A molecular dynamics study of water nucleation using the TIP4P/2005 model

Alejandro Pérez[a)] and Angel Rubio[b)]

Nano-bio Spectroscopy Group and ETSF Scientific Development Center, Department of Materials Science, Faculty of Chemistry, University of the Basque Country (UPV/EHU), Centro Joxe Mari Korta, Avenida de Tolosa 72, E-20018 Donostia-San Sebastián, Spain

Extensive molecular dynamics simulations were conducted using the TIP4P/2005 water model of Abascal and Vega [J. Chem. Phys. **123**, 234505 (2005)] to investigate its condensation from super-saturated vapor to liquid at 330 K. The mean first passage time method [J. Wedekind, R. Strey, and D. Reguera, J. Chem. Phys. **126**, 134103 (2007); L. S. Bartell and D. T. Wu, **125**, 194503 (2006)] was used to analyze the influence of finite size effects, thermostats, and charged species on the nucleation dynamics. We find that the Nosé–Hoover thermostat and the one proposed by Bussi *et al.* [J. Chem. Phys. **126**, 014101 (2007)] give essentially the same averages. We identify the maximum thermostat coupling time to guarantee proper thermostating for these simulations. The presence of charged species has a dramatic impact on the dynamics, inducing a marked change towards a pure growth regime, which highlights the importance of ions in the formation of liquid droplets in the atmosphere. It was found a small but noticeable sign preference at intermediate cluster sizes (between 5 and 30 water molecules) corresponding mostly to the formation of the second solvation shell around the ion. The TIP4P/2005 water model predicts that anions induce faster formation of water clusters than cations of the same magnitude of charge. © *2011 American Institute of Physics*. [doi:10.1063/1.3672063]

I. INTRODUCTION

Nucleation is a fundamental process in nature that is responsible for many important life phenomena. Familiar manifestations of nucleation include the formation of clouds, liquid droplets from saturated vapor, fog, and crystals. It is also very important in the chemical industry as it plays a critical role in the preparation of metallic powder-based catalysts and thin layers, among other applications. Physically, nucleation is a rare event that emerges from the sudden local fluctuations in particle density in an otherwise homogeneous phase. Nucleation, which is the first step in many phase transitions, must be distinguished from the so-called spinodal decomposition. The latter is typically characterized by a fast, barrierless growth process in which any fluctuation, however small, will drive the phase segregation in the system. Water condensation is arguably one of the most important nucleation processes due to its biological and environmental implications, such as cloud formation. Creating rain on demand and in general controlling the weather locally is a long standing dream that have captivated human thinking since ancient times. Inducing cloud formation on arid regions of the earth has hitherto mainly focused on cloud seeding with silver iodide or dry ice. Very recently, however, it was demonstrated[1,2] an experimental technique that could artificially trigger rain on demand with enormous potential environmental and economic benefits. That study[1] reported water condensation in sub-saturated

air conditions using high energy femtosecond infrared laser pulses. The researches conducted the experiment under controlled conditions (laboratory chamber) and reproduced it in free atmosphere above Berlin.[2] This finding is remarkable due to the activated nature of the nucleation processes and the simplicity of inducing rain locally using a portable laser device.[1] This experiment paves the way for artificially triggering rain on demand with potentially enormous benefits. The mechanism of this phenomenon was hypothesized to involve electrostatics via the formation of ions, which then would act as nucleation seeds. More recently, the same group put forward a mechanism whereby hygroscopic HNO_3 would stabilize the formation of water droplets,[3] however, the significance of other factors is yet to be elucidated.

Although the thermodynamics is well understood, the nucleation dynamics of polar liquids remains to some extend an unexplored field in molecular physics. The aim of our study is to unveil the mechanism for water condensation observed in the aforementioned experiment[1] and, more generally, to gain atomistic insight of water nucleation via molecular dynamics. Apart from the obvious environmental implications of water nucleation, there are other reasons that motivate our study: (1) Many computational studies have been devoted to study the vapor–liquid transition of simple systems such as Lennard–Jones argon[4–11] fluids but fewer to water condensation.[12–14] (2) Few theoretical and computational studies have been reported on the effect of ions on water condensation.[15–19] (3) Vapor–liquid equilibrium is known to be very sensitive to the interaction potential and the many-body nature of the interaction. Recently, it was reported[20] the impact of various

---

a)Electronic mail: alejandroperezpaz@yahoo.com.
b)Electronic mail: angel.rubio@ehu.es.

intermolecular potentials on the vapor–liquid equilibrium of argon. Thus, it would be interesting to address the effect of polarization for the case of water condensation. (4) Finally, most of the existing literature on this topic uses force fields. It would be interesting to address the importance of many body effects using *ab initio* or First Principles methods.

The goal of this paper is to gain atomistic insight of the nucleation process of water via molecular dynamics (MD). Clearly, addressing all the above points in a single study is an impossible task and in this first paper we address mostly point (1). To this end, we conducted extensive MD simulations of the nucleation process of water in presence and absence of ions. We have also explored how simulation parameters (finite size effects and thermostat coupling) affect the dynamics. The nucleation phenomenon may take place on the millisecond-second time scale, which is far too long for ordinary MD calculations (without enhanced sampling techniques). Computational simulations of nucleation are therefore performed under high supersaturated conditions owing to the extremely low density of vapor and the activated nature of the nucleation process.

This paper is organized as follows: In Sec. II, we review the main results of classical nucleation theory for homogeneous systems (II A). Then, in Subsection II B, we recall the Thomson model, which describes the heterogeneous nucleation in the presence of charged species. The mean first passage method used in the analysis of the nucleation dynamics is reviewed in Subsection II C. To finish with this section, the TIP4P/2005 water model used here is described (II D). In Sec. III, the computational details are given. Then, Sec. IV presents and discusses the main results. The finite size effects including a discussion of the modified liquid drop model are presented in Subsection IV A. Thermostats influence (IV B), and the impact of charged species on the nucleation dynamics (IV C) are also presented. Finally, Sec. V summarizes the main conclusions and future work.

II. THEORY

A. Review of classical nucleation theory

In this section, we review the main results of homogeneous classical nucleation theory (CNT).[21–24] The first step in nucleation is the formation a microscopic cluster, which is assumed to have the same properties as a macroscopic droplet. The free energy difference for the formation of a spherical cluster of radius R in unstable equilibrium with a surrounding vapor phase is

$$\Delta G\,(R) = \left(\frac{4}{3}\pi R^3\right)\Delta g + \gamma_{lv}(4\pi R^2), \tag{1}$$

where γ_{lv} is the surface tension (assumed independent of the curvature of the drop) of the liquid–vapor interface and Δg is the free energy density change, which provides the driving force for the process. If the vapor behaves ideally, we can write the difference in free energy density as

$$\Delta g = -(\rho_l - \rho_v)(\mu_{ss} - \mu_s)$$
$$\approx -\rho_l k_B T \ln \frac{\rho_{ss}}{\rho_s} = -\rho_l k_B T \ln S, \tag{2}$$

where the supersaturation ratio in the nucleating vapor $S \approx \rho_{ss}/\rho_s$ was introduced. In Eq. (2), μ_s and μ_{ss} are the chemical potentials at vapor–liquid coexistence and supersaturation, respectively. Also, the density of the vapor ρ_v was neglected with respect to the liquid density ρ_l because the former is typically three orders of magnitude smaller than the latter.

The critical radius R^* is found from the extremum condition of the Gibbs free energy, $\frac{d\Delta G(R)}{dR} = 0$, leading to the so-called Kelvin radius,

$$R^* = \frac{2\gamma_{lv}}{|\Delta g|} \approx \frac{2\gamma_{lv}v_l}{k_B T \ln S}, \tag{3}$$

where v_l is the molecular volume calculated from the bulk liquid density. The number of molecules in the critical embryo is

$$n^* = \frac{32\pi}{3}\frac{v_l^2\gamma_{lv}^3}{(k_B T \ln S)^3}. \tag{4}$$

Large excursions beyond this critical size are needed to result in a continuous cluster growth to macroscopic sizes. We note that slightly bigger post-critical clusters can still disaggregate and go back to the vapor phase. We mention in passing that the number of molecules n^* in the critical cluster can be estimated experimentally from the slope of the log-log scale of J-S isothermal plots[25] (J is the nucleation rate),

$$\left(\frac{\partial \ln J}{\partial \ln S}\right)_T = 1 + n^*. \tag{5}$$

This relation is often called the "the first fundamental nucleation theorem".[25–27]

The activation free energy to form the critical embryo is obtained from Eq. (1) evaluated at $R = R^*$,

$$\Delta G^* \approx \frac{16\pi}{3}\frac{v_l^2\gamma_{lv}^3}{(k_B T \ln S)^2}. \tag{6}$$

Equation (6) erroneously predicts a nonzero work for the monomer and was later corrected by Girshick and Chiu.[28]

