# Reviewer's Comments for gmd-2019-132

After reading the paper, and clearly having in mind your research questions: Is an online EnKF really faster than an offline EnKF? Can an offline EnKF be as fast as, if not faster than, an online EnKF with a good framework and algorithms using advanced techniques of parallel IO? I think the paper can be improved in the following directions:

1. In section 1 (Introduction), you discuss efficient EnKF formulations: EnKF methods that exploit the rank-deficiency of ensemble covariances to come up with efficient EnKF formulations (i.e., by using Sherman Morrison, SVD, etc.), and efficient EnKF formulations which account for localization. You can enrich your paper by distinguishing between these two classes of filter derivations; it is not a good idea two present both families as a single one.

2. In section 2, I do not agree with the statement: *the Cholesky decomposition is more efficient than the SVD, but the SVD is more robust if the matrix is significantly ill-conditioned.* For unlocalized EnKF formulations, via the SVD decomposition, we can obtain EnKF implementations whose computational effort reads $O(n \cdot N^2)$, where $n$ is the number of model components and $N$ is the ensemble size. On the other hand, by employing a direct solution (i.e., by employing a Cholesky decomposition), we can get EnKF formulation whose analysis steps can be computed with $O(n^3)$ long computations.

3. As you may know, **localization methods** are crucial to getting accurate analysis states. These methods mitigate the impact of spurious correlations in EnKF based formulations. In operational contexts, when domain decomposition is employed during assimilation steps, the dimension of local boxes can still be much larger than that of ensemble sizes. Therefore, local analysis increments can be poorly estimated at each sub-domain. How does your proposed method deal with this? I would like you to increase the discussion in this direction, you have already cited two papers that deal with this:

   (a) Anderson, J. L. (2001). An ensemble adjustment Kalman filter for data assimilation. Monthly weather review, 129(12), 2884-2903.

   (b) Nino-Ruiz, E. D., Sandu, A., & Deng, X. (2019). A parallel implementation of the ensemble Kalman filter based on modified Cholesky decomposition. Journal of Computational Science, 36, 100654.

4. While localization methods reduce the impact of spurious correlations, **covariance inflation** mitigates the impact of the under-estimation of sample variances. Most of the square-root based formulations (i.e., ETKF, and LETKF) suffer from under-estimation of sample variances, but, you do not mention this relevant topic in your entire paper.

5. When domain decomposition is employed, it is very common to send boundary information to obtain physically consistent solutions, in this manner, solutions at different processors do not look like independent domains but, they are "*connected*" since neighboring sub-domains share boundary information. I may misunderstand but, are you sharing boundary information among neighboring sub-domains? Thus far, it seems like not, right? If this is correct, how do you guaranty that global solutions (once all local analysis are mapped back onto the global domain) are consistent with the physics and the dynamics of numerical models?.

6. What observational network do you employ during experiments? are employing a full observational network (all model components are observed)? What is the performance of your method (in terms of accuracy) as the observational network becomes sparse?

I hope these comments help to improve your paper.