Spanning the Gap from Bulk to Bin: A Novel Spectral Microphysics Method

E. K. de Jong[1*], T. Bischoff[1], A. Nadim[2], T. Schneider[1,3]

[1]California Institute of Technology, Pasadena, CA, USA
[2]Institute of Mathematical Sciences, Claremont Graduate University, Claremont, CA, USA
[3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Key Points:

- A new microphysics method using collocation of basis functions is presented.
- The method improves spectral accuracy and precipitation predictions over bulk and bin methods.
- The method applies to a flexible range of computational complexity, providing a way to unify microphysics models.

*

Corresponding author: Emily de Jong, edejong@caltech.edu

Abstract

Microphysics methods for climate models typically track one, two, or three moments of a droplet size distribution for various categories of liquid, ice, and aerosol. Such methods rely on conversion parameters between these categories, which introduces uncertainty into predictions. While higher-resolution options such as bin and Lagrangian schemes exist, they require too many degrees of freedom for climate modeling applications and introduce numerical challenges. Here we introduce a flexible spectral microphysics method based on collocation of basis functions. This method generalizes to a linear bulk scheme at low resolution and a smoothed bin scheme at high resolution. Tested in an idealized box setting, the method improves spectral accuracy for droplet collision-coalescence and improves precipitation predictions relative to bulk methods; furthermore, it generalizes well to multimodal distributions with less complexity than a bin method. The potential to extend this collocation representation to multiple hydrometeor classes suggests a path forward to unify liquid, ice, and aerosol microphysics in a single, flexible, computational framework for climate modeling.

Plain Language Summary

Clouds and aerosols affect global warming by reflecting and absorbing radiation and by storing and transporting water. Climate models need a way to efficiently track the size and number of cloud droplets, ice, and aerosols in order to accurately predict the impact that these "microphysical" particles have on climate. Existing methods of microphysics rely on many uncertain parameters and are either too complicated or too simple to take advantage of today's computational resources. We propose a new way to represent cloud droplets that can both reduce uncertainties and make use of increased computing power.

1 Introduction

Droplets, aerosols, and ice particles, collectively a subset of atmospheric microphysical particles, affect planetary-scale climate, yet the processes that govern their behavior occur at the microscale. This extreme range of scales, from droplets to clouds to atmospheric dynamics, makes it challenging to computationally represent microphysics. There are simply too many particles to represent directly, yet the microphysics processes involved are highly nonlinear and do not lend themselves easily to simplifications. Instead, microphysics schemes in climate and numerical weather models predict the particle size distribution (PSD) present at various locations in the atmosphere: the PSD and number concentration determine the macroscopic behavior of the system, such as cloud albedo or precipitation rates. Historically, methods to represent the PSD developed along two trajectories: bulk methods, which predict aggregate properties of the droplet population, and spectral methods, which explicitly track the PSD. Both of these representations make assumptions about the droplet distribution and the microphysical process rates, with spectral methods being the more flexible of the two options. Unfortunately, these parameterizations and assumptions contribute a major yet difficult-to-quantify source of uncertainty in climate predictions (Intergovernmental Panel on Climate Change, 2014; Morrison et al., 2020; Randall et al., 2003; Khain et al., 2015; Arakawa, 2004).

Bulk schemes, originating with Kessler (1969), explicitly track one or more prognostic moments of the PSD and therefore are very compact representations suitable for global climate applications. However, by abstracting a droplet population to one, two, or three variables, bulk methods make two fundamental simplifications. First, many single-droplet processes such as sedimentation or aerosol activation require parameterizations to approximate how the process impacts the prognostic moments. Second, because many such process rates depend on higher-order moments which are not explicitly tracked, moment-based methods require a closure to relate these higher order moments back to the prog-

nostic variables. Frequently this closure is accomplished by relating the prognostic moments back to an underlying assumed size distribution such as a gamma or exponential (e.g., Morrison & Grabowski, 2008; Seifert & Beheng, 2006; Milbrandt & Yau, 2005), which corresponds well to data in many empirical settings. However, in the case of a multimodal distribution, for instance, when both small cloud droplets and larger rain droplets are present, this closure assumption introduces significant structural uncertainty into the microphysics scheme. There is no physical reason, a priori, to restrict a droplet population to maintaining a particular size distribution as they coalesce, break up, grow, sediment, and change phases. Unfortunately, inverting a multimodal distribution analytically is frequently ill-posed (Morrison et al., 2019). Most traditional bulk methods avoid the issue by representing several categories of hydrometeors (rain, cloud droplets, and several categories of aerosols) through separate prognostic moments, assuming a simple unimodal distribution for each of these categories. However, these categories of condensed water, while intuitive, are artificial: in reality, liquid hydrometeors are distributed across a continuous spectrum, from small chemically-active aerosol particles, to large liquid cloud droplets, to droplets which are large enough to fall as rain. Conversion between these categories adds further complexity and uncertainty to the model.

On the other hand, spectral or "bin" microphysics schemes directly evolve the PSD in time through discrete bins, or particle size ranges (e.g., Tzivion (Tzitzvashvili) et al., 1987; Berry, 1967; Berry & Reinhardt, 1974; Young, 1974). Bin methods have made a great impact in understanding aerosol-cloud interactions (e.g., Morrison & Grabowski, 2007; Khain et al., 2015), but at a higher computational cost that currently makes them infeasible for climate simulations. For example, Gettelman et al. (2021) ran a general circulation model (GCM) with bin microphysics, incurring a factor of five cost penalty over a bulk scheme. Furthermore, while bin methods avoid the closure assumptions of bulk schemes, they suffer from numerical challenges (Morrison et al., 2019) as well as from sensitivity to the bin discretization (Ghan et al., 2011). The purpose of the method presented here is to target the middle ground of complexity, between traditional bulk and bin methods, using more sophisticated numerical techniques.

To meet the needs of future climate and weather models, a microphysics scheme should maintain enough flexibility to function with a wide range of degrees of freedom and minimal structural uncertainty in the PSD representation. While bin-scheme complexity may be unattainable for GCMs in the near future, we still need a microphysics method that can maintain spectral details without the closure assumptions and conversion parameterizations required by moment-based bulk methods. Some recent efforts in microphysics modeling have focused on relaxing assumptions about the size distribution and process rates to reduce these structural uncertainties. One option, Lagrangian microphysics, directly tracks tracer particles known as superdroplets (Riechelmann et al., 2012; Andrejczuk et al., 2010, 2008; Shima et al., 2009), but it is far too computationally expensive for global or even regional models. A different moment-based method, the BOSS scheme proposed by Morrison et al. (2019) leaves all process rates and closures as generalized power series whose parameters are learned from data. Bieli et al. (n.d.) present a more efficient way to learn these parameters within a similar bulk microphysics framework that still relies on closures. More complex yet, Rodríguez Genó and Alfonso (2022) tackle the challenge of inverting multimodal distribution closures using a machine-learning based method, which could avoid the necessity for cloud-rain conversion rate parameterizations. However, these bulk methods cannot function in a wide range of computational degrees of freedom, nor do they provide complete spectral details about the PSD that might alleviate uncertainties about conversion between hydrometeor types. One solution is to think beyond the classical bulk versus bin representations of the PSD, leveraging numerical techniques developed for fluid mechanics.