In experiments, the property that is typically measured is the nucleation rate J or the number of macroscopic clusters appearing per unit of volume and time at constant supersaturation. The classical nucleation rate follows the usual Arrhenius expression, $J = K \exp\left(-\frac{\Delta G^*}{k_B T}\right)$, where ΔG^* is given by Eq. (6). The kinetic prefactor K was first calculated by Becker and Döring[23] to be $K = Zr_f\rho_v$, where the quantity

$$Z = \sqrt{-\frac{1}{2\pi k_B T}\left(\frac{\partial^2 \Delta G\,(n)}{\partial n^2}\right)_{n=n^*}}, \tag{7}$$

is the Zeldovich factor and

$$r_f = \sqrt{\frac{2\gamma_{lv}}{\pi m}}\rho_g, \tag{8}$$

is the classical forward rate at the critical embryo derived from kinetic theory (m is the molecular mass). Evaluating explicitly the kinetic prefactor using Eq. (1) leads to the classical nucleation rate (P_{ss} is the supersaturated pressure),

$$J = \sqrt{\frac{2\gamma_{lv}}{\pi m}}\left(\frac{P_{ss}}{k_B T}\right)^2 v_l \exp\left(-\frac{\Delta G^*}{k_B T}\right). \tag{9}$$

B. Nucleation in presence of charged species: the Thomson model

A model for the reduction in the free energy barrier for ion-induced nucleation was first proposed by Thomson[29] who used the reversible work to solvate an ion as given by the Born solvation model,[30]

$$W_E = \frac{\epsilon_0}{2} \int \epsilon(\mathbf{r}) E(\mathbf{r})^2 \, d\mathbf{r} = \frac{q^2 e^2}{8\pi\epsilon_0} \left(\frac{1}{\epsilon_v} - \frac{1}{\epsilon_l} \right) \left(\frac{1}{R} - \frac{1}{R_{ion}} \right), \tag{10}$$

where E is the electric field strength, e is the elementary charge, q is the charge of the ion in units of the elementary charge, ϵ_0 is the the permittivity of the vacuum, ϵ_v and ϵ_l are the static relative dielectric constant of the vapor and condensing liquid, respectively. Finally, R_{ion} is the van der Waals radius of the ion core, assumed to be at the center of a spherical cluster. For the case of condensing water, the parameter R_{ion} can be also taken from the onset of the ion–oxygen radial distribution function.

The combination of the Born reversible work, Eq. (10), with CNT, Eq. (1), defines the Thomson model for heterogeneous nucleation in presence of charged species. Thus, the free energy of formation of a spherical cluster about an ion at its center is

$$\Delta G(R) = - \left(\frac{4\pi k_B T}{3 v_l} \ln S \right) R^3 + (4\pi \gamma_{lv}) R^2$$
$$+ \frac{(qe)^2}{8\pi\epsilon_0} \left(\frac{1}{\epsilon_v} - \frac{1}{\epsilon_l} \right) \left(\frac{1}{R} - \frac{1}{R_{ion}} \right). \tag{11}$$

The equilibrium radius R^* of the droplet formed by liquid condensed around the ion is given by the Thomson equation,

$$\left(\frac{d\Delta G(R)}{dR} \right)_{R=R^*} = 0 = -\frac{4\pi k_B T \ln S}{v_l} R^{*2} + 8\pi \gamma_{lv} R^*$$
$$- \frac{(qe)^2}{8\pi\epsilon_0 R^{*2}} \left(\frac{1}{\epsilon_v} - \frac{1}{\epsilon_l} \right). \tag{12}$$

If we now define the Kelvin radius, $R_K = (2\gamma_{lv} v_l)/(k_B T \ln S)$, which was already introduced by Eq. (3), and the Rayleigh radius, $R_R^3 = (qe)^2 (1/\epsilon_v - 1/\epsilon_l) / (64\pi^2 \epsilon_0 \gamma_{lv})$, we can rewrite the last equation concisely as[16–18,31]

$$\frac{R^*}{R_K} = 1 - \left(\frac{R_R}{R^*} \right)^3. \tag{13}$$

Thomson found that R_R is the radius that minimizes the sum of the second and third terms in Eq. (11). Since heterogeneous nucleation is faster than homogeneous nucleation, the radii obey not surprisingly the relation $R_R < R_K$. Thomson estimated that the parameter R_R was of the order of few angstroms for a monovalent ion in water vapor at ambient conditions, which is inconsistent with the macroscopic quantities involved in its calculation. If we now define the dimensionless radius $x = R^*/R_R$ in Eq. (13), we have the following nonlinear equation:

$$\alpha = \frac{1}{x} - \frac{1}{x^4}, \tag{14}$$

which defines the location of equilibrium reduced radius, $x = R^*/R_R$, for the charged drop-vapor system as a function of the rescaled supersaturation,

$$\alpha = \frac{R_R}{R_K} = \frac{R_R k_B T}{2\gamma_{lv} v_l} \ln S. \tag{15}$$

Although the general solution to Eq. (14) is involved, we can easily determine the maximum value of parameter α by taking derivatives with respect to the dimensionless variable x on both sides of this expression to get $\alpha^* = 3/4^{4/3} \approx 0.472$, which occurs at $x^* = 4^{1/3} \approx 1.587$. The critical size x^* divides the equilibrium $\alpha - x$ plot in two regions: for $x > x^*$, we have an unstable branch where an increase in the supersaturation S (i.e., α is increased) leads to an indefinite growth of the cluster size until vapor is depleted. For $x < x^*$, however, we have a stable branch characterized by an increase of the embryo size with the supersaturation to new stable values given by Eq. (14).[16]

We remark that the Thomson model employs macroscopic quantities and complete agreement with atomistic simulations cannot be expected. This is due to the fact that the latter resolves microscopic details regarding molecular configurations neglected by the former. Finally, we mention that the Thomson model was extended[32] for the case of polar liquids condensing around ions to include the dipole-charge interaction explicitly.

C. The mean first passage time method

In this section, we review the mean first passage time (MFPT) method for rare events proposed in Refs. 9, 33, and 34 For activated processes, such as nucleation events at sufficiently low supersaturations, the average time of first appearance of a cluster of n particles or MFPT is given approximately by

$$\tau(n) = \frac{\tau_J}{2} \{ 1 + \text{erf}[c(n - n^*)] \}, \tag{16}$$

where n^* is the critical number of particles in the embryo (corresponding to the maximum in free energy), and the quantity

$$c = \sqrt{ -\frac{1}{2k_B T} \left(\frac{\partial^2 \Delta G(n)}{\partial n^2} \right)_{n=n^*} }, \tag{17}$$

is proportional to the local curvature at the transition state, which is in turn connected to the Zeldovich factor $Z = c/\sqrt{\pi}$ introduced in Eq. (7). The parameter $\tau_J = 1/(JV)$ is inversely related to the nucleation rate J (V is the volume) and erf() is the error function.

The derivation of Eq. (16) starts from the Fokker–Planck equation and proceeds through a series of approximations including a steepest descent evaluation, which is strictly valid only for high barriers along the reaction coordinate.[33] Eq. (16) can be also derived using the Becker–Döring theory[23] as shown in Ref. 34. The MFPT method was successfully applied to investigate the homogeneous condensation of Lennard–Jones argon systems,[6,7,33,35] the nucleation and growth of zinc clusters from supersaturated vapor,[36] the

formation of $FeCl_2$/NaCl nanoparticles in supercritical water,[37] the melting process in copper,[38] among others.

The MFPT method is very easy to implement in MD simulations of nucleation processes: from a given MD trajectory t, one simply records the time $\tau_t(n)$ at which a cluster of size n makes the first appearance in the system. Then, the MFPT is simply obtained from the average over many trajectories (the number of them, $\mathcal{T}$, is usually over tens or hundreds), $\tau(n) = (1/\mathcal{T})\sum_{t=1}^{\mathcal{T}}\tau_t(n)$, each starting from a different initial condition. The resulting curve has sigmoidal character and can be fitted to Eq. (16). The MFPT method considers only the time of first appearance regardless the number of clusters in the system and whether the successive MFPT are reached or not by the same cluster. Thus, the MFPT analysis is less sensitive to finite size effects than other approaches, such as the Yasuoka–Matsumoto method,[12] where a large box is needed to allow the coexistence of a distribution of clusters of varying sizes. To apply the MFPT method, it suffices a box large enough to contain several times the critical number of molecules $n*$. Another advantage is that the analysis is very simple: the fitting of the $\tau(n)$ data to Eq. (16) yields directly all relevant quantities (τ_J, $n*$, c) of the process. In particular, the parameter τ_J is obtained from the flat plateau region of the plot of $\tau(n)$ versus n. At large values of n, the MFPT curve levels out owing to the mathematical property of the error function $erf(x) \to 1$ as its argument x increases. The critical cluster size $n*$ may be obtained from the inflexion point (zero second derivative) of the ascending part of the sigmoidal curve. At this point, the critical cluster can either evaporate on one side or continue growing on the other side of the free energy profile. Furthermore, the free energy profile of any activated process (including nucleation) can be reconstructed using the MFPT values and the time-independent probability distribution function as proposed in Ref. 39. Such free energy reconstruction method was later applied to the Lennard–Jones system to investigate its condensation from the vapor phase[8] and to characterize its crystallization from the liquid state.[35]

The limitation of the MFPT approach (especially for large systems) is that a large number of trajectories is needed for proper statistics owing to the activated nature of the process under study. Usually, the most difficult part to converge is the plateau region where large variance is often encountered. It is emphasized that the MFPT method is strictly valid for rare events, such a nucleation or crystallization, and only in these cases a good fit to Eq. (16) can be obtained. Deviations of the plateau behavior at larger times indicate that an additional growth phase occurs and is typical of processes with low activation barriers. Nonetheless, even in such low barrier cases, the quantity τ_J (and from it the nucleation rate) can be estimated from twice the MFPT value at the inflection point of the MFPT curve, that is, $\tau_J \approx 2 \times \tau(n = n*)$.[33] In practice, the fitting procedure is as follows: first, the entire MFPT curve is fitted to estimate $n*$. Then, this is followed by another fit of $\tau(n)$ up to $2 \times n*$. If the parameters (τ_J, $n*$, c) do not change, then the fitting is converged, otherwise the previous step is repeated until self-consistency. In our fittings, we found that 3–4 iterations are normally sufficient to achieve convergence.

