Breakups are Complicated: An Efficient Representation of Collisional Breakup in the Superdroplet Method

Emily de Jong1, J. Ben Mackay2,3, Anna Jaruga2, and Sylwester Arabas4

1Department of Mechanical and Civil Engineering, California Institute of Technology, Pasadena, CA, USA
2Department of Environmental Science and Engineering, California Institute of Technology, Pasadena, CA, USA
3Scripps Institution of Oceanography, San Diego, CA, USA
4Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland

Correspondence: Emily de Jong (edejong@caltech.edu)

Abstract. A key constraint of particle-based methods for modeling cloud microphysics is the conservation of total particle number, which is required for computational tractability. The process of collisional breakup poses a particular challenge to this framework, as breakup events often produce many droplet fragments of varying sizes, which would require creating new particles in the system. This work introduces a representation of collisional breakup in the so-called "superdroplet" method which conserves the total number of superdroplets in the system. This representation extends an existing stochastic collisional-coalescence scheme and samples from a fragment-size distribution in an additional Monte Carlo step. This method is demonstrated in a set of idealized box model and single-column warm-rain simulations. We further discuss the effects of the breakup dynamic and fragment-size distribution on the particle size distribution, hydrometeor population, and microphysical process rates. This representation of collisional breakup is able to produce a stationary particle-size distribution, in which breakup and coalescence rates are approximately equal, and it recovers expected behavior such as precipitation suppression in the column model. Furthermore, representing breakup has potential benefits that extend beyond warm rain processes, such as the ability to capture mechanisms of secondary ice production in the superdroplet method. The breakup algorithm presented here contributes to an open-source pythonic implementation of the superdroplet method, ‘PySDM’, which will facilitate future research using particle-based microphysics.

1 Introduction

The superdroplet method (SDM) for cloud microphysics is a high-fidelity particle-based (Lagrangian) representation of aerosols and hydrometeors that offers notable advantages over traditional bulk and bin microphysics schemes. Particle-based methods were initially used in atmospheric simulations to represent ice nucleation (Paoli et al., 2004; Jensen and Pfister, 2004; Shirgaonkar and Lele, 2006; Sölch and Kärcher, 2010), and were later extended to study aerosol indirect effects with a superdroplet approach (Andrejczuk et al., 2008) in which each "superdroplet" represents a multiplicity of modeled particles with identical attributes, such as size and chemical properties. Later, the SDM was extended to include a stochastic representation of collisional coalescence (Shima et al., 2009; Riechelmann et al., 2012) and ice-phase processes (Shima et al., 2020), making the SDM a nearly-complete Monte Carlo representation of cloud microphysics. The burgeoning field of particle-based cloud microphysics
uses SDM implementations in large-eddy simulations (LES) to understand microphysical processes that are underresolved in traditional bulk and bin methods (e.g., Chandrakar et al., 2021; Andrejczuk et al., 2010; Morrison et al., 2019; Dziekan et al., 2019; Grabowski, 2020; Hoffmann, 2017). Furthermore, a growing literature of machine learning in microphysics utilizes the SDM as a source of high-fidelity training data from which to "learn" microphysical tendencies and properties (Bieli et al., 2022; Seifert and Rasp, 2020). However, without a complete representation of microphysical processes in the SDM, its predictive and benchmarking power for cloud feedbacks is limited.

Many implementations of the SDM do not include the process of collisional breakup of droplets. Not only is collisional breakup a highly uncertain process in existing bin and bulk parameterizations (Morrison et al., 2020; Grabowski et al., 2019), but these uncertainties have been found to impact rain rates and other macroscale quantities in bin microphysics studies (Seifert et al., 2005). Studying collisional breakup in the SDM is not straightforward – a single breakup event is likely to produce fragments of multiple different sizes. A naive representation of all fragments in the SDM would require the creation of new superdroplet tracers in the system, which can lead to an explosive growth of superdroplet quantity and dramatically inhibit performance of the SDM. Scalability of the SDM for parallel applications such as LES requires a conservation of the total number of superdroplets. This work proposes a superdroplet-conserving SDM algorithm for the representation of collisional breakup, using a Monte Carlo step that samples from a fragment size distribution. This superdroplet-conserving breakup implementation draws inspiration from an analogous "superparticle" representation of phytoplankton (Jokulsdottir and Archer, 2016): individual phytoplankton aggregates spontaneously break uniformly into a number of fragments determined by a power law probability distribution. We apply a similar spontaneous breakup principle to an intermediate coalesced state resulting from the collision of two droplets. (While spontaneous breakup of water droplets has also been investigated (Kamra et al., 1991), it has not been observed in in-situ studies of cloud droplet collisions (Testik and Rahman, 2017) and is widely considered insignificant for atmospheric microphysics (Rogers and Yau, 1989).) The presented collisional breakup algorithm utilizes empirical collection/breakup efficiencies (such as Schlottke et al. (2010); Beard and Ochs (1995); Berry (1967)) to determine whether a colliding droplet pair is likely to break-up, and then samples from a corresponding empirical fragment size distribution (such as Low and List (1982); Schlottke et al. (2010); Beard and Ochs (1995); McFarquhar (2004)) to determine the properties of the resulting fragmented superdroplet. Breakup parameterizations are typically very complex and aim to summarize multiple physical mechanisms of breakup. This work addresses how the proposed SDM breakup algorithm samples from such complex fragment size distributions, but leaves evaluation and analysis of these empirical distributions to future work.

