores, such as near free-energy barriers. This can be achieved, for example, with umbrella-sampling molecular dynamics.

From the computational point of view, the Langevin dynamics simulations for a few relevant degrees of freedom \( q \) are trivial compared to carrying out nucleation rate calculations directly in atomistic simulations. The relative efficiency in obtaining the parameters required for the Langevin equation from an atomistic simulation compared to simulating the rate directly is less clear. We note however that the quantities we report below were obtained from a MD trajectory of 50,000 time steps, and we estimate the relative statistical error in them and in the nucleation rate as less than 10%. This can be compared with the direct atomistic simulations of the rate constant using the reactive flux method by ten Wolde et al.\(^{17}\) who used 5000 independent trajectories of 4000 steps, or \( 2 \times 10^7 \) MD time steps in total, which gave a nucleation rate with a relative statistical error of 67%. On the basis of these numbers, to get the same statistical accuracy one would require \( 2 \times 10^4 \) times as many MD steps for the direct simulation of the transmission coefficient than for the diffusion constant and inertial mass. In other words, the present Langevin method is more than four orders of magnitude more efficient than the conventional direct simulation of the nucleation rate. Further it may be possible to estimate the Langevin parameters by analytic approximations or other means as functions of the coordinates \( q \) as well as of the thermodynamic conditions. If this can be done, the Langevin dynamics method offers even greater savings. As a first step towards implementing the method and understanding how \( m_i \) and \( \zeta_{ij} \) vary with \( q \) and the thermodynamic conditions, in this work these quantities have been obtained from atomistic simulations.

### IV. UMBRELLA-SAMPLING MOLECULAR DYNAMICS

#### A. General equations

For a system with total potential energy

\[
U_f[\mathbf{r}^N] = U_0[\mathbf{r}^N] + V[\mathbf{q}(\mathbf{r}^N)],
\]

where \( U_0[\mathbf{r}^N] \) is the potential energy of the reference (unbiased) system as a function of atomic coordinates \( \mathbf{r}^N \) and \( V[\mathbf{q}(\mathbf{r}^N)] \) is a biasing (umbrella) potential, which is a function of the coordinates \( q \), the force acting on a particle \( i \) in the system is

\[
\mathbf{F}_i = -\nabla_q U_f[\mathbf{r}^N] = -\nabla_q U_0[\mathbf{r}^N] - \nabla_q V[\mathbf{q}(\mathbf{r}^N)] = \mathbf{F}_{0i} + \mathbf{F}_{Vi}.
\]

#### B. Continuous cluster definition

In the nucleation problem, we are interested in coordinates like \( q=(\rho, R) \), or equivalently \( q=(N, R) \), where \( \rho, R, \) and \( N \) are the droplet density, radius, and particle number, respectively.

In order to apply Eq. (17), \( N \) and \( R \) must be differentiable functions of the atomic coordinates \( \mathbf{r}^N \). Therefore, continuous analogs of the more conventional “discrete” cluster definitions\(^{8,9}\) are required. We have taken

\[
N = \sum_i N_{2i}
\]

for the cluster particle number and

\[
R = \left( \frac{1}{N} \sum_i N_{2i} w_i |\mathbf{r}_i - \mathbf{r}_{cm}|^M \right)^{1/M}
\]

for the cluster radius, where the center of mass is

\[
\mathbf{r}_{cm} = \frac{1}{N} \sum_i N_{2i} w_i \mathbf{r}_i,
\]

the weight of an atom in the cluster is

\[
w_i = \frac{1}{N_{2i}} \sum_{j \neq i} f(|\mathbf{r}_i - \mathbf{r}_j|; q_c),
\]

and the continuous analog of the Heaviside function is

\[
f(r; q_c) = \{ 1 + \exp[-(q_c - r)/d] \}^{-1}.
\]

The weight \( w_i \) is a measure of the “contribution” of particle \( i \) to the cluster particle number (\( w_i=1 \) for a particle near the center of the cluster and \( 0 < w_i < 1 \) for a particle near the surface).