TABLE I. The TIP4P/2005 force field parameters.[40]

d_{OH} (Å)	d_{OM} (Å)	H–O–H (^{0})	σ_{OO} (Å)	ϵ_{OO} (kJ mol^{-1})	$q_H(e)$	q_M (e)
0.9572	0.1546	104.52	3.1589	0.77490	0.5564	$-2q_H$

D. The TIP4P/2005 water model

The TIP4P/2005 water model of Abascal and Vega[40] is a re-parametrization of the popular TIP4P rigid water model.[41] TIP4P/2005 was applied to predict the phase diagram[42–45] and other physico-chemical properties of water with great success.[46–49] The TIP4P/2005 is a rigid water model with OH bond lengths of 0.9572 Å and a bond angle HOH of 104.52°. The intermolecular potential interaction consists of the usual Lennard–Jones and Coulomb terms,

$$U_{inter} = \sum_{n,l<n} 4\epsilon_{OO}\left[\left(\frac{\sigma_{OO}}{r_{nl}}\right)^{12} - \left(\frac{\sigma_{OO}}{r_{nl}}\right)^6\right]$$
$$+ \frac{1}{4\pi\epsilon_0}\sum_{i,j<i}\frac{q_i q_j}{r_{ij}}. \tag{18}$$

The Lennard–Jones intermolecular interaction is only between oxygen atoms pairs. The partial charges are placed on the hydrogen atoms and on an additional site called "M", which is located midway along the HOH angle bisector, 0.1546 Å away from the oxygen atom. The partial charge on the oxygens is zero. In Eq. (18), r_{nl} and r_{ij} denote the distances between oxygen atoms and charged sites of two water molecules, respectively. The TIP4P/2005 force field parameters are listed in Table I. The physico-chemical properties of TIP4P/2005 water are collected in the Appendix A.

Three ingredients, taken from different water models, were used as a guiding principles in the design of the TIP4P/2005 force field:[50] (1) Nearly the same charge distribution as the TIP4P model[41] was adopted owing to its adequate prediction of the phase diagram of water. (2) The SPC/E polarization energy term due to Berendsen *et al.*[51] to correct the vaporization enthalpy of the SPC model[52] was used. Namely, the correction term is $E_{pol}/N = (d - d_g)^2/(2\alpha)$, where α is the polarizability of the water molecule, d_g is its dipole moment in the gas phase, and d is the dipole moment of the model. (3) Following the TIP5P model,[53] the maximum in the density of experimental water at room pressure was chosen as a target property in the parametrization. In addition, the density of several ice polymorphs and the melting point of ice Ih was included in the the fitting.

Overall, TIP4P/2005 represents one of the best rigid non-polarizable water models. It predicts a correct phase diagram, which is very important in the simulation of nucleation processes as various phases are involved in the phase transition. Furthermore, recently[54] SPC/E and TIP3P models were shown to give wrong structures for water clusters in gas phase, while the TIP4P related models seem to give qualitatively correct geometries. This issue is important as it may lead to the wrong initial embryo and alter the subsequent dynamics of the nucleation process. For all these reasons, we adopted the TIP4P/2005 water model in all our MD simulations.

However, at present there is no perfect water model, and therefore TIP4P/2005 has also its limitations: (1) The agreement with the experimental heat capacity of water is poor due to the partial neglect of nuclear quantum effects. Nonetheless, this limitation has been partially alleviated in a new re-parametrization of the model called TIP4PQ/2005,[55] to be used in path integral calculations.[56] (2) As in many water models, the TIP4P/2005's electric dipole moment (2.3054 D) is designed to mimic the one in bulk water rather than the one in isolated gas phase water molecule (1.85 D). As a result, the TIP4P/2005 fails to describe some vapor properties, including second virial coefficient, critical pressure, and vapor pressure. We should mention that there exists another re-parametrization of the popular TIP4P model called TIP4P-Ew[57] suited for its use with Ewald techniques and was shown to have slightly better vapor properties than the TIP4P/2005 model.[45,58]

III. COMPUTATIONAL DETAILS

All MD calculations were carried out using the GROMACS (version 4.5.3) package[59] compiled in double precision. The TIP4P/2005 water molecules were placed in a cubic box and periodic boundary conditions were used in all directions.[60] Supersaturated water conditions were used (as is usual in practically all MD simulations of nucleation) due to the extremely low density of saturated water vapor and the activated nature of the critical embryo formation. Although any quantitative comparison of simulation with experimental results impossible, nonetheless, valuable conclusions can still be drawn from the simulations. After some initial tests, we found that a particle density of 2.7×10^{-5} Å^{-3} at 330 K has the appropriate combination of parameters for the MFPT analysis. Since the MFPT procedure requires a large number of trajectories, the physical conditions were chosen so that the plateau region in MFPT curve could be observed in a reasonable simulation time scale. We remark that lower supersaturations would require not only longer simulation times, which would affect negatively energy conservation, but also larger computational cells to minimize finite size effects. Of course, we note that there are many possible combinations of physical conditions that yield the same MFPTs and there is no real preference in our choice. At our physical conditions, the parameters predicted by CNT for experimental water are $S = 7.146$, $\Delta G^*/k_B T = 12.6012$, $n^* = 12.82$, whereas the corresponding values for the TIP4P/2005 model are $S = 25.735$, $\Delta G^*/k_B T = 4.238$, and $n^* = 2.61$. The physico-chemical properties of TIP4P/2005 and experimental water needed for CNT and Thomson models are listed in the Appendices A and B, respectively. A minimum of 216 water molecules in a cubic box of 200 Å^3 is appropriate starting size for MD simulations. According to CNT applied to experimental water, the size of the critical embryo is about 16 times smaller than the number of molecules in the box.

A cutoff of 99 Å was used for all non-bonded interactions (van der Waals and electrostatics) and systems investigated, except for the smallest one ($N = 108$, side $L = 158.7401$ Å) where a cutoff of 79.3 Å was used to be consistent with the minimum image convention. Such a large cutoff was carefully chosen to reproduce the same energies, pressure, and forces as a fully converged Ewald calculation. This choice of cutoff was tested on several configurations (vapor and condensed phases). We also checked that the obtained MFPT plots are very similar in either case (Ewald versus cutoffs). The use of cutoffs (instead of Ewald/PME techniques) allows us to carry out nanosecond-long simulations in few hours on a single processor. In contrast, the use of Ewald/PME techniques on the same system increases the computational effort by two orders of magnitude (more than a week on a single processor). We also verified that the 99 Å cutoff was also converged when simulating systems of larger size and in the presence of ions. We remark that energy conservation issues arise at long simulation times if the cutoffs used in the parametrization of the water models are employed in our simulations. The interaction energy and forces were calculated at each time step. The translation of the center of mass was removed every time step.

The OH bonds and HOH angles were constrained using the SETTLE algorithm[61] to allow for larger MD time steps. The neglect of intramolecular modes has hardly any impact on the nucleation process, which takes place on a much longer time scale. Interestingly, however, the experimental study of Ref. 62 reported that the nucleation rates of D_2O is 2500 times larger than those of H_2O at the same vapor pressure and temperature. However, the same study showed that the D_2O rates superimpose within experimental error with those of H_2O when compared at the same supersaturation S and temperature T. It is remarked that the TIP4P/2005 was parametrized by fitting classical simulation data directly to experimental data. Therefore, this empirical force field implicitly includes (although partially) some quantum effects.

First, a 20 ns long equilibration of the vapor phase was carried out in the canonical (NVT) ensemble at 1000 K. We checked stability of the average total energy and other thermodynamic properties by performing a short MD simulation in the microcanonical (NVE) ensemble. The average pressure was stable at 3.657 atm (3.7056 bar) and very close to its ideal value of 3.679 atm (3.7277 bar). These conditions are well below the calculated critical point for the TIP4P/2005 model ($T_c = 640$ K, $P_c = 146$ bar)[45] and for the experimental critical point (647.096 K and 221.2 bar). Equilibration conditions away from the critical point are necessary to avoid excessive formation of small clusters in the vapor phase. These clusters might act as nucleation centers for the formation of liquid droplets. At 1000 K, TIP4P/2005 water vapor is mostly found in monomers with some occasional dimers. We output hundreds of vapor configurations spaced evenly every 100 ps. This time interval is sufficiently long to avoid any structural correlation between adjacent configurations and ensure a good sampling of the initial conditions. Since the trajectories are assumed totally independent, the error bars were computed using the standard deviation formula.

