



Cite this: *Phys. Chem. Chem. Phys.*,  
2021, **23**, 4517

# Homogeneous nucleation of carbon dioxide in supersonic nozzles II: molecular dynamics simulations and properties of nucleating clusters†

Roope Halonen, <sup>‡\*</sup> Valtteri Tikkanen, <sup>‡\*</sup> Bernhard Reischl, <sup>a</sup>  
Kayane K. Dingilian, <sup>b</sup> Barbara E. Wyslouzil <sup>bc</sup> and Hanna Vehkamäki <sup>a</sup>

Large scale molecular dynamics simulations of the homogeneous nucleation of carbon dioxide in an argon atmosphere were carried out at temperatures between 75 and 105 K. Extensive analyses of the nucleating clusters' structural and energetic properties were performed to quantify these details for the supersonic nozzle experiments described in the first part of this series [Dingilian *et al.*, *Phys. Chem. Chem. Phys.*, 2020, **22**, 19282–19298]. We studied ten different combinations of temperature and vapour pressure, leading to nucleation rates of  $10^{23}$ – $10^{25}$  cm<sup>−3</sup> s<sup>−1</sup>. Nucleating clusters possess significant excess energy from monomer capture, and the observed cluster temperatures during nucleation – on both sides of the critical cluster size – are higher than that of the carrier gas. Despite strong undercooling with respect to the triple point, most clusters are clearly liquid-like during the nucleation stage. Only at the lowest simulation temperatures and vapour densities, clusters containing over 100 molecules are able to undergo a second phase transition to a crystalline solid. The formation free energies retrieved from the molecular dynamics simulations were used to improve the classical nucleation theory by introducing a Tolman-like term into the classical liquid-drop model expression for the formation free energy. This simulation-based theory predicts the simulated nucleation rates perfectly, and improves the prediction of the experimental rates compared to self-consistent classical nucleation theory.

Received 29th October 2020,  
Accepted 4th February 2021

DOI: 10.1039/d0cp05653g

rsc.li/pccp

## 1 Introduction

The nucleation of clusters from a supersaturated vapour has a great significance in atmospheric and engineering processes. Nucleation of CO<sub>2</sub> is of particular importance, for example, in the context of carbon capture and storage (CCS) technologies,<sup>1</sup> as well as in cloud formation in the Martian mesosphere.<sup>2</sup> Modelling nucleation, or analysing experimental results, usually assumes that clusters are in a well defined phase, either solid or liquid, and that the respective bulk properties can be used to calculate nucleation rates, critical cluster sizes, *etc.* However, if nucleation takes place under extreme physical conditions in terms of temperature and pressure, the properties of the nucleating

clusters may deviate substantially from those of the equilibrium phase. In the first part of this series,<sup>3</sup> we studied the homogeneous nucleation of carbon dioxide in an argon atmosphere in supersonic nozzles at temperatures between 75 and 92 K with corresponding CO<sub>2</sub> partial pressures of 39 to 793 Pa, using multiple experimental techniques. Position-resolved pressure trace measurements, small angle X-ray scattering (SAXS) measurements, and Fourier transform infrared (FTIR) spectroscopy were used to determine the onset of nucleation, particle size distributions and aerosol number densities, and the phase of the largest clusters, respectively. While the FTIR spectra suggested that nanometer sized particles had well-defined crystal structures, as expected at these temperatures and pressures, molecular dynamics (MD) simulations revealed that the small clusters critical to the initial stages of nucleation were liquid-like, despite temperatures over 100 K below the triple point of CO<sub>2</sub>,  $T_t = 216.55$  K. This second paper in the series is devoted to a detailed analysis of the properties of the nucleating clusters, and the implications regarding theoretical modelling of the nucleation process.

The theoretical treatment of nucleation has traditionally been a unification of kinetic and equilibrium thermodynamic factors; however cluster energetics should also be considered to account for possible non-equilibrium effects. Recently, the

<sup>a</sup> Institute for Atmospheric and Earth System Research/Physics, University of Helsinki, P.O. Box 64, FI-00014, Helsinki, Finland.

E-mail: roope.halonen@helsinki.fi

<sup>b</sup> William G. Lowrie Department of Chemical and Biomolecular Engineering, The Ohio State University, Columbus, Ohio 43210, USA

<sup>c</sup> Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0cp05653g

‡ Both authors contributed equally to this work.



standard kinetic scheme behind the majority of nucleation theories has been validated by atomistic simulations.<sup>4,5</sup> In classical nucleation theory (CNT), the thermodynamics related to nucleation, *i.e.* the clusters' formation free energies, are treated in a simple fashion using the bulk properties of the nucleating substance. While the CNT formalism often fails drastically to predict measured nucleation rates quantitatively, the theory is able to provide qualitative trends matching the observed data. Several extensions and modifications have therefore been introduced to the classical theory. Many of them are phenomenological, while some require a model of molecular interactions.<sup>6–13</sup>

A very general expression for both the cluster formation free energy  $\Delta W_n$  and potential energy  $U_n$  of a cluster containing  $n$  monomers is

$$\Delta W_n \quad \text{or} \quad U_n = An + Bn^{2/3} + Cn^{1/3} + \dots, \quad (1)$$

which includes size-dependent contributions related to volume (coefficient  $A$ ), surface ( $B$ ) and radius ( $C$ ). Generally, the contributions from the volume and surface terms compete and are the only terms considered in CNT. Yet, it is widely acknowledged that a more realistic expression should include a radius-related term. This term is often associated with the size-dependent surface tension, first proposed by Tolman *et al.*<sup>14</sup> Expression (1) is able to describe large cluster sizes relatively well; however, it can fail for small clusters, especially in the crystalline phase. For example, clusters containing only some tens or hundreds of monomers (translating to cluster radii below 1 nm) may have special geometric and crystallographic structures, or irregular energies and melting points.

Since many nucleation experiments (and simulations) are carried out at temperatures below the triple point, the clusters formed as a result of a rapid, non-equilibrium process may not exhibit the thermodynamically most stable phase. The lack of empirical data for the relevant bulk properties at these low temperatures also means that important properties such as surface tension need to be extrapolated from measurements at higher temperatures, or obtained directly from simulation. Finally, the formation of a cluster releases a considerable amount of latent heat, and due to the limited number of thermalising collisions with the carrier gas, the temperature of the nucleating clusters may deviate from the bath temperature. The classical treatment of this non-isothermal effect, developed by Feder *et al.*,<sup>15</sup> captures the deviation from thermodynamic equilibrium in a simple correction factor applied to the nucleation rate.

The main difficulty in theoretically estimating nucleation rates is modelling the formation free energy of clusters with sufficient accuracy.<sup>16</sup> For the proper treatment of small clusters, atomistic modelling is often unavoidable. For very small clusters, minimum energy structures at 0 K can be calculated using quantum chemical methods, and free energies can be obtained using principles of statistical mechanics.<sup>17</sup> For larger clusters, atomistic interactions can be described using classical force fields, and free energies can be obtained from Monte Carlo methods.<sup>16,18</sup> These free energetics can then be combined with

the standard kinetic scheme of nucleation to calculate actual nucleation rates. In both approaches, it is assumed that the clusters are in perfect thermal equilibrium. However, in this study we are modelling the time evolution of the nucleation process directly using molecular dynamics simulations to account for possible non-equilibrium properties of emerging clusters, which are not captured in the indirect modelling limited to equilibrium conditions. The drawback of direct nucleation simulations is the restriction to conditions where nucleation occurs on a time scale accessible in the simulation. This typically corresponds to nucleation rates that are many orders of magnitude higher than in the experimental processes being modelled.

While many studies on the homogeneous nucleation of the Lennard-Jones fluid and H<sub>2</sub>O have been carried out using MD simulations,<sup>5,19–23</sup> direct nucleation simulations of carbon dioxide remain rare. To our knowledge, only Horsch *et al.*<sup>24,25</sup> have used MD simulations to study homogeneous CO<sub>2</sub> nucleation from the vapour phase. Even though these simulations were carried out in a similar manner to this work, several major differences exist. In particular, the authors used a different force field, and the simulations were performed at higher temperatures corresponding to atmospheric rather than the experimental nozzle conditions. Furthermore, temperature control was achieved by a global velocity scaling scheme, which may result in an unphysical removal of latent heat from the nucleating clusters.<sup>4,21</sup>

The outline of the paper is as follows. Section 2 explains the theoretical framework of nucleation and the details of the molecular dynamics simulations. Section 3 presents detailed results of nucleation simulations and the analyses of the nucleating clusters. Section 4 describes the simulation-based improvement of classical nucleation theory. Section 5 summarizes and concludes the work.

## 2 Theories and methods

### 2.1 Theoretical framework of nucleation

The classical nucleation theory (CNT), or one of its variants, is often applied to predict nucleation rates and to interpret results obtained from experiments and simulations. In CNT, the formation of a new phase from a supersaturated vapour is assumed to proceed through a critical cluster, and clusters can only grow by monomer addition and decay by monomer loss. By combining equilibrium thermodynamics with kinetic gas theory, the classical expression for the nucleation rate in a supersaturated vapour is given as<sup>26,27</sup>

$$J = f\beta_n Z S^{n^*} N_n^{\text{eq}}, \quad (2)$$

where  $n^*$  is the number of monomers in the critical cluster,  $N_n^{\text{eq}}$  is the equilibrium critical cluster density,  $f$  is the non-isothermal correction factor,  $Z$  is the Zeldovich factor, and  $S$  is the saturation ratio defined as the ratio between the vapour pressure of the monomers and equilibrium pressure over the bulk liquid. The collision rate between monomers and an



$n$ -cluster with radii  $r_1$ , and  $r_n$ , respectively, can be obtained from kinetic gas theory at temperature  $T$  as

$$\beta_n = N_1 \sqrt{\frac{8\pi RT(n+1)}{nm}} (r_1 + r_n)^2, \quad (3)$$

where  $N_1$  is the monomer number density,  $R$  is the gas constant and  $m$  is the monomer mass. Assuming the clusters are perfectly spherical with bulk liquid number density  $\rho_l$ , the cluster radius can be calculated using the liquid-drop radius model as

$$r_n = \left( \frac{3n}{4\pi\rho_l} \right)^{1/3}. \quad (4)$$

While neglecting long-range interactions between molecules systematically underestimates the collision rates,<sup>28–31</sup> the enhancement factor for weakly interacting neutral molecules, such as CO<sub>2</sub>, is likely to be rather small.