The contents of this paper proceed as follows: Section 2 begin with a conceptual description of the proposed breakup algorithm, followed by a mathematical description of its implementation. Section 3 then presents several idealized simulations including various parameterizations of collisional breakup to demonstrate the behavior of this implementation in the SDM. Section 4 concludes the discussion and poses additional scientific questions which may be within reach given this novel implementation.
2 Superdroplet-conserving Collisional Breakup

2.1 Conceptual description

Two colliding liquid hydrometeors in the atmosphere can break-up via several physical pathways, including filament, sheet, and disc breakup (Barros et al., 2008). The colliding droplets, referred to as "parents", typically lose mass to newly-formed tiny "satellite" droplets that result from the collision, thereby resulting in several differently sized droplet fragments (see figure 1, left). As noted previously, scaling of the SDM relies on preserving the number of tracers in the system. In order to preserve the number of superdroplets in a binary collisional breakup event, breakup is treated as a two step process based on superdroplet-conserving coalescence (figure 1, right). First, the two superdroplets collide and coalesce: the superdroplet of higher multiplicity acts as a "donor" by donating mass and multiplicity while maintaining its attributes; the other superdroplet acts as a "receiver" by growing in mass and maintaining its multiplicity to form a "coalesced transition state." This unstable coalesced transition state immediately breaks up into fragments of uniform size: the attributes and multiplicity of this fragmented receiver are updated, with multiplicity increasing and mass of the individual droplets represented by the superdroplet decreasing. Uniform fragmentation is required to maintain conservation of superdroplets. Furthermore, uniform fragmentation requires the assumption that all superparticle attributes are extensive quantities and undergo equipartitional splitting (not ap-
3a. Coalesce with probability $E!$
4. Random number of fragments $n"$
1. Random pair
3b. Breakup with probability $E#(1-E!)$
3c. Bounce with probability $1-E!(1-E")-E"$
3. Random dynamic $\gamma$ collisions occur in time $\Delta t$
2. Randomly accept collision

Figure 2. Diagram of the Monte Carlo decision pathway during a collision-coalescence-breakup event in the proposed algorithm.

The product of a collisional breakup event is therefore two superdroplets: the donor maintains its attributes but donates multiplicity, and the fragmented receiver represents (uniform) fragments that result from the breakup event following a coalesced transition state. As in the original Monte Carlo step that determines whether a collision occurs, the fragment size is sampled at random from a fragment size distribution, which may depend on the properties of the colliding particles.

2.2 Mathematical description

The superdroplet-conserving method of collisional breakup is illustrated in figure 2 and formulated below using notation following work of Shima et al.. A single superdroplet with label $i$ has a position $x_i(t)$ and extensive physical attributes $a_i(t)$, such as volume or mass ($V_i(t), M_i(t)$, respectively). Each superdroplet corresponds to a multiplicity $\xi_i(t)$ of "real" droplets which exist in the same gridbox and have identical such attributes.

The proposed breakup algorithm unifies the representation of collisional coalescence and breakup and builds on the original coalescence Monte Carlo steps in Shima et al.. As in this original SDM, we begin by selecting pairs of superdroplets to consider collisions:

1. All superdroplets within a cell are randomly ordered in a list of non-overlapping pairs $(j_\alpha, k_\alpha)$ where $j$ and $k$ are the superdroplet indices, and $\alpha$ refers to the pair index.

Next, we determine how many collisions, $\gamma_\alpha$, occur for the pair $\alpha$ in the time step:
2. The probability of collision between droplets $i$ and $j$ is given by

$$P_{i,j} = K_{i,j} \Delta t$$  \hspace{1cm} (1)$$

where $K_{i,j}$ is the rate of collisions based on the properties of droplets $i$ and $j$, and $\Delta t$ is the model time step. The scaled probability of collision $P^{(s)}_{\alpha}$ for this pair $\alpha$ accounts for the multiplicities of the colliding superdroplets:

$$P^{(s)}_{\alpha} = \max(\xi_j, \xi_k) P_{\alpha}.$$  \hspace{1cm} (2)$$

Only a subset $\lfloor n_s/2 \rfloor$ of possible SD pairs are considered out of all possible superdroplet pairs at each time step. Therefore, the probability is further scaled up to form the corrected probability of collision:

$$p_{\alpha} = \frac{n_s(n_s - 1)}{2} / \lfloor \frac{n_s}{2} \rfloor P^{(s)}_{\alpha}.$$  \hspace{1cm} (3)$$

The number of collisions that occur in this time step, $\gamma_{\alpha}$, is then determined in a Monte Carlo step based on $p_{\alpha}$. Taking $\phi_{\alpha} \in (0, 1)$ to be a uniform random number,

$$\gamma_{\alpha} = \min([p_{\alpha} - \phi_{\alpha}], [\xi_{j,\alpha}/\xi_{k,\alpha}]).$$  \hspace{1cm} (4)$$

Here, we assumed the superdroplets are ordered such that $\xi_{j,\alpha} \geq \xi_{k,\alpha}$. If $\gamma_{\alpha} = 0$, then no collisions occur.

The collision rate is then $\gamma_{\alpha}$ collisions per gridbox and per time step. Due to the constraint in equation (4) based on droplet multiplicity, some collisions which should occur probabilistically cannot if the donor superdroplet has insufficient multiplicity to collide $p_{\alpha}$ times. Therefore, a collision deficit $p_{\alpha} - \gamma_{\alpha}$ may be tracked as a tool to assess whether the model time step is sufficiently small (elimination of the collision deficit is used for adaptive step size control in the SDM implementation used herein (Arabas et al., 2022)).