These equilibrated vapor configurations were subsequently used as initial conditions for nonequilibrium MD simulations. Velocities were re-sampled from the Maxwell–Boltzmann distribution at a temperature of 330 K. Canonical ensemble (NVT) conditions were maintained using the thermostat proposed by Bussi et al.,[63] unless otherwise stated. The typical length of the nonequilibrium MD simulations

was 3 ns and snapshots of the system were output every 0.5 ps for the MFPT analysis. Longer simulation times (up to 6 ns) were needed to investigate the dynamics in the weakly-thermostatted regime (see later, Subsection IV B). As nucleation proceeds, large density fluctuations are expected, which may affect adversely the energy conservation. A small time step of 0.5 fs was chosen in all MD simulations to ensure stable integration of the equations of motion. The typical energy drift after 3 ns of MD simulation was small and about 8 kJ/mol per 2000 degrees of freedom.

A cluster analysis code was written to post-process the MD trajectories. Different cluster definitions are available and two of the most commonly used[5,64] were discussed in Ref. 11 for the simple case of Lennard–Jones argon systems. The choice of cluster criterion is irrelevant as far as attractive pair potentials are concerned. The popular Stillinger criterion[64] was mostly used in our analysis to determine whether a pair of particles is connected or not. In this definition, two water molecules form a cluster if the distance between the oxygen sites of each molecule is less than $r_S = 3.36$ Å. This threshold distance (or Stillinger radius, r_S) corresponds approximately to the first minimum in the oxygen–oxygen radial distribution function of the TIP4P/2005 water model (see Appendix A).[40] In Appendix C, we compare the Stillinger versus the ten Wolde and Frenkel cluster definitions for the TIP4P/2005 water model. In addition, we show the effect of other value for r_S on the MFPT curves.

IV. RESULTS AND DISCUSSION

A. Finite-size effects

In this section, we analyze the finite-size effects in the water nucleation following the discussion of the modified liquid drop (MLD) model in Ref. 9. Studying bulk systems undergoing a phase transition remains a challenging task in chemical physics due to long-ranged nature of the correlations in the system. The resolution of geometrical quantities is typically limited to the cell size used in the simulations. The low frequency Fourier modes are suppressed and one has to make sure that the property of interest does not depend critically on these components by performing simulations at increasing values of the computational cell dimensions.

The CNT expression given by Eq. (1) is strictly valid for open systems, where the supply of vapor molecules is assumed unlimited and the condensation of molecules does not lead to a significant decrease in the supersaturation. If we express Eq. (1) in terms of the number of particles n in the cluster, we have

$$\Delta G(n) = -nk_{\mathrm{B}}T \ln \frac{P_{ss}}{P_s} + \left(36\pi v_l^2\right)^{1/3} \gamma_{lv} n^{2/3}$$
$$+ n\left(k_{\mathrm{B}}T - P_s v_l\right), \quad (19)$$

where the quantity $P_{ss} = Nk_{\mathrm{B}}T/V$ is the supersaturated pressure of the gas (V is the volume of the computational cell) and P_s is the equilibrium vapor pressure over a flat surface. The additional third factor in Eq. (19) is the $P - V$ work required to form a cluster against the vapor and usually is a negligible quantity.

In closed systems, however, the supersaturation S is a function of the cluster size n because as the cluster growths, the vapor phase is depleted from water molecules. This is the typical situation of MD calculations in the NVT ensemble where one has to decide *a priori* the number of molecules N to place in a computational cell of volume V. References 9 and 65 introduced the MLD model, which gives the reversible work of formation of a cluster of n particles corrected for finite size effects,

$$\Delta F_{\mathrm{MLD}}(n) = -nk_{\mathrm{B}}T \ln \frac{P_{act}(n)}{P_s} + \left(36\pi v_l^2\right)^{1/3} \gamma_{lv} n^{2/3}$$
$$+ n\left(k_{\mathrm{B}}T - P_s v_l\right) + Nk_{\mathrm{B}}T \ln \frac{P_{act}(n)}{P_{ss}}$$
$$= \Delta G(n) + (N - n)k_{\mathrm{B}}T \ln \frac{P_{act}(n)}{P_{ss}}, \quad (20)$$

where $P_{act}(n) = (N - n)k_{\mathrm{B}}T/(V - nv_l)$ is the actual ideal vapor pressure in a closed system of $(N - n)$ vapor molecules and a cluster of n particles. From Eq. (20), we see that usually $\Delta F_{\mathrm{MLD}}(n) > \Delta G(n)$, which means that finite size effects typically overestimate activation free energy barriers. We also note that the critical size n^* of the MLD model approaches the CNT critical size as the finite size effects are reduced. As far as nucleation rates is concerned, convergence in size is assessed by checking how close the MLD curves are to the CNT curve at the maximum of free energy. Differences beyond this point are irrelevant for the nucleation dynamics because it is assumed that once passed this point the condensation proceeds down-hill.

Figure 1 (top) shows the free energy barriers for experimental water at 330 K and fixed supersaturation $S = 7.146$ for different simulation box sizes. At this supersaturation, the critical size is approximately the same by all curves ($n^* \approx 12.8$), except that for the smallest system ($N = 108$, $L = 158.7401$ Å) which is clearly unconverged with respect to finite size effects. For experimental water, Fig. 1 (top), we expect that the free energy barrier to be reasonably converged (within 0.2 kcal/mol) for a system size of 432 water molecules in a cubic box of length $L = 251.9842$ Å. Clearly, simulating a larger system, such as the ($N = 1728$, $L = 400$ Å) would be a waste computationally and would not yield any new information. Fig. 1 (bottom) also shows the free energy barriers for TIP4P/2005 water at 330 K and fixed supersaturation $S = 25.735$ for different simulation box sizes. As mentioned before, the enhanced value of the dipole moment of the TIP4P/2005 model results in a poor description of the vapor properties of water. The TIP4P/2005 vapor pressure at 330 K ($P_s \approx 0.0478$ bar) is significantly lower than the experimental value ($P_s \approx 0.1721$ bar). As a result, the nucleation free energy barriers are severely underestimated for TIP4P/2005 model and CNT predicts a smaller critical cluster size $n^* = 2.6$. We note that according to CNT, finite size effects are predicted to be of less relevance for TIP4P/2005 water due to the small value of the free energy barrier ($\Delta F^*/k_{\mathrm{B}}T = 4.238$) [see Fig. 1 (bottom)].

Figure 2 shows the effect of system size on the MFPT for TIP4P/2005 water nucleation at 330 K and fixed supersaturation $S = 25.735$. For each system size, we averaged
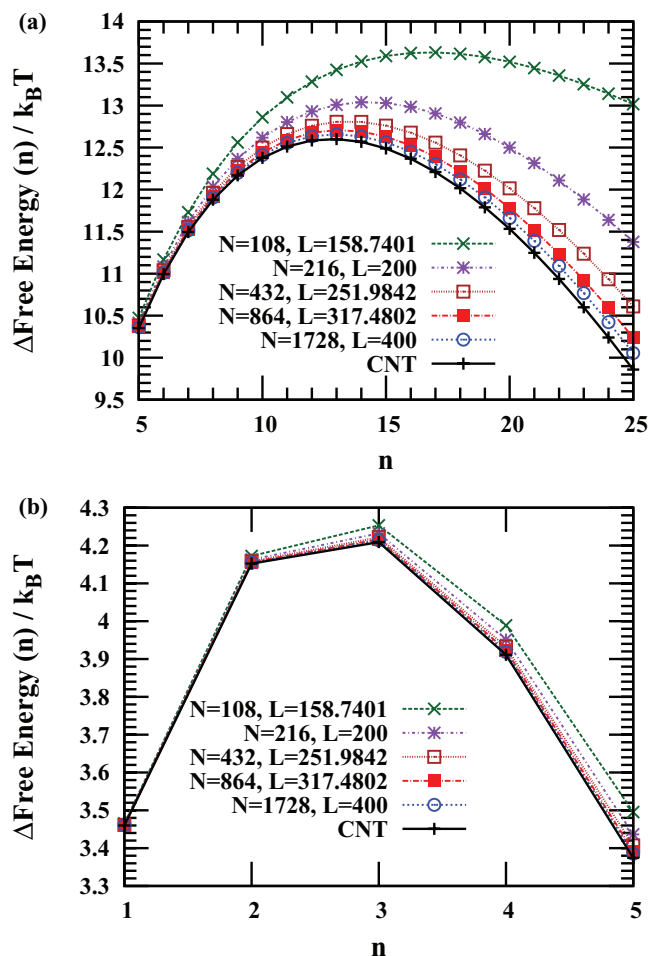
FIG. 1. Assessing finite size effects in experimental (top) and TIP4P/2005 (bottom) water nucleation at 330 K and fixed particle density 2.7×10^{-5} Å^{-3}. Free energy profiles as a function of the cluster size n according to classical nucleation theory [Eq. (19)] (CNT in legend) and the modified liquid drop model [Eq. (20)] (N and L values, in legend). At this particle density, the experimental water has a supersaturation of $S = 7.146$ (top) while the TIP4P/2005 water model exhibits a supersaturation of $S = 25.735$ (bottom).
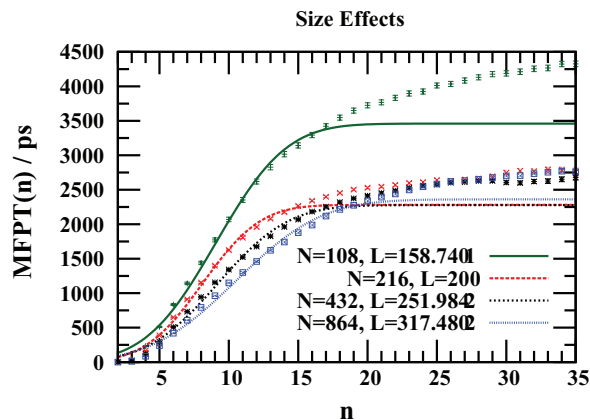
FIG. 2. MFPTs as a function of the cluster size n for the TIP4P/2005 water model at 330 K and fixed supersaturation $S = 25.735$ ($S = 7.146$, for experimental water) at various system sizes. In legend, N is the number of water molecules in a cubic box of side length L (given in Å).

