Response to Reviewers of Manuscript **acp-2015-144**

First, I would like to thanks the anonymous referees for their comments that will improve the quality of the paper. Our revised version will include several of their suggestions.

Response to Anonymous Referee # 1

The curse of dimensionality (Bellman, R.; 1961) is a problem of both the data and the algorithm being applied. Solutions to this problem involve changing the algorithm or processing the data into a lower dimensional form. This can be done in many problems since high dimensional data sets can be reduced to lower-dimensional without significant information loss. For example, in Bayesian statistics where posterior distributions have multiple dimensions, this problem was overcome with the implementation of the Markov Chain Monte methods. **This will be discussed shortly in connection with our concrete problem, as suggested by the referee.**

Reference:

Bellman, R. (1961), Adaptive Control Processes: A Guided Tour, Princeton University Press.

Response to Anonymous Referee # 2

All the referee's observations regarding the Stochastic Simulation Algorithm (SSA) of Gillespie (1975a) **will be incorporated in the revised version of the paper.**

In our paper we did just want to remark the fact that in Gillespie (1972), the evolution equation for the probability of finding a given number of *m*-drops of a particular size at time *t*, does not allow to solve for $P(n,m;t)$. This is due to the fact that the right hand side contains a set of unknown conditional probabilities (Gillespie, 1972). Then, the simplest way to proceed in order to close the system was to ignore the correlations.

1. **Comparison between the Gillespie's SSA and the numerical algorithm.**

As we know, in the Gillespie's SSA, the ensemble mean for the number of droplets of each droplet mass is calculated from the expression (Gillespie, 1975a):

$$N(m;t) = \frac{1}{N_r}\sum_{i=1}^{N_r} N^i(m;t) \qquad (1)$$

where N_r is the number of realizations of the stochastic algorithm, $N^i(m;t)$ is the number of droplets of mass *m* in the *i*-realization at time *t*, and $N(m;t)$ is the ensemble mean. From expression (1), it is clear that in order to obtain the correct expected values $(N(m;t))$ at the large end of the droplet size distribution; we will need a huge number of realizations of the stochastic algorithm (> 10^9).

To further investigate this question, the evolution of a cloud system with an initial monodisperse droplet size distribution of N_0=30 droplets of 14 *μm* in radius (droplet mass 1.1494×10^{-8}g) at t_0, and a volume of 1cm^3 was calculated with both the numerical algorithm and the SSA of Gillespie.

The results obtained by the two methods, were then compared with the analytical solution of the master equation (Eq.(13) our paper) obtained by Tanaka and Nakazawa (1993) for the same conditions.

The averages calculated from the Gillespie's method for N_r=10^3 realizations, and the analytical solution at t=1200 are displayed in Fig. 1. As can be observed both the Monte Carlo averages and the analytical solution are closely coincident for the small end of the droplet size distribution. However, due to the small number of realizations, the SSA fails to reproduce the distribution for the expected values at the large end (See Table 1).