In this study, we present and test a novel way to span the gap in complexity between bin and bulk microphysics methods by applying the collocation method with ba-

sis functions (BFs) to represent the particle size distribution. (For simplicity, it will be referred to going forward as the BF method.) Finite element methods such as collocation have been historically overlooked for microphysics applications, with the exception of Gelbard and Seinfeld (1978)'s demonstration using collocation of quartic or cubic polynomials, which was never widely adopted in favor of contemporaneous bin methods. More recent results from the applied math community suggest that combining collocation with radial basis functions, rather than polynomials, is a promising numerical technique for advection problems (Zhang et al., 2000; Franke & Schaback, 1998). This work extends the basis function collocation technique to the integro-differential equations encountered in microphysics. Beyond retaining spectral details of the PSD, the BF method has appealing extremes of complexity: at low resolutions, the method is effectively a linear closure, as in the context of bulk schemes; at moderate or high resolutions, it converges toward a smoothed bin scheme (replicating a bin scheme exactly if constant piecewise BFs and appropriate numerics are used). Therefore collocation of basis functions promises greater flexibility than either bulk or bin methods alone, while retaining desirable aspects such as low-to-moderate complexity and spectral predictions. This paper describes the method and presents results of applying the method to droplet collision and coalescence, benchmarked against commonly used bulk, bin, and Lagrangian frameworks. We additionally address some limitations posed by the method that are specific to the context of tracking a PSD, such as mass non-conservation and a finite size range. Overall, the BF method improves spectral PSD predictions in a box model as well as simple precipitation predictions, measured as a size exceedance, compared to a three-moment bulk method, and with fewer degrees of freedom than a bin method. Furthermore, the runtime complexity of the method scales quadratically with the number of degrees of freedom, making it just as efficient as or faster than a bin method.

The remainder of this paper is organized as follows: section 2 describes the method of collocation of basis functions to approximately solve the population balance equation for collision-coalescence in microphysics, and section 3 describes a set of microphysics box model case studies. Section 4 compares the accuracy of these case studies solved using basis functions, bulk, bin, and Lagrangian schemes, and discusses the computational complexity of these methods. Finally, section 5 concludes the paper and suggests potential improvements and applications.

2 Method Description

2.1 Key Equations

The governing equations for microphysics describe a population balance for the droplet size distribution. The governing equation for collision-coalescence, also called the Smoluchowski or Stochastic Collection Equation (SCE), is given by

$$\partial_t n(x,t) = \frac{1}{2}\int_0^x n(x-y,t)n(y,t)K(x-y,y)E_c(x-y,y)dy \\ -n(x)\int_0^\infty n(y,t)K(x,y)E_c(x,y)dy, \tag{1}$$

where $n(x,t)$ represents the number density of particles of mass x at time t, $K(x,y)$ is the collision rate of particles of masses x and y, and $E_c(x,y)$ is the coalescence efficiency of said collision. The first integral represents production of droplets of size x from two smaller droplets, and the second integral represents loss of droplets of size x due to coalescence with other droplets.

Other microphysical processes such as condensation, evaporation, sedimentation, and aerosol activation also affect the PSD. To demonstrate the proposed BF method for microphysics, we initially focus on only the coalescence process as in equation (1). The SCE is notoriously difficult to solve numerically, as it is an integro-partial differential equa-

tion and frequently involves rapid acceleration of particle growth, yet this mechanism is crucial to determining the onset of rain and drizzle (Stephens et al., 2010). Later, we will also consider two non-collisional processes of sedimentation and injection of new particles. Sedimentation is defined as removal all particles above a size threshold $x_{\max}$, which can prevent unphysically rapid acceleration of collisions. Sedimentation is enforced by limiting the upper bound of each integral to $x_{\max}$, effectively truncating the PSD to have a value of $n(x > x_{\max}, t) = 0$. We can alternatively prevent particles larger than the maximum size $x_{\max}$ from forming by rejecting those collisions in a mass conserving manner. The appropriate upper bound for the second integral in this case is $x_{\max} - x$ (Filbet & Laurençot, 2004). When such collisions are not rejected and particles exit the system, we introduce new droplets to the system, mimicking entrainment or activation of new particles. The rate of particle injection $P_{\text{inj}}(x, t)$ is given by

$$P_{\text{inj}}(x, t) = \dot{P} I(x) \tag{2}$$

where $I(x)$ represents a normalized size distribution of the injected droplets, which might be smaller than the average droplet in the system, and $\dot{P}$ is the rate of particle injection.

2.2 Collocation of Basis Functions with Positivity Constraint

In our proposed method, based on the work of Zhang et al. (2000), the PSD is approximated by a weighted sum of n_{BF} basis functions:

$$n(x, t) \approx \tilde{n}(x, t) = \sum_{k=1}^{n_{\text{BF}}} c_k(t) \phi(x; \theta_k) = \mathbf{c}(t) \cdot \boldsymbol{\phi}(x). \tag{3}$$

We denote the approximate solution $\tilde{n}(x, t)$, the collocation weights $c_k(t)$, and the basis functions $\phi(x|\theta_k)$ where ϕ is the functional form and θ_k are the parameters of the k-th BF (for instance, mean and variance of a Gaussian). In the collocation method, one such parameter is the center or mean of the basis function, $\mu_k \in \theta_k$, known as the collocation points. In the context of microphysics, these collocation points refer to particle masses, which locate the mode of each basis function. In equation (3), we have also compactly rewritten the BFs and weights in vector form as $\boldsymbol{\phi}(x) = \big(\phi(x|\theta_1), \phi(x|\theta_2), \ldots, \phi(x|\theta_{n_{\text{BF}}})\big)$ and $\mathbf{c}(t) = \big(c_1(x), c_2(x), \ldots, c_{n_{\text{BF}}}(x)\big)$.

Since the basis functions have a fixed shape over the droplet size range, evolving the approximate PSD reduces to solving for $\mathbf{c}(t)$ in time as a system of ordinary differential equations. Because liquid water is a conserved quantity in the absence of evaporation/condensation, we consider the evolution of the local mass density $m(x, t) = x\, n(x, t)$ rather than the local number density. Thus although we use basis functions to approximate the number density, the equations are evolved in time based on local mass density, as in a one-moment bulk method or a standard flux-method bin scheme.

Denote the vector of approximate mass density at the collocation points μ_k to be $\tilde{\mathbf{m}}(t) = \big(\mu_1 \tilde{n}(\mu_1, t), \ldots, \mu_p \tilde{n}(\mu_p, t)\big)$. At each timestep, recovering the weights from the interpolated collocation points requires solving for $\mathbf{c}(t)$ in the linear system

$$\tilde{\mathbf{m}}(t) = \boldsymbol{\Phi} \cdot \mathbf{c}(t) \tag{4}$$

where $\boldsymbol{\Phi}$ is a $n_{\text{BF}} \times n_{\text{BF}}$ matrix, with elements $\Phi_{jk} = \mu_j \phi_k(\mu_j)$ representing the mass density of the k-th basis function evaluated at the jth collocation point. For a linearly independent set of basis functions, this system is well-posed and guarantees a unique solution. However, it may be ill-conditioned, particularly when the choice of basis function has global rather than compact support (Zhang et al., 2000).

The approximate solution is initialized by projecting the initial mass distribution onto the basis space. This projection comes from solving an optimization problem:

$$\min_{\mathbf{c}(0)} \quad \|\boldsymbol{\Phi} \cdot \mathbf{c}(0) - \tilde{\mathbf{m}}(0)\|^2 \qquad \text{s.t.} \quad \mathbf{c}(0) \geq 0\,. \tag{5}$$

The positivity constraint mathematically enforces the fact that the PSD should be non-negative at all points. Equation 5 is formulated as a quadratic optimization, and therefore can be solved efficiently numerically.

This projection could additionally incorporate a mass conservation constraint, both initially and at every future time step, but at significantly higher cost than solving the linear system in equation 4. Additionally, since the exact solution to the equation does not necessarily exist as a projection of the basis functions, the mass and positivity constraints in the optimizer can lead to unphysical solutions as the approximate PSD evolves in time. While relaxing this constraint might lead to an artificial reduction or increase in mass throughout the simulation time, it allows a more efficient nonnegative least-squares solution. In developing this method, we observed that evolving the linear system in mass density with a positivity constraint, rather than using number density directly, led to more physical and realistic PSDs compared to including a mass-conserving constraint at all times.

2.2.1 Interpretability and design choices

The method described above generalizes to solve many categories of differential equation, but selecting the basis functions and parameters θ_k requires care in order to preserve physical properties of a droplet distribution. To model a droplet PSD, we choose to let the basis functions themselves be distributions, in contrast to the cubic splines employed by Gelbard and Seinfeld (1978) or spectral element methods. If we choose Gaussian or lognormal BF's collocated on a grid of droplet sizes, each BF effectively represents a droplet size mode. This feature provides a useful analogy to aerosol size modes, or cloud versus rain droplet distributions, much as a typical bin scheme will distinguish between aerosol, cloud, and rain size bins, or how a moment scheme will have a separate set of moments for cloud and rain water. In fact, this representation is a generalization of bin schemes, which can be considered piecewise constant basis functions: $\phi_k(x) = 1, x \in \{x_k, x_{k+1}\}$ (see figure 1). At low resolution, the BF representation can similarly be thought of as approximating a linear closure, as in the method of moments (MOM), where the prognostic variable is the first moment calculated over sub-intervals of the particle size range.