The upper limit on the sums, \( N_{2i} \), is the number of particles in the largest discrete cluster in the system, defined according to Stillinger’s criterion\(^5\) with cluster cutoff distance \( d_{c2} \). According to this criterion, two particles belong to the same cluster if separated by less than the cutoff distance (see Fig. 2). The quantities \( q_c, d, n_c, \) and \( M \) are parameters to be chosen. To obtain a physically meaningful value for \( N \), \( n_c \) is chosen such that \( w_i=1 \) for particles in the center of the cluster. \( q_c \) and \( d_{c2} \) are chosen such that the discontinuity in the force is small when a particle joins the discrete cluster.
This definition of the continuous cluster particle number \( N \) is the continuous analog of the “liquidlike” cluster particle number used by ten Wolde and Frenkel. Their liquidlike cluster criterion is almost the same as Stillinger’s criterion, except that only liquidlike particles are counted when determining whether two particles are part of the same cluster. A particle is liquidlike if it has more than a certain number of nearest neighbors (in their work this number was four), with particles defined as being nearest neighbors if separated by less than the cluster cutoff distance. According to this liquidlike cluster criterion, particles near the center of the cluster contribute a value of 1 to the particle number while particles near the surface may not contribute at all. Similarly, according to our continuous cluster definition, particles near the center contribute \( w_i \approx 1 \) to the particle number while particles near the surface contribute \( 0 < w_i < 1 \), with \( w_i \) approaching zero as the particle moves away from the cluster.

C. Force on the cluster

For the quadratic biasing potential

\[
V(N,R) = \frac{c_N}{2} (N-N_0)^2 + \frac{c_R}{2} (R-R_0)^2, \tag{23}
\]

the contribution from the biasing potential to the force on particle \( i \) is

\[
F_{vi} = -c_N(N-N_0)\nabla_i N - c_R(R-R_0)\nabla_i R, \tag{24}
\]

where

\[
\nabla_i N = \frac{1}{2} \sum_{n_c j \neq i} \nabla_i f'(r_{ij}) \tag{25}
\]

and

\[
\nabla_i R = \frac{1}{MNR^{M-1}} \left\{ Mw_i r_{ic}^{M-2} - \frac{1}{N} \sum_{j} w_j r_{ic}^{M-2} r_{jc} \right\} \\
+ \frac{1}{N} \sum_{n_c j \neq i} r_{ij} f'(r_{ij}) \left[ r_{ic}^{M-1} + r_{jc}^{M-1} - 2R^{M} \right] \\
- \frac{M}{N} \sum_{k} w_k r_{kc}^{M-2} (r_{ic} + r_{jc}) \right\}. \tag{26}
\]

Here, \( r_{ij} = r_i - r_j \), \( r_{ic} = r_i - r_{cm} \) (and similarly for \( r_{jc} \) and \( r_{kc} \)), \( r_{ij} \equiv |r_{ij}| \), and \( f'(r_{ij}) = df(r_{ij})/dr_{ij} \). Note that the force only acts on the \( N_{d2} \) particles in largest discrete cluster in the system.

V. APPLICATION TO THE LJ FLUID

The umbrella-sampling MD method of the previous section were applied to the LJ fluid with intermolecular potential truncated and shifted at 2.5\( \sigma \). All quantities hereafter are in reduced LJ units unless otherwise stated.

The MD simulations of 864 particles at a constant temperature of \( T = 0.741 \) and a constant pressure of \( P = 0.01205 \) were carried out using the extended system methods of Nosé, Hoover, and Andersen, and the reversible integration algorithm of Martyna et al. The MD time step used was 0.01\( \tau \), where \( \tau = \sqrt{\varepsilon/(m_0\sigma^2)} \).

It should be noted that Eq. (7) for the internal pressure \( P_{int} \) in Ref. 13 must be modified to include the contribution of the biasing potential \( \Delta P_{int} \). For \( V(N,R) \) given by Eq. (23), this contribution is

\[
\Delta P_{int} = \frac{1}{dV} \left\{ \frac{c_N(N-N_0)}{n_c} \sum_{i} \sum_{j
n_c} r_{ij} f'(r_{ij}) \right\} \\
+ \frac{c_R(R-R_0)}{MN\sigma^{M-1}} \left( M \sum_i w_i r_{ic}^M + \frac{1}{n_c} \sum_i \sum_{j
n_c} r_{ij} f'(r_{ij}) \right) \times \left[ r_{ic}^M - R^M - \frac{M}{N} \sum_k w_k r_{kc}^M (r_{ic} + r_{kc}) \right]. \tag{27}
\]

where \( V \) is the volume of the cubic simulation box and \( d \) is the dimensionality.

