Responses to Reviewer #2

We thank the reviewer for taking the time to review our paper and for the constructive comments and suggestions. The page and line numbers that we quote for indicating where we changed the manuscript refer to the revised marked-up version.

(2.1) I do not believe the flux-based approach is essentially different from particle-based Lagrangian tracking schemes. For example, a recent study has applied a quantization approach—similar to methods used in machine learning—to particle advection, effectively reducing computational costs and memory usage by quantizing particle positions within a cell (Matsushima et al., 2023). They round the positions of advected particles to the nearest possible locations within a cell, but it is also possible to use stochastic rounding instead of rounding to the nearest. Furthermore, the level of quantization can be optimized for the scientific objectives. Conversely, your method could also reconstruct Lagrangian particle trajectories probabilistically. I suggest investigating the impact of numerical diffusion more clearly by examining the variance in particle distribution, specifically how many particles move significantly faster or slower than the mean flow field. Such evaluations, particularly in two-dimensional test cases, could help clarify how your scheme differs from exact Lagrangian particle tracking.

Thank you for bringing this paper to our attention. We agree with the reviewer that when moving particles, one can either track exact positions or apply different levels of quantization of location. The choice to quantize results in a trade off of errors. The errors are either made in the mean speed, in the variance or some combination of mean speed and variance. The methods presented in our manuscript end up with the all the error in the variance. This error is equivalent to the numerical diffusion of the finite volume scheme. As a consequence of our choice, this means that some particles travel faster than mean speed while some travel slower, i.e., some particles will hop too many grid cells while some persist in a grid cell longer than they should.

We have added the following to the manuscript to address this comment:

- Line 238 (new Subsection 2.9 "Comparison to Lagrangian particle tracking"): "With particles transported by deterministic advection there is no variance in the final position of particles that start in the same initial position. However, when we quantize space and only store which grid cell a particle is in, we can no longer move particles to the exact position where they should be located. That is, we are forced to incur some error. In a classic bias/variance tradeoff, we could achieve zero variance by moving all collocated particles to the same new grid cell, but this would result in an incorrect average position of the particles and a large bias. Alternatively, as we do in this paper, we can move some particles and not others, resulting in the correct mean velocity (zero bias) at the expense of introducing variance in particle position. Consequently, some particles will move faster and some slower than the mean velocity. To quantify the magnitude of this effect, see the example in Section 3.2 and Fig. 7."
- Line 358 (end of Section 3.2): "Finally, to study the effect of spatial quantization where some particles move faster and some slower, causing variance in particle velocity and position (Sec. 2.9), let us consider the following example. If we assume a constant solution at all times (as in Appendix C), then the probability that a particle moves k grid cells is Binom $\big(k; N_\mathrm{t}, p\big)$, where N_t is the number of time steps and p is the probability of moving each step, which will be equal to the Courant number. To investigate this, we refined the grid spacing and time step both by a factor of 10 to be $\Delta x = 0.002$ ($N_x = 500$) and $\Delta t = 0.0008$, which preserves the Courant number of $C = p = 0.4$ of the original simulation, and we took $T = 1$ ($N_t = 1250$) for one revolution. Then, using the binomial distribution, the mean number of grid cells moved in one revolution is $\mu = N_t\, p = 500$, which is an exact approximation (zero bias), while the standard deviation is $\sigma = \sqrt{N_\mathrm{t} p(1-p)} = 17.3$. This corresponds to a physical distance of $x_\sigma = \sigma \Delta x = 0.035$. To understand the limiting behavior, we can use

$N_t = T/\Delta t$ and $p = C = u\Delta t/\Delta x$ to rewrite x_σ as

$$x_\sigma = \Delta x \sqrt{\frac{T}{\Delta t}\frac{u\Delta t}{\Delta x}(1-C)} = \sqrt{Tu(1-C)\Delta x}. \quad (1)$$

Now consider refining the grid ($\Delta x \to 0$) and time step ($\Delta t \to 0$) while keeping constant the Courant number C, the final time T and the velocity u. In this limit, we can see that $x_\sigma \to 0$, so that the numerical diffusion of particles caused by the stochastic method vanishes.

