



Cite this: *Environ. Sci.: Atmos.*, 2023, 3, 552

## Reducing chemical complexity in representation of new-particle formation: evaluation of simplification approaches†

Tinja Olenius, \*<sup>a</sup> Robert Bergström, <sup>a</sup> Jakub Kubečka, <sup>b</sup> Nanna Myllys <sup>c</sup> and Jonas Elm <sup>b</sup>

Adequate representation of new-particle formation from vapors in aerosol microphysics and atmospheric transport models is essential for providing reliable predictions of ambient particle numbers and for interpretation of observations. Atmospheric particle formation processes involve multiple species, which complicates the derivation and implementation of data and parameterizations to describe the processes. Ideally, the representation of multi-compound mechanisms should be reduced, but remain accurate. Here, we evaluate common approaches to simplify the description of representative multi-compound acid–base chemistries by applying different theoretical molecular cluster data sets, focusing on simplifications (1) in formation rates that are used as input in aerosol process models and large-scale models, and (2) in cluster models that are applied to assess particle formation dynamics and survival to larger sizes. We test the following approaches: assuming non-interactive additive formation pathways, lumping of similar species, and application of quasi-unary approximations. We assess the possible biases of the simplifications for different types of chemistries and propose best practices for reducing the chemical complexity. We demonstrate that simplifications in formation rates are most often justifiable, but the choice of the preferred simplification method depends on the types of species and their similarity. Simplifications in cluster growth dynamics by quasi-unary approaches, on the other hand, are reasonable mainly for strong cluster formation involving very low-evaporating species and at excess concentration of the implicitly treated stabilizing compound.

Received 6th December 2022  
Accepted 21st January 2023

DOI: 10.1039/d2ea00174h

rsc.li/esatmospheres

### Environmental significance

New-particle formation is a major source of atmospheric aerosol particles, and thus an important factor affecting aerosol–cloud–climate processes. While model representation of particle formation mechanisms should preferably be concise and robust, the participation of multiple chemical species complicates the description. This work evaluates common approaches to simplify multi-compound chemistries, focusing on sulfuric acid–base systems. We provide assessments of reasonable simplifications for (1) incorporating formation rates in atmospheric models, and (2) detailed studies of new-particle formation and growth dynamics. This contributes to finding the best approaches for representing secondary particle sources, and mapping and reducing model uncertainties.

## 1 Introduction

Formation of new aerosol particles from condensable vapors makes a major contribution to atmospheric aerosol numbers, and has significant—but uncertain—effects on aerosol–cloud–climate interactions, the global radiation budget, and air quality.<sup>1–3</sup> The initial formation of nanometer-sized particles, occurring through clustering of gas-phase molecules to

molecular clusters, is affected by multiple chemical compounds, including acid and base species and potentially also oxidized organic compounds.<sup>4,5</sup> The clustering can proceed through multi-compound chemical mechanisms or parallel pathways,<sup>6,7</sup> which need to be adequately represented in atmospheric aerosol models in order to capture the effects on particle concentrations.

The main drivers of the initial particle formation process include mixtures of acid and base compounds.<sup>8–10</sup> Specifically, sulfuric acid (H<sub>2</sub>SO<sub>4</sub>) is a central particle formation agent together with stabilizing base species, the most abundant of which is ammonia (NH<sub>3</sub>). In addition, atmospheric amines efficiently form molecular clusters with H<sub>2</sub>SO<sub>4</sub>.<sup>11</sup> These include monoamines, *i.e.* mono-, di- and trimethylamine,<sup>12</sup> of which dimethylamine has often been studied as a representative

<sup>a</sup>Swedish Meteorological and Hydrological Institute, SE-60176 Norrköping, Sweden.  
E-mail: tinja.olenius@alumni.helsinki.fi; Tel: +46 76 495 7787

<sup>b</sup>Department of Chemistry, Aarhus University, Langelandsgade 140, 8000 Aarhus C, Denmark

<sup>c</sup>Department of Chemistry, University of Helsinki, FI-00014 Helsinki, Finland

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d2ea00174h>



amine species,<sup>13</sup> as well as diamines.<sup>14</sup> Amines are typically more efficient clustering agents than  $\text{NH}_3$  but their atmospheric concentrations are generally much lower than that of  $\text{NH}_3$ , although elevated close to the sources.<sup>15–17</sup> Depending on the ambient conditions, these different base species can contribute to  $\text{H}_2\text{SO}_4$ -driven particle formation either alone or together.

The presence of multi-compound formation pathways increases the complexity of the description of the formation process in several ways. The functional form of the formation rate rapidly becomes very complex as the number of independent parameters increases, which complicates both the experimental and theoretical efforts to study the formation mechanisms, as well as their implementation in aerosol modeling. Experimental assessment of formation rates over a wide range of combinations of precursor vapors at different concentrations,<sup>18</sup> as well as deriving robust parameterizations of the data, is laborious. Theoretical predictions of the properties of multi-compound molecular clusters of various compositions are likewise demanding. Such predictions are obtained by quantum chemical calculations,<sup>19</sup> which require substantial computational resources when performed over large sets of cluster compositions. Simulations of molecular cluster population dynamics,<sup>20,21</sup> generally applied together with quantum chemical input data to yield cluster concentrations and formation rates, also become more complex. A very large number of different molecular compositions can both increase the simulation time and make interpretation of results less straight-forward. Finally, implementation of formation rates in atmospheric models requires more sophisticated interfaces when multiple data sets are needed to determine the rates for different vapor species and their combinations. Overall, reducing chemical complexity in the representation of nanoparticle formation mechanisms by appropriate simplifications is highly desirable

- (1) due to limitations in available experimental and theoretical data,
- (2) for an optimized description of the process in large-scale models, and
- (3) for robust interpretation of measurements by molecular cluster dynamics modeling.

A common simplification that is routinely applied to account for several particle formation pathways is to assume that the formation rates  $J$  determined for different chemical systems are additive.<sup>3,22</sup> For instance, if there exist data for particle formation from compounds A and B, and from A and C, the rates are summed to yield the total formation rate as  $J_{\text{tot}} = J_{\text{AB}} + J_{\text{AC}}$ . This corresponds to assuming that clustering pathways AB and AC do not interfere, meaning that compounds B and C compete to form clusters with A and/or do not form bonds with each other. While a simplified treatment is unavoidable when there exist no data for the three-component system ABC, potential errors arising from the sum approach have not been assessed for typical atmospheric clustering agents.

Another approach to reduce the number of compounds is to approximate chemically similar compounds as a lumped species. This has been assumed, for example, for amines.<sup>23</sup> If the chemical properties of clusters in systems AB, AC, and ABC are very similar, it is justifiable to approximate the ABC system as AD, where D is

a representative species with a concentration of  $[\text{D}] = [\text{B}] + [\text{C}]$ . However, in this case the sum approach may give distorted results, as it cannot generally be assumed that  $J_{\text{AD}}([\text{D}] = [\text{D}]_1 + [\text{D}]_2) = J_{\text{AD}}([\text{D}] = [\text{D}]_1) + J_{\text{AD}}([\text{D}] = [\text{D}]_2)$ . That is, the choice of simplification should depend on the properties of the chemical systems.

Multi-compound systems can also be reduced to quasi-unary systems by implicit treatment of compounds.<sup>24–28</sup> This corresponds to assuming that one or more compounds do not need to be explicitly included, but only affect the cluster properties, most importantly the cluster evaporation rates which generally strongly depend on the molecular composition.<sup>29,30</sup> This corresponds to representing, for example, all cluster compositions  $\text{A}_3\text{B}_x$  that consist of three molecules of compound A and any number of molecules of compound B, as a single cluster  $\text{A}_3$  with effective properties that are affected by the presence of B. Here, B is an enhancing species that cannot drive clustering on its own. This is justifiable, for instance, when the rate constants for collisions and evaporations of compound B are significantly higher than for A. The dynamic processes related to B thus occur in such short time scales that the cluster distribution is always instantaneously equilibrated with respect to B.<sup>25,26</sup> Representative properties may also be a justifiable assumption when the clusters are likely to exist in specific, thermodynamically favorable compositions. For example,  $\text{H}_2\text{SO}_4$ -dimethylamine clusters have been observed to favor a close to 1 : 1 molar ratio<sup>27,30</sup> and have thus been approximated as a quasi-unary compound consisting of  $\text{H}_2\text{SO}_4$ -dimethylamine pairs.<sup>10,27</sup>

In this work, we evaluate the approaches to simplify the representation of multi-compound new-particle formation. We test the validity of the typical simplifications by comparison against explicit simulations for atmospherically relevant  $\text{H}_2\text{SO}_4$ -base chemistries. The purpose is to answer the following questions:

- (1) Can the simplified approaches be expected to represent the multi-compound systems with reasonable accuracy?
- (2) How significant errors may the simplifications cause, in terms of the magnitude and direction of the possible errors?
- (3) Does the applicability of the simplifications depend on the properties of the chemical system and/or the ambient conditions?

## 2 Validation of simplifications against explicit simulations

The simplification approaches, depicted in Fig. 1, are validated against explicit simulations of representative multi-compound chemical systems over a wide range of atmospheric conditions. We apply available sets of quantum chemical data for the thermodynamic properties of  $\text{H}_2\text{SO}_4$ -base clusters together with the cluster population dynamics model ACDC,<sup>31</sup> as detailed below.

### 2.1 Chemical systems and thermodynamic data sets

We focus on  $\text{H}_2\text{SO}_4$ -base chemistries where the primary bases of interest include ammonia ( $\text{NH}_3$ ) and dimethylamine (DMA). In order to ensure the generalizability of the results, we apply



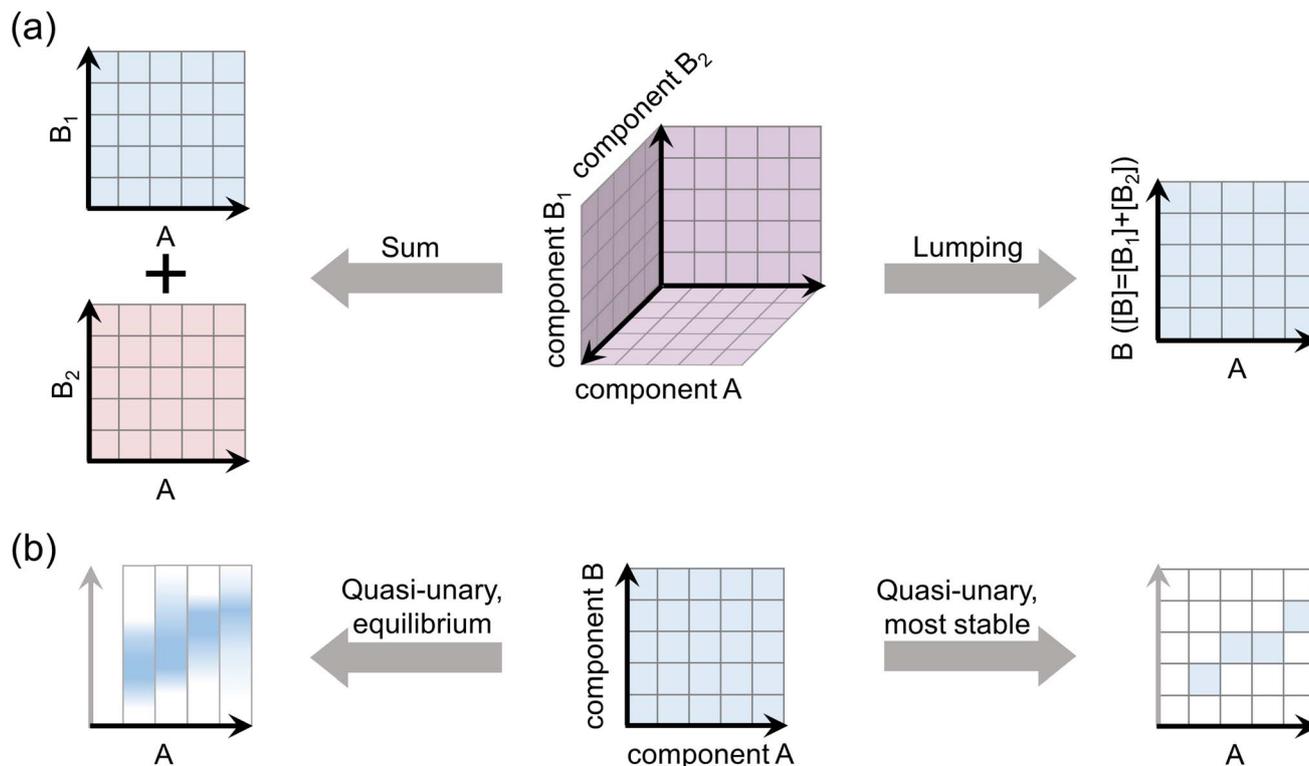


Fig. 1 Schematic presentation of the approaches to simplify multi-component chemistries for (a) assessments of particle formation rate in the presence of multiple species, and (b) modeling of cluster growth by a quasi-unary framework. The axes depict cluster compositions as numbers of molecules of different species.