TABLE II. Finite size effects on the nucleation of TIP4P/2005 water model: parameters obtained from the fitting to Eq. (16).

N, L (Å)	$J \times 10^{25}$ (cm^{-3}s^{-1})	τ_J (ps)	n^*	c
108, 158.7401	7.22	3459.4	9.1	0.176805
216, 200.0000	5.48	2280.7	8.1	0.214934
432, 251.9842	2.74	2276.8	9.2	0.177014
864, 317.4802	1.32	2360.5	10.5	0.145692

the MFPT values over 600 independent trajectories, which is sufficient to obtain converged data. The target temperature was maintained using the stochastic thermostat proposed by Bussi *et al.*[63] (see Subsection IV B) with a coupling time of ζ = 5 ps. The MFPT curves were fitted according to Eq. (16) and the resulting parameters are collected in Table II.

As predicted in Fig. 1, we observe that for $N = 108$, $L = 158.7401$ Å (which overestimates the CNT free energy barrier) leads not surprisingly to longer MFPT at intermediates sizes. This behavior, as explained before, has its origin in the the depletion of the water molecules in the vapor phase. The case $N = 108$, $L = 158.7401$ Å shows that the MFPT on the "plateau" region is indeed affected, showing a drastic increase in MFPT values due to depletion effects. We observe that the asymptotic behavior of MFPT values at larger sizes n is generally well reproduced for the rest of cases (except for the smallest system $N = 108$). However, we do not find the expected trend that the rate should increase with the number of particles (see Table III).[9]

The most likely explanation of this odd behavior was pointed out already by Wedekind *et al.* in Ref. 9. That study reported an occasional decrease in the rates for higher number of particles. The authors attributed this anomaly to the appearance of other competing factors such as thermostating issues (see also Ref. 66). The thermostat problem arises in large systems where there exists a coexistence of various clusters, especially of large size, while the thermostat conditions are kept the same. This results in a poor dissipation of latent heat for larger systems (the heat transfer becomes inefficient) which leads to a decrease in the rates. Another possible explanation lies in uncertainties in the fitting due to a superimposed phase of growth that complicates the accurate determination of rates. Finally, we would like to point out that similar trend in nucleation rates was found in other studies of Lennard–Jones systems: namely, see Table I (at supersaturation $S = 12.83$) in Ref. 8 and Table II (at $S = 15.4$) in Ref. 7. We note that larger fluctuations are typically observed in the plateau

TABLE III. Effect of thermostat coupling time constant ζ on the TIP4P/2005 water nucleation ($N = 216$, $L = 200$ Å): fitted parameters according to Eq. (16).

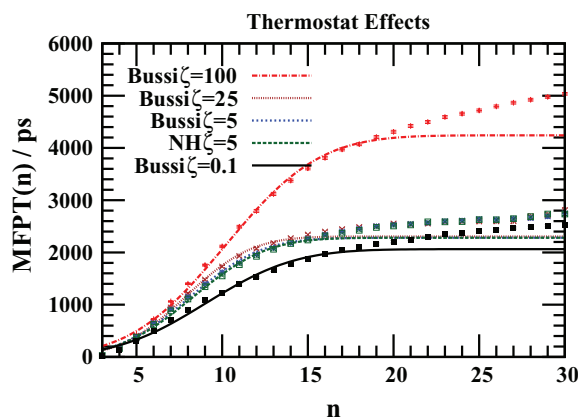| ζ (ps)	$J \times 10^{25}$ (cm^{-3}s^{-1})	τ_J (ps)	n^*	c
0.1 (Bussi)	6.06	2063.2	9.1	0.17260
5 (NHC)	5.48	2281.9	8.3	0.20821
5 (Bussi)	5.48	2280.7	8.1	0.21493
25 (Bussi)	5.42	2306.2	7.9	0.23398
100 (Bussi)	2.94	4241.9	10.2	0.16329

FIG. 3. MFPTs versus cluster sizes n for 216 TIP4P/2005 water molecules ($L = 200$ Å) at 330 K and supersaturation $S = 25.735$ ($S = 7.146$ for experimental water) at different values of the thermostat coupling time constant ζ (in ps). In the legend, NH refers to the Nosé–Hoover thermostat.

region (that is, a wide spread of nucleation rates) and are due to the activated nature of the nucleation events investigated here.

B. Influence of thermostats

Computational studies of nucleation dynamics face two conflicting goals: on the one hand, we need to maintain constant temperature as most experiments are performed at these conditions, and on the other hand, we want the thermostats not to disrupt much the nonequilibrium dynamics of the process we are interested in. Therefore, the choice of thermostat parameters in the condensation simulations of vapor to liquid must be carefully evaluated. The effect of different thermostat techniques on the nucleation of Lennard–Jones argon system was discussed in Ref. 10. That study found that nucleation dynamics of that system does not depend critically on the thermostat details and concluded that direct thermostat methods are equally effective for thermalization and less expensive than the carrier gas technique.

Figure 3 shows the MFPT curves for TIP4P/2005 water at a supersaturation $S = 25.735$ ($S = 7.146$ for experimental water) as a function of the cluster size n using different values of the thermostat time coupling constant ζ. This parameter determines the efficiency of the thermostat; that is, how often the thermostat acts on the system. To this end, we followed the nucleation of 216 TIP4P/2005 water molecules in vapor phase placed in a cubic box of side length 200 Å after a velocity quench to 330 K. For each thermostat ζ parameter, we averaged the MFPT curves over 600 independent trajectories, which is sufficient to converge the error bars. We fitted the averaged MFPT curves according to Eq. (16) and the resulting parameters are collected in Table III. First, we show that for $\zeta = 5$ ps, the deterministic Nosé–Hoover thermostat[67] gives essentially the same MFPT results as the stochastic thermostat proposed by Bussi et al.[63] Therefore, either the Nosé–Hoover or Bussi thermostats can be used to maintain the temperature constant in simulations of water nucleation processes. We recall that both thermostats were shown to reproduce rigorously

the NVT ensemble. The nucleation rates in Table III agree reasonably well with the ones reported in the MD nucleation studies on SPC/E water at 330 K[14] on the TIP4P model at 350 K.[13] However, the nucleation rate predicted by CNT from Eq. (9) for TIP4P/2005 model is 3.757×10^{22} cm^{-3}s^{-1}, and 8.895×10^{18} cm^{-3}s^{-1} for experimental water. These CNT rates are several orders of magnitude smaller than the calculated from MD simulations (see Table III). This discrepancy is a common finding of MD studies of nucleation processes[4, 13] and is usually attributed to the neglect of the cluster-size dependence of the thermodynamic quantities in CNT expressions.

We repeated the MD simulations for increasing values of ζ parameter using the Bussi thermostat and monitor its impact on the MFPT values. Fig. 3 shows that the $\zeta = 25$ ps curve yields very similar nucleation rates (and also n^*, τ_J, c parameters) as the $\zeta = 5$ ps case and both ζ values could be equally used in the simulation. If we further increase the coupling time to $\zeta = 100$ ps, we observed that the thermalization deteriorates due to the infrequent exchange of energy between the physical system and the thermostat. Poor thermalization results in reduced dissipation of latent heat, which inhibits further growth of larger clusters as is evident by the greater MFPT values for the $\zeta = 100$ ps curve. Greater values of the parameter ζ were not investigated due to the increasing difficulty of thermostating the system. Finally, we have also investigated the strongly thermostated case ($\zeta = 0.1$ ps) and found the expected result that the cluster formation occurs more rapidly.

C. Effect of charged species

In this section, we consider the heterogeneous condensation of water in the presence of a single charged species. Interest in ion-induced condensation in supersaturated water vapor goes back to the pioneering work of Wilson a century ago.[68] The most common nucleation species found in atmosphere are NO_3^-, O_2^-, and HSO_4^- among the anions and various forms of solvated proton $H^+ \cdot (H_2O)_n$ for cations, although their concentration varies significantly with the altitude.[69, 70] We remark that the nature of the species in atmospheric nucleation is still a matter of debate.

