

# Impact of Quantum Chemistry Parameter Choices and Cluster Distribution Model Settings on Modeled Atmospheric Particle Formation Rates

Vitus Besel,<sup>\*,†</sup> Jakub Kubečka,<sup>†</sup> Theo Kurtén,<sup>‡</sup> and Hanna Vehkamäki<sup>†</sup>

<sup>†</sup>*University of Helsinki, Physicum, Gustaf Hällströmin Katu 2, 00560 Helsinki*

<sup>‡</sup>*University of Helsinki, Chemicum, A. I. Virtasen aukio 1, 00560 Helsinki*

E-mail: vitus.besel@helsinki.fi

Phone: +49 175 8819882

## Abstract

We tested the influence of various parameters on the new particle formation rate predicted for the sulfuric - acid ammonia system using quantum chemistry and cluster distribution dynamics simulations, in our case ACDC (Atmospheric Cluster Dynamics Code). We found that consistent consideration of the rotational symmetry number of monomers (sulfuric acid and ammonia molecules, and bisulfate and ammonium ions) leads to a significant rise in the predicted particle formation rate, whereas inclusion of the rotational symmetry number of the clusters only changes the results slightly, and only in conditions where charged clusters dominate the particle formation rate. This is because most of the clusters stable enough to participate in new particle formation have a rotational symmetry number of one, and the few exceptions to this rule are positively charged. In contrast, the application of the quasi-harmonic correction for low-frequency

vibrational modes tends to generally decrease predicted new particle formation rates, and also significantly alters the slope of the formation rate curve plotted against the sulfuric acid concentration which is a typical convention in atmospheric aerosol science. The impact of the maximum size of the clusters explicitly included in the simulations depends on the simulated conditions. The errors arising from a limited set of clusters are higher for higher evaporation rates, and thus tend to increase with temperature. Similarly, the errors tend to be higher for lower vapor concentrations. The boundary conditions for outgrowing clusters (which clusters are counted as formed particles) have only a small influence on the results, provided that the definition is chemically reasonable, and that the set of simulated clusters is sufficiently large. A comparison with data from the CLOUD (Cosmics Leaving OUtdoor Droplets) chamber and a cluster distribution dynamics model using older quantum chemistry input data shows improved agreement when using our new input data, and the proposed combination of symmetry and quasi-harmonic corrections.

## Introduction

Aerosol particles can be directly emitted to the atmosphere, but also form from gas phase molecules in the atmosphere (new particle formation)<sup>1-3</sup>. Aerosols play a significant role in the health threat posed by poor air quality since particles below diameters of 2.5  $\mu\text{m}$  can enter the respiratory system and be deposited in the lungs, causing a number of diseases<sup>4-6</sup>. Furthermore, aerosols play a crucial role in weather and climate as cloud condensation nuclei<sup>7</sup>. The effect of aerosols via clouds is one of the largest uncertainties in the current models of climate change<sup>8</sup>, and it is estimated that 40% to 70% of all cloud condensation nuclei are a product of new particle formation<sup>9,10</sup>. New particle formation is commonly referred to as "nucleation": this is technically incorrect as the process may not always involve energy barriers. A large number of studies have been devoted to unraveling the precise molecular-level mechanisms responsible for the first steps of atmospheric aerosol formation<sup>11-17</sup>. However,

due to the large number of potentially participating compounds, the process is exceedingly complex, and both experimental and modeling tools still require development in order to reach a quantitative understanding of atmospheric new particle formation.

Sulfuric acid and ammonia have been identified to be key players in the formation of aerosols<sup>16</sup> and have been studied extensively in the past both experimentally and theoretically. The main reason for this is that the equilibrium vapour pressure of sulfuric acid in the presence of bases, of which ammonia is the most abundant in the atmosphere, is very close to zero<sup>18</sup>. Furthermore, in contrast to for example most low-volatility organic compounds, sulfuric acid forms practically in a single oxidation step from the very volatile precursor SO<sub>2</sub>. The subsequent reaction steps involve high-concentration reactants, H<sub>2</sub>O and O<sub>2</sub>, and are thus essentially instantaneous. Mixing ratios of sulfuric acid in the lower atmosphere very seldom exceed ten parts per trillion (i.e. about  $2 \times 10^8$  molecules per cm<sup>3</sup>) even in highly polluted conditions, and are often much lower<sup>19,20</sup>. Ammonia mixing ratios, on the other hand, can reach tens of parts per billion (on the order of  $10^{11}$  molecules per cm<sup>3</sup>)<sup>21</sup>.

The process of single molecules forming new particles in the atmosphere is difficult to measure experimentally. Molecule and cluster concentrations are low, and therefore they have to be measured with mass spectrometry (MS). In order for molecules to be detectable by MS they have to be charged, but measurements have shown that the new particle formation in the atmospheric boundary layer is usually dominated by neutral particles<sup>22</sup> consisting of multiple other compounds in addition to sulfuric acid and ammonia. To complement experiments, computational studies are needed, and in order to capture all relevant factors, such as proton transfers, quantum chemistry tools are necessary: Density functional theory (DFT) is able to give accurate information about cluster structures<sup>23-25</sup>, and recently high level of theory wavefunction methods have been employed to study binding energies of atmospheric clusters<sup>26,27</sup>. The Atmospheric Cluster Dynamics Code (ACDC)<sup>28,29</sup> is a model that simulates molecular cluster formation for a given set of compounds, and has been used in

many atmospheric studies<sup>30-34</sup>. ACDC uses Gibbs free energies provided by quantum chemistry to calculate evaporation coefficients and collision rates from the kinetic gas theory to eventually compute formation rates of atmospheric clusters and trace the cluster formation pathways. The purpose of this study is to revisit sulfuric acid - ammonia clustering in the light of recent advances in the modeling schemes.<sup>27,35</sup> We use a newly sampled, extended set of molecular cluster structures for the sulfuric acid - ammonia system and review the effect of some common approximations or assumptions on modeled particle formation rates. We also conducted a systematic study on the effect of different approximations used in computing thermodynamic properties and consideration of symmetry in quantum chemistry programs on the outcome of ACDC calculations. Table 1 in the Conclusions section gives a short summary of all quantum chemistry parameters and simulation settings examined and the respective findings. Moreover, we compare our modeled particle formation rates with experimental data from the CLOUD (Cosmics Leaving OUTdoor Droplets) chamber at CERN<sup>36</sup> and previous modeling of this experiment<sup>30</sup>.

## Methodology

### Molecular cluster structures

Initially, it is necessary to find energetically low-lying structures for the sulfuric acid - ammonia clusters. We proceeded as described by Kubečka et al.<sup>35</sup>:

1. We optimized bisulfate ion, ammonia, ammonium, *cis*- and *trans*- sulfuric acid structures and obtained their partial charges at a MP2/6-31++G(d,p)<sup>37-39</sup> level of theory. We used these partial charges together with Lennard-Jones parameters from the CHARMM (Chemistry at Harvard Molecular Mechanics) force field<sup>40,41</sup> to model the interaction at rigid monomeric units. Further, we utilized the ABCluster program<sup>42,43</sup> which is based on the artificial bee colony algorithm for the exploration of the whole configurational space of

each studied cluster. We configured the program so that it generates 10 000 local minima for each monomer combination fulfilling the charge constraints (e.g., 1  $\text{H}_2\text{SO}_4\text{-NH}_3$  cluster can be made of *cis*- $\text{H}_2\text{SO}_4 + \text{NH}_3$ , *trans*- $\text{H}_2\text{SO}_4 + \text{NH}_3$  or  $\text{HSO}_4^- + \text{NH}_4^+$ , for each of those 10 000 local minima were created). The program used 5 scout bees within 200 generations and optimizes generated cluster structures with molecular mechanics using the CHARMM force field, where structures of single molecules within the cluster are kept rigid.

2. The structures were optimized semi-empirically with the GFN-*x*TB program<sup>44</sup>.
3. Redundant structures and outliers were removed as described by Kubečka et al.<sup>35</sup> and the roughly 1000 remaining structures were passed to a density functional theory (DFT) calculation conducted with the Gaussian 16<sup>45</sup> program. We used the  $\omega\text{B97X-D}$  functional<sup>46</sup> and employed the 6-31++g\*\* basis set. These parameters have been extensively studied in previous publications and shown to result in sufficiently accurate structures and vibrational modes<sup>26,27</sup>.

First, we optimized the structures with *loose* convergence criteria and selected energetically lowest-lying minima plus a representative subset of all the structures<sup>47</sup> as described by Kubečka et al.<sup>35</sup>. Then, we optimized the selected structures with *verytight* convergence criteria.

4. We chose the ten energetically lowest lying structures and calculated the harmonic vibrational frequencies, making sure that no imaginary frequencies were encountered. Then we calculated enthalpies  $H$ , entropies  $S$  and Gibbs free energies  $G$  for a given temperature. Up to five lowest-lying structures were selected for further processing.

5. These finally selected structures were sent to single point DLPNO-CCSD(T) (domain-based local pair natural orbital - CCSD(T))<sup>48,49</sup> calculations with the augmented correlation consistent polarized valence triple  $\zeta$  basis set<sup>50</sup>. The approach combining DLPNO-CCSD(T) single point calculations with DFT frequency calculations has been shown to yield high accuracy<sup>26,27</sup>.

The configurational sampling procedure was conducted for all sulfuric acid - ammonia clusters except those that Olenius et al.<sup>34</sup> found not to be on the particle formation path. These clusters are very far from the "diagonal" (diagonal = clusters with the same number of sulfuric acid and ammonia monomers), or positively charged clusters with more sulfuric acid than ammonia, or negatively charged cluster with more ammonia than sulfuric acid molecules and they are deemed unstable by any reasonable computational method.

## Thermodynamic properties of the global minima

The evaporation coefficients used by ACDC depend exponentially on the Gibbs free energies computed by quantum chemistry<sup>28</sup> and, therefore, it is crucial to determine the free energies extremely accurately. However, there are assumptions in the calculation of the Gibbs free energy which can make a difference of several kcal/mol, but have not been accounted for properly in past studies. As previously discussed<sup>35,51</sup>, we focus on the global minimum in order to calculate the Gibbs free energy of formation of a cluster of a certain composition, rather than averaging over all found local minima. Accordingly, all the following calculations were done using the *trans*-sulfuric acid.

We investigated the effect of symmetry on the Gibbs free energy  $G$  and, consequently, on ACDC simulation results. Symmetry thresholds in many quantum chemistry programs are so tight that, the structures of especially clusters cannot be optimized with a numerical precision accurate enough to be identified as symmetric. Therefore, Gaussian or other quantum chemistry programs might not by default recognize the  $C_2$  symmetry of (*trans*)-sulfuric acid, whereas it does assign the  $C_{3v}$  point group for ammonia,  $T_d$  for ammonium and  $C_s$  for bisulfate. Previous studies used these defaults and thus included inconsistencies. The section "Symmetry" discusses the impact of symmetry considerations.