several quantum chemical data sets, as listed in Table 1. These are the most complete data sets available in terms of included cluster sizes and compositions, and they have been evaluated against laboratory measurements in previous works (see references in Table 1). These widely studied chemical systems represent weaker ( $\text{NH}_3$ ) and stronger (DMA) clustering agents that are also expected to interact and exhibit synergistic effects.<sup>32–34</sup>

In addition, we perform tests with  $\text{H}_2\text{SO}_4$ –guanidine data. Guanidine, here referred to as GUA, is a very strong organobase that is able to form extremely low-evaporating clusters with  $\text{H}_2\text{SO}_4$ .<sup>30,35,36</sup> Finally, we apply a set of multi-base data to simulate  $\text{H}_2\text{SO}_4$ –base<sub>1</sub>–base<sub>2</sub> chemistries with different combinations of bases, including  $\text{NH}_3$ , DMA, methylamine (MA), trimethylamine (TMA), and ethylenediamine (EDA).<sup>37</sup> These amine species are observed in the ambient atmosphere, and cover weaker (MA) and stronger (DMA, TMA, EDA) clustering strengths.<sup>14,38</sup> Further details of all data sets and simulation settings can be found in the ESI.†

## 2.2 Cluster dynamics simulations

Cluster concentrations and formation rates are simulated by the cluster population model ACDC,<sup>20,31</sup> which generates and solves the time derivatives of all cluster concentrations  $C_i$  for a given set of cluster compositions  $i$ :

$$\frac{dC_i}{dt} = \frac{1}{2} \sum_{j < i} (\beta_{j,i-j} C_j C_{i-j} - \gamma_{i \rightarrow j,i-j} C_i) - \sum_j (\beta_{ij} C_i C_j - \gamma_{i+j \rightarrow ij} C_{i+j}) - S_i C_i \quad (1)$$

In eqn (1), the first sum includes all collisions that create composition  $i$ , and the corresponding reverse evaporations. The second sum includes collisions of cluster  $i$  with other clusters, and the reverse evaporations that produce composition  $i$ .  $\beta_{i,j}$  is the rate constant for a collision between clusters  $i$  and  $j$ , and  $\gamma_{i \rightarrow j,i-j}$  is the rate constant for evaporation of cluster  $i$  into products  $j$  and  $i - j$ .  $S_i C_i$  is a loss term corresponding to scavenging of clusters. Here, the collision constants are calculated using the kinetic gas theory, and the evaporation constants are obtained from the quantum chemical formation free energies.<sup>20</sup> The loss rate constants correspond to scavenging by larger aerosol particles, characterized by the  $\text{H}_2\text{SO}_4$  condensation sink CS as  $S_i = \text{CS} \times (d_{p,i}/d_{p,\text{H}_2\text{SO}_4})^{-1.6}$ , where  $d_{p,i}$  and  $d_{p,\text{H}_2\text{SO}_4}$  are the diameters of cluster  $i$  and  $\text{H}_2\text{SO}_4$ , respectively.<sup>41</sup> Ionic molecules are included when available (see ESI Section 1†) by adding collisions with ionizing species in eqn (1), corresponding to ionization and recombination processes.<sup>20</sup>

The solution to eqn (1) gives the cluster concentrations  $C_i$ , and the particle formation rate  $J$  is defined as

$$J = \frac{1}{2} \sum_i \sum_j \beta_{ij} C_i C_j \text{ for } \{i+j|\text{condition}\}, \quad (2)$$



**Table 1** Chemical systems and quantum chemical data sets applied in this work. Data sets "RICC2" and "DLPNO" are computed at RICC2/aug-cc-pV(T+d)Z//B3LYP/CBSB7 and DLPNO-CCSD(T)/aug-cc-pVTZ// $\omega$ B97X-D/6-31++G(d, p) levels of theory, respectively. Different data sets also involve different configurational sampling approaches and theoretical approximations (see the ESI for more details)

Chemical system	Quantum chemical data sets	Cluster size range (max. numbers of H <sub>2</sub> SO <sub>4</sub> and base molecules)	References
H <sub>2</sub> SO <sub>4</sub> -NH <sub>3</sub> , H <sub>2</sub> SO <sub>4</sub> -DMA, H <sub>2</sub> SO <sub>4</sub> -NH <sub>3</sub> -DMA	RICC2	4 H <sub>2</sub> SO <sub>4</sub> , 4 base, 5 H <sub>2</sub> SO <sub>4</sub> , 5 NH <sub>3</sub> <sup>a</sup>	Almeida <i>et al.</i> (2013) <sup>13</sup> Olenius <i>et al.</i> (2013) <sup>20</sup> Kürten <i>et al.</i> (2016) <sup>39</sup>
	DLPNO_Myllys	4 H <sub>2</sub> SO <sub>4</sub> , 4 base	Myllys <i>et al.</i> (2019a) <sup>30</sup> Myllys <i>et al.</i> (2019b) <sup>32</sup>
	DLPNO_Besel <sup>b</sup>	6 H <sub>2</sub> SO <sub>4</sub> , 6 NH <sub>3</sub>	Besel <i>et al.</i> (2020) <sup>40</sup>
	DLPNO_Kubečka	4 H <sub>2</sub> SO <sub>4</sub> , 4 base	Kubečka <i>et al.</i> (2022) <sup>37</sup>
H <sub>2</sub> SO <sub>4</sub> -GUA	DLPNO_Myllys	4 H <sub>2</sub> SO <sub>4</sub> , 4 base	Myllys <i>et al.</i> (2018) <sup>35</sup> Myllys <i>et al.</i> (2019a) <sup>30</sup>
H <sub>2</sub> SO <sub>4</sub> -base <sub>1</sub> -base <sub>2</sub> with all combinations of NH <sub>3</sub> , MA, DMA, TMA, EDA	DLPNO_Kubečka	4 H <sub>2</sub> SO <sub>4</sub> , 4 base	Kubečka <i>et al.</i> (2022) <sup>37</sup>

<sup>a</sup> For comparisons with H<sub>2</sub>SO<sub>4</sub>-NH<sub>3</sub>-DMA chemistry (Section 2.3.1), a smaller cluster set with up to 4 H<sub>2</sub>SO<sub>4</sub> and 4 NH<sub>3</sub> molecules is used for consistency. The 5 H<sub>2</sub>SO<sub>4</sub>, 5 NH<sub>3</sub> set is used in the quasi-unary simulations (Section 2.3.2). <sup>b</sup> DLPNO\_Besel is primarily applied for the H<sub>2</sub>SO<sub>4</sub>-NH<sub>3</sub> chemistry in the quasi-unary simulations (Section 2.3.2) as it covers larger cluster sizes than other DLPNO-based data, which are used for the particle formation rate assessments (Section 2.3.1).

corresponding to the rate at which stable clusters grow beyond the simulated size range. In eqn (2), the summation goes over all collisions that lead to a stable collision product  $i + j$  according to the given stability criteria, denoted by  $\{i + j \mid \text{condition}\}^{20}$  (Table S1†). We focus on steady-state conditions, and perform simulations over sets of atmospherically relevant vapor concentrations, temperatures, and CS, from clean to highly polluted environments. As H<sub>2</sub>SO<sub>4</sub> molecules tend to cluster with strong bases, we define the H<sub>2</sub>SO<sub>4</sub> vapor concentration as the total concentration of H<sub>2</sub>SO<sub>4</sub> monomers and H<sub>2</sub>SO<sub>4</sub>-base, where "base" refers to any base molecule. This is consistent with measurements, as the clustered molecules are expected to contribute to the H<sub>2</sub>SO<sub>4</sub> signal.<sup>42</sup>

## 2.3 Simplified set-ups and comparison to explicit results

### 2.3.1 Representation of the total particle formation rate in atmospheric models.

First, the H<sub>2</sub>SO<sub>4</sub>-NH<sub>3</sub>-DMA data sets are applied to simulate a multi-compound system with chemically different but possibly interacting species with weaker and stronger cluster formation potentials. The assumption of non-interacting particle formation pathways is tested by comparing the three-compound formation rate  $J_{3\text{-comp}}$  to the sum of the formation rates  $\sum J_{2\text{-comp}}$  for the two-compound systems H<sub>2</sub>SO<sub>4</sub>-NH<sub>3</sub> and H<sub>2</sub>SO<sub>4</sub>-DMA (Fig. 1a). Second, ideal cases of chemically similar systems are constructed by applying the same thermochemical data for two separate simulation systems H<sub>2</sub>SO<sub>4</sub>-base<sub>1</sub> and H<sub>2</sub>SO<sub>4</sub>-base<sub>2</sub>, where both base<sub>1</sub> and

base<sub>2</sub> have the properties of either NH<sub>3</sub> or DMA. In this case, the three-compound system H<sub>2</sub>SO<sub>4</sub>-base<sub>1</sub>-base<sub>2</sub> is correctly described by the lumping approach, that is, assuming only one base with a concentration of [base<sub>1</sub>] + [base<sub>2</sub>]. However, the sum approach may yield erroneous results. Finally, the DLPNO\_Kubečka multi-base data set is applied to test if other base combinations behave similarly to the H<sub>2</sub>SO<sub>4</sub>-NH<sub>3</sub>-DMA case, and to examine if there are different bases within the data set that can be lumped with reasonable accuracy.

The test cases assess the effects of the sum assumption for particle formation mechanisms that involve interacting species that are either (1) different, or (2) very similar in terms of their clustering efficiency. This addresses the question of potential errors in total particle formation rates in aerosol microphysics and large-scale atmospheric models, which generally apply the sum approach. Furthermore, tests with the DLPNO\_Kubečka data provide assessments on whether relatively similar bases are better represented by the sum or the lumping approach.

### 2.3.2 Representation of cluster concentrations for assessments of formation and growth dynamics.

Next, approaches to describe the cluster size distribution through a quasi-unary model are tested (Fig. 1b). First, the systems are reduced to a quasi-unary H<sub>2</sub>SO<sub>4</sub> system by assuming that H<sub>2</sub>SO<sub>4</sub>-containing clusters are equilibrated with respect to the base species. That is, each  $n$ -mer (H<sub>2</sub>SO<sub>4</sub>) <sub>$n$</sub> ,  $n = 1, 2, 3, \dots$  exists in its equilibrium distribution for which the relative concentration of an  $n$ -mer containing  $m$  base molecules is<sup>43</sup>



$$\frac{C_m}{\sum_{m_n=0}^{m_{n,\max}} C_{m_n}} = \frac{\left(\frac{P_{\text{base}}}{P_{\text{ref}}}\right)^m \exp\left(-\frac{\Delta G_m}{k_B T}\right)}{\sum_{m_n=0}^{m_{n,\max}} \left(\frac{P_{\text{base}}}{P_{\text{ref}}}\right)^{m_n} \exp\left(-\frac{\Delta G_{m_n}}{k_B T}\right)}, \quad (3)$$

where the summation goes over all possible numbers  $m_n$  of base molecules in the  $(\text{H}_2\text{SO}_4)_n$  cluster,  $\Delta G_m$  is the binding free energy of the cluster with  $m$  base molecules,  $P_{\text{ref}}$  is the pressure at which the binding free energies are calculated, and  $P_{\text{base}}$  is the partial pressure of the base vapor.  $k_B$  and  $T$  are the Boltzmann constant and temperature, respectively. Effective collision and evaporation rate constants between  $n$ -mers  $i$  and  $j$  become<sup>25</sup>

$$\beta_{i,j}^{\text{eff}} = \sum_{m_i=0}^{m_{i,\max}} \sum_{m_j=0}^{m_{j,\max}} \beta_{i,m_i;j,m_j} f_{i,m_i} f_{j,m_j}, \quad (4)$$

and

$$\gamma_{i+j \rightarrow i,j}^{\text{eff}} = \sum_{m_i=0}^{m_{i,\max}} \sum_{m_j=0}^{m_{j,\max}} \gamma_{i+j,m_i+m_j \rightarrow i,m_i;j,m_j} f_{i+j,m_i+m_j}. \quad (5)$$

In eqn (4) and (5),  $m_i$  and  $m_j$  are the numbers of base molecules in clusters  $i$  and  $j$ , respectively, and  $f_{i,m_i}$  is the relative fraction of the composition with  $m_i$  bases in the total concentration of cluster  $i$  (eqn (3)).  $\beta_{i,m_i;j,m_j}$  is the collision rate of composition  $(i, m_i)$  with composition  $(j, m_j)$ , and  $\gamma_{i+j,m_i+m_j \rightarrow i,m_i;j,m_j}$  is the evaporation rate of  $(i+j, m_i+m_j)$  into  $(i, m_i)$  and  $(j, m_j)$ . The equilibrium assumption is generally applied for water, for which the rate constants are typically orders of magnitude higher than for other species. Applicability on other stabilizing agents depends on whether the time scale of equilibration is sufficiently fast so that the distribution of different base contents  $m$  depends only on the thermodynamics (eqn (3)). The effects of the exact distribution on the effective collision constants  $\beta$  are typically minor, while the effective evaporation constants  $\gamma$  may be very sensitive to it.<sup>25</sup>

Second, a more simplified approach for a quasi-unary substance is applied by assuming that the base content of a given  $n$ -mer corresponds to the most stable composition. The substance is assumed to consist of  $\text{H}_2\text{SO}_4$ -base heterodimers, and cluster evaporation rates are set to those corresponding to the lowest total evaporation rate for a given  $n$  (ESI Section 2.2†). In this case, it is assumed that the base content can be represented by a single composition, which further simplifies the model treatment and interpretation of results.