Additional design choices include selecting the collocation points and additional hyperparameters of the BFs, such as the variance for lognormal or Gaussian distributions. An in-depth description and justification of the BF setup used in following sections can be found in Appendix A. Notably, we introduce a compactly-supported BF that approximates a lognormal distribution (CSLBF1: equation A1), use exponentially-spaced collocation points, and set the geometric standard deviation as the distance between adjacent collocation points.

2.3 Application to the SCE and microphysical processes

The equations involved in applying the BF method to the SCE are derived in appendix B, with the result summarized by equation 6 below:

$$\begin{cases} d_t \tilde{\mathbf{m}}(t) = \mathbf{c}(t) \cdot \mathbf{Q} \cdot \mathbf{c}(t) + \sum_{l=1}^{N_{proc}} \mathbf{P}_l \\ \mathbf{\Phi} \cdot \mathbf{c}(t) = \tilde{\mathbf{m}}, \qquad \text{with} \quad \mathbf{c}(t) \geq 0 \end{cases} . \tag{6}$$

In this equation, third-order tensor $\mathbf{Q}$ and vectors $\mathbf{P}_l$ are obtained by taking various inner products of the collision kernel and additional process rates (respectively) with the basis functions. All integrals for this collision-coalescence term can be pre-computed for a fixed set of basis functions, defining these tensors through numerical integration and projection of rate processes onto the basis space. (The required precomputations and scaling of these computations with the number of BFs are described in Appendix B. In summary, the precomputation steps scale at most cubically with the number BFs, and

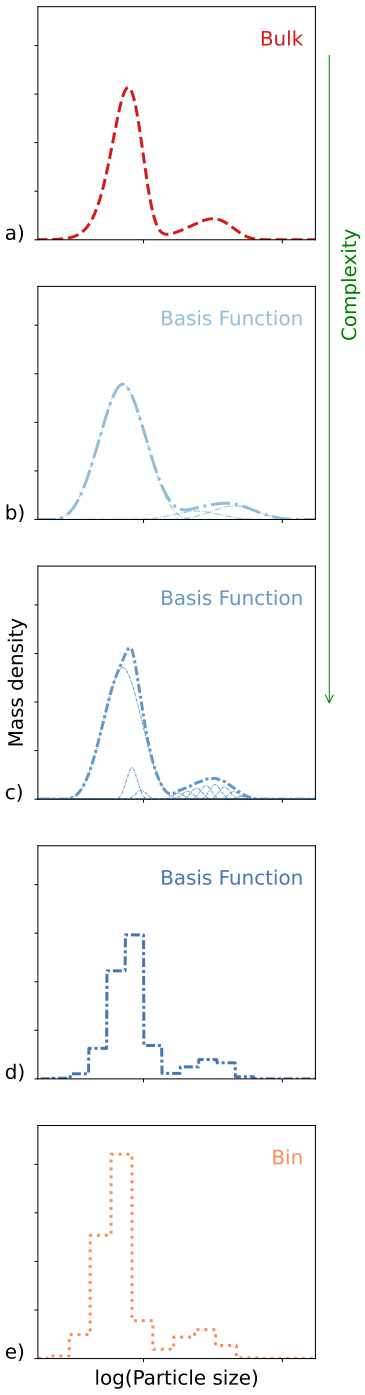

Figure 1. Illustration of the way that the collocation of basis functions can span the gap from bulk to bin microphysics. The PSD for a two-mode gamma mixture of particles, corresponding, for instance, to a cloud and rain mode, is plotted as it would be represented in a: (a) 3-moment bulk scheme with gamma closure (one set of moments for each mode); (b) 4 lognormal basis functions; (c) 16 lognormal basis functions; (d) 16 piecewise-constant basis functions; (e) bin method with 32 bins.

Case	Dynamics	Parameters	Duration	Initial/Injection Distribution
1C	Constant kernel	$A = 10^{-4}\text{cm}^3\text{s}^{-1}$	360s	Gamma
1G	Golovin kernel	$B = 1500\text{s}^{-1}$	4hr	$\theta = 100\mu\text{m}^3$
1H	Hydrodynamic kernel	$C = \pi \times 10^{-9}\text{cm}^{-3}\mu\text{m}^{-4}\text{s}^{-1}$	360s	$N_0 = 100\text{cm}^{-3}, k = 3$
2	Golovin kernel	$B = 1500\text{s}^{-1}$	4hr	Gamma mixture $N_{0,a} = 100\text{cm}^{-3}$ $k_a = 4, \theta_a = 100\mu\text{m}^3$ $N_{0,b} = 100\text{cm}^{-3}$ $k_b = 2, \theta_b = 15\mu\text{m}^3$
3	Golovin kernel, Injection, Precipitation	$B = 1500\text{s}^{-1}$ $P_{\text{inj}} = 1\text{cm}^{-3}\text{s}^{-1}$ $x_{\max} = 1000\mu\text{m}^3$	4hr	Gamma distribution, $\theta = 100\mu\text{m}^3$ $N_0 = 0, k = 3$

Table 1. Summary of the dynamics, parameters, and initial or injection distributions employed for each box model test case.

the computation at each time step scales cubically or quadratically depending on the basis chosen.) The result is a simple set of quadratic coupled ordinary differential equations for the mass density at the collocation points, $\tilde{\mathbf{m}}(t)$, and the BF weight vector $\mathbf{c}(t)$.

3 Test Cases

To compare the accuracy and efficiency of the proposed BF method with bin, bulk, and Lagrangian microphysics schemes, we use three sets of dynamics in a zero-dimensional box. The parameters for these experiments are summarized in table 1.

The first test case evolves a PSD with collision-coalescence dynamics only, beginning from a single droplet size mode following a gamma distribution with number density N_0, shape parameter k and scale parameter θ. We consider a constant collection efficiency $E_c = 1$ and three separate collision kernels: (1C) a constant rate of collision $K_C(x, y) = A$; (1G) a Golovin linear kernel (Golovin, 1963) $K_G(x, y) = B(x+y)$, and (1H) a hydrodynamic kernel $K_H(x, y) = C(r(x) + r(y))^2|a(x) - a(y)|$, where $r(x)$ and $a(x)$ represent the particle radius and area respectively. Collision kernel parameters and time of simulations are chosen such that the final droplet spectrum has approximately 1/3 the number density of the initial spectrum. We investigate the PSD (mass density) following collisions, as well as the first three moments of the PSD which correspond to total number density, mass density, and radar reflectivity. Spectral errors are calculated as a sum of squared differences in the approximated profiles and a reference solution from Lagrangian microphysics. In addition, we calculate the percent mass exceedance over a droplet-size threshold of $x_{\max} = 1000\mu\text{m}^3$. This exceedance can be considered a proxy for precipitation, even though all mass remains in the box.

The second test case retains the Golovin collision kernel but uses a two-mode initial distribution: a sum of two gamma distributions. This initial distribution can be thought of as representing two aerosol modes, or alternatively a cloud mode and rain droplet mode. A simple closure-based bulk representation cannot capture multiple modes without an additional set of prognostic moments and autoconversion rates; therefore, this test case highlights the information gained from using a more flexible PSD representation.