In equilibrium with bulk liquid (*i.e.*  $S = 1$ ), the cluster distribution  $N_n^{\text{eq}}$  is related to the total number density of monomers and clusters  $N_{\text{tot}}^{\text{eq}}$  in the system and the equilibrium formation free energy  $\Delta W_n^{\text{eq}}$  of an  $n$ -cluster as

$$N_n^{\text{eq}} = N_{\text{tot}}^{\text{eq}} \exp\left(-\frac{\Delta W_n^{\text{eq}}}{RT}\right). \quad (5)$$

The equilibrium cluster distribution in the supersaturated vapour is given as

$$\bar{N}_n = S^n N_n^{\text{eq}}, \quad (6)$$

and assuming that the actual monomer density in the nucleating vapour  $N_1 = \bar{N}_1$ , eqn (2) can be written as

$$J = f \beta_n Z N_1 \exp\left(-\frac{\Delta W_n^*}{RT}\right). \quad (7)$$

Here, the formation free energy of an  $n$ -cluster in the supersaturated vapour is

$$\Delta W_n^* = \Delta W_n^{\text{eq}} - \Delta W_1^{\text{eq}} - (n^* - 1)RT \ln S, \quad (8)$$

and this construction satisfies both the law of mass action and the requirement for self-consistency.<sup>6,7,32</sup> In the liquid-drop model, the equilibrium formation free energy is characterised by the clusters' surface tension  $\sigma_n$  and surface area  $A_n = 4\pi r_n^2$ , as

$$\Delta W_n^{\text{eq}} = A_n \sigma_n = A_1 \sigma_n n^{2/3}, \quad (9)$$

where  $A_1$  is the monomer surface area. In CNT, based on the capillary approximation, the surface tension of a cluster equals that of an infinite planar interface, *i.e.*,  $\sigma_n = \sigma$ . Setting the derivative of eqn (8) with respect to  $n$  to zero results in the classical expressions for the critical cluster size,

$$n_{\text{CNT}}^* = \frac{32\pi\sigma^3}{3\rho_l^2(RT \ln S)^3}, \quad (10)$$

and the critical cluster radius,

$$r_{\text{CNT}}^* = \frac{2\sigma}{\rho_l RT \ln S}. \quad (11)$$

The Zeldovich factor is defined as

$$Z = \sqrt{\frac{Q}{2\pi RT}}, \quad (12)$$

where  $Q$  is the second derivative of the formation free energy at the critical size, and in the context of CNT,

$$Q_{\text{CNT}} = -\frac{\partial^2 \Delta W_n^*}{\partial n^2} = \frac{2}{9} A_1 \sigma (n_{\text{CNT}}^*)^{-4/3}. \quad (13)$$

Various corrections or extensions to the CNT liquid-drop model have been suggested, many of them concentrate on the radius-related term in eqn (1). A common modification to CNT is to include the size-dependency of the droplet's surface tension in the formation free energy expression by using the Tolman equation,<sup>14</sup>

$$\sigma_n = \sigma \left( 1 + \frac{2\delta_T}{r_1} n^{-1/3} \right)^{-1}, \quad (14)$$

where  $\delta_T$  is the characteristic Tolman length. Expanding in a Taylor series and truncating after the second term yields

$$\sigma_n = \sigma \left( 1 - \frac{2\delta_T}{r_1} n^{-1/3} \right). \quad (15)$$

Inserting this into eqn (9) creates an additional term  $\propto n^{1/3}$  in the liquid-drop model with  $C = 8\pi r_1 \sigma \delta_T$ . The difficulty in routinely using this equation is the uncertainty in both the magnitude and sign of the Tolman length.<sup>14,33–35</sup>

Classical nucleation theory also assumes perfect thermal equilibrium throughout the system. However, cluster formation involves the release of substantial amounts of latent heat, which requires subsequent collisions with the carrier gas to be dissipated. To account for any non-isothermal effects, we use the classical non-isothermal correction factor of Feder *et al.*,<sup>15</sup>

$$f = \frac{b^2}{b^2 + q^2}, \quad (16)$$

where  $b^2$  is the mean square of the energy fluctuation of the colliding monomers and carrier gas atoms, denoted with subscript c, given by

$$b^2 = 2(RT)^2 \left( 1 + \frac{N_c}{N_1} \sqrt{\frac{m}{m_c}} \right), \quad (17)$$

and  $q$  is the energy released upon the addition of a monomer to the critical cluster. In the original theory by Feder *et al.*,<sup>15</sup>  $q$  is given as

$$q = \Delta H_{\text{vap}} - \frac{RT}{2} - \frac{\partial A_n \sigma}{\partial n}. \quad (18)$$

However, as noted by Kuni *et al.*<sup>36</sup> and Barrett,<sup>37</sup> the last term incorrectly introduces surface entropy in this expression. In this study, we have therefore estimated the size-dependent  $q_n$  directly from the atomistic simulations as

$$q_n = U_{n-1} - U_n, \quad (19)$$



where  $U_n$  is the potential energy of an  $n$ -cluster (see ESI† for more detail). The magnitude of the non-isothermal correction is typically between 0.01 and 1.

In this work, the nucleation rates calculated using eqn (7)–(13), (16), (17), and (19) are denoted by  $J_{\text{CNT}}$ .

## 2.2 Molecular dynamics simulation of carbon dioxide

Classical molecular dynamics simulations (MD) are employed to study the homogeneous nucleation of  $\text{CO}_2$  clusters in an atmosphere of  $\text{CO}_2$  and Ar.  $\text{CO}_2$  molecules are described by the TraPPE potential,<sup>38</sup> using a rigid geometry with a C–O bond length of 1.16 Å and O–C–O angle of 180°. This potential has been fitted to reproduce the experimental liquid–vapour coexistence line and the Clausius–Clapeyron curve of the saturated vapour pressure, and it also reproduces the low-temperature vapour–liquid equilibria and solid–vapour equilibria reasonably well.<sup>39</sup> Argon is modelled as a nonpolar atom with only pairwise interactions with C, O and other Ar atoms. The pairwise interactions between atoms  $i$  and  $j$  with partial charges  $q_i$  and  $q_j$ , separated by a distance  $r_{ij}$ , are described by the sum of Lennard-Jones and Coulomb interactions as

$$U_{ij} = 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}. \quad (20)$$

The force field parameters  $\varepsilon_{ii}$ ,  $\sigma_{ii}$ , and  $q_i$ , for C, O and Ar atoms are shown in Table 1. For the interactions between dissimilar atoms, Lorentz–Berthelot mixing rules are applied,

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}, \quad \sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2}. \quad (21)$$

All simulations have been carried out with the LAMMPS MD code,<sup>41</sup> using a velocity-Verlet integrator and a time step of 5 fs. Lennard-Jones and Coulomb interactions were cut off at 10 Å and 100 Å, respectively. For large systems with low density, evaluating the Coulomb interaction in real space only, but with a large cut-off, was found to be an efficient yet accurate alternative to the usual Ewald-type approach. The standard approach of using a global thermostatting algorithm is not suitable for directly simulating a nucleating vapour, as it leads to an unphysical removal of latent heat from the nucleating clusters. It has been shown that the best way to avoid this artefact is to introduce explicit carrier gas molecules in the simulation, and apply a thermostat only to the latter, even though this increases the computational cost significantly.<sup>4,5,19–21</sup> Here, we have applied a Nosé–Hoover thermostat with time constant 0.1 ps to the Ar atoms. This approach prevents heating of the carrier gas due to latent heat release. However, at the studied conditions,

this is a minor effect and we chose to neglect it in the simulation setup.

Clusters are distinguished from the vapour by using a Stillinger criterion<sup>42</sup> with a carbon–carbon distance cut-off of 6 Å, roughly corresponding to the position of the second peak in the radial distribution function of liquid  $\text{CO}_2$ .

To study the structural and energetic properties of the emerging clusters, we saved the atomic positions and velocities of clusters larger than four monomers from the nucleation simulations for further analysis. The structural and energetic properties of the isolated clusters were then evaluated in separate microcanonical MD simulation runs. These *NVE* simulations were carried out for 1 ns or until the cluster evaporated: the integrity of the cluster was checked every 50 ps and, in case one or more molecules had evaporated, the last 50 ps of the simulation data were discarded.

Moreover, a subset of these isolated clusters were also equilibrated by coupling them directly to a Nosé–Hoover thermostat with a time constant of 5 ps. In these simulations, the Nosé–Hoover thermostat mimics heat exchange with carrier gas of a well-defined temperature. By choosing a relatively weak coupling, we ensure that the clusters are able to explore phase space efficiently and have a chance to undergo a liquid-to-solid phase transition. The duration of these *NVT* simulations varied by cluster size. The simulations were run until a clear transition from the initial state was observed as a drop in potential energy.

## 3 Results and discussion

### 3.1 Direct nucleation simulations

We have studied 10 individual systems of supersaturated  $\text{CO}_2$  vapour at different densities and temperatures. The selected bath temperatures,  $T = 75, 90$  and  $105$  K, correspond to extremely undercooled vapour and overlap with the experimental temperature range in part I of this study.<sup>3</sup> The simulated  $\text{CO}_2$  number densities are slightly higher than the experimental ones. To observe nucleation on a computationally accessible time-scale, nucleation rates in the simulations have to be on the order of  $10^{24} \text{ cm}^{-3} \text{ s}^{-1}$ , whereas the supersonic nozzle experiments<sup>3</sup> were constrained to rates of about  $10^{17} \text{ cm}^{-3} \text{ s}^{-1}$ . All the simulated systems consisted of 64 000  $\text{CO}_2$  molecules and an equal number of thermalizing Ar gas atoms. This corresponds to the upper end of  $\text{CO}_2$ :Ar ratios in the experiments where  $\text{CO}_2$  concentrations reached 39 mol%.