The  $(\text{H}_2\text{SO}_4)_n$  concentrations of the quasi-unary systems are compared to the total concentrations  $\sum (\text{H}_2\text{SO}_4)_n \cdot (\text{base})_m$  of the explicit systems for each  $n$ -mer. Here, the cluster concentrations are compared instead of formation rates as such one-compound model substances are typically applied to interpret observations of nanoparticle formation and growth<sup>27,44</sup> and to assess cluster survival to larger sizes.<sup>45</sup>

## 3 Results and discussion

### 3.1 Simplifications in formation rates by assuming non-interacting or lumped species

The ratio of the summed formation rates  $\sum J_{2\text{-comp}}$  to the explicit three-component rate  $J_{3\text{-comp}}$  for  $\text{H}_2\text{SO}_4$ - $\text{NH}_3$ -DMA at typical

atmospheric base concentrations (Table 2) is shown in Fig. 2. The results are here shown only for the DLPNO\_Myllys input data, as the RICC2 and DLPNO\_Kubečka data give similar trends (see below). For such chemistry, the omission of base synergy can cause underprediction, but the errors are not very significant: the largest bias occurs at the combination of high  $[\text{NH}_3]$  (mainly above  $\sim 10^2$ - $10^3$  ppt) and low  $[\text{DMA}]$  (mainly below a few ppt) when the bases contribute together in an optimal way by enhanced initial cluster stabilization (DMA) and enhanced cluster growth ( $\text{NH}_3$ ). Under other conditions,  $[\text{NH}_3]$  is not high enough for significant enhancement, and/or the formation process becomes dominated by the stronger base DMA. While the high  $[\text{NH}_3]$ -low  $[\text{DMA}]$  conditions are common due to the different sources of the bases, the underprediction is typically clearly less than a factor of  $\sim 10$  at an average atmospheric scavenging sink of  $\text{CS} = 10^{-3} \text{ s}^{-1}$ . The bias increases with CS, and is here up to a factor of  $\sim 30$  under polluted conditions with  $\text{CS} = 10^{-2} \text{ s}^{-1}$  at which efficient growth becomes more critical.

The corresponding results for the RICC2 and DLPNO\_Kubečka data sets are given in the ESI.† While the absolute formation rates predicted by these data are mostly very different (Fig. S1†), the qualitative biases in the simplifications are consistent (Fig. S2 and S3†). The absolute biases are affected by the predicted quantitative cluster stabilities: lower stabilities (DLPNO\_Kubečka) mostly result in more bias and higher stabilities (RICC2) in less bias, with DLPNO\_Kubečka generally giving highest biases up to factors of  $\sim 40$ - $60$ . As effective cluster stabilities depend on vapor concentrations and temperature (*i.e.* molecular growth *vs.* evaporation), the conditions with largest biases may be somewhat shifted in the vapor concentration space or  $T$  space for the different data sets.

Fig. 3 shows  $\sum J_{2\text{-comp}}/J_{3\text{-comp}}$  for the synthetic test chemistries  $\text{H}_2\text{SO}_4$ -base<sub>1</sub>-base<sub>2</sub>, where the two bases have exactly the same properties. That is, clusters with equal amounts of  $\text{H}_2\text{SO}_4$  and base are assumed to be similar regardless of whether they contain base<sub>1</sub>, base<sub>2</sub>, or both, and the cases are thus correctly represented by lumping. Here the largest effects occur around the diagonal  $[\text{base}_1]:[\text{base}_2] = 1:1$ . For the weaker base represented by  $\text{NH}_3$ , the sum approach leads to underprediction by factors of  $\leq 5$  (260 K) and  $\leq 10$  (280 K, 300 K). For the stronger base DMA, underprediction (up to factors of  $\leq 10$ ) and possibly overprediction (up to a factor of  $\sim 2$ ) occur at low and high  $[\text{base}]$ , respectively. Base concentrations with the largest overprediction, however, correspond to very polluted environments. These trends are understood through the non-linear behavior of  $J$  as a function of vapor concentration, here  $[\text{base}]$ : at low  $[\text{base}]$ , increase in  $[\text{base}]$  causes a stronger increase in  $J$  than at high  $[\text{base}]$  where the system eventually becomes saturated with respect to  $[\text{base}]$ . That is, as the slope of  $J$  decreases with increasing  $[\text{base}]$ , the sum rates  $\sum J_{2\text{-comp}}$  change from too low to too high. Under saturated conditions, the rate is double-counted by  $\sum J_{2\text{-comp}}$  at  $\text{base}_1 \approx \text{base}_2$ , as the explicit formation rate remains the same upon doubling of base concentration. These patterns are similar for RICC2 (Fig. S4†) and DLPNO\_Kubečka (not shown; very similar to DLPNO\_Myllys), with some shifts in the vapor concentration space due to the



Table 2 Typical atmospheric concentration ranges of  $\text{NH}_3$  and amines from measurements

Base species	Clean and average environments (e.g. remote and marine sites)	Polluted environments (e.g. agricultural and urban sites, megacities)	References
$\text{NH}_3$	Tens of ppt to hundreds of ppt	$\sim 1$ to $\geq 10$ ppb	Nair and Yu (2020) <sup>46</sup> Croft <i>et al.</i> (2016) <sup>47</sup> Makkonen <i>et al.</i> (2014) <sup>48</sup> Xiao <i>et al.</i> (2021) <sup>9</sup>
Amines (MA, DMA, TMA)	<1 ppt	$\sim 1$ to $\geq 10$ ppt	Yu and Luo (2014) <sup>38</sup> Yao <i>et al.</i> (2018) <sup>10</sup>

differences in predicted stabilities. For example, as RICC2 predicts higher stabilities, the underprediction is often less pronounced compared to the other data sets and also slight overprediction may occur, since  $J$  is less sensitive to [base] and saturates at lower [base].

Assessments of lumping of different bases, conducted with the DLPNO\_Kubečka multi-base data, suggest that DMA and EDA are similar enough to be better described by lumping than summing. Fig. 4 shows  $J_{\text{lumped}}/J_{3\text{-comp}}$ , with  $J_{\text{lumped}}$  obtained

assuming the properties of DMA and a total base concentration of  $[\text{DMA}] + [\text{EDA}]$  (see Fig. S6† for summing). Errors due to lumping are minor when  $[\text{DMA}]$  and  $[\text{EDA}]$  are of the same order, and become notable only at  $[\text{DMA}]:[\text{EDA}] \ll 1$  when EDA is the dominating species.

However, here the lumping of TMA with DMA/EDA gives errors of similar order or higher than the sum approach (Fig. S7†). This is likely due to TMA being less efficient in cluster growth, as its three methyl groups cause steric

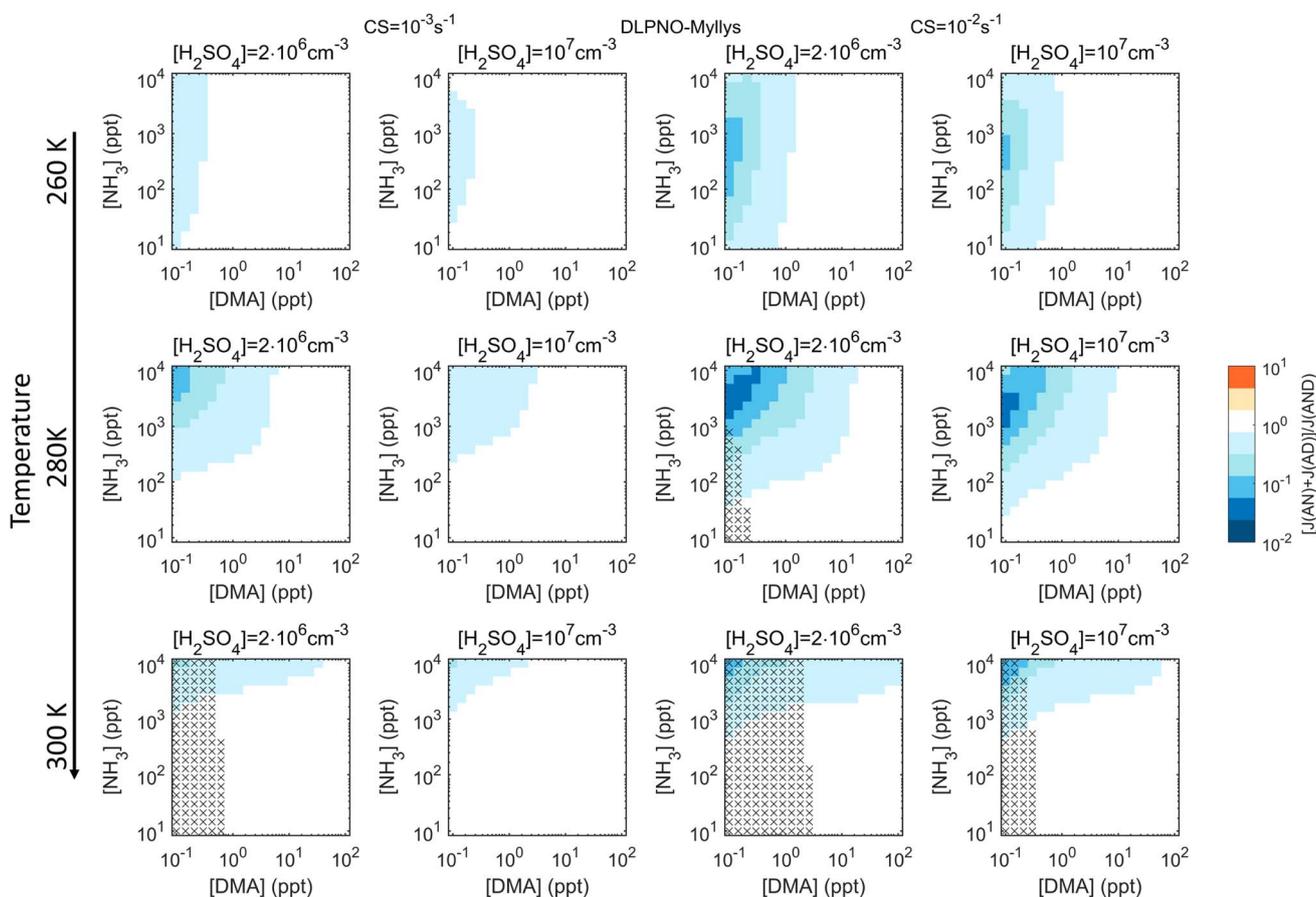


Fig. 2 Comparison of simulated particle formation rates assuming non-interacting base species ( $\text{H}_2\text{SO}_4\text{-NH}_3$  and  $\text{H}_2\text{SO}_4\text{-DMA}$ ;  $\sum J_{2\text{-comp}} = J(\text{AN}) + J(\text{AD})$ ) or the fully interacting three-component  $\text{H}_2\text{SO}_4\text{-NH}_3\text{-DMA}$  system ( $J_{3\text{-comp}} = J(\text{AND})$ ). Ratios of  $\sum J_{2\text{-comp}}/J_{3\text{-comp}}$  at different concentrations of  $\text{NH}_3$  and DMA at two different  $[\text{H}_2\text{SO}_4]$  and CS of  $10^{-3} \text{ s}^{-1}$  (two left columns) and  $10^{-2} \text{ s}^{-1}$  (two right columns) at  $T = 260, 280$  and  $300 \text{ K}$  are illustrated for simulations using the DLPNO\_Mylyls quantum chemistry data. The crosses mark conditions with both  $\sum J_{2\text{-comp}}$  and  $J_{3\text{-comp}} < 10^{-3} \text{ cm}^{-3} \text{ s}^{-1}$ .