The third test case incorporates additional dynamics of particle injection and precipitation from the box. Given a constant prescribed injection rate, this set of dynam-

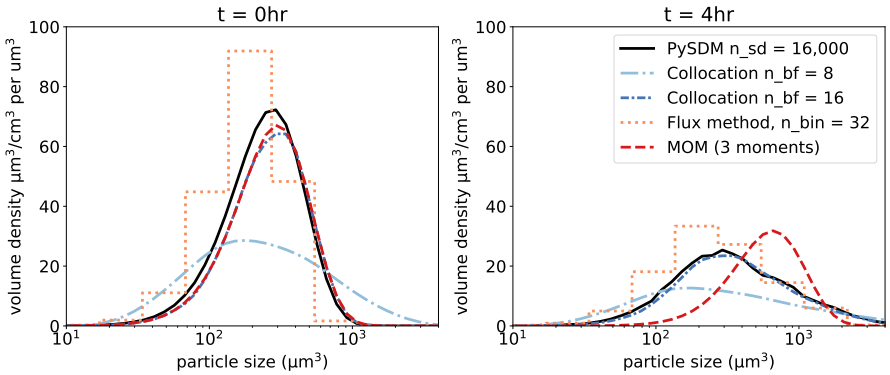

Figure 2. Initial spectrum (left) and post-collision spectrum (right) resulting from a Golovin kernel collision-coalescence (1G) for bulk (MOM), bin (flux), and Lagrangian methods, and using the BF (collocation) method with 8 or 16 degrees of freedom.

ics will drive the PSD to a steady state in which particles enter the system, collide, grow, and precipitate out of the system. While modeling collision-coalescence by itself is a useful numerical test, it requires that the microphysics scheme be able to represent arbitrarily large particles with an accelerating rate of growth. Using a simplified proxy for the introduction of small droplets and removal of large droplets allows for a more physically realistic particle size distribution and time scale of dynamics.

We solve each test case numerically using the flux method for spectral bin microphysics with 32 bins (Bott, 1998), a three-moment gamma-closure method of moments (Bieli et al., n.d.), and collocation of BFs with varying numbers of basis functions, referred to as the degrees of freedom. The bin method used follows the original setup from Bott (1998), spanning a range of $1.06\mu m^3$ to $2.28\times10^9\mu m^3$ with mass doubling between bins. Additionally we include results from a Lagrangian particle-based code called PySDM (v2.5) (Bartman et al., 2022) as a high-resolution reference for the first three cases. The PySDM simulations use 16,384 superdroplets to represent the particle population in a box of volume $1m^3$.

The BF method as demonstrated here uses 8 or 16 CSLBF1 basis functions to span a particle size range of $8\mu m^3$ to $125{,}000\mu m^3$, which corresponds to 15 of the 32 bins used in the bin approach. Collocation points are logarithmically spaced over this size range. Particles are assumed spherical with liquid water density. BF shape parameters θ_k are chosen such that the basis functions overlap with their nearest neighbors: $\theta_k = \mu_k - \mu_{k-2}$ and $\theta_1 = \theta_2 = \mu_2$. The method is implemented in the Julia programming language and uses a variable time-step with the DifferentialEquations.jl package (Rackauckas & Nie, 2017). The inversion is solved using NonNegLeastSquares.jl v0.4.0 (non-negative least squares). Numerical integrals are computed using Cubature.jl v1.5.1.

4 Results

4.1 Case 1: Unimodal collision-coalescence

For the collision-coalescence only box case, we are interested in the ability of each microphysics method to accurately predict: (1) the PSD; (2) the first three moments of the distribution; and (3) the number of particles above a particular size threshold. The spectra in figure 2 reveal that more than eight basis functions are necessary for this particular BF configuration to approximate the initial condition's primary size mode. From

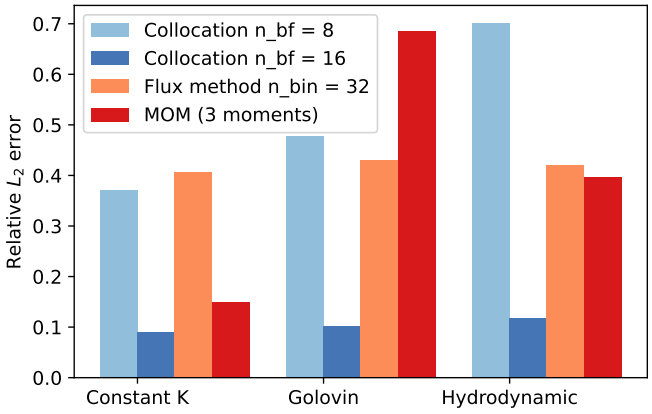

Figure 3. Spectral error (L_2) for the bulk, bin (flux), and BF methods with 8 or 16 basis functions, computed relative to a Lagrangian PySDM result. Errors are shown for each of three coalescence-only experiments using a constant, Golovin, and hydrodynamic kernel (case 1C, 1G, and 1H, respectively).

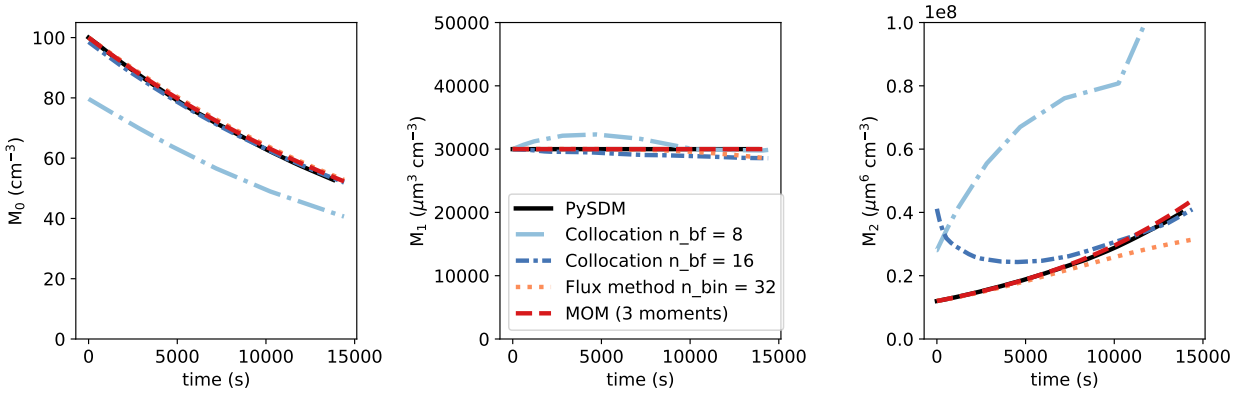

Figure 4. Evolution of the first three moments of the PSD over time for bulk, bin, and BF method with 8 or 16 degrees of freedom for the Golovin collision kernel (1G).

the final spectra in figure 2 for the Golovin kernel, as well as the summarized spectral errors in figure 3, we find that the collocation method with 8 BFs performs on par with a flux bin method, and with 16 BFs it outperforms both a bin and bulk method in predicting the post-collision spectra. The bin method consistently has a spectral error of around 40% relative to the Lagrangian results, owing in part to numerical diffusion, while the bulk method of moments has an error which varies significantly according to the complexity of the collision kernel. While the 8-BF collocation approach suffers from this same challenge, using the BF approach with 16 degrees of freedom results in consistently small spectral errors less than 15% for all three collision kernels investigated. These results demonstrate the potential for the collocation method to resolve realistic droplet spectra using the same or fewer degrees of freedom than a traditional bin method.