A. Equilibrium free energies

The umbrella-sampling MD method was tested by comparing the change in the Gibbs free energy for nucleation as a function of cluster particle number \( N \) with results obtained by ten Wolde and Frenkel from Monte Carlo (MC) simulations under the same thermodynamic conditions. For these calculations, the biasing potential acted only on \( N \) and not on \( R \); i.e., \( c_R = 0 \) in Eq. (23) or

\[
V(N) = \frac{c_N}{2} (N-N_0)^2. \tag{28}
\]

As explained in Sec. IV B, ten Wolde and Frenkel used a different, liquidlike, definition of the cluster particle number than the continuous definition used here. The liquidlike cluster number will be denoted hereafter as \( N_l \).

The parameters in Eqs. (21) and (22) were chosen so that \( N \) and \( N_l \) in our simulations matched as closely as possible over the entire range on \( N \) [\( n_c = 0.9 \), \( q = 1.5\sigma \) (same as the cutoff used by ten Wolde and Frenkel), \( q_{d2} = 2.0 \), and \( d = 0.02 \)]. Figure 3 shows a typical configuration near the free-energy barrier to nucleation from our simulations.

Following an equilibration period of 1 000 000 MD cycles, simulations of 3 000 000 MD cycles in length were carried out in each of 17 approximately evenly spaced windows for \( N \) spanning the range from 0 to approximately 350.
The umbrella-sampling MD algorithm does not work for small $N$, where the identity of the largest cluster can fluctuate between different clusters. So, for small $N$, a 10 000 000-cycle simulation without the biasing potential was carried out. The length of these simulations was chosen to match the statistical accuracy of the MC simulations of ten Wolde and Frenkel.\textsuperscript{9} They carried out simulations of 250 000 MC cycles in length in approximately the same number of umbrella-sampling windows as we used. However, the “diffusion constant” in the $N_l$ coordinate in the MC simulations was estimated to be 7–8 times as large as that in MD simulations with the same time step as ours.\textsuperscript{9} Therefore, it was necessary to carry out MD simulations approximately eight times as long to obtain the same accuracy.

Figure 4 shows contour plots of two-dimensional histograms $n(N, N_l)$ from two different simulations, one with the quadratic biasing potential centered at small $N$ and the other at large $N$. It can be seen from Fig. 4 that $N$ and $N_l$ match very closely at small $N$. At large $N$, $N$ is generally larger than $N_l$ and for a given value of $N$, $N_l$ can take quite a large range of values.

The joint probability distribution function $p_2(N, N_l)$ for the reference system without the biasing potential was calculated from the distribution function with the biasing potential, $\bar{p}_2(N, N_l)$, using

$$p_2(N, N_l) = \bar{p}_2(N, N_l) \frac{\exp[\beta V(N)]}{\langle \exp[\beta V(N)] \rangle_1}, \quad (29)$$

where subscript “1” notes an average for the biased system.

The full joint probability distribution function $p_2(N, N_l)$ was obtained for the entire range of $N$ and $N_l$ by using the multiple-histogram method\textsuperscript{14} generalized to two dimensions. The singlet distribution functions were then obtained by using

$$p_1(N) = \int dN_l p_2(N, N_l), \quad p_1(N_l) = \int dN p_2(N, N_l). \quad (30)$$

The Gibbs free energy was calculated using

$$\Delta G(q) = -k_B T \ln[p_1(q)], \quad q = N \text{ or } N_l. \quad (31)$$

Almost identical results were obtained when the singlet distribution function for the reference system, $p_1(q)$, where $q = N$ or $N_l$ was calculated directly from the singlet distribution with the biasing potential, $\tilde{p}_1(q)$, and joined together using the multiple-histogram method in one dimension. Since the joint distribution function $\tilde{p}_2(N, N_d)$ was not calculated in the simulations while the singlet distribution function $\tilde{p}_1(N_d)$ was, this latter method was used to obtain the free energy for $N_d$.

The Gibbs free energy as a function of $N$, $N_s$, and $N_l$ is shown in Fig. 5. The free energy as a function of $N$ does not extend to zero because we only calculated $N$ for the largest cluster in the system. It was instead assumed that $\Delta G(N)$ coincided with $\Delta G(N_l)$ for $N=N_l=15$.

$$\Delta G(q) = -k_B T \ln[p_1(q)], \quad q = N \text{ or } N_l. \quad (31)$$

Almost identical results were obtained when the singlet distribution function for the reference system, $p_1(q)$, where $q = N$ or $N_l$ was calculated directly from the singlet distribution with the biasing potential, $\tilde{p}_1(q)$, and joined together using the multiple-histogram method in one dimension. Since the joint distribution function $\tilde{p}_2(N, N_d)$ was not calculated in the simulations while the singlet distribution function $\tilde{p}_1(N_d)$ was, this latter method was used to obtain the free energy for $N_d$.