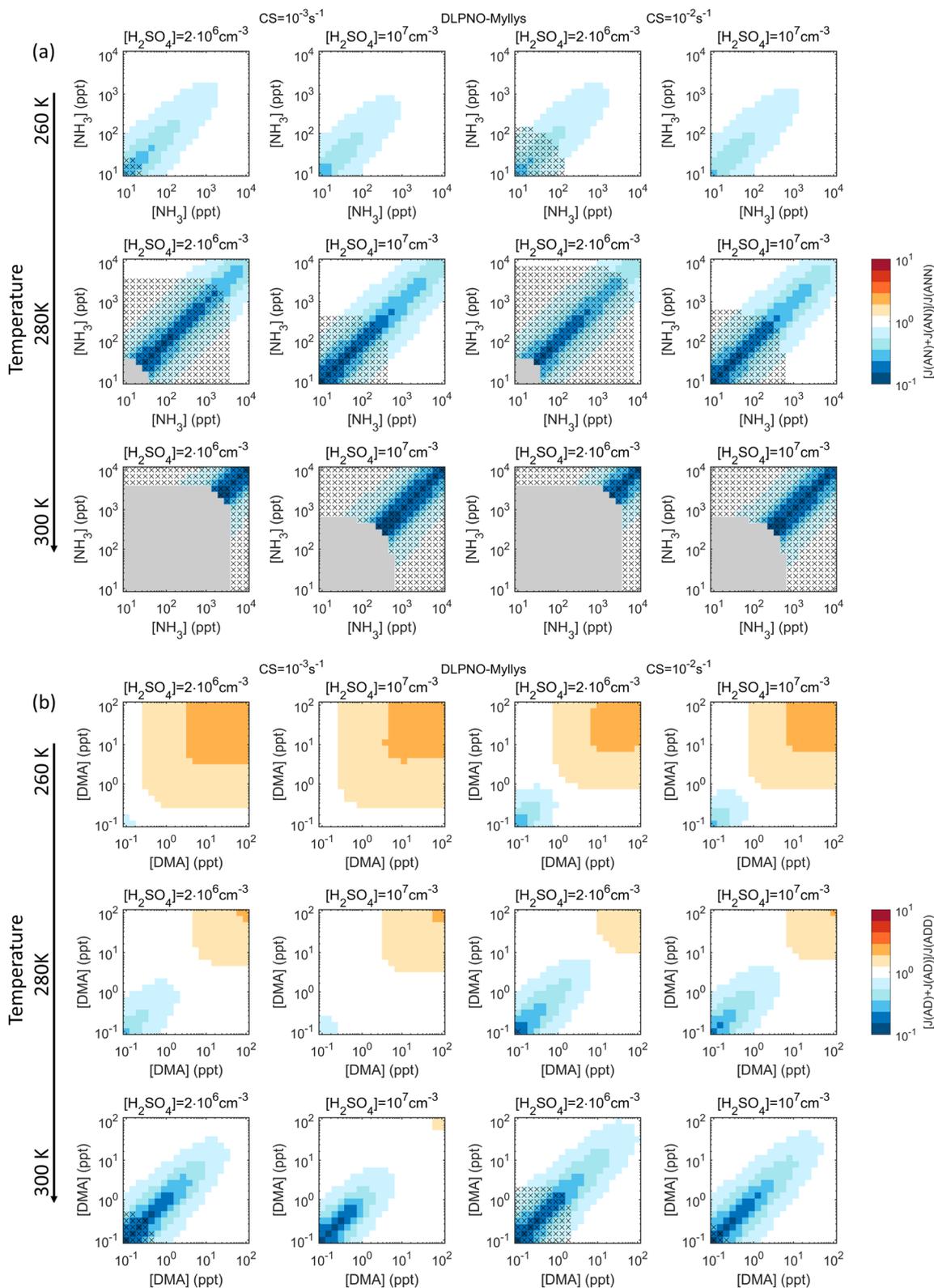


Fig. 3 Comparison of simulated particle formation rates assuming non-interacting base species ( $\sum J_{2\text{-comp}}$ ) or the fully interacting three-component system ( $J_{3\text{-comp}}$ ) for similar bases, here using "ideal cases" of bases with identical thermochemical data (either NH<sub>3</sub> data or DMA data). Ratios of  $\sum J_{2\text{-comp}}/J_{3\text{-comp}}$  at different concentrations of the bases at two different [H<sub>2</sub>SO<sub>4</sub>] and CS of 10<sup>-3</sup> s<sup>-1</sup> (two left columns) and 10<sup>-2</sup> s<sup>-1</sup> (two right columns) at T = 260, 280 and 300 K are illustrated for simulations using the DLPNO\_Myllys quantum chemistry data: (a) for weakly clustering bases (NH<sub>3</sub>); (b) for strongly clustering bases (DMA). Grey areas indicate that both  $\sum J_{2\text{-comp}}$  and  $J_{3\text{-comp}}$  are <10<sup>-10</sup> cm<sup>-3</sup> s<sup>-1</sup>. The crosses mark conditions with both  $\sum J_{2\text{-comp}}$  and  $J_{3\text{-comp}}$  <10<sup>-3</sup> cm<sup>-3</sup> s<sup>-1</sup>.



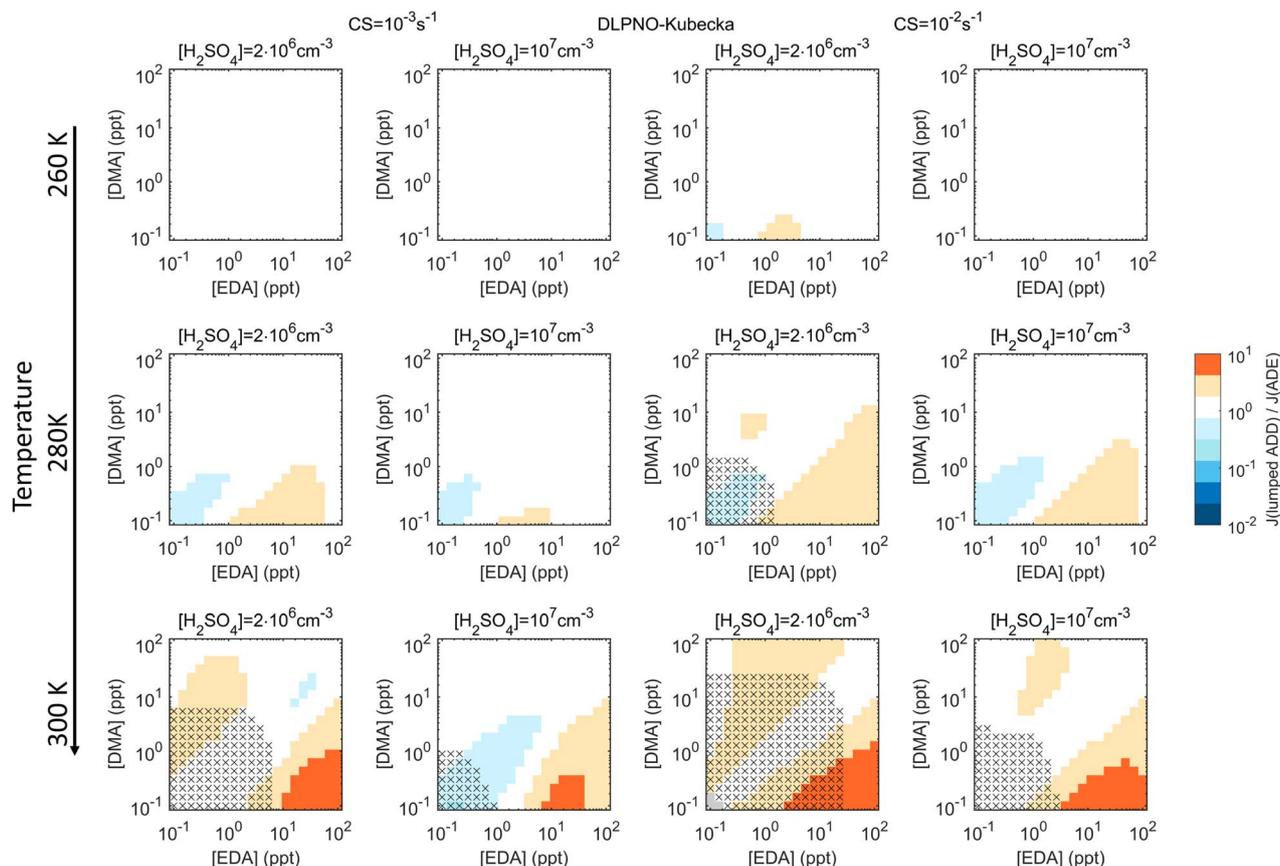


Fig. 4 Comparison of simulated particle formation rates using "lumped" base species, with EDA treated as DMA ( $J_{\text{lumped}}$ ), or the fully interacting three-component  $\text{H}_2\text{SO}_4$ -DMA-EDA system ( $J_{3\text{-comp}}$ ). Ratios  $J_{\text{lumped}}/J_{3\text{-comp}}$  at different concentrations of the bases at two different  $[\text{H}_2\text{SO}_4]$  and CS of  $10^{-3} \text{ s}^{-1}$  (two left columns) and  $10^{-2} \text{ s}^{-1}$  (two right columns) at  $T = 260, 280$  and  $300 \text{ K}$  are illustrated for simulations using the DLPNO\_Kubečka quantum chemistry data. Grey areas indicate that both  $J_{\text{lumped}}$  and  $J_{3\text{-comp}}$  are  $< 10^{-10} \text{ cm}^{-3} \text{ s}^{-1}$ . The crosses mark conditions with both  $J_{\text{lumped}}$  and  $J_{3\text{-comp}} < 10^{-3} \text{ cm}^{-3} \text{ s}^{-1}$ .

hindrance. MA lies between  $\text{NH}_3$  and the other amines and thus cannot be lumped with any of these bases (e.g. Fig. S8†). Comparisons of lumping vs. summing for DLPNO\_Kubečka are summarized in ESI Section 3†: for  $\text{H}_2\text{SO}_4$ - $\text{NH}_3$ -amine chemistries, trends in the biases of the sum approach are similar to those for  $\text{H}_2\text{SO}_4$ - $\text{NH}_3$ -DMA. For  $\text{H}_2\text{SO}_4$ -amine combinations, underprediction in summed  $J$  peaks when the amine concentrations are of the same order (except for MA for which biases peak at higher  $[\text{MA}]$ ). For lumped  $J$ , patterns are less regular and wider ranges of conditions become biased—both low and high—when the species are not lumpable (e.g. Fig. S7†).

To find indicators for similarity, we compared the formation rates  $J_{2\text{-comp}}$  of the isolated systems (ESI Section 3†). The differences in  $J_{2\text{-comp}}$  generally follow the patterns of the error in the lumped rate  $J_{\text{lumped}}$  (e.g. Fig. 4); small differences (less than a factor of 10 between  $\text{H}_2\text{SO}_4$ -DMA and  $\text{H}_2\text{SO}_4$ -EDA) coincide with low errors in  $J_{\text{lumped}}$ . This suggests that the rates can be used as an indicator to assess similarity. In addition, a simple theoretical proxy based on acid-base heterodimer stability<sup>36</sup> is in line with the comparisons (ESI Section 3†), and can thus serve as an additional measure.

The quantitative biases may be affected by the presence of ionic species. Ion effects are assessed using the RICC2 data set, which includes also charged clusters, assuming an average boundary layer ion production rate of  $3 \text{ cm}^{-3} \text{ s}^{-1}$ . For the  $\text{H}_2\text{SO}_4$ - $\text{NH}_3$ -DMA case, ions do not have notable effects on the bias at  $\text{CS} = 10^{-3} \text{ s}^{-1}$  but significantly decrease the maximum biases at  $\text{CS} = 10^{-2} \text{ s}^{-1}$ , especially at  $T = 300 \text{ K}$  (Fig. S2†). For the synthetic cases of identical bases, there are only minor effects for the strong base, but substantially less bias for the weaker base (Fig. S5†). This is due to  $J$  being less sensitive to vapor concentrations when the charge increases cluster stability. In general, ions are important for weaker clustering species and/or under conditions at which clusters are relatively less stable, such as low vapor concentrations and higher temperatures. Based on the RICC2 tests, ions can be expected to decrease the biases under such conditions.

