Test of classical nucleation theory and mean first-passage time formalism on crystallization in the Lennard-Jones liquid

Sarah E. M. Lundrigan and Ivan Saika-Voivod[a)]
Department of Physics and Physical Oceanography, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador A1B 3X7, Canada

(Received 9 April 2009; accepted 13 August 2009; published online 8 September 2009)

We perform molecular dynamics (MD) and Monte Carlo computer simulations to test the ability of the recently developed formalism of mean first-passage time (MFPT) [J. Wedekind, R. Strey, and D. Reguera, J. Chem. Phys. **126**, 134103 (2007); J. Wedekind and D. Reguera, J. Phys. Chem. B **112**, 11060 (2008)] to characterize crystal nucleation in the Lennard-Jones liquid. We find that the nucleation rate, critical embryo size, Zeldovich factor, attachment rate, and the nucleation barrier profile obtained from MFPT all compare very well to the same quantities calculated using other methods. Furthermore, we find that the nucleation rate obtained directly through MD closely matches the prediction of classical nucleation theory. © *2009 American Institute of Physics*. [doi:10.1063/1.3216867]

I. INTRODUCTION

Molecular dynamics (MD) and Monte Carlo (MC) computer simulations play an increasingly important role in studying nucleation and testing classical nucleation theory (CNT).[1–6] While straightforward atomistic simulations can provide a great deal of microscopic detail of a system and are used when achieving equilibrium is not problematic, large free energy and kinetic barriers often encountered in studying nucleation necessitate the use of special computational methods to more efficiently sample phase space. To this end, various techniques such as metadynamics,[7] aggregation-volume-bias MC,[8] and transition path sampling[9] can be employed.

Particularly notable in the computational study of nucleation is the work of Frenkel and co-workers,[10–15] who developed techniques to calculate quantities pertinent to nucleation, such as the size of a crystal-like embryo, the work required to form a critical embryo (the nucleation barrier height), embryo composition, and the rate of particle attachment to the critical embryo. The techniques, in part and with variations, have been applied to several cases, including nucleation of globular proteins near a metastable critical point,[16] liquids near a wall,[17] hard spheres,[13,14] sodium chloride,[15] a system near an isostructural phase transition,[18] and silica.[19]

A subset of Frenkel's techniques used in concert forms what can be regarded as a standard approach in making a CNT-based prediction of the rate. A suitable set of criteria is used to define crystal-like embryos, after which biased sampling MC is used to find the equilibrium number distribution $N(n)$ of embryos of size n, using the largest embryo size in the system $n_{\max}$ as the biasing order parameter in order to sample rare states. To more efficiently establish equilibrium, parallel tempering, both in temperature T and cluster size, is often employed. This procedure yields the size of the critical embryo n^*, the work required to form an embryo of size n, $\Delta F(n)$, and the Zeldovich factor Z. The attachment rate of particles to the critical embryo $f^+_{n^*}$ is calculated separately with MD simulations of systems containing a critically sized embryo.

Recently, Wedekind *et al.*[20–22] developed a formalism useful in the regime where nucleation can be observed directly in MD simulations, namely, the mean first-passage time (MFPT) method. From an ensemble of nucleating runs, the MFPT method yields a clear determination of the rate, the size of the critical embryo, and the Zeldovich factor. Additionally, the same steady-state MD simulation data can be used to determine free energy barrier profiles as well as attachment rates. Thus, in the regime of applicability, i.e., where nucleation occurs on simulation time scales, the MFPT method provides a straightforward way of determining all the quantities necessary for characterizing the nucleation process.

In developing the MFPT method and applying it to liquid nucleation from the vapor,[23] Reguera and co-workers made use of the probability $P(n)$ of observing the largest embryo in the system to be of size n. $P(n)$ is not the distribution directly pertinent to CNT, $N(n)$ is. $P(n)$ is extensive, in that in a larger system, larger embryos are less rare, and hence the most common size for the largest embryo in the system increases with increasing system size, eventually leading to an apparent loss of liquid metastability at sufficiently large system size.[24–26] Despite this extensivity, $P(n)$ is related to $N(n)$; for sufficiently high barrier, for a given system size, the two distributions become approximately equal beyond a sufficiently large embryo size. This enables us to determine $\Delta F(n)$ from $P(n)$ and other steady-state data.

In this work, we compare the rate of nucleation determined from the MFPT method to a more conventional method employing MD simulations. The MFPT method then allows us to use this same set of data to determine n^*, Z, $\Delta F(n)$, and f^+_n, the attachment rate as a function of embryo

a)Electronic mail: saika@mun.ca.

size. We determine separately $\Delta F(n)$ through biased sampling MC and compare to the MFPT result and to the functional forms commonly used in CNT. Further, we calculate $f_{n^*}^+$ through size fluctuations of critical embryos and then compare the rate predicted by CNT to the directly determined one.

II. NUCLEATION RATE FROM MD SIMULATIONS

We simulate $N_p=4000$ particles of mass m under constant T and volume V at number density $\rho=0.95$, using the Lennard-Jones pair potential $U_{\mathrm{LJ}}(r)=4\epsilon[(\sigma/r)^{12}-(\sigma/r)^6]$ in a periodic cubic box of length $L=16.147\,754\,2$. This density was recently used to study nucleation by Wang *et al.*[27] We report all quantities in dimensionless units, with the length scale given by σ, energy scale by ϵ, and time scale by $\sqrt{m\sigma^2/\epsilon}$. We truncate interactions at a radius of 2.5 and we employ a time step of $\Delta t=0.001$. Temperature is held fixed with the Nosé–Hoover thermostat, with the mass parameter chosen so that the large fluctuations in the potential energy per particle U damp down after 1000 time steps following quenches from $T=1.2$ to $T=0.58$ ($Q=20$ in Eq. 6.1.26 in Ref. 28), and no unreasonable fluctuations occur in the steady state. We confirm that the conserved quantity in the Nosé–Hoover ensemble exhibits fluctuations much smaller than those in the potential energy and shows no drift over relevant time scales for our work.