Figure 7 shows the numerical result of the diffusion after one revolution for the particles originating in grid cell 250 (in blue), with the analytical binomial model shown in red. During sampling, some particles will travel faster and some will travel slower, resulting in the binomial distribution of particles around the mean position."

(2.2) The impact of resampling within your scheme requires further improvements. There are concerns that over time, the high-dimensional information held by particles may degenerate into overly similar states due to repeated resampling. Please consider adding a test case where 2–3 tracers are internally mixed to assess such effects on attribute-space dynamics. A simple yet effective analysis could compare the differences between the initial and final distributions in attribute space, providing insights into the extent of particle degeneracy. In addition, please provide details on variations of areas in each cell and atmospheric density and the results for the vertical cross sections of the tracer mixing ratio for the three-dimensional case to clarify the range of applicability of your scheme.

Thank you for these good suggestions. We added a new test case (Figure 15) to compare the initial and final distributions in attribute space (we used the 1D particle size as the attribute). We provided additional details on the cell variations and added a vertical profile to Figure 14 (used to be Figure 12). Specifically:

- Line 218: "A potential concern is that the repeated resampling due to varying computational volumes, grid cell volumes, and air densities may cause the high-dimensional infomation carried by particles (see Section 2.4) to degenerate into overly similar representations. For example, if the particles carry a diameter sampled from a size distribution, the repeated resampling may cause the particles to converge to a single diameter. In Sec. 3.4 we investigate this numerically and see that it is not a significant issue in practice."
- Page 24: Added new Figure 15.
- Line 429: "In Sec. 2.7, we discussed sampling complexities due to different computational volumes, grid cell volumes and air densities. When these quantities substantially differ in adjacent grid cells, it could lead to undersampling of rare particle types. In our three-dimensional example, the largest ratio in density was 1.29, and the largest grid cell volume ratio was 1.96. For most of the grid cells, these ratios were closer to 1, indicated by domain average ratios of 1.01 and 1.11, respectively, at $t = 12$ h. To investigate whether undersampling occurred in practice, we ran the same scenario but sampled the particle diameter (a 1D attribute carried by particles) from a log-normal size distribution so that both rare large and small particles existed while most computational particles resided in the center of the size distribution. We then compared the final size distributions with the initial size distributions to determine to what extent the rare large and small particles were systematically lost due to undersampling.

  Figure 15(a) shows the locations for the initial and final size distribution plots. The locations of the initial and final points were chosen so that the final point is downwind of the initial point. All grid cells were initialized with 100 computational particles drawn from a single log-normal mode, all with a constant geometric mean diameter and geometric standard deviation where only the magnitude of the distribution was adjusted. Figure 15(b) shows the normalized mean particle size distribution at the initial time and the final time at two single grid cells. Each distribution was averaged over five ensemble runs.

As we see from Fig. 15(b), the size distribution at the final time was similar to that at the initial time, with some stochastic noise. To reduce the stochastic noise, Fig. 15(c) shows the normalized mean particle size distribution at the initial time and the final time for two 15 × 15 grid cell patches surrounding the points chosen for Fig. 15(b). Here the normalized size distributions were nearly identical, indicating that the size distribution information was not lost in the sampling procedure."

- Added a new subfigure to Figure 14 that shows a vertical profile of the ensemble mean mixing ratio for 10, 100, 1000 computational particles.
- Line 424: "Fig. 14(b)–(d) show different transects through the three-dimensional space and time. The star in Fig. 14(a) marks the location of the vertical mixing ratio profile (log-scaled) in Fig. 14(b) and the time series shown in Fig. 14(d). The red line denotes the transect shown in Fig. 14(c). The finite volume solution is compared to the ensemble mean of 10, 100 and 1000 computational particles with error bars denoting the 95% confidence interval. As the number of particles increased, the variance decreased and the solution converged to the finite volume solution."

(2.3) The dependency of the number of computational particles on the results is well presented in your scheme. However, convergence with respect to space and time resolution is not addressed. Additionally, it would be beneficial to specify the Courant number used in your test cases and discuss the limits of the Courant number that your scheme can accommodate.