A common approach to calculate the rotational and vibrational contributions to thermody-

dynamic properties is using the Rigid Rotor-Harmonic Oscillator (RRHO) approximation, but the harmonic approximation often does not describe correctly the lowest energy "vibrational" modes which actually correspond to large amplitude intermolecular motions. Therefore, we recalculated the partition functions, and consequently the enthalpies and entropies, with the free-rotor model for low-lying frequencies as suggested by Grimme<sup>52</sup> and tested for molecular clusters by Myllys et al.<sup>27</sup>. We used the GOODVIBES python script<sup>53</sup> with a cut-off of 100 cm<sup>-1</sup> for this so called quasi-harmonic approximation. Throughout the article, we refer to the change introduced by the quasi-harmonic approach as the *quasi-harmonic correction*. Moreover, we apply the scaling factor of 0.9960 in order to account for the actual vibrational anharmonicity of the modes with wavenumbers above 100 cm<sup>-1</sup>, as suggested by Myllys et al.<sup>27</sup>.

We obtain the enthalpy  $H_{\text{DLPNO-CCSD(T),}i}$  for the cluster  $i$  resulting from the configurational sampling process by combining the electronic energy  $E_{\text{elec,DFT},i}$  and enthalpy  $H_{\text{DFT},i}$  of the DFT calculations and the DLPNO-CCSD(T) single point energy  $E_{\text{elec,DLPNO-CCSD(T),}i}$  as

$$H_{\text{DLPNO-CCSD(T),}i} = H_{\text{DFT},i} + E_{\text{elec,DLPNO-CCSD(T),}i} - E_{\text{elec,DFT},i}. \quad (1)$$

This approach combines the precision of DLPNO-CCSD(T) calculations and the speed of a DFT frequency analysis. Generally, using DFT structures and frequencies in combination with high-level wavefunction based energies is a broadly accepted approach in quantum chemistry<sup>54-56</sup>.

Subsequently, we calculate the enthalpy  $\Delta H$  and entropy  $\Delta S$  of cluster formation with respect to the monomers:

$$\Delta H_i = H_{\text{DLPNO-CCSD(T),}i} - \sum H_{\text{DLPNO-CCSD(T),monomer}} \quad (2)$$

and

$$\Delta S_i = S_i - \sum S_{\text{monomer}}. \quad (3)$$

The monomers are ammonia and sulfuric acid plus one bisulfate ion for negative clusters or plus one ammonium ion for positive clusters.

We correct  $\Delta G$ s for symmetry using a correction factor

$$C_{\text{sym},i} = R \cdot T \cdot \ln(\sigma_i), \quad (4)$$

$$\Delta G_i = G_i + C_{\text{sym},i} - \sum (G_{\text{monomer}} + C_{\text{sym,monomer}}). \quad (5)$$

where  $R$  is the molar gas constant,  $T$  is temperature, and  $\sigma$  is the rotational symmetry number of a cluster or molecule. The  $\Delta G$  given by Eq. (5) corresponds to the reference pressure at which the monomer and cluster free energies are computed, typically 1 atm. The reference pressure cancels out in the calculation of evaporation rates, and thus does not actually affect cluster distributions. For illustrative purposes, and by analogy to the free energy surfaces used in classical nucleation theory, so-called "actual  $\Delta G$ " values can be calculated by correcting for the actual partial pressures of the monomers, which are typically much lower than 1 atm.

Further, we test the influence of a variation of the set of simulated clusters, in other words the maximum size of clusters explicitly included in the simulation, for ACDC and compare our data to previously published results.

## Atmospheric Cluster Dynamics Code (ACDC)

The Atmospheric Cluster Dynamics Code (ACDC)<sup>28,29</sup> is a program which constructs and numerically solves the birth-death equations for a set of clusters. It calculates collision coefficients  $\beta$  from the kinetic gas theory<sup>57</sup> and evaporation coefficients  $\gamma$  from the Gibbs



free energy of formation as<sup>28</sup>

$$\gamma_{(i+j)\rightarrow i} = \beta_{ij} \cdot \frac{c_i^e c_j^e}{c_{i+j}^e} = \beta_{ij} \cdot c_{\text{ref}} \cdot \exp\left(\frac{\Delta G_{i+j} - \Delta G_i - \Delta G_j}{k_b T}\right) \quad (6)$$

where  $i$  and  $j$  are the daughter clusters,  $\beta_{ij}$  is the collision coefficient between clusters  $i$  and  $j$ ,  $c_i^e$  is the equilibrium concentration of cluster  $i$  and  $c_{\text{ref}} = p_{\text{ref}}/RT$  ( $\approx 2.5 \cdot 10^{23} \text{ cm}^{-3}$ ) is the concentration corresponding to the reference pressure for which the  $\Delta G$ s were calculated. We run ACDC with constant monomer concentrations, which in effect assumes that cluster formation does not constitute a significant sink for the vapors. This is the case for  $\text{H}_2\text{SO}_4$  and  $\text{NH}_3$  under most atmospheric conditions as well as in the CLOUD experiment. These monomers are allowed to collide and form clusters. Once clusters have formed, they can collide and evaporate. When a cluster that is not explicitly included in the simulation is formed, it evaporates immediately unless it forms a predefined outgrowing cluster deemed stable as shown in Fig. 1. Any collision that forms a cluster in the "outgrowing brown area" is counted towards the particle formation rate. In a time independent steady-state simulation, the particle formation rate is the number of all outgrowing clusters forming per second per volume [ $\text{cm}^{-3}\text{s}^{-1}$ ].

Outgrowing areas are determined by the direction in which we expect the clusters to keep growing if they leave the set of simulated clusters. This is estimated by the ratio of the collision rate  $R_{\text{col},i,x}$  of cluster  $i$  colliding with monomer  $x$  (either sulfuric acid or ammonia) to the evaporation rate  $R_{\text{evap},i}$  of cluster  $i$

$$\frac{R_{\text{col},i,x}}{R_{\text{evap},i}} = \frac{\beta[i][x]}{\gamma[i]} = \frac{\beta[x]}{\gamma} \quad (7)$$

If Eq. 7 yields a result  $\gg 1$  we assume the corresponding cluster  $i$  to be stable with respect to the monomer  $x$  within the simulation and we define  $i + x$  as an outgrowing cluster. The assumption is made that  $i + x$  is stable itself, which is valid because at the chosen cluster set size  $i + x$  has surpassed the critical cluster size and will continue growing (The

concept of critical cluster is explained in more detail in section "Choice of set of simulated clusters"). Additionally, the contributions of marginal outgrowing clusters is generally very small and has little influence on the simulation results as long as the main outgrowth routes are captured. As Eq. 7 is a function of the monomer concentration this "stability" can be different with respect to different monomers which is the reason why outgrowth by ammonia addition and by sulfuric acid addition are determined independently from each other. Also any cluster containing more sulfuric acid monomers and more ammonia monomers than cluster  $i$  is then in the outgrowing area. In practice, multiple of these areas can be defined. All outgrowing clusters have been chosen following this rule. The only exception to this is the  $(\text{HSO}_4^-)(\text{H}_2\text{SO}_4)_5(\text{NH}_3)_2$  cluster, where the definition would have led to the counterintuitive situation of having a "pocket" where this cluster is considered stable and outgrowing, whereas larger clusters such as  $(\text{HSO}_4^-)(\text{H}_2\text{SO}_4)_5(\text{NH}_3)_5$  would still be part of the cluster set with a chance to evaporate. Figure S4 in the appendix shows what this "pocket" looked like and the caption states differences introduced by including it. Eventually, it is important to note that in theory outgrowing cluster choice has to depend on the particular monomer concentration of each simulation. In practice, for the range at conditions chosen in the present simulations the set of clusters for which the value of Eq. 7 exceeds unity remains the same for all simulation runs.

The model includes parameters for explicit modeling of different types of external losses of the molecular clusters in the CLOUD (Cosmics Leaving Outdoor Droplets) chamber<sup>57</sup>. We consider wall losses, dilution losses, and account for the enhancement of loss rates arising from electrostatic interaction of charged clusters and molecules with the walls<sup>57</sup>. A charge enhancement factor of 3.3 has been determined for the CLOUD chamber<sup>57</sup>. We also account for coagulation losses with a constant coefficient of  $2.6 \cdot 10^{-3} \text{s}^{-1}$  derived from field measurement data<sup>58</sup> as no value has been reported for the CLOUD experiment. This inconsistency, however, has a minor impact on predicted particle formation rates as wall losses make a difference of two orders of magnitude opposed to a factor of 2 introduced by coagulation losses.

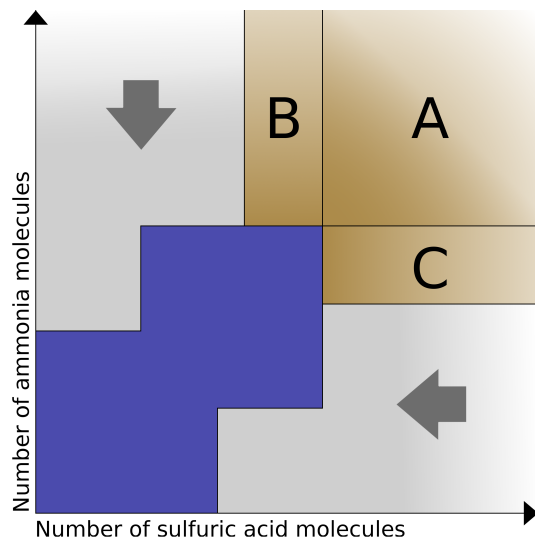


Figure 1: An illustration of cluster areas of the ACDC simulation: **Blue** is the set of simulated clusters, **brown** are possible areas of outgrowing clusters. If a cluster outside of these two areas (**grey**) is formed, it is brought back to the blue area by evaporating either single ammonia or sulfuric acid molecules. The outgrowing areas are defined by their lower-left corner and span an infinite region to the right and up. Possible areas are A, A+B (defined by the lower left corner of B), A+C (defined by the lower left corner of C) and A+B+C (defined by the lower left corner of B and C).

