

## Response to Referee2

We thank Referee 2 for the careful and constructive review. We especially appreciate the reviewer's assessment that documenting tracking algorithms in SCALE-SDM is useful for the SDM community, and we agree that the original manuscript did not sufficiently separate the methodological contribution from the case-study interpretation. The review also correctly identified three major weaknesses of the original manuscript: the scientific novelty was overstated, the practical 3D limitation of full-population backward tracking was not addressed strongly enough, and the computational evaluation was not controlled enough to support the algorithmic framing.

In response to these comments, and to related comments from Referee 1, we made a substantial revision of both the code workflow and the manuscript. The revised manuscript differs substantially from the discussion version. The title and abstract now emphasise implementation, computational evaluation, and Two-Pass Hybrid Tracking (TPHT). The abstract now states that the framework includes forward tracking, backward tracking, a shared target-selection and sampling subsystem, coalescence-event diagnostics, and TPHT (lines 13-21). It also states that controlled computational tests are used to characterise runtime, memory, storage, output-frequency sensitivity, and sampling behaviour, and that a three-dimensional DYCOMS-II RF02 case is used to demonstrate TPHT for target-restricted backward reconstruction (lines 20-29).

We also changed the manuscript framing. The revised introduction now explicitly states that the lucky-droplet concept, stochastic growth toward raindrop formation, and entrainment/mixing effects on droplet histories have already been studied previously, and that these processes are used in the present paper primarily to define targets and demonstrate the diagnostic capability of the tracking framework rather than to claim a new general mechanism of warm-rain formation (lines 75-80). The revised introduction further distinguishes full-lineage backward tracking from target-restricted

reconstruction and introduces TPHT as a target-restricted alternative for practical 3D applications (lines 81-89).

At the implementation level, we added and documented the TPHT workflow. TPHT uses a forward-discovery pass to record interest-defined target IDs, an offline merge/deduplication step to form a compact target-set handoff, and a backward-reconstruction pass that reconstructs only the deduplicated target set from the same physical initial state (Sect. 2.5, lines 347-386; Algorithm 3, lines 387-395). The revised manuscript also adds a practical comparison of backward tracking, forward tracking, and TPHT, including the explicit statement that simply increasing the forward-tracked sample is not equivalent to TPHT (Sect. 2.6, lines 403-453).

Finally, the revised case-study section now begins by stating that its purpose is to demonstrate TPHT in a stratocumulus simulation rather than to establish a new general mechanism of warm-rain formation (lines 466-467). The discussion and conclusions repeat this moderated framing: the stratocumulus case is described as a methodological demonstration, and coalescence-linked records are described as diagnostic event-log links rather than a complete causal process budget (lines 864-868).

**General comments:**

Overall, I believe this work makes a useful effort to document tracking algorithms that will be of interest to the SDM community. However, I have several major concerns regarding the way the manuscript frames its scientific novelty. In particular, some of the scientific implications emphasized by the authors are already broadly known in the literature, but these prior studies are not adequately acknowledged. In addition, the currently proposed backward algorithm does not appear to be practically applicable to the most important simulation setting, namely, realistic 3D simulations, which are required to address many of the scientific questions the authors themselves propose.

This limitation makes it difficult for me to view the present framework as a complete and practically ready method for broad community use. At this stage, it reads more as a promising conceptual capability that still requires further methodological development before it can be deployed for the more ambitious applications suggested by the authors.

If this latter interpretation is closer to the authors' intent, then I encourage them to state this purpose more clearly and to frame the paper accordingly, as an important step toward future development rather than as a method that is already practically ready for the full range of applications discussed. I therefore recommend that the authors address the concerns raised below and clarify the intended scope of the paper.

**Reply:** We thank the reviewer for this clear assessment. We agree that the original manuscript over-emphasised scientific implications and did not define the practical scope of the framework carefully enough. We have therefore revised the manuscript in three main ways.

First, we changed the paper's scope and framing. The revised manuscript is now framed as a development and technical paper documenting a Lagrangian particle-tracking framework in SCALE-SDM, rather than as a paper claiming new general cloud-physics mechanisms. The revised abstract states the framework components and the computational evaluation directly (lines 13-29), and the introduction now states that the case studies are methodological demonstrations rather than claims of new general cloud-physics mechanisms (lines 96-103).