The system is equilibrated at $T=1.2$ and subsequently 200 independent configurations are harvested by taking a snapshot configuration every 20 000 timesteps, a time large enough for particles to diffuse a few diameters. Each of these harvested configurations is then used to start a new simulation, a crystallization run, by simply changing the target temperature to 0.58. Each simulation proceeds until well past nucleation, i.e., well after the significant drop in U associated with crystallization. At this density, $T=1.2$ is near the upper boundary of the liquid-crystal coexistence region for the Lennard-Jones system with inifinite size and full range of potential,[29] while 0.58 is below the coexistence region, but above the temperature of 0.54 where stability of the liquid and the nature of nucleation are called into question.[27,30]

Figure 1 shows U as a function of time t for a few crystallization runs. Qualitatively, we see that the system easily establishes a metastable liquid state prior to crystallizing, and that crystallization is a fairly sharp event. The time required for the crystallite to grow to a large fraction of the system is short compared to the nucleation time and so it is highly unlikely to encounter multiple nucleation events.

One way to determine the rate J, the number of nucleation events per unit time per unit volume, is to model the fraction of non-nucleated members of the crystallizing ensemble as a function of time $R(t)$ with $\ln R(t)=-JV(t-t_0)$, where $V=L^3=4210.526\,35$ and t_0 is the time lag, or the minimum time to construct a critical embryo.[23] In Fig. 2, we plot $R(t)$ using different criteria for determining the time at which the system is nucleated. U drops below -6 for the first time, $n_{\max}$ first reaches 80, and the last time $n_{\max}$ is less than 64. All three sets of curves yield similar slopes, with $JV=(3.9\pm0.3)\times10^{-4}$, or $J=(9.0\pm0.7)\times10^{-8}$.

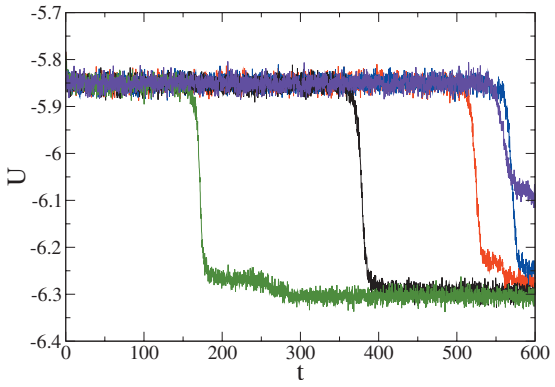

FIG. 1. A sampling of potential energy time series for crystallizing MD runs. Equilibrium configurations drawn from $T=1.2$ are quenched to $T=0.58$ at $t=0$ by changing the Nosé–Hoover thermostat setting. For the data shown here, it is clear that a metastable liquid state is attained prior to nucleation.

To define crystal-like embryos, we employ the method developed by Frenkel and co-workers,[10] based on Steinhardt bond-ordering analysis,[31] with cutoff criteria taken from Ref. 27. Namely, we define for particle i a 13-component complex vector $q_{6m}(i)=\Sigma_j Y_{6m}(\hat{r}_{ij})$, where Y_{6m} is the spherical harmonic function of order six, $\hat{r}_{ij}$ is the unit normal giving the direction of the bond between particle i and its neighbor j, and the sum is carried out over all neighboring particles, i.e., those within a distance of 1.4 of particle i. If the dot product $\Sigma_{m=-6}^6 \hat{q}_{6m}(i)\hat{q}_{6m}(j)^*$ between two neighboring particles is greater than 0.5, where $\hat{q}_{6m}(i)=q_{6m}(i)/[\Sigma_{m=-6}^6|q_{6m}(i)|^2]^{1/2}$ and q^* is the complex conjugate of q, then particle j is a *connected* neighbor of i. If a particle has at least 11 connected neighbors, then it is considered crystal-like. Finally, and this definition differs from Refs. 10 and 27, two crystal-like particles are part of the same crystal-like embryo if they are connected (and not just neighbors).

In the MFPT formalism, the average time that an embryo of size n appears for the first time is given by

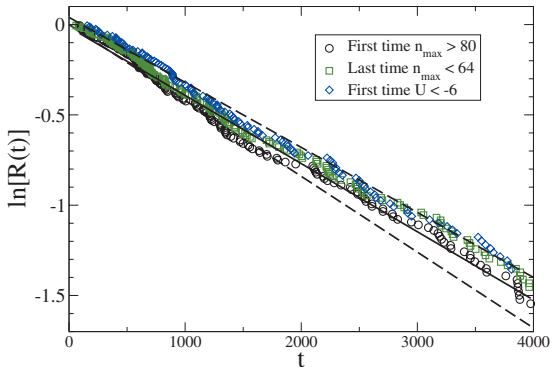

FIG. 2. Determining J from $R(t)$, the fraction of non-nucleated runs from an ensemble of MD simulations. Shown are $\ln[R(t)]$ using different criteria for determining the time at which a system has nucleated: First time $n_{\max}>80$ (circles), last time $n_{\max}<64$ (squares), and first time $U<-6$ (diamonds). The solid line is a representative line of best fit with $JV=3.8\times10^{-4}$, while the dashed lines are guides to the eye in estimating uncertainties with $JV=3.6\times10^{-4}$ and $JV=4.2\times10^{-4}$.