The nucleation rates can be calculated from the time evolution of the cluster density using the Yasuoka–Matsumoto threshold method:<sup>19</sup>

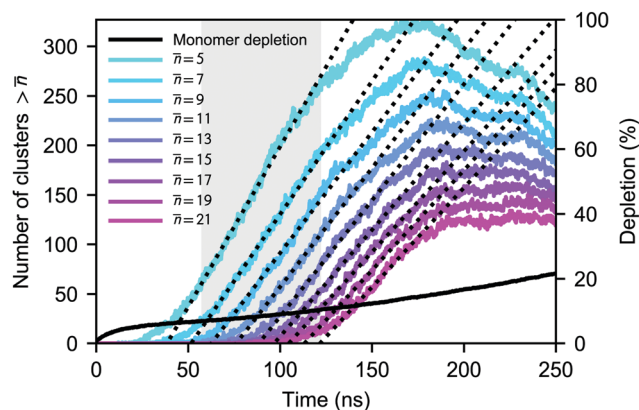
$$J_{\text{sim}} = \frac{1}{\tau V} = \frac{dN_{n>\bar{n}}}{dt}, \quad (22)$$

where  $t$  is the elapsed simulation time,  $V$  is the volume of the simulated system, and  $N_{n>\bar{n}}$  is the number density of clusters larger than some threshold size  $\bar{n}$ . Here,  $\tau$  is proportional to the slope of the linear part of the curve, shown in Fig. 1 (equivalent graphs are shown for the other simulation runs in ESI,† Fig. S1), during the nucleation stage. This nucleation stage is

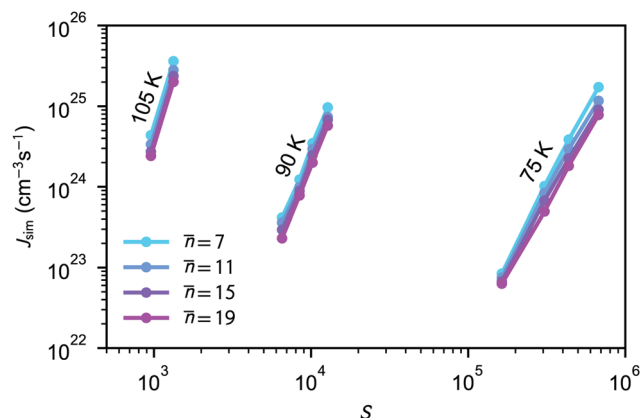
**Table 1** Lennard-Jones parameters ( $\varepsilon_{ii}$  and  $\sigma_{ii}$ ), partial charges  $q_i$  and masses  $m_i$  for the simulated atoms

Atom	$\varepsilon_{ii}$ (kcal mol <sup>−1</sup> )	$\sigma_{ii}$ (Å)	$q_i$ (e)	$m_i$ (g mol <sup>−1</sup> )	Ref.
C	0.05365	2.8	0.7	12.0	38
O	0.15697	3.05	−0.35	16.0	38
Ar	0.238	3.4	0.0	40.0	40





**Fig. 1** Simulated time-evolution of the total amount of clusters larger than the threshold size  $\bar{n}$  at a  $\text{CO}_2$  number density of  $N = 2 \times 10^{-6} \text{ Å}^{-3}$  and  $T = 75 \text{ K}$  (colored lines). The linear fits whose slopes correspond to  $\tau$  are shown as dotted lines. The black solid line shows the level of monomer depletion  $D$  during the simulation. The “nucleation stage” for  $\bar{n} = 5$ , i.e., the time period during which the average monomer density was calculated, is indicated by the shaded gray area.



**Fig. 2** Simulated nucleation rates for four different threshold values  $\bar{n}$  shown as a function of saturation ratio  $S$ . According to the nucleation theorem, eqn (23), for constant temperature, the slope of a presented line on this graph equals approximately the critical cluster size  $n^*$ .

arbitrarily estimated as the time period when the number of clusters  $n > \bar{n}$  lies between 20% and 80% of the observed maximum. However, the main results of this study are rather insensitive to this definition. Ideally,  $\tau$  should be constant if  $\bar{n}$  is sufficiently larger than the critical size. In our case, however, the monomer density is decreasing too quickly as discussed below and nucleation, consequently, slows down. We have estimated the nucleation rates using multiple threshold sizes, and verified that the magnitude of the rate for an individual run remains largely unaffected by that choice.

The average nucleation rates  $J_{\text{sim}}$  are presented in Table 2 and Fig. 2. The results clearly show that the nucleation rate increases with saturation ratio  $S$ . Additional information about the nucleation rate calculations and equilibrium pressure can be found in the ESI†.

As the number of molecules is kept constant during the simulations, the monomer pool is continuously depleted due to cluster formation. The simulated nucleation process is relatively

aggressive, clusters are formed quickly and in large numbers, and most of the simulation runs do not reach a steady monomer concentration as shown in Fig. 1 for one case. To estimate the monomer density (and saturation ratio) at which the nucleation actually takes place, we have calculated the average monomer density at the nucleation stage corresponding to  $n > 5$ . In the case of the densest system at 75 K, almost one quarter of the monomers are lost to clusters already after 5 ns. At higher temperatures, the number of clusters formed and consequently the extent of depletion are smaller, yet the nucleation rates are on the same order of magnitude as in the simulations carried out at 75 K.

During the nucleation stage, the possibility for cluster-cluster collisions is very minor; the ratio of dimers to free monomers is at most  $\sim 0.05$ . However, after the nucleation stage, the reduction in the number of small clusters in Fig. 1 and Fig. S1 (ESI†) is due to the abundance of very large clusters, which eventually start to scavenge other clusters *via* coagulation, in addition to the increase in monomer depletion.

Even though we find a systematic decrease of  $\tau$  with threshold size, as shown in Fig. 1 and Fig. S1 (ESI†), the slopes suggest

**Table 2** Nucleation simulation results:  $N$  is the total density of  $\text{CO}_2$  molecules,  $N_1$  is the average monomer density during the nucleation stage, and  $D$  is the extent of depletion. The presented saturation ratios are calculated as  $S = N_1/N^{\text{eq}}$ , where  $N^{\text{eq}}$  is the ideal vapour density corresponding to equilibrium vapour pressure of a liquid TraPE- $\text{CO}_2$  (see Table S1 in ESI). The nucleation rates are presented according to the Yasuoka–Matsumoto method *i.e.* eqn (22) ( $J_{\text{sim}}$ ), implementing the barrier height calculated from the cluster distributions into eqn (7) ( $J_{\text{AW}}$ ), and simulation-based ( $J_{\text{ST}}$ ) and classical nucleation theory ( $J_{\text{CNT}}$ )

$T$ (K)	$N$ ( $\text{Å}^{-3}$ )	$N_1$ ( $\text{Å}^{-3}$ )	$D$ (%)	$\ln S$	$J_{\text{sim}}$ ( $\text{cm}^{-3} \text{ s}^{-1}$ )	$J_{\text{AW}}$ ( $\text{cm}^{-3} \text{ s}^{-1}$ )	$J_{\text{ST}}$ ( $\text{cm}^{-3} \text{ s}^{-1}$ )	$J_{\text{CNT}}$ ( $\text{cm}^{-3} \text{ s}^{-1}$ )
75	$2.0 \times 10^{-6}$	$1.83 \times 10^{-6}$	8.5	12.01	$7.5 \times 10^{22}$	$1.9 \times 10^{23}$	$6.7 \times 10^{22}$	$1.2 \times 10^{-2}$
75	$4.0 \times 10^{-6}$	$3.42 \times 10^{-6}$	14.6	12.63	$8.4 \times 10^{23}$	$1.9 \times 10^{24}$	$6.4 \times 10^{23}$	$2.0 \times 10^{-1}$
75	$6.0 \times 10^{-6}$	$4.88 \times 10^{-6}$	18.7	12.99	$2.9 \times 10^{24}$	$6.5 \times 10^{24}$	$2.1 \times 10^{24}$	$7.7 \times 10^{-1}$
75	$1.0 \times 10^{-5}$	$7.55 \times 10^{-6}$	24.5	13.42	$1.2 \times 10^{25}$	$2.8 \times 10^{25}$	$8.2 \times 10^{24}$	$3.2 \times 10^0$
90	$6.0 \times 10^{-6}$	$5.48 \times 10^{-6}$	8.6	8.79	$3.6 \times 10^{23}$	$7.4 \times 10^{23}$	$2.9 \times 10^{23}$	$4.5 \times 10^2$
90	$8.0 \times 10^{-6}$	$7.08 \times 10^{-6}$	11.5	9.04	$1.0 \times 10^{24}$	$2.1 \times 10^{24}$	$1.0 \times 10^{24}$	$3.0 \times 10^3$
90	$1.0 \times 10^{-5}$	$8.57 \times 10^{-6}$	14.3	9.23	$2.9 \times 10^{24}$	$5.6 \times 10^{24}$	$2.7 \times 10^{24}$	$1.6 \times 10^4$
90	$1.3 \times 10^{-5}$	$1.07 \times 10^{-5}$	17.9	9.45	$7.4 \times 10^{24}$	$1.5 \times 10^{25}$	$7.5 \times 10^{24}$	$7.6 \times 10^4$
105	$2.0 \times 10^{-5}$	$1.71 \times 10^{-5}$	14.4	6.86	$2.8 \times 10^{24}$	$6.0 \times 10^{24}$	$3.3 \times 10^{24}$	$5.9 \times 10^6$
105	$3.0 \times 10^{-5}$	$2.38 \times 10^{-5}$	20.7	7.19	$2.2 \times 10^{25}$	$4.5 \times 10^{25}$	$2.9 \times 10^{25}$	$3.0 \times 10^8$





that the critical sizes are below 7, 9, and 11 monomers at  $T = 75$ , 90, and 105 K, respectively (see ESI†). If the cluster growth follows classical theory, the critical cluster size can be estimated from the first nucleation theorem:<sup>43</sup>

$$\left(\frac{\partial \ln J}{\partial \ln S}\right)_T \approx n_{\text{NT}}^* \quad (23)$$

Thus, the critical size can be directly determined from a linear fit to a double logarithmic plot of  $J$  versus  $S$ , as shown in Fig. 2. Simulated nucleation rates at  $T = 75$  and 90 K show clear linear behaviour over the studied range of saturation ratios, indicating similar critical sizes at different monomer densities. The average critical sizes are  $n_{\text{NT}}^* \approx 4$ , 5, and 7 at  $T = 75$ , 90, and 105 K, respectively. These values are compared with results derived from theoretical models in Section 4 and Table 3. The selected threshold size has only a slight impact on the results obtained through the nucleation theorem.