To summarize, the commonly applied sum approach can be expected to give lower-limit assessments of the overall formation rate, here with the largest underprediction by factors of up to  $\sim 10$ – $100$  (Table S2†) occurring at vapor level combinations at which all species are able to contribute (i.e. high and low concentrations for weak and strong clustering agents,



respectively). The direction of the bias is consistent, as it follows from the omission of the total vapor concentration and similar behavior can thus be expected for other interacting chemistries. However, if model data and/or observations suggest very similar particle formation rates for stabilizing species with concentrations of the same magnitude, lumping is likely to involve less errors than summing of  $J$ .

### 3.2 Simplifications in cluster dynamics by assuming quasi-unary chemical systems

**3.2.1 Applying effective rate constants obtained by the equilibrium assumption.** To verify the applicability of the equilibrium approach in a case where the assumptions can be expected to be valid, we first performed test simulations for  $\text{H}_2\text{SO}_4\text{-H}_2\text{O}$  as water is an ideal example of such case. These tests show a perfect comparison (Fig. S9†), indicating that the implementation is robust and we can thus proceed to the  $\text{H}_2\text{SO}_4\text{-base}$  chemistries. Fig. 5 and 6 show the  $C_{\text{eq}}/C_{\text{expl}}$  ratios for each  $n$ -mer containing  $n$   $\text{H}_2\text{SO}_4$  molecules for  $\text{NH}_3$  and DMA, respectively, for the RICC2 and DLPNO\_Besel data. Here,  $C_{\text{eq}}$  is the concentration in the quasi-unary model system, and  $C_{\text{expl}} = \sum C_m$  is the total  $n$ -mer concentration summed over all base contents  $m$  in the explicit 2-compound simulation. The [base] ranges are here wider in order to recognize concentrations required for equilibration (in case it occurs).

The main conclusions can be summarized as follows: first, all data sets show a consistent tendency towards overprediction of  $n$ -mer concentrations by the equilibrium assumption by factors of up to  $\sim 10\text{-}10^2$  and even  $\geq 10^3$  especially at low [base]. The overprediction typically emerges and/or increases with increasing  $n$  and decreasing [base], although  $C_{\text{eq}}/C_{\text{expl}}$  may also exhibit maxima along the base axis, likely linked to shifts in the most abundant compositions (*i.e.* shifts in the growth paths; Fig. S10†). The positive bias can be understood through base dynamics: the equilibrium approach assumes that cluster formation and growth are not limited by the availability of base vapor, which is not the case at low [base], leading to overpredicted  $C_{\text{eq}}$ .

Second, the comparison is often seemingly better for the more strongly clustering species  $\text{H}_2\text{SO}_4\text{-DMA}$ . However, this is not due to the clustering kinetics (most importantly the evaporation rates; eqn (5)) being well described by the effective equilibrium rate constants. Instead, the evaporation rates of the  $\text{H}_2\text{SO}_4\text{-DMA}$  clusters are often so low compared to vapor collision rates that their exact values do not have major effects. This was tested by re-running the equilibrium simulations with all clusters stabilized by 2 kcal mol<sup>-1</sup>, which decreases all evaporation rates by approximately two orders of magnitude. The conditions at which the  $n$ -mer concentrations  $C_{\text{eq,test}}$  of these test runs are within a factor of two from  $C_{\text{eq}}$  are marked in the figures with crosses, implying that the quantitative evaporation rates are not very important. Such cases include some of the low-temperature conditions for  $\text{H}_2\text{SO}_4\text{-NH}_3$ , and most conditions for  $\text{H}_2\text{SO}_4\text{-DMA}$ . This type of condition can be experimentally identified based on high cluster concentrations: for the crossed areas,  $C_{\text{expl}} \gtrsim 10^4\text{-}10^5$  cm<sup>-3</sup>.

The figures can thus be interpreted as follows: under conditions with no bias and no crosses, the equilibrium assumption can be considered reasonable. Under conditions that have no bias but are marked with crosses, the quasi-unary approximation is reasonable but independent of the validity of the assumption, as well as of the exact rate constants. Under conditions with visible bias the clusters are not equilibrated with respect to base due to low [base] and/or evaporation effects, resulting in distorted predictions.

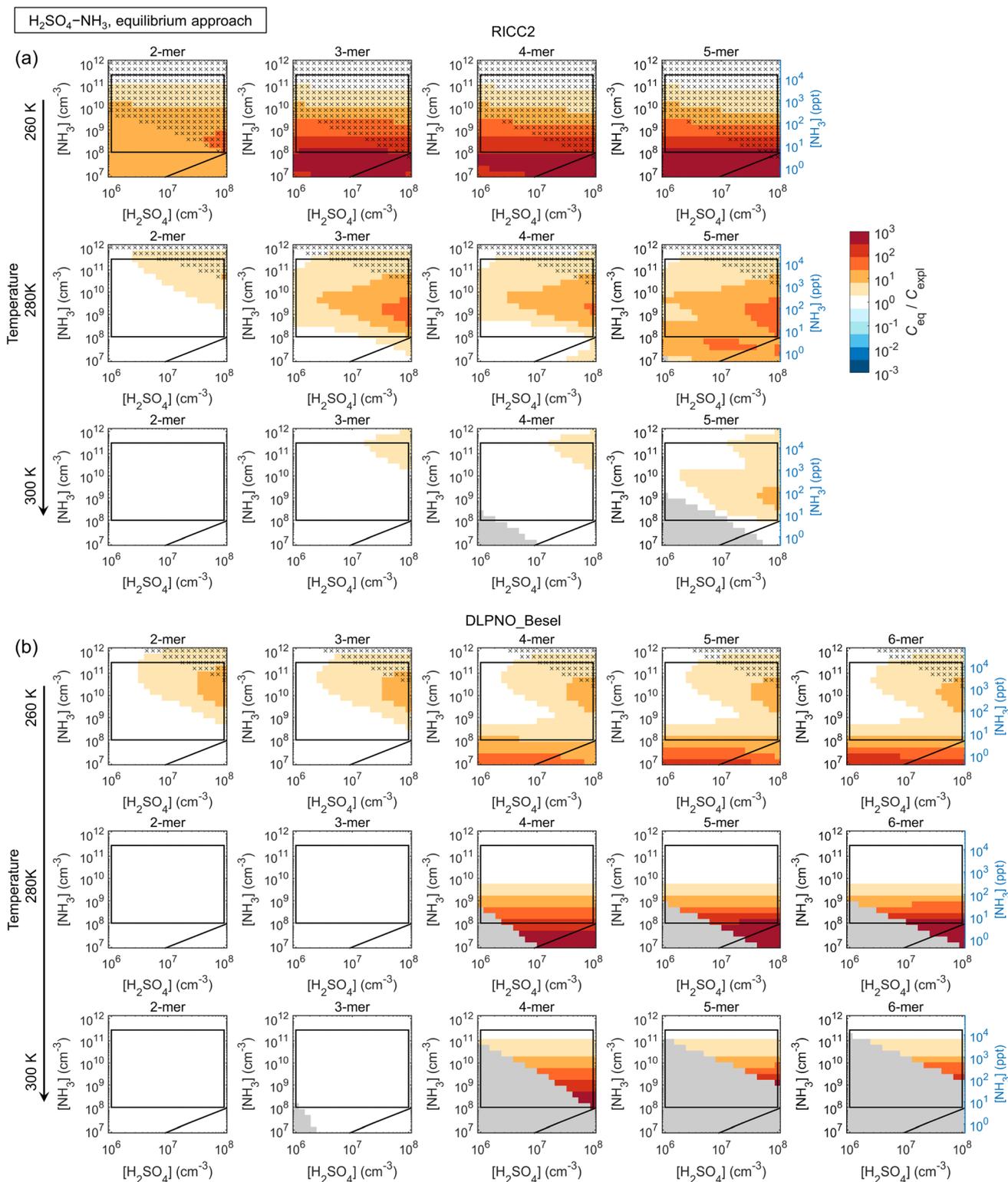
Also the weakly clustering  $\text{H}_2\text{SO}_4\text{-NH}_3$  case may show a good performance for the equilibrium assumption, but this depends strongly on the conditions, the cluster size ( $n$ -mer), and data set (*i.e.* the exact evaporation rates). For such chemistries, the assumption is thus more uncertain and unreliable. Overall, the effects of temperature on the quantitative  $C_{\text{eq}}/C_{\text{expl}}$  vary somewhat with the species and data sets. Increasing temperature can facilitate equilibration by increasing evaporation rates and thus shortening evaporation time scales. On the other hand, increasing  $T$  may also introduce notable evaporation in a chemical system that has insignificant evaporation and thereby less bias at lower  $T$ . The quantitative  $C_{\text{eq}}/C_{\text{expl}}$  ratios are also affected by the scavenging sink: here, we apply a reference sink of  $\text{CS} = 10^{-3}$  s<sup>-1</sup>, which is of the order of an average atmospheric sink or a wall loss sink in a clean chamber experiment.<sup>49</sup> Test simulations with  $10^{-4}$  s<sup>-1</sup> and  $10^{-2}$  s<sup>-1</sup> show increasing overprediction with increasing CS (not shown). In general, scavenging effects become more significant at lower cluster concentrations which can occur, for example, at higher  $T$ , which may also contribute to the observed temperature trends.

The DLPNO\_Kubečka data set that generally involves higher evaporation rates gives significantly more overprediction especially for the weaker clustering agents  $\text{NH}_3$  and MA, but the quasi-unary approach becomes less reliable also for DMA, as well as for EDA (not shown). Within DLPNO\_Kubečka, TMA at low  $T$  gives the least bias for the quasi-unary assumption (Fig. S11†). Simulations with  $\text{H}_2\text{SO}_4\text{-GUA}$  (DLPNO\_Myllys) are in line with these results: the guanidine-enhanced system with extremely low evaporation rates shows a good comparison even at high  $T$  (Fig. S12†).

The bias and its size-dependence are directly linked to cluster survival probability assessments. Fig. 7 shows an example of a cluster survival parameter, here defined as the ratio  $\text{SP} = C_n/C_{\text{ref}}$  where “ref” is a reference size smaller than  $n$ , for  $\text{H}_2\text{SO}_4\text{-NH}_3$  within DLPNO\_Besel data. Here the survival parameter is overpredicted by the equilibrium approach, converging to the explicit results only at very high  $[\text{NH}_3]$ . Quantitative survival estimates obtained by such a quasi-unary model should thus be interpreted with caution especially if  $\text{NH}_3$  is not in large excess.