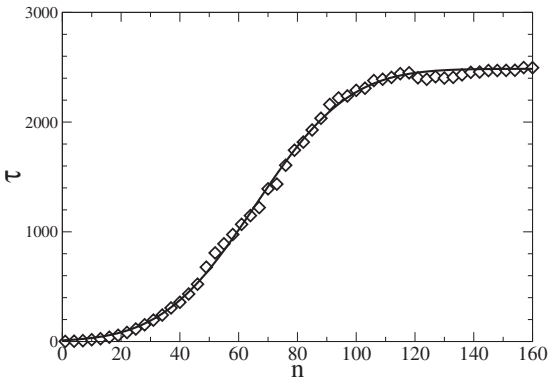

FIG. 3. Plot of mean first-passage times. Plotted are $\tau(n)$ (diamonds), along with a fit of $\tau(n)$ to Eq. (1).

$$\tau(n) = \frac{1}{2JV}\{1 + \mathrm{erf}[c(n - n^*)]\}, \tag{1}$$

where n^* is the size of the critical embryo, erf(x) is the error function, and $c = \sqrt{\Delta F^{*\prime\prime}(n^*)/k_B T} = Z\sqrt{\pi}$ characterizes the curvature at the top of the nucleation barrier profile $\beta\Delta F^*(n) = -\ln[P(n)] + \kappa$, where $\beta = (k_B T)^{-1}$ and k_B is Boltzmann's constant. We will clarify the meaning of $\Delta F^*(n)$ in Sec. III.

For each crystallization run, we determine $\tau(n)$ by saving configurations every 1000 time steps and finding the largest crystal-like embryo in each saved configuration. Because of the discrete sampling of configurations, the nature of embryo growth and decay, and possible amalgamation of embryos of size greater than one, the time series $n_{\max}(t)$ need not be monotonic, and often decreases or increases by more than one.

For example, let us assume that the values of $n_{\max}(t)$ for the first four sampling times for a particular nucleation run are $n_{\max}(t=1)=0$, $n_{\max}(2)=2$, $n_{\max}(3)=4$, and $n_{\max}(4)=1$. The resulting first-passage times are $\tau(n_{\max}=0)=1$, $\tau(1)=2$, $\tau(2)=2$, $\tau(3)=3$, and $\tau(4)=3$, where we filled in entries for skipped embryo sizes, i.e., when an embryo of a given size is first observed at a given time, that same time is assigned to all smaller sizes that have not yet been assigned a time. Sampling more frequently than once every 1000 time steps could minimize the effects of this "filling in" procedure.

We plot $\tau(n)$ averaged over all 200 crystallization runs in Fig. 3, along with a fit of $\tau(n)$ to Eq. (1). From the fit to $\tau(n)$, we obtain $J_{\mathrm{MFPT}}V = 4.0 \pm 0.1 \times 10^{-4}$, or $J_{\mathrm{MFPT}} = 9.4 \pm 0.3 \times 10^{-8}$, $c_{\mathrm{MFPT}} = 0.028 \pm 0.001$ or $Z_{\mathrm{MFPT}} = 0.0158 \pm 0.0006$, and $n^*_{\mathrm{MFPT}} = 65 \pm 1$. The uncertainties in the rate appear to be larger than what may arise from the bias of at most 1 in τ introduced by the filling in procedure used in getting τ. To confirm the criticality of n^*-sized embryos, we choose 27 configurations containing such embryos, randomize velocities, and find that 13 of the configurations continue to crystallize, while the other half decays.

III. BARRIER RECONSTRUCTION

In this work, and particularly in this section, we make use of two distributions, $N(n)$ and $P(n)$, upon which free energies $\Delta F(n)$ and $\Delta F^*(n)$ are based. For clarity, we reiterate their meanings and differences. $N(n)$ is the average number of embryos of size n in the equilibrium liquid. Generally, $N(0)$ (the number of liquidlike particles) is only slightly less than the total number of particles and $N(n)$ decreases monotonically for $0 \le n \le n^*$. $P(n)$ is the probability that the largest embryo in a configuration taken from the equilibrium liquid is of size n. When there is a large free energy barrier to nucleation, larger embryos are rare (when a rare embryo is present, there is approximately no other embryo of that size or larger in the system) and for rare embryos sizes $N(n)$ and $P(n)$ are approximately equal. For not-so-rare smaller clusters, the two distributions are quite different. For example, we find for our system that most of the time there are some crystal-like particles present, and so it is rare to find all the particles in the system to be liquidlike, i.e., $P(0)$ is smaller than $P(1)$. In our case we find that $P(4)$ is a maximum, i.e., the most common largest embryo size in the system is 4. Whereas $N(n)$ decreases monotonically from $n=0$, $P(n)$ need not do so. For a discussion on how the two distributions are formally related, see Ref. 26.

Wedekind and Reguera outline their procedure in Refs. 21 and 22 for determining the barrier to nucleation as a function of an order parameter that characterizes the system. In particular, they choose the largest embryo in the system as the appropriate order parameter to track the nucleation process and obtain

$$\beta\Delta F^*_{\mathrm{MFPT}}(n_{\max}) = \ln[B(n_{\max})] - \int_{a'}^{n_{\max}} \frac{dx'}{B(x')} + C, \tag{2}$$

where a' is a formal lower limit of integration the explicit choice of which is absorbed into C, which is fixed by choosing an appropriate reference state. The function $B(x)$ is given by

$$B(x) = \frac{1}{P_{\mathrm{st}}(x)}\left[\int_a^x P_{\mathrm{st}}(x')dx' - JV\tau(x)\right], \tag{3}$$