**3.1.1 Cluster formation free energies.** In order to understand the thermodynamics of homogeneous  $\text{CO}_2$  nucleation, the formation free energies are derived from the cluster data generated by the nucleation simulations. As the cluster distribution obtained from direct MD simulation corresponds to the actual cluster densities,  $N_n$ , at the nucleation stage, we can calculate the formation free energies using the concept of detailed balance at the supersaturated equilibrium (zero net flux),

$$\beta_{n-1}\tilde{N}_{n-1} - \alpha_n\tilde{N}_n = 0, \quad (24)$$

and in the nucleation stage (non-zero net flux  $J$ ),

$$\beta_{n-1}N_{n-1} - \alpha_nN_n = J, \quad (25)$$

where  $\alpha_n$  is the evaporation rate of an  $n$ -cluster, which depends only on the bath temperature. Since, as implied in eqn (7),  $\tilde{N}_n = \tilde{N}_1 \exp(-\Delta W_n/RT)$ , we can derive the following relationship for the free energy change upon monomer addition using eqn (24) and (25):

$$\begin{aligned} \delta\Delta W_n &= \Delta W_n - \Delta W_{n-1} \\ &= -RT \left[ \ln N_n - \ln \left( N_{n-1} - \frac{J}{\beta_{n-1}} \right) \right]. \end{aligned} \quad (26)$$

**Table 3** Critical cluster sizes of the simulated systems estimated with different approaches: threshold size  $\tilde{n}$  (the critical size should be smaller than this), nucleation theorem  $n_{\text{NT}}^*$ , free energy profile calculated from the simulated cluster distributions  $n_{\text{sim}}^*$ , simulation-based theory  $n_{\text{ST}}^*$ , and classical theory  $n_{\text{CNT}}^*$ . The planar surface tensions extrapolated from Fig. 3 are used to calculate  $n_{\text{ST}}^*$  and  $n_{\text{CNT}}^*$

$T$ (K)	$N$ ( $\text{\AA}^{-3}$ )	$\tilde{n}$	$n_{\text{NT}}^*$	$n_{\text{sim}}^*$	$n_{\text{ST}}^*$	$n_{\text{CNT}}^*$
75	$2.0 \times 10^{-6}$	$<7$	3.6	8.1	6.5	16.5
75	$4.0 \times 10^{-6}$	$<7$	3.6	7.6	5.9	15.6
75	$6.0 \times 10^{-6}$	$<7$	3.6	7.3	5.6	15.1
75	$1.0 \times 10^{-5}$	$<7$	3.6	7.0	5.3	14.7
90	$6.0 \times 10^{-6}$	$<9$	4.7	8.9	8.0	18.9
90	$8.0 \times 10^{-6}$	$<9$	4.7	9.1	7.5	18.1
90	$1.0 \times 10^{-5}$	$<9$	4.7	9.0	7.1	17.5
90	$1.3 \times 10^{-5}$	$<9$	4.7	7.9	6.7	16.9
105	$2.0 \times 10^{-5}$	$<11$	6.5	10.3	9.1	20.7
105	$3.0 \times 10^{-5}$	$<11$	6.5	8.6	7.9	18.9

As we have assumed that  $N_1 = \tilde{N}_1$ , the monomer formation free energy is by definition zero. The  $n$ -cluster formation free energy is now easily computed as  $\Delta W_n = \sum_{i=1}^n \delta\Delta W_i$ . Note, that possible size-independent liquid-drop model terms of  $\Delta W_n$  are now neglected, but this is equivalent to the “self-consistent” free energy treatment in eqn (8).

In order to retrieve the formation free energies from the simulations, we have analysed the cluster distributions during the nucleation stage (as defined earlier, using the “20–80%” rule) to calculate  $J$  using different threshold values  $\tilde{n}$ . The free energy curves obtained using eqn (26) are shown in the upper panel of Fig. 3. At each temperature studied, a clear barrier can be seen, whose height decreases as the vapour density increases. The shape of the tail of the curve at the largest cluster sizes is affected by the threshold size  $\tilde{n}$  used; this might be a consequence of the shift of the nucleation stage in time (eventually nucleation is hindered by depletion, and coagulation becomes significant). The value used for the collision rate  $\beta_n$  also affects the shape somewhat: the barrier becomes sharper if the collision rate is reduced. However, both the barrier height and the subcritical free energies are rather robust against changes in threshold size and collision rate.

To reduce the calculated formation free energies at supersaturation ( $S > 1$ ) to equilibrium values ( $S = 1$ ), the volume-dependent term shown in eqn (9) must be subtracted. According to the extended liquid-drop model, the equilibrium formation energy of an  $n$ -cluster is

$$\Delta W_n^{\text{eq}} = \Delta W_n + (n - 1)RT \ln S = A_1 \sigma n^{2/3} + Cn^{1/3}. \quad (27)$$

To study possible contributions to the free energies related to  $n^{1/3}$ , we have plotted  $\Delta W_n^{\text{eq}}/A_n$  as a function of  $n^{-1/3}$  in the bottom panel of Fig. 3, where we have used the liquid-drop radius model (eqn (4)) to calculate the surface area. Fig. 3 clearly shows that the equilibrium free energy per surface area of the simulated  $\text{CO}_2$  clusters depends linearly on  $n^{-1/3}$ . The same behaviour was observed by Tanaka *et al.*<sup>22</sup> for Lennard-Jones clusters. We use this observation to formulate a so-called “simulation-based” theory in Section 4.

The values extrapolated from the data give quite reasonable values for the surface tension ( $\sigma = \lim_{n \rightarrow \infty} \Delta W_n^{\text{eq}}/A_n$ ), of about 80.2, 70.9 and 64.7  $\text{mN m}^{-1}$  at  $T = 75$ , 90 and 105 K, respectively. These simulation-based “planar surface tensions” are higher than those obtained from the thermophysical relations of Quinn<sup>44</sup> (55.2, 50.8, 46.4  $\text{mN m}^{-1}$ ) and Lielmezs–Herrick<sup>3,45</sup> (67.9, 60.2 and 53.4  $\text{mN m}^{-1}$ ), and there is a slight dependence on vapour density. Notably, the simulated cluster distributions used in eqn (26) correspond to non-isothermal conditions, hence the effect of insufficient thermalisation is embodied in the  $\Delta W_n$  (the barrier is higher as nucleation is suppressed). The nonisothermality is inherently a thermodynamic effect and thus related to the free energies, yet in the theory of Feder *et al.*,<sup>15</sup> the effect of excess energy is reduced to a prefactor in the nucleation rate expression.

To test the reliability of the kinetic scheme of the classical theory (eqn (7)), we have used the barrier heights shown in the upper panel



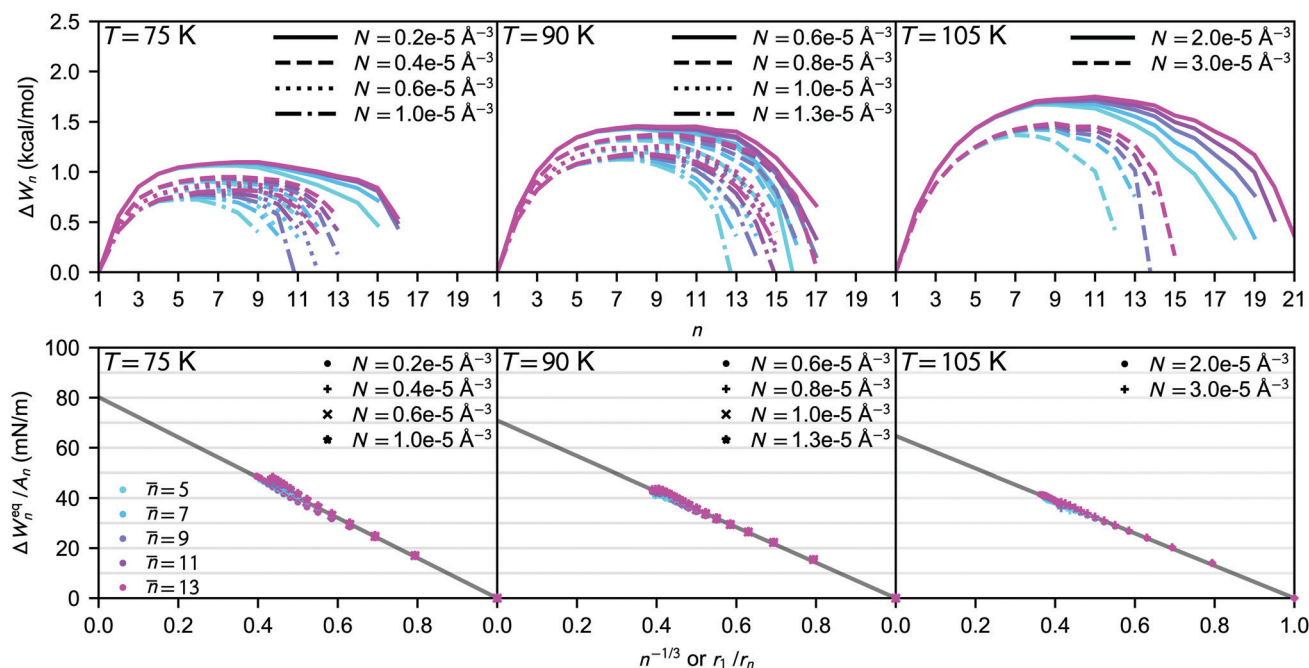


Fig. 3 (upper panel) Formation free energies retrieved from the MD simulations at 75, 90 and 105 K. The different line types (solid, dashed, dotted and dash-dotted) represent the various total number densities of  $\text{CO}_2$ , and the different colours correspond to different threshold sizes used to determine the nucleation rate, as shown in the legend in the lower left-most figure. (lower panel) The equilibrium formation free energy per cluster area as a function of  $n^{-1/3}$  (or  $r_1/r_n$ ). The different number densities are shown with different symbols, while the colours correspond to those in the upper panel.

of Fig. 3 to back-calculate the simulated nucleation rates. As the planar surface tension can be evaluated from the results shown in Fig. 3, we have used eqn (12) and (13) to estimate the value of the Zeldovich factor. The results are shown in Table 2 as  $J_{\Delta W}$ . The kinetic scheme seems to work rather well as the back-calculated nucleation rates are within a factor of two of the directly calculated ones, corresponding to an uncertainty of  $RT$  in the barrier height.

