# Study of homogeneous bubble nucleation in liquid carbon dioxide by a hybrid approach combining molecular dynamics simulation and density gradient theory

K. Langenbach[1,*], M. Heilig[2], M. Horsch[1], and H. Hasse[1]

[1]*Laboratory of Engineering Thermodynamics, University of Kaiserslautern, Kaiserslautern, D-67663, Germany*

[2]*ROM, Digitalization in Research and Development, BASF SE, Ludwigshafen, D-67056, Germany*

A new method for predicting homogeneous bubble nucleation rates of pure compounds from vapor-liquid equilibrium (VLE) data is presented. It combines molecular dynamics simulation on the one side with density gradient theory using an equation of state (EOS) on the other. The new method is applied here to predict bubble nucleation rates in metastable liquid carbon dioxide ($CO_2$). The molecular model of $CO_2$ is taken from previous work of our group. PC-SAFT is used as EOS. The consistency between the molecular model and the EOS is achieved by adjusting the PC-SAFT parameters to VLE data obtained from the molecular model. The influence parameter of density gradient theory is fitted to the surface tension of the molecular model. Massively parallel molecular dynamics simulations are performed close to the spinodal to compute bubble nucleation rates. From these simulations the kinetic prefactor of the hybrid nucleation theory is estimated, whereas the nucleation barrier is calculated from density gradient theory. This enables the extrapolation of molecular simulation data to the whole metastable range including technically relevant densities. The results are tested against available experimental data and found to be in good agreement. The new method does not suffer from typical deficiencies of classical nucleation theory concerning the thermodynamic barrier at the spinodal and the bubble size dependence of surface tension, which is typically neglected in classical nucleation theory. In addition, the density in the center of critical bubbles and their surface tension is determined as a function of their radius. The usual linear Tolman correction to the capillarity approximation is found to be invalid.

## I. INTRODUCTION

Bubble nucleation is important in many applications. It occurs for example during ultrasonic cleaning or upon cavitation in technical equipment. In the present work, homogeneous bubble nucleation in metastable liquids is studied. That process is not spontaneous, but inhibited by a free energy barrier. In metastable liquids small bubbles stochastically form and break down. If a bubble exceeds the critical size, where the free energy of formation is maximal, it is likely to grow and become a macroscopic bubble. In this paper, a hybrid approach for describing homogeneous bubble nucleation is developed.

* Author to whom correspondence should be addressed. Electronic mail: kai.langenbach@mv.uni-kl.de

As homogeneous bubble nucleation requires activation energy, the rate of forming bubbles $J$ can formally be written in the form of an Arrhenius type expression

$$J = J_0 \exp\left( -\frac{\Delta\Omega^*}{kT} \right) \tag{1}$$

where $J_0$ is the so-called kinetic prefactor, $\Delta\Omega^*$ is the nucleation barrier height, i.e. the activation energy, $k$ is Boltzmann's constant and $T$ is the temperature. There are several theories of homogeneous bubble nucleation, which differ in the way $J_0$ and $\Delta\Omega^*$ are calculated.

The most common approach to bubble nucleation is classical nucleation theory (CNT). In CNT several strong simplifications are used[1]. First it is assumed that the capillary approximation holds, which implies that the width of the interface is zero and that the curvature of bubbles does not play a role, neither for $J_0$ nor for $\Delta\Omega^*$. This means that nanobubbles are assumed to have the same density and pressure as the saturated bulk vapor at the same temperature and that their surface tension is the same as that of the planar interface[2–4]. Furthermore, the free energy is considered to be the equilibrium free energy of the bulk. This, for example, results in a wrong limiting behavior of the nucleation barrier height close to the spinodal. Transport is usually not considered in CNT. In many cases CNT fails drastically[5–8], while in other cases, it gives fair predictions[4]. This behavior is not fully understood, but it is argued sometimes that in nucleation different phenomena combine and sometimes errors cancel out[9,10]. Many empirical and semi-empirical modifications of CNT have been proposed[11–14]. Of these the most commonly used is to describe the dependence of the surface tension on the nucleus size with the Tolman equation[15]. However, recent work has shown that a Tolman correction with a constant value of the Tolman length is unphysical and, therefore, introduces new artifacts[16–19].

Another method used for studies of nucleation is density functional theory (DFT) [20–22] or simplifications thereof like the density gradient theory (DGT)[23]. DFT and DGT have the advantage of yielding the correct limiting behavior upon approaching the binodal as well as the spinodal. Both theories also inherently consider the curvature dependence of surface tension and the finite width of the interface. Laaksonen et al.[20] as well as Oxtoby[21] give reviews of the application of DFT in nucleation theory. Using DFT, the free energy of the heterogeneous system must be known as a functional of density, which is often not the case. We therefore use DGT here, which can be used with bulk equations of state. As equation of state we choose PC-SAFT[24]. Both DFT and DGT can only be used to calculate the thermodynamic factor $\exp(-\Delta\Omega^* / (kT))$, but not the kinetic prefactor $J_0$ in Eq. (1). In the literature on the application of DGT or DFT for nucleation studies, therefore the

prefactor $J_0$ is often simply adopted from some version of CNT [e.g. 20–22]. This is dangerous, as CNT makes assumptions on the exponential term, which may be inconsistent with the DGT/DFT results.

A different approach is proposed here: DGT is used for determining the thermodynamic factor $\exp(-\Delta\Omega^* / (kT))$, while the kinetic prefactor $J_0$ is determined from molecular dynamics simulations instead of the usual CNT expressions, thereby hybridizing both theories. Presently studies of nucleation by molecular simulation can only be carried out for states close to the spinodal, where nucleation events occur frequently. Hence, the hybrid nucleation theory (HyNT) developed here can be seen as an extrapolation of nucleation rates determined by molecular dynamics close to the spinodal to state points in the entire metastable range by DGT.

As an example, the bubble nucleation in metastable liquid $CO_2$ is studied. The molecular model of $CO_2$ is taken from Merker et al.[25]. It is known to give good results for vapor pressure, liquid and vapor coexistence densities, second virial coefficients as well as transport properties[25]. The parameters of the PC-SAFT EOS are fitted to vapor-liquid equilibrium (VLE) data of Merker et al.[25] in order to ensure consistency between both models. Molecular simulations are carried out for homogenous metastable states and compared to results from PC-SAFT to verify that the PC-SAFT model is valid in the range of interest for nucleation. The influence parameter of DGT is obtained from a fit to surface tension data obtained with the molecular model of Merker et al.[25] taken from[26]. It should be noted that the experimental surface tension of $CO_2$ is overpredicted by roughly 20%[26] by the model of Merker et al.[25]. That deviation is accepted here, as no molecular model of $CO_2$ is presently available that would yield better results for the entire set of thermodynamic properties which are relevant for the present application. Finally, direct molecular dynamics simulations of bubble nucleation are carried out for two temperatures and several bulk phase densities close to the spinodal. The bubble formation statistics are evaluated with the approach introduced by Yasuoka and Matsumoto[27] for droplet nucleation and later on applied by Diemand et al.[28] for bubble nucleation. Using molecular simulation, no assumptions are made regarding the separation of kinetic prefactor and thermodynamic factor. However, in the range, where direct molecular simulation of nucleation is possible, the kinetic prefactor is dominant. This dominance is stronger the closer simulations are performed to the spinodal. Therefore, the prefactor of HyNT is calculated from the simulated nucleation rate closest to the spinodal. The other simulated nucleation rates are used to test the accuracy of the extrapolation using HyNT. The results of HyNT for homogeneous bubble nucleation are further compared to experimental data for the bubble nucleation of $CO_2$[29,30].

## II. THEORY

### A. Hybrid Nucleation Theory

In hybrid nucleation theory (HyNT), with the nucleation rate formally written as in Eq. (1), the free energy $\Delta\Omega^*$ of forming a critical bubble nucleus in a surrounding mother phase of density $\rho^h$ and temperature $T$ is calculated from DGT.

DGT[23,31,32] can be used in conjunction with a suitable EOS to investigate systems with a heterogeneous density distribution. In this work the PC-SAFT EOS[24] is used, which has been shown before to yield good results for both bulk fluid phase behavior and interfacial properties when combined with DGT [e.g. 33]. For a heterogeneous one-component system, the grand potential is written as a functional of the spatially varying density $\rho(\vec{x})$ and a function of the temperature $T$ according to[31]

$$\Delta\Omega(T)[\rho(\vec{x})] = \int_V \left( \Delta\omega_0(T, \rho(\vec{x})) + \frac{1}{2}\kappa\left(\vec{\nabla}\rho(\vec{x})\right)^2 \right) \mathrm{d}V \tag{2}$$

where $\Delta\omega_0 = f_0(T, \rho(\vec{x})) - \mu(T, \rho^h)\rho(\vec{x}) + p(T, \rho^h)$ is the grand potential per volume. Here, $f_0$ is the homogeneous free energy density calculated from the PC-SAFT EOS at the local density $\rho(\vec{x})$, $\mu$ is the chemical potential calculated for the density of the mother phase $\rho^h$, $p$ is the pressure, and $\kappa$ is the so-called influence parameter. Since we are interested here in the grand potential of critical bubbles, which corresponds to the free energy of forming them in the mother phase of the molar density $\rho^h$, radial symmetry is assumed. In spherical coordinates, the grand potential is given by

$$\Delta\Omega(T)[\rho(\mathrm{r})] = 4\pi \int_r \left( \Delta\omega_0(T, \rho(r)) + \frac{1}{2}\kappa\left(\frac{\mathrm{d}\rho}{\mathrm{d}r}\right)^2 \right) r^2 \mathrm{d}r \tag{3}$$