Second, we added the missing prior literature and moderated the scientific claims. The revised introduction now cites the lucky-droplet and stochastic-growth literature, including Telford (1955), Dziekan and Pawlowska (2017), Hoffmann et al. (2017) and Li et al. (2022), and states that these processes are used primarily to define targets and

demonstrate diagnostic capability rather than to claim a new general mechanism of warm-rain formation (lines 75-80). We also added Lim and Hoffmann (2023, 2024) as examples of event- or history-based Lagrangian diagnostics for entrainment and mixing (lines 78 and 349-351).

Third, we addressed the practical 3D limitation directly. We agree that brute-force full-population backward tracking can be storage-limited in realistic 3D simulations. The revised manuscript now explicitly distinguishes full-lineage backward tracking from target-restricted reconstruction (lines 90-95 and 610-612). To make the 3D workflow practical, we added TPHT: a forward-discovery pass identifies interest-defined targets, and a backward-reconstruction pass reconstructs only the deduplicated target set (lines 440-4542). The revised 3D DYCOMS-II RF02 demonstration uses this TPHT workflow in a full  $2 \text{ km} \times 2 \text{ km} \times 1.5 \text{ km}$  domain (lines 506-511).

### **Major comments:**

1. **Comment:** For the purpose of documenting particle-tracking algorithms within an SDM-based model, I am generally satisfied with the presentation of the forward algorithm. However, I do not see strong algorithmic novelty in this part, as similar forward-tagging approaches have already been used more widely in the superdroplet community than the papers that the author cited, and are relatively straightforward to implement in existing models that employ the superdroplet method. In my view, the manuscript should therefore place greater emphasis on the backward algorithm, which appears to be the more distinctive methodological contribution.

Moreover, if the data size of forward sampling is about 76 MB for 400K particles, wouldn't it be more straightforward to simply employ more particles to track using the forward method?

**Reply:** We agree. The original manuscript over-emphasised the novelty of the forward algorithm. In the revised manuscript, forward tracking is no longer presented as the main algorithmic novelty. It is now described as selected-target output and as a necessary implementation/workflow component within SCALE-SDM. The more distinctive methodological elements are now the predecessor-link backward reconstruction and the TPHT workflow.

The forward-tracking section now explicitly states that the forward algorithm should be interpreted primarily as an implementation and workflow component rather than as a claim of general algorithmic novelty (lines 254-262). We also added a practical workflow comparison that explains when forward tracking is useful and what its limitations are (lines 427-440).

We also agree that increasing the number of forward-tracked particles can be useful when the target class is known before or during the simulation. However, simply increasing the forward-tracked sample is not equivalent to TPHT. Forward tracking writes full selected-output histories for particles selected before or during the run. Increasing the sampling fraction reduces the chance of missing rare particles, but the output volume still scales with the number of retained particles, output times, and variables written. In the limit, forward tracking approaches full-population selected-output. TPHT instead separates target discovery from trajectory reconstruction: the discovery pass records compact target IDs only when interest conditions are met, and the reconstruction pass writes selected-output histories only for the deduplicated target set. This distinction is now stated explicitly in the revised manuscript (lines 436-446). Forward tracking remains useful and is retained in the framework, but we no longer claim that it is strongly novel as a generic Lagrangian-tagging algorithm. Its contribution here is its documented integration into SCALE-SDM, its shared sampling subsystem, its

coalescence-event diagnostics, and its role in TPHT discovery.

2. **Comment:** A fair comparison of computational cost for the two algorithms should be conducted under appropriately similar conditions (Sec.3.2.3). At present, the comparison is not fully convincing because the forward and backward algorithms are demonstrated under substantially different configurations, while the manuscript is framed as an algorithmic contribution.

More importantly, the backward algorithm, which seems to be the more novel component of the paper, is not currently demonstrated to be practical for 3D simulations because of the storage bottleneck associated with storing lineage pointers for all superdroplets. This is a substantial limitation. The scientific applications highlighted by the authors, including lucky-droplet growth, eddy hopping, and inhomogeneous mixing, fundamentally require realistic 3D simulations. In this sense, the current inability to apply the backward method efficiently in 3D limits the practical impact and generality of the proposed framework.

While the authors briefly note this in Sec. 4.2, and I appreciate the authors' effort to construct a method with efficient lookup and strong post-processing capability, I believe the manuscript would be significantly strengthened by a more explicit discussion of this 3D limitation and by a more controlled performance comparison under similar simulation settings. In a methodological paper, practical scalability is not a secondary detail but part of the central novelty claim.