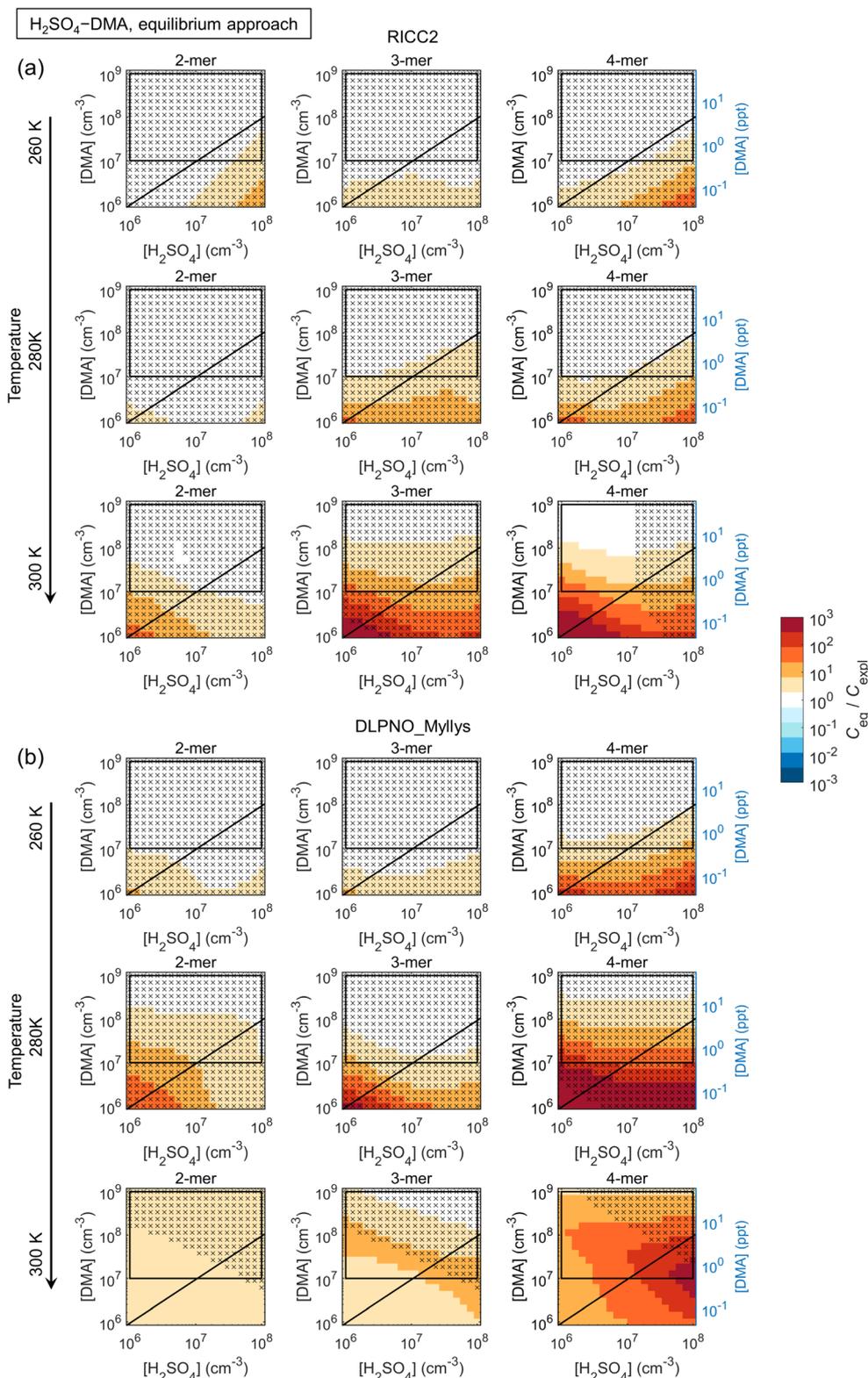
**3.2.2 Approximating cluster properties by the most stable multi-compound composition.** Despite the overprediction trend, the equilibrium approach cases suggest that the  $\text{H}_2\text{SO}_4\text{-base}$  chemistries can be reduced under specific conditions, and possibly even further simplified when the distributions with respect to base are dominated by a single composition (*e.g.* Fig. S10†). Fig. S13 and S14† show relative  $n$ -





**Fig. 5**  $C_{eq}/C_{exp}$  ratios of  $n$ -mer concentrations  $\sum_m([H_2SO_4]_n \cdot [base]_m)$  obtained by assuming equilibration with respect to the base species ( $C_{eq}$ ) and explicit multi-compound kinetics ( $C_{exp}$ ) for H<sub>2</sub>SO<sub>4</sub>-NH<sub>3</sub> chemistry, modeled applying (a) the RICC2 and (b) the DLPNO\_Besel quantum chemistry data set. The black diagonal lines correspond to [H<sub>2</sub>SO<sub>4</sub>]:[base] = 1:1, the rectangles give an approximative vapor concentration range for typical laboratory and ambient measurements, and the grey areas indicate  $n$ -mer concentrations below 10<sup>-10</sup> cm<sup>-3</sup>. The crosses mark the conditions at which  $C_{eq}$  is insensitive to the quantitative evaporation rates ( $C_{eq}$  within a factor of two upon cluster stabilization by 2 kcal mol<sup>-1</sup>).





**Fig. 6**  $C_{\text{eq}}/C_{\text{expl}}$  ratios of  $n$ -mer concentrations  $\sum_m ([\text{H}_2\text{SO}_4]_n \cdot [\text{base}]_m)$  obtained by assuming equilibration with respect to the base species ( $C_{\text{eq}}$ ) and explicit multi-compound kinetics ( $C_{\text{expl}}$ ) for  $\text{H}_2\text{SO}_4$ -DMA chemistry, modeled applying (a) the RICC2 and (b) the DLPNO\_Mylyls quantum chemistry data set. The black diagonal lines correspond to  $[\text{H}_2\text{SO}_4] : [\text{base}] = 1 : 1$ , and the rectangles give an approximative vapor concentration range for typical laboratory and ambient measurements. The crosses mark conditions at which  $C_{\text{eq}}$  is insensitive to the quantitative evaporation rates ( $C_{\text{eq}}$  within a factor of two upon cluster stabilization by  $2 \text{ kcal mol}^{-1}$ ).



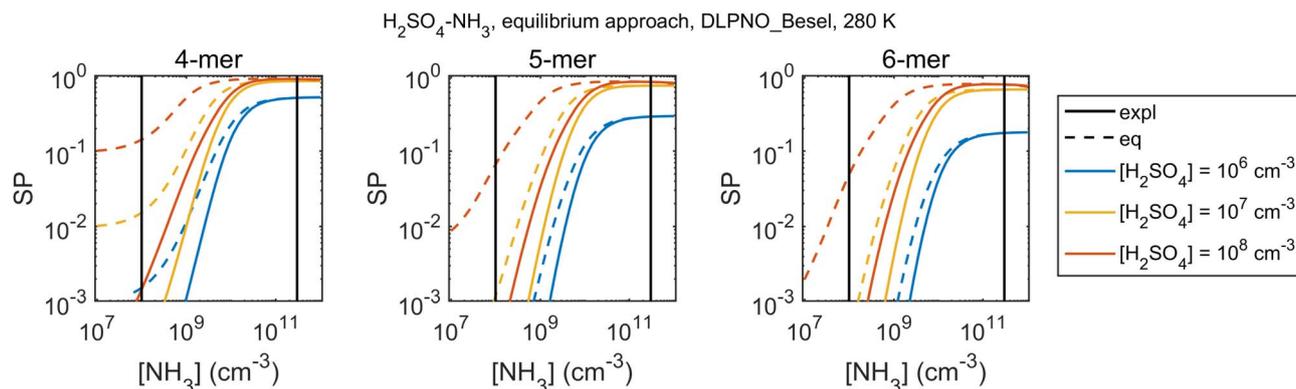


Fig. 7 Cluster survival parameter, defined as  $SP = C_n/C_{ref}$  where the reference size is 3-mer, for the explicit and equilibrium simulations for  $H_2SO_4-NH_3$  and DLPNO\_Besel quantum chemistry data sets at 280 K. The vertical lines give an approximative vapor concentration range for typical laboratory and ambient measurements.

mer concentrations as in Fig. 5 and 6, now applying the assumption of the most stable composition (Section 2.3.2). However, the simplified approach leads to substantially worse overprediction in case of weaker cluster formation: for  $H_2SO_4-NH_3$ , cluster concentrations are overpredicted by factors of up to  $10^3$  and above. The comparisons are better only at very high [base], mainly under conditions at which the simple quasi-unary results  $C_{st}$  are close to the equilibrium results  $C_{eq}$  and concurrently exact evaporation rates are unimportant (marked with plus signs and crosses, respectively). Similarly,  $C_{st}$  are close to  $C_{eq}$  for the low-evaporating case  $H_2SO_4-DMA$ , and positive bias mainly occurs at  $[H_2SO_4]:[DMA] \geq 1$  when clustering becomes base-limited, and at higher  $T$ .

Overall, while quasi-unary approximations are convenient for studying nanoparticle formation and growth dynamics, the results show that their applicability is limited to very

strongly interacting, *i.e.* low-evaporating, species and/or low temperatures at excess base vapor. For weaker clustering agents such as  $H_2SO_4-NH_3$ , which can be expected to be important on the global scale, the formation process is affected by multi-compound kinetics. Representing effective cluster properties through the distribution of base contents results in less bias than using a single composition, but the concentrations are often still significantly overpredicted. *Vice versa*, fitting a quasi-unary model to reproduce measurements cannot be expected to give rate constants that correspond to actual effective cluster properties. While the comparisons improve with increasing [base], the base concentrations at which the approximation becomes valid are often beyond typical atmospheric ranges, corresponding to very polluted environments. Application of quasi-unary approaches for such clustering agents generally leads to overprediction of nanoparticle concentrations, growth rates and survival

Table 3 Qualitative assessment of the applicability of different simplifications in multi-component particle formation modeling based on the representative  $H_2SO_4$ -base chemistries<sup>a,b</sup>

Model parameter and simplification	Applicability for different types of compounds	
	Different clustering properties	Similar clustering properties
Formation rate		
Summing of rates	✓*	✗**
Lumping of species	✗	✓**
Cluster concentrations		
	Weakly clustering compounds	Strongly clustering compounds
Quasi-unary approximation by equilibrium assumption	✗	✓***
Quasi-unary approximation by most stable composition	✗	✓***

<sup>a</sup> A checkmark implies that the simplification can in principle be applied without major errors, and a cross indicates that the simplification leads to distorted results and should not be applied. The applicability can depend on ambient conditions, as indicated in footnote *b*. <sup>b</sup> \*Underprediction is expected at  $[base_1] \gg [base_2]$ , where  $base_1$  and  $base_2$  are weaker and stronger stabilizers, respectively, and  $[base_2]$  is at the ppt or sub-ppt level. \*\*Lumping is reasonable especially at similar concentrations of the compounds; for summing, underprediction is expected when the base concentrations are of the same order:  $[base_1] \sim [base_2]$ . \*\*\*Applicable at low temperatures (here mainly  $\leq 280$  K) and excess concentration of the implicitly treated compound.



probabilities. Overpredicted concentrations may also result in, for instance, overpredicted contribution of cluster coagulation to nanoparticle growth.

## 4 Conclusions

The evaluations for H<sub>2</sub>SO<sub>4</sub>-base chemistries indicate that some typical approaches to simplify multi-component particle formation can often be applied without sacrificing much accuracy, while others are reasonable only for specific, limited conditions. The trends in the biases are consistent and similar for different quantum chemical data sets. We tested simplifications for different parameters that are central in particle formation modeling: formation rates that are used in atmospheric aerosol models, and molecular cluster properties that are applied to assess cluster growth and survival, as summarized in Table 3.

The main results for formation rates  $J$  are as follows:

- For stabilizing compounds base<sub>1</sub> and base<sub>2</sub> with very different clustering efficiencies, the sum assumption  $J(\text{H}_2\text{SO}_4\text{-base}_1\text{-base}_2) \approx J(\text{H}_2\text{SO}_4\text{-base}_1) + J(\text{H}_2\text{SO}_4\text{-base}_2)$  is biased towards underprediction for the combination of high [base<sub>1</sub>] and low [base<sub>2</sub>], where base<sub>1</sub> and base<sub>2</sub> are weaker and stronger stabilizers, respectively. The errors are minor especially for clean environments with low CS, but may increase to factors of up to ~10–100 as CS increases to  $10^{-3}\text{--}10^{-2} \text{ s}^{-1}$ .

- As the clustering efficiencies of the species approach each other, the underprediction tendency shifts towards similar concentrations [base<sub>1</sub>] ~ [base<sub>2</sub>].

- For very similar compounds,  $J$  is reasonably approximated by lumping the species through an effective vapor concentration [base] = [base<sub>1</sub>] + [base<sub>2</sub>]. The sum approach can result in errors of factors of up to ~10–100 when [base<sub>1</sub>] and [base<sub>2</sub>] are of the same order. Under- and overprediction by the sum approach typically occur at lower and higher [base], respectively; the overprediction is minor and mainly occurs for strongly clustering species when  $J$  is saturated with respect to base.

- Similarity can be assessed based on formation rates, and additionally by a theoretical proxy involving the free energies of acid-base heterodimers. Minor differences in these indicators, here approximately by factors of  $\leq 10$ , suggest similarity.

- A recent multi-base data set suggests that dimethylamine and ethylenediamine are similar enough to be lumped. The properties of trimethylamine are somewhat different compared to those of these two other amines, and lumping does not generally give better results than the sum approach.

- The presence of ions is expected to decrease the biases in the simplified formation rates. Charge enhances cluster binding, which reduces the differences in the relative stabilization efficiencies of different species.

Reducing H<sub>2</sub>SO<sub>4</sub>-base chemistries to quasi-unary H<sub>2</sub>SO<sub>4</sub> systems for modeling molecular cluster concentrations is, in contrast, much more uncertain:

- The quasi-unary approaches are mainly applicable for strong stabilizing species with low-evaporating clusters at vapor concentrations that are high enough to not limit the cluster

formation. Such a case is, for example, an efficiently clustering H<sub>2</sub>SO<sub>4</sub>-amine chemistry at [amine]:[H<sub>2</sub>SO<sub>4</sub>] > 1.

- Especially for weaker clustering species such as H<sub>2</sub>SO<sub>4</sub>-NH<sub>3</sub> that are common particle formation agents on the global scale, and/or for warm conditions, cluster concentrations are overpredicted by up to 1–2 orders of magnitude, or even more, in the quasi-unary frameworks.