Here, we consider the effect of the charge on the condensation dynamics, without worrying much about the true identity of the ion. The TIP4P/2005 water model was simulated in the presence of a single ion to avoid simulation artifacts related to image ions getting too close. The ion parameters were taken from the OPLS force field[71] (entry: opls_401). The Lennard–Jones parameters of the ion are $\sigma = 0.441724$ Å and $\epsilon = 0.492833$ kJ/mol, which correspond to the chloride anion. Although, these quantities probably need to be reparametrized for the present water model, the purpose here is solely to assess the role of the charge in the nucleation dynamics. Both positively and negatively charged ions were considered, while the Lennard–Jones parameters were kept constant. We also performed a control MD simulation where the charge of the ion was set to zero (heterogeneous condensation in the presence of uncharged species). In this case, we could
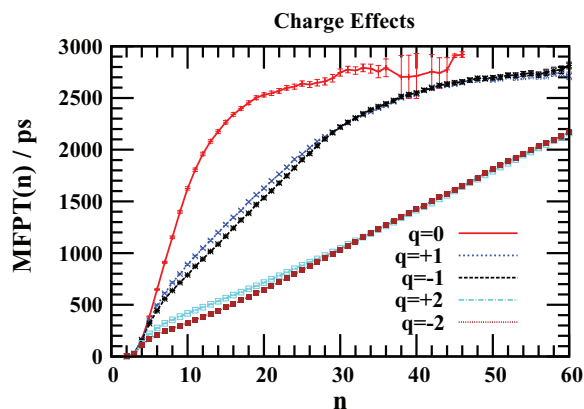
Charge Effects

FIG. 4. MFPTs as a function of cluster size n for 216 TIP4P/2005 water molecules at 330 K in a cubic box of side length 200 Å in presence of a single ion of varying charge. Each curve is an average over 600 independent trajectories.

reproduce the MFPT results already presented in previous sections on the homogeneous water nucleation.

Figure 4 shows the MFPT values versus cluster size n for the TIP4P/2005 model in presence of a single species of varying charge. When the charge is different of zero, the system is mostly growing than nucleating, which is more evident for the divalent ion case, where a marked change in behavior is observed. The Thomson model predicts a barrierless process for the nucleation of TIP4P/2005 water in presence of both monovalent and divalent ions. The divalent ion induces a stronger electric field than the monovalent ion and leads to even shorter MFPT values during aggregation. Fig. 4 suggests that marked time scales are involved in the condensation process induced in absence of ions: first, a short period corresponding to the formation of the critical embryo ($n \approx 9$ molecules) occurs, which is then followed by a transient evolution, and finally by a superposition of nucleation and growth of this cluster. For divalent ions, on the other hand, only one time scale is observed in this size interval and characterizes the continuous growth of water clusters. We expect, nonetheless, that a steady state is reached eventually at very large cluster sizes where the charge of the ion is screened. This behavior is observed for the monovalent ion, which features a situation "in between" these two cases, and its curve levels out at about $n \approx 45$. Thus, the presence of a single ion does have a huge impact on the condensation of supersaturated water.

Since the activation barriers are much reduced in the presence of charged species, the application of Eq. (16) to the fitting of MFPT values is invalid because it requires that the event to be rare. The critical radii according to the Thomson model described in Subsection II B were calculated. For the monovalent ion, we have $R_R = 4.07588$ Å, $R_K = 4.53037$ Å, $\alpha = 0.899679$, whereas for the divalent ion, we have $R_R = 6.47006$ Å, $R_K = 4.53037$ Å, $\alpha = 1.42815$. These values for α are greater than the critical scaled supersaturation, $\alpha* = 0.47247$. As a result, it is not surprising that the system growths rather than nucleates.

The asymmetry in condensation rates for water and other polar liquids has been widely discussed in the past. The Thomson model predicts that the electrostatic interaction of

a water molecule with a positive charge should be similar to that with a negative charge. That is, due to the square dependence on the charge, the Thomson (Born) model predicts the same the reduction in the free energy barrier for positive and negatively charged ions. However, the Thomson model employs macroscopic quantities, and agreement with atomistic simulations cannot be expected because latter resolves microscopic details regarding molecular configurations neglected by the former. In reality, asymmetries have been reported about the sign preference of water nucleation around ionic centers. The molecular geometry of water molecule leads to sign preference on the nucleation and was found to decrease as the supersaturation increases.[72,73] Using the rigid non-polarizable TIP4P/2005 force field, we do observe some noticeable differences at intermediate times (see Fig. 4). In particular, negatively-charged ion seems to induce faster (lower MFPT values) formation of water clusters than positively-charged ion at intermediate times. Using umbrella-sampling Monte Carlo method, Oh *et al.*[74] also found that anions are better water nucleators than cations with the same charge. This can be explained qualitatively due to the presence of two positive centers in water as opposed of only one negative center. The sign preference is indistinguishable at longer times due to the screening of first and second solvation shells. That is, the sign preference does not extend beyond the second solvation shells. A strong ion field compresses the first hydrogen shell making the water–water interaction repulsive. It was found that the local water structure in the clusters becomes perturbed to a larger degree around negative ions compared to ions carrying a corresponding positive charge.[19] The obtained work of cluster formation for the anions was found to be consistently less than that for the cations.[19]

Figure 5 shows snapshots after 2 ns of MD simulation. The water molecules solvate the ion preferentially to minimize the ion-dipole Coulomb interactions. The biggest water cluster was found to form around the ions. The shape of the embryo is more ellipsoidal than spherical. The solvation of anions resembles more the natural structure of bulk water than the cations.

We state some limitations of our approach: TIP4P/2005 water model lacks flexibility and true polarization effects. Previous studies have underscored the importance of including polarization effects in ion-induced nucleation studies.[15]
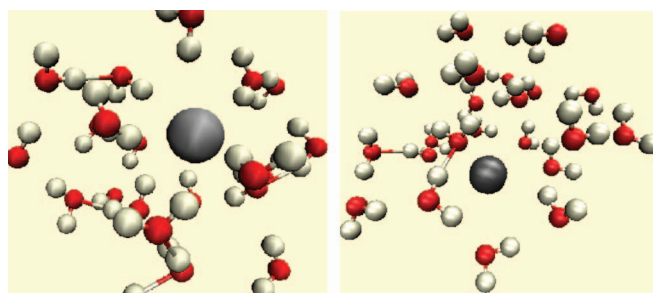

FIG. 5. Snapshots (created with the visualization program VMD[75]) of the embryo after 2 ns of MD simulation at 330 K. Nucleation centers: monovalent anion (left), monovalent cation (right). The "M" sites and other TIP4P/2005 water molecules have been removed for clarity. The ion is depicted as a big gray sphere.

Therefore, to describe the charged cluster-vapor dynamics properly, a fully polarizable force field is required. To this end, we performed preliminary MD simulations using the fully polarizable AMOEBA[76] force field as implemented in the TINKER MD package.[77] We found that the AMOEBA water nucleates on a much longer time scale than the TIP4P/2005 water model. In our preliminary MD simulations, we did not observe clusters larger than 3 or 4 AMOEBA water molecules after passed 2 ns using the same setup as the TIP4P/2005. We noticed than the total potential energy was positive, whereas in the TIP4P/2005 the interaction was attractive. The AMOEBA water system remains frozen in the vapor phase and is reluctant to nucleate, having very few water molecules engaged in hydrogen bonding after 2 ns simulation. Currently, we are investigating this issue, as well as the difference in condensation dynamics for various water force fields. Obviously, the faster nucleation of the TIP4P/2005 model compared to AMOEBA is related to the differences in vapor pressure. Finally, we remark that the AMOEBA simulations are much more expensive (typically two orders of magnitude) than non-polarizable MD calculations.

V. CONCLUSIONS AND FUTURE WORK

Extensive molecular dynamics simulations were performed using the rigid non-polarizable TIP4P/2005[40] water model to investigate its condensation from supersaturated vapor to liquid at 330 K. The mean first passage time method[33, 34] was used to analyze the influence of finite size effects, thermostats, and charged species on the nucleation dynamics. At supersaturation conditions of $S = 25.735$ ($S = 7.146$ for experimental water), it was found that a cubic box of 432 molecules and side length $L = 251.9842$ Å is practically converged with respect to finite size effects. We find that the Nosé–Hoover thermostat and the one proposed by Bussi *et al.*[63] give essentially the same MFPT averages. We found that the coupling time constant of the thermostat does not have a very significant effect on the water nucleation, unless very large (or small) values are used, and we identified adequate values for MD studies of water nucleation. The presence of a single charged species, however, has a dramatic effect on the condensation dynamics, which underscores the importance of ions in the formation of liquid droplets in the atmosphere. It was found a small but noticeable sign preference

at intermediate times (n between 5 and 30 water molecules) corresponding mostly to the formation of the second solvation shell around the ion. The TIP4P/2005 water model predicts that anions are found better cluster inducers than cations of the same magnitude of charge in agreement with experiment.

It will be also interesting to address the effect of polarizability in the nucleation rates. In particular, we will investigate nucleation dynamics of water using the AMOEBA polarizable force field.[76] The effect of an external electric fields, presence of foreign gases, lower supersaturations, and ions are worth further investigation. Work is underway in our group to elucidate their role in the nucleation dynamics of water.