The Gibbs free energy as a function of $N$, $N_s$, and $N_l$ is shown in Fig. 5. The free energy as a function of $N$ does not extend to zero because we only calculated $N$ for the largest cluster in the system. It was instead assumed that $\Delta G(N)$ coincided with $\Delta G(N_l)$ for $N=N_l=15$.
Figure 5 shows that there is a discrepancy between our results for $\Delta G(N_i)$ and those of ten Wolde and Frenkel. However, this discrepancy is likely due to the extreme sensitivity of $\Delta G(N_i)$ to supersaturation under the particular set of thermodynamic conditions studied. (Fitting their data for the free-energy barrier $\beta \Delta G(N'_s)$ as a function of supersaturation $S$ to $\beta \Delta G(N'_s) = -397.59 + 2341.71 S - 4635.03 S^2 + 3240.05 S^3$, it can be seen that changing $S$ by just 2% from $S=1.53$ (where $P=0.01202$) to $S=1.50$ changes $\beta \Delta G(N'_s)$ from 57.5 to 63.5. Another possible contribution to the discrepancy between our results of $\Delta G(N_i)$ and those of ten Wolde and Frenkel could be the accumulation of small errors arising from the joining of nonoverlapping umbrella-sampling windows by eye in Ref. 9. This could be particularly problematic at small $N_i$ due to the steepness of the free-energy curve and to the statistical error in the data.

The crucial point to note from Fig. 5 is that $\Delta G(N_i) = \Delta G(N)$, for all $N$ and $N_i$. On the other hand $\Delta G(N'_i)$ is significantly different. The free-energy barrier as a function of $N_d$ is also lower than that for $N$ and $N_i$. There is, in fact, no reason why $\Delta G(N'_i)$, $\Delta G(N'_i)$, and $\Delta G(N')$ should be the same, since there is no one-to-one correspondence between these coordinates.

VI. NUCLEATION RATE

As a simple first test of the Langevin dynamics method for calculating nucleation rates, the calculations were restricted to those of the Langevin equation for one degree of freedom ($N$) with constant coefficients; i.e.,

$$m\dddot{N}(t) = -m \gamma \dot{N}(t) + f[N(t)] + a\Gamma(t),$$

where $a = k_B T m \gamma$ and $\gamma = \xi / m$.

The coefficients $m$ and $\gamma$ were estimated from the autocorrelation function for the velocity of $N$, $\dot{N} = dN/dt$,

$$\langle \dot{N}(t) \dot{N}(0) \rangle = \frac{k_B T}{m} \exp(-\gamma t).$$

It should be noted that this equation does not hold for $t=0$, where $d\langle \dot{N}(t) \dot{N}(0) \rangle / dt = 0$.

For the constant $NPT$ MD simulations carried out in this work using the algorithm in Ref. 13, $\dot{N}$ can be calculated analytically by using

$$\dot{N} = \frac{1}{n_c} \sum_{i}^{N_d} \sum_{j \neq i}^{N_d} f'(r_{ij}) \left( \mathbf{v}_{ij} \cdot \left( \mathbf{v}_{ij} + \frac{p_e}{W} \mathbf{r}_{ij} \right) \right),$$

where $\mathbf{v}_{ij} = \mathbf{r}_{ij}/dt - \mathbf{r}_{ij}/dt$ and $p_e$ and $W$ are the momentum and mass, respectively, of the barostat in the MD algorithm (see p. 1119 of Ref. 13 for notation). An advantage of using a continuous coordinate such as $N$ rather than a discrete coordinate like $N_i$ or $N_d$ is that the velocity can be calculated analytically. This avoids any ambiguity or error that may arise from taking a numerical difference to calculate the velocity of the discrete coordinate.

The simulations in which $m$ and $\gamma$ were calculated were carried out for various values of $c_N$ and $N_0$ in the biasing potential (see Table I). The simulations were 50,000 MD cycles in length. They were effectively carried out at constant volume by setting the inertia on the barostat at a very large number. This was done because the fluctuations in the simulation box volume were found to produce a lot of noise in $\langle \dot{N}(t) \dot{N}(0) \rangle$. The simulation box volume was set at its average value in constant $NPT$ simulations using the same biasing potential. Since these simulations were effectively at constant volume, then the second term in Eq. (34) can be ignored.