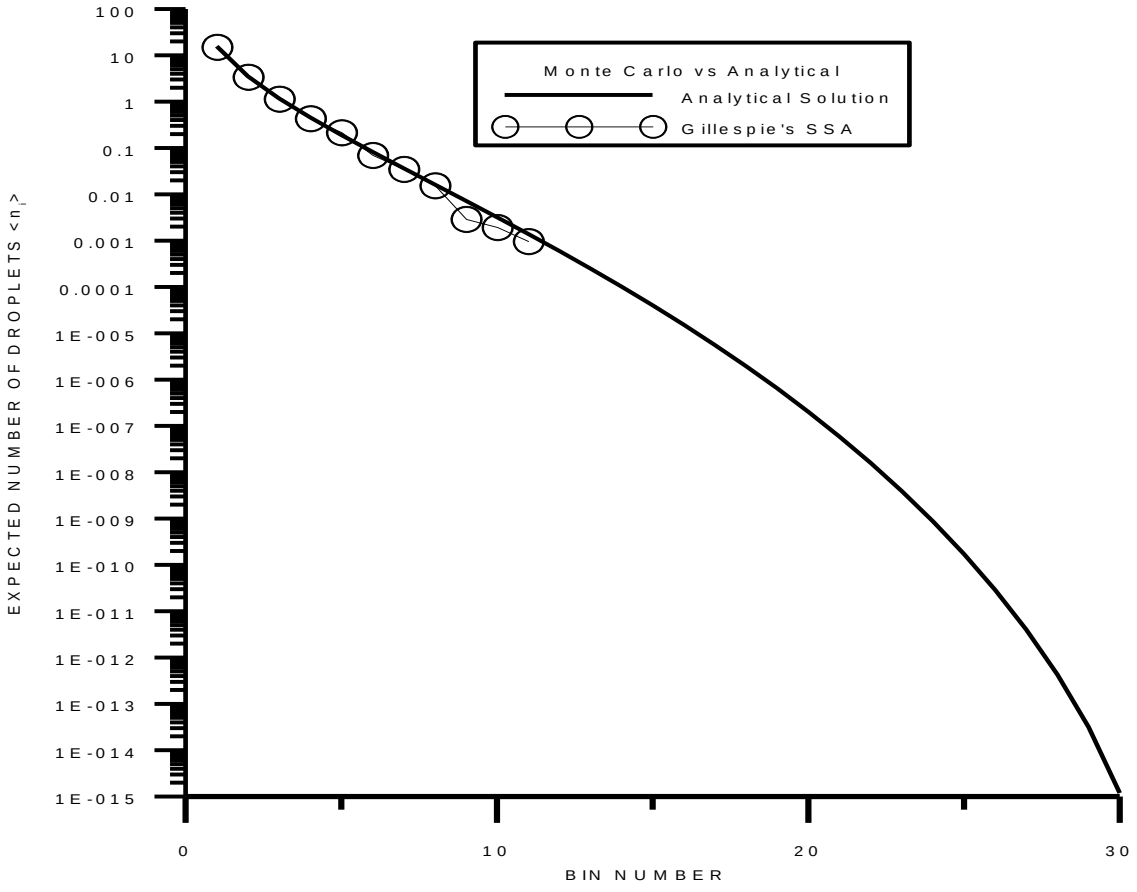

FIG. 1. For the sum kernel, size distribution obtained from the analytical solution of the master equation (line) and the SSA of Gillespie for 10^3 realizations (circles) at *t*=1200 sec. Calculations were performed with the initial condition $P(30,0,0,0,...,0;0)=1$ and the sum kernel $K(i,j)=B(x_i+x_j)$, with *B*=8.82×10^2 cm^3 sec^{-1}.