where $r$ is the radial coordinate. Functional minimization leads to the Euler-Lagrange equation

$$\frac{\mathrm{d}^2\rho}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}\rho}{\mathrm{d}r} = \kappa^{-1}\left(\frac{\partial\Delta\omega_0}{\partial\rho}\right)_T \tag{4}$$

which cannot be integrated directly as is possible in the case of planar interfaces in one component systems. The boundary conditions arise from symmetry ($\mathrm{d}\rho/\mathrm{d}r|_{r=0} = 0$) and from the mother phase density, which must be reached at infinite distance ($\rho(r \to \infty) = \rho^h$). The differential Eq. (4) is numerically solved using a finite difference scheme and the shooting method, for details see Appendix A. The solution of Eq. (4) corresponds to the density profile of the critical bubble in a mother phase of temperature $T$ and density $\rho^h$. With this density profile the grand potential or bubble nucleation barrier height is calculated from[34]

$$\Delta\Omega^* = -8\pi \int_0^\infty \Delta\omega_0\left(T, \rho(r)\right) r^2 \mathrm{d}r \tag{5}$$

where the integration is carried out numerically. Other properties of the critical bubble can be calculated once the radial density profile is known, for an overview see Hrubý et al.[34]. Here, we use the interfacial tension of the bubble $\gamma$, the Laplace pressure $\Delta p$, the equimolar radius $r_e$, and the radius of tension $r_s$ [34]:

$$\gamma = \left(\frac{3}{16\pi}\Delta\Omega^* \Delta p^2\right)^{\frac{1}{3}} \tag{6}$$

$$\Delta p = p\left(T, \rho(r \to \infty)\right) - p\left(T, \rho(r=0)\right) \tag{7}$$

$$r_e = \left(\frac{3\int_0^\infty \left(\rho(r) - \rho^h\right) \mathrm{r}^2 \mathrm{d}r}{\rho(r=0) - \rho^h}\right)^{\frac{1}{3}} \tag{8}$$

$$r_s = \left(\frac{3}{2\pi}\frac{\Delta\Omega^*}{\Delta p}\right)^{\frac{1}{3}} \tag{9}$$

The influence parameter $\kappa$ is obtained from a fit to data for the planar surface tension[35] using the expression[31]

$$\gamma_\infty = \sqrt{2\kappa}\int_{\rho''}^{\rho'} \sqrt{\Delta\omega_0\left(T, \rho\right)}\, \mathrm{d}\rho \tag{10}$$

where $\rho'$ and $\rho''$ are the molar liquid and gas density in equilibrium respectively.

The kinetic prefactor in Eq. (1) is defined here as

$$J_0 = k_0 \hat{\rho}^* \tag{11}$$

where $k_0$ is the kinetic part determined from molecular simulations close to the spinodal as elaborated in paragraph IV and $\hat{\rho}^*$ is the unit cavity density defined here as

$$\hat{\rho}^* = N_{av}\left(\rho^h - \rho'(T)\right) \tag{12}$$

where $N_{av}$ is Avogadro's number.

### B. Classical nucleation theory

Classical nucleation theory (CNT) was introduced by Volmer and Weber[36] and further developed by Farkas[37], Becker and Döring[38], and Zeldovich[39]. CNT has been modified empirically and semi-empirically in many ways, cf. Tanaka et al.[14] for a recent discussion. Since CNT is used here only for reference purposes, we simply use the most widespread variant for bubble nucleation, which is the one of Blander and Katz[4]. They give the bubble nucleation rate for a pure fluid as

$$J = N_{av}\rho^{h}\left(\frac{2\gamma_{\infty}}{\pi m}\right)^{\frac{1}{2}}\exp\left(\frac{-16\pi(\gamma_{\infty})^{3}}{3kT\left(p\left(T,\rho''\right)-p\left(T,\rho^{h}\right)\right)^{2}}\right) \tag{13}$$

where $m$ is the mass of the $CO_2$ molecule. Since the surface tension of the planar interface $\gamma_{\infty}$ is used, this is also called the capillarity approximation. We use PC-SAFT to connect the pressure, temperature and density.

## III. MODELING AND SIMULATION

### A. Molecular simulation

The simulations were carried out using the *ls1 mardyn*[40] program, which was modified by the authors to evaluate bubble nucleation statistics. The $CO_2$ model is the rigid three Lennard-Jones sites + point quadrupole model of Merker et al.[25]. It is given in Appendix B for completeness. In all simulations a cutoff at $r_c$ = 13.2 Å was used. The long range correction for the Lennard-Jones force field was treated with the method proposed by Lustig[41], while the quadrupole does not need such correction since it vanishes by angular integration. In the method of Lustig[41], a center of mass cutoff is employed adding constant mean field correction terms to the potential energy and the virial, where contributions from individual interaction sites are determined by angular integration. This is a simplification for bubbles, however, as long as the bubbles are small it should be a good approximation. The leapfrog integrator[42] was used with a time step of 2.5 fs. Simulations were conducted in the *NVT* ensemble. The occurrence of gaseous regions was sampled on a regular grid. The number of grid points $M$ was of the order of $N/4$ in all cases, see Appendix C for details. A point was considered to be gaseous, if $N_{g,max}$ = 5 or less molecular centers of mass were in a sphere of the radius $r_b$ = 6.88 Å around the point. This local density criterion corresponds to a threshold density of roughly 6.1 mol $l^{-1}$. By variation of both the number of molecular centers $N_{g,max}$ and the radius $r_b$, it was established that the results are not particularly sensitive to the particular choices, as long as the corresponding density is gaseous, i.e. below the critical density. This is the case for the present choice of parameters. A cluster $c$ is an aggregation of unit cavities. The cluster size is determined by connecting gaseous grid points that are immediate neighbors and counting the number $M_c$ of connected grid points per cluster. The volume of these clusters corresponding to bubble volume is calculated via

$$V_{bub} = \frac{M_c}{M} V \tag{14}$$

where $V$ is the simulation box volume. The number of such bubbles exceeding a certain threshold volume $V_{min}$ was counted and used to determine the nucleation rate as explained below.

A crystal like structure, with superimposed random displacements of the individual molecules, was used as the initial configuration. No pre-equilibration was conducted. In this way two series of simulations close to the liquid spinodal were carried out in which for $T$ = 220 K and $T$ = 280 K several densities were studied (cf. Table 4 and Table 5 in Appendix C).

In each simulation the bubble nucleation rates are determined using the Yasuoka-Matsumoto[27] (YM) method. In FIG. 1 a snapshot of such a simulation in the final stage of bubble nucleation is shown in an iso-density surface representation. The surfaces are shown for bubbles exceeding 1 nm in size in one direction. The bubbles are not spherical, but rather fluctuate around a spherical shape. In the YM method[27] the number of occurrences of bubbles that exceed a threshold volume $V_{min}$ is counted over time. A typical result is shown in FIG. 2. For a given threshold volume, a rate of bubble formation is obtained from the slope of this curve in the region of its inflection point, where the population of bubbles of the studied class increases at an approximately stationary rate. Numerical values for the rate of bubble formation obtained with this method for different threshold values obtained in the present simulations are given in Tables 4 and Table 5 in Appendix C. Additional information on the simulations is supplied in Table 6 of Appendix C.

[FIG. 1 about here]

[FIG. 2 about here]

From the rates of bubble formation obtained by the YM method, the macroscopic nucleation rate can be estimated if two different finite-size effects (FSE) are taken into account: First, if the threshold size $V_{min}$ is of the order of the critical size of the nucleus $V^*$ or smaller, the rate of formation overestimates the macroscopic nucleation rate, since the rate of formation contains a contribution due to relatively small bubbles that do not grow to macroscopic size. This finite-size effect (FSE 1) is caused by the finite threshold volume. It becomes significant for small values of $V_{min}$ and can be neglected if $V_{min}$ is much larger than $V^*$. Second, if the total volume of the considered bubbles approaches the volume occupied by the vapor phase in equilibrium, the rate of formation from the YM method obviously underestimates the macroscopic nucleation rate. This is basically caused by coalescence of the growing bubbles. This finite-size effect (FSE2) is caused by the finite simulation

volume. It can be neglected for large simulation volumes with comparably small values of $V_{min}$ and small nucleation rates, but becomes significant as $V_{min}$ and $J$ increase, or if systems with a small volume are simulated.