That coagulation losses are small compared to wall losses is also stated by Dunne et al.<sup>36</sup>. Our simulated system consists of clusters up to  $(\text{H}_2\text{SO}_4)_6(\text{NH}_3)_6$ ,  $(\text{H}_2\text{SO}_4)_6(\text{NH}_4^+)(\text{NH}_3)_5$ , and  $(\text{HSO}_4^-)(\text{H}_2\text{SO}_4)_5(\text{NH}_3)_6$  (*cf.* Fig. 2). We consider only clusters that have been shown to be relevant by Olenius et al.<sup>34</sup> and do not include hydration of clusters. The effect of hydration may have an impact on the overall performance of the model compared to the experimental results<sup>59</sup>, but this effect is outside the scope of this study. The model includes charging and recombination of clusters via generic charger ions  $\text{O}_2^-$  and  $\text{H}_3\text{O}^+$ , which transfer their charge upon collision.

We compared our ACDC simulations with experimental data from the CLOUD chamber<sup>36</sup>. We took all data points with temperature  $T \pm 1$  K for temperatures 292 K, 278 K, 248 K, 223 K, or 208 K. Since the experiment was aiming for an ionization rate of  $3 \text{ cm}^{-3}\text{s}^{-1}$ , we excluded all data points with an ionization rate outside of the interval  $2\text{--}4 \text{ cm}^{-3}\text{s}^{-1}$ . The sul-

furic acid concentration in the simulations ranged from  $10^5$  to  $10^9$   $\text{cm}^{-3}\text{s}^{-1}$  and the ammonia concentrations were 2, 10, 100, or 1000 ppt (this is a range of  $6.6 \cdot 10^7$   $\text{cm}^{-3}$  to  $2.5 \cdot 10^{10}$   $\text{cm}^{-3}$  depending on temperature). These concentrations were chosen in accordance with Kürten et al.<sup>30</sup> in order to stay within relevant concentrations for the CLOUD experiment and allow comparability. The temperature as well as the ionization rate in the simulation was set to the respective mean of the measured data points.

## Results and discussion

### Simulation set-up

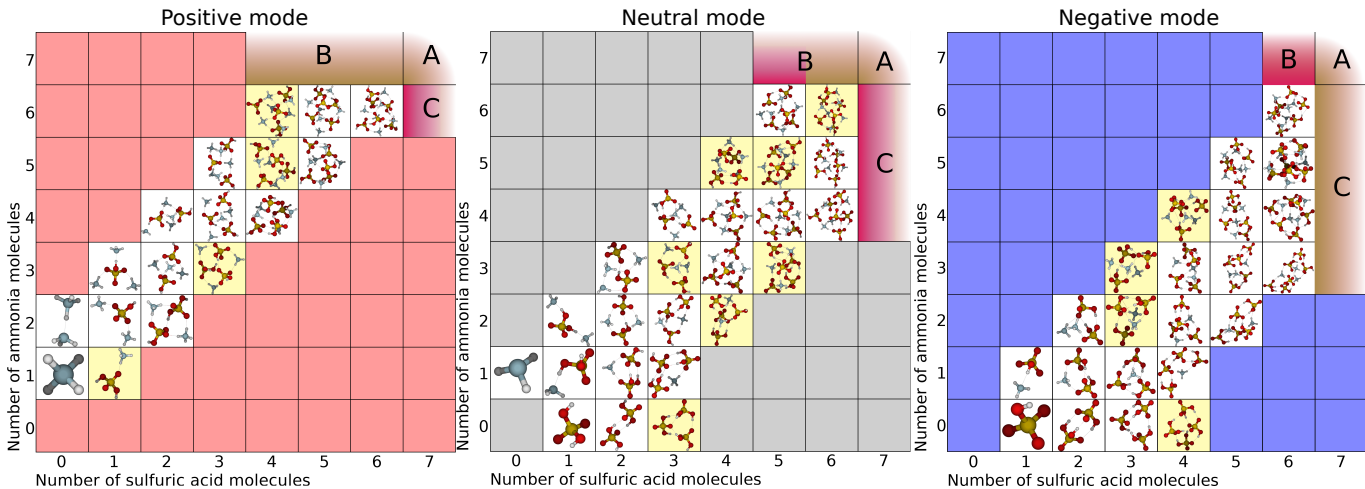


Figure 2: The global minima structures making up a set of simulated clusters. The yellow highlighted clusters are clusters for which the global minimum structure changes depending on whether we apply the quasi-harmonic correction by the GOODVIBES script or not. The brown is the area in which formed cluster are counted towards the particle formation rate in all simulations except the simulation for "Determination of boundary conditions for outgrowing cluster". For the "Determination of boundary conditions for outgrowing cluster" section the brown and the magenta area are outgrowing areas. This selection had to be made to obtain a understanding of the outgrowing cluster direction where the outgrowing areas A, B and C are kept independent.

Figure 2 shows the set of simulated clusters in ACDC. Each cluster type is represented by the global minimum structure, with respect to their Gibbs free energy at 298.15 K found in this study. As already pointed out, cluster types far from the diagonal in these diagrams have

been omitted as they play a minor role in cluster fluxes<sup>34</sup>. The structures and thermodynamic parameters corresponding to clusters in Fig. 2 can be found in the supporting information. Depending on whether we apply the quasi-harmonic correction, the structure representing the global free energy minimum may differ. Unless otherwise denoted, we use the global minima corresponding to the case with quasi-harmonic approximation. Generally, the global minimum structure may change with temperature. We checked all of our global minimum structures and found that it is the case for three of the molecular clusters once we are at 248.4 K or below. We updated the cluster set accordingly for these simulations. These three clusters are shown in the SI, Fig. S3.

## Determination of boundary conditions for outgrowing clusters

First, we want to estimate the impact of variations of boundary conditions for outgrowing clusters, i.e. the definition of outgrowing areas. We set up a series of simulations with successively extending outgrowing areas: brown and magenta areas A, A+B, A+C, and A+B+C, as shown in Fig. 2. We tested each area for 292 K and a ammonia concentration of 100 ppt, and for 223 K and a ammonia concentration of 10 ppt with sulfuric acid concentrations ranging from  $10^5$  to  $10^9$   $\text{cm}^{-3}$ . The resulting particle formation rates as a function of acid concentration are shown in Fig. 3. Outgrowth through area A produces the lowest rates because it cannot be reached by monomer addition, but only through collision with already formed clusters, e.g.,  $(\text{H}_2\text{SO}_4)(\text{NH}_3)$ .

The definition which only contains the outgrowing area A is thus a poor choice if monomer concentrations exceed cluster concentrations by several orders of magnitude. In all present cases, monomer concentrations are significantly higher than cluster concentrations and outgrowth occurs primarily through monomer addition. All other reasonable choices of boundary definitions show, especially in conditions of atmospheric relevance (roughly with particle formation larger than  $0.01$   $\text{cm}^{-3}\text{s}^{-1}$ ) negligibly small differences. The effect of including area B is larger than that of including area C as in most of the simulated conditions cluster

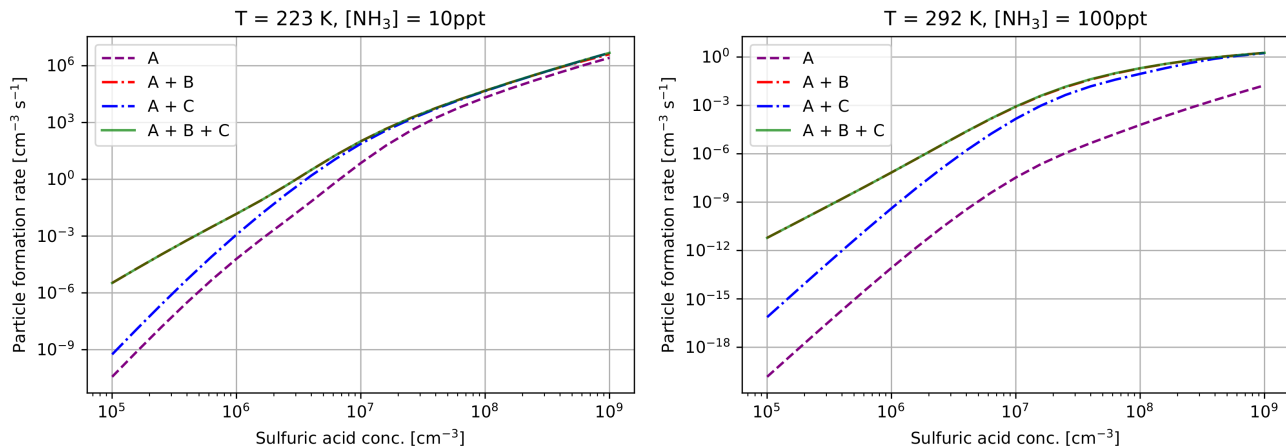


Figure 3: Particle formation rates from simulation runs with different declarations of outgrowing areas. The (A+B+C) line is overlapping the (A+B) line.

encounter ammonia molecules much more often than sulfuric acid molecules.

## Comparison to previous global minima

We compared particle formation rates calculated using our new global minima to rates calculated using previously found global minima. We extracted the previous minima from the Atmospheric Cluster Database (ACDB)<sup>24,60,61</sup> (November 2019). These data are based on quantum chemistry calculations, which generally do not consider symmetry, except for the ammonium molecule. We recalculated all the structures taken from ACDB with the DFT and DLPNO-CCSD(T) methods described in section 2.1. Not all molecular cluster types we sampled were available in the database, but also not all cluster types in the database have been sampled by us as explained at the end of section 2.1. Therefore, these comparison simulations were run using only these clusters which were available both in ACDB and in our resampling. We also adjusted the outgrowing areas to the used set of simulated clusters. The two data sets only differed in the global minima of the molecular clusters, and consequently their  $\Delta S$  and  $\Delta H$  but were calculated at the same level of theory. The results shown in this section should not be compared to simulation runs presented in the following sections due to the differences in considered clusters and symmetry treatment.

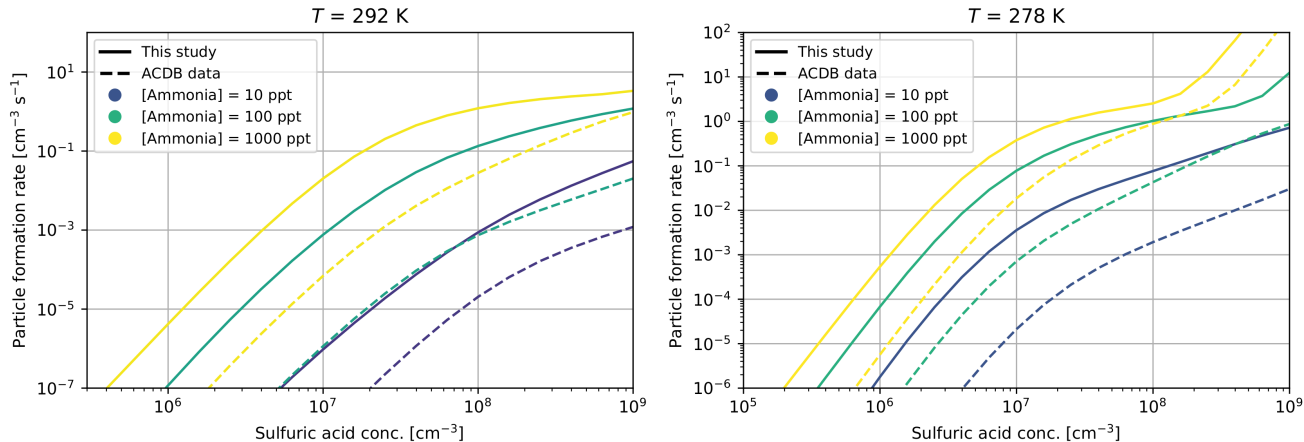


Figure 4: Particle formation rates from simulation runs with cluster structures from ACDB and this study.