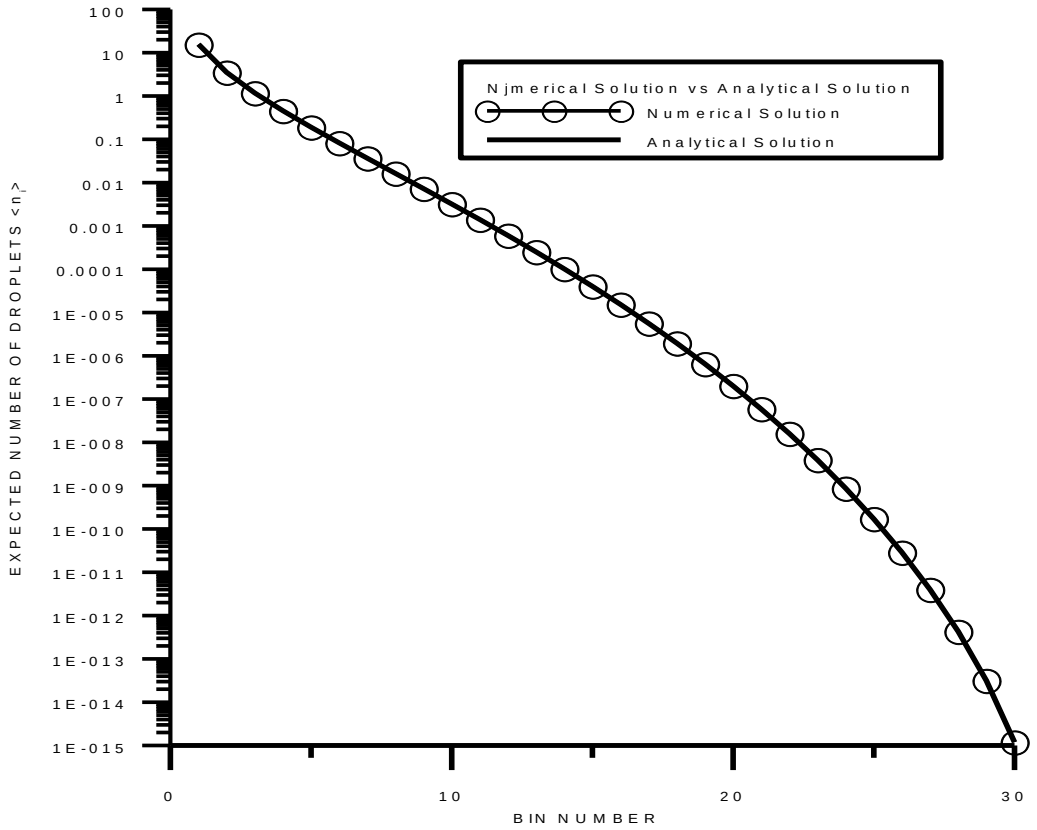

FIG. 1. For the sum kernel, size distribution obtained from the analytical solution of the master equation (line) and the numerical algorithm proposed in this paper (circles) at *t*=1200 sec. Calculations were performed with the initial condition $P(30,0,0,0,...,0;0)=1$ **and the sum kernel** $K(i,j)=B(x_i+x_j)$ **, with *B*=8.82×10^2 cm^3 sec^{-1}.**

For a more detailed analysis, the expected number of particles for each droplet size calculated from the analytical solution, the numerical algorithm and the SSA of Gillespie (for 1000 and 10,000 realizations) are displayed in Table 1. As can be checked in the table, the size distributions are almost identical for the small end, but they differ substantially at the large end since the SSA produces no particles larger than $12v_0$ and $16v_0$ for 1000 and 10,000 realizations respectively (v_0 =1.1494×10^{-8}g, mass of a 14 *μm* droplet).

For 1000 realizations, the Monte-Carlo averages differ from the analytical solution for bin numbers larger than 8. For 10,000 realizations we have the same situation for bin numbers larger than 13.

Table 1. Expected values for each droplet mass obtained at *t*=1200 sec. for the analytical solution, the numerical algorithm proposed in this work, and the Gillespie's SSA (for 1000 and 10,000 realizations). Calculations were performed with the initial condition $P(30,0,0,0,...,0;0)=1$ and the sum kernel $K(i,j)=B(x_i+x_j)$, with *B*=8.82×10^2 cm^3 sec^{-1}.

Expected values for each droplet size: <n_i>

t=1200 sec.

Bin Number	Analytical Solution	Numerical algorithm	SSA (N_r= 1000)	SSA (N_r= 10,000)
1.000	1.5633E+01	1.5622E+01	1.5612E+01	1.5619E+01
2.000	3.5302E+00	3.5303E+00	3.5250E+00	3.5425E+00
3.000	1.1754E+00	1.1762E+00	1.1870E+00	1.1712E+00
4.000	4.5543E-01	4.5609E-01	4.4800E-01	4.5050E-01
5.000	1.9017E-01	1.9057E-01	2.2300E-01	1.9600E-01
6.000	8.2592E-02	8.2824E-02	7.2000E-02	8.2000E-02
7.000	3.6583E-02	3.6709E-02	3.6000E-02	3.6800E-02
8.000	1.6320E-02	1.6387E-02	1.6000E-02	1.6100E-02
9.000	7.2696E-03	7.3034E-03	3.0000E-03	6.5000E-03
10.000	3.2117E-03	3.2284E-03	2.0000E-03	3.5000E-03
11.000	1.3997E-03	1.4077E-03	1.0000E-03	1.2000E-03
12.000	5.9891E-04	6.0263E-04	0.0000E+00	4.0000E-04
13.000	2.5049E-04	2.5216E-04	0.0000E+00	4.0000E-04
14.000	1.0197E-04	1.0269E-04	0.0000E+00	3.0000E-04
15.000	4.0229E-05	4.0529E-05	0.0000E+00	0.0000E+00
16.000	1.5312E-05	1.5431E-05	0.0000E+00	1.0000E-04
17.000	5.5954E-06	5.6404E-06	0.0000E+00	0.0000E+00
18.000	1.9526E-06	1.9687E-06	0.0000E+00	0.0000E+00
19.000	6.4672E-07	6.5217E-07	0.0000E+00	0.0000E+00
20.000	2.0189E-07	2.0361E-07	0.0000E+00	0.0000E+00
21.000	5.8917E-08	5.9419E-08	0.0000E+00	0.0000E+00
22.000	1.5913E-08	1.6048E-08	0.0000E+00	0.0000E+00
23.000	3.9295E-09	3.9622E-09	0.0000E+00	0.0000E+00
24.000	8.7349E-10	8.6634E-10	0.0000E+00	0.0000E+00
25.000	1.7127E-10	1.7176E-10	0.0000E+00	0.0000E+00
26.000	2.8809E-11	2.8765E-11	0.0000E+00	0.0000E+00
27.000	3.9922E-12	3.9906E-12	0.0000E+00	0.0000E+00
28.000	4.2746E-13	4.2803E-13	0.0000E+00	0.0000E+00
29.000	3.1450E-14	3.1525E-14	0.0000E+00	0.0000E+00
30.000	1.1930E-15	1.1962E-15	0.0000E+00	0.0000E+00