The autocorrelation functions $\langle \dot{N}(t) \dot{N}(0) \rangle$ obtained from these simulations are shown in Fig. 6. $k_B T/m$ was taken to be the value of $\langle \dot{N}(t) \dot{N}(0) \rangle$ at $t=0$, while $\gamma$ was obtained by fitting $\langle \dot{N}(t) \dot{N}(0) \rangle$ for small but nonzero $t$ ($t=0.02–0.06\tau$) to an exponential function of the form of Eq. (33). The values of $k_B T/m$ and $\gamma$ obtained from the simulations are shown in Table I. The data for $k_B T/m$ from Table I are plotted as a function of $\langle N \rangle$ in Fig. 7.

It can be seen from Table I that the friction coefficient $\gamma$ acting on the $N$ coordinate is roughly independent of $N$ and approximately equal to $50\tau^{-1}$. On the other hand, from Fig. 7, it can be seen that $k_B T/m$ varies linearly with $N$ (or at least

| $N_0$ | $c_N$ [eV] | $\langle N \rangle$ [eV] | $k_B T/m$ [meV/\tau] | $\gamma$ [eV/\tau] |
|------|---------|-----------------|-------------------|----------|
| 110  | 0.015   | 92.9            | 363.7             | 44.829   |
| 230  | 0.014   | 228.7           | 842.8             | 49.144   |
| 230  | 0.030   | 226.96          | 856.0             | 53.172   |
| 290  | 0.014   | 291.63          | 1066.2            | 47.556   |
| 350  | 0.015   | 345.926         | 1301.0            | 51.986   |

FIG. 6. Autocorrelation function for the velocity of $N$, $\langle \dot{N}(t) \dot{N}(0) \rangle$, as a function of time $t$ from simulations using various biasing potentials: $c_N = 0.015 eV$, $N_0 = 110$ (circles); $c_N = 0.014 eV$, $N_0 = 230$ (squares); $c_N = 0.030 eV$, $N_0 = 230$ (triangles up); $c_N = 0.014 eV$, $N_0 = 290$ (triangles down). The solid lines are exponential fits to the data points for $t=0.02–0.06\tau$.

FIG. 7. $k_B T/m$ vs $\langle N \rangle$ from Table I. (The linear fit to the points is $k_B T/m = 17.86 + 3.66(\langle N \rangle)$.)
with \( \langle N \rangle \) in our simulations. Therefore, the “mass” of the \( N \) coordinate is inversely proportional to the actual mass of the cluster.

The reason for this dependence follows directly from the definition Eq. (34). Ignoring the barostat term, it is straightforward to show that

\[
\frac{k_B T}{m} = \langle \dot{\mathbf{N}}^2 \rangle = \frac{4 \rho T \delta_{i3} \sigma_{ii}^s \langle \mathbf{g}^{(3)}(q_s) \rangle}{n_i^2 m_{11}} - \frac{2 \rho \delta_{i3} \sigma_{ii}^s \langle \mathbf{g}^{(3)}(q_s, q_s) \rangle}{n_i^2 m_{11}},
\]

where \( f_1 = \int d\mathbf{r} f'(r) \), \( f_2 = \int d\mathbf{r} f'(r)^2 \), and the pair and first Legendre transform of the triplet distribution function appear. This result uses the facts \( \langle \mathbf{v}_i \rangle = 0 \) and \( \langle \mathbf{v}_i \cdot \mathbf{v}_j \rangle = 3 \delta_{i3} \delta_{ij} T / m_{11} \), and assumes that \( f'(r) \) is sharply peaked at \( q_s \), that the droplet is of liquid density \( \rho_0 \), that the correlation functions are bulk quantities, and that the contribution from the surface of the droplet may be ignored. The particular consequence of this last assumption is that the mass of the coordinate \( N \) is inversely proportional to \( N \), in agreement with Fig. 7. One concludes from these results that the dominant contribution to \( \dot{\mathbf{N}} \) at small times comes from changing the number of bonds of atoms in the interior of the droplet rather than from atoms joining or leaving the surface of the droplet. This in turn suggests that the behavior of the liquid-like cluster definition and its continuum analog used here may not be as appealing as the original Stillinger definition.