In the original SDM, particles coalesce as long as $\gamma_{\alpha} > 0$, as the rate of collisions is taken to refer only to collisional coalescence. However, when we consider collisional breakup, an additional Monte Carlo step must be taken to determine whether the particles coalesce or break up. This is determined based on a coalescence efficiency (or collection efficiency) $E_c$, which generally depends on properties of the colliding particles such as their fall speed, mass, and surface tension. We additionally account for the fact that in some collisions, droplets may bounce off of one another elastically by including an optional additional parameter for the breakup efficiency, $E_b$. This second Monte Carlo step is summarized as follows.

3. Compute the dynamic that occurs: coalescence, breakup, or bounce (nothing). A second uniform random number $\phi'_{\alpha}$ determines the outcome:

$$\text{dynamic}_{\alpha} = \begin{cases} 
\text{coalescence,} & \phi'_{\alpha} \leq E_c(a_j, a_k) \\
\text{breakup,} & E_c(a_j, a_k) < \phi'_{\alpha} \leq E_b(a_j, a_k)(1 - E_c(a_j, a_k)) + E_c(a_j, a_k) \\
\text{bounce,} & \phi'_{\alpha} > E_b(a_j, a_k)(1 - E_c(a_j, a_k)) + E_c(a_j, a_k)
\end{cases}$$  \hspace{1cm} (5)$$
Once the dynamic is determined, a fragment size is sampled if necessary:

4. Sample a fragment size $M_{f,\alpha}$ (mass) from a fragment size distribution, $P_{f,\alpha}$, with cumulative distribution function (CDF) $C_{f,\alpha}(\phi)$ that depends on the colliding particle attributes. A related variable, $N_{f,\alpha}$, is taken to denote the number of fragments that would form in a collision between droplets of mass $M_j$ and $M_k$: $N_{f,\alpha} = \frac{M_k + M_j}{M_{f,\alpha}}$.

Finally, updating of multiplicities and attributes proceeds based on the selected dynamic, number of collisions, and sampled fragment size (if applicable):
(a) For coalescence:

\[
\begin{align*}
\xi'_j &= \xi_j - \gamma \alpha \xi_k \\
\mathbf{a}'_k &= \mathbf{a}_k + \gamma \alpha \mathbf{a}_j \\
\text{if } \xi_j = 0, \text{ then } \xi'_j, \xi'_k &= \xi_k/2, \mathbf{a}'_j = \mathbf{a}_k
\end{align*}
\]

The coalescence rate is incremented by \(\gamma \alpha \xi_k\).

(b) For breakup:

In some cases, only \(\gamma_{jk} \leq \gamma \alpha\) breakups can occur for a given superdroplet pair without encountering negative multiplicities. We compute this maximum possible number of breakup steps and update the superdroplet properties using a recurrence relation (assuming \(\gamma \alpha > 0\)), and track a breakup deficit rate of \(\gamma \alpha - \gamma_{jk}\). (Alternatively, one may perform substepping of the breakup event.) The particle attributes are updated such to be consistent with the result of several breakup steps with \(\gamma \alpha = 1\) occurring in sequence, always producing fragments of size \(M_{f,\alpha}\).

\[
\begin{align*}
\gamma_{jk} &= 0 \\
\xi'_j &= \xi_j \\
\xi'_k &= 0 \\
\text{while } \gamma_{jk} < \gamma \alpha \text{ and } \xi'^{\text{transfer}}_j < \xi_j : \\
\xi'^{\text{transfer}}_j &= \xi'^{\text{transfer}}_j + \xi^{\text{new}}_k \\
\xi'^{\text{new}}_k &= \xi'^{\text{new}}_k \left( \frac{M_j}{M_{f,\alpha}} \right) + \xi_k \frac{M_j + M_k}{M_{f,\alpha}} \\
\gamma_{jk} &= \gamma_{jk} + 1
\end{align*}
\]

\[
\begin{align*}
\xi'_j &= \xi_j - \xi'^{\text{transfer}}_j \\
\xi'_k &= \xi'^{\text{new}}_k \\
\mathbf{a}'_k &= \frac{\xi_k \mathbf{a}_k + \xi'^{\text{transfer}} \mathbf{a}_j}{\xi'^{\text{transfer}}_k} \\
\text{if } \xi_j = 0, \text{ then } \xi'_j, \xi'_k &= \xi_k/2, \mathbf{a}'_j = \mathbf{a}_k
\end{align*}
\]

The breakup rate is incremented by \(\gamma_{jk} \xi_k\). The breakup deficit rate is incremented by \((\gamma \alpha - \gamma_{jk}) \xi_k\).

(c) For bounce:

No update is made to droplet multiplicities or attributes, and only the collision counter is incremented.
2.3 Additional Implementation Details

This method of breakup allows for the splitting of a coalesced transition state into a non-integer number of fragments \( N_{f,\alpha} = \frac{M_k + M_j}{M_{f,\alpha}} \) need not be integer), depending on the sampled fragment size. For instances where it may be desirable to preserve superdroplet multiplicities as integers, we recommend rescaling the multiplicities after the breakup step by a factor of \( r_k = \lceil \xi_k \rceil / \xi_k \), and the multiplicities correspondingly by \( 1/r_k \) such that extensive attributes (including mass) are conserved.