The dependence of the rate of formation $J$ on the threshold volume $V_{min}$ is shown in FIG. 3 ($T$ = 220 K) and FIG. 4 ($T$ = 280 K). Results for two homogeneous densities $\rho^h$ are plotted in each diagram. Depending on the temperature, the density, and the system size, FSE 1 or FSE 2 are found to be present or not. The four different combinations lead to four different shapes. At $T$ = 220 K (FIG. 3), for conditions close to the spinodal line, the critical volume is of the order of 1 nm$^3$ or smaller, so that the influence of the finite threshold volume, which was at least 1 nm$^3$ in all cases, can be neglected and FSE 1 is absent. For $T$ = 220 K and $\rho^h$ = 23.2 mol l$^{-1}$, the nucleation rate is comparably large, so that at large values of $V_{min}$, a decay of the rate of formation is found due to the influence of the finite simulation volume. Hence, FSE 2 is present. In this case, the macroscopic nucleation rate corresponds to the rate of formation for small bubbles. It is labeled as case A (FSE 1: no; FSE 2: yes). For $T$ = 220 K and $\rho^h$ = 23.9 mol l$^{-1}$ also FSE 2 is absent, since $J$ is small, and no finite-size effects are observed. A flat diagram is obtained, and any of the values can be used directly to estimate the macroscopic nucleation rate. This is labeled as case B here (FSE 1: no; FSE 2: no). At $T$ = 280 K (FIG. 4), for conditions close to the spinodal line, the critical volume is of the order of 100 nm$^3$. Hence, with decreasing $V_{min}$, the rate of bubble formation increases due to the increasing influence of the finite threshold volume. For $T$ = 280 K and $\rho^h$ = 17.35 mol l$^{-1}$, the nucleation rate is comparably small, so that FSE 1 is the only relevant finite-size effect and the macroscopic nucleation rate can be estimated from the rate of formation for large bubbles. This case is labeled case C (FSE 1: yes; FSE 2: no). For $T$ = 280 K and $\rho^h$ = 16.9 mol l$^{-1}$, both finite-size effects are relevant, and an estimate for the macroscopic nucleation rate is obtained from the rate of formation in the region of the inflection point of the plot, where both effects cancel out. This is labeled case D (FSE 1: yes; FSE 2: yes). In the above statement yes or no regarding the FSE must not be taken mathematically. Both effects are in principle always present, but may not be important. From this, we infer that the plot $J(V_{min})$ (at constant $T$ and $\rho^h$) always has an inflection point. It may, however, be outside the studied range. From the curvature of $J(V_{min})$ it is always possible to infer if the inflection point is in, below, or above the investigated range. If it is in the investigated range the simulation corresponds to case D. If it is above the investigated range it corresponds to case A. If it is below the investigated range, case B or case C can be present. In case C, almost constant nucleation rates are obtained. All present MD simulations correspond to one of these four cases and were evaluated accordingly.

[FIG. 3 about here]

[FIG. 4 about here]

For obtaining a nucleation rate from a given simulation not only a single threshold value $V_{min}$ is used. Rather, results obtained using several values of $V_{min}$ from the region in which the influence of both FSE 1 & FSE 2 is small are used. The highest and lowest of these values for one simulation are considered the confidence interval and their geometric mean is the $J$ reported in Table 1. All values for $J$ are given in Table 4 where those inside the confidence interval are underlined.

Table 1. Average nucleation rates as a function of temperature and density from molecular simulation and thermodynamic barrier from DGT for the states where $J_0$ is fitted.

| $\rho$ / mol l$^{-1}$ | $\overline{J}$ / m$^{-3}$s$^{-1}$ | $\exp\left(-\Delta\Omega^* / (kT)\right)$ |
|---|---|---|
| | $T$ = 220 K | |
| 22.6 | $6.92 \cdot 10^{34}$ | $2.85 \cdot 10^{-02}$ |
| 23.2 | $1.59 \cdot 10^{34}$ | - |
| 23.6 | $1.74 \cdot 10^{33}$ | - |
| 23.9 | $1.93 \cdot 10^{32}$ | - |
| | $T$ = 280 K | |
| 16.3 | $9.63 \cdot 10^{32}$ | - |
| 16.6 | $4.99 \cdot 10^{32}$ | $9{,}71 \cdot 10^{-01}$ |
| 16.9 | $2.55 \cdot 10^{32}$ | - |
| 17.2 | $5.71 \cdot 10^{31}$ | - |
| 17.35 | $2.18 \cdot 10^{31}$ | - |

In addition to bubble nucleation simulations, canonical ensemble equilibrium MD simulations were performed with comparably small systems. Particle numbers were between 800 and 1500. These simulations were conducted to capture the relation between density and pressure for the employed $CO_2$ model at $T$ = 220 K and $T$ = 280 K. Both stable and metastable states were studied. The same local density criterion as for the large scale simulations of bubble nucleation was applied to a regular grid with size $M$ = 10, where only simulations were taken into account with more 999.95 grid points on average belonging to a single metastable or stable phase. The simulation results are listed in Table 7.

### B. PC-SAFT + DGT

$CO_2$ is modeled here with the original PC-SAFT[24] without any polar terms[43–45]. Hence, the PC-SAFT $CO_2$ model has three parameters: the number of segments $m$, the segment diameter $\sigma$, and the interaction energy $\varepsilon$. They were fitted to the molecular simulation data of Merker[25] for the vapor pressure, saturated liquid density, and saturated vapor volume, who used the grand equilibrium method by Vrabec and Hasse[46] as simulation technique. A weighted least squares method based on the relative deviations for all data points reported by Merker et al.[25] was used for the fit. The influence parameter $\kappa$ of DGT was fitted to a single data point for the planar surface tension of the $CO_2$ model of Merker[25] as reported by Werth et al[35] for $T$ = 260 K. The resulting parameters are presented in Table 2, a comparison between the molecular simulation data for $CO_2$[25,35] and the results from SAFT + DGT is shown in FIG. 5. Also results for $\Delta h_V$ are presented, which were not included in the fit. Good agreement is observed for all properties. FIG. 5, panel b includes also predictions of the spinodal obtained with both methods. To retrieve data for the spinodal from the molecular simulation results, the method of Imre et al.[26] was used, which is based on heterogeneous simulation data for the density profile and the diagonal elements of the stress tensor. For a

detailed explanation of the method, the reader is referred to ref. 26. While the agreement between both methods for the spinodal on the liquid side is excellent, deviations occur on the vapor side, which are, however, not important here, as bubble nucleation is studied.

[FIG. 5 about here]

Table 2. Pure component parameters of the $CO_2$ model in PC-SAFT and DGT.

| $m$ | $\sigma$ / Å | $(\varepsilon / k)$ / K | $(\kappa 10^{20})$ / Jm$^5$mol$^{-2}$ |
|---|---|---|---|
| 2.6779 | 2.5253 | 149.4366 | 3.4331 |

FIG. 6 shows results for the thermal behavior of $CO_2$ obtained from the molecular simulations and from PC-SAFT. Besides data for the vapor-liquid equilibrium, which agree very well, data for isotherms at 220 and 280 K is shown. That data covers both the stable, metastable and unstable region. Molecular simulation data is available for stable and metastable states. It agrees very well with the PC SAFT predictions.

[FIG. 6 about here]

In addition to thermodynamic properties, structural ones can be compared. FIG. 7 shows density profiles for planar interfaces between vapor and liquid in equilibrium obtained from MD simulations[35] and PC-SAFT + DGT for several temperatures from 220 to 290 K. The results agree very well with each other.

[FIG. 7 about here]

## IV. RESULTS & DISCUSSION

### A. Evaluation of kinetic prefactor and comparison of nucleation rates from theory and simulation

The prefactor $k_0$ of Eq. (11) as used in Eq. (1) is fitted such that to the nucleation rate of Table 1 for the lowest density above the spinodal density (obtained from PC-SAFT) for the respective temperature and that from the HyNT nucleation rate agree. The kinetic prefactor is temperature dependent. We choose a simple Arrhenius type dependence

$$\ln\left(\frac{k_0}{\mathrm{s}^{-1}}\right) = a + \frac{b}{(T/K)} \tag{15}$$

with coefficients $a$ = -18.36 and $b$ = 8.611 to describe that dependence. This is somewhat arbitrary, but it has been argued in literature[14,47] that the kinetic prefactor depends on several quantities (e.g. viscosity, vapor pressure) that themselves have an exponential temperature dependence, hence our choice. Another sensible option would have been a power law dependence, however, with only two measured data points for $k_0$ both choices seemed likewise plausible. In FIG. 8 bubble nucleation rates for $T$ = 220 K and $T$ = 280 K from molecular simulation, CNT according to Eq. (13) and HyNT are shown. The molecular simulation results show that the nucleation rate close to the spinodal is larger for lower temperatures. Since

HyNT is fitted to that nucleation rate it shows the same temperature dependence. CNT predicts the inverse trend. With increasing density, nucleation rates drop super-exponentially. For $T$ = 280 K, HyNT shows a very good agreement with simulation. CNT overestimates the simulation data by roughly two orders of magnitude at that temperature, even though the same EOS and surface tension data is used. For $T$ = 220 K the agreement between HyNT and simulation is still reasonable, however, not as good as for the higher temperature. This may be due to the simplifications in DGT, namely the assumptions that the influence parameter $\kappa$ is a constant and the truncation of the series expansion (cf. Eq. (2) & (3)), which are both strictly valid only for vanishing gradients, i.e. at the critical point. Hence, increasing deviations for lower temperatures are not unexpected. CNT predicts nucleation rates about 5 orders of magnitude too low. The better agreement of HyNT is of course expected, since the kinetic prefactor is fitted to simulation data.

[FIG. 8 about here]

### B. Comparison to experimental data

For $CO_2$ only limited experimental data on metastable states and nucleation is available. Straub[29] has measured the limit of supercooling for isochoric cooling of liquid $CO_2$ below the coexistence temperature. This was also done by Huang et al.[30]. In both cases measurements were made close to the critical temperature of $CO_2$. In order to compare results from such experimental data to any nucleation theory both the sample volume and the relaxation time, i.e. the average time until the first macroscopic bubble is observed, need to be known. Neither Straub[29] nor Huang et al.[30] report both these values. However, McGraw[48] assessed that a nucleation barrier height of roughly 50 – 60 $kT$ is an appropriate threshold for the observation of macroscopic bubble nucleation near a critical point. This threshold corresponds to a nucleation rate of about $J = 1\ \mathrm{cm^{-3}s^{-1}}$ for the temperatures studied by Huang et al.[30]. For simplicity, this constant nucleation rate is adopted here for the evaluation of the supercooling from HyNT. The results of the comparison of the experimental data by Straub[29] and Huang et al.[30] to the predictions of HyNT and CNT are shown in FIG. 9. For the comparison the cavitation temperature $T_{\mathrm{cav}}$ is set and the required homogenous density $\rho^{\mathrm{h}}$ to achieve the prescribed $J$ is calculated. Then the vapor-liquid equilibrium temperature $T^{\mathrm{LV}}$ corresponding to that temperature is calculated. These temperatures are also known from the experiment. The comparison is carried out in a plot in reduced units, see FIG. 9. The agreement between the experimental data and that from HyNT is good. However, FIG. 9 also shows that the CNT according to Eq. (13) yields a similar agreement. The available supercooling experiments near the critical point are unfortunately not well suited for discriminating the theories.