In order to calculate the rate coefficient for nucleation \( k \), the reactive flux method\(^{15,16} \) was used, in conjunction with Langevin dynamics. In this method, the rate coefficient is given by a time-dependent function,

\[
k(t) = \frac{\langle q(0) \dot{\mathbf{q}}(0) - q^* \dot{\mathbf{q}}(t) - q^* \dot{\mathbf{q}}^+(t) \rangle_{\text{eq}}}{\langle \dot{\mathbf{q}}^2 \rangle_{\text{eq}}},
\]

which should be approximately constant and equal to the forward rate of reaction for time \( t \) intermediate between the molecular relaxation time and the reaction time. Here, \( q \) is an order parameter that discriminates between the reactant and product states, \( \dot{\mathbf{q}} \) is the transition state separating the reactant and product states, \( q < q^* \) denotes reactants and \( q > q^* \) denotes products, and the subscript “eq” denotes an equilibrium average.

The transition state theory (TST) estimate of the rate coefficient \( k_{\text{TST}} \), which assumes that trajectories that cross the transition state \( q^* \) from reactants to products never recross the transition state surface, is the \( t \to 0^+ \) limit of \( k(t) \):

\[
k_{\text{TST}} = \lim_{t \to 0^+} k(t) = \frac{\langle q^* \rangle_{\text{eq}} \langle \dot{\mathbf{q}}^2 \rangle_{\text{eq}}}{2 \langle \dot{\mathbf{q}}^2 \rangle_{\text{eq}}}.
\]

The transmission coefficient is defined as the ratio of \( k \) to \( k_{\text{TST}} \):

\[
\kappa = k/k_{\text{TST}}.
\]

The rate \( k \) was taken as the plateau value of \( k(t) \).

The transmission coefficient \( \kappa \) should be governed largely by diffusive behavior near the free-energy barrier. So \( \kappa \) from Langevin dynamics simulations on different free-energy surfaces, \( \Delta G(N) \) that are the same near the barrier but differ in the stable states should be approximately the same. Since we had accurate free-energy surfaces for \( \Delta G(q) \), where \( q = N \) or \( N_i \) from both our atomistic simulations and those of ten Wolde and Frenkel,\(^9 \) these surfaces were used in the Langevin dynamics calculations rather than a free-energy surface derived from the equation of state in Eq. (10). The barrier region of our simulation curve for \( \Delta G(N) \) was fitted to a couple of different polynomial functions that differed in the stable states [curves (4) and (5) in Fig. 8]. Due to the discrepancy between our results for \( \Delta G(N) \) and those of ten Wolde and Frenkel,\(^9 \) we did the same for their simulation curve for \( \Delta G(N) \) [curves (1)–(3) in Fig. 8].

In order to use Eq. (32), both \( \gamma \) and \( m \) had to be fixed at constant values. \( \gamma \) was fixed at its approximately constant value of 50.0. Since \( \kappa \) should depend mostly on what happens at the barrier, \( m \) was fixed at its value at the barrier in the free-energy curves from the atomistic simulations. Therefore, we took \( m = 7.0 \times 10^{-4} e \cdot a^2 \) in the simulations on curves (1)–(3) (since \( N^* \approx 280 \)) and \( m = 6.0 \times 10^{-4} e \cdot a^2 \) in the simulations on curves (4) and (5) (since \( N^* \approx 330 \)).

The calculated values of \( k \), \( k_{\text{TST}} \), and \( \kappa \) on the free-energy surfaces (1)–(6) in Fig. 8 are shown in Table II. As expected, \( \kappa \) is approximately independent of the differences in the stable states. The similarity between the value \( \kappa \) for surfaces (1)–(3) and (4) and (5) is because, despite the discrepancy between our results for \( \Delta G(N) \) and those of ten Wolde and Frenkel,\(^9 \) the curvature of our surface and theirs near the barrier is approximately the same.

The sensitivity of \( k \), \( k_{\text{TST}} \), and \( \kappa \) to the values of \( m \) and \( \gamma \) chosen can be estimated from Kramers’ rate theory in the high-friction limit, which predicts that \( k \sim 1/m \gamma \), \( k_{\text{TST}} \sim 1/\sqrt{m} \), and \( \kappa \sim 1/\sqrt{m} \gamma \).