The presence of a "breakup deficit" in the case where \( \gamma_{jk} < \gamma_\alpha \) can be averted by substepping, though this is inadvisable for highly parallel applications of the SDM. Furthermore, superdroplet multiplicities may increase without bound according to the algorithm as presented above, which can lead to numerical artifacts and instability within a simulation. A set of limiters preventing runaway multiplicity is discussed in Appendix A. Finally, a method for sampling a fragment size from a highly nonlinear empirical distribution, such as Straub 2010, is discussed in Appendix B.

3 Numerical Experiments and Discussion

To demonstrate the behavior and impact of the proposed breakup algorithm on particle size spectra and process rates, we present a set of zero-dimensional box and one-dimensional rainshaft cases that include collisional breakup, implemented in the open-source Pythonic superdroplet code ‘PySDM’ (Bartman et al., 2022b). All simulations use a geometric collision kernel, where the rate of collisions \( K_{jk} \) between droplets with the properties of superdroplets \( j \) and \( k \) is given by

\[
K_{jk} = \pi(R_j + R_k)^2|v_j - v_k|
\]

where \( R_j \) is the radius of particle \( j \) and \( v_j \) is the terminal velocity/fall speed of particle \( j \), computed using the parameterization of Gunn and Kinzer. Collisions within a superdroplet (i.e. collisions between droplets represented by the same superdroplet) are not considered in ‘PySDM’, though the use of a geometric collision kernel results in zero collisions between equally-sized droplets, as they have the same terminal velocity.

The coalescence efficiency is specified to be either a constant value (for sensitivity studies), or the empirical coalescence efficiency of Straub et al. which depends on the Weber number of the colliding droplet pair. (The Weber number is a ratio of kinetic collisional energy and surface tension, and relates to the stability of a droplet pair under collision.) We consider three types of fragmentation functions: (1) a constant fragment number \( N_f \), in which the particle-size distribution (PSD) is a delta function \( P_f(M_{f,\alpha}) = \delta(M_{f,\alpha} - \frac{M_j + M_k}{N_f}) \); (2) a normal distribution \( M_{f,\alpha} \sim N(\mu, \sigma) \) where the mean \( \mu \) and variance \( \sigma^2 \) are specified; and (3) the empirically derived fragmentation function of Straub et al., which uses four categories of fragmentation and lognormal or normal subdistributions.

3.1 Particle Size Distribution

The zero-dimensional box simulations include collisional-coalescence and breakup dynamics only. The droplet size distribution is initialized to an exponential distribution in mass \( x \), given by \( N(x) = x_0 \exp(-x/x_0) \) with the characteristic size \( x_0 = (4\pi/3)R_0^3 \) set using \( R_0 = 30.531\mu m \) as in Shima et al. (2009). The simulations employ \( 2^{13} = 8192 \) superdroplets to represent...
Figure 3. Particle size distribution with varying coalescence efficiencies under a geometric collision kernel after 100s (left) and 200s (right). The breakup fragmentation function is deterministic, with the fragment size determined as 1/8th the sum of the colliding droplet sizes. The dashed black line represents the initial PSD, and solid lines represent various fixed values of the coalescence efficiency. The pink line corresponds to a size-dependent coalescence efficiency from Straub et al..

a number density of 100/cm$^3$ in a box of volume 1m$^3$ with a fixed time step of 1s. This choice of superdroplet quantity is sufficient to produce consistent results in the PSD across realizations using a different random seed, and was shown by Shima et al. to closely match the exact PSD in a similar box model simulation of collisional coalescence. All particle size distributions are displayed as the marginal mass distribution $g(R) = \frac{dm}{d\ln(R)} = 3x^2n(x)$ where $n(x)$ is the particle size distribution. This mass distribution is computed by binning the resulting superdroplets into 128 logarithmically-spaced size bins between particle radius 1µm and radius 10mm. We separate the simulations into those which use a deterministic fragmentation function, in which breakups result in a constant number of fragments in any given collision; a stochastic fragmentation function with fragment sizes sampled from a specific distribution; and a size-dependent fragmentation function, where the fragment sizes are sampled from a distribution whose parameters depend on the colliding particles. We further include experiments exploring the use of a fixed coalescence efficiency versus a particle-attribute-derived coalescence efficiency. This separation elucidates which aspects of the particle population behavior are attributable to stochastic stochastic sampling of the fragmentation function, or related to particle-property-dependent parameters such as Weber number.

3.1.1 Deterministic and Size Independent Fragmentation

First we investigate the sensitivity of the PSD evolution to the coalescence efficiency, using four values of a constant-valued efficiency $E_c$ between 0.8 and 1.0 ($E_c = 1.0$ corresponds to coalescence-only) and a particle-size dependent $E_c$ parameterization (Straub et al., 2010). All simulations use a deterministic fragmentation function in which all single-step collisional breakups
Figure 4. Sensitivity to fragmentation function of PSDs following collisions with a geometric kernel and fixed coalescence efficiency in a deterministic and stochastic fragmentation function case. (Left) The fragment size is fixed by a divisor of the sum of colliding particle volumes; (right) fragment size is sampled from a Gaussian distribution with varying means determined as a multiple of the initial distribution mean, with variance $\sigma^2 = \mu^2/4$. The initial distribution is shown as a black dashed line in each figure.

result in $N_f = 8$ fragments:

$$P_{f,\alpha}(M_f) = \delta(M_f - \frac{M_j + M_k}{N_f}).$$  \hspace{1cm} (10)

Figure 3 displays two snapshots of the PSD under this set of dynamics, demonstrating the impact of stochasticity in selecting whether coalescence or breakup occurs, independent of sampling a fragment size. As expected, the initial PSD broadens and shifts toward larger droplets at 100s, with the largest values of fixed $E_c$ leading to the largest increase in average particle mass. However, after 200s, the PSD for the $E_c = 0.8$ case remains approximately steady with a mean size that is smaller than the initial distribution mean, demonstrating that coalescence and breakup are approximately balanced in this case.