**Reply:** We agree. This comment motivated one of the largest revisions. The revised manuscript no longer relies on the original comparison between forward and backward tracking under substantially different configurations. Instead, we added a controlled three-dimensional computational benchmark using comparable no-tracking, coalescence-log-only, sampled forward-tracking, and sampled

backward-tracking configurations.

In the revised benchmark, all cases use the DYCOMS-II RF02 configuration in a full three-dimensional domain of  $2 \text{ km} \times 2 \text{ km} \times 1.5 \text{ km}$ , with grid spacing of  $50 \text{ m} \times 50 \text{ m} \times 5 \text{ m}$  and an integration time of 1800 s (Appendix A, lines 875-883). The sampled forward and backward tracking cases both use 5% stratified target selection and selected-super-droplet output every 60 s (lines 880-883). This provides a controlled basis for comparing the practical overhead of sampled tracking diagnostics.

The main computational results are now summarised in Sect. 3.2.3. In this benchmark, 5% sampled forward and backward tracking showed no large wall-clock or peak-memory penalty relative to the no-tracking baseline: wall-clock times were 8754-8750 s for sampled tracking and 8982 s for no tracking, while peak memory per rank remained around 649-653 MiB compared with 650 MiB for the baseline (lines 544-550). The main additional burden was output volume, not identifier bookkeeping (lines 596-601 and 885-899). Appendix A further states that values slightly below the no-tracking case should not be interpreted as speed-up, but as indicating that the sampled-tracking overhead is small relative to run-to-run variability in that controlled setup (lines 892-893).

To address scalability, we added Appendix C. It reports SDNC-scaling tests with SDNC values of 10, 20, 40, and 80 per cell and output-interval sensitivity tests with selected-output intervals of 30, 60, and 120 s (lines 983-990). The text explicitly cautions that these results describe finite-range empirical behaviour for the tested machine configuration and should not be interpreted as general asymptotic scaling laws (lines 991-996).

Most importantly, we agree that full-population backward tracking remains storage-limited for realistic 3D simulations. We therefore added TPHT as a target-restricted 3D backward-reconstruction workflow. TPHT does not store arbitrary full-population histories for later unrestricted analysis. Instead, it discovers target IDs using event- or threshold-based conditions and reconstructs only the deduplicated target set (lines 348-356). The revised 3D TPHT demonstration generated 27,826,660,866 all-time interest-ID records, deduplicated them to 3,554,256 target identities, and recovered all deduplicated targets in the backward-reconstruction output with no additional targets (lines 560-565). The deduplicated target set represented about 19.3% of estimated valid super-droplet identities, and the manuscript now explains how the estimated full-population backward output was obtained by scaling the measured backward selected-output size by the inverse of the retained-target fraction (lines 577-583).

We also revised the storage interpretation to distinguish temporary working storage from retained analysis storage. The raw interest-ID streams are intermediate files used for offline deduplication and can be deleted after the deduplicated target-set handoff has been generated. In the TPHT demonstration, the raw streams occupied 245 GiB, whereas the deduplicated handoff occupied only 32.0 MiB and the backward selected-super-droplet output occupied 69 GiB. If raw streams are retained, the total TPHT working storage gives an estimated  $1.1\times$  reduction relative to estimated full-population backward output. After raw-ID deletion, the retained output corresponds to an estimated  $5.2\times$  reduction. We now state this distinction explicitly in Sect. 3.2.3 (lines 584-592) and in the summary (lines 849-856).

The revised manuscript is explicit that TPHT is not a complete end-to-end performance optimum and does not eliminate all storage cost. The offline

merge/deduplication wall-clock time was not separately instrumented in the present analysis; therefore, the TPHT storage numbers quantify the model-output and retained-analysis burden, while total end-to-end workflow cost also includes offline parsing, deduplication, repartitioning, and handoff-file writing (lines 599-605). This limitation is now stated directly.

3. **Comment:** Furthermore, if the suggested future work for backward tracing is based on sampling, I would suggest exploring event-based particle backward tracking, as employed by Lim and Hoffmann (2023) and Lim and Hoffmann (2024). They tracked backward when the event of interest happened (entrainment and mixing in their paper). The authors could apply a similar approach for lucky droplets, triggering backward tracking when either a collision or mixing occurs. This might be a more appropriate and efficient solution for future work.

**Reply:** We agree. This suggestion strongly influenced the revised manuscript. We added Lim and Hoffmann (2023, 2024) to the discussion of event- and history-based Lagrangian diagnostics and introduced TPHT as a practical event- or threshold-restricted workflow (lines 3804-806).