Next we investigate bulk quantities predicted by each method in figures 4 and 5, which illustrate the time evolution of the first three moments and exceedance mass, re-

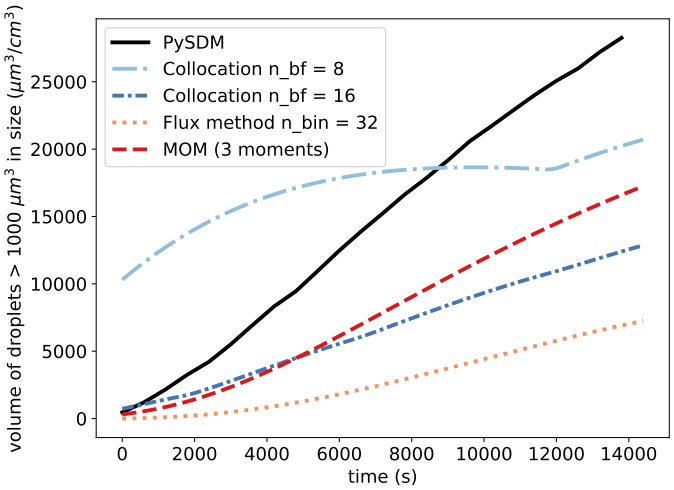

Figure 5. Volume of droplets exceeding $1000\mu\text{m}^3$ in size for Lagrangian, bulk, bin, and collocation methods as a function of time for a Golovin collision kernel (1G).

spectively. The bulk method of moments outperforms the BF method in predicting the time evolution of the PSD moments, as the first two moments are predicted as prognostic moments analytically, and the gamma closure approximation is only employed in computing the second moment. The BF method does not exactly conserve mass in the lower-resolution case, in part because the use of compactly supported basis functions prevents the representation of particles larger than the support of the basis functions ($125{,}000\mu\text{m}^3$ in this case). Larger particles may form according to the physics of the collision-coalescence equation; therefore the BF method encounters error in the tail of the spectral representation, and especially in the higher-order moments as a result. Furthermore, the matrix inversion in equation 4 does not guarantee conservation of mass, particularly where the system of equations might be large and ill-conditioned. The second moment is overestimated by the BF method initially due to error in projecting the initial PSD onto the basis space: the initial projection slightly overpredicts the size of some droplets, but not so much as to miscategorize them in the exceedance regime larger than $x_{\max} = 1000\mu\text{m}^3$ in the higher-resolution BF case, as indicated in figures 2 and 5. Indeed, despite shortcomings in predicting PSD moments, the BF method does comparably well or better than bin or bulk methods at capturing the mass of particles which lie in the tail of the distribution (figure 5). It is apparent that the conversion of small particles to medium or larger particles is adequately captured by the BF method. All methods underpredict the exceedance volume relative to the Lagrangian superdroplet method at longer times, but the BF approach displays comparable accuracy to the bulk method of moments and outperforms the bin method.

4.2 Case 2: Multimodal collision-coalescence

One strength of the BF method is its ability to represent up to n_{BF} modes of a PSD, where n_{BF} is the number of basis functions used. By contrast, bulk methods can represent at most one droplet mode, and bin methods lose spectral detail of the modes due to the piecewise constant representation of the PSD. We demonstrate in figure 6 an example of collision-coalescence with an initially bimodal distribution: the second mode initially has a narrower and more peaked structure, which broadens and extends as these larger particles collide more rapidly according to the Golovin collision dynamics. The

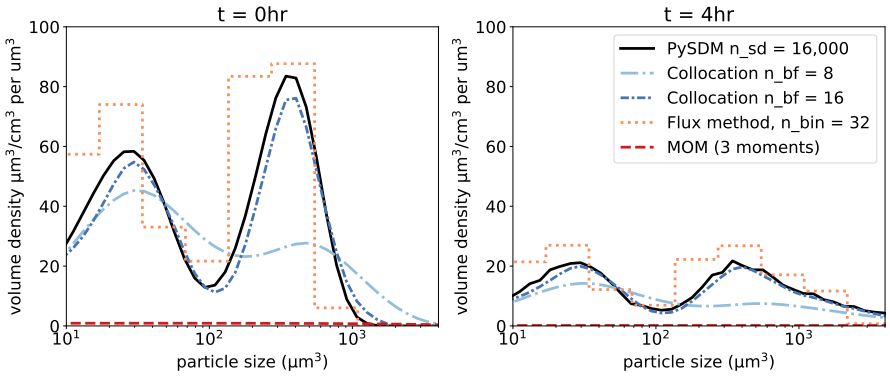

Figure 6. Spectra following collision-coalescence of a bimodal droplet population using Lagrangian, bulk, bin, and BF methods with a Golovin kernel (case 2).

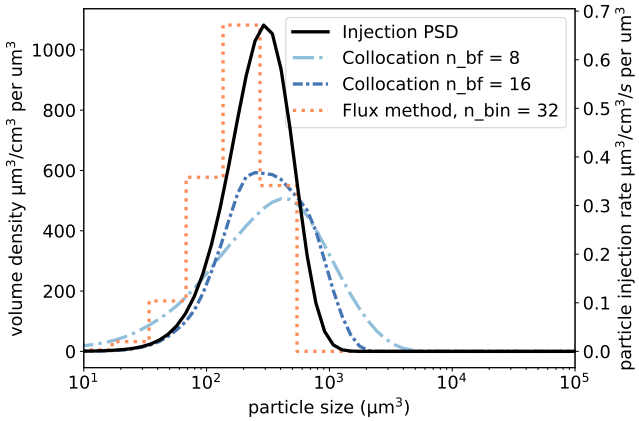

Figure 7. Steady state PSD for the third case with collisions, sedimentation, and injection, using a bin method and the BF method with 8 or 16 basis functions. The PSD of injected particles is plotted as a solid black line with units on the right y axis.

BF method accurately captures both of these modes during the PSD evolution, while the bulk method with a gamma-closure cannot represent the initial or final PSD due to the underlying unimodal closure assumption. Furthermore, while the bin method accurately predicts droplets in both size ranges, the BF method is able to do so with fewer degrees of freedom and yields a closer and more interpretable spectral match to the Lagrangian results.

4.3 Case 3: Collision-coalescence with injection and removal

When including removal of large particles and introduction of small particles, we investigate the steady-state PSD as well as the time evolution of the PSD moments. The Lagrangian and method of moments simulations are excluded in this case, as the removal and injection process rates used are not applicable in these frameworks. As seen in figure 7, the BF method solution is a broadened image of the injected PSD, as expected: particles enter the system, grow through collisions, and exit once they reach $1000\mu m^3$ in size. As in previous cases, the BF solution is slightly narrower when more degrees of

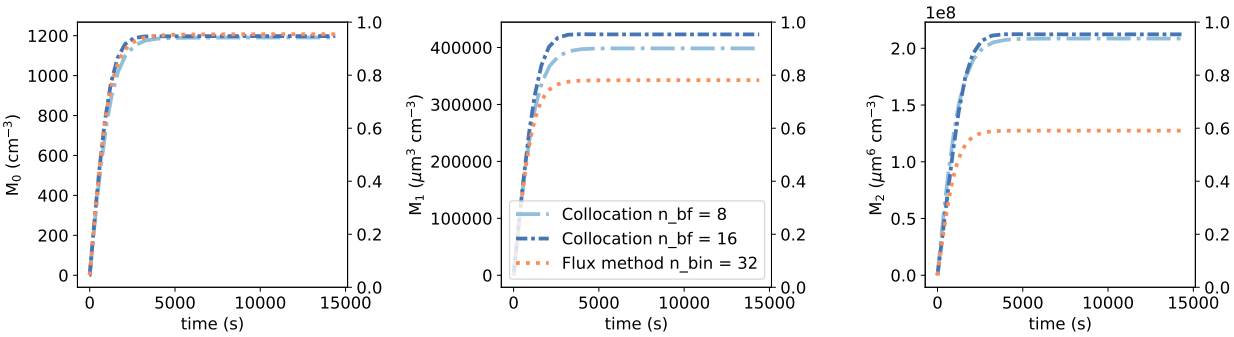

Figure 8. Time-series evolution of first three moments of the distribution for the collocation and bin methods with collisions, precipitation, and injection.

freedom are used, but the lower-resolution BF case does not display as large of a discrepancy when large particles are removed. The bin approach underpredicts the steady state distribution of larger particles, in part because the piecewise constant representation leads to over-removal of particles in the largest bin. This shortcoming is further demonstrated in the higher-order moments of the system in figure 8. Both the BF and bin methods converge to a steady state on the same time scale, with the same number density, but the bin method underpredicts the first and second moments relative to even the lower-resolution BF method, demonstrating the improvement possible from using nonlinear distributions as a basis, rather than piecewise bins. This more realistic set of dynamics, which removes large particles from the system, demonstrates that the BF method using collocated compactly supported basis functions is well-suited for representing a complete microphysical system at longer GCM-relevant time scales. These results illustrate a tradeoff between higher accuracy in the spectrum and moments from using more basis functions when coalescence dominates, and the ability to use a lower complexity setup with fewer (e.g., 8 versus 16) degrees of freedom when additional compensating dynamics are considered.

