Review on the manuscript entitled

A Continuous-Time Ensemble Kalman–Bucy Smoother for Causal Inference and Model Discovery

August 30, 2026

Generality

The authors propose continuous-time and derivative-free schemes, namely ensemble Kalman–Bucy filter (EnKBF) / smoother (EnKBS), in which the conditional (smoothing) distributions are reconstructed from ensemble moments, avoiding tangent-linear and adjoint models. The smoother is claimed to converge to the exact smoothing solution in linear dynamic systems. The proposed approaches have employed covariance localization and inflation to address high-dimensional problems. In this manuscript, the authors illustrated the performance of EnKBF/EnKBS on state estimation for high-dimensional chaotic systems (Lorenz 96, Lorenz 84) and iterative learning for causal relationships and their temporal influence range in a dyadic trigger–feedback model and in a reduced-order model of midlatitude atmospheric circulation.

The topic suits the NPG scope as continuous-time smoothers for nonlinear systems are likely under-explored, and the link between smoothing and backward tracing of causal drivers to extreme events is an appealing idea. However, several methodological points in the current manuscript could be elaborated, and the performance of the proposed approaches relative to existing ensemble smoothers should be explicitly demonstrated. I recommend addressing the following points before further consideration.

Major comments

1. In the abstract, the authors state that integrating future observations reveals the causal mechanism and supports retrospective updates. The Kalman–Bucy smoother in the linear-Gaussian case already provides this benefit. For the dyadic trigger–feedback model and the reduced-order atmospheric model, please clarify i) whether the dynamics are genuinely nonlinear and non-Gaussian in the regimes considered, and what the EnKBS provides that the analytical KS (or a linearized/unscented RTS smoother) does not.

2. The introduction of assimilation time s in pages 5-6 needs a proper definition. As written, s is called a "time interval" although it appears to be a pseudo-time coordinate. Please clarify: does s depend on t? How exactly does s in Eqs. (2) and (3), which currently show no explicit s-dependence? Is $\mathbf{P}$ in these equations the propagated filtering covariance, and if so, is it a function of t, of s, or of both?

3. Are backward dynamics reversible in time as always? Equation (9) uses reverse prediction of $\mathbf{x}(t)$ given $\mathbf{x}(t+1)$. Simply integrating the forward drift $\mathbf{f}$ backward in time is not, in general, the time reversal of a stochastic process: the reverse-time SDE has a modified drift and the forward noise process is not reversible as such [Anderson(1982)]. The authors should state precisely what backward model is used, whether $\mathbf{f}$ is transformed, and how the noise is handled.

Moreover, Eq. (11) discretizes a reversed stochastic (difference) equation. It is known that (i) reversing an approximation introduces additional error, (ii) discretizing a reverse SDE to sufficient fidelity can be computationally very demanding even for moderate state dimensions, and (iii) there exist SDEs with bounded solutions whose discretizations are unbounded. Please discuss the discretization error, the time-step constraints, and the numerical stability of the backward pass, and comment on whether the asymptotic convergence claim survives these approximations.

4. Part of the current parameter-estimation literature relies on online, filter-based (state-augmentation) approaches, while the manuscript motivates offline parameter estimation; i.e., smoothing by the need for retrospective analysis once $\mathbf{y}_{[0,T]}$ is available. The author please explain why offline parameter estimation is preferable for the model-learning task.

5. Inconsistent experimental choices (pages 11–13):

- Table 2 reports the best RMSE for $\sigma^2 = 1.01$ and $r_0 = 4$, yet the time series in Figure 2 of x_1 from EnKBF/EnKBS are shown for $\sigma^2 = 1.005$ and $r_0 = 3$. Please justify this choice or use the best-performing configuration.
- Figure 3: the text states that inflation "does not noticeably change the results in this regime" and is therefore not applied, while elsewhere inflation is presented as needed to improve RMSE. Please reconcile these statements and specify in which regimes inflation is (not) required.
- Section 3.1.2 mentions an ensemble size $m \leq 13$; this bound is unexplained. Please clarify.

6. No observation operator or observation model is specified for experimental models (e.g. Lorenz-96 in session 3.1.1), so it is unclear how the ground truth of hidden state and observations were generated (which frequency, which noise distribution).

7. Novelty relative to existing ensemble smoothers: the first case study is presented as "highlighting the necessity of regularization and demonstrating the applicability of EnKBS" to a relatively high-dimensional chaotic system. Covariance localization and inflation are, however, standard approaches in EnKS and other extended ensemble smoothers, which already work with small ensemble sizes in high dimensions. The manuscript should state explicitly what the continuous-time formulation adds in the experiments (e.g. accuracy, stability for small ensemblesize, cost, or theoretical convergence), ideally through a direct comparison with a discrete-time EnKS using the same regularization.

Other specific comments

1. Introduction: the claim that smoothing can "mitigates the detection delay of filters at the onset of extreme events" would benefit from a reference or the numerical evidence in the paper.

2. Section. 2: Equations (3) and (4) both use $d\mathbf{x}^{(i)}$ for what appear to be different quantities (pseudo-time update with s versus physical-time increment with t). Please make this clear.

3. Section 3.1.1: two equations use the same symbol σ for different noise levels. Either distinguish them (σ_1, σ_2) or remove the redundant equation.

4. Localization (page 8): the distance $d(i, j)$ depends on the state values x_i, x_j or on their indices i, j (e.g. cyclic index distance). If the latter, please correct the definition in the Localization subsection.

5. Section 3.1.2: the statement that EnKBF remains stable down to $m = 3$ and that localized EnKBS remains stable down to $m = 5$ "(not shown)" is central to the small-ensemble claim of the abstract. Please show the EnKBF and EnKBS trajectories for both scenarios (with and without localization) so the reader can assess performance.

6. Metrics (ACI and CIR): please give explicit formulas together with references, and state how the confidence/credible intervals used in these metrics are computed from the ensemble.

References

[Anderson(1982)] B. D. Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326, 1982.