A comparison of our structures with ACDB shows that we identified the same global minima only for the smallest clusters. Our structures have a generally lower  $\Delta G$ , and especially for clusters with three or more sulfuric acid molecules, the differences are greater than 1 kcal/mol. Our structures have a higher Gibbs free energy only in five cases, of which the highest difference is 0.67 kcal/mol for the  $(\text{NH}_3)_3\text{NH}_4^+(\text{H}_2\text{SO}_4)_2$  cluster. As Fig. 4 shows, these differences result in significantly higher particle formation rates when using the new data. Generally, this illustrates the importance of performing systematic and thorough conformational sampling. The new molecular cluster data will be updated to ACDB.

## Symmetry

We prepared three different sets of  $\Delta G$ s as input for ACDC differing in their consideration of symmetry. In the first set, we made sure that the Gaussian program does not use any symmetry at all, in the second set, we only turned on symmetry for the monomers, and in the third set, we additionally turned on symmetry for clusters. Cluster symmetry was identified visually and with the SYMMOL<sup>62</sup> program and we corrected for it according to Eqs. (4) and (5).

Regarding the Gibbs free energy of formation, the treatment of symmetry results in a differ-

ence of roughly 1 kcal/mol between *"no symmetry at all"* and *"only monomer symmetry"* for each addition of a  $(\text{NH}_3)(\text{H}_2\text{SO}_4)$  pair into the cluster (*cf.* dashed lines in Fig. 5). Fig. 6 shows the molecular clusters for which the program Symmol found symmetry which was not found by the quantum chemistry programs due to tight classification thresholds and results in a rotational symmetry number  $>1$ . Several molecular clusters were symmetric but still had a rotational symmetry number of 1, e.g. with point group  $C_s$ . If the rotational symmetry number is 1 it has no impact on the  $\Delta G$ s. Cluster (b) is a good example for the difficulties of determining symmetry only by eye: it has a vertical symmetry axis.

Fig. 7 shows particle formation rates obtained from ACDC for different symmetry inclusion cases. Taking symmetry into account results in more stable clusters and thus higher formation rates. As Eq. (5) shows, accounting for the symmetry of monomers changes  $\Delta G$  for all involved clusters leading to a significant increase in formation rates. On the other hand, the particle formation rates for *"only monomer symmetry"* and *"cluster and monomer symmetry"* differ only slightly at low concentrations of sulfuric acid and converge for higher concentrations. In order to explain this behaviour, we have to understand the shape of the particle formation rate vs. sulfuric acid concentration curve.

In general, the particle formation rate exhibits a inverted sigmoid slope. This slope is due to two different particle formation modes and a transition region between them. The first mode is in a region with a relatively high concentration of charged clusters compared to that of neutral clusters, and the particle formation rate continuously increases with increasing sulfuric acid concentration. Then intermediate sulfuric acid concentrations correspond to a transition region where the particle formation rate increases only weakly with acid concentration. The second mode is found at even higher acid concentrations where the formation rate again increases strongly.

The cluster fluxes show that positive clusters grow into larger positive clusters, negative clusters grow into larger negative clusters, and that neutral clusters generally grow through collisions with charged monomers into larger negative and positive clusters at all monomer



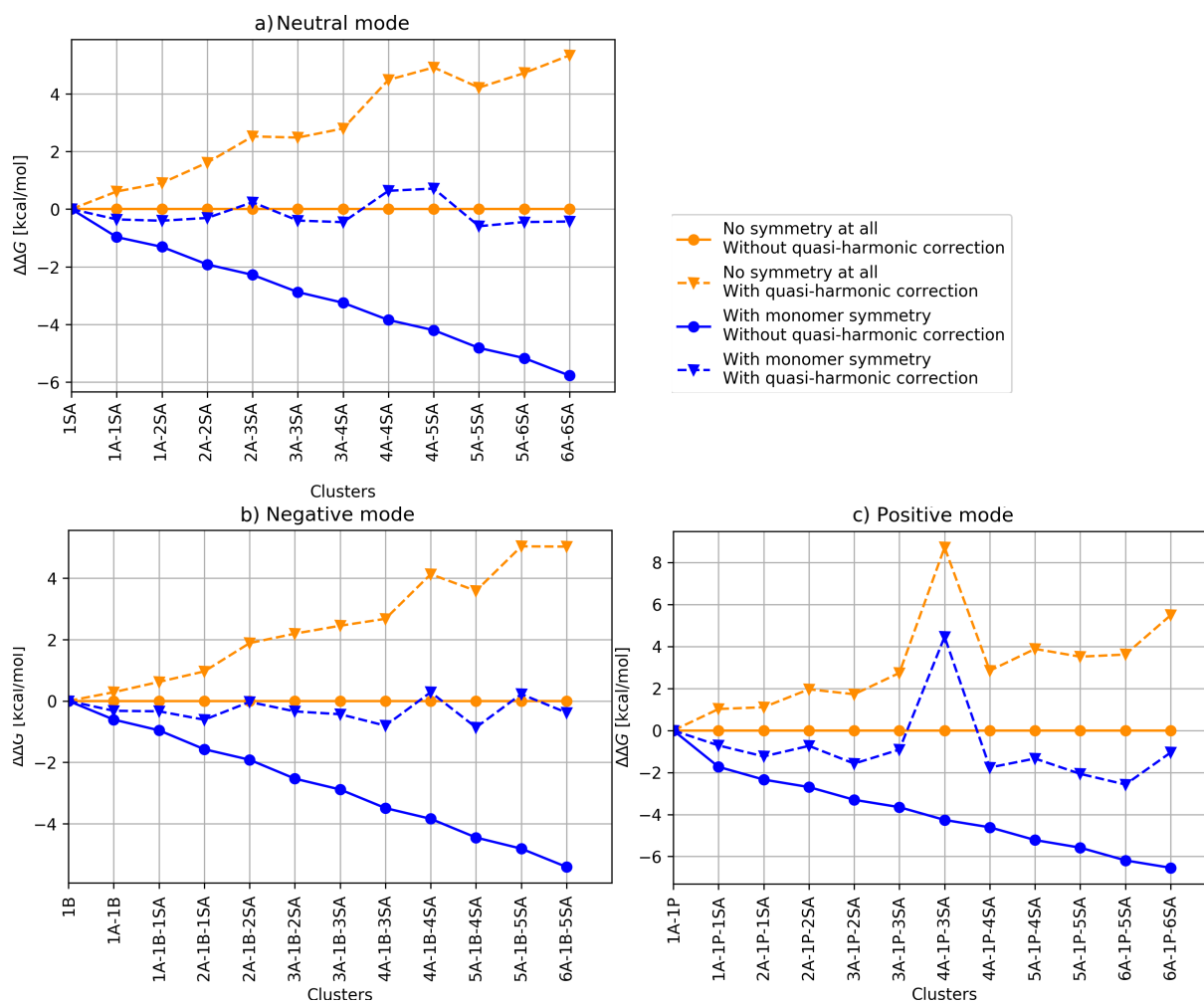


Figure 5: The differences of Gibbs free energies of formation  $\Delta\Delta G$  in kcal/mol at 298 K and reference pressure 1 atm for (a) neutral clusters and (b) negative clusters with the same number of acid and ammonia molecules or one excess acid, and (c) positive clusters with the same number of acid and ammonia molecules or one excess ammonia: SA = sulfuric acid, A = ammonia, B = bisulfate, P = proton. The points are connected to guide the eye.

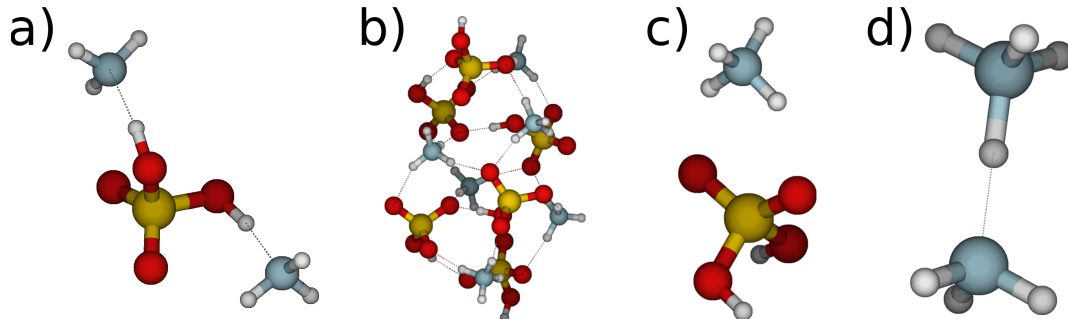


Figure 6: Molecular clusters with symmetry and a rotational symmetry number  $\sigma$  larger than 1. Cluster (a)  $(\text{NH}_3)_2(\text{H}_2\text{SO}_4)$ , (b)  $(\text{NH}_4^+)_6(\text{HSO}_4^-)_6$  and (c)  $(\text{NH}_4^+)(\text{H}_2\text{SO}_4)$  belong to the  $C_2$  point group ( $\sigma = 2$ ) and cluster (d)  $(\text{NH}_4^+)(\text{NH}_3)$  belongs to the  $C_3$  point group ( $\sigma = 3$ ).

concentrations. Additionally, negative and positive clusters may recombine into neutral clusters. However, the growth of neutral clusters into larger neutral clusters is not observed unless the monomer concentrations are set high as they are unstable at lower vapor concentrations where evaporation is more significant. Also, since the ionization rate stays constant, a higher concentration of sulfuric acid and/or ammonia means less charged clusters relative to neutral ones. Accordingly, the plateau in the transition region is located around the ionization rate on the y-axis after which neutral particle formation starts dominating. Therefore, the fraction of purely neutral cluster growth increases with increasing monomer concentrations. The switch from charged clustering pathways to the neutral pathways gives rise to the two modes.