### 3.2 Cluster analysis

We performed a detailed analysis of the structural and energetic properties of  $\text{CO}_2$  clusters during the nucleation stage and compared them to their equilibrated counterparts.

From the direct nucleation simulations, a large set of clusters in the size range  $n = 5 \dots 315$ , was extracted and examined in isolation. For each combination of temperature and vapour pressure, the collected sample consisted of at least tens of thousands of statistically different configurations for clusters containing 10 monomers or less, and, for example, at least hundreds of configurations for 50-clusters. For sizes exceeding 100 molecules, the sample did not contain configurations for each individual cluster size.

To study the size-specific properties of nucleation-stage clusters, all of the isolated clusters were simulated in *NVE* for a maximum of 1 ns, and the potential energy,  $U$ , and kinetic energy,  $K$ , were stored. As evaporation events are rather frequent in the studied size range, the runs were divided into 50 ps sub-runs. In case of evaporation, the simulation was terminated and the last 50 ps of data discarded, to avoid spurious contributions to the average from fragmented clusters. As the resulting energy

distributions were near-Gaussian, they are well-described by the mean values of the energy.

The properties of equilibrated clusters of different sizes were investigated by simulating a subset of the clusters isolated from the nucleation simulations in *NVT* for a few to hundreds of nanoseconds, depending on the cluster size and temperature. For small clusters, at least tens of different cases were studied for each size, with sample standard deviations of  $U$  being very small (a few percent at most). For clusters exceeding the size  $n = 50$ , even a single cluster could yield useful information, as all large clusters exhibited well-defined crystalline configurations.

**3.2.1 Cluster radii.** The radii of  $\text{CO}_2$  clusters during the nucleation stage were analysed based on the *NVE* simulation data. Eqn (4) gives an estimate for the equimolar cluster radius assuming it to be a sphere with bulk liquid density. However, especially the radii of the smallest clusters might differ significantly from the liquid-drop radius model. To evaluate the actual size of the clusters, we have calculated the radial distribution function (RDF) between the carbon atoms and the cluster's centre of mass. The averaged RDFs for cluster sizes  $n = 10 \dots 34$  (from simulations carried out at 75 K) are shown in Fig. S2 (ESI†); the shapes of the RDFs indicate a liquid-like inner structure for the clusters, as no sharp peaks are present. The obtained density profiles match rather well with the experimental bulk liquid densities<sup>46</sup> (for clusters  $n > 150$ , the agreement is within 2%), and the location of the equimolar dividing surface is quite accurately predicted by the liquid-drop radius model.

As shown in Fig. 4, the deviation between the equimolar radius  $r_n^{\text{MD}}$  obtained from the analysis of the clusters' radial density



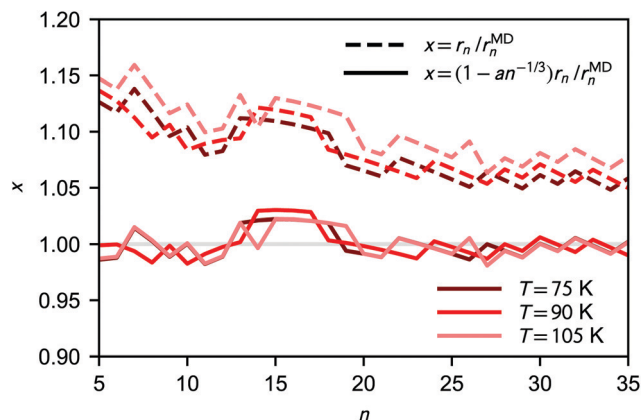


Fig. 4 Ratio of liquid-drop radius and simulated equimolar radius (dashed lines) as a function of cluster size. The solid lines shows the fitted correction to the liquid-drop model ( $a \approx 0.25$ ).

profiles and the liquid-drop radius  $r_n$  increases for the smaller clusters. This difference can be fitted quite accurately as

$$r_n^{\text{MD}} = (1 - an^{-1/3})r_n, \quad (28)$$

where  $a$  is a positive coefficient, and for all studied temperatures,  $a \approx 0.25$ .

The consequences of a deviation in cluster size compared to classical theory are mostly thermodynamical, *i.e.*, related to the surface area in eqn (9); the effect on the collision rate is very minor. In line with kinetic theory, possible effects due to long-range interactions on collision rates<sup>28,30,31</sup> are neglected. The implications of the radius correction is further discussed in Section 4.

**3.2.2 Kinetic energy and temperature.** Due to the insufficient thermalization through collisions with carrier gas, the nucleating clusters accumulate a considerable amount of latent heat released by each monomer capture. The cluster temperature during nucleation,  $T_n$ , is estimated from the kinetic energy data gathered from the *NVE* simulations. After the translational motion of the simulated clusters is removed, the cluster temperature is obtained from the average kinetic energy  $K_n$  as

$$T_n = \frac{2K_n}{(5n - 3)R}. \quad (29)$$

We have used a simple arithmetic average for  $K_n$  as the energy distributions are close to Gaussian over the range of studied cluster sizes (see Fig. S3 in ESI†). As shown in Fig. 5, the cluster temperatures are overall higher than the bath temperature, and the cluster temperature increases with size. The heating of the clusters with respect to the bath temperature,  $\Delta T_n = T_n - T$ , is more pronounced for the densest simulated systems due to more frequent growth events. The efficiency of thermalisation is similar in simulations at different densities and same temperature, when the ratio of  $\text{CO}_2$  to carrier gas remains roughly constant, as seen in eqn (17). Fig. 5 reveals a clear temperature-dependency in the magnitude of  $\Delta T_n$ , with larger heating observed at lower bath temperature, which can be due to differences in the amount of latent heat released. In Fig. S4 in ESI†, we have estimated cluster-specific values of the energy

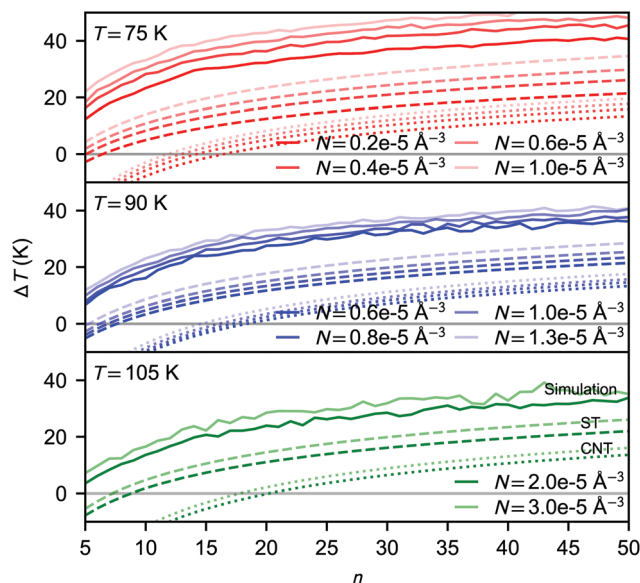


Fig. 5 Average cluster temperature difference relative to the bath temperature as a function of cluster size: for simulations carried out at 75 K (upper panel), 90 K (middle panel) and 105 K (lower panel). The solid lines correspond to average kinetic energies calculated from the *NVE* simulations. The predicted values using eqn (30) are shown for CNT and the simulation-based theory as dotted and dashed lines, respectively.

released upon monomer addition,  $q$ , from the *NVE* simulations. The difference between  $q$  of clusters simulated at 75 and 90 K (and similarly between 90 and 105 K) is about  $0.2 \text{ kcal mol}^{-1}$  which, for  $n = 5$ , corresponds to about 9 K. This matches the observed  $\Delta T_n$  difference between systems with roughly equal number densities at different bath temperatures.

The temperature difference expression related to the non-isothermal theory is given as<sup>15,21</sup>

$$\Delta T_n = -T \frac{q}{b^2 + q^2} \left( \frac{\partial \Delta W_n}{\partial n} \right). \quad (30)$$

Note, here the excess energy  $q$  is given for the critical cluster size, as in eqn (16). The theoretical values of  $\Delta T_n$  with  $\Delta W_n$  taken from both CNT and simulation-based theory (see Section 4) are shown in Fig. 5. Eqn (30) yields a much lower temperature difference than the simulations by construction, as  $\Delta T_n = 0$  at the critical size. Yet, the shape of the curve and the relative increase in cluster temperature as the saturation ratio rises is rather well predicted by eqn (30), especially when using the simulation-based theory.

In rather similar nucleation simulations of Lennard-Jones vapour, Wedekind *et al.*<sup>21</sup> found that the cluster temperature distributions were skewed, diverging from Gaussian, and they focused on the peak value instead of the mean of the cluster temperature distribution. In our case, as shown in Fig. S3 (ESI†), the temperature distributions of  $\text{CO}_2$  clusters are clearly Gaussian, and there is little difference between mean and peak values.