4.4 Computational Complexity

The BF method offers similar or improved computational scaling relative to the bin spectral method, but higher complexity than a traditional multimoment bulk method. Bulk methods with a closure assumption scale with the number of moments, $\mathcal{O}(N_{\text{mom}})$ when the relationship between the prognostic moments and PSD parameters is known, but more complex PSD closures may require nonlinear operations or even optimization, leading to a more computationally intensive operation at each time step. Spectral bin methods such as the flux method used here (Bott, 1998) scale quadratically with the number of bins, $\mathcal{O}(N_{\text{bin}}^2)$, as each pair of bins is considered sequentially. The basis function method scales either cubically or quadratically depending on the choice of basis (see appendix B). While the initial precomputation for the BF method is cubic in the number of basis functions, a compactly supported basis will lead to quadratic operations in the forward time-marching of equation 6, as the third-order tensor $\mathbf{Q}$ is sparse. This places the BF method at the same order of complexity as other spectral methods, $\mathcal{O}(N_{\text{BF}}^2)$.

5 Discussion and Conclusions

This paper describes and demonstrates a novel method to represent the particle size distribution of droplets for atmospheric microphysics. Collocation of basis functions

provides a more flexible PSD approximation than either bin microphysics or the method of moments with closure (bulk microphysics). In particular, selecting BFs which are themselves distributions generalizes traditional spectral bin methods to a smoothed representation that can be interpreted as the sum of droplet size modes. The method is also appropriate for applications where more than three degrees of freedom (the most usually provided in a bulk scheme) are desired, but where full bin complexity is infeasible. In this low-resolution limit, collocation of basis functions can be considered a form of linear closure relating the mass density at the collocation points to a BF weight vector.

Tested in a variety of box model settings, we find that the BF method improves spectral accuracy under collision-coalescence dynamics compared to a three-moment bulk method, while using fewer degrees of freedom than a bin method. The spectral detail from the BF approach allows for a more precise calculation of water mass in the tail of the distribution (exceedance), which could avert the need for precipitation parameterizations that are required by bulk methods. Another strength of the method is its ability to represent multimodal distributions, unlike 3-moment bulk methods. At short time scales with rapidly accelerating collisions, the BF method suffers from numerical challenges and compact support that cannot represent arbitrarily large particles; therefore, we propose an additional set of dynamics that allows removal of large particles from the box. In this collision-injection-removal case study, the BF method outperforms a bin scheme in computing the steady-state distribution, and it requires fewer degrees of freedom.

In general, the BF method is a more flexible framework than bulk or bin methods: the suggested implementation can receive an arbitrary set of microphysical processes and automatedly perform all required numerical integrations. This is in contrast to bin methods, which require tabulated collision and breakup kernels that are dependent on the bin discretization, and in contrast to bulk methods, which frequently include hard-coded parameterizations and closures. This ability to specify arbitrary functional process rates for the BF method will be especially useful for reducing microphysics parameter uncertainty while also improving the structural PSD representation.

The BF method does have limitations. First, although the linear system in equation 4 is solved in mass density space with a positivity constraint, the method does not exactly conserve mass for a collision-coalescence-only set of dynamics. When employed with compactly supported basis functions, the method can only represent particles up to a maximum size, unlike bulk or Lagrangian methods. This shortcoming manifests in errors in the higher order moments of the PSD, including some mass loss from the system (figure 4). Solutions could involve allowing for some globally supported basis functions, or periodically rescaling the weight vector to exactly conserve mass in the system. Nevertheless, despite this limitation, the method is able to predict both spectral details and moments when particle removal and injection are considered; therefore, further refinement may be unnecessary to describe a full set of microphysical processes. Future work to improve and test this novel microphysics method will involve incorporating additional microphysical processes, as well as employing one, two, and three-dimensional simulations to test the ability of the method to reproduce mesoscale cloud properties. Further testing of the method in a one-dimensional setting with spatial advection will be necessary to assess how susceptible the collocation implementation is to numerical diffusion, as is often observed with bin schemes.

The BF method presented here improves spectral accuracy at a lower cost per degree of freedom than bulk or bin methods, and it has the potential to reduce the computational cost of microphysics even further. Using inspiration from proposed moving bin schemes, the locations or shapes of BFs could be periodically updated to maximize the information potential provided by only n_{BF} degrees of freedom. While this approach would impose a higher cost of recomputing numerical integrals, it would cluster basis functions near the most-weighted droplet modes, improving the accuracy-complexity tradeoff. Another potential benefit of the collocation representation is the ability to use mul-

tidimensional basis functions: one independent variable could be the droplet size, as in this work, while other particle properties such as aerosol hygroscopicity, ice riming fraction, or surface tension could occupy additional inputs. This multidimensional representation has been explored for aerosol bin schemes (Lebo & Seinfeld, 2011), as well as for ice bulk methods (Morrison & Milbrandt, 2015). However, it may be more computationally efficient to represent multiple particle properties in the BF framework due to the flexibility of selecting basis functions as well as using compact support to generate a sparse system and lessen the computational burden. Such a representation could eliminate the uncertainties of conversion parameterizations and of information loss from aggregating particles into categories with distinct sets of dynamics. This potential provides a path toward unifying the numerical representation of all microphysical particles in a single, consistent framework.

Appendix A Basis functions, collocation points, and hyperparameters

The BF collocation parameters demonstrated in this study are briefly explained. As the collocation points correspond to the droplet mode represented by each BF, we should not assume a priori any particular initial or final distribution of particles. However, we can use the inherent length scales of the physical system to aid the setup. For cloud droplets and aerosols, the size domain should extend from $x_{\min} \geq 0\mu$m to the size of the largest particles $x_{\max}$ that do not sediment out of the system or instantaneously break up, hence making a finite domain approximation reasonable. Furthermore, we draw inspiration from bin microphysics to suggest logarithmically spaced collocation points over the domain.

The basis function family and their hyperparameters should then be selected to ensure a few criteria:

1. The entire domain $[0, x_{\max}]$ is spanned with some minimum probability.
2. There should be no particles with negative or infinite mass; that is, $\phi_k(x < 0), \phi_k(x \to \infty) = 0$ for all basis functions.
3. BF hyperparameters should be selected to minimize oscillations and jumps in the approximated distribution.

The first condition is equivalent to requiring either globally-supported BFs, such that $\phi(x) > 0\ \forall x$, or sufficient overlap of compactly-supported BFs, which are positive over some interval and zero elsewhere. The second condition cannot be met exactly for any BFs that are globally supported over $(-\infty, \infty)$, therefore we suggest using either compactly-supported BFs (CSBFs) or exponentially decaying BFs. CSBFs are additionally recommended due to their favorable numerical properties: Zhang et al. (2000) demonstrate that CSBFs result in a better conditioned system of equations (as in equation 5). The third criterion is the trickiest and will depend on the family of BFs chosen. As a simple heuristic for a two-parameter family such as Gaussians, we suggest setting the scale factors as some multiple of the spacing between collocation points to ensure support and smoothness over the domain. More sophisticated methods of setting the hyperparameters, such as optimization over a set of potential distributions or constraints on fluctuations in the second derivatives, are possible but beyond the scope of this paper.

Several families of basis functions are suitable to approximate a droplet size distribution, such as Gaussian, gamma, and lognormal distributions. In order to obtain a compactly supported basis, however, we propose to use a version of the CSRBF1, a compactly supported Gaussian approximation proposed by Wu (1995), modified to instead uses a logarithmic argument. This basis function, which we will refer to as CSLBF1 (compactly supported lognormal BF 1) takes the form:

$$\phi(r) = \begin{cases} \frac{12}{35}(1-r)^4(4+16r+12r^2+3r^3)\frac{dr}{dx} & r \leq 1 \\ 0 & r > 1 \end{cases} \tag{A1}$$

with argument

$$r = \frac{|\log(x) - \mu|}{\theta}$$

where μ is the collocation point and θ is a scale factor. Given that CSRBF1 approximates a normal distribution, CSLBF1 approximates a lognormal distribution, which is better suited to particle distributions as it is right skewed.