where a is the left (reflecting) boundary of the order parameter domain, which in our case we take to equal zero, and $P_{\mathrm{st}}(n_{\max})$ is the steady-state probability of finding a configuration in the ensemble of crystallization runs with largest crystal-like embryo size equal to $n_{\max}$. We obtain $P_{\mathrm{st}}(n_{\max})$ by constructing a histogram in $n_{\max}$, considering configurations from all 200 runs with $n_{\max} \le h$ and $t > 10$, and then dividing by the number of configurations considered. $h = 100$ is the upper limit on $n_{\max}$ for the purposes of determining $P_{\mathrm{st}}(n)$, as well as $N_{\mathrm{st}}(n)$, the steady-state counterpart of $N(n)$, i.e., $\Sigma_{n=0}^{n=h} P_{\mathrm{st}}(n) = 1$ and $N_{\mathrm{st}}(0) + \Sigma_{n=1}^{n=h} n N_{\mathrm{st}}(n) = N_p$. In this work, we actually evaluate the integrals in Eqs. (2) and (3) as discrete sums.[22]

It is perhaps worth pointing out that the steady state differs from equilibrium. Steady state refers to the case where the system is in the process of nucleating (and the nucleation rate is constant in time). A nucleating liquid will not sufficiently sample near-critical states, since it is apt to quickly slide down the free energy landscape to the crystal once it reaches the top of the free energy barrier. The equi-

librium liquid can only be obtained if the system is formally prevented from nucleating, e.g., by some constraint on the largest embryo.

In equilibrium, $\beta\Delta F^*(n) = -\ln P(n) + \kappa$, and it is essentially the goal of the MFPT formalism to obtain $P(n)$ from $P_{st}(n)$ available from dynamic simulations without placing constraints on the system as is done with, e.g., umbrella sampling MC. $\Delta F^*(n_2) - \Delta F^*(n_1)$ can be interpreted as the work required to change the system from having $n_{max} = n_1$ to having $n_{max} = n_2$.

Recently, Wedekind and Reguera[22] carried out their procedure in the case of liquid nucleation from the vapor. Here, we are working with crystal nucleation and there appears to be an additional subtlety. The central thermodynamic quantity relevant to predicting the rate via CNT is $N(n)$ (again, the equilibrium number of crystal-like embryos of size n in the system). For rare embryo sizes $N(n) \approx P(n)$ [provided both $N(n)$ and $P(n)$ are defined on the same range of n, otherwise they differ by a normalization constant]. As mentioned above, at small n they are different in general. Depending on system size (and perhaps to some extent on the definition of a crystal-like embryo), it is not surprising to find it rare for the system to be devoid of crystal-like particles. In this case $P(n)$ should have a maximum at n_{min} corresponding to a most likely largest embryo size for the system [or a minimum in $-\ln P(n)$].

Despite the differences at small cluster sizes between $P(n)$ and $N(n)$, the MFPT formalism recovers the CNT expression for the rate, in the case that $N(n^*) = P(n^*)$ and $\Delta F^*(n^*)$ is a pronounced local maximum (so that $\exp[\beta\Delta F^*(n)]$ dominates near n^*). In this case, n^* is now also consistent with being defined as the size at which the work required to form an n-sized embryo from liquid particles $\beta\Delta F(n) = -\ln N(n) + \kappa$ has its maximum. The constant $\kappa = \ln N(0)$ is a constant chosen so that $\Delta F(0) = 0$ [$N(0)$ is the number of liquidlike particles]. In this case, we can use Eq. A4 from Ref. 20, which we rewrite here as

$$J_{\mathrm{MFPT}} V = f^+_{n^*} \frac{\exp[-\beta\Delta F^*(n^*)]}{\int_{a=0}^{n^*} dz \exp[-\beta\Delta F^*(z)]} \sqrt{\frac{\Delta F^{*\prime\prime}(n)}{2\pi k_B T}}, \tag{4}$$

and note that the last factor is the Zeldovich factor. The numerator of the middle factor, given the definition of $\Delta F^*(n)$, is $N(n^*)/N(0)$ and the denominator, given the normalization of $P(n)$ and a choice of h to be near n^*, equals

$$\int_{a=0}^{n^*} dz e^{-\beta\Delta F^*(z)} = \frac{1}{N(0)} \left(1 - \int_{n^*}^{h} dz P(z) \right) \approx \frac{1}{N(0)}, \tag{5}$$

since near n^*, $P(n)$ is very small. We briefly discuss the h dependence of this result in Sec. VI. Thus, we recover the CNT result,

$$J_{\mathrm{MFPT}} = J_{\mathrm{CNT}} = f^+_{n^*} Z \frac{N(n^*)}{V}, \tag{6}$$

$$\approx \rho f^+_{n^*} Z \exp\left(-\frac{\Delta F(n^*)}{k_B T} \right). \tag{7}$$