[FIG. 9 about here]

### C. Discussion

The HyNT method presented above relies both on molecular simulation and an EOS + DGT. The employed molecular model and the EOS should be consistent and they should describe the available experimental data on VLE and surface tension of the studied fluid well. In CNT an EOS describing the VLE and surface tension data is needed so that both HyNT and CNT require the same input. As shown above, CNT predicts the wrong temperature dependence of the nucleation rate $J$, whereas in HyNT that problem does not occur. Nucleation rate is calculated from two different factors, the kinetic prefactor $J_0$ and the thermodynamic factor $\exp(-\Delta\Omega^* / (kT))$. The first one is shown in FIG. 10 and the nucleation barrier heights $\Delta\Omega^* / (kT)$ are shown in FIG. 11. The prefactors (cf. FIG. 10) from both theories differ by several orders of magnitude and their density dependence is different. The temperature dependence is qualitatively the same. The nucleation barrier height (cf. FIG. 11) is similar in the binodal limit for both temperatures. This is not surprising, since the capillarity approximation of CNT is expected to hold there. However, even at moderate supersaturations the behavior of HyNT and CNT differs drastically. CNT does not converge to the correct zero barrier height limit at the spinodal. This results in wrong thermodynamic factors, which are the reason for the wrong temperature dependence of CNT in the spinodal limit (cf. FIG. 8).

[FIG. 10 about here]

[FIG. 11 about here]

Since the capillarity approximation of the CNT of Blander and Katz[4] (Eq. (13)) breaks down, the Tolman correction[15] to the surface tension can be used as a modification. This correction is linear in the inverse of the equimolar radius of the critical bubble. Using DGT, the Tolman correction can be checked by calculating the surface tension of critical nuclei. For $T$ = 220 K and $T$ = 280 K, this radius dependent surface tension divided by the planar surface tension at the respective temperature is shown as a function of the inverse of equimolar radius in FIG. 12. The equimolar density is a function of the homogeneous density $\rho^h$ which is limited between the liquid spinodal density and the saturated liquid density. At both limits, the radius diverges to infinity. Hence, its inverse is zero. At the liquid spinodal density there is no stable nucleus. The surface tension approaches zero. At saturated liquid density, the planar limit is reached. It may seem odd that for a given critical radius two solutions for the surface tension exist. However, this is a direct consequence of the above limits. The two solutions correspond to two different bulk densities. For large radii, the linear Tolman correction captures the behavior of surface tension well, but for nanometer sized bubbles, the second order correction cannot be neglected for both temperatures. In case of the isotherm $T$ = 280 K, the surface tension even goes through a minimum rendering a truncated Taylor series expansion inappropriate. In summary, for high temperatures the linear Tolman correction breaks down much earlier than for low temperatures.

[FIG. 12 about here]

The Tolman length $\delta = r_e - r_s$ is depicted in FIG. 13 as a function of the inverse equimolar radius. Here the lower branch corresponds to solutions close to the binodal, while the upper branch, where the Tolman length seems to diverge, corresponds to solutions close to the spinodal. At the binodal the Tolman length is of the order of Ångström and positive.

[FIG. 13 about here]

One of the assumptions of CNT is that the density inside the bubble is the same as the saturated vapor density. This assumption has been shown to be wrong for the Lennard-Jones truncated and shifted (LJTS) fluid by Horsch and Hasse[49] before. These authors have shown that decreasing the bubble radius, first the density at the center of the critical bubble decreases. Upon further decrease of the bubble radius, the density increases again. The trend is confirmed here for $CO_2$. The results for the density at the center of the critical bubble from DGT are shown in FIG. 14 over the inverse of the equimolar radius for both studied temperatures. For comparison to the earlier results of Horsch and Hasse[49] the density at the center of the bubble $\rho(r = 0)$ is reduced by the saturated vapor density in the planar case $\rho''$. The radius is reduced by the Lennard-Jones diameter in case of the LJTS fluid and by the hard sphere diameter of one segment in case of the DGT results. At the binodal the radius of the critical bubble becomes infinite ($1/r_e = 0$), because only the planar interface is stable. In this case the saturated vapor density is reached. Starting from the planar case ($1/r_e = 0$), a decrease of density inside the bubble is predicted by molecular simulation as well as DGT + EOS for decreasing $r_e$. But for very small bubbles, the density increases sharply due to the finite size of the interface.

[FIG. 14 about here]

## V. CONCLUSIONS

A hybrid theory of bubble nucleation (HyNT) is developed. This theory is based on DGT + EOS and a molecular model, which should be consistent with each other and experimental data for VLE and interfacial properties. This can be achieved by parameter fitting. In this article the parameters of EOS and DGT are fitted to results from molecular simulation obtained with a model that is known to describe $CO_2$ well[25,35]. Bubble nucleation rates close to the spinodal are predicted using massively parallel molecular dynamics simulations with up to 100 million sites. These simulations are evaluated using the method of Yasuoka and Matsumoto[27]. The nucleation rate is split into two parts, the thermodynamic barrier $\exp(-\Delta\Omega^* / (kT))$ and the kinetic prefactor $J_0$. The former is calculated using DGT + PC-SAFT and the latter is fitted to simulated nucleation rates for two temperatures, as close to the spinodal as possible, because there the thermodynamic barrier vanishes. HyNT shows a good agreement with nucleation rates determined from simulation, also for simulations that were not included in the fit. CNT in the version of Blander and Katz[4], using the

same EOS, shows the wrong temperature dependence of the nucleation rate. This is explained in terms of the non-zero limiting nucleation barrier height of CNT at the spinodal. Hence, CNT should not be used to interpret data from molecular simulation close to the spinodal.

Further improvements of HyNT might be possible using a density functional instead of the DGT used here, however, at the expense of increased complexity. Results for the isochoric limit of supercooling for liquid $CO_2$ obtained from HyNT are compared to experimental data[29,30] and show good agreement. However, also CNT agrees well with that data, which is only available in a small temperature range close to the critical point and therefore not suitable for discriminating the results of both theories. The new theory should be tested for temperatures closer to the triple point of $CO_2$ preferably using direct bubble nucleation rate measurements, if such data becomes available. Additionally, the linear Tolman correction to the surface tension is found to be qualitatively wrong for very small bubbles, i.e. in the regime, where molecular simulation takes place. The qualitative dependency of the density in the center of critical bubbles on the bubble size that was observed earlier for the Lennard-Jones truncated and shifted fluid is confirmed here for $CO_2$.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from BMBF within the SkaSim project (grant no. 01H13005A) and from DFG within IRTG 2057 (subproject UKL3.5). The large-scale MD simulations were carried out on hermit at HLRS (Stuttgart) with specially allocated resources, on hazel hen at HLRS within the supercomputing project MMHBF2, on SuperMUC at LRZ (Garching) within the supercomputing project SPARLAMPE (pr48te), and on Elwetritsch at RHRK within the supercomputing project TUKL-TLMV. The authors would like to thank C.W. Glass for facilitating access to HPC infrastructure.

## REFERENCES

[1] J.F. Lutsko, in *New Perspectives on Mineral Nucleation and Growth*, edited by A.E.S.V. Driessche, M. Kellermeier, L.G. Benning, and D. Gebauer (Springer International Publishing, 2017), pp. 25–41.
[2] M. Blander and J.L. Katz, J. Stat. Phys. **4**, 55 (1972).
[3] H. Vehkamäki, *Classical Nucleation Theory in Multicomponent Systems* (Springer, Heidelberg, 2006).
[4] M. Blander and J.L. Katz, AIChE J. **21**, 833 (1975).
[5] V.G. Baidakov and K.S. Bobrov, J. Chem. Phys. **140**, 184506 (2014).
[6] M.A. Gonzalez, J.L.F. Abascal, C. Valeriani, and F. Bresme, J. Chem. Phys. **142**, 154903 (2015).
[7] F. Römer and T. Kraska, J. Chem. Phys. **127**, 234509 (2007).
[8] S. Sinha, A. Bhabhe, H. Laksmono, J. Wölk, R. Strey, and B. Wyslouzil, J. Chem. Phys. **132**, 064304 (2010).
[9] F. Römer, B. Fischer, and T. Kraska, Soft Mater. **10**, 130 (2012).
[10] J. Wedekind, A.-P. Hyvärinen, D. Brus, and D. Reguera, Phys. Rev. Lett. **101**, 125703 (2008).
[11] M. Horsch, J. Vrabec, and H. Hasse, Phys. Rev. E **78**, 011603 (2008).
[12] V.I. Kalikmanov, J. Chem. Phys. **124**, 124505 (2006).
[13] A. Laaksonen, I.J. Ford, and M. Kulmala, Phys. Rev. E **49**, 5517 (1994).
[14] K.K. Tanaka, H. Tanaka, R. Angélil, and J. Diemand, Phys. Rev. E **92**, 022401 (2015).
[15] R.C. Tolman, J. Chem. Phys. **17**, 333 (1949).
[16] B.J. Block, S.K. Das, M. Oettel, P. Virnau, and K. Binder, J. Chem. Phys. **133**, 154702 (2010).
[17] A. Malijevský and G. Jackson, J. Phys.: Condens. Matter **24**, 464121 (2012).