- The biases in cluster concentrations affect particle survival probability assessments, which can also be biased high, as the overprediction often emerges and/or increases with cluster size.

Simplifications in multi-compound chemistries are highly desirable for convenient representation, and the tests suggest that they can be possible especially for formation rates. The present results provide assessments of how and when to simplify and of the possible biases, thereby contributing to improved implementation of input data and understanding of model uncertainties.

## Author contributions

Conceptualization, software (kinetic modeling), methodology: TO; formal analysis, investigation, visualization: RB and TO; resources (quantum chemical data and analysis): JK, NM, and JE; writing – original draft: TO; writing – review & editing: all authors.

## Conflicts of interest

There are no conflicts of interest to declare.

## Acknowledgements

T. O. and R. B. acknowledge the Swedish Research Council VR (grant no. 2019-04853) and the Swedish Research Council for Sustainable Development FORMAS (grant no. 2019-01433) for financial support. J. K. and J. E. thank the Independent Research Fund Denmark (grant no. 9064-00001B). N. M. thanks the Academy of Finland for funding (grant no. 347775). We thank the CSC-IT Center for Science in Espoo, Finland, for providing computational resources.

## References

- 1 V.-M. Kerminen, X. Chen, V. Vakkari, T. Petäjä, M. Kulmala and F. Bianchi, Atmospheric new particle formation and growth: review of field observations, *Environ. Res. Lett.*, 2018, **13**, 103003.
- 2 S. Fuzzi, U. Baltensperger, K. Carslaw, S. Decesari, H. Denier van der Gon, M. C. Facchini, D. Fowler, I. Koren, B. Langford, U. Lohmann, E. Nemitz, S. Pandis, I. Riipinen, Y. Rudich, M. Schaap, J. G. Slowik, D. V. Spracklen, E. Vignati, M. Wild, M. Williams and S. Gilardoni, Particulate matter, air quality and climate: lessons learned and future needs, *Atmos. Chem. Phys.*, 2015, **15**, 8217–8299.
- 3 H. Gordon, J. Kirkby, U. Baltensperger, F. Bianchi, M. Breitenlechner, J. Curtius, A. Dias, J. Dommen, N. M. Donahue, E. M. Dunne, J. Duplissy, S. Ehrhart,



- R. C. Flagan, C. Frege, C. Fuchs, A. Hansel, C. R. Hoyle, M. Kulmala, A. Kürten, K. Lehtipalo, V. Makhmutov, U. Molteni, M. P. Rissanen, Y. Stozhkov, J. Tröstl, G. Tsagkogeorgas, R. Wagner, C. Williamson, D. Wimmer, P. M. Winkler, C. Yan and K. S. Carslaw, Causes and importance of new particle formation in the present-day and preindustrial atmospheres, *J. Geophys. Res.: Atmos.*, 2017, **122**, 8739–8760.
- 4 S.-H. Lee, H. Gordon, H. Yu, K. Lehtipalo, R. Haley, Y. Li and R. Zhang, New Particle Formation in the Atmosphere: From Molecular Clusters to Global Climate, *J. Geophys. Res.: Atmos.*, 2019, **124**, 7098–7146.
- 5 B. R. Bzdek and J. P. Reid, Perspective: Aerosol microphysics: From molecules to the chemical physics of aerosols, *J. Chem. Phys.*, 2017, **147**, 220901.
- 6 K. Lehtipalo, C. Yan, L. Dada, F. Bianchi, M. Xiao, R. Wagner, D. Stolzenburg, L. R. Ahonen, A. Amorim, A. Baccarini, P. S. Bauer, B. Baumgartner, A. Bergen, A.-K. Bernhammer, M. Breitenlechner, S. Brilke, A. Buchholz, S. B. Mazon, D. Chen, X. Chen, A. Dias, J. Dommen, D. C. Draper, J. Duplissy, M. Ehn, H. Finkenzeller, L. Fischer, C. Frege, C. Fuchs, O. Garmash, H. Gordon, J. Hakala, X. He, L. Heikkinen, M. Heinritzi, J. C. Helm, V. Hofbauer, C. R. Hoyle, T. Jokinen, J. Kangasluoma, V.-M. Kerminen, C. Kim, J. Kirkby, J. Kontkanen, A. Kürten, M. J. Lawler, H. Mai, S. Mathot, R. L. Mauldin, U. Molteni, L. Nischman, W. Nie, T. Nieminen, A. Ojdanic, A. Onnela, M. Passananti, T. Petäjä, F. Piel, V. Pospisilova, L. L. J. Quéléver, M. P. Rissanen, C. Rose, N. Sarnela, S. Schallhart, S. Schuchmann, K. Sengupta, M. Simon, M. Sipilä, C. Tauber, A. Tomé, J. Tröstl, O. Väisänen, A. L. Vogel, R. Volkamer, A. C. Wagner, M. Wang, L. Weitz, D. Wimmer, P. Ye, A. Ylisirniö, Q. Zha, K. S. Carslaw, J. Curtius, N. M. Donahue, R. C. Flagan, A. Hansel, I. Riipinen, A. Virtanen, P. M. Winkler, U. Baltensperger, M. Kulmala and D. R. Worsnop, Multicomponent new particle formation from sulfuric acid, ammonia, and biogenic vapors, *Sci. Adv.*, 2018, **4**, eaau5363.
- 7 J. Elm, Unexpected Growth Coordinate in Large Clusters Consisting of Sulfuric Acid and C<sub>8</sub>H<sub>12</sub>O<sub>6</sub> Tricarboxylic Acid, *J. Phys. Chem. A*, 2019, **123**, 3170–3175.
- 8 M. Kulmala, J. Kontkanen, H. Junninen, K. Lehtipalo, H. E. Manninen, T. Nieminen, T. Petäjä, M. Sipilä, S. Schobesberger, P. Rantala, A. Franchin, T. Jokinen, E. Järvinen, M. Äijälä, J. Kangasluoma, J. Hakala, P. P. Aalto, P. Paasonen, J. Mikkilä, J. Vanhanen, J. Aalto, H. Hakola, U. Makkonen, T. Ruuskanen, R. L. Mauldin, J. Duplissy, H. Vehkamäki, J. Bäck, A. Kortelainen, I. Riipinen, T. Kurtén, M. V. Johnston, J. N. Smith, M. Ehn, T. F. Mentel, K. E. J. Lehtinen, A. Laaksonen, V.-M. Kerminen and D. R. Worsnop, Direct Observations of Atmospheric Aerosol Nucleation, *Science*, 2013, **339**, 943–946.
- 9 M. Xiao, C. R. Hoyle, L. Dada, D. Stolzenburg, A. Kürten, M. Wang, H. Lamkaddam, O. Garmash, B. Mentler, U. Molteni, A. Baccarini, M. Simon, X.-C. He, K. Lehtipalo, L. R. Ahonen, R. Baalbaki, P. S. Bauer, L. Beck, D. Bell, F. Bianchi, S. Brilke, D. Chen, R. Chiu, A. Dias, J. Duplissy, H. Finkenzeller, H. Gordon, V. Hofbauer, C. Kim, T. K. Koenig, J. Lampilahti, C. P. Lee, Z. Li, H. Mai, V. Makhmutov, H. E. Manninen, R. Marten, S. Mathot, R. L. Mauldin, W. Nie, A. Onnela, E. Partoll, T. Petäjä, J. Pfeifer, V. Pospisilova, L. L. J. Quéléver, M. Rissanen, S. Schobesberger, S. Schuchmann, Y. Stozhkov, C. Tauber, Y. J. Tham, A. Tomé, M. Vazquez-Pufleau, A. C. Wagner, R. Wagner, Y. Wang, L. Weitz, D. Wimmer, Y. Wu, C. Yan, P. Ye, Q. Ye, Q. Zha, X. Zhou, A. Amorim, K. Carslaw, J. Curtius, A. Hansel, R. Volkamer, P. M. Winkler, R. C. Flagan, M. Kulmala, D. R. Worsnop, J. Kirkby, N. M. Donahue, U. Baltensperger, I. El Haddad and J. Dommen, The driving factors of new particle formation and growth in the polluted boundary layer, *Atmos. Chem. Phys.*, 2021, **21**, 14275–14291.
- 10 L. Yao, O. Garmash, F. Bianchi, J. Zheng, C. Yan, J. Kontkanen, H. Junninen, S. B. Mazon, M. Ehn, P. Paasonen, M. Sipilä, M. Wang, X. Wang, S. Xiao, H. Chen, Y. Lu, B. Zhang, D. Wang, Q. Fu, F. Geng, L. Li, H. Wang, L. Qiao, X. Yang, J. Chen, V.-M. Kerminen, T. Petäjä, D. R. Worsnop, M. Kulmala and L. Wang, Atmospheric new particle formation from sulfuric acid and amines in a Chinese megacity, *Science*, 2018, **361**, 278–281.
- 11 W. A. Glasoe, K. Volz, B. Panta, N. Freshour, R. Bachman, D. R. Hanson, P. H. McMurry and C. Jen, Sulfuric acid nucleation: an experimental study of the effect of seven bases, *J. Geophys. Res.: Atmos.*, 2015, **120**, 1933–1950.
- 12 C. N. Jen, P. H. McMurry and D. R. Hanson, Stabilization of sulfuric acid dimers by ammonia, methylamine, dimethylamine, and trimethylamine, *J. Geophys. Res.: Atmos.*, 2014, **119**, 7502–7514.
- 13 J. Almeida, S. Schobesberger, A. Kürten, I. K. Ortega, O. Kupiainen-Määttä, A. P. Praplan, A. Adamov, A. Amorim, F. Bianchi, M. Breitenlechner, A. David, J. Dommen, N. M. Donahue, A. Downard, E. Dunne, J. Duplissy, S. Ehrhart, R. C. Flagan, A. Franchin, R. Guida, J. Hakala, A. Hansel, M. Heinritzi, H. Henschel, T. Jokinen, H. Junninen, M. Kajos, J. Kangasluoma, H. Keskinen, A. Kupc, T. Kurtén, A. N. Kvashin, A. Laaksonen, K. Lehtipalo, M. Leiminger, J. Leppä, V. Loukonen, V. Makhmutov, S. Mathot, M. J. McGrath, T. Nieminen, T. Olenius, A. Onnela, T. Petäjä, F. Riccobono, I. Riipinen, M. Rissanen, L. Rondo, T. Ruuskanen, F. D. Santos, N. Sarnela, S. Schallhart, R. Schnitzhofer, J. H. Seinfeld, M. Simon, M. Sipilä, Y. Stozhkov, F. Stratmann, A. Tomé, J. Tröstl, G. Tsagkogeorgas, P. Vaattovaara, Y. Viisanen, A. Virtanen, A. Vrtala, P. E. Wagner, E. Weingartner, H. Wex, C. Williamson, D. Wimmer, P. Ye, T. Yli-Juuti, K. S. Carslaw, M. Kulmala, J. Curtius, U. Baltensperger, D. R. Worsnop, H. Vehkamäki and J. Kirkby, Molecular understanding of sulphuric acid-amine particle nucleation in the atmosphere, *Nature*, 2013, **502**, 359–363.
- 14 C. N. Jen, R. Bachman, J. Zhao, P. H. McMurry and D. R. Hanson, Diamine-sulfuric acid reactions are a potent source of new particle formation, *Geophys. Res. Lett.*, 2016, **43**, 867–873.