Fig. 6 shows that two small positively charged molecular clusters are symmetric,  $(\text{NH}_4^+)(\text{H}_2\text{SO}_4)$  (c) and  $(\text{NH}_4^+)(\text{NH}_3)$  (d), of which cluster (d) has the highest symmetry number  $\sigma$  of all clusters. Inclusion of cluster symmetry increases  $\Delta G$  for these specific clusters located at the very beginning of the positive growth path. Therefore, clustering in the positive mode is influenced by the symmetry stronger than clustering in the neutral mode, and the differences caused by introducing cluster symmetry decrease with increasing sulfuric acid concentration as the importance of the neutral particle formation pathway increases. The positive mode particle formation rate increases when symmetry is taken into account as positive clusters

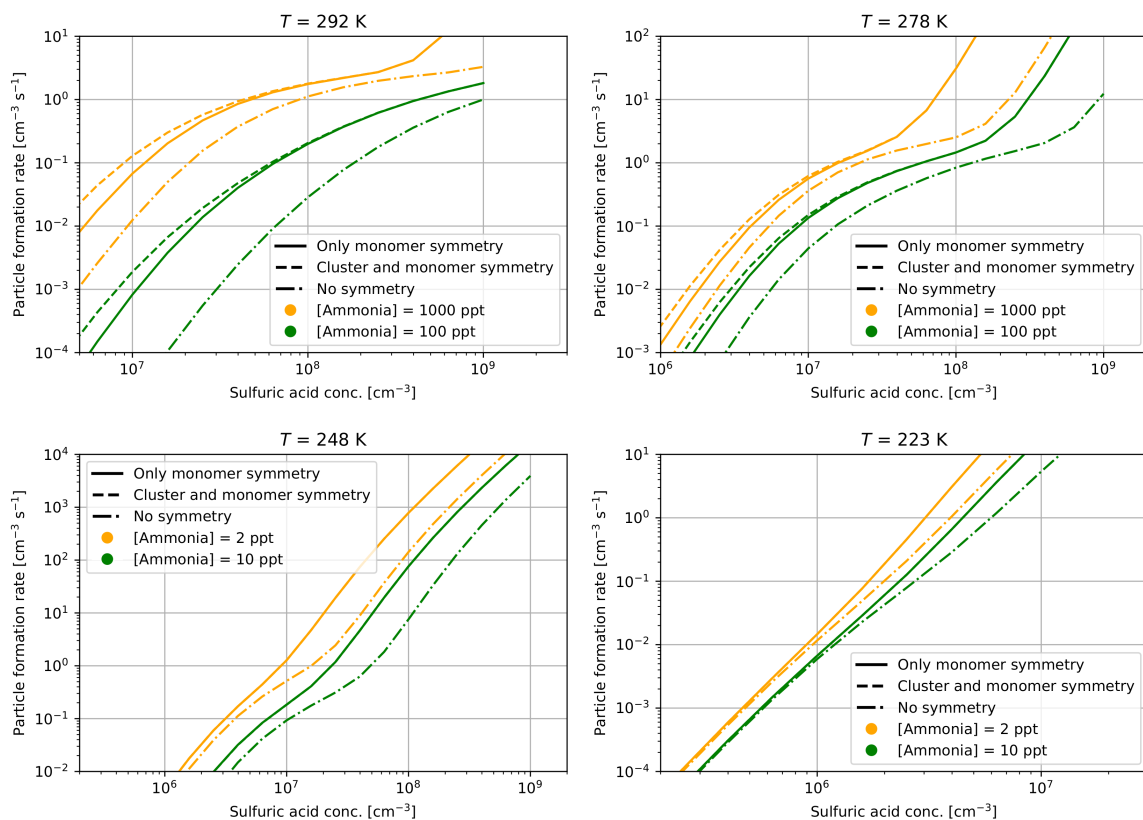


Figure 7: Particle formation rates in  $\text{cm}^{-3}\text{s}^{-1}$  for different concentrations of sulfuric acid and ammonia at four different temperatures. Comparison of different choices of including symmetry. At a temperature of 248 K and 223 K "Only monomer symmetry" and "Cluster and monomer symmetry" overlap.

evaporate back to cluster (c) and (d) at a lower rate, i.e., grow at a higher net rate. The differences in the particle formation rates caused by taking into account cluster symmetry are only significant for rather low formation rates of charged clusters and thus have little practical relevance in the boundary layer conditions. It is also worth recalling that, as mentioned in the introduction, mixing ratios of sulfuric acid rarely exceed concentrations about  $2 \cdot 10^8 \text{ cm}^{-3}$  in the lower atmosphere, whereas lab experiments often go to higher concentrations. Therefore, in the real atmospheric conditions inclusion of symmetry does not result in significant differences in particle formation rates.

### Quasi-harmonic correction

We prepared two different sets of  $\Delta G$ s, one with the quasi-harmonic correction applied and one without it. The orange lines in Fig. 5 show clearly the difference of  $\Delta G$ s for selected clusters in these two sets. While the quasi-harmonic correction increases all  $\Delta G$ s, it affects some clusters more than others. The "4A-1P-3SA" cluster is an outlier and very strongly impacted by the quasi-harmonic correction due to the presence of a exceptionally low frequency mode at  $14.24 \text{ cm}^{-1}$ . The dashed blue line in Fig. 5 shows that the quasi-harmonic correction and the inclusion of monomer symmetry almost cancel each other out in terms of their effect on the  $\Delta G$  where only in positive mode the decrease introduced by the symmetry is consistently larger than the increase introduced by the quasi-harmonic correction. It should be noted that although this cancellation is present for our sulfuric acid - ammonia system, the balance of the quasi-harmonic correction and the symmetry correction can be very different for other systems. Moreover, these two corrections are computationally very cheap, and thus do not constitute bottlenecks, and therefore we suggest that future studies perform both corrections rather than rely on uncertain error cancellation.

Charged clusters are, in general, more strongly bound than the corresponding neutral ones, as long as we compare clusters with roughly similar numbers of acid and base molecules. Therefore, neutral clusters tend to have more low vibrational frequencies. Subsequently, as

shown in Fig. 8, the impact of the quasi-harmonic correction is stronger for higher sulfuric acid concentrations, where neutral clusters dominate the overall particle formation path. Overall, the quasi-harmonic correction decreases cluster stability. Thus, the quasi-harmonic correction slightly lowers the particle formation rate by shifting the curve to the right and down. This effect is more pronounced for higher temperatures, where the effect of entropy (including vibrational entropy) on free energy is greater. However, as pointed out for inclusion of symmetry in the end of the previous section, the effect of quasi harmonic correction is significant at higher vapor concentrations and at higher temperatures, and thus has little practical relevance in terms of the real troposphere.

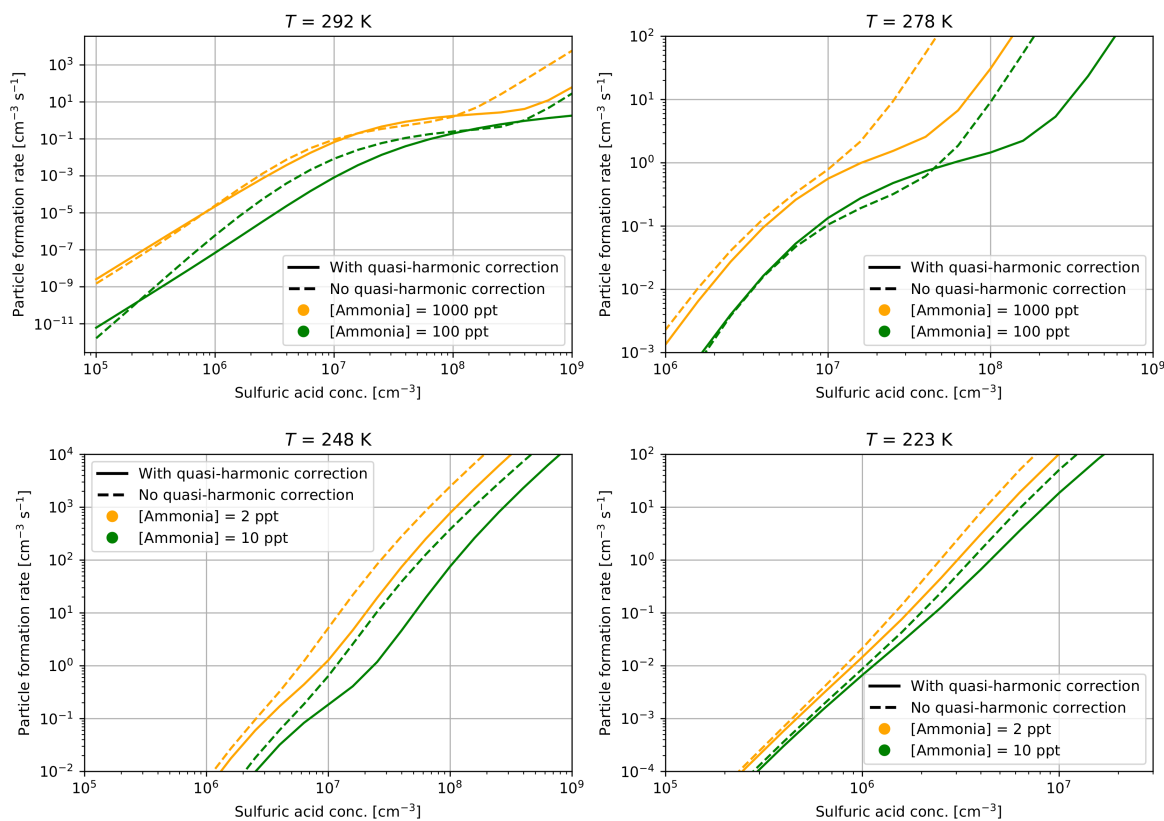


Figure 8: Particle formation rates in  $\text{cm}^{-3} \text{s}^{-1}$  for different concentrations of sulfuric acid and ammonia at four different temperatures. The particle formation rates have been modeled with and without the quasi-harmonic correction. The  $\Delta G$ s were calculated with monomer symmetry.

## Choice of the set of simulated clusters

In this section, we investigate the influence of the set of simulated clusters on the particle formation rates. The 6x6 set of clusters is that shown in Fig. 2. The 5x5 set of clusters consists only of the clusters with less than six ammonia molecules and less than six sulfuric acid molecules. The 4x4 set of clusters consists only of the clusters with less than five ammonia molecules and less than five sulfuric acid molecules. The outgrowing areas are determined for each simulation set size using the same philosophy as in Fig. 2.

The results for different temperatures and ammonia concentrations are shown in Fig. 9. All the simulated particle formation rates converge with increasing sulfuric acid concentration analogous to the comparison of "only monomer" and "cluster and monomer" symmetries shown in Fig.7. At the lower sulfuric acid concentration, where the particle formation rates differ for different simulation sets, clustering predominantly occurs through charged clusters.