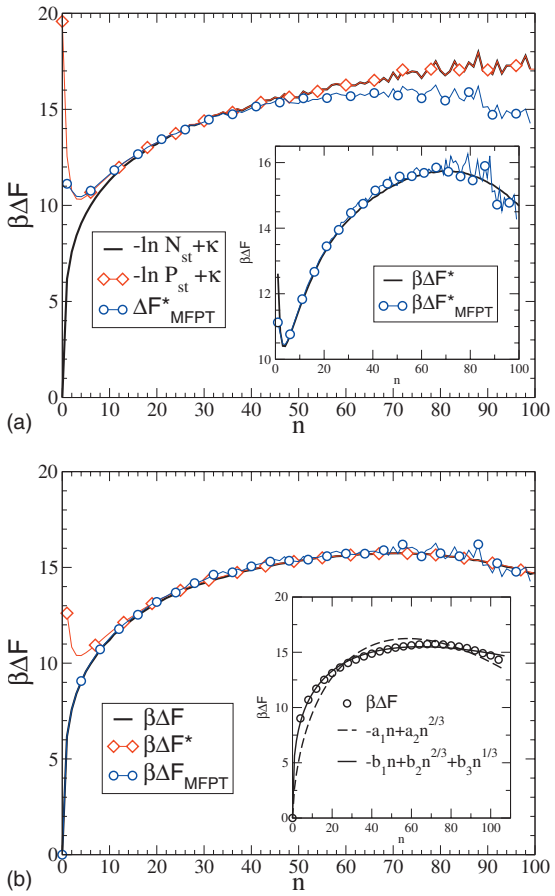

FIG. 4. Steady-state distributions and free energies at $T = 0.58$. (a) Shown are negative logarithms of steady-state distributions $P_{st}(n)$ (thin red line with diamonds) and $N_{st}(n)$ (thick black line), and $\Delta F^*_{\mathrm{MFPT}}(n)$ (blue line with circles), the free energy obtained from $P_{st}(n)$ and $\tau(n)$ through Eqs. (2) and (3). Inset to (a): comparison of $\Delta F^*_{\mathrm{MFPT}}(n)$ and $\Delta F^*(n)$, obtained through MC simulations. (b) Shown are $\Delta F^*(n)$ (red curve with diamonds) and $\Delta F(n)$ (black curve), both obtained through MC simulations, as well as $\Delta F_{\mathrm{MFPT}}(n)$ (blue line with circles), obtained by splicing the portion of $-\ln N_{st}$ below $n = 26$ and the portion of $\beta\Delta F^*_{\mathrm{MFPT}}(n)$ for $n \geq 26$. Inset to (b): $\Delta F(n)$ (circles) and fitting functions $-a_1 n + a_2 n^{2/3}$ and $-b_1 n + b_2 n^{2/3} + b_3 n^{1/3}$, with best fit parameters $a_1 = 0.558$, $a_2 = 3.242$, $b_1 = 0.123$, $b_2 = 0.0275$, and $b_3 = 5.733$.

Thus, we see that $\Delta F^*(n)$ can be used to determine the rate in so far as it can be used to determine $N(n^*)$. Although tempting, it does not appear to be the case that the free energy barrier $\Delta F^*(n^*) - \Delta F^*(n_{min})$ is of immediate value in determining rates, at least for bulk systems when barriers are high. For low barriers or small systems the scenario might be different.[24,32]

In Fig. 4(a) we show the results of using Eqs. (2) and (3) to determine $\Delta F^*_{\mathrm{MFPT}}(n)$ and compare the results against those obtained through constrained, parallel tempered MC simulations for the equilibrium quantity $\Delta F^*(n)$ (see Sec. IV). We find very good agreement notwithstanding some noise in the MD data.

Consistent with large clusters being rare, we find that $P_{st}(n)$ and $N_{st}(n)$ are practically the same for $n \gtrsim 25$. Moreover, the n dependence of $\Delta F^*_{\mathrm{MFPT}}(n)$ is not appreciably different from that of $-k_B T \ln P_{st}(n)$ for $n_{max} \lesssim 30$, i.e., the steady-state distribution equals the equilibrium one for this lower range of embryo sizes.

This suggests defining and determining $\Delta F_{\mathrm{MFPT}}(n)$ as a splicing of $-k_BT \ln N_{\mathrm{st}}(n)$ at small n and $\Delta F^*_{\mathrm{MFPT}}(n)$ at larger n. In particular, we follow $-k_BT \ln[N_{\mathrm{st}}(n)/N_{\mathrm{st}}(0)]$ for $n<26$ and then follow $\Delta F^*_{\mathrm{MFPT}}(n)$ after shifting it to match $-k_BT \ln[P_{\mathrm{st}}(n)/N_{\mathrm{st}}(0)]$ at small n. In Fig. 4(b) we compare $\Delta F_{\mathrm{MFPT}}(n)$ against $\Delta F(n)$ obtained from MC simulations (see Sec. IV) and find very good agreement.

What is also very encouraging is the excellent agreement between $\beta\Delta F(n)$ and $-\ln[N_{\mathrm{st}}(n)/N_{\mathrm{st}}(0)]$ at small n. In principle, this should not be surprising since for small embryo sizes the steady-state distribution of embryos should match the equilibrium one. Nonetheless, here we have MD simulations with a Nosé–Hoover thermostat to get one and MC simulations to get the other. It seems that the simulations produce the canonical distribution of embryos.

IV. MC SIMULATIONS

To determine $\Delta F(n)$, we use biased sampling MC with parallel tempering in both temperature and cluster size. Our biasing potential is $\phi(n_{\max};n_0)=k_0(n_{\max}-n_0)^2/2$, with $k_0=0.05$,[27] and n_0 is the target size for the largest cluster. We carry out 22 simulations in parallel using n_0 values of 0, 10, 20, …, and 100, each at two temperatures $T=0.57$ and $T=0.58$. The simulations are seeded with configurations taken from crystallizing MD runs with appropriately sized embryos.