By contrast, the PSD for the Straub 2010 parameterization of $E_c$ is initially comparable to the $E_c = 0.95$ simulation at 100s, but narrows without shifting toward much larger droplets after further time has elapsed, leading to a dominant mode that is more similar to the $E_c = 0.9$ case. This empirical parameterization also shows evidence of approaching a steady state distribution, in which coalescence and breakup rates are matched on average, driving the PSD to a stationary state. The Straub 2010 parameterization decreases exponentially with the colliding particle Weber number, which is correlated with the size and relative velocity of the colliding particles. Two colliding particles with a similar size have a low relative velocity, therefore as the PSD shifts toward larger coalesced droplets, there is a competing effect between a larger particle size increasing the Weber number, and decreased relative velocity reducing it. This competition produces the stationary behavior and narrowing of the PSD observed in this case.
Figure 5. Initial PSD (black dashed) and PSD’s following collisions with a geometric kernel, the Straub 2010 collection efficiency, and the Straub 2010 fragmentation function Straub et al. (2010) after several elapsed times (colors).

Next we consider the PSD evolution when the coalescence efficiency is held fixed at a constant value and the fragmentation function is varied. In figure 4(a), we consider a deterministic fragmentation function where the number of fragments from a single breakup event is fixed (as in figure 3), as well as a Gaussian fragment size distribution with mean $\mu$ specified as a multiple of the initial mean particle mass $x_0$, and variance $\sigma^2 = \mu^2/4$. When the number of fragments is fixed, results using the largest number of fragments display the smallest mean particle size and broadest spectra. The first behavior is expected, as a larger value of $N_f$ results in smaller typical fragment sizes. The broadening of the spectrum can be attributed to a wider range of collision rates between very small droplets (which result from fragmentation), and is generally an expected outcome of including collisional breakup.

When the fragment size is sampled from a normal distribution (figure 4(b)), the resulting spectra are bimodal, with a large-droplet mode that is similar between different choices of the mean $\mu$, and a narrow small-droplet mode that depends on the distribution parameters. The appearance of a second mode occurs when the fragment size is sampled from the left tail of the fragment size distribution, whereas the large mode corresponds to a droplets undergoing coalescence only, as in the $N_f = 1$ case. This behavior indicates that through stochastic sampling of the dynamic and fragment size together, the droplet population splits into one mode which fragments into smaller droplets, and a second mode which primarily undergoes coalescence and grows in size. Because larger droplets collide at much quicker rates than small droplets, the fragmented mode is less likely to collide and re-coalesce to form medium-sized droplets, while the coalesced-mode retains some probability of colliding and either growing (coalescing), or breaking up into smaller droplets. Thus we observe that the small-droplet-mode grows in this instance, with particles effectively become "stuck" in this dynamical regime due to the separation of scales in collision rates.
3.1.2 Stochastic Size-Dependent Fragmentation

Finally, we consider an empirically derived coalescence efficiency (Straub et al., 2010, as above) in conjunction with a complex empirical fragmentation function whose parameters depend on the colliding droplet properties. In figure 5, we consider the evolution of the PSD under the Straub 2010 efficiency and fragmentation dynamics, beginning from the same initial distribution as previous experiments. At first, the PSD broadens and shifts towards larger droplet sizes, as in figure 3, but the PSD after 540s is virtually unchanged from the PSD at 180s. These results indicate the stationarity of the particle size distribution after sufficient time has elapsed: coalescence and breakup are balanced, as in the previous example. Contrasted with figure 3, which used a deterministic size-independent fragmentation function, the stationary PSD resulting from the Straub et al. (2010) parameterization of fragmentation is broader and less symmetric. This difference reflects the use of sampling from a distribution of fragment sizes, contrasting with the symmetric PSDs found from using a fixed number of fragments.

3.2 Cloud and Precipitation Properties

Next we consider the impact of collisional breakup in a one-dimensional warm rain setting that includes condensation/evaporation (including aerosol activation/deactivation), collisions, and transport of particles within the column through advection and sedimentation/precipitation. These 1D simulations are based on the kinematic framework of Shipway and Hill, using a fixed profile of dry-air potential temperature and dry-air density $\rho_d(z)$, and a resolved budget of water vapor (advection and coupling with vapor uptake and release by particles). The vapor advection is solved using the MPDATA algorithm on a columnar grid with spatial grid step of 25m (employing the PyMPDATA implementation Bartman et al., 2022a). An aerosol population with hygroscopicity $\kappa = 0.9$ is initialized throughout the vertical domain with $2^{10} = 1024$ superdroplets per gridbox. This choice of 1024 superdroplets per gridbox reflects the higher computational demands of the one dimensional simulation compared to the box model, and still produces statistically convergent results in the mesoscale quantities investigated across simulation instances. For the first 600s of spin-up, condensation-evaporation (including aerosol activation) and particle advection are the only active dynamics, with a time-varying updraft momentum flux of $\rho_d w(t) = 3 \text{kg m}^{-3} \text{ms}^{-1} \sin(\pi t/600\text{s})$. After this spin-up time, the updraft velocity is set to 0, particle displacement due to sedimentation is enabled, and collision-coalescence-breakup is allowed to occur. The time step is fixed at 1s throughout the simulation.