Generally, an increase of the cluster set size can either decrease the particle formation rate or leave it unchanged. Once a cluster grows into the outgrowing area, it is counted towards the particle formation rate. When the set size is increased, some outgrowing clusters of the previous, smaller set become "regular" clusters in the enlarged set. These, now "regular", clusters can now either collide with other clusters/monomers or evaporate back to smaller clusters. The particle formation rate decreases if they evaporate at a significant rate.

Deeper insight is provided by the "actual  $\Delta G$ "s of clusters which take the real partial pressures of the monomers into account. For most of the systems and conditions studied here, the actual  $\Delta G$  surface contains one or more barriers between the monomers and the outgrowing clusters. The highest barrier on the lowest-energy path connecting the monomers to the outgrowing clusters (a saddle point on the actual  $\Delta G$  surface) represents the so-called "critical cluster". Due to the strong bonding between acids and bases, this lowest-energy path is typically close to the diagonal, i.e. the clusters on it tend to have roughly equally

many acids and bases. Clusters smaller than the critical cluster are generally more likely to evaporate than to grow, while clusters larger than it are more likely to grow than to evaporate. (A common but false misconception is that clusters larger than the critical cluster do not evaporate at all.) For this reason, the explicitly simulated set of clusters should always include the critical cluster, and preferably some slightly larger clusters as well.

For charged clusters, the actual  $\Delta G$  surface almost inevitably contains also a deep well (local minimum) corresponding to the core ion and a small number of neutral molecules ("ligands" in the language of co-ordination chemistry), located prior to the maximum corresponding to the critical cluster. The actual  $\Delta G$  surface at 278 K, [ammonia] = 1000 ppt and [sulfuric acid] =  $10^5 \text{ cm}^{-3}$  is given as an example in the SI.

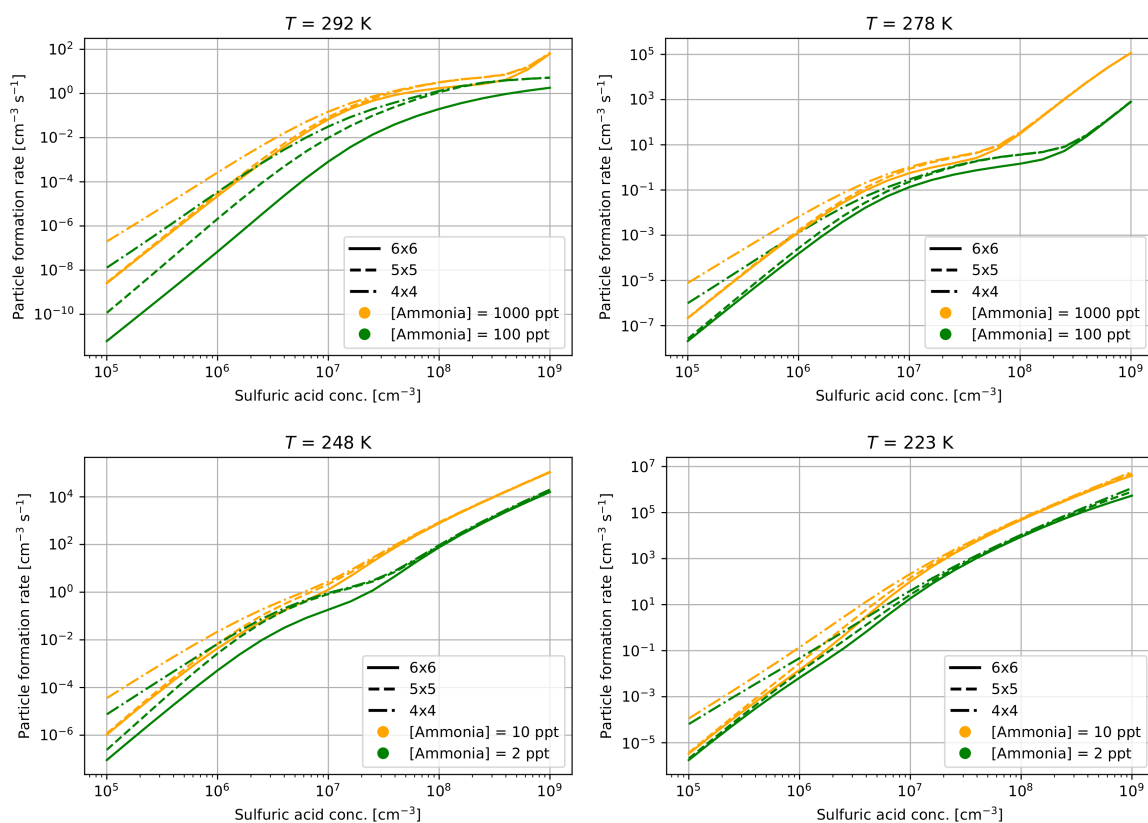


Figure 9: Particle formation rates in  $\text{cm}^{-3} \text{ s}^{-1}$  for different concentrations of sulfuric acid and ammonia at four different temperatures for various maximum sizes of clusters included in the simulation.

When extending the set of simulated clusters, there is a drop in the particle formation rates at all temperatures and ammonia concentrations for low sulfuric acid concentrations where cluster growth happens through charged modes, predominantly through the positive mode. The actual  $\Delta G$ s in positive mode for the charged clusters with **four** sulfuric acid or **four** ammonia monomers are mostly lower or only slightly higher than those for the clusters with **five** sulfuric acid or **five** ammonia. Therefore, there is a clear drop in the formation rate since the set is extended to clusters that have significant evaporation rates. Going from the 5x5 set to the 6x6 set the added clusters have lower actual  $\Delta G$ s than the largest clusters in the 5x5 set. Therefore, the formation rate drops only very little as clusters growing out of the 5x5 set also grow out of the 6x6 set. One exception is at 292 K and [ammonia] = 100 ppt for which the actual  $\Delta G$  surface continuously rises in both the charged and the neutral modes (*cf.* SI): When the cluster set is too small and does not contain the critical cluster the simulation can only give an upper bound for the particle formation rate, which will drop with each extension of the cluster set.

The particle formation rates in the transition region also decrease slightly when extending the set of simulated clusters. The reason for this is that the effect of the simulation set size is generally greater when there are multiple competing growth pathways, and thus multiple potential outgrowing clusters affected by the boundary conditions. We also report that the effect of losses on the modeled formation rates are basically the same for all cluster set sizes, as the bulk of losses occurs in the smaller cluster sizes. Eventually, for high sulfuric acid concentrations cluster growth occurs through neutral clusters for which a critical cluster is within the simulated cluster set at all temperatures and ammonia concentrations. Accordingly, the particle formation rates do not change when changing the set size. We conclude that in atmospherically relevant conditions a 5x5 cluster set is adequate for predicting the particle formation in the sulfuric acid - ammonia system.



## Comparison to CLOUD

In Fig. 10 we plot our simulation results alongside the experimental data points selected and processed as described in *Methodology* and alongside ACDC simulations reported by Kürten et al.<sup>30</sup> based on quantum chemistry data from Ortega et al.<sup>24,63</sup>. The data from Ortega et al. corresponds to B3LYP/CBSB7<sup>64,65</sup> structures and vibrational frequencies combined with RICC2/aug-cc-pvtz energy corrections.<sup>66</sup> The outgrowing areas were chosen as in previous sections. The experimental data points correspond to particles with a mobility diameter of 1.7 nm which is chosen in order to make sure that detected clusters are above critical cluster size<sup>30</sup>. Using the equation used by Kürten et al.<sup>67</sup> we obtain a mobility diameter of 1.36 nm for our 6 sulfuric acid - 6 ammonia cluster. This corresponds to a diameter of 1.4 nm obtained from a quantum chemistry calculation with the Gaussian<sup>45</sup> program with the "Volume" keyword after the addition of 0.3 nm to convert from a geometric diameter to the mobility diameter as suggested by Ku and Fernandez de la Mora<sup>68</sup>. However, the differences in particle formation rates introduced by the "detection" size difference are small as long as the set of simulated clusters contains the critical cluster. Only at 292 K for a ammonia concentration of 100 ppt or less this is not the case, and thus in these conditions the simulated particle formation rate overestimates the true rate.

Some of the experimental data exhibit an inverted sigmoid slope with a plateau-like transition region consistent with the model runs. The present simulations match the experiment reasonably well for low temperatures and for moderate to high ammonia concentrations. This suggests that the presented model is especially suitable for cases with low evaporation rates. For these cases, our model matches the experiments better than the ACDC simulations reported by Kürten et al., and generally displays particle formation rates up to several orders of magnitude lower than those in Kürten et al.'s. In contrast, the particle formation rates of Kürten et al. perform well for the lowest ammonia concentrations for which our new model is approximately one order of magnitude lower than that of Kürten et al.. The results of Kürten et al. are systematically higher than ours because their electronic energies in their

$\Delta G$ s were calculated with RICC2 which is known to introduce significant overbinding. In contrast, DLPNO-CCSD(T) used for our data is known to introduce slight underbinding, which is a possible reason for why some of our particle formation rates are lower than the CLOUD data points.