One regular MC step consists of five Metropolis *NVT* MC moves per particle to generate a new configuration, followed by acceptance of the new configuration with probability $\exp[-\beta\phi(n_{\max};n_0)]$.[27] After nine regular MC steps, we carry out a "switch" MC step, for which an attempt is made to switch configurations with nearest-neighbor simulations in T or n_0, alternating between the two kinds of switches, i.e., a T-switch is attempted every 20 MC steps. A switch of configurations between two simulations running at different T (but same n_0) is accepted with probability $\min\{1,\exp[(\beta_i-\beta_j)(U_i-U_j)]\}$, where β_i is the Boltzmann factor for simulation i and U_i is its instantaneous potential energy (including the contribution form the constraint potential). For a switch of configurations between two simulations with different values of n_0 (but same T), the acceptance probability is $\min\{1,\exp[-\beta(w_n-w_o)]\}$, where $w_n=\phi(n_i;n_{0j})+\phi(n_j;n_{0i})$ and $w_o=\phi(n_i;n_{0i})+\phi(n_j;n_{0j})$, n_i is the size of the largest cluster in simulation i. After equilibrating the system, we collect data for 10^6 MC steps, over which time the energy for the state points simulated appears very stable, showing no appreciable drift.

For each simulation at a given T and n_0, we keep track of the histogram $P_C(n)$, which counts the number of times the largest embryo in the system is of size n, and the histogram $N_C(n)$, which counts the number of times an embryo of size n appears in the system. We transform the histograms in the constrained ensemble to the unconstrained *NVT* ensemble via, $N(n)=\langle N_C(n)\exp[\beta\phi(n;n_0)]\rangle_C$const, where $\langle\cdots\rangle_C$ denotes an average in the simulation using the constraining potential, and $P(n)=P_C(n)\exp[\beta\phi(n;n_0)]$const (see Ref. 14 for details), thus determining sections of ΔF and ΔF^* up to an additive constant for each T and n_0. We then trim $N(n)$ and $P(n)$, only keeping histogram entries for which $P_C(n)$ is greater than 1.5% of the total number of samples. We make an exception for the case of $n_0=0$, for which we keep the $n=0$ entry in $N(n)$, even though the $n_{\max}=0$ state is sampled less than 1.5% of the time.

To obtain the free energies for all n, we shift each section in order minimize the difference between overlapping parts and then average the overlapping values. Differences in free energy at overlapping points of adjacent curves prior to averaging are about $0.05k_BT$ or less. We then shift the curves so that $\Delta F(0)=0$ and $\Delta F^*(70)=\Delta F(70)$.

We show our results and compare to those obtained through MFPT at $T=0.58$ in Fig. 4. We find that the results of the two methods agree very well. In the inset to Fig. 4(b), we show a fit of $\Delta F(n)$ to the usual CNT form $-|\Delta\mu|n+sn^{2/3}$, as well as to a modified form which includes an extra term to account for curvature effects, $kn^{1/3}$. The extra parameter improves the fit dramatically and the fit parameters seem to indicate that the curvature term is an important contribution. This warrants further detailed investigation.

From a quadratic fit to $\Delta F(n)$ near the top of the barrier, we find $n^*=71\pm1$, $\beta\Delta F(n^*)=15.74\pm0.25$, and $c_{\mathrm{MC}}=\sqrt{\beta\Delta F''(n^*)/2}=0.031\pm0.002$, or $Z_{\mathrm{MC}}=0.0175\pm0.0011$. This value of the curvature is quite close to that obtained from MFPT.

Comparing our results to Ref. 27, we obtain a similar value for $\Delta F^*(n^*)-\Delta F^*(n_{\min})$, but n^* was reported to be about 40. We have not accounted for the discrepancy but do note that at least in our case, $\Delta F(n)$ [or equivalently $\Delta F^*(n)$] is rather flat, with roughly a size range of $n=40$ to $n=90$ corresponding to about a $1k_BT$ difference in free energy. It could be that the small difference in embryo definition gives rise to the difference in our value of n^*.

V. KINETIC FACTORS AND RATE PREDICTION

To complete the calculation for the CNT prediction of the rate, we obtain the attachment rate of particles to critically sized embryo based on the analysis presented in Ref. 14. We take

$$f^+_{n^*}=\text{slope of } \langle[n_{\max}(t)-n_{\max}(0)]^2\rangle/2, \tag{8}$$

where $n_{\max}(0)$ is close to n^*.

Specifically, we harvest 50 configurations from our crystallizing MD runs at $T=0.58$ with $n_{\max}$ near n^* and use them as seeds for new MD runs after randomizing velocities. We define time origins to occur whenever $n_{\max}$ is near n^*, i.e., $63\le n_{\max}(0)\le 79$ (the barrier is quite flat in this region), with a minimal time interval of 1 between origins. Since we wish to study properties near the top of the barrier, we eliminate time origins where $|n_{\max}(20)-n_{\max}(0)|>70$, as these trajectories do not correspond to fluctuations near the top of the barrier. The results are not sensitive to the particular size and time cutoffs. We plot $\langle[n_{\max}(t)-n_{\max}(0)]^2\rangle$ in Fig. 5 and obtain from the slope $f^+_{n^*}=43\pm3$.

Having calculated all the factors in Eq. (7), namely, Z and $\Delta F(n^*)$ from MC simulation and $f^+_{n^*}$ from MD simulations of critically sized embryos, we calculate the CNT prediction for the rate to be $J_{\mathrm{CNT}}=(10\pm3)\times10^{-8}$. This value of

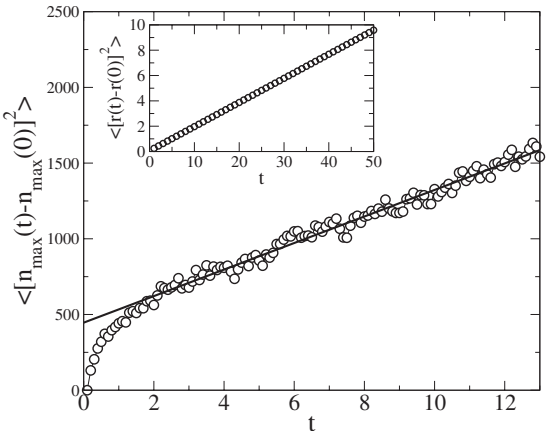

FIG. 5. Determination of $f_{n^*}^+$ at $T=0.58$ from the time dependence of size fluctuations of near-critical embryos. Shown is a line of best fit obtained by fitting the data (circles) starting from $t=4$. Inset: mean squared displacement as a function of t for the metastable liquid, also at $T=0.58$.

the rate agrees quite well with those obtained by direct simulation. The largest source of uncertainty stems from our determination of $\Delta F(n^*)$.