- 15 Y. You, V. P. Kanawade, J. A. de Gouw, A. B. Guenther, S. Madronich, M. R. Sierra-Hernández, M. Lawler, J. N. Smith, S. Takahama, G. Ruggeri, A. Koss, K. Olson, K. Baumann, R. J. Weber, A. Nenes, H. Guo, E. S. Edgerton, L. Porcelli, W. H. Brune, A. H. Goldstein and S.-H. Lee, Atmospheric amines and ammonia measured with a chemical ionization mass spectrometer (CIMS), *Atmos. Chem. Phys.*, 2014, **14**, 12181–12194.
- 16 J. Zheng, Y. Ma, M. Chen, Q. Zhang, L. Wang, A. F. Khalizov, L. Yao, Z. Wang, X. Wang and L. Chen, Measurement of atmospheric amines and ammonia using the high resolution time-of-flight chemical ionization mass spectrometry, *Atmos. Environ.*, 2015, **102**, 249–259.
- 17 J. Mao, F. Yu, Y. Zhang, J. An, L. Wang, J. Zheng, L. Yao, G. Luo, W. Ma, Q. Yu, C. Huang, L. Li and L. Chen, High-resolution modeling of gaseous methylamines over a polluted region in China: source-dependent emissions and implications of spatial variations, *Atmos. Chem. Phys.*, 2018, **18**, 7933–7950.
- 18 L. Dada, K. Lehtipalo, J. Kontkanen, T. Nieminen, R. Baalbaki, L. Ahonen, J. Duplissy, C. Yan, B. Chu, T. Petäjä, K. Lehtinen, V.-M. Kerminen, M. Kulmala and J. Kangasluoma, Formation and growth of sub-3-nm aerosol particles in experimental chambers, *Nat. Protoc.*, 2020, **15**, 1013–1040.
- 19 J. Elm, J. Kubečka, V. Besel, M. J. Jääskeläinen, R. Halonen, T. Kurtén and H. Vehkamäki, Modeling the formation and growth of atmospheric molecular clusters: A review, *J. Aerosol Sci.*, 2020, **149**, 105621.
- 20 T. Olenius, O. Kupiainen-Määttä, I. K. Ortega, T. Kurtén and H. Vehkamäki, Free energy barrier in the growth of sulfuric acid–ammonia and sulfuric acid–dimethylamine clusters, *J. Chem. Phys.*, 2013, **139**, 084312.
- 21 J. Kontkanen, D. Stolzenburg, T. Olenius, C. Yan, L. Dada, L. Ahonen, M. Simon, K. Lehtipalo and I. Riipinen, What controls the observed size-dependency of the growth rates of sub-10 nm atmospheric particles?, *Environ. Sci.: Atmos.*, 2022, **2**, 449–468.
- 22 E. M. Dunne, H. Gordon, A. Kürten, J. Almeida, J. Duplissy, C. Williamson, I. K. Ortega, K. J. Pringle, A. Adamov, U. Baltensperger, P. Barmet, F. Benduhn, F. Bianchi, M. Breitenlechner, A. Clarke, J. Curtius, J. Dommen, N. M. Donahue, S. Ehrhart, R. C. Flagan, A. Franchin, R. Guida, J. Hakala, A. Hansel, M. Heinritzi, T. Jokinen, J. Kangasluoma, J. Kirkby, M. Kulmala, A. Kupc, M. J. Lawler, K. Lehtipalo, V. Makhmutov, G. Mann, S. Mathot, J. Merikanto, P. Miettinen, A. Nenes, A. Onnela, A. Rap, C. L. S. Reddington, F. Riccobono, N. A. D. Richards, M. P. Rissanen, L. Rondo, N. Sarnela, S. Schobesberger, K. Sengupta, M. Simon, M. Sipilä, J. N. Smith, Y. Stozkhov, A. Tomé, J. Tröstl, P. E. Wagner, D. Wimmer, P. M. Winkler, D. R. Worsnop and K. S. Carslaw, Global atmospheric particle formation from CERN CLOUD measurements, *Science*, 2016, **354**, 1119–1124.
- 23 T. Bergman, A. Laaksonen, H. Korhonen, J. Malila, E. M. Dunne, T. Mielonen, K. E. J. Lehtinen, T. Kühn, A. Arola and H. Kokkola, Geographical and diurnal features of amine-enhanced boundary layer nucleation, *J. Geophys. Res.: Atmos.*, 2015, **120**, 9606–9624.
- 24 F. Gelbard and J. H. Seinfeld, Simulation of multicomponent aerosol dynamics, *J. Colloid Interface Sci.*, 1980, **78**, 485–501.
- 25 H. Henschel, T. Kurtén and H. Vehkamäki, 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.
- 26 F. Yu, A. B. Nadykto, J. Herb, G. Luo, K. M. Nazarenko and L. A. Uvarova, H<sub>2</sub>SO<sub>4</sub>–H<sub>2</sub>O–NH<sub>3</sub> ternary ion-mediated nucleation (TIMN): kinetic-based model and comparison with CLOUD measurements, *Atmos. Chem. Phys.*, 2018, **18**, 17451–17474.
- 27 A. Kürten, T. Jokinen, M. Simon, M. Sipilä, N. Sarnela, H. Junninen, A. Adamov, J. Almeida, A. Amorim, F. Bianchi, M. Breitenlechner, J. Dommen, N. M. Donahue, J. Duplissy, S. Ehrhart, R. C. Flagan, A. Franchin, J. Hakala, A. Hansel, M. Heinritzi, M. Hutterli, J. Kangasluoma, J. Kirkby, A. Laaksonen, K. Lehtipalo, M. Leiminger, V. Makhmutov, S. Mathot, A. Onnela, T. Petäjä, A. P. Praplan, F. Riccobono, M. P. Rissanen, L. Rondo, S. Schobesberger, J. H. Seinfeld, G. Steiner, A. Tomé, J. Tröstl, P. M. Winkler, C. Williamson, D. Wimmer, P. Ye, U. Baltensperger, K. S. Carslaw, M. Kulmala, D. R. Worsnop and J. Curtius, Neutral molecular cluster formation of sulfuric acid–dimethylamine observed in real time under atmospheric conditions, *Proc. Natl. Acad. Sci.*, 2014, **111**, 15019–15024.
- 28 J. S. Johnson and C. N. Jen, A sulfuric acid nucleation potential model for the atmosphere, *Atmos. Chem. Phys.*, 2022, **22**, 8287–8297.
- 29 I. K. Ortega, O. Kupiainen, T. Kurtén, T. Olenius, O. Wilkman, M. J. McGrath, V. Loukonen and H. Vehkamäki, From quantum chemical formation free energies to evaporation rates, *Atmos. Chem. Phys.*, 2012, **12**, 225–235.
- 30 N. Myllys, J. Kubečka, V. Besel, D. Alfaouri, T. Olenius, J. N. Smith and M. Passananti, Role of base strength, cluster structure and charge in sulfuric-acid-driven particle formation, *Atmos. Chem. Phys.*, 2019, **19**, 9753–9768.
- 31 T. Olenius, *Atmospheric Cluster Dynamics Code: Software Repository*, <https://github.com/tolenius/ACDC>, 2021.
- 32 N. Myllys, S. Chee, T. Olenius, M. Lawler and J. Smith, Molecular-Level Understanding of Synergistic Effects in Sulfuric Acid–Amine–Ammonia Mixed Clusters, *J. Phys. Chem. A*, 2019, **123**, 2420–2425.
- 33 B. Temelso, E. F. Morrison, D. L. Speer, B. C. Cao, N. Appiah-Padi, G. Kim and G. C. Shields, Effect of Mixing Ammonia and Alkylamines on Sulfate Aerosol Formation, *J. Phys. Chem. A*, 2018, **122**, 1612–1622.
- 34 H. Li, A. Ning, J. Zhong, H. Zhang, L. Liu, Y. Zhang, X. Zhang, X. C. Zeng and H. He, Influence of atmospheric conditions on sulfuric acid–dimethylamine–ammonia-based new particle formation, *Chemosphere*, 2020, **245**, 125554.
- 35 N. Myllys, T. Ponkkonen, M. Passananti, J. Elm, H. Vehkamäki and T. Olenius, Guanidine: A Highly



- Efficient Stabilizer in Atmospheric New-Particle Formation, *J. Phys. Chem. A*, 2018, **122**, 4717–4729.
- 36 S. Chee, K. Barsanti, J. N. Smith and N. Myllys, A predictive model for salt nanoparticle formation using heterodimer stability calculations, *Atmos. Chem. Phys.*, 2021, **21**, 11637–11654.
- 37 J. Kubečka, I. Neeffjes, V. Besel, F. Qiao, H.-B. Xie and J. Elm, Atmospheric sulfuric acid–multi-base new particle formation revealed through quantum chemistry enhanced by machine learning, *Environ. Sci. Technol.*, under review.
- 38 F. Yu and G. Luo, Modeling of gaseous methylamines in the global atmosphere: impacts of oxidation and aerosol uptake, *Atmos. Chem. Phys.*, 2014, **14**, 12455–12464.
- 39 A. Kürten, F. Bianchi, J. Almeida, O. Kupiainen-Määttä, E. M. Dunne, J. Duplissy, C. Williamson, P. Barmet, M. Breitenlechner, J. Dommen, N. M. Donahue, R. C. Flagan, A. Franchin, H. Gordon, J. Hakala, A. Hansel, M. Heinritzi, L. Ickes, T. Jokinen, J. Kangasluoma, J. Kim, J. Kirkby, A. Kupc, K. Lehtipalo, M. Leiminger, V. Makhmutov, A. Onnela, I. K. Ortega, T. Petäjä, A. P. Praplan, F. Riccobono, M. P. Rissanen, L. Rondo, R. Schnitzhofer, S. Schobesberger, J. N. Smith, G. Steiner, Y. Stozhkov, A. Tomé, J. Tröstl, G. Tsagkogeorgas, P. E. Wagner, D. Wimmer, P. Ye, U. Baltensperger, K. Carslaw, M. Kulmala and J. Curtius, Experimental particle formation rates spanning tropospheric sulfuric acid and ammonia abundances, ion production rates, and temperatures, *J. Geophys. Res.: Atmos.*, 2016, **121**(12), 377–412.
- 40 V. Besel, J. Kubečka, T. Kurtén and H. Vehkamäki, Impact of Quantum Chemistry Parameter Choices and Cluster Distribution Model Settings on Modeled Atmospheric Particle Formation Rates, *J. Phys. Chem. A*, 2020, **124**, 5931–5943.
- 41 K. E. J. Lehtinen, M. Dal Maso, M. Kulmala and V.-M. Kerminen, Estimating nucleation rates from apparent particle formation rates and vice versa: Revised formulation of the Kerminen–Kulmala equation, *J. Aerosol Sci.*, 2007, **38**, 988–994.
- 42 L. Rondo, S. Ehrhart, A. Kürten, A. Adamov, F. Bianchi, M. Breitenlechner, J. Duplissy, A. Franchin, J. Dommen, N. M. Donahue, E. M. Dunne, R. C. Flagan, J. Hakala, A. Hansel, H. Keskinen, J. Kim, T. Jokinen, K. Lehtipalo, M. Leiminger, A. Praplan, F. Riccobono, M. P. Rissanen, N. Sarnela, S. Schobesberger, M. Simon, M. Sipilä, J. N. Smith, A. Tomé, J. Tröstl, G. Tsagkogeorgas, P. Vaattovaara, P. M. Winkler, C. Williamson, D. Wimmer, U. Baltensperger, J. Kirkby, M. Kulmala, T. Petäjä, D. R. Worsnop and J. Curtius, Effect of dimethylamine on the gas phase sulfuric acid concentration measured by Chemical Ionization Mass Spectrometry, *J. Geophys. Res.: Atmos.*, 2016, **121**, 3036–3049.
- 43 A. Jaeger-Voirol, P. Mirabel and H. Reiss, Hydrates in supersaturated binary sulfuric acid–water vapor: A reexamination, *J. Chem. Phys.*, 1987, **87**, 4849–4852.
- 44 C. Li and P. H. McMurry, Errors in nanoparticle growth rates inferred from measurements in chemically reacting aerosol systems, *Atmos. Chem. Phys.*, 2018, **18**, 8979–8993.
- 45 S. Tuovinen, R. Cai, V.-M. Kerminen, J. Jiang, C. Yan, M. Kulmala and J. Kontkanen, Survival probabilities of atmospheric particles: comparison based on theory, cluster population simulations, and observations in Beijing, *Atmos. Chem. Phys.*, 2022, **22**, 15071–15091.
- 46 A. A. Nair and F. Yu, Quantification of Atmospheric Ammonia Concentrations: A Review of Its Measurement and Modeling, *Atmosphere*, 2020, **11**, 1092.
- 47 B. Croft, G. R. Wentworth, R. V. Martin, W. R. Leitch, J. G. Murphy, B. N. Murphy, J. K. Kodros, J. P. D. Abbatt and J. R. Pierce, Contribution of Arctic seabird-colony ammonia to atmospheric particles and cloud-albedo radiative effect, *Nat. Commun.*, 2016, **7**, 13444.
- 48 U. Makkonen, A. Virkkula, H. Hellén, M. Hemmilä, J. Sund, M. Äijälä, M. Ehn, H. Junninen, P. Keronen, T. Petäjä, D. R. Worsnop, M. Kulmala and H. Hakola, Semi-continuous gas and inorganic aerosol measurements at a boreal forest site: seasonal and diurnal cycles of NH<sub>3</sub>, HONO and HNO<sub>3</sub>, *Boreal Environ. Res.*, 2014, **19**(suppl. B), 311–328.
- 49 A. Kürten, C. Williamson, J. Almeida, J. Kirkby and J. Curtius, On the derivation of particle nucleation rates from experimental formation rates, *Atmos. Chem. Phys.*, 2015, **15**, 4063–4075.