Appendix B Collocation of BFs for the SCE

Evaluating equation 1 with arbitrary additional processes $\mathbf{P_l}$ in mass density at collocation point μ_j, we find:

$$\begin{aligned} \partial_t \mu_j n(\mu_j, t) =& 1/2\mu_j \int_0^{\mu_j} n(\mu_j - y, t) n(y, t) K(\mu_j - y, y) E(\mu_j - y, y) dy \\ & - \mu_j n(\mu_j, t) \int_0^{x_{\max} - \mu_j} n(y, t) E(\mu_j, y) K(\mu_j, y) dy + \sum_{l=1}^{N_{proc}} P_l(\mu_j, n(\mu_j, t)) \end{aligned} \tag{B1}$$

Substituting the collocation approximate solution for local mass density, $x\tilde{n}(x,t) = \sum_{k=1}^{p} x\phi_k(x)c_k(t)$, this time derivative becomes:

$$\begin{aligned} \partial_t \tilde{m}_j(t) =& 1/2 \sum_{k=1}^{n_{\mathrm{BF}}} \sum_{l=1}^{n_{\mathrm{BF}}} \mu_j c_k(t) c_l(t) \int_0^{\mu_j} \phi_k(\mu_j - y) \phi_l(y, t) K(\mu_j - y, y) E(\mu_j - y, y) dy \\ & - \sum_{k=1}^{n_{\mathrm{BF}}} \sum_{l=1}^{n_{\mathrm{BF}}} \mu_j c_k(t) c_l(t) \phi_k(\mu_j) \int_0^{x_{\max} - \mu_j} \phi_l(y) K(\mu_j, y) E(\mu_j, y) dy + \sum_{l=1}^{N_{proc}} \mu_j P_l(\mu_j, \tilde{n}(\mu_j, t)) \end{aligned} \tag{B2}$$

The collision-coalescence dynamics are summarized via a third-order tensor in mass density: $\mathbf{Q}$, with

$$Q_{jkl} = 1/2\mu_j \int_0^{\mu_j} \phi_k(\mu_j - y) \phi_l(y, t) K(\mu_j - y, y) E(\mu_j - y, y) dy - \mu_j \phi_k(x\mu_j) \int_0^{x_{\max} - \mu_j} \phi_l(y) K(\mu_j, y) E(\mu_j, y) dy \tag{B3}$$

The overall dynamics are then summarized by cubic collision-coalescence dynamics plus the additional processes projected onto the basis space as in equation 5 to obtain the terms $\mathbf{P}_l = \left(\mu_1 P_l(\mu_1), \mu_2 P_l(\mu_2), \ldots, \mu_k P_l(\mu_k)\right)$ in equation 6.

Many of the quantities in equation 6 can be precomputed and stored for a given set of basis functions. These precomputations include:

- The linear system, $\mathbf{\Phi}$;
- The third order tensor $\mathbf{Q}$ which can be computed numerically via quadrature or Monte Carlo integration, given a functional form of the kernel.
- Appropriate projection of additional processes onto the basis space to obtain $\mathbf{P}_l$. For the purpose of ensuring mass conservation, this may require computing the first moments of the basis functions over the integration window $[0, x_{\max}]$.
- The initial condition at the collocation points $\tilde{\mathbf{m}}(0)$.

The computation of $\mathbf{Q}$ scales cubically with the number of collocation points for globally supported basis functions, and quadratically for partially overlapping compactly supported basis functions. The dynamical system in equation 6 involves at most cubic

vector-tensor multiplication and function evaluations for the tensor-vector inner products, and therefore a small system of basis functions is more likely to be limited by the time-stepping scheme or matrix inversion than by the precomputation. Another advantage of choosing compactly supported basis functions is that the constant-collocation matrix $\boldsymbol{\Phi}$ can be N-diagonal (CSBF's that only overlap their nearest neighbors will result in a tridiagonal system, for example) thus making the inversion much more computationally efficient. Finally, using CSBFs limits the range of particle sizes to a finite domain, making numerical integration more straightforward.

Acronyms

BF Basis function (method)
CSBF Compactly supported basis function
CSLBF1 Compactly support lognormal basis function 1
GCM General circulation model
MOM Method of moments
PSD Particle size distribution
SCE Stochastic collection equation

Notation

$\boldsymbol{x}$ Particle mass or volume
$\boldsymbol{n(x,t)}$ Particle size distribution: number of particles of mass x in a volume of air at time t
$\boldsymbol{K(x,y)}$ Collision kernel: rate of collisions between particles of mass x and y
$\boldsymbol{E_c(x,y)}$ Coalescence efficiency for particles of mass x and y
$\boldsymbol{x_{max}}$ Particle size threshold; particles above this mass are removed from the system
$\boldsymbol{P_{inj}(x,t)}$ Injection rate of particles of size x at time t, given in number of particles per air volume per time
$\boldsymbol{\dot{P}}$ Injection rate, in number of particles per air volume per time
$\boldsymbol{I(x)}$ Normalized size distribution of injected particles
$\boldsymbol{\tilde{n}(x,t)}$ Approximate PSD using a basis function representation
$\boldsymbol{c(t)}$ Vector of basis function weights at time t
$\boldsymbol{\phi(x)}$ Vector of basis functions
$\boldsymbol{\theta_k}$ Hyperparameters of the k-th basis function
$\boldsymbol{\mu_k}$ Collocation point of the k-th basis function
$\boldsymbol{\tilde{m}(t)}$ Mass density of the k-th weighted basis function
$\boldsymbol{\Phi}$ Basis function mass density tensor: $\Phi_{jk} = \mu_j \phi_k(\mu_j)$
$\boldsymbol{Q}$ Third order collision kernel tensor in basis function space
$\boldsymbol{P_l}$ Vector of process rate l projected onto basis function space

Acknowledgments

We thank Anna Jaruga, Melanie Bieli, Clare Singer, Zach Lebo, and John Seinfeld for feedback, insights, and discussion. Additional thanks go to Jakob Shpund for providing access to the Bott flux method bin implementation. E. de Jong was supported by a Department of Energy Computational Sciences Graduate Fellowship. This research was additionally supported by Eric and Wendy Schmidt (by recommendation of Schmidt Futures) and the Heising-Simons Foundation. The implementation of basis function collocation and examples used in this work can be found in the package RBFCloud.jl at `https://doi.org/10.5281/zenodo.6536677`. The 3-moment bulk scheme uses the package Cloudy.jl,

available at `https://github.com/CliMA/Cloudy.jl`, and the Lagrangian microphysics package PySDM is available at `https://github.com/atmos-cloud-sim-uj/PySDM`.

References

Andrejczuk, M., Grabowski, W. W., Reisner, J., & Gadian, A. (2010). Cloud-aerosol interactions for boundary layer stratocumulus in the Lagrangian Cloud Model. *Journal of Geophysical Research: Atmospheres*, *115*(22). doi: 10.1029/2010JD014248

Andrejczuk, M., Reisner, J. M., Henson, B., Dubey, M. K., & Jeffery, C. A. (2008). The potential impacts of pollution on a nondrizzling stratus deck: Does aerosol number matter more than type? *Journal of Geophysical Research: Atmospheres*, *113*(19). doi: 10.1029/2007JD009445

Arakawa, A. (2004). The Cumulus Parameterization Problem: Past, Present, and Future. *Journal of Climate*, *17*(13), 2493–2525. doi: 10.1175/1520-0442(2004)017⟨2493:RATCPP⟩2.0.CO;2

Bartman, P., Bulenok, O., Górski, K., Jaruga, A., Łazarski, G., Olesik, M. A., ... Arabas, S. (2022). PySDM v1: particle-based cloud modeling package for warm-rain microphysics and aqueous chemistry. *Journal of Open Source Software*, *7*(72), 3219. doi: 10.21105/joss.03219

Berry, E. X. (1967). Cloud Droplet Growth by Collection. *Journal of Atmospheric Sciences*, *24*(6), 688–701. doi: 10.1175/1520-0469(1967)024⟨0688:CDGBC⟩2.0.CO;2