The test cases demonstrated here include a no-breakup case, a property-independent breakup case where the coalescence efficiency is fixed and fragment sizes are sampled from a fixed distribution, and the particle-property-dependent empirical coalescence efficiency and fragmentation parameterizations from Straub et al.. All simulations use a geometric collision rate (equation 9) and the Gunn and Kinzer terminal velocity parameterization. In the no-breakup case, all collisions result in coalescence. In the property-independent breakup case, we fix $E_c = 0.95$ for all superdroplet collisions based on the correspondence in figure 3 to the empirical coalescence efficiency. This case samples fragment sizes from a Gaussian distribution in particle volume with mean radius 30$\mu$m and standard deviation 15$\mu$m. In contrast to the property-independent case, in which the fragmentation parameters are hand-selected, the property-dependent setting is based on empirical evidence, and is expected to be more reflective of the variability of real clouds. In both the property-independent and -dependent cases, the breakup efficiency
Figure 6. Hydrometeor concentrations without breakup (left column), with breakup using a property-independent coalescence efficiency (middle column), and with breakup following the property-dependent Straub et al. (2010) parameterizations (right column). Included are cloud water mixing ratio (top row), rain water mixing ratio (middle row), and aerosol number concentration (bottom row).

is set to $E_b = 1$ such that all collisions result in either coalescence or breakup. To contrast the behavior of the three cases, we consider the hydrometeor population at various altitudes throughout the simulation, as well as collision process rates and aerosol processing rates.

3.2.1 Hydrometeor and Cloud Quantities

The mixing ratio of cloud droplets (activated droplets of no more than 50µm radius), rain droplets (radius greater than 50µm), and the number concentration of unactivated aerosols are displayed for the three test cases in figure 6. The no-breakup simulation forms a cloud due to activation of aerosols between 500m and 2200m altitude until the updraft is terminated after 600s.
Larger rain-range droplets form from collisional coalescence and begin to sediment out of the system in clusters, visible as distinct streaks in the \((t, z)\) plane, with surface precipitation beginning around 1200s into the simulation, depleting the cloud droplet population.

When property-independent breakup is included, a higher concentration of cloud-sized droplets persists at cloud base, and the surface precipitation is delayed and spread out relative to the case with no breakup. This behavior indicates that rain droplets favorably break up within the cloud and especially near cloud base, fragmenting into smaller cloud droplets (the mean of the fragment size distribution is 30µm radius, only slightly lower than the rain size range) with a lower sedimentation rate. Furthermore, the aerosol population below cloud base is not depleted as quickly in this property-independent case, indicating a reduction in aerosol scavenging and washout that is consistent with the lower precipitation rates. These phenomena are consistent with documented impacts of collisional breakup such as suppressed surface precipitation (Seifert et al., 2005), and show that the proposed algorithm can meaningfully represent the breakup process.

The empirical property-dependent breakup case using the Straub et al. parameterizations displays hydrometeor populations that are more similar to the no-breakup case, indicating that the choice of \(E_c = 0.95\) in the property-independent case likely overestimates the rate of collisional breakup when condensation and evaporation are present (contrasted with figure 3). As in the no-breakup case, the property-dependent empirical case displays distinct streaks of precipitation, with surface precipitation initiated around 1200s.

The 3m/s updraft velocity used in this setup leads to a relatively shallow cloud, with drizzle-range precipitable hydrometeors which are less likely to undergo collisional breakup than very large droplets in a deeper convective setting. The likelihood of breakup in the Straub parameterization is strongly correlated with the size of the colliding droplets, therefore we expect to see a stronger impact of including SDM breakup in a strongly precipitating convective case. Due to challenges and complexity of representing mixed-phase hydrometeors in the superdroplet method, we do not present any such deep convective experiment. However, it has been noted that collisional breakup of supercooled liquid droplets upon impact with ice is a potentially important secondary ice production mechanism (Zhao and Liu, 2022; James et al., 2021; Phillips et al., 2018), thus we suggest that the novel SDM breakup representation presented in this work could be an instrumental tool in further research on secondary ice production and mixed-phase processes.

### 3.2.2 Process Rates

Figure 7 displays the local rates of superdroplet collision (scaled by multiplicity), as well as distinguishing between rates of coalescence and rates of breakup. We see an expected correlation between the time and location of collisions in all three cases with the location of hydrometeors (outlined in black for cloud and red for rain)--as expected, a higher concentration of hydrometeors, particularly large rain-range hydrometeors corresponds to higher rates of all collisional dynamics. The rate of collisions increases throughout the simulation time, particularly near cloud base where the largest droplets are sedimenting and colliding at higher rates. The property-independent case is consistent with the other cases in displaying higher collision rates at cloud base, even though the droplets in this region are slightly smaller and fall in the cloud category.
Figure 7. Collisional dynamic rates for the 1-dimensional case with (left) no breakup, (middle) breakup with a fixed coalescence efficiency, and (right) breakup with the Straub 2010 parameterizations. The dynamics shown include (top to bottom): collision rate, coalescence rate, and breakup rate. Dashed contour lines represent the level of $q_c = 0.2g/kg$ (black) and $q_r = 0.2g/kg$ (red), representing a cloudy and rainy region of the time-space domain, respectively.
All three cases display similar rates of collision and coalescence, with the highest of these rates occurring in the cloud among rain droplets, and below cloud base among precipitating rain droplets. In the no-breakup case, every feasible collision results in a coalescence, and the breakup rate is zero. When property-independent breakup is included, the time-space distribution of the breakup rate is nearly identical to that of the collision rate. This trend is consistent with the use of a uniform coalescence efficiency $E_c = 0.95$, which is agnostic to the size of the colliding particles. In contrast, the empirical property-dependent case sees collisional breakup primarily where larger rain droplets are present, consistent with the Straub et al. parameterization based on Weber number. In this case, breakup events are much less frequent, and thus breakup plays less of a role in determining the hydrometeor populations of this simulation case. These results demonstrate that the SDM breakup algorithm can produce expected process behavior in both a property-independent setting, where the collision dynamics result in strong breakup, and in an empirically parameterized setting.