It has been a subject of a long debate whether the subcritical clusters are colder than the bath temperature or hotter just like





the rest of the clusters in a non-isothermal system. Following the seminal study by Feder *et al.*,<sup>15</sup> the “cool cluster” hypothesis has both been supported<sup>21,47</sup> and rejected.<sup>48–51</sup> The observed discrepancy between the simulated and theoretical cluster temperature shift may at first appear as a failure of eqn (16), however, as already briefly discussed by Feder *et al.*,<sup>15</sup> the observed  $\Delta T_n$  might always be positive. The reason is that even if the majority of the subcritical clusters are above the bath temperature, nucleation is mainly promoted by the rare “cold” and stable clusters. A more elaborate discussion about the cluster temperatures is presented by Barrett.<sup>50</sup>

**3.2.3 Potential energy and surface energy.** The average potential energy per molecule  $U_n/n$  as a function of  $n^{-1/3}$  for both nucleating (NVE) and equilibrated (NVT simulated) clusters are shown in the inset of Fig. 6 for a bath temperature of 75 K (similar plots at  $T = 90$  and 105 K are presented in Fig. S5 and S6, ESI†). To better compare these two sets, we have rescaled the vibrational energy contribution to the potential energy of the NVE simulated clusters at their intrinsic temperature  $T_n$  to the bath temperature  $T$ , using the harmonic approximation in the high temperature limit:

$$U_n(T) = U_n(T_n) - \frac{5n - 6}{2} R \Delta T_n. \quad (31)$$

When the results are presented as a function of  $n^{-1/3}$ , the energies of nucleating clusters clearly follow a liquid-drop model:

$$U_n = An + Bn^{2/3}. \quad (32)$$

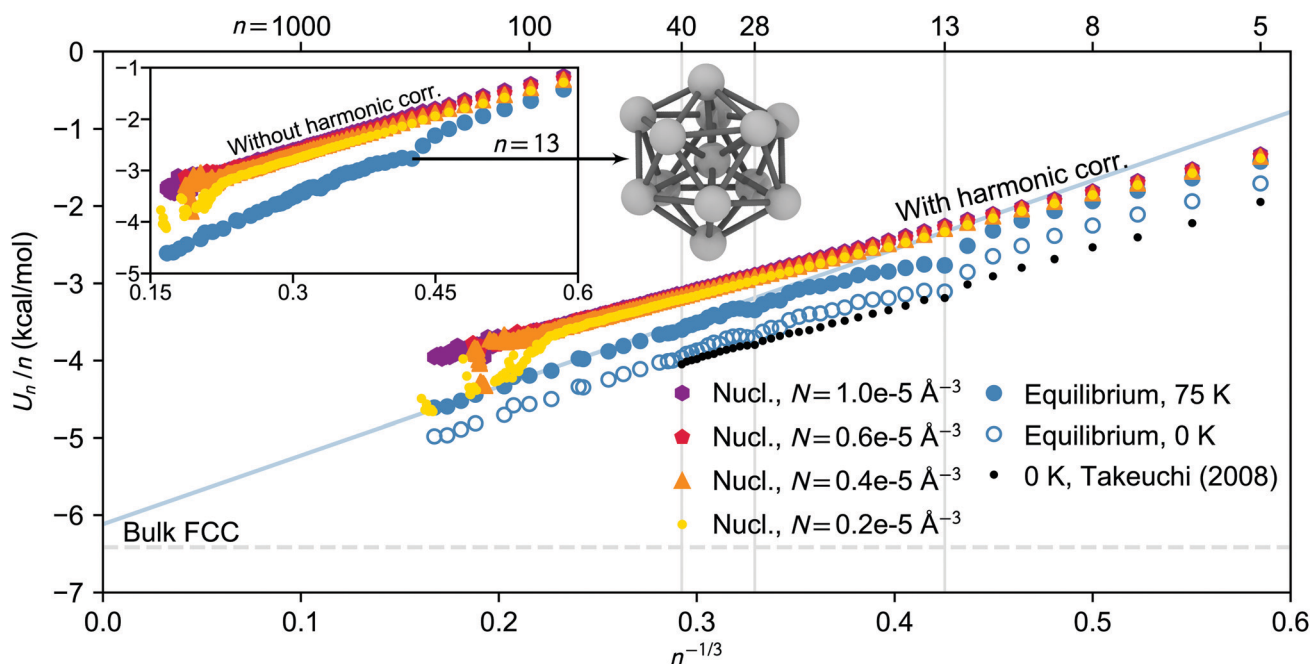
Interestingly, the term proportional to  $n^{1/3}$  is negligible here, in contrast to the case of the formation free energy.

The surface energies, or surface enthalpies, retrieved from NVE results using the liquid-drop model given in eqn (32) are not directly comparable to the surface tension, *i.e.*, surface free energy, which also includes surface entropy. In fact, molleman and Hiemstra<sup>53</sup> recently found Tolman lengths of different magnitudes and even different signs for surface free energy and surface enthalpy of water. Using the liquid-drop radius model, the surface energy can be estimated from eqn (32) as

$$\Delta H_{\text{surf}} = B/A_1. \quad (33)$$

The obtained  $\Delta H_{\text{surf}}$  is about 74 mN m<sup>−1</sup> for all three studied temperatures, and is quite close to the value of 77 mN m<sup>−1</sup>, calculated from the surface tension relation of Quinn<sup>44</sup> as  $\Delta H_{\text{surf}} = \sigma - T(\partial\sigma/\partial T)$ .

While the potential energies of the nucleating clusters show an almost perfectly linear dependence on  $n^{-1/3}$ , the equilibrated cluster potential energies manifest an irregular size dependency in Fig. 6. At the studied temperatures, the face centred cubic (FCC) phase is the stable crystallographic structure of bulk CO<sub>2</sub>. Yet, the energetically most favorable structure of small clusters can be very different due to the large fraction of under-coordinated molecules at the surface. In case of small clusters of simple liquids such as CO<sub>2</sub>, it has been widely reported that an icosahedral packing is preferred over a cubo-octahedral



**Fig. 6** Potential energy per molecule as a function of  $n^{-1/3}$  for clusters isolated from nucleation simulation carried out at 75 K. The energies of the nucleating clusters at different vapour densities (from NVE simulations) are shown as colourful dots, triangles, pentagons and hexagons, and the equilibrated cluster energies (from NVT simulations) are indicated as filled blue circles. In the inset the energies of the nucleating clusters are given without harmonic correction, while the energies in the main figure are harmonically scaled to 75 K (the blue solid line shows the extrapolated  $n^{-1/3}$  dependence for the large equilibrated clusters). The bulk FCC cohesive energy value at 75 K is shown by the horizontal dashed line. The open circles illustrate the equilibrated potential energies harmonically scaled to 0 K, and the black dots are the geometry optimized global minimum energies obtained by Takeuchi.<sup>52</sup> In addition, the equilibrium structure of the icosahedral 13-cluster is illustrated (the oxygen atoms are omitted) and the exceptionally stable cluster sizes (13, 28 and 40) are marked by vertical gray lines.



shape.<sup>52,54,55</sup> Despite these structural ambiguities, we have compared the potential energy extrapolated linearly to  $n \rightarrow \infty$  from the largest clusters in the simulation, with cohesive energies of bulk FCC systems at equal temperatures. The extrapolated values are relatively close to the bulk value even though the largest clusters analysed contain only about 200 molecules, and clusters this small may still have mixed structures deviating from perfect FCC. The surface enthalpies are 109 and 104 mN m<sup>-1</sup> at 75 and 90 K, respectively, which surprisingly correspond exactly to the values obtained from the parameterisation by Lielmezs and Herrick.<sup>3,45</sup> Interestingly, for both *NVE* and *NVT* simulated clusters, the ratio of the regression coefficients follows rather well the empirical relation<sup>56</sup>  $B/A \approx -2/3$ .

In the first part of this series, we only analysed the structural differences between nucleating and equilibrated clusters of size  $n > 10$  by calculating the carbon-carbon radial distribution functions and visualising the cluster geometries, as shown in Fig. 11 in ref. 3.

The potential energies presented in Fig. 6, as well as Fig. S5 and S6 (ESI†), confirm the clear difference between nucleating and equilibrated clusters for  $n \gtrsim 10$  at temperatures of 75 and 90 K. However, nucleating clusters are energetically comparable with the equilibrated ones for  $n < 10$ , due to the large fraction of under-coordinated and therefore loosely bound molecules. At the highest studied temperature of 105 K, the discrepancy in energy between nucleating and equilibrated clusters only appears at sizes  $n \gtrsim 60$ .

We observe that the largest clusters considered in this analysis,  $n \gtrsim 100$ , may undergo a phase transition in the nucleation simulation when both temperature and vapour density are low enough ( $T = 75$  K and  $N = 0.2 \times 10^{-5}$  Å<sup>-3</sup>). Even though the relative amount of carrier gas is about equal in all simulations, at low density the clusters accumulate less latent heat and have more time for relaxation between growth events, enabling the transition. Indeed, in the first part of this series,<sup>3</sup> Fourier transform infrared spectroscopy identified that the large, nano-scale CO<sub>2</sub> particles had a cubic crystal structure. Furthermore, Ramos *et al.*<sup>57</sup> showed that the Raman spectra of growing CO<sub>2</sub> clusters in a free-jet expansion indicate a cross-over size  $n \approx 150 \dots 400$  between liquid and solid phases at very similar conditions as discussed here.

We have compared our equilibrium cluster energies with values by Takeuchi,<sup>52</sup> who calculated the global minimum (0 K) energies for CO<sub>2</sub> clusters  $n = 4 \dots 40$  using the MOM potential. This model uses five partial charge sites, but the LJ parameters of MOM are less than 5% smaller than the corresponding TraPPE values.<sup>58</sup> To our knowledge, this is the only other study systematically probing the structures and energies of CO<sub>2</sub> clusters  $n > 13$ . After scaling the potential energies obtained in the *NVT* simulations to 0 K using the harmonic approximation, the scaled potential energies are in relatively good agreement with the global minimum values even though the force fields differ slightly. However, the scaled potential energies are still systematically higher than the reported global minimum values, which is partly due to the ensemble averaging over multiple minima. The difference in energy may also arise from thermal fluctuations in the configuration and the

vibrational anharmonicities which are usually more pronounced for loosely bound systems. Indeed, the difference is larger at higher temperatures (see Fig. S5 and S6, ESI†). Based on the Lindemann index analysis, the thermal motion of the smallest equilibrated clusters at 75 K is comparable to the thermal motion of the nucleating clusters of the same size (see ESI† and Fig. S7 for details). This supports the assumption that the equilibrium phase of small pre-icosahedral (*i.e.*  $n < 13$ ) clusters can be characterized as liquid-like.