Thank you for raising this issue. The error in the stochastic solutions consists of two portions: (1) the stochastic error between the stochastic solution and the finite volume solution, which decreases with increasing number of computational particles, and (2) the error between the finite volume solution and the true solution, which decreases as the time step and grid spacing become smaller. The focus of this paper is on analyzing the convergence of the stochastic method to the finite volume method, while assuming that the standard finite volume methods presented here in WRF have the correct properties to converge as Δx and Δt are refined. We made changes to clarify this in the paper and added more details about the simulation:

- Line 248: "The total error of a stochastic transport scheme can be bounded by two error terms that can be evaluated independently: (1) the stochastic error between the the stochastic solution and the finite volume solution, and (2) the deterministic error due to the space-time discretization of the finite volume scheme. That is, for a stochastic solution n^{stoc}, a finite volume solution n^{FV}, and an exact true solution n^{true}, we can write:

$$\underbrace{\|n^{\text{stoc}} - n^{\text{true}}\|}_{\text{total error}} \leq \underbrace{\|n^{\text{stoc}} - n^{\text{FV}}\|}_{\text{stochastic error}} + \underbrace{\|n^{\text{FV}} - n^{\text{true}}\|}_{\text{deterministic error}} .'' \tag{2}$$

- Line 253: "We do not consider the refinement of $\Delta x \to 0$ or $\Delta t \to 0$ as it is well understood how the finite volume methods converge to the true solution (deterministic error goes to zero) in these limits."
- Line 275: "resulting in a Courant number of 0.4."
- Line 379: "The maximum Courant number was 0.5."
- Line 395: "The model time step was set to $\Delta t = 20$ s, ensuring that the sum of particle cell transfer probabilities did not exceed 1."

(2.4) L152: The extension to multiple dimensions could be more clearly improved. In your method, unlike methods such as the corner transport upstream method or the conservative semi-Lagrangian method (Lin et al., 1996), it appears that you assume particles do not move in diagonal directions. Is my understanding correct? If so, would it be more reasonable to employ a flux that better aligns with the actual transport of particles?

It is true that we do not consider particles moving in the diagonal directions. And it is correct that other finite volume schemes would result in an improved solution. However, the aim of this work was to develop methods that specifically mimic the advection schemes used in the WRF model. The reason for this is that it allows us to isolate the impact of aerosol representation when comparing WRF-PartMC to WRF-Chem.

In future work, the framework for simulating advection may be applied to other finite volume approaches. We made the following changes to clarify our choice:

- Line 55: "While we only present the development of stochastic advection schemes based on the finite volume methods in WRF, the methodology described here is applicable to any finite volume scheme or transport scheme such as Corner-Transport Upwind (Colella, 1990; LeVeque, 2002) or Flux-Form Semi-Lagrangian (Lin and Rood, 1996, 1997) that can be found in other host models."
- Line 128: "Although we are presenting the method using the particular discretizations above, it is straightforward to derive stochastic versions of other spatial and temporal discretizations in the same way."

(2.5) L300: To check whether the schemes can be applied to simulate the mixing process, it would be better to adopt a swirling shear flow and a steeper initial condition like those of a cosine bell-type distribution and verify how well the tracer filament structures are preserved (see p.264 in Durran 2010).

As described in the response to comment (2.3), our intention is to construct a stochastic method that mimics the existing finite volume methods in WRF. We are not aiming to improve the finite volume methods, but rather to understand the impact of the stochastic representation of aerosols on the model results. While test cases such as a swirling shear flow are interesting, our stochastic method will treat them in the same way as the finite volume method. Such examples are well-understood from the perspective of finite volume methods, and we thus feel that such examples are not necessary for this paper.

References

Colella, P.: Multidimensional upwind methods for hyperbolic conservation laws, J. Comput. Phys., 87, 171–200, 1990.

LeVeque, R. J.: Multidimensional Scalar Equations, p. 447–468, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2002.

Lin, S.-J. and Rood, R. B.: Multidimensional flux-form semi-Lagrangian transport schemes, Mon. Weather Rev., 124, 2046–2070, 1996.

Lin, S.-J. and Rood, R. B.: An explicit flux-form semi-Lagrangian shallow-water model on the sphere, Q. J. R. Meteorolog. Soc., 123, 2477–2498, 1997.