As noted in the discussion of figure 6, the property-independent breakup case experiences a persistent population of aerosols below cloud base, while the no-breakup and property-dependent cases demonstrate washout upon the onset of precipitation. Collision breakup resulting in very small droplet fragments could potentially introduce cloud droplets so small that they deactivate in their environment. In figure 8, we investigate differences in aerosol processing rates under these collisional breakup scenarios, looking at rates of aerosol activation, deactivation, and ripening.

The property-independent and no-breakup cases have nearly identical behavior in aerosol processing, consistent with the correspondence between their hydrometeor concentrations and collision process rates. In all three cases, a few superdroplets at cloud top encounter humidity close to their critical supersaturation, which results in the "ripening" processes of fluctuation between an activated and deactivated state due to competition when the supersaturation is insufficient to activate all aerosols (e.g., Arenberg, 1939; Wood et al., 2002). (We define ripening rate as the number of activated droplets growing through condensation per unit of time within a grid cell in which deactivation simultaneously occurs on other particles). Several instances of ripening occur within the extended cloud base of the property-dependent case, indicating that breakup events in this time-space domain produce fragments whose critical supersaturation is close to the environment supersaturation (i.e., hinting breakup-induced deactivation, conceptually analogous to collision-induced activation discussed in Hoffmann, 2017). Aerosols activate primarily at the start of the simulation when an updraft is present, defining the altitude boundaries of the cloud. No additional activation is seen in either instance including collisional breakup. Deactivation occurs among a few aerosols which activate and then rise in altitude beyond cloud top initially, and more strongly below cloud base as droplets sediment out of the cloud and evaporate. The property-independent case experiences much stronger deactivation at cloud base, which corresponds to the higher rate of fragmentation of droplets at this altitude.

The more realistic property-dependent case shows a close match in aerosol processing rates to the no-breakup case, indicating that collisional breakup is not a significant process for shallow clouds such as this kinematic one-dimensional setting. However, the enhanced rates of ripening and deactivation in the property-independent case indicate that collisional breakup could be a relevant process for future studies of aerosol-cloud effects, particularly in deeper-convective cases where collision rates are likely to be higher. The SDM representation of collisional breakup presented in this work can capture such potential effects, making it a useful tool for future studies of microphysics processes.
Figure 8. Aerosol processing rates for (left to right) no breakup, property-independent breakup, and Straub et al. (2010) parameterizations. Included are (top to bottom): ripening rate, activating rate, and deactivating rate. Dashed contour lines represent the level of $q_c = 0.2g/kg$ (black) and $q_r = 0.2g/kg$ (red), representing a cloudy and rainy region of the time-space domain, respectively.
4 Conclusions

This work presents a superdroplet algorithm for collisional breakup that is both scalable in avoiding creation of new superdroplets, and physical in its ability to produce results in a box and one-dimensional setting that are consistent with the expected suppression of rain. Furthermore, the algorithm produces hydrometeor populations and process rates that differ between a property-independent approach (with a fixed coalescence efficiency and fixed fragment size distribution), and a property-dependent approach using empirical parameterizations. These differences indicate the importance of random, stochastic events in warm rain microphysics, a trait which has also been documented in other microphysical phenomena such as giant CCN (Feingold et al., 1999; Yin et al., 2000). Without such a scalable representation, the superdroplet method has heretofore been unable to capture these additional stochastic impacts of breakup, nor has it been applied to compare empirical parameterizations of coalescence and breakup, which contribute uncertainties to operational process, weather, and climate models.

This work provides the basis for a more complete representation of microphysical processes in particle-based simulations. For instance, when combined with collisional breakup, a superdroplet representation of ice-phase hydrometeors could probe processes of secondary ice production. For instance, when supercooled water droplets collide with solid phase hydrometeors, they could break up, with small fragments freezing following impact or freezing onto the surface of the colliding ice crystal (Phillips et al., 2018; James et al., 2021). Similarly, collisional breakups between two ice crystals can lead to splintering and the formation of many smaller ice crystals, which then grow by vapor deposition (Harris-Hobbs and Cooper, 1987). These and other mixed-phase processes are poorly understood due to challenges in obtaining direct observational or laboratory measurements, thus a high-fidelity particle-based representation such as the superdroplet method provides an ideal means for studying these phenomena. While the collisional breakup representation presented here does not address underlying uncertainties in parameterization of processes such as collision rates and phase change, it provides a path forward for more rigorous and complete studies of cloud microphysics.