[18] S. Werth, S.V. Lishchuk, M. Horsch, and H. Hasse, Physica A **392**, 2359 (2013).
[19] Ø. Wilhelmsen, D. Bedeaux, and D. Reguera, J. Chem. Phys. **142**, 064706 (2015).
[20] A. Laaksonen, V. Talanquer, and D.W. Oxtoby, Annu. Rev. Phys. Chem. **46**, 489 (1995).
[21] D.W. Oxtoby, Ann. Rev. Mater. Res. **32**, 39 (2002).
[22] V.G. Baidakov, J. Chem. Phys. **144**, 074502 (2016).
[23] J.W. Cahn and J.E. Hilliard, J. Chem. Phys. **31**, 688 (1959).
[24] J. Gross and G. Sadowski, Ind. Eng. Chem. Res. **40**, 1244 (2001).
[25] T. Merker, C. Engin, J. Vrabec, and H. Hasse, J. Chem. Phys. **132**, 234512 (2010).
[26] A.R. Imre, G. Mayer, G. Házi, R. Rozas, and T. Kraska, J. Chem. Phys. **128**, 114708 (2008).
[27] K. Yasuoka and M. Matsumoto, J. Chem. Phys. **109**, 8451 (1998).
[28] J. Diemand, R. Angélil, K.K. Tanaka, and H. Tanaka, Phys. Rev. E **90**, 052407 (2014).
[29] J. Straub, in *Proceedings of the 3rd International Conference on Chemical Thermodynamics and Physico-Chemical Techniques at High Temperatures, Baden, Austria* (Butterworth, London, 1973).
[30] J.S. Huang, W.I. Goldburg, and M.R. Moldover, Phys. Rev. Lett. **34**, 639 (1975).
[31] J.W. Cahn and J.E. Hilliard, J. Chem. Phys. **28**, 258 (1958).
[32] J.W. Cahn, J. Chem. Phys. **30**, 1121 (1959).
[33] S. Enders, H. Kahl, and J. Winkelmann, Fluid Phase Equilibr. **228**–**229**, 511 (2005).
[34] J. Hrubý, D.G. Labetski, and M.E.H. van Dongen, J. Chem. Phys. **127**, 164720 (2007).
[35] S. Werth, M. Kohns, K. Langenbach, M. Heilig, M. Horsch, and H. Hasse, Fluid Phase Equilibria **427**, 219 (2016).
[36] M. Volmer and A. Weber, Z. Phys. Chem. **119**, 277 (1926).
[37] L. Farkas, Z. Phys. Chem. **125**, 236 (1927).
[38] R. Becker and W. Döring, Ann. Phys. **416**, 719 (1935).
[39] J.B. Zeldovich, Zh. Eksp. Teor. Fiz. **12**, 525 (1942).
[40] C. Niethammer, S. Becker, M. Bernreuther, M. Buchholz, W. Eckhardt, A. Heinecke, S. Werth, H.-J. Bungartz, C.W. Glass, H. Hasse, J. Vrabec, and M. Horsch, J. Chem. Theory Comput. **10**, 4455 (2014).
[41] R. Lustig, Mol. Phys. **65**, 175 (1988).
[42] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, New York, NY, USA, 1989).
[43] J. Gross, AIChE J. **51**, 2556 (2005).
[44] J. Gross and J. Vrabec, AIChE J. **52**, 1194 (2006).
[45] M. Kleiner and J. Gross, AIChE J. **52**, 1951 (2006).
[46] J. Vrabec and H. Hasse, Mol. Phys. **100**, 3375 (2002).
[47] B.V. Deryagin, A.V. Prokhorov, and N.N. Tunitskii, Soviet Journal of Experimental and Theoretical Physics **46**, 962 (1977).
[48] R. McGraw, J. Chem. Phys. **91**, 5655 (1989).
[49] M. Horsch and H. Hasse, Chem. Eng. Sci. **107**, 235 (2014).
[50] M.V. Bulatov, P.M. Lima, and D.T. Thanh, Vestnik YuUrGU. Ser. Mat. Model. Progr. **8**, 5 (2015).
[51] M. Krone, J. Stone, T. Ertl, and K. Schulten, in *EuroVis - Short Papers* (2012).
[52] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).

## APPENDIX

### APPENDIX A: Numerical Procedure for Determining Radial Density Profiles

Density profiles are calculated according to the differential Eq. (4). This equation is discretized using finite differences with a step width of $h$ = 0.2 Å using a central difference scheme. Rewriting Eq. (4) for point $i$ in the finite difference scheme results in

$$(r_i + h)\rho_{i+1} = (h - r_i)\rho_{i-1} + 2r_i\rho_i + r_i h^2 \kappa^{-1} \left( \frac{\partial \Delta\omega_0}{\partial \rho} \right)_T (\rho_i, T) \tag{C.1}$$

for $i > 0$ and in

$$\rho_1 = \rho_0 + \frac{h^2}{2} \kappa^{-1} \left( \frac{\partial \Delta\omega_0}{\partial \rho} \right)_T (\rho_0, T) \tag{C.2}$$

for the bubble center, ensuring the condition of vanishing derivative of $\rho$ at $r = 0$. Since the second condition is at infinite radius, also there the derivative of $\rho$ must vanish in addition to the condition $\rho(r \to \infty) = \rho^h$. Since boundary conditions on a semi-infinite domain are difficult to handle, we use the fact that if both conditions are fulfilled Eq. (4) is exactly zero. This means that at vanishing derivative at some large $r$, the homogenous density $\rho^h$ must be recovered for the non-trivial solution of this boundary value problem. Using the iteration formulas above, we guess a starting density $\rho_0$ and calculate the profile until the absolute change of density from $i$ to $i$+1 divided by the total density difference between center of the bubble and homogeneous phase is smaller than $10^{-6}$. If the density at that point $n$ is larger than the homogeneous phase, $\rho_0$ is reduced and if larger it is increased and the density profile recalculated. After that initial step, the bisection method is used to find the correct $\rho_0$. Convergence is assumed if

$$\frac{\rho_n - \rho^h}{\rho_0} < 0.001 \frac{\rho^h - \rho_0}{\rho^h} \tag{C.3}$$

Using this method, the calculation of profiles is limited by the machine precision. If the lower and upper bound in the bisection method differ by less than double the machine precision, no better solution is possible. This occurs approaching the binodal, where the radius of the bubble grows to infinity. There have been suggestions in literature[50] to work around this, however, we do not apply such methods here, since we are mainly interested in the range closer to the spinodal.

### APPENDIX B: Molecular Model for $CO_2$

The molecular model for $CO_2$ by Merker et al.[25] consists of three linearly arranged Lennard-Jones sites, one for each oxygen and the carbon and a point quadrupole on the carbon aligned with the molecular axis. The parameters of the model are the Lennard-Jones diameters of the oxygen $\sigma_O$ and carbon $\sigma_C$ as well as their interaction energies $\varepsilon_O$ and $\varepsilon_C$ The distance between carbon and oxygen $r_{CO}$ and the quadrupole moment $Q$, all of which are listed in Table 3.

Table 3. Molecular model of $CO_2$ from Merker et al.[25]

| $\sigma_C$ / Å | $\varepsilon_C$ / K | $\sigma_O$ / Å | $\varepsilon_O$ / K | $Q$ / DÅ | $r_{CO}$ / Å |
|---|---|---|---|---|---|
| 2.8137 | 12.3724 | 2.9755 | 100.493 | 4.0739 | 1.2869 |

### APPENDIX C: Simulation Data on Bubble Nucleation, Metastable and Stable States of $CO_2$

In Table 4 and Table 5 the nucleation rate from molecular simulation, evaluated as described in paragraph III.A, is given for the two investigated temperatures as indicated in the heading for different densities close to the spinodal. The numbers of particles used for the different simulations as well as the numbers of grid points for bubble evaluation in one dimension are given in Table 6. Simulation results for metastable and stable $p$-$\rho$-$T$ behavior of the $CO_2$ model are given in Table 7.

Table 4. Nucleation rates from the present MD simulations at $T$ = 220 K, evaluated according to the YM method with varying values of the threshold bubble volume $V_{min}$. The rates that were used for the data of $J$ (cf. Table 1 and text) are underlined. "n/a" indicates that the nucleation rate could not be obtained reliably by the YM method.

| $V_{min}$ / nm$^3$ | 1 | 2 | 5 | 10 | 20 | 50 | 100 | 200 | 500 |
|---|---|---|---|---|---|---|---|---|---|
| $\rho^h$ / mol l$^{-1}$ | $J$ / m$^{-3}$s$^{-1}$ | | | | | | | | |
| 22.6 | $\underline{8.4 \cdot 10^{34}}$ | $\underline{7.0 \cdot 10^{34}}$ | $\underline{5.7 \cdot 10^{34}}$ | $4.7 \cdot 10^{34}$ | $3.6 \cdot 10^{34}$ | $2.3 \cdot 10^{34}$ | $1.4 \cdot 10^{34}$ | $5.9 \cdot 10^{34}$ | n/a |
| 23.2 | $\underline{1.8 \cdot 10^{34}}$ | $\underline{1.8 \cdot 10^{34}}$ | $\underline{1.6 \cdot 10^{34}}$ | $\underline{1.4 \cdot 10^{34}}$ | $1.1 \cdot 10^{34}$ | $9.4 \cdot 10^{33}$ | $7.8 \cdot 10^{33}$ | $5.3 \cdot 10^{33}$ | $1.1 \cdot 10^{33}$ |
| 23.6 | $\underline{1.9 \cdot 10^{33}}$ | $\underline{1.9 \cdot 10^{33}}$ | $\underline{1.6 \cdot 10^{33}}$ | $\underline{1.7 \cdot 10^{33}}$ | $\underline{1.7 \cdot 10^{33}}$ | $\underline{1.6 \cdot 10^{33}}$ | $1.3 \cdot 10^{33}$ | $1.1 \cdot 10^{33}$ | $1.1 \cdot 10^{33}$ |
| 23.9 | $\underline{2.0 \cdot 10^{32}}$ | $\underline{1.7 \cdot 10^{32}}$ | $\underline{1.9 \cdot 10^{32}}$ | $\underline{2.2 \cdot 10^{32}}$ | $\underline{2.0 \cdot 10^{32}}$ | $\underline{2.0 \cdot 10^{32}}$ | $\underline{2.0 \cdot 10^{32}}$ | $\underline{1.8 \cdot 10^{32}}$ | $\underline{2.0 \cdot 10^{32}}$ |