The equilibrium simulations at 75 and 90 K identify enhanced stability of 13, 28 and 40-clusters, a result that is also found by Takeuchi<sup>52</sup> (see Fig. S8, ESI† for more detail), and consequently for these sizes the difference between average equilibrium potential energies and global minimum values is very small (less than 3% at 75 K). In many studies of solid clusters, “magic numbers” such as 13, 55, 147, ... are of particular interest as they correspond to very tightly bound and highly symmetric icosahedral shapes.<sup>55</sup> As mentioned before, at 105 K the icosahedral shape and configurational energy minimum of the 13-, or 55-clusters is not observed during the *NVT* simulation, where all clusters below size  $\sim 60$  are found to be liquid-like. This is in line with results by Maillet *et al.*,<sup>54,59</sup> who observed that using the MOM potential, both 13- and 55-clusters undergo a phase transition between 90 and 100 K.

In summary, the nucleating clusters are energetically and structurally very different from the equilibrium, or global minima configurations. Despite the strong undercooling, the nucleating clusters are more liquid-like, and should be treated accordingly when predicting nucleation rates at these conditions.

## 4 Simulation-based improvement on CNT

In Section 3.1.1, we found reasonable agreement between theory and atomistic simulations by importing the nucleation barrier height obtained from the MD simulation into the theoretical kinetic equation of nucleation. The validity of the kinetic scheme of nucleation has been also recently reported by Halonen *et al.*<sup>4</sup> and Ayuba *et al.*<sup>5</sup> In addition, in Sections 3.1.1 and 3.2.3, we have demonstrated that the nucleating CO<sub>2</sub> clusters are liquid-like and a radius-related term should be taken into account while applying CNT-like theory to predict homogeneous nucleation. Here we present an improvement to the classical theory, in similar fashion as Tanaka *et al.*,<sup>22</sup> by expressing the critical radius and the  $Q$  parameter in the Zeldovich factor in terms of their classical counterparts.

If the requirement for self-consistency is followed (*i.e.*  $\Delta W_1^{\text{eq}} = 0$ ), the linear behaviour of  $\Delta W_n^{\text{eq}}/A_n$  seen from the lower panel of Fig. 3 would imply that the liquid-drop model coefficient  $C = -A_1\sigma$ , which translates into a Tolman-like constant  $\delta_T = r_1/2$ . Applying this result, we can introduce a simulation-based theory (ST) which states that the critical cluster size can be written as

$$r_{\text{ST}}^* = \frac{r_{\text{CNT}}^*}{2} \left( 1 + \sqrt{1 - \frac{2r_1}{r_{\text{CNT}}^*}} \right), \quad (34)$$



and the  $Q$  parameter in the Zeldovich factor (see eqn (13)) becomes

$$Q_{\text{ST}} = Q_{\text{CNT}} \left( \frac{r_{\text{CNT}}^*}{r_{\text{ST}}^*} \right) \left( 1 - \frac{r_1}{r_{\text{ST}}^*} \right). \quad (35)$$

Unlike in the CNT formalism, eqn (34) internally restricts the critical size from being smaller than the monomer radius.

It would be imprudent to claim that the Tolman length for pure  $\text{CO}_2$  clusters is positive and its value to be  $r_1/2$  (actually, it is recently estimated to be negative<sup>35</sup>). The origin of the  $n^{1/3}$  term is not necessarily only due to the effect of curvature on surface tension, but other physical effects with the same functional form might also be present. As discussed in Section 3.2.1, the simulated cluster radii are related to the liquid-drop radii by eqn (28). This deviation from the liquid-drop radii gives a rather clear indication of a non-Tolman-like effect. As both the surface area ( $A_n = 4\pi r_n^2$ ) and the “Tolman” term ( $2\delta_{\text{T}}/r_n$ ) depend on the cluster radius, the correction to the cluster radius introduces an additional term in the formation free energies:

$$\Delta W_n^{\text{eq}} \approx A_1 \sigma \left( n^{2/3} - \left( 2a + \frac{2\delta_{\text{T}}}{r_1} \right) n^{1/3} \right). \quad (36)$$

Here, a small  $n$ -independent term is neglected. This is justified, as when the self-consistent theory is applied to calculate the nucleation rates, this term cancels out. According to the RDF calculations, roughly half of the coefficient  $C$  in eqn (1) is due to the size correction, as  $a \approx 0.25$ .

We have also used the effective planar surface tensions extrapolated from Fig. 3 to calculate the theoretical estimates for nucleation rates in the simulated systems using eqn (27), (34) and (35), instead of eqn (9), (11) and (13), respectively. As these surface tensions are obtained from non-equilibrium simulations, the non-isothermal correction should not be applied here.

The nucleation rates from simulation and both classical (CNT) and simulation-based (ST) theories are reported in Table 2 and illustrated in Fig. 7. The agreement between simulation and ST is perfect, whereas the values from CNT are several orders of magnitude off. The critical cluster sizes are reported in Table 3. Again, the theoretical values from ST are close, but still smaller, than the estimated threshold values used to calculate the nucleation rates from direct MD simulation data. The agreement between critical cluster size  $n_{\text{ST}}^*$  and the locations of the top of the nucleation barriers shown Fig. 3 is also quite good, whereas CNT overpredicts the critical size by about a factor of two. The critical cluster size obtained from the first nucleation theorem is systematically smaller than observed in the simulation, this is in line with findings of wedekind and Reguera,<sup>60</sup> and Napari *et al.*,<sup>61</sup> who used alternative simulation methods to study the nucleation in super-saturated Lennard-Jones vapour.

It may seem at first glance that the reason for the excellent agreement between simulation and ST predictions is that the nucleation rate estimated directly from the nucleation simulations is inserted in eqn (26) to compute the formation free energies. However, the theoretical prediction does not strongly depend on the value of  $J$ . For example, the position of the top of the free energy barrier, the linear behaviour shown in the

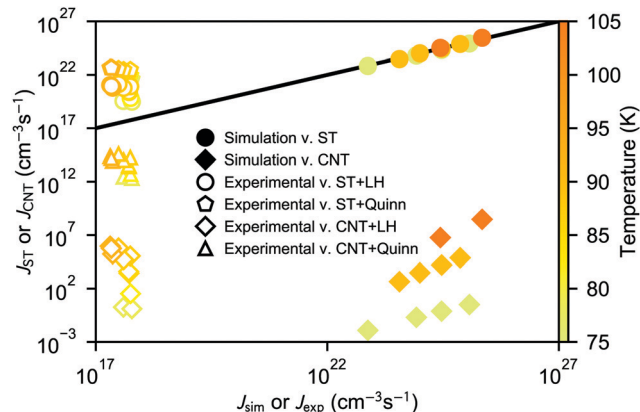


Fig. 7 Comparison between theoretical nucleation rates and those observed in simulation and experiment. For the comparison with simulation, full circles and diamonds are used for simulation-based theory (ST) and self-consistent version of classical theory (CNT), respectively. For the comparison with experiment, open circles, pentagons, diamonds, and triangles, are used to represent the combinations of ST, and CNT, with surface tension parameterizations by Lielmezs–Herrick (LH), or Quinn, respectively. The temperature is indicated by the symbol colour. The black line indicates perfect agreement.

bottom panels of Fig. 3, and the extrapolated surface tension, remain the same when  $J$  is chosen to be several orders of magnitude higher. However, such a radical change of  $J$  alters the shape of the free energy barrier for over-critical clusters. The perfect agreement between ST predictions and simulations is due to the fact that both the size dependence of the formation free energy (*i.e.*, the thermodynamics), as well as the kinetic scheme are well described.

The theoretical nucleation rates corresponding to the experimental conditions of Dingilian *et al.*<sup>3</sup> are also presented in Fig. 7. The thermophysical properties of liquid  $\text{CO}_2$  as well as surface tension parameterisations by Quinn, and Lielmezs and Herrick, needed to calculate the nucleation rates are given in Table S1 (ESI†). Here the non-isothermal correction is included, and the value of  $q$  is estimated from the cluster potential energies (see ESI†). The nucleation rate values from ST are a couple orders of magnitude higher than the experimental rates, and using the Lielmezs–Herrick relation for surface tension yields a slightly better agreement than using the Quinn relation. The CNT predictions using the Quinn relation are slightly lower than the experimental rates, but still quite close, whereas using CNT with the Lielmezs–Herrick relation underpredicts the experimental rates by over ten orders of magnitude.

While the uncertainty in surface tension of small, severely undercooled liquid clusters is the main concern in the context of this work, other sources of error exist as well. Some additional terms might still be missing in the liquid-drop model, other than the ones related to radius. Accounting for the effect of vapour–cluster interactions in highly supersaturated systems can increase the nucleation barrier,<sup>4,62</sup> and this can have a substantial effect on the results. Also the effect of long-range interactions and possible non-accommodation of collision rates are neglected in the classical view of nucleation.<sup>28–31</sup>



## 5 Conclusions

We have carried out large scale molecular dynamics simulations of the homogeneous nucleation of CO<sub>2</sub> in Ar carrier gas in a similar range of temperatures (75 to 105 K) as in the supersonic nozzle experiments in the first part of this series.<sup>3</sup> We calculated nucleation rates using the threshold method for ten different combinations of temperature and vapour pressure and derived the respective cluster formation free energies from the obtained cluster distributions. Analysis of the size dependence of the free energies using a liquid-drop model clearly indicated the need to incorporate an additional Tolman-like term  $\propto n^{1/3}$  into the free energy expression, as proposed by Tanaka *et al.*<sup>22</sup> for Lennard-Jones fluid. Using this finding, we were able to formulate a simulation-based improvement on classical nucleation theory, leading to a highly accurate prediction of the simulated nucleation rates. Here, values of the planar surface tension obtained from an extrapolation of the size dependence of the equilibrium formation free energies were used. At these deeply undercooled temperatures, the surface tension values were significantly higher compared to extrapolated values using parameterizations for liquid CO<sub>2</sub>. The simulation-based theory did not predict the experimental nucleation rates as accurately as the simulated nucleation rates. However, the difference was only 1–5 orders of magnitude, while the predictions of self-consistent classical nucleation theory were off by 3–18 orders of magnitude, indicating that the simulation-based theory offers a significant improvement.