APPENDIX A: PROPERTIES OF THE TIP4P/2005 WATER MODEL

Table IV lists the thermo-physical properties of the TIP4P/2005 water model as obtained from computational studies. All values were taken from Ref. 45 unless otherwise indicated. In that study, the vapor pressure of TIP4P/2005 water model at 330 K was computed to be $P_s = 0.0478$ bar and the liquid density in equilibrium was $\rho_l = 0.9841$ g/cm^3 using the Gibbs–Duhem method. At 330 K, the surface tension was estimated to be $\gamma_{lv} = 64.88$ dyn/cm.[49] From our water/interface MD slab calculations, we obtained

TABLE IV. TIP4P/2005 water properties.[40]

| Property	Value ($x = 1 - T/T_c$)
Electric dipole	d = 2.3054 Debyes
Melting temperature (at 1 bar)	$T_m = 252.5$ K
Boiling temperature (at 1 bar)	$T_b = 401$ K
Critical temperature	$T_c = 640$ K
Critical pressure	$P_c = 146$ bar
Critical density	$\rho_c = 0.337$ g/cm^3
Vapor pressure	$P = \exp[12.4612 - 4476.552/(T - 41.49840)]$ bar, (T in K)
Surface tension[a]	$\gamma_{lv} = 227.86(1 - \frac{T}{641.4})^{11/9}[1 - 0.6413(1 - \frac{T}{641.4})]$ dyn/cm, (T in K)
Relative permittivity[b]	$\epsilon = 58$. At 298 K and 1 bar.

[a]Reference 49.
[b]Reference 50.
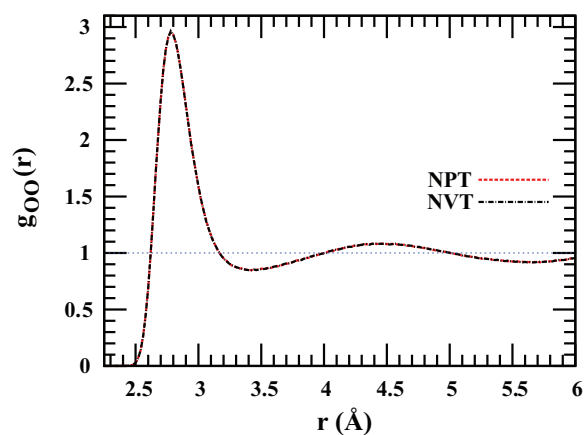
FIG. 6. The oxygen–oxygen radial distribution function for $N = 360$ TIP4P/2005 water molecules at 330 K in different ensembles.

results that agree well with the aforementioned values: γ_{lv} = 63.0 dyn/cm (including a long-range dispersion correction), $\rho_l = 0.9845$ g/cm^3, and $P_s = 0.0463$ bar. We note that converging such low vapor pressures is difficult in ordinary MD simulations. We estimated the relative permittivity of TIP4P/2005 to be about 52.42 at 330 K.

We computed the oxygen–oxygen radial distribution function for $N = 360$ TIP4P/2005 water molecules at $T = 330$ K and is given in Fig. 6. The target pressure for the isothermal-isobaric (NPT) ensemble was set to 1 bar. Identical results are obtained for both NVT and NPT ensembles.

APPENDIX B: EXPERIMENTAL PROPERTIES OF WATER

The experimental physico-chemical properties of water are listed in Table V. Using this table, the properties of experimental water take the following values at 330 K: $P_s = 0.1721$ bar, $\rho_l = 0.9843$ g/cm^3, $\gamma_{lv} = 66.78$ dyn/cm, and the relative permittivity $\epsilon = 67.59$.

TABLE VI. Cluster definitions for 216 TIP4P/2005 water molecules in a cubic box of 200 Å side at 330 K: fitted parameters according to Eq. (16).

n_{tWF}, r_S (in Å)	$J \times 10^{25}$ (cm^{-3} s^{-1})	τ_J (ps)	n^*	c
1, 3.36	5.48	2280.7	8.1	0.21493
1, 3.50	5.47	2282.7	8.2	0.21586
2, 3.36	5.65	2210.8	7.3	0.22771
2, 3.50	5.54	2256.2	7.5	0.21736
4, 3.36	4.96	2517.9	3.2	0.24160
4, 3.50	4.89	2557.7	3.9	0.20775

APPENDIX C: CLUSTER DEFINITIONS

In this Appendix, we compare the ten Wolde and Frenkel[5] (tWF) and the Stillinger[64] cluster definitions for the TIP4P/2005 water model. The tWF definition,[5] which derives from the Stillinger criterion, considers that a particle belongs to a cluster if such a particle has a minimum number of neighbors (n_{tWF}) within a Stillinger distance r_S. The tWF criterion is known to underestimate the size of the physical cluster (including its critical size) because water molecules on the surface of the forming cluster are neglected. This poses a problem especially when the free energy barrier is small and therefore the critical radius is also small. Nonetheless, both definitions should yield essentially the same nucleation rates in the limit of large cluster sizes and large free energy barriers.

In Fig. 7, we observe the aforementioned behavior. As expected, the MFPT times for the $r_S = 3.36$ Å are slightly larger than for $r_S = 3.50$ Å which is more evident for n_{tWF} = 4. Table VI displays the fitted parameters according to Eq. (16) and shows how the critical cluster size n^* depends significantly on the cluster definition. The tWF criterion systematically predicts smaller cluster sizes (especially for n_{tWF} = 4) than the Stillinger definition. We note that the tWF curves are subject to larger fluctuations, which makes the fitting of the MFPT for large cluster sizes somewhat difficult. We notice that nucleation process is accompanied by some degree of surface growth, which is more evident when the Stillinger definition is used where significant deviations of the
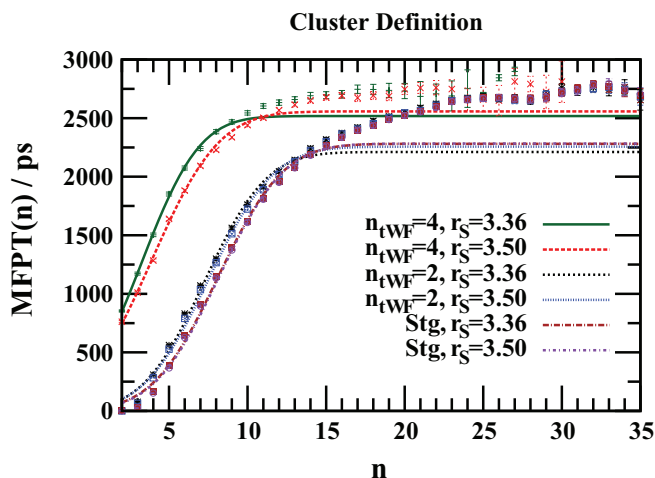
TABLE V. Experimental water properties.

Property	Value ($x = 1 - T/T_c$)
Mass	M = 18.0152833 kg/mol
Electric dipole	d = 1.850 Debyes
Critical temperature	T_c = 647.096 K
Critical pressure	P_c = 221.5 bar
Critical density	ρ_c = 0.322 g/cm^3
Vapor pressure[a]	P = $P_c \exp[(p_1 x + p_2 x^{1.5} + p_3 x^3 + p_4 x^{3.5} + p_5 x^4 + p_6 x^{7.5})\frac{T_c}{T}]$ bar, $p_1 = -7.85951783$, $p_2 = 1.84408259$, $p_3 = -11.7866497$, $p_4 = 22.6807411$, $p_5 = -15.9618719$, $p_6 = 1.80122502$.
Liquid density[b]	$\rho_l = 0.08\tanh[(T - 225.0)/46.2] + 0.7415x^{0.33} + 0.32$ g/cm^3, (T is in K).
Surface tension[c]	$\gamma_{lv} = 235.8(1.0 - 0.625x)x^{1.256}$ dyn/cm
Relative permittivity[d]	$\epsilon = 5321.0/T + 233.76 - 0.9297T + 0.001417T^2 - 0.0000008292T^3$, ($T$ is in K).

[a]Equation (1) of Reference 78.
[b]Table I of Reference 62.
[c]Equation (1) of Reference 79.
[d]Equation (1) of Reference 80.

Cluster Definition

FIG. 7. MFPTs as a function of the cluster size n for $N = 216$ TIP4P/2005 water molecules in a cubic box of 200 Å side at 330 K (thermostat time coupling constant $\zeta = 5$ ps). In legend, the ten Wolde and Frenkel (tWF) minimum number of neighbors (n_{tWF}) with different Stillinger cutoff radius (r_S, in Å). Note that the tWF definition coincides with the Stillinger (Stg) criterion for the case of $n_{tWF} = 1$.

plateau region are observed. The $n_{tWF} = 4$ curves are less affected by surface growth.

[1] P. Rohwetter, J. Kasparian, K. Stelmaszczyk, Z. Hao, S. Henin, N. Lascoux, W. M. Nakaema, Y. Petit, M. Queißer, R. Salamé, E. Salmon, L. Wöste, and J.-P. Wolf, Nat. Photonics **4**, 451 (2010).