Table 5. Same as Table 4 but for $T$ = 280 K.

| $V_{min}$ / nm$^3$ | 2 | 5 | 10 | 20 | 50 | 100 | 200 | 500 |
|---|---|---|---|---|---|---|---|---|
| $\rho^h$ / mol l$^{-1}$ | $J$ / m$^{-3}$s$^{-1}$ | | | | | | | |
| 16.3 | $8.1 \cdot 10^{34}$ | $1.7 \cdot 10^{34}$ | $\underline{5.8 \cdot 10^{33}}$ | $\underline{2.7 \cdot 10^{33}}$ | $\underline{1.1 \cdot 10^{33}}$ | $\underline{5.6 \cdot 10^{32}}$ | $\underline{3.5 \cdot 10^{32}}$ | $\underline{1.6 \cdot 10^{32}}$ |
| 16.6 | $6.0 \cdot 10^{34}$ | $1.1 \cdot 10^{34}$ | $\underline{2.8 \cdot 10^{33}}$ | $\underline{1.3 \cdot 10^{33}}$ | $\underline{5.6 \cdot 10^{32}}$ | $\underline{3.2 \cdot 10^{32}}$ | $\underline{1.8 \cdot 10^{32}}$ | $\underline{8.9 \cdot 10^{31}}$ |
| 16.9 | $3.8 \cdot 10^{34}$ | $4.0 \cdot 10^{33}$ | $\underline{1.3 \cdot 10^{33}}$ | $\underline{4.9 \cdot 10^{32}}$ | $\underline{2.9 \cdot 10^{32}}$ | $\underline{1.9 \cdot 10^{32}}$ | $\underline{1.2 \cdot 10^{32}}$ | $\underline{5.0 \cdot 10^{31}}$ |
| 17.2 | $2.3 \cdot 10^{34}$ | $2.0 \cdot 10^{33}$ | $3.5 \cdot 10^{32}$ | $1.4 \cdot 10^{32}$ | $\underline{7.1 \cdot 10^{31}}$ | $\underline{5.8 \cdot 10^{31}}$ | $\underline{4.6 \cdot 10^{31}}$ | n/a |
| 17.35 | $1.4 \cdot 10^{34}$ | $1.2 \cdot 10^{33}$ | $1.8 \cdot 10^{32}$ | $7.0 \cdot 10^{31}$ | $3.3 \cdot 10^{31}$ | $\underline{2.8 \cdot 10^{31}}$ | $\underline{1.9 \cdot 10^{31}}$ | $\underline{1.7 \cdot 10^{31}}$ |
| 17.575 | $9.7 \cdot 10^{33}$ | $5.9 \cdot 10^{32}$ | $5.0 \cdot 10^{31}$ | n/a | n/a | n/a | n/a | n/a |
| 17.65 | $7.8 \cdot 10^{33}$ | $3.6 \cdot 10^{32}$ | n/a | n/a | n/a | n/a | n/a | n/a |

Table 6. Density $\rho^h$, temperature $T$, number of molecules $N$, and grid resolution $M^{1/3}$ for the present MD simulations.

| $\rho^h$ / mol l$^{-1}$ | $N$ | $M^{1/3}$ |
|---|---|---|
| | $T$ = 220 K | |
| 22.6 | 13 000 000 | 160 |
| 23.2 | | |
| 23.6 | | |
| 23.9 | | |
| | $T$ = 280 K | |
| 16.3 | 20 000 000 | 170 |
| 16.6 | | |
| 16.9 | | |
| 17.2 | 25 000 000 | 180 |
| 17.35 | 7 000 000 | 120 |
| 17.575 | 20 000 000 | 170 |
| 17.65 | 7 000 000 | 120 |

Table 7. Thermal properties of homogeneous states of $CO_2$ as described by the model of Merker et al.[25]: Pressure $p$ as a function of temperature $T$ and density $\rho$. The number in parenthesis is the uncertainty of the last digit of the number reported for the pressure. Due to data loss, no uncertainty for the simulations at $T$ = 280 K can be reported here.

| $\rho$ / mol l$^{-1}$ | $p$ / MPa |
|---|---|
| $T$ = 220 K | |
| 0.1 | 0.17878(9) |
| 0.2 | 0.3496(2) |
| 0.3 | 0.5121(7 |
| 0.4 | 0.666(3) |
| 24 | -32(1) |
| 25.4 | -17.8(4) |
| 26.4 | 0.8(3) |
| 26.5 | 5(1) |
| 27.4 | 28.3(2) |
| 29 | 89.9(6) |
| 31.5 | 255.7(4) |
| $T$ = 280 K | |
| 1.75 | 3.18 |
| 2 | 3.5 |
| 2.25 | 3.77 |
| 2.5 | 4.03 |
| 2.7663 | 4.27 |
| 3 | 4.47 |
| 3.25 | 4.65 |
| 3.75 | 4.91 |
| 17 | -1.24 |
| 18.5 | -0.518 |
| 20.12 | 4.19 |
| 21.5 | 12.9 |
| 23 | 28.1 |
| 24.5 | 56.2 |
| 26 | 95.9 |

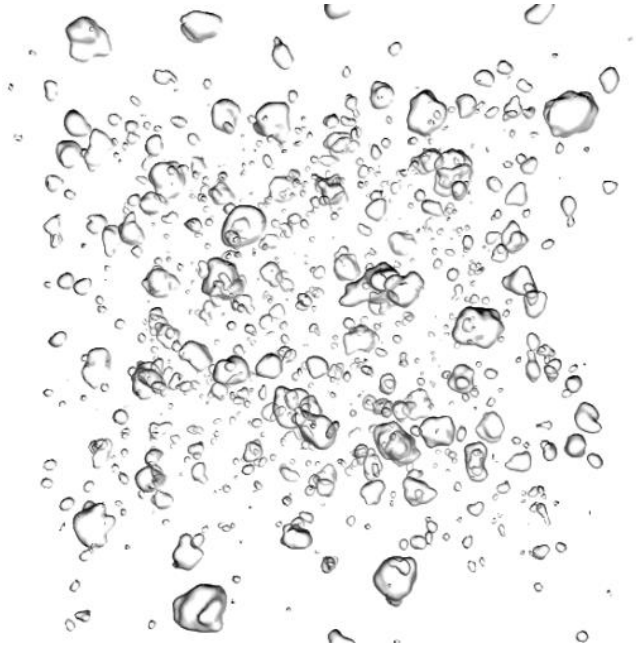

FIG. 1. Iso-density surface representation of a snapshot during the final stages of bubble nucleation at $T$ = 280 K and $\rho^{h}$ = 17.2 mol $l^{-1}$. The surfaces are generated by Gaussian density mapping using the quicksurf algorithm[51] in vmd[52] and mark aggregations of unit cavities as defined in paragraph III.A.

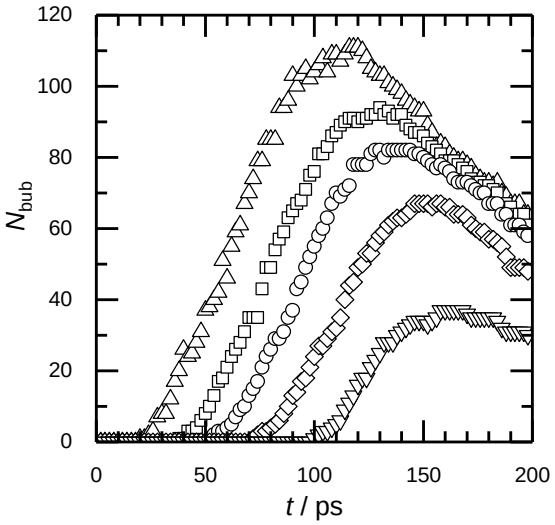


FIG. 2. Number of bubbles $N_{bub}$ exceeding the threshold volumes $V_{min}$ = 1 (triangles up), 5 (squares), 20 (circles), 100 (diamonds), and 500 nm³ (triangles down) in a bubble nucleation simulation at $T$ = 220 K and $\rho^h$ = 23.6 mol $l^{-1}$ over simulation time.