As expected, for 1000 and 10000 realizations, no states with droplets 30 times larger than monomer-sized ones were realized. A minimum number of N_r realizations must be performed in order to obtain expected values larger or equal to 10^{-Nr}.

At the same time, the numerical algorithm performed very well, with expected values that are very close to the analytical solution.

We can conclude that our method will be suitable if we need to accurately calculate the large end of the droplet spectrum for small systems, with <50 monomer droplets in the initial state. As for that case the SSA requires a large number of realizations, it will be computationally very expensive. Then, our method will be a good alternative, as it provides the desired accuracy to detect the possible small differences between different numerical approaches. It can also work as a benchmark for different Monte Carlo methods for the collision coalescence process.

A comparison of results and efficiency with Gillespie's method will be included in the revised version of the paper.

2. On Seeβelberg's variation of Gillespie's SSA.

In its original version, the Gillespie's Stochastic Simulation Algorithm (Gillespie, 1975a) was formulated as a particle accounting algorithm, as it keeps track of every individual particle in the system. Unfortunately, this type of algorithm can be computationally very demanding, as a system of N particles requires storage of $N(N+1)/2$ transition probabilities. Then, although it yields exact realizations of the coalescence process, the requirements of computational storage strongly limit its applicability.

The algorithm of Seeβelberg's et al. (1996) found a way to overcome this problem by dividing the drop population into classes and storing only the transitional probabilities between classes. The problem with that approach was the method of removal of the coalescing drops from "colliding classes", a process that introduces an error.

The numerical difficulties of Gillespie (1975a) were significantly overcome in the modified version proposed by Laurenzi et al. (2002). Within their approach, they define collisions between species (that are hydrometeors with the same attributes: mass and composition).By using this framework, there is only the need to store the probabilities of "aggregation reaction" between species, with a considerable reduction of computer memory requirements. After a collision, the reactant (colliding) and product species are updated without any approximation, and not any kind of systematic errors are introduced.

3. On the finite volume approach

In defense of the finite system approach, it might be argued that, in the early stages of precipitation formation, due to small terminal velocities of the droplets, the coalescence process is a fairly localized process. Then, two droplets in widely separated parts of the cloud are not going to be coalescing with each other. This was the approach followed by Bayewitz et al. (1974) (and endorsed in Gillespie, 1975a). In their paper, for comparing the stochastic and kinetic approaches, they partitioned the cloud into many sub-volumes, with no collisions being permitted for two droplets of different sub-volumes.

A more complex model that uses the master equation formalism, and introduces the interactions between the sub-volumes was developed by Merkulovich and Stepanov (1990, 1994). This model is based on a scheme proposed by Nicolis and Prigogine (1984). Within this theory, the whole system is subdivided into subsystems that can be considered spatially homogeneous, and interactions between neighbors occur through particle exchange. That leads to a very complex set of master equations for each sub-volume. Although very complex, it could be a starting point in order to consider the interactions between cells through sedimentation or other processes.

4. On the impact of fluctuations on cloud development

As was discussed, in systems of small populations statistical fluctuations become important and the outcome from kinetic equations may differ from the stochastic means. However, fluctuations will also be very important when the system is near a critical point. In the cloud physics context, a critical point could be related with the formation of a droplet with mass comparable to the mass of the initial system (Alfonso et al., 2013; Lushnikov, 2004). This could be a mechanism responsible for the formation of droplet embryos that trigger precipitation formation.

At this moment, fluctuations will have a significant impact in the development of the system and the total mass predicted by the KCE will start to decrease. This is usually interpreted to mean that a super particle has formed (known as a *gel*) and the system exhibits a sol-gel transition (also called *gelation*).