TPHT is conceptually related to the event-based strategy suggested by the reviewer, but it is implemented differently. Rather than triggering an online backward trace at the exact event time, TPHT first performs a forward-discovery pass that records compact target IDs when user-defined interest conditions are met, followed by an offline deduplication step and a backward-reconstruction pass for the deduplicated target set. In the present demonstration, the interest conditions are radius-threshold exceedance and binary coalescence participation. This is described in the TPHT methodology (lines 348-356) and in the TPHT experiment setup (lines 658-656).

We also agree that entrainment and mixing events would be valuable target definitions. The current TPHT demonstration does not yet implement a full entrainment/mixing-event detector. We therefore added this explicitly as future work. The revised limitations section states that other questions would require different event definitions, including entrainment, mixing, supersaturation exposure, sedimentation, or residence in cloud-top and cloud-base regions, and that the same target-ID handoff mechanism could then be used to reconstruct histories for particles satisfying these more specialised diagnostic conditions (lines 850-814).

4. **Comment:** When discussing the scientific implications, apart from algorithmic novelty, the authors do not sufficiently cite existing literature on lucky droplet and raindrop growth (e.g., Hoffmann et al., 2017; Dziekan & Pawlowska, 2017; Li et al., 2022), which has already been reported in the community using the superdroplet method. The 15-20  $\mu\text{m}$  threshold and the lucky-droplet pathway are interesting, but they appear more as a demonstration of the framework's capability rather than broadly general new physics. Therefore, the authors need to cite foundational works done with superdroplets and the canonical papers from as early as the 1950s that established this theory, such as Telford (1955).

**Reply:** We agree. The original manuscript did not sufficiently acknowledge the prior lucky-droplet and stochastic-growth literature, and it overstated the scientific interpretation of the 15-20  $\mu\text{m}$  threshold and lucky-droplet pathway. We have added the requested literature and revised the framing.

The introduction now cites Telford (1955), Dziekan and Pawlowska (2017), Hoffmann et al. (2017), Li et al. (2022) and Lim and Hoffmann (2023, 2024). It also states that these processes are used primarily to define targets and

demonstrate the diagnostic capability of the tracking framework, rather than to claim a new general mechanism of warm-rain formation (lines 75-80).

The practical workflow comparison also cites previous studies on lucky-droplet growth, eddy hopping, spectrum broadening, and event-based Lagrangian diagnostics when explaining possible target definitions for backward tracking and TPHT (lines 414-420). The case-study section now begins by stating explicitly that its purpose is to demonstrate TPHT in a stratocumulus simulation, not to establish a new general mechanism of warm-rain formation (lines 466-467).

The TPHT target diagnostics are now framed as target-set characterisation. The manuscript reports the recovered target classes, first recovered  $r \geq 15 \mu\text{m}$  crossing heights, and near-threshold radius behaviour (lines 683-741), but it does not present these as a new general theory of lucky-droplet formation.

5. **Comment:** In Sect. 4.1, the authors state that the algorithm provides “novel insights” into the microphysical pathway. However, the backward case is not used to analyze actual collision or coalescence events.

Rather, it is used to demonstrate condensational growth pathways that lead droplets into the size range favorable for coalescence. This distinction should be made more explicit. In its current form, the backward analysis supports pre-collisional growth diagnostics rather than reconstruction of actual collision histories.

More importantly, I do not find the scientific framing sufficiently justified. The results presented in this section appear to illustrate processes that are already broadly known, rather than providing genuinely new physical insight. For this reason, I encourage the authors to reconsider the use of phrases such as “novel insights” and to moderate the framing of this section accordingly.

**Reply:** We agree. The revised manuscript no longer uses the framing, such as “novel insights”, for the case study. Instead, the results are presented as diagnostic examples showing what TPHT can reconstruct and how target histories can be linked to event-log context and Eulerian fields.

We removed the original Sect. 4.1 “Synthesis of Findings”, which was the section containing the strongest claims about novel microphysical pathways. The revised Discussion no longer synthesises the case study as a set of new cloud-physics findings. Instead, it discusses the practical implications of target-restricted tracking and the limitations of storage, predefined target criteria, sampling interpretation, and coalescence-event diagnostics.

Yes, we agree that the backward analysis supports pre-collisional growth diagnostics rather than reconstruction of actual collision histories. We have also clarified the interpretation of coalescence-event diagnostics. The revised text states that `if_coal` is a binary interval marker and not an event counter, and that target-linked coalescence records should not be treated as a complete causal process budget unless all required event branches, multiplicities, and output variables have been retained and analysed (lines 793-797). The Figure 5 caption similarly states that plotted coalescence records are diagnostic event-log links rather than a complete causal process budget (lines 690-700).

