Review of the manuscript gmd-2021-306 “A fast, single-iteration ensemble Kalman smoother for sequential data assimilation” by Colin Grudzien and Marc Bocquet

Pavel Sakov

November 7, 2021

1 General comments

The paper gives an overview of the iterative ensemble Kalman smoother methods; introduces a new scheme called single-iteration ensemble Kalman smoother (SIEnKS); and runs a number of tests on it with Lorenz-96 model. It provides algorithms for the most significant in the context of the paper methods, and uses the open-source Julia package DataAssimilationBenchmarks.jl for the benchmarking.

It is not an easy paper to review, mainly due to its sheer length, but also because of some vagueness in formulating the purpose and results, and some language used. Just to be more specific – it is impossible to get an idea about the new method from a rather lengthy abstract apart from that it is new.

It would be a too large effort for me to give a proper review of the paper of this length; instead I will list a few points that may or may not be accepted by the authors.

2 Issues

1. My preference (perhaps contrary to the established practice) is to avoid characterising EnKF methods as “ensemble variational”. The Kalman filter *is* a variational method, even if formulated in a sequential way. To me, it can make sense to talk about “ensemble variational” if the method explicitely uses the model adjoint.

2. Further, unlike to 3/4D-Var, I can not get any sense of the "outer loop" terminology in the paper. How is it different to the iterative minimisation?

   (Also, I am not a big fan of the "4D-MAP" abbreviation.)

3. **L. 158-159: "ensemble is drawn", "columns sampled".**

   In deterministic EnKF methods the ensemble *is not* a stochastic ensemble, but rather a (possibly, lossly compressed) factorisation of the state error covariance. Indeed, there can be some stochastic elements even in mainly deterministic systems, e.g. due to random perturbations of forcing etc., but using that statistical terminology is largely misleading, I belive.

4. **L. 363: "Raanes (2016) demonstrates the equivalence of the EnKS and the Rauch-Tung-Striebel (RTS)."**

   I am surprised that this needs to be demonstrated.

5. **L. 410: "In the perfect, linear-Gaussian model, this formulation of the IEnKS is actually equivalent to the 4D-EnKF ..."**

   This may be true in the specific context, but does not make sense on its own: how can a smoother be equivalent to a filter?

6. **. L. 449: "A revised and simplified form of the Gauss-Newton IEnKS, transform variant is presented for the first time in Algorithm 7."**

   Hmm... I trust the authors that this algorithm must be a substantial achievement. Just wanted to note that it has 35 lines (with some functions), while a similar one takes only 19 lines in Table 2 of Bocquet and Sakov (2014).

7. **The paragraph l. 449-454.**

   I would add "similarly to MLEF" somewhere in this paragraph.

8. **The SIEnKS.**

   I struggle to understand what is actually new in the SIEnKS compared to the Lin-IEnKS (sorry). It would help if the authors explained it explicitely and/or put the two algorithms side by side to see the difference.

9. **The paragraph l. 491-504.**

   Because this paragraph writes more than just a few words on RIP, it may be useful to note for a non-specialist reader that RIP does not minimise the same cost function as the IEnKF. It makes the best fit to observations in the model subspace by assimilating them multiple times until convergence. This is equivalent to minimising the cost function with the forecast covariance multiplied by a large number.

   Further, in regard to l. 502-504 – there is no such thing as a single iteration RIP, I guess.

3 Conclusion

Overall, the paper is well written and in a way represents a nice overview of the ensemble smoother methods. Perhaps, it could read a bit easier if a more straightforward terminology was adopted. I have no comments on the experimental part of the paper, partly because it failed for me to generate the excitement of benchmarking a new extension, due my failure to understand the essence of the novelty of the SIEnKS. Having said that, the experiments look well done.

In my view, the paper aligns with the goals of the GMD and will be interesting to the growing EnKF data assimilating community. I leave it to the authors' discretion whether and how revise it to address the issues raised in this review. This probably translates as recommendation to **accept with minor revision**.

References

Bocquet, M. and P. Sakov, 2014: An iterative ensemble Kalman smoother. *Q. J. R. Meteorol. Soc.*, **140**, 1521–1535.