Another quantity of interest is the atomic jump distance λ that enters into the expression for the attachment rate,

$$f_{n^*}^+ = 24Dn^{*2/3}/\lambda^2,$$

where $D=0.0317$ is the diffusion coefficient, evaluated from the slope of the mean squared displacement in the metastable liquid at $T=0.58$ shown in the inset of Fig. 5. We find that $\lambda=0.55\pm0.03$, a physically reasonable value and comparable to that obtained for hard spheres.[14]

According to Ref. 22, the attachment rates as a function of n can also be obtained from MFPT and P_{st} via

$$f_n^+ = B(n)/\left(\frac{\partial\tau(n)}{\partial n}\right). \tag{9}$$

Using the fit to $\tau(n)$ to calculate the derivative, we plot f_n^+ so obtained in Fig. 6, along with the single point for $f_{n^*}^+$ obtained from Eq. (8). Although there is some noise stemming from P_{st}, the agreement near n^* between the two methods is quite good.

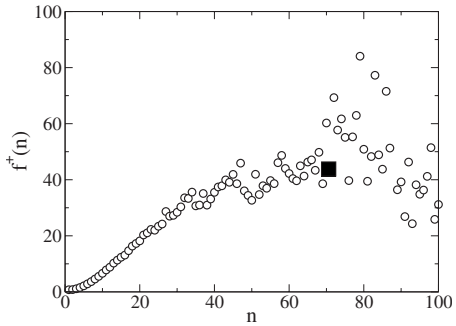

FIG. 6. Attachment rates from the MFPT formalism, Eq. (9), as well as the attachment rate at critical size obtained from fluctuations of the critical embryos, Eq. (8) (filled square).

TABLE I. Summary of calculated quantities for $T=0.58$.

Quantity	Value
N_p	4000
ρ	0.95
n^*_{MFPT}	65 ± 1
n^*	71 ± 1
$\beta\Delta F(n^*)$	15.74 ± 0.25
$f_{n^*}^+$	43 ± 3
D	0.0317
Z_{MFPT}	0.0158 ± 0.0006
Z_{MC}	0.0175 ± 0.0011
λ	0.55 ± 0.03
J	$(9.0\pm0.7)\times10^{-8}$
J_{MFPT}	$(9.4\pm0.3)\times10^{-8}$
J_{CNT}	$(10\pm3)\times10^{-8}$

VI. DISCUSSION AND CONCLUSIONS

Our aims in this work are to test the method of mean first-passage times and to compare the CNT rate prediction to a direct determination of the rate in fairly deeply supercooled Lennard-Jones liquid. We are working in a regime where the barrier to nucleation is high enough so that metastable equilibrium is readily achieved, but low enough so that nucleation is readily observable through unbiased MD. The diffusivity of the liquid is also fairly high, reducing concerns about kinetic barriers to equilibration, although the dynamics of the crystal-like embryos are likely much slower than that of the liquid.

Our main results are that there is generally very good agreement between quantities calculated through MFPT and other methods, and that the CNT rate prediction holds very well for the system studied in the sense that putting in values for the barrier height, Zeldovich factor, and attachment rate yields a rate equal to the directly determined one within error.

Table I gives a summary of some of the quantities calculated. The uncertainty estimates on n^* and n^*_{MFPT} come from using different subsets of data to perform the fits. However, we also see that the barrier curve is quite flat near the top, and so it is difficult to attribute a significance to the difference in critical sizes obtained through MC and MFPT.

The barrier obtained from MFPT is based on $P(n)$, whereas the CNT barrier comes from the embryo size distribution $N(n)$. As $N_{st}(n)\approx N(n)$ for small n and $\Delta F^*(n)\approx\Delta F(n)$ for large (rare) clusters, it is possible to reconstruct the CNT barrier using $N_{st}(n)$ and $\Delta F^*_{\mathrm{MFPT}}(n)$. The result agrees very well with the constrained MC result. If we fit the top of ΔF_{MFPT} with a quadratic, even though there is some noise, we obtain a critical size of 72 ± 2, Zeldovich factor of 0.020 ± 0.006, and barrier height of 15.90 ± 0.05 (for a rate of 9.9×10^{-8}). These estimates are all in line with MC results. It seems, then, that MFPT based on 200 nucleating runs gives quite good estimates on the properties related to the barrier, as well as the barrier profile itself. We note that the determination of $\Delta F_{\mathrm{MFPT}}(n)$ does not hinge upon approximations relating to the height of the barrier, in contrast

with Eq. (1). So determining quantities from the barrier itself may provide truer, though perhaps noisier, estimates.

One unsettled point is the choice of the upper bound on embryo sizes in determining $P(n)$ and the impact this has on formally connecting J_{MFPT} and J_{CNT} through Eqs. (4)–(6). Here, we have chosen h to be near enough to n^* so that the area under $P(n)$ [or $P_{\mathrm{st}}(n)$] beyond n^* is negligible in order to obtain $J_{\mathrm{MFPT}} \approx J_{\mathrm{CNT}}$. Similarly, the normalization of P_{st} affects $\Delta F^*_{\mathrm{MFPT}}(n)$ through the difference appearing in Eq. (3). Interestingly, $\tau(n)$ is independent of the normalization of $P(n)$; see, e.g., Eq. 2 in Ref. 20. Although we obtain excellent agreement between MFPT and MC simulations in determining the barrier, we wish to understand this point in the future.