[2] Laser creates clouds over Germany, 2010, see http://www.newscientist.com/article/dn18848-laser-creates-clouds-over-germany.html.

[3] S. Henin, Y. Petit, P. Rohwetter, K. Stelmaszczyk, Z. Q. Hao, W. M. Nakaema, A. Vogel, T. Pohl, F. Schneider, J. Kasparian, K. Weber, L. Wöste, and J.-P. Wolf, Nat. Commun. **2**, 456 (2011).

[4] K. Yasuoka and M. Matsumoto, J. Chem. Phys. **109**, 8451 (1998).

[5] P. R. ten Wolde and D. Frenkel, J. Chem. Phys. **109**, 9901 (1998).

[6] J. Wedekind, J. Wölk, D. Reguera, and R. Strey, J. Chem. Phys. **127**, 154515 (2007).

[7] G. Chkonia, J. Wölk, R. Strey, J. Wedekind, and D. Reguera, J. Chem. Phys. **130**, 064505 (2009).

[8] J. Wedekind, G. Chkonia, J. Wölk, R. Strey, and D. Reguera, J. Chem. Phys. **131**, 114506 (2009).

[9] J. Wedekind, D. Reguera, and R. Strey, J. Chem. Phys. **125**, 214505 (2006).

[10] J. Wedekind, D. Reguera, and R. Strey, J. Chem. Phys. **127**, 064501 (2007).

[11] J. Wedekind and D. Reguera, J. Chem. Phys. **127**, 154516 (2007).

[12] K. Yasuoka and M. Matsumoto, Fluid Phase Equilib. **144**, 369 (1998).

[13] K. Yasuoka and M. Matsumoto, J. Chem. Phys. **109**, 8463 (1998).

[14] H. Matsubara, T. Koishi, T. Ebisuzaki, and K. Yasuoka, J. Chem. Phys. **127**, 214507 (2007).

[15] S. J. Keasler, H. Kim, and B. Chen, J. Phys. Chem. A **114**, 4595 (2010).

[16] M. Gamero-Castaño and J. F. de la Mora, J. Chem. Phys. **117**, 3345 (2002).

[17] H. Kitamura and A. Onuki, J. Chem. Phys. **123**, 124513 (2005).

[18] S. P. Fisenko, D. B. Kane, and M. S. El-Shall, J. Chem. Phys. **123**, 104704 (2005).

[19] E. Brodskaya, A. P. Lyubartsev, and A. Laaksonen, J. Chem. Phys. **116**, 7879 (2002).

[20] M. J. McGrath, J. N. Ghogomu, N. T. Tsona, J. I. Siepmann, B. Chen, I. Napari, and H. Vehkamäki, J. Chem. Phys. **133**, 084106 (2010).

[21] M. Volmer and A. Weber, Z. Phys. Chem. **119**, 227 (1926).

[22] L. Farkas, Z. Phys. Chem. **125**, 236 (1927).

[23] R. Becker and W. Döring, W. Ann. Phys. **24**, 719 (1935).

[24] J. W. Gibbs, Collected Works. Vol. 1. Thermodynamics (Longmans, Green & Co., London, 1878).

[25] D. Kashchiev, J. Chem. Phys. **76**, 5098 (1982).

[26] I. J. Ford, Phys. Rev. E **56**, 5615 (1997).

[27] D. W. Oxtoby and D. Kashchiev, J. Chem. Phys. **100**, 7665 (1994).

[28] S. L. Girshick and C. Chiu, J. Chem. Phys. **93**, 1273 (1990).

[29] J. J. Thomson, Conduction of Electricity Through Gases (Dover, New York, 1969).

[30] M. Born, Z. Phys. A: Hadrons Nucl. **1**, 45 (1920).

[31] A. G. Nasibulin, J. Fernandez de la Mora, and E. I. Kauppinen, J. Phys. Chem. A **112**, 1133 (2008).

[32] A. B. Nadykto and F. Yu, Phys. Rev. Lett. **93**, 016101 (2004).

[33] J. Wedekind, R. Strey, and D. Reguera, J. Chem. Phys. **126**, 134103 (2007).

[34] L. S. Bartell and D. T. Wu, J. Chem. Phys. **125**, 194503 (2006).

[35] S. E. M. Lundrigan and I. Saika-Voivod, J. Chem. Phys. **131**, 104503 (2009).

[36] F. Römer and T. Kraska, J. Chem. Phys. **127**, 234509 (2007).

[37] N. Lummen and B. Kvamme, Phys. Chem. Chem. Phys. **10**, 6405 (2008).

[38] L. Zheng, Q. An, Y. Xie, Z. Sun, and S.-N. Luo, J. Chem. Phys. **127**, 164503 (2007).

[39] J. Wedekind and D. Reguera, J. Phys. Chem. B **112**, 1160 (2008).

[40] J. L. F. Abascal and C. Vega, J. Chem. Phys. **123**, 234505 (2005).

[41] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

[42] J. L. F. Aragones, M. M. Conde, E. G. Noya, and C. Vega, Phys. Chem. Chem. Phys. **11**, 543 (2009).

[43] J. L. Aragones and C. Vega, J. Chem. Phys. **130**, 244504 (2009).

[44] J. L. F. Abascal and C. Vega, J. Chem. Phys. **133**, 234502 (2010).

[45] C. Vega, J. L. F. Abascal, and I. Nezbeda, J. Chem. Phys. **125**, 034503 (2006).

[46] H. L. Pi, J. L. Aragones, C. Vega, E. G. Noya, J. L. Abascal, M. A. Gonzalez, and C. McBride, Mol. Phys. **107**, 365 (2009).

[47] M. A. González and J. L. F. Abascal, J. Chem. Phys. **132**, 096101 (2010).

[48] J. Alejandre and G. A. Chapela, J. Chem. Phys. **132**, 014701 (2010).

[49] C. Vega and E. de Miguel, J. Chem. Phys. **126**, 154707 (2007).

[50] C. Vega, J. L. F. Abascal, M. M. Conde, and J. L. Aragones, Faraday Discuss. **141**, 251 (2009).

[51] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).

[52] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, Intermolecular Forces,; edited by B. Pullman (Reidel, Dordrecht, 1981).

[53] M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).

[54] P. T. Kiss and A. Baranyai, J. Chem. Phys. **134**, 054106 (2011).

[55] C. McBride, C. Vega, E. G. Noya, R. Ramírez, and L. M. Sesé, J. Chem. Phys. **131**, 024506 (2009).

[56] C. Vega, M. M. Conde, C. McBride, J. L. F. Abascal, E. G. Noya, R. Ramirez, and L. M. Sesé, J. Chem. Phys. **132**, 046101 (2010).

[57] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon, J. Chem. Phys. **120**, 9665 (2004).

[58] H. W. Horn, W. C. Swope, and J. W. Pitera, J. Chem. Phys. **123**, 194504 (2005).

[59] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, J. Chem. Theory Comput. **4**, 435 (2008).

[60] GROMACS files for the TIP4P/2005 model, see http://www.sklogwiki.org.

[61] S. Miyamoto and P. A. Kollman, J. Comput. Chem. **13**, 952 (1992).

[62] J. Wölk and R. Strey, J. Phys. Chem. B **105**, 11683 (2001).

[63] G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).

[64] F. H. Stillinger, J. Chem. Phys. **38**, 1486 (1963).

[65] D. Reguera, R. K. Bowles, Y. Djikaev, and H. Reiss, J. Chem. Phys. **118**, 340 (2003).

[66] J. Wedekind, A.-P. Hyvärinen, D. Brus, and D. Reguera, Phys. Rev. Lett. **101**, 125703 (2008).

[67] G. J. Martyna, M. L. Klein, and M. E. Tuckerman, J. Chem. Phys. **97**, 2635 (1992).

[68] C. T. R. Wilson, Proc. R. Soc. London, Ser. A **85**, 285 (1911).

[69] F. Arnold and G. Henschen, Nature (London) **257**, 521 (1978).

[70] F. Arnold, A. A. Viggiano, and H. Schlager, Nature (London) **297**, 371 (1982).

[71] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225 (1996).

[72] I. Kusaka, Z. Wang, and J. H. Seinfeld, J. Chem. Phys. **102**, 913 (1995).

[73] I. Kusaka, Z. Wang, and J. H. Seinfeld, J. Chem. Phys. **103**, 8993 (1995).

[74] K. J. Oh, G. T. Gao, and X. C. Zeng, Phys. Rev. Lett. **86**, 5080 (2001).

[75] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).

[76] P. Y. Ren and J. W. Ponder, J. Comput. Chem. **23**, 1497 (2002).

[77] J. W. Ponder, TINKER: Software Tools for Molecular Design (version 5.1). Washington University School of Medicine, Saint Louis, MO, 2010; see http://dasher.wustl.edu/tinker/.

[78] W. Wagner and A. Pruss, J. Phys. Chem. Ref. Data **22**, 783 (1993).

[79] N. B. Vargaftik, B. N. Volkov, and L. D. Voljak, J. Phys. Chem. Ref. Data **12**, 817 (1983).

[80] A. Catenaccio, Y. Daruich, and C. Magallanes, Chem. Phys. Lett. **367**, 669 (2003).