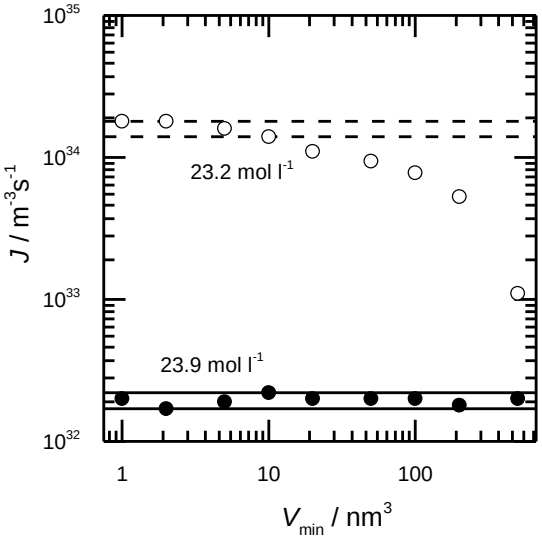


FIG. 3. Rate of bubble formation over threshold volume $V_{min}$ at $T$ = 220 K and $\rho^h$ = 23.2 (open circles) and 23.9 mol $l^{-1}$ (filled circles) obtained from the present MD simulations, evaluated with the YM method, and confidence intervals for the macroscopic nucleation rate at $\rho^h$ = 23.2 (dashed lines) and 23.9 mol $l^{-1}$ (solid lines).

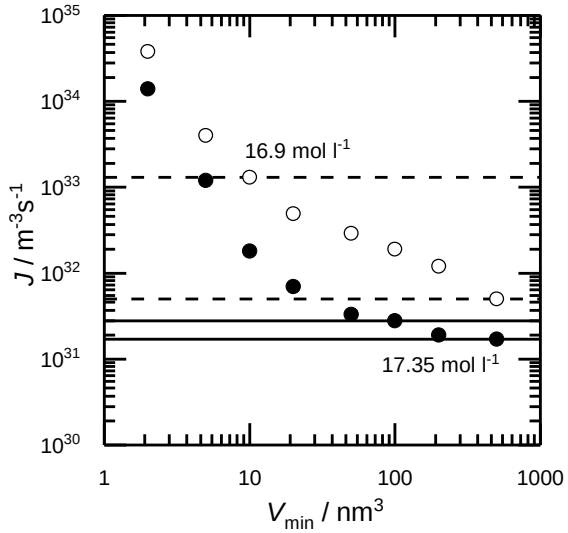


FIG. 4. Rate of bubble formation over threshold volume $V_{min}$ at $T$ = 280 K and $\rho^h$ = 16.9 (open circles) and 17.35 mol l$^{-1}$ (filled circles) obtained from the present MD simulations, evaluated with the YM method, and confidence intervals for the macroscopic nucleation rate at $\rho^h$ = 16.9 (dashed lines) and 17.35 mol l$^{-1}$ (solid lines).

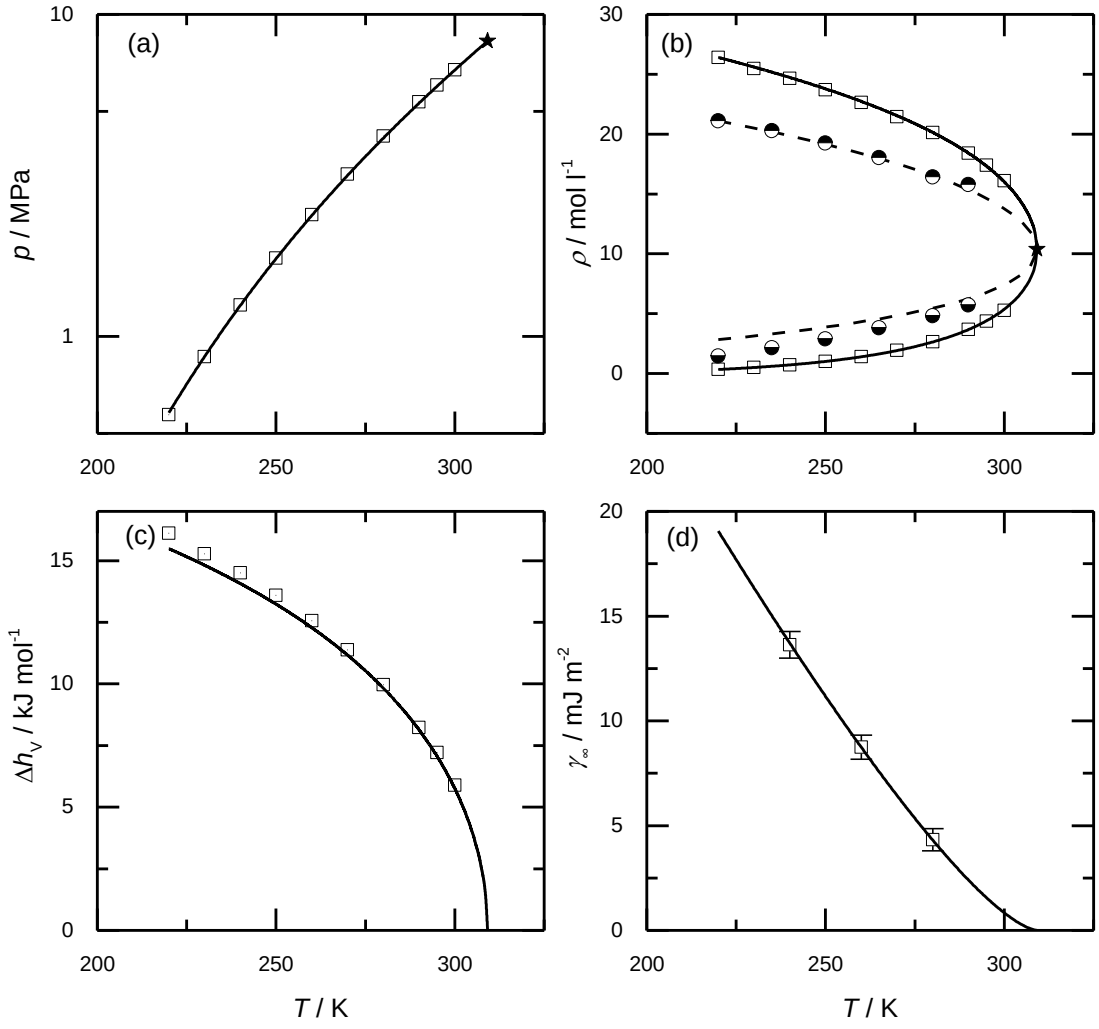


FIG. 5. Vapor pressure (a), coexisting and spinodal densities (b), enthalpy of vaporization (c) and planar surface tension (d) of $CO_2$ over temperature from PC-SAFT (lines) and molecular simulation (symbols) of Merker et al.[25] (a-c) and Werth et al.[35] (d).

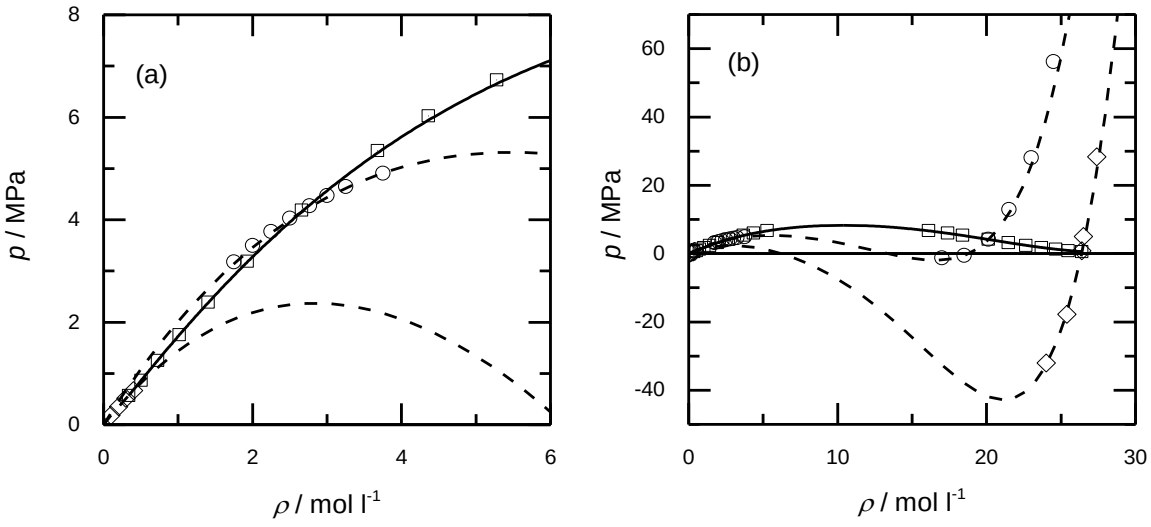


FIG. 6. Thermal properties of $CO_2$ as described by the molecular model of Merker et al.[25] and the PC-SAFT model from the present work. Both data for the vapor-liquid equilibrium and data for isotherms at 220 K and 280 K in the stable metastable and unstable range are shown. Full line: binodal PC-SAFT; dashed line: isotherm PC-SAFT; squares: VLE molecular simulation Merker et al.[25]; symbols are data for the isotherms obtained from molecular simulation (diamonds: 220 K; circles: 280 K). Panel a) shows the gas side while panel b) shows the entire studied range.

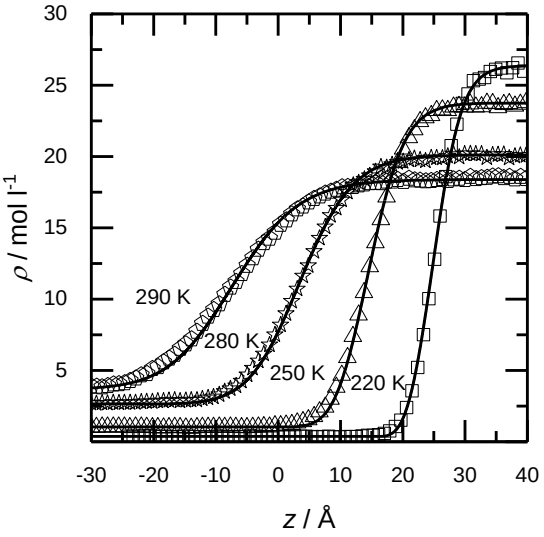


FIG. 7. Density profiles through the vapor-liquid interface for $CO_2$ as described by the molecular model of Merker et al. [25] and the PC-SAFT + DGT model from the present work for isotherms as indicated in the figure. Symbols are molecular simulation data[35] and lines PC-SAFT + DGT calculations.