The barrier $\Delta F(n)$ we calculate is not well described by the usual phenomenological form typically used in CNT, i.e., competing bulk and surface free energy term $(-an+bn^{2/3})$. Adding a curvature correction term $(n^{1/3})$ improves the fit and moreover curvature effects seem to be strong. Furthermore, as mentioned earlier, our estimate of n^* is larger than that of Ref. 27, necessarily indicating a different barrier shape that perhaps arises from a small difference in embryo criteria. However, the states we identify as containing critical embryos must not be very far from the truth, since we do get (albeit for a fairly small sample size) roughly equal probability for growth and decay. If the true critical embryo size were nearer to 40, instead of 70, then the probability of decay would imply a flat barrier at the top, which we have. Furthermore, the rate determined from considering the energy is consistent with those obtained by considering embryo sizes and all are consistent with the CNT prediction. However, this consistency may not be particularly sensitive to the shape of the barrier, nor to n^*, since the important quantity is $N(n^*)$. Regardless, a more detailed study relating the structure of the embryos to $\Delta F(n)$ and also one relating the dependence of the shape of $\Delta F(n)$ on the criteria used to define crystal-like embryos would be useful.

The attachment rate is also recovered effectively by MFPT, although we only have one point of comparison, at n^*, with an independent method. In conclusion, we find the method of mean first-passage times to be an effective computational tool in characterizing crystal nucleation in the regime where nucleation occurs on simulation time scales.

ACKNOWLEDGMENTS

We thank Richard K. Bowles for useful discussions, as well as for computing resources. Additional computational resources and support are provided by SHARCNET, ACEnet, and WESTGRID, and we acknowledge financial support from NSERC.

[1] J. W. Gibbs, *The Scientific Papers of J. Willard Gibbs* (Dover, New York, 1961).
[2] M. Volmer and A. Weber, Z. Phys. Chem. **119**, 227 (1926).
[3] L. Farkas, Z. Phys. Chem. **125**, 236 (1927).
[4] R. Becker and W. Döring, Ann. Phys. **24**, 719 (1935).
[5] K. F. Kelton, *Crystal Nucleation in Liquids and Glasses* (Academic, Boston, 1991), Vol. 45, pp. 75–177.
[6] P. G. Debenedetti, *Metastable Liquids: Concepts and Principles* (Princeton University Press, Princeton, 1996).
[7] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. U.S.A. **99**, 12562 (2002); A. Laio and F. L. Gervasio, Rep. Prog. Phys. **71**, 126601 (2008).
[8] B. Chen and J. I. Siepmann, J. Phys. Chem. B **104**, 8725 (2000); **105**, 11275 (2001); B. Chen, J. I. Siepmann, K. J. Oh, and M. L. Klein, J. Chem. Phys. **115**, 10903 (2001).
[9] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).
[10] P. R. ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, J. Chem. Phys. **104**, 9932 (1996).
[11] P. R. ten Wolde and D. Frenkel, J. Chem. Phys. **109**, 9901 (1998).
[12] P. R. ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, J. Chem. Phys. **110**, 1591 (1999).
[13] S. Auer and D. Frenkel, Nature (London) **409**, 1020 (2001).
[14] S. Auer and D. Frenkel, J. Chem. Phys. **120**, 3015 (2004).
[15] C. Valeriani, E. Sanz, and D. Frenkel, J. Chem. Phys. **122**, 194501 (2005).
[16] P. R. ten Wolde and D. Frenkel, Science **277**, 1975 (1997).
[17] S. Auer and D. Frenkel, Phys. Rev. Lett. **91**, 015703 (2003).
[18] A. Cacciuto, S. Auer, and D. Frenkel, Phys. Rev. Lett. **93**, 166105 (2004).
[19] I. Saika-Voivod, R. K. Bowles, and P. H. Poole, J. Chem. Phys. **124**, 224709 (2006).
[20] J. Wedekind, R. Strey, and D. Reguera, J. Chem. Phys. **126**, 134103 (2007).
[21] J. Wedekind and D. Reguera, *17th International Conference on Kinetic Reconstruction of the Nucleation Free Energy Landscape in Nucleation and Atmoshperic Aeorsols*, Galway, Ireland, 2007 (Springer, Netherlands, 2008).
[22] J. Wedekind and D. Reguera, J. Phys. Chem. B **112**, 11060 (2008).
[23] G. Chkonia, J. Wölk, R. Strey, J. Wedekind, and D. Reguera, J. Phys. Chem. B **130**, 064505 (2009).
[24] P. Bhimalapuram, S. Chakrabarty, and Biman Bagchi, Phys. Rev. Lett. **98**, 206104 (2007).
[25] L. Maibaum, Phys. Rev. Lett. **101**, 019601 (2008).
[26] S. Chakrabarty, M. Santra, and Biman Bagchi, Phys. Rev. Lett. **101**, 019602 (2008).
[27] H. Wang, H. Gould, and W. Klein, Phys. Rev. E **76**, 031604 (2007).
[28] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic, San Diego, 2002).
[29] E. A. Mastny and J. de Pablo, J. Chem. Phys. **127**, 104504 (2007).
[30] F. Trudu, D. Donadio, and M. Parrinello, Phys. Rev. Lett. **97**, 105701 (2006).
[31] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, Phys. Rev. B **28**, 784 (1983).
[32] E. Mendez-Villuendas, I. Saika-Voivod, and R. K. Bowles, J. Chem. Phys. **127**, 154703 (2007).