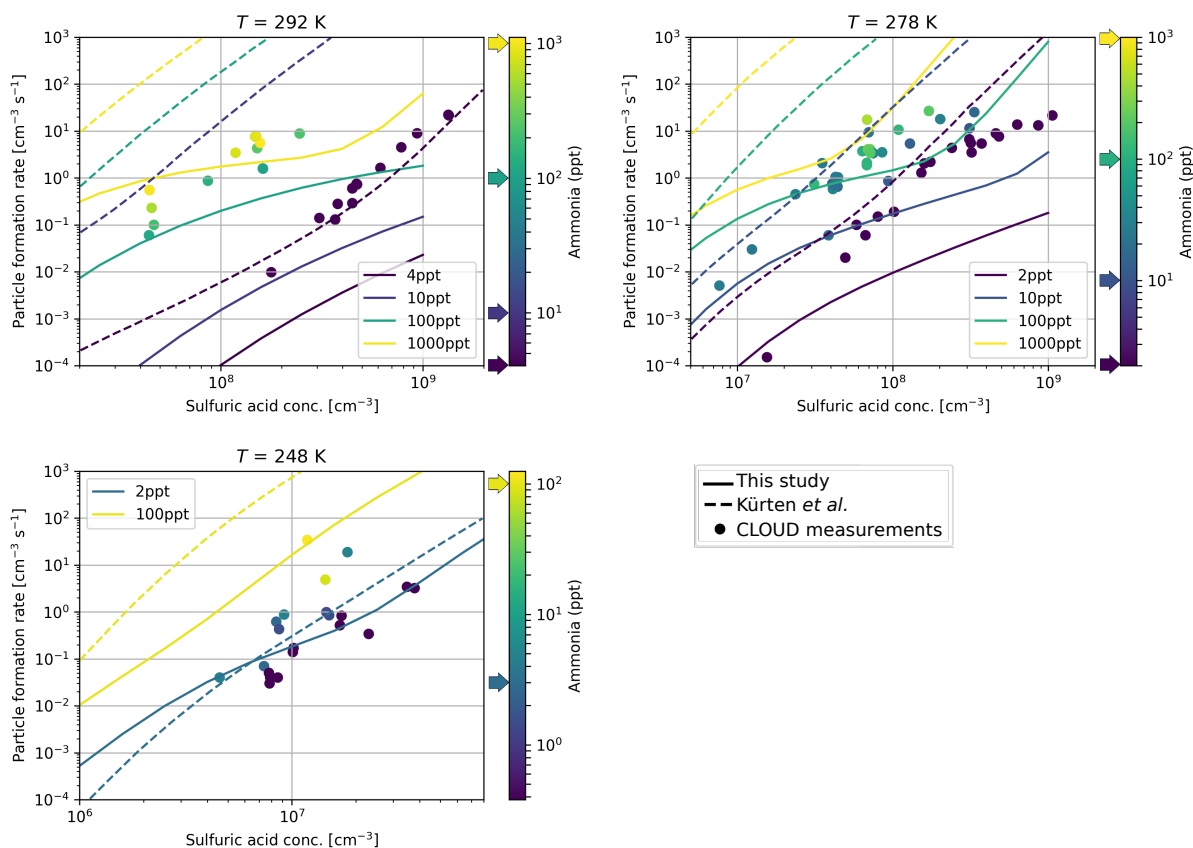


Figure 10: Particle formation rates in cm<sup>-3</sup>s<sup>-1</sup> for different concentrations of sulfuric acid and ammonia modeled by us, Kürten et al.<sup>30</sup> alongside CLOUD measurements<sup>36</sup>. Arrows on the color bar depict the ammonia concentrations where these simulations were conducted.

## Conclusions

Table 1 provides a quick summary of our findings. We performed extensive systematic resampling for electrically neutral, positive and negative sulfuric acid - ammonia clusters with up to six acid and six base molecules. For most cluster compositions, we found significantly more stable cluster conformers than reported in previous studies. This set of clusters was then

used to explore the effect of various assumptions, approximations and boundary conditions on the particle formation rates predicted using the cluster distribution dynamics code ACDC. First, we show that the precise definition of the outgrowing clusters (i.e. which clusters are considered stable and thus counted toward the particle formation rate) is rather insignificant, as long as a chemically reasonable definition is used, and provided that the set of simulated cluster reaches large enough sizes. Second, we showed that including the correct rotational symmetry number of all monomers increases particle formation rates significantly. On the other hand, accounting for the rotational symmetry of the clusters is not necessary for the sulfuric acid - ammonia system due to the relatively low symmetry of almost all minimum-energy clusters. We note that cluster symmetry may still be important for other systems, where clusters exhibit higher degrees of symmetry. Third, we assessed the effect of the quasi-harmonic correction for low-frequency vibrational modes. This correction leads to a general increase in  $\Delta G$ , but also alters the slope of the particle formation curve as a function of the sulfuric acid concentration, as it affects neutral clusters and thus neutral formation pathways more strongly than charged clusters and pathways. The quasi-harmonic correction and the inclusion of monomer symmetry cancel each other out to some degree, implying that previous studies which have typically omitted both corrections may nevertheless give adequate results. However, since neither correction is computationally expensive to perform, we strongly recommend that future studies carry out both. Fourth, we conducted simulations with different maximum limits for the size of the included clusters (6x6, 5x5 and 4x4). The results illustrate the importance of including the critical cluster (defined here as the highest barrier on the lowest-energy path connecting the monomers to the outgrowing clusters) within the explicitly simulated set of clusters. We note that as the composition of the critical cluster depends both on the temperature and on the monomer concentrations, so does the minimum size of the required cluster set. Furthermore, smaller cluster sets are sufficient to model situations where the cluster growth is dominated by a small number of growth pathways, while larger sets are required to capture situations where multiple

growth pathways contribute significantly. Finally, we compared the particle formation rates predicted using our new cluster data set, and the optimal settings revealed by this work, to experimental results from the CLOUD chamber. Compared to previous ACDC simulations, we find substantially better agreement with experiments especially for lower temperatures and/or moderate to high ammonia concentrations.

Table 1: A quick overview of the investigated parameters and the respective findings.

Parameter/Setting	Findings
Symmetry	Consideration of monomer symmetry has significant impact on modeled particle formation rates, whereas symmetry of the molecular clusters has negligible influence.
Quasi-harmonic correction	Using a quasi-harmonic correction generally decreases predicted particle formation rates.
Choice of set of simulated clusters	Impact of the size of the set of simulated clusters explicitly included in the simulation is dependent on the simulated conditions, where the impact increases with increasing evaporation rates.
Boundary conditions for outgrowing clusters	The choice of which clusters are counted as formed particles once they grow out of the model has little impact on the particle formation rate if chosen appropriately.

## Acknowledgement

This research is supported by the European Research Council project 692891-DAMOCLES, Academy of Finland, University of Helsinki, Faculty of Science ATMATH project. We also thank CSC - Finnish IT Centre for access to computer clusters.

## Supporting Information Available

- "ACDC" directory: This contains all input, energies, dipoles and polarizabilities needed to run an ACDC simulation

- "cluster\_ structures" directory: This contains subdirectories with .xyz files and condensed .log files for all molecular clusters and a table containing all corresponding energies, dipoles and polarizabilities
- "SuppInfo.pdf": This contains two actual  $\Delta G$  surfaces, the three structures that are only global minima below 248 K, a description of the neglected outgrowing cluster, and more detailed information on the ACDC input files and on all files in the supporting information

## References

- (1) Brock, C. A.; Hamill, P.; Wilson, J. C.; Jonsson, H. H.; Chan, K. R. Particle formation in the upper tropical troposphere: A source of nuclei for the stratospheric aerosol. *Science* **1995**, *270*, 1650–1653.
- (2) Kulmala, M.; Laakso, L.; Lehtinen, K. E. J.; Riipinen, I.; Dal Maso, M.; Anttila, T.; Kerminen, V.-M.; Hörrak, U.; Vana, M.; Tammet, H. Initial steps of aerosol growth. *Atmos. Chem. Phys.* **2004**, *4*, 2553–2560.
- (3) Clarke, A. D. Atmospheric nuclei in the remote free-troposphere. *J. Atmos Chem.* **1992**, *14*, 479–488.
- (4) Shiraiwa, M.; Ueda, K.; Pozzer, A.; Lammel, G.; Kampf, C. J.; Fushimi, A.; Enami, S.; Arangio, A. M.; Fröhlich-Nowoisky, J.; Fujitani, Y. et al. Aerosol health effects from molecular to global scales. *Environ. Sci. Technol.* **2017**, *51*, 13545–13567.
- (5) Pöschl, U. Atmospheric aerosols: Composition, transformation, climate and health effects. *Angew. Chem. Int. Ed.* **2005**, *44*, 7520–7540.
- (6) Mauderly, J. L.; Chow, J. C. Health effects of organic aerosols. *Inhalation Toxicol.* **2008**, *20*, 257–288.

- (7) Merikanto, J.; Spracklen, D. V.; Mann, G. W.; Pickering, S. J.; Carslaw, K. S. Impact of nucleation on global CCN. *Atmos. Chem. Phys.* **2009**, *9*, 8601–8616.
- (8) Myhre, G.; Shindell, D.; Bréon, F.-M.; Collins, W.; Fuglestedt, J.; Huang, J.; Koch, D.; Lamarque, J.-F.; Lee, D.; Mendoza, B. et al. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Doschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P. M., Eds.; Cambridge University Press: Cambridge, UK, 2013; pp 659–740.
- (9) Yu, F.; Luo, G. Simulation of particle size distribution with a global aerosol model: contribution of nucleation to aerosol and CCN number concentrations. *Atmos. Chem. Phys.* **2009**, *9*, 7691–7710.
- (10) Wang, M.; Penner, J. E. Aerosol indirect forcing in a global model with particle nucleation. *Atmos. Chem. Phys.* **2009**, *9*, 239–260.
- (11) Jacobson, M. C.; Hansson, H. C.; Noone, K. J.; Charlson, R. J. Organic atmospheric aerosols: Review and state of the science. *Rev. Geophys.* **2000**, *38*, 267–294.
- (12) Sun, J.; Ariya, P. A. Atmospheric organic and bio-aerosols as cloud condensation nuclei (CCN): A review. *Atmos. Environ.* **2006**, *40*, 795–820.
- (13) Lohmann, U.; Feichter, J. Global indirect aerosol effects: a review. *Atmos. Chem. Phys.* **2005**, *5*, 715–737.
- (14) Mäkelä, J. M.; Aalto, P.; Jokinen, V.; Pohja, T.; Nissinen, A.; Palmroth, S.; Markkanen, T.; Seitsonen, K.; Lihavainen, H.; Kulmala, M. Observations of ultrafine aerosol particle formation and growth in boreal forest. *Geophys. Res. Lett.* **1997**, *24*, 1219–1222.
- (15) Metzger, A.; Verheggen, B.; Dommen, J.; Duplissy, J.; Prevot, A. S. H.; Weingartner, E.; Riipinen, I.; Kulmala, M.; Spracklen, D. V.; Carslaw, K. S. et al. Evidence for

- the role of organics in aerosol particle formation under atmospheric conditions. *PNAS* **2010**, *107*, 6646–6651.
- (16) Schobesberger, S.; Franchin, A.; Bianchi, F.; Rondo, L.; Duplissy, J.; Kürten, A.; Ortega, I. K.; Metzger, A.; Schnitzhofer, R.; Almeida, J. et al. On the composition of ammonia-sulfuric-acid ion clusters during aerosol particle formation. *Atmos. Chem. Phys.* **2015**, *15*, 55–78.
- (17) Schobesberger, S.; Junninen, H.; Bianchi, F.; Lönn, G.; Ehn, M.; Lehtipalo, K.; Dommen, J.; Ehrhart, S.; Ortega, I. K.; Franchin, A. et al. Molecular understanding of atmospheric particle formation from sulfuric acid and large oxidized organic molecules. *PNAS* **2013**, *110*, 17223–17228.
- (18) Marti, J. J.; Jefferson, A.; Cai, X. P.; Richert, C.; McMurry, P. H.; Eisele, F. H<sub>2</sub>SO<sub>4</sub> vapor pressure of sulfuric acid and ammonium sulfate solutions. *J. Geophys. Res. D: Atmos.* **1997**, *102*, 3725–3735.
- (19) Weber, R. J.; Marti, J. J.; McMurry, P. H.; Eisele, F. L.; Tanner, D. J.; Jefferson, A. Measured atmospheric new particle formation rates: implications for nucleation mechanisms. *Chem. Eng. Commun.* **1996**, *151*, 53–64.
- (20) Petäjä, T.; Mauldin, R. L., III; Kosciuch, E.; McGrath, J.; Nieminen, T.; Paasonen, P.; Boy, M.; Adamov, A.; Kotiaho, T.; Kulmala, M. Sulfuric acid and OH concentrations in a boreal forest site. *Atmos. Chem. Phys.* **2009**, *9*, 7435–7448.
- (21) Ziereis, H.; Arnold, F. Gaseous ammonia and ammonium ions in the free troposphere. *Nature* **1986**, *321*, 503–505.
- (22) Rose, C.; Sellegri, K.; Asmi, E.; Hervo, M.; Freney, E.; Colomb, A.; Junninen, H.; Duplissy, J.; Sipilä, M.; Kontkanen, J. et al. Major contribution of neutral clusters to new particle formation at the interface between the boundary layer and the free troposphere. *Atmos. Chem. Phys.* **2015**, *15*, 3413–3428.