Berry, E. X., & Reinhardt, R. L. (1974). An Analysis of Cloud Drop Growth by Collection: Part I. Double Distributions. *Journal of the Atmospheric Sciences*, *31*(7), 1814–1824. doi: 10.1175/1520-0469(1974)031⟨1814:AAOCDG⟩2.0.CO;2

Bieli, M., Dunbar, O., de Jong, E. K., Jaruga, A., Schneider, T., & Bischoff, T. (n.d.). *An efficient Bayesian approach to learning droplet collision kernels: Proof of concept using "Cloudy", a new n-moment bulk microphysics scheme.* doi: 10.1002/essoar.10510248.1

Bott, A. (1998). A Flux Method for the Numerical Solution of the Stochastic Collection Equation. *Journal of the Atmospheric Sciences*, *55*(13), 2284–2293. doi: 10.1175/1520-0469(1998)055⟨2284:AFMFTN⟩2.0.CO;2

Filbet, F., & Laurençot, P. (2004). Numerical Simulation of the Smoluchowski Coagulation Equation. *SIAM Journal on Scientific Computing*, *25*(6), 2004–2028. doi: 10.1137/S1064827503429132

Franke, C., & Schaback, R. (1998). Solving partial differential equations by collocation using radial basis functions. *Applied Mathematics and Computation*, *93*(1), 73–82. doi: 10.1016/S0096-3003(97)10104-7

Gelbard, F., & Seinfeld, J. H. (1978). Numerical solution of the dynamic equation for particulate systems. *Journal of Computational Physics*, *28*(3), 357–375. doi: 10.1016/0021-9991(78)90058-X

Gettelman, A., Gagne, D. J., Chen, C.-C., Christensen, M. W., Lebo, Z. J., Morrison, H., & Gantos, G. (2021). Machine Learning the Warm Rain Process. *Journal of Advances in Modeling Earth Systems*, *13*(2), e2020MS002268. doi: 10.1029/2020MS002268

Ghan, S. J., Abdul-Razzak, H., Nenes, A., Ming, Y., Liu, X., Ovchinnikov, M., ... Shi, X. (2011). Droplet nucleation: Physically-based parameterizations and comparative evaluation. *Journal of Advances in Modeling Earth Systems*, *3*(4). doi: 10.1029/2011MS000074

Golovin, A. M. (1963). The Solution of the Coagulation Equation for Raindrops. Taking Condensation into Account. *Izv. Akad. Nauk. SSSR. Ser. Geofiz.*, *5*, 783–791.

Intergovernmental Panel on Climate Change. (2014). *Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assess-*

ment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press. (doi: 10.1017/CBO9781107415324)

Kessler, E. (1969). On the Distribution and Continuity of Water Substance in Atmospheric Circulations. In E. Kessler (Ed.), *On the Distribution and Continuity of Water Substance in Atmospheric Circulations* (pp. 1–84). Boston, MA: American Meteorological Society. (doi: 10.1007/978-1-935704-36-2_1)

Khain, A. P., Beheng, K. D., Heymsfield, A., Korolev, A., Krichak, S. O., Levin, Z., ... Yano, J.-I. (2015). Representation of microphysical processes in cloud-resolving models: Spectral (bin) microphysics versus bulk parameterization. *Reviews of Geophysics*, *53*(2), 247–322. doi: 10.1002/2014RG000468

Lebo, Z. J., & Seinfeld, J. H. (2011). A continuous spectral aerosol-droplet microphysics model. *Atmospheric Chemistry and Physics*, *11*(23), 12297–12316. doi: 10.5194/acp-11-12297-2011

Milbrandt, J. A., & Yau, M. K. (2005). A Multimoment Bulk Microphysics Parameterization. Part I: Analysis of the Role of the Spectral Shape Parameter. *Journal of the Atmospheric Sciences*, *62*(9), 3051–3064. doi: 10.1175/JAS3534.1

Morrison, H., & Grabowski, W. W. (2007). Comparison of Bulk and Bin Warm-Rain Microphysics Models Using a Kinematic Framework. *Journal of the Atmospheric Sciences*, *64*(8), 2839–2861. doi: 10.1175/JAS3980

Morrison, H., & Grabowski, W. W. (2008). Modeling Supersaturation and Subgrid-Scale Mixing with Two-Moment Bulk Warm Microphysics. *Journal of Atmospheric Sciences*, *65*(3), 792–812. doi: 10.1175/2007JAS2374.1

Morrison, H., Lier-Walqui, M. v., Fridlind, A. M., Grabowski, W. W., Harrington, J. Y., Hoose, C., ... Xue, L. (2020). Confronting the Challenge of Modeling Cloud and Precipitation Microphysics. *Journal of Advances in Modeling Earth Systems*, *12*(8). doi: 10.1029/2019MS001689

Morrison, H., & Milbrandt, J. A. (2015). Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part I: Scheme Description and Idealized Tests. *Journal of the Atmospheric Sciences*, *72*(1), 287–311. doi: 10.1175/JAS-D-14-0065.1

Morrison, H., van Lier-Walqui, M., Kumjian, M. R., & Prat, O. P. (2019). A Bayesian Approach for Statistical–Physical Bulk Parameterization of Rain Microphysics. Part I: Scheme Description. *Journal of the Atmospheric Sciences*, *77*(3), 1019–1041. doi: 10.1175/JAS-D-19-0070.1

Rackauckas, C., & Nie, Q. (2017). DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia. *Journal of Open Research Software*, *5*(1), 15. doi: 10.5334/jors.151

Randall, D., Khairoutdinov, M., Arakawa, A., & Grabowski, W. (2003). Breaking the Cloud Parameterization Deadlock. *Bulletin of the American Meteorological Society*, *84*(11), 1547–1564. doi: 10.1175/BAMS-84-11-1547

Riechelmann, T., Noh, Y., & Raasch, S. (2012). A new method for large-eddy simulations of clouds with Lagrangian droplets including the effects of turbulent collision. *New Journal of Physics*, *14*(6). doi: 10.1088/1367-2630/14/6/065008

Rodríguez Genó, C. F., & Alfonso, L. (2022). Parameterization of the collision–coalescence process using series of basis functions: COLNETv1.0.0 model development using a machine learning approach. *Geoscientific Model Development*, *15*(2), 493–507. doi: 10.5194/gmd-15-493-2022

Seifert, A., & Beheng, K. D. (2006). A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description. *Meteorology and Atmospheric Physics*, *92*(1-2), 45–66. doi: 10.1007/s00703-005-0112-4

Shima, S., Kusano, K., Kawano, A., Sugiyama, T., & Kawahara, S. (2009). The super-droplet method for the numerical simulation of clouds and precipitation: a particle-based and probabilistic microphysics model coupled with a non-hydrostatic model. *Quarterly Journal of the Royal Meteorological Society*,

135(642), 1307–1320. doi: 10.1002/qj.441

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettelmen, A., Golaz, J.-C., Bodas-Salcedo, A., ... Haynes, J. (2010). Dreary state of precipitation in global models. *Journal of Geophysical Research: Atmospheres*, *115*(24). doi: 10.1029/2010JD014532

Tzivion (Tzitzvashvili), S., Feingold, G., & Levin, Z. (1987). An Efficient Numerical Solution to the Stochastic Collection Equation. *Journal of Atmospheric Sciences*, *44*(21), 3139–3149. doi: 10.1175/1520-0469(1987)044⟨3139:AENSTT⟩2.0.CO;2

Wu, Z. (1995). Compactly supported positive definite radial functions. *Advances in Computational Mathematics*, *4*(1). doi: 10.1007/BF03177517

Young, K. C. (1974). A Numerical Simulation of Wintertime, Orographic Precipitation: Part I. Description of Model Microphysics and Numerical Techniques. *Journal of the Atmospheric Sciences*, *31*(7), 1735–1748. doi: 10.1175/1520-0469(1974)031⟨1735:ANSOWO⟩2.0.CO;2

Zhang, X., Song, K. Z., Lu, M. W., & Liu, X. (2000). Meshless methods based on collocation with radial basis functions. *Computational Mechanics*, *26*(4), 333–343. doi: 10.1007/s004660000181