As demonstrated through the thermophysical properties retrieved from the MD simulations, the standard Szilard–Farkas kinetic scheme of nucleation describes the simulated homogeneous nucleation of CO<sub>2</sub> rather well. As already discussed in the first part of this series, the major problem in theoretically estimating the nucleation rate of extremely undercooled CO<sub>2</sub> is the inaccuracy of the available surface tension parameterisation. The experiments are carried out over 100 K below the triple point, conditions in which liquid is a highly unstable phase. The analysis of the TraPPE–CO<sub>2</sub> clusters shows considerably higher planar surface tension than for liquid CO<sub>2</sub>. On the other hand, the obtained surface enthalpies of liquid clusters are in good agreement with the surface tension relation of Quinn.<sup>44</sup>

From the nucleation simulations we isolated clusters and studied their structural and energetic properties. Due to insufficient thermalization of the nucleating clusters through collisions with the carrier gas, the clusters accumulate a significant amount of latent heat from individual monomer captures. Based on the analysis of the kinetic energy, we observed cluster temperatures systematically above the temperature of the carrier gas. The energetic analysis revealed that during the nucleation stage, clusters containing fewer than 100 molecules were liquid-like at all studied temperatures and pressures. While the majority of these clusters undergoes a phase transition to the solid state upon equilibration, some of the smallest clusters remain liquid-like despite the strong undercooling. Over a century ago, Ostwald<sup>63</sup> observed in the context of crystal growth from solution that the phase transition occurred through intermediate metastable phases

until the most stable phase was reached. Here we have demonstrated that this empirical step rule is equally valid for clusters in the gas phase in extremely undercooled conditions, where the stable cluster structure is crystalline.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

This work was completed with the financial support of the ERC grant no. 692891-DAMOCLES, the University of Helsinki, Faculty of Science ATMATH project, and the National Science Foundation under grant number CBET-1511498. Supercomputing resources were provided by the CSC-IT Center for Science, Ltd, Finland, and the Finnish Grid and Cloud Infrastructure (urn:nbn:fi:research-infras-2016072533) at the University of Helsinki. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

## References

- 1 J. Gibbins and H. Chalmers, *Energy Policy*, 2008, **36**, 4317–4322.
- 2 A. Määttänen, F. Montmessin, B. Gondet, F. Scholten, H. Hoffmann, F. González-Galindo, A. Spiga, F. Forget, E. Hauber and G. Neukum, *et al.*, *Icarus*, 2010, **209**, 452–469.
- 3 K. K. Dingilian, R. Halonen, V. Tikkanen, B. Reischl, H. Vehkamäki and B. E. Wyslouzil, *Phys. Chem. Chem. Phys.*, 2020, **22**, 19282–19298.
- 4 R. Halonen, E. Zapadinsky and H. Vehkamäki, *J. Chem. Phys.*, 2018, **148**, 164508.
- 5 S. Ayuba, D. Suh, K. Nomura, T. Ebisuzaki and K. Yasuoka, *J. Chem. Phys.*, 2018, **149**, 044504.
- 6 W. G. Courtney, *J. Chem. Phys.*, 1961, **35**, 2249–2250.
- 7 S. L. Girshick and C.-P. Chiu, *J. Chem. Phys.*, 1990, **93**, 1273–1277.
- 8 H. Reiss, W. K. Kegel and J. I. Katz, *Phys. Rev. Lett.*, 1997, **78**, 4506–4509.
- 9 A. Dillmann and G. Meier, *J. Chem. Phys.*, 1991, **94**, 3872–3884.
- 10 D. Reguera and J. Rubí, *J. Chem. Phys.*, 2001, **115**, 7100–7106.
- 11 R. McGraw and A. Laaksonen, *Phys. Rev. Lett.*, 1996, **76**, 2754.
- 12 D. Reguera and H. Reiss, *Phys. Rev. Lett.*, 2004, **93**, 165701.
- 13 G. K. Schenter, S. M. Kathmann and B. C. Garrett, *Phys. Rev. Lett.*, 1999, **82**, 3484–3487.
- 14 R. C. Tolman, *J. Chem. Phys.*, 1949, **17**, 333–337.
- 15 J. Feder, K. Russell, J. Lothe and G. Pound, *Adv. Phys.*, 1966, **15**, 111–178.





- 16 J. Merikanto, E. Zapadinsky, A. Lauri and H. Vehkamäki, *Phys. Rev. Lett.*, 2007, **98**, 145702.
- 17 J. Elm, J. Kubečka, V. Besel, M. J. Jääskeläinen, R. Halonen, T. Kurtén and H. Vehkamäki, *J. Aerosol Sci.*, 2020, **149**, 105621.
- 18 B. Chen, J. I. Siepmann, K. J. Oh and M. L. Klein, *J. Chem. Phys.*, 2001, **115**, 10903–10913.
- 19 K. Yasuoka and M. Matsumoto, *J. Chem. Phys.*, 1998, **109**, 8451–8462.
- 20 K. Yasuoka and M. Matsumoto, *J. Chem. Phys.*, 1998, **109**, 8463–8470.
- 21 J. Wedekind, D. Reguera and R. Strey, *J. Chem. Phys.*, 2007, **127**, 064501.
- 22 K. K. Tanaka, J. Diemand, R. Angélil and H. Tanaka, *J. Chem. Phys.*, 2014, **140**, 194310.
- 23 R. Angélil, J. Diemand, K. K. Tanaka and H. Tanaka, *J. Chem. Phys.*, 2015, **143**, 064507.
- 24 M. Horsch, J. Vrabec, M. Bernreuther, S. Grottel, G. Reina, A. Wix, K. Schaber and H. Hasse, *J. Chem. Phys.*, 2008, **128**, 164510.
- 25 M. Horsch, Z. Lin, T. Windmann, H. Hasse and J. Vrabec, *Atmos. Res.*, 2011, **101**, 519–526.
- 26 F. F. Abraham, *Homogeneous Nucleation: The Pretransition Theory of Vapor Condensation*, Academic Press, New York, 1974.
- 27 H. Vehkamäki, *Classical nucleation theory in multicomponent systems*, Springer Science & Business Media, 2006.
- 28 R. Halonen, E. Zapadinsky, T. Kurtén, H. Vehkamäki and B. Reischl, *Atmos. Chem. Phys.*, 2019, **19**, 13355–13366.
- 29 E. Goudeli, J. Lee and C. J. Hogan Jr, *J. Aerosol Sci.*, 2020, 105558.
- 30 O. H. Vasilev and H. Reiss, *J. Chem. Phys.*, 1996, **105**, 2946–2947.
- 31 O. H. Vasilev and H. Reiss, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1996, **54**, 3950–3954.
- 32 D. Saltz, *J. Chem. Phys.*, 1994, **101**, 6038–6051.
- 33 Y. A. Lei, T. Bykov, S. Yoo and X. C. Zeng, *J. Am. Chem. Soc.*, 2005, **127**, 15346–15347.
- 34 N. Bruot and F. Caupin, *Phys. Rev. Lett.*, 2016, **116**, 056102.
- 35 P. Rehner, A. Aasen and Ø. Wilhelmsen, *J. Chem. Phys.*, 2019, **151**, 244710.
- 36 F. M. Kuni, A. P. Grinin and A. K. Shchekin, *Physica A*, 1998, **252**, 67–84.
- 37 J. Barrett, *J. Phys. A: Math. Gen.*, 1994, **27**, 5053–5068.
- 38 J. J. Potoff and J. I. Siepmann, *AIChE J.*, 2001, **47**, 1676–1682.
- 39 B. Chen, J. J. Potoff and J. I. Siepmann, *J. Phys. Chem. B*, 2001, **105**, 3093–3104.
- 40 J. O. Hirschfelder, C. F. Curtiss and R. B. Bird, *Molecular theory of gases and liquids*, Wiley, New York, 1954.
- 41 S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1–19.
- 42 F. H. Stillinger Jr, *J. Chem. Phys.*, 1963, **38**, 1486–1494.
- 43 D. Kashchiev, *J. Chem. Phys.*, 1982, **76**, 5098–5102.
- 44 E. L. Quinn, *J. Am. Chem. Soc.*, 1927, **49**, 2704–2711.
- 45 J. Lielmezs and T. Herrick, *Chem. Eng. J.*, 1986, **32**, 165–169.
- 46 R. Span and W. Wagner, *J. Phys. Chem. Ref. Data*, 1996, **25**, 1509–1596.
- 47 J. Ter Horst, D. Bedeaux and S. Kjelstrup, *J. Chem. Phys.*, 2011, **134**, 054703.
- 48 B. E. Wyslouzil and J. H. Seinfeld, *J. Chem. Phys.*, 1992, **97**, 2661–2670.
- 49 J. Barrett, C. Clement and I. Ford, *J. Phys. A: Math. Gen.*, 1993, **26**, 529–548.
- 50 J. C. Barrett, *J. Chem. Phys.*, 2011, **135**, 096101.
- 51 J. W. Schmelzer, G. S. Boltachev and A. S. Abyzov, *J. Chem. Phys.*, 2013, **139**, 034702.
- 52 H. Takeuchi, *J. Phys. Chem. A*, 2008, **112**, 7492–7497.
- 53 B. Molleman and T. Hiemstra, *Phys. Chem. Chem. Phys.*, 2018, **20**, 20575–20587.
- 54 J.-B. Maillet, A. Boutin, S. Buttefy, F. Calvo and A. H. Fuchs, *J. Chem. Phys.*, 1998, **109**, 329–337.
- 55 F. Baletto and R. Ferrando, *Rev. Mod. Phys.*, 2005, **77**, 371.
- 56 K. Hansen, *Statistical physics of nanoparticles in the gas phase*, Springer International Publishing, 2nd edn, 2018.
- 57 A. Ramos, J. Fernández, G. Tejeda and S. Montero, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2005, **72**, 053204.
- 58 C. Murthy, S. O'Shea and I. McDonald, *Mol. Phys.*, 1983, **50**, 531–541.
- 59 J.-B. Maillet, A. Boutin and A. H. Fuchs, *J. Chem. Phys.*, 1999, **111**, 2095–2102.
- 60 J. Wedekind and D. Reguera, *J. Chem. Phys.*, 2007, **127**, 154516.
- 61 I. Napari, J. Julin and H. Vehkamäki, *J. Chem. Phys.*, 2009, **131**, 244511.
- 62 K. J. Oh and X. C. Zeng, *J. Chem. Phys.*, 2000, **112**, 294–300.
- 63 W. Ostwald, *Z. Phys. Chem.*, 1897, **22**, 289–330.