The total nucleation rate, (number of droplets produced per unit volume per unit time), as shown by ten Wolde et al.,\(^{17} \) can be calculated by using
they found that their atomistic MD simulations.

The errors in their value of energy barrier with respect to the vapor phase, $0.011 \pm 0.009$, and $0.004 \pm 0.01$.

particle number coordinate Langevin simulations, the vapor density are very small and it takes a lot of costly simulation time to improve the statistics.

of alizations of the reactive flux method, they calculated values of $5.4 \pm 3.5$.

VII. DISCUSSION

This simple first test of the Langevin dynamics method for calculating the nucleation rate demonstrates its potential utility. In this work the free-energy surface and mass and friction coefficients used in the Langevin calculations were obtained from atomistic computer simulations. Compared to the direct atomistic simulation of the rate coefficient using the reactive flux method, it appears that more than four orders of magnitude less computer time is required to obtain the Langevin coefficients for comparable statistical accuracy. In addition, there are potentially other means for estimating the free-energy surface and mass and friction coefficients that may even avoid the atomistic simulations.

Our first test of the method for the LJ fluid demonstrated that, for at least one set of thermodynamic conditions, the nucleation rate is accurately estimated by assuming the diffusive behavior of a single coordinate, the number of particles in the cluster. In simulations of nucleation dynamics of the Ising model under similar thermodynamic conditions to the LJ fluid studied in this work (away from the critical point and at low supersaturation), Pan and Chandler found that the number of particles in the cluster was “good” reaction coordinate and obtained a reasonable estimate of $\kappa$ with a simple equation based on diffusive behavior of $N$ near barrier. (This analysis was analogous to ours in terms of Langevin dynamics).

However, as discussed in Sec. II, other degrees of freedom may be important under other conditions (e.g., near the critical point and at high supersaturation). These degrees of freedom can readily be incorporated into our Langevin dynamics calculations, along with the coordinate dependence of mass and friction coefficients.

ACKNOWLEDGMENT

We would like to thank Dr. Pieter Rein ten Wolde for useful discussions and data.

1. Y. Viisanen, R. Strey, and H. Reiss, J. Chem. Phys. 99, 4680 (1993).
2. P. Attard, Thermodynamics and Statistical Mechanics: Equilibrium by Entropy Maximisation (Academic, London, 2002).
3. M. P. Moody and P. Attard, Phys. Rev. Lett. 91, 056104 (2003).
4. M. P. Moody and P. Attard, J. Chem. Phys. 120, 1892 (2004).
5. J. K. Johnson, J. A. Zollweg, and K. E. Gubbins, Mol. Phys. 78, 591 (1993).
6. M. P. Allen and D. J. Tildesley, Computer Simulation of Liquids (Clarendon, Oxford, 1987).
7. D. L. Ermak and J. A. McCammon, J. Chem. Phys. 69, 1352 (1978).
8. F. H. Stillinger, J. Chem. Phys. 38, 1486 (1963).
9. P. R. ten Wolde and D. Frenkel, J. Chem. Phys. 109, 9901 (1998).
10. S. Nosé, J. Chem. Phys. 81, 511 (1984).
11. W. G. Hoover, Phys. Rev. A 31, 1695 (1985).
12. H. C. Andersen, J. Chem. Phys. 72, 2384 (1980).
13. G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, Mol. Phys. 87, 1117 (1996).
14. A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. 63, 1195 (1989).
15. D. Chandler, J. Chem. Phys. 68, 2959 (1978).
16. P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. 62, 251 (1990).
17. P. R. ten Wolde, M. J. Ruíz-Montero, and D. Frenkel, J. Chem. Phys. 110, 1591 (1998).
18. A. C. Pan and D. Chandler, J. Phys. Chem. B 108, 19681 (2004).

| Surface | $10^{5}k_{m}$ | $10^{7}k$ | $10^{6}k_{\text{TST}}$ | $\kappa$ |
|---------|---------------|--------------|----------------|---------|
| (1)     | 7.0           | 50           | 24             | 0.015   |
| (2)     | 7.0           | 50           | 3.6            | 0.013   |
| (3)     | 7.0           | 50           | 0.42           | 0.014   |
| (4)     | 6.0           | 50           | 5.1            | 0.015   |
| (5)     | 6.0           | 50           | 1.2            | 0.014   |

Rate = $\kappa \langle |\dot{N}|^2 \rangle / 2 \rho_v \exp[-\beta \Delta G(N^*)]$. (39)