Because the master equation employs the stochastic approach without any approximations, it can predict the behavior of the collection process at all times. This way, the expected values at the large end of the droplet size distribution can be calculated after the total mass predicted by the KCE starts to decrease (Lushnikov, 2004). By using this method, an accurate comparison between the kinetic and stochastic methods can be performed after the super particle is formed. Then, for this case, it is expected to obtain broader droplet mass distributions by using the stochastic approach.

All this questions will be discussed in the revised version of the paper, as suggested by the referee. A follow-up paper (that is under preparation) will be devoted to a more detailed analysis of all this problems.

5. **Minor comments:**

a) Referee Comment: In Section 2, l. 14 author writes that to solve master equation directly, arrays of the size of 3 x 10^{20} elements would have to be used, where does this number come from?

Answer to the comment:

The number of elements in the array will be 1.34 x 10^{16}, and not 3 x 10^{20} as was written in the paper. **This number will be corrected in the revised version**. The explanation follows:

If we attempt to solve the master equation by brute force by using a typical discretization, then, for an initial monodisperse droplet distribution with *N* monomers, we will need to define an *N*-dimensional array (due to the fact that we have configurations of the form $P(n_1, n_{2,}, ..., n_k, ...; t)$). If we have *n1* =N monomers of size k=1 at t=0, then, the maximum number of particles for size k=2 will be [*N*/2], for size k=3, [*N*/3], and so on, where [*N*/k] denotes the closest smaller integer. (For example, the maximum possible number of droplets of size k=*N* is [*N*/*N*]=1). Of course, due to the mass conservation relation:

$$\sum_{i=1}^{N} x_i n_i = M_T \qquad (2)$$

these maximum values for each droplet size can be attained only if the number of droplets for the rest of the bins are set equal to zero or close to zero. The number of elements of the resulting array then will be the product: $N*[N/2]*[N/3]*[N/(N-1)]*[N/N]$. If evaluate this expression for *N*=50, the value of 1.34x10^{16} will be obtained.

b) Referee Comment: - In the Abstract it should be more clearly stated that the kinetic collection equation and Smoluchowski coagulation equation are different names for the same equation.

Answer to the comment: The corresponding correction will be made in the revised version as suggested.

6. **Technical Comments:**

The corresponding modifications will be made.

7. References

Alfonso, L., Raga, G.B., Baumgardner, D.: The validity of the kinetic collection equation revisited. Part II: Simulations for the hydrodynamic kernel, *Atmos. Chem. Phys*., 10, 6219-6240, 2010.

Alfonso, L., Raga, G. B., Baumgardner, D.: The validity of the kinetic collection equation revisited–Part 3: Sol–gel transition under turbulent conditions. *Atmospheric Chemistry and Physics*, vol. 13, no 2, p. 521- 529, 2013.

Bayewitz, M.H., Yerushalmi, J., Katz, S., and Shinnar, R.: The extent of correlations in a stochastic coalescence process, *J. Atmos. Sci*., 31, 1604-1614, 1974.

Gillespie, D. T.: The stochastic coalescence model for cloud droplet growth, *Journal of the Atmospheric Sciences*, *29*(8), 1496-1510, 1972.

Gillespie, D.T.: An Exact Method for Numerically Simulating the Stochastic Coalescence Process in a Cloud, *J. Atmos. Sci*. 32, 1977-1989, 1975a.

Laurenzi, I. J., Bartels, J. D., & Diamond, S. L. (2002). A general algorithm for exact simulation of multicomponent aggregation processes. *Journal of Computational Physics*, *177*(2), 418-449.

Lushnikov, A. A.: From sol to gel exactly. *Physical review letters*, vol. 93, no 19, p. 198302, 2004.

Merkulovich, V. M., and A. S. Stepanov. "Statistical description of coagulation in finite spatially inhomogeneous systems." *Atmospheric research* 25.5 (1990): 431-444.

Merkulovich, V. M., and A. S. Stepanov. "Statistical description of coagulation in finite spatially inhomogeneous systems (Part 2)." *Atmospheric Research* 26.4 (1991): 311-327.

Nicolis, Grégoire, and M. Malek Mansour. "Onset of spatial correlations in nonequilibrium systems: a master-equation description." *Physical Review A* 29.5 (1984): 2845.

Valioulis, I.A., List, E.J.: A numerical evaluation of the stochastic completeness of the kinetic coagulation equation, J. Atmos. Sci., 41, 2516-2529, 1984.