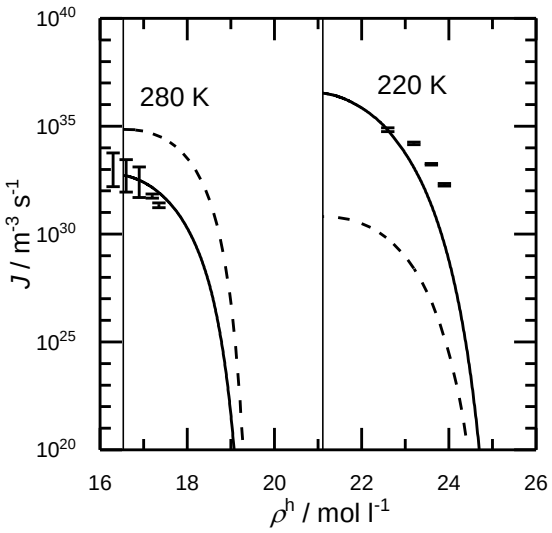


FIG. 8. Nucleation rate of $CO_2$ over liquid density at 220 K (right) and 280 K (left). Symbols are molecular simulation results, the dashed lines are CNT calculations, and the solid lines HyNT calculations The symbols show the confidence intervals for the macroscopic nucleation rate determined from the rate of formation of nanobubbles in the present MD simulations. Thin vertical lines are the respective PC-SAFT spinodals.

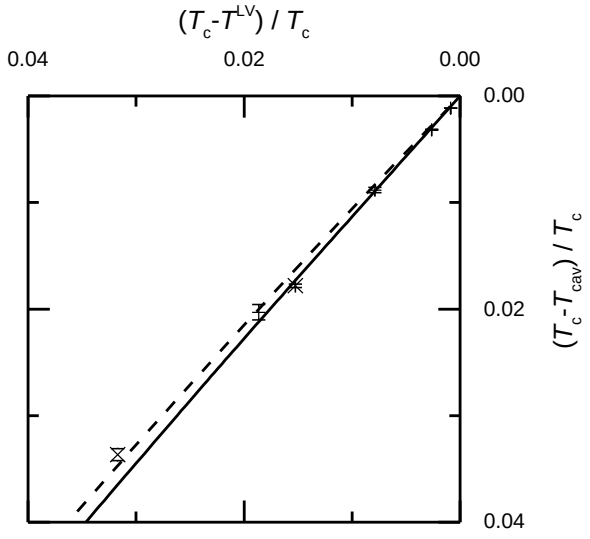


FIG. 9. Reduced isochoric cavitation temperature for nucleation rates of $J = 1$ cm$^{-3}$s$^{-1}$ over reduced equilibrium temperature from HyNT (solid line), CNT (broken line) and experiments of Straub[29] (x) and Huang et al.[30] (+) together with experimental confidence intervals given by the respective authors. Axis are reversed to have lower temperatures left and at the bottom.

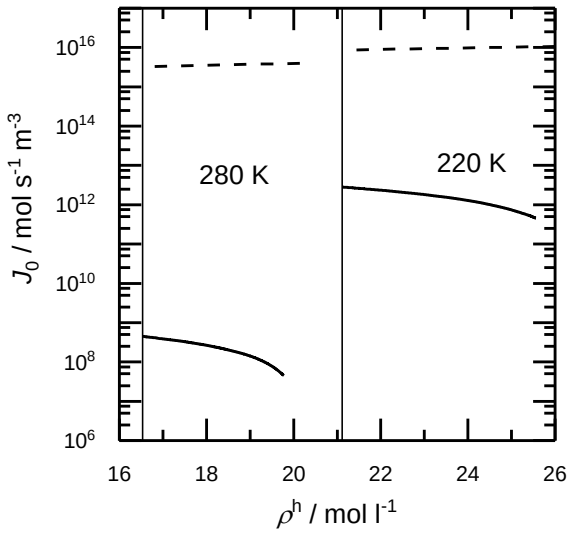


FIG. 10. Kinetic prefactor of the nucleation of $CO_2$ over liquid density at 220 K (right) and 280 K (left). Dashed lines are CNT calculations and solid lines HyNT calculations Thin vertical lines are the respective PC-SAFT spinodals.

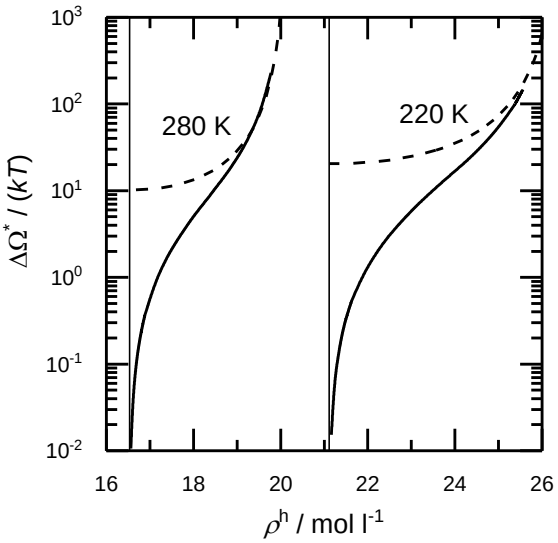


FIG. 11. Nucleation barrier height of $CO_2$ over liquid density at 220 K (right) and 280 K (left). Dashed lines are CNT calculations and solid lines HyNT calculations Thin vertical lines are the respective PC-SAFT spinodals.

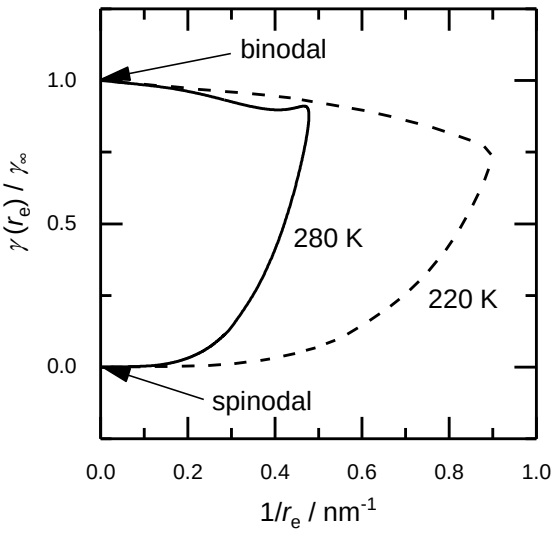


FIG. 12. Surface tension of critical bubbles divided by planar surface tension over the inverse equimolar radius of these bubbles for $T$ = 280 K (full line) and $T$ = 220 K (dashed line) obtained from DGT for $CO_2$.

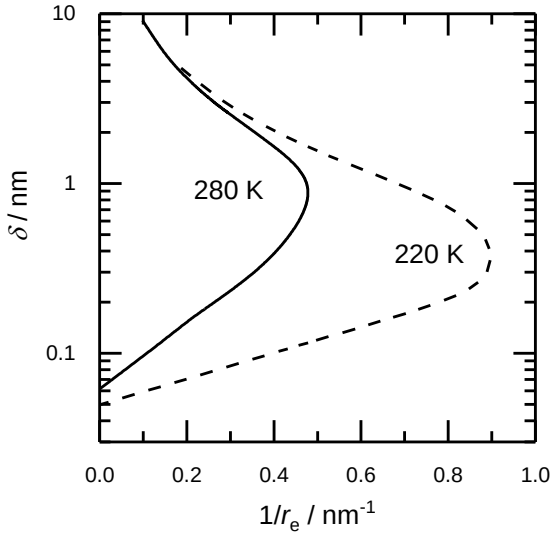


FIG. 13. Tolman length of critical bubbles over the inverse equimolar radius of these bubbles for $T$ = 280 K (full line) and $T$ = 220 K (dashed line) obtained from DGT for $CO_2$.

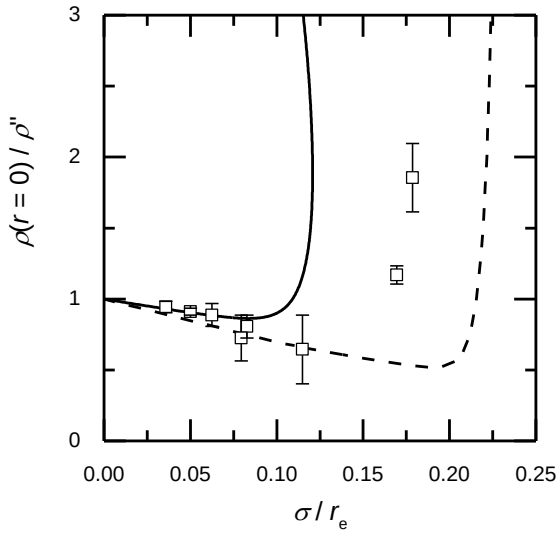


FIG. 14. Density at the center of critical bubbles for $CO_2$ predicted with DGT ($T / T_c$ =0.91: solid line; $T / T_c$ =0.71: broken line) from this work and for the Lennard-Jones truncated shifted fluid (LJTS) from molecular simulation[49] ($T / T_c$ =0.69: symbols) over the inverse equimolar radius. Units are reduced as explained in the text to improve comparability.