- (23) Kurtén, T.; Torpo, L.; Ding, C.-G.; Vehkamäki, H.; Sundberg, M. R.; Laasonen, K.; Kulmala, M. A density functional study on water-sulfuric acid-ammonia clusters and implications for atmospheric cluster formation. *J. Geophys. Res. D: Atmos.* **2007**, *112*, D04210.
- (24) Ortega, I. K.; Kupiainen, O.; Kurtén, T.; Olenius, T.; Wilkman, O.; McGrath, M. J.; Loukonen, V.; Vehkamäki, H. From quantum chemical formation free energies to evaporation rates. *Atmos. Chem. Phys.* **2012**, *12*, 225–235.
- (25) Leverentz, H. R.; Siepmann, J. I.; Truhlar, D. G.; Loukonen, V.; Vehkamäki, H. Energetics of atmospherically implicated clusters made of sulfuric acid, ammonia, and dimethyl amine. *J. Phys. Chem. A* **2013**, *117*, 3819–3825.
- (26) Kildgaard Jens Vive,; Mikkelsen Kurt V.,; Bilde Merete,; Elm Jonas, Hydration of atmospheric molecular clusters: A new method for systematic configurational sampling. *J. Phys. Chem. A* **2018**, *122*, 5026–5036, doi: 10.1021/acs.jpca.8b02758.
- (27) Myllys, N.; Elm, J.; Kurtén, T. Density functional theory basis set convergence of sulfuric acid-containing molecular clusters. *Comput. Theor. Chem.* **2016**, *1098*, 1 – 12.
- (28) McGrath, M. J.; Olenius, T.; Ortega, I. K.; Loukonen, V.; Paasonen, P.; Kurtén, T.; Kulmala, M.; Vehkamäki, H. Atmospheric Cluster Dynamics Code: a flexible method for solution of the birth-death equations. *Atmos. Chem. Phys.* **2012**, *12*, 2345–2355.
- (29) Olenius, T.; Schobesberger, S.; Kupiainen-Määttä, O.; Franchin, A.; Junninen, H.; Ortega, I. K.; Kurtén, T.; Loukonen, V.; Worsnop, D. R.; Kulmala, M. et al. Comparing simulated and experimental molecular cluster distributions. *Faraday Discuss.* **2013**, *165*, 75–89.
- (30) Kürten, A.; Bianchi, F.; Almeida, J.; Kupiainen-Määttä, O.; Dunne, E. M.; Duplissy, J.; Williamson, C.; Barnet, P.; Breitenlechner, M.; Dommen, J. et al. Experimental particle formation rates spanning tropospheric sulfuric acid and ammonia abundances, ion



- production rates, and temperatures. *J. Geophys. Res. D: Atmos.* **2016**, *121*, 12,377–12,400.
- (31) Zhang, H.; Li, H.; Liu, L.; Zhang, Y.; Zhang, X.; Li, Z. The potential role of malonic acid in the atmospheric sulfuric acid - Ammonia clusters formation. *Chemosphere* **2018**, *203*, 26 – 33.
- (32) Bork, N.; Elm, J.; Olenius, T.; Vehkamäki, H. Methane sulfonic acid-enhanced formation of molecular clusters of sulfuric acid and dimethyl amine. *Atmos. Chem. Phys.* **2014**, *14*, 12023–12030.
- (33) Myllys, N.; Elm, J.; Halonen, R.; Kurtén, T.; Vehkamäki, H. Coupled cluster evaluation of the stability of atmospheric acid-base clusters with up to 10 molecules. *J. Phys. Chem. A* **2016**, *120*, 621–630.
- (34) Olenius, T.; Kupiainen-Määttä, O.; Ortega, I. K.; Kurtén, T.; Vehkamäki, H. Free energy barrier in the growth of sulfuric acid-ammonia and sulfuric acid-dimethylamine clusters. *J. Chem. Phys.* **2013**, *139*, 084312.
- (35) Kubečka, J.; Besel, V.; Kurtén, T.; Myllys, N.; Vehkamäki, H. Configurational sampling of noncovalent (atmospheric) molecular clusters: sulfuric acid and guanidine. *J. Phys. Chem. A* **2019**, *123*, 6022–6033.
- (36) Dunne, E. M.; Gordon, H.; Kürten, A.; Almeida, J.; Duplissy, J.; Williamson, C.; Ortega, I. K.; Pringle, K. J.; Adamov, A.; Baltensperger, U. et al. Global atmospheric particle formation from CERN CLOUD measurements. *Science* **2016**, *354*, 1119–1124.
- (37) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. A direct MP2 gradient method. *Chem. Phys. Lett.* **1990**, *166*, 275 – 280.
- (38) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. Semi-direct algorithms for the MP2 energy and gradient. *Chem. Phys. Lett.* **1990**, *166*, 281 – 289.

- (39) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 energy evaluation by direct methods. *Chem. Phys. Lett.* **1988**, *153*, 503 – 506.
- (40) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I. et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.
- (41) Yu, W.; He, X.; Vanommeslaeghe, K.; MacKerell Jr., A. D. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem.* **2012**, *33*, 2451–2468.
- (42) Zhang, J.; Dolg, M. ABCluster: the artificial bee colony algorithm for cluster global optimization. *Phys. Chem. Chem. Phys.* **2015**, *17*, 24173–24181.
- (43) Zhang, J.; Dolg, M. Global optimization of clusters of rigid molecules using the artificial bee colony algorithm. *Phys. Chem. Chem. Phys.* **2016**, *18*, 3003–3010.
- (44) Grimme, S.; Bannwarth, C.; Shushkov, P. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems parametrized for all spd-block elements ( $Z = 1-86$ ). *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.
- (45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian 16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.
- (46) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

- (47) Chaudhuri, B. How to choose a representative subset from a set of data in multi-dimensional space. *Pattern Recognit. Lett.* **1994**, *15*, 893 – 899.
- (48) Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 034106.
- (49) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013**, *139*, 134101.
- (50) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (51) Partanen, L.; Vehkamäki, H.; Hansen, K.; Elm, J.; Henschel, H.; Kurtén, T.; Halonen, R.; Zapadinsky, E. Effect of conformers on free energies of atmospheric complexes. *J. Phys. Chem. A* **2016**, *120*, 8613–8624.
- (52) Grimme, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. European J.* **2012**, *18*, 9955–9964.
- (53) Funes-Ardoiz, I.; Paton, R. S. GoodVibes: GoodVibes v1.0.1. **2016**,
- (54) Bauschlicher, C. W.; Partridge, H. A modification of the Gaussian-2 approach using density functional theory. *J. Chem. Phys.* **1995**, *103*, 1788–1791.
- (55) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. *J. Chem. Phys.* **1999**, *110*, 2822–2827.
- (56) Baboul, A. G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-3 theory using density functional geometries and zero-point energies. *J. Chem. Phys.* **1999**, *110*, 7650–7657.

- (57) Kupiainen-Määttä, O.; Olenius, T. ACDC: Atmospheric Cluster Dynamics Code technical manual. University of Helsinki, 2017.
- (58) Maso, M. D.; Hyvärinen, A.; Komppula, M.; Tunved, P.; Kerminen, V.-M.; Lihavainen, H.; Öviisanen, Y.; Hansson, H.-C.; Kulmala, M. Annual and interannual variation in boreal forest aerosol particle number and volume concentration and their connection to particle formation. *Tellus B: Chem. Phys. Meteorol.* **2008**, *60*, 495–508.
- (59) Henschel, H.; Kurtén, T.; Vehkamäki, H. Computational study on the effect of hydration on new particle formation in the sulfuric acid/ammonia and sulfuric acid/dimethylamine systems. *J. Phys. Chem. A* **2016**, *120*, 1886–1896.
- (60) Elm, J. An atmospheric cluster database consisting of sulfuric acid, bases, organics, and water. *ACS Omega* **2019**, *4*, 10965–10974.
- (61) Loukonen, V.; Kurtén, T.; Ortega, I. K.; Vehkamäki, H.; Pádua, A. A. H.; Sellegri, K.; Kulmala, M. Enhancing effect of dimethylamine in sulfuric acid nucleation in the presence of water - a computational study. *Atmos. Chem. Phys.* **2010**, *10*, 4961–4974.
- (62) Pilati, T.; Forni, A. Symmol. 1998; <https://www.mtg.msm.cam.ac.uk/files/symmol.zip/view>.
- (63) Ortega, I. K.; Olenius, T.; Kupiainen-Määttä, O.; Loukonen, V.; Kurtén, T.; Vehkamäki, H. Electrical charging changes the composition of sulfuric acid-ammonia/dimethylamine clusters. *Atmos. Chem. Phys.* **2014**, *14*, 7995–8007.
- (64) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (65) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

- (66) Hättig, C.; Weigend, F. CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. *J. Chem. Phys.* **2000**, *113*, 5154–5161.
- (67) Kürten, A.; Li, C.; Bianchi, F.; Curtius, J.; Dias, A.; Donahue, N. M.; Duplissy, J.; Flagan, R. C.; Hakala, J.; Jokinen, T. et al. New particle formation in the sulfuric acid–dimethylamine–water system: reevaluation of CLOUD chamber measurements and comparison to an aerosol nucleation and growth model. *Atmos. Chem. Phys.* **2018**, *18*, 845–863.
- (68) Ku, B. K.; de la Mora, J. F. Relation between electrical mobility, mass, and size for nanodrops 1–6.5 nm in diameter in air. *Aerosol Sci. Technol.* **2009**, *43*, 241–249.

# Graphical TOC Entry