Code and data availability. Implementation of this breakup algorithm in the SDM is available at https://doi.org/10.5281/zenodo.7306034. The simulations presented in this work (and all necessary input information) are available in the folder ‘deJong_Mackay_2022’ at https://doi.org/10.5281/zenodo.7308668. The notebooks in this folder reproduce all results and figures presented in this study, with no external datasets required. The scripts run the relevant model configuration in a matter of minutes and plot the resulting output. All results presented in this paper can be reproduced by one of two means: (1) downloading and installing ’PySDM’ and ’PySDM-examples’ (e.g. using ’pip install’), and running the notebooks locally; (2) accessing the PySDM-examples repository online and running the examples notebooks in the folder ‘deJong_Mackay_2022’ on Google Colab. These codes, PySDM and PySDM-examples, are continuously under development at https://github.com/atmos-cloud-sim-uj/PySDM and https://github.com/atmos-cloud-sim-uj/PySDM-examples, and are further documented in a software publication (Arabas et al., 2022).
Appendix A: Limiters

In implementing collisional breakup for superdroplets, we suggest imposing a few limiters to enforce physical constraints and maintain stability of the code. If the user-selected time step for the SDM implementation is too large, collisional breakup may quickly become a runaway process with superdroplet multiplicities increasingly rapidly and unphysically, leading to numerical overflow. As an example, suppose a droplet of multiplicity $10^2$ should undergo 6 collisional breakups ($\gamma_\alpha = 5$) into 5 fragments each time ($N_{f,\alpha} = 5$): then $\gamma_{jk} = 3906$ and its new multiplicity is $15,525 = \mathcal{O}(10^4)$. Successive collisional breakups between droplets whose multiplicities have grown so rapidly would then lead to exponentially booming multiplicities, and could quickly exceed the maximum representable quantity for the computing machine (overflow). One solution is to set a maximum allowable multiplicity for any superdroplet, and to reject any collisional breakups that would produce a superdroplet exceeding this multiplicity.

In addition, the process of collisional breakup is physically constrained such that the resulting superdroplet (the "fragment") volume should not exceed the volume of either colliding droplets, nor should it drop below a realistic size for a liquid water droplet (molecule scale, for instance). These physical constraints can be imposed by setting a minimum and maximum allowable fragment size resulting from breakup.

The first of these constraints can then be imposed during computation of $\gamma_{jk}$ within the while loop in equation 7:

$$\xi_k^{\text{new}} \leq \xi_{\text{max}}$$

where $\xi_{\text{max}}$ is a maximum multiplicity set to prevent overflow. The second two constraints are imposed during the sampling of a fragment size:

$$M_{\min} \leq M_{f,\alpha} \leq \max(M_j, M_k)$$

where $M_{\min}$ is a minimum physically allowed fragment size, and the final constraint restricts the resulting fragment to be no larger than either colliding droplet.

Appendix B: Sampling from empirical fragment size distributions

Sampling a fragment size $M_{f,\alpha}$ requires the CDF of the fragment size distribution $P_f(\phi)$, which can be challenging for an empirical fragmentation function that is piecewise and lacks a closed form CDF. For instance, the commonly-used fragmentation function of Low and List partitions the fragment size distribution into three categories of distinct functional form, corresponding to filament, sheet, and disk breakup. Similarly, Straub et al. distinguish four categories of fragmentation, with the fragment size distribution within each category following a lognormal or normal distribution. We will demonstrate how a uniform random number $\phi''_\alpha$ can be used to sample a fragment size from such complex distributions, following the notation of Straub et al..
Suppose the unnormalized fragment size distribution $P_f(D)$ in droplet diameter $D$ is described as a sum of $k$ subdistributions:

$$P_f(D) = \sum_{r=1}^{k} N_r p_r(D) \quad (B1)$$

where $N_r$ is the expected number of fragments from mode $r$, and $p_r(D)$ is the normalized fragment size distribution for mode $r$. Note that $\int_0^\infty P_f(D) dD = \sum_{r=1}^{k} N_r \neq 1$ is the expected total number of fragments, and thus sampling a fragment size from the distribution requires normalization.

To sample a single fragment size $D_f$, we first use the random number $\phi''_\alpha$ to determine which mode of fragmentation occurs by finding $s$ such that

$$\sum_{r=1}^{s-1} N_r \leq \phi''_\alpha < \sum_{r=1}^{s} N_r \quad (B2)$$

The fragment size is then chosen by sampling at random from the CDF of $p_s(D)$, which is assumed to be approximable by a closed form equation (as in the case of a Gaussian or lognormal distribution). This second step of sampling can be accomplished by selecting a new random number, reusing the random number from a different colliding droplet pair, or simply rescaling $\phi''_\alpha$ to the selected mode $s$:

$$\tilde{\phi''}_\alpha = \frac{\phi''_\alpha - \sum_{r=1}^{s-1} N_r}{\sum_{r=1}^{s} N_r - \sum_{r=1}^{s-1} N_r} \quad (B3)$$

The fragment size $D_f$ is then chosen such that

$$\tilde{\phi''}_\alpha = p_s(D_f) \quad (B4)$$

**Author contributions.** EdJ led the code development, generation of results, interpretation, and writing. BM contributed to code development and the underlying methodology. AJ contributed to underlying methodology and interpretation of results. SA contributed to the code development, interpretation, and leads maintenance of the PySDM codebase.

**Competing interests.** The authors declare there to be no competing interests present related to this work.

**Acknowledgements.** We thank Tapio Schneider and Shin-ichiro Shima for feedback, insights, and discussion. Additional thanks go to Piotr Bartman for ongoing support of PySDM and to Oleksii Bulenok for assistance in GPU backend and code review for the collisional breakup dynamics. E. de Jong was supported by a Department of Energy Computational Sciences Graduate Fellowship. S. Arabas acknowledges support from the Polish National Science Center (grant no. 2020/39/D/ST10/01220). This research was additionally supported by Eric and Wendy Schmidt (by recommendation of Schmidt Futures) and the Heising-Simons Foundation.
References