The current TPHT example does include coalescence-event context, but it does not claim to reconstruct a complete collision history. The coalescence-context section compares the Hall gravitational collection kernel with TPHT-linked event records only as radius-pair context for the coalescence-participation criterion. The text explicitly states that the comparison should not be interpreted as validation

of the Hall kernel or as a closed collision-coalescence budget (lines 640-646).

**Minor comments:**

6. **Comment:** The manuscript would benefit from a clearer separation between the “framework description” and the “case-study science results.” I am not certain if these test cases actually add new scientific knowledge beyond what is already known. While the framework itself shows potential, the results the authors presented here do not clearly demonstrate it.

**Reply:** We agree and have restructured the manuscript. The revised manuscript now separates the framework description, workflow comparison, computational evaluation, and case-study demonstration.

The algorithmic and workflow material is now concentrated in Sect. 2: backward tracking, forward tracking, shared target-selection and sampling, TPHT, and the practical workflow comparison (lines 403-454). The case-study section now starts with a clear statement that its purpose is to demonstrate TPHT rather than establish new warm-rain physics (lines 466-467). The computational evaluation and scaling tests are separated into Sect. 3.2.3 and Appendices A-C (lines 438-612 and 875-1003). The target-specific case-study diagnostics are then presented as coalescence-event context, target-set diagnostics, and predecessor-chain examples.

7. **Comment:** Several wording issues and typos currently hinder the readability of the paper. I encourage the authors to carefully proofread the manuscript and revise it accordingly.

**Reply:** We have revised the manuscript extensively to improve readability and

reduce overstatement. In particular, we moderated language such as “novel insights”, “complete lifecycle”, “powerful and versatile”, and similar phrasing from the original version. The revised text now uses more cautious terms such as “selected-target diagnostics”, “methodological demonstration”, “diagnostic context”, “estimated storage reduction”, and “target-restricted reconstruction”.

We also corrected several technical and formatting issues during the broader revision. The new manuscript now includes clearer definitions of identifier variables, `if_coal`, `selected-super-droplet` output, `coalescence-event` output, and TPHT handoff files (lines 162-164, 174-184, and 396-402). The code/data and author-contribution statements were also updated.

8. **Comment:** In Table 1, MP setup indicates `cnd + coal` for BKW `cond`. Please fix this if this is an error.

**Reply:** This issue is no longer applicable because the original experiment table and the `BKW_cnd` experiment framing have been removed in the revised manuscript. The revised manuscript now contains a new workflow-comparison Table 1 and a new TPHT configuration Table 2.

The new Table 1 compares the three workflows — backward tracking, forward tracking, and TPHT — in terms of best use case, main caveat, and practical strength (lines 404-464). The new Table 2 describes the forward-discovery and backward-reconstruction passes in the TPHT experiment. It states that `coalescence-event` logging is enabled in both passes and that event rows are appended only when recorded events occur (lines 531-536).

## Reference

Dziekan, P. and Pawlowska, H.: Stochastic coalescence in Lagrangian cloud microphysics, *Atmospheric Chemistry and Physics*, 17, 13509–13520,

<https://doi.org/10.5194/acp-17-13509-2017>, 2017.

Hoffmann, F., Noh, Y., and Raasch, S.: The Route to Raindrop Formation in a Shallow Cumulus Cloud Simulated by a Lagrangian Cloud Model, *Journal of the Atmospheric Sciences*, 74, 2125–2142, <https://doi.org/10.1175/JAS-D-16-0220.1>, 2017.

Li, X.-Y., Mehlig, B., Svensson, G., Brandenburg, A., and Haugen, N. E. L.: Collision Fluctuations of Lucky Droplets with Superdroplets, *Journal of the Atmospheric Sciences*, 79, 1821–1835, <https://doi.org/10.1175/JAS-D-20-0371.1>, 2022.

Lim, J.-S. and Hoffmann, F.: Between broadening and narrowing: How mixing affects the width of the droplet size distribution, *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037900, 2023.

Lim, J.-S. and Hoffmann, F.: Life cycle evolution of mixing in shallow cumulus clouds, *Journal of Geophysical Research: Atmospheres*, 129, e2023JD040393, <https://doi.org/10.1029/2023JD040393>, 2024.

Telford, J. W.: A new aspect of coalescence theory, *Journal of Atmospheric Sciences*, 12, 436–444, 1955.