

## **Response to Reviewer #1**

Dear Reviewer,

We would like to thank the reviewer's instrumental suggestions, which have greatly improved our manuscript. We have conducted new experiments to fully explore the performance of the new method. Our manuscript is significantly revised to include the results of new experiments and to address your comments and suggestions. Please see our point-by-point responses to your comments as follows.

### **Major comments:**

1. If it's beneficial to artificially increases the ensemble size, why did the authors only consider to increase the size by one or two in their implementation? In general, how many more pseudomembers can the authors include before the EnKF's performance saturates (or perhaps deteriorates).

Thank you very much for this important comment. We have conducted new experiments with a large ensemble size (20 members). As shown in the new Table 3, the benefit of adding pseudomembers based on IESVs increases as the number of pseudomember increases, particularly when the standard LETKF (M20\_CNTL) has large analysis RMSE. Using more than 10 IESVs can reduce the analysis RMSE by more than 45% and the IESV1 (M20\_Orth\_ESV1) provides the dominant effect. The improvement rate with 15 IESVs becomes saturated for the following 30-step forecast. Such a performance (M20\_Orth\_ESV1-15) is better than the 25-member standard LETKF in general, and even better than the 30-member standard LETKF for the group of mildly large analysis errors.

We have included the following figure (new Figure 6) and relevant discussion at lines 155-159 in the revised manuscript.

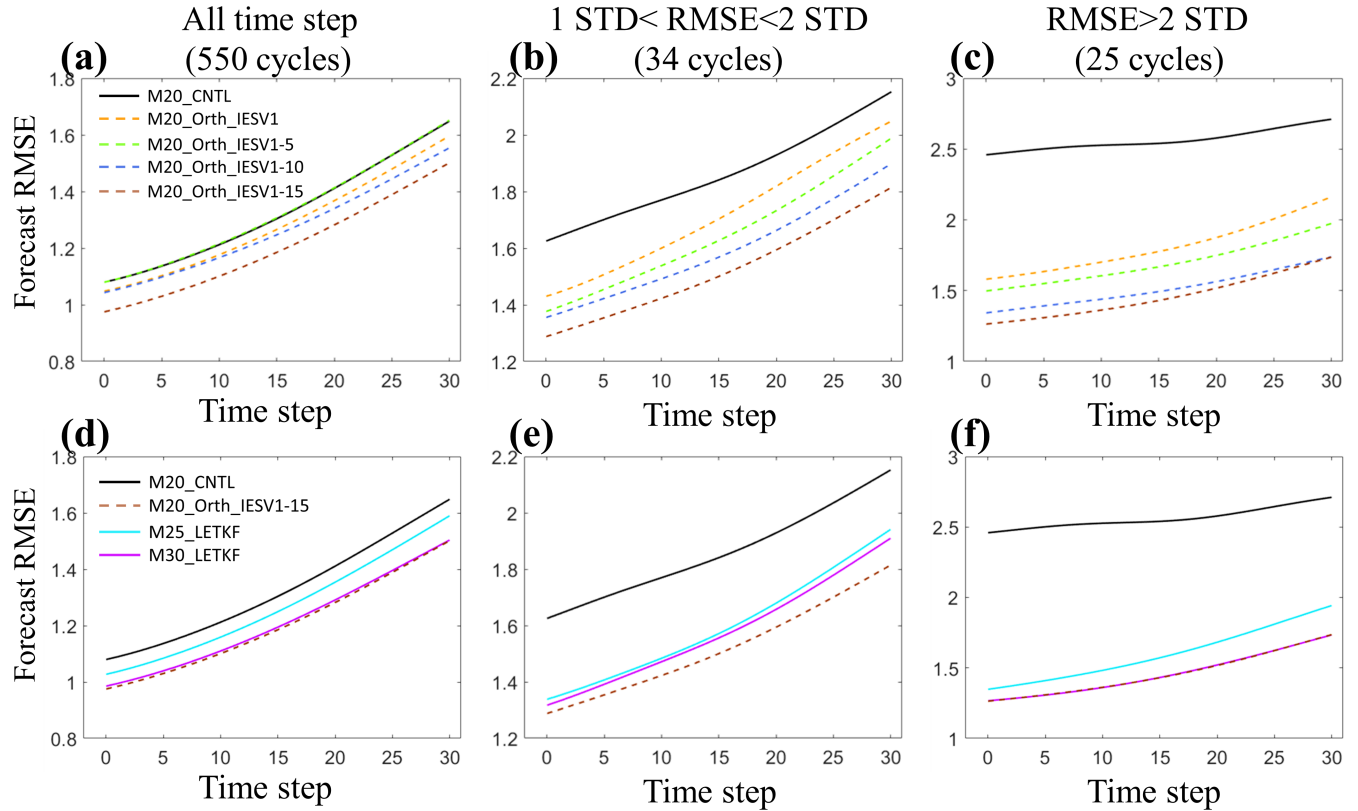


Figure: The mean RMSE of the forecast error during the 30-step integration. CNTL\_M20, M25-mem\_LETKF and M30\_LETKF use the standard LETKF with 20, 25 and 30 members, respectively. M20\_Orth\_IESV1, M20\_Orth\_IESV1-5, M20\_Orth\_IESV1-10 and M20\_Orth\_IESV1-15 use 20 members and orthogonal components of IESVs as the pseudomembers.

2. The authors considered a few different ways (EMV, IESV, RSV) to add ensemble members and concluded that EMV appears to be better. Any reason to explain the better performance of EMV?

In dynamic systems, the ensemble perturbations evolve upon the mean structure. In other words, the dynamic errors are expected to relate to the instabilities existing in the ensemble mean. Therefore, the structure of the ensemble mean is effective in expanding the ensemble space to capture the error. A better performance with EMV as the pseudomember suggests that the background error also projects on the structure of ensemble mean. We observed that EMV is particularly useful when the ensemble size is small. With a small ensemble size, the ensemble can not capture the forecast error well and thus EMV is particularly useful. With a large ensemble size (and longer localization), the large-scale error in the background mean state is much reduced and the structure of the mean state is less effective as the pseudomember, compared with IESV1. We have provided the explanation at lines 168-170.

3. There are many factors, such as inflation factor, frequency and density of the observation operator, the variance of observation errors, the length of the assimilation time window and even random seeds, which could potentially affect the performance of an EnKF. I'd suggest that the authors conduct an extensive investigation on the impacts of these factors on the performance of the EnKF.

We agree that we should investigate the performance of the newly proposed method under different assimilation configurations. To address this comment, we have carried out a series of experiments to investigate how the new framework is sensitive to assimilation frequency, density of observation, assimilation interval and random seeds.

Our conclusion remains under all different assimilation configurations, but with different degrees of improvement rate. Therefore, we can conclude that adding pseudomembers can effectively improve the analysis accuracy, in particular when the standard LETKF performs poorly. We note that the inflation factors used in all experiments are optimized.

Please see the new Table 3 and the relevant discussions section 3.2 in the revised manuscript.

4. In particular, I'd suggest the authors repeat one experiment for a number of times (each time with a different random seed) to reduce the effect of statistical fluctuations. In the current manuscript, I got the impression that the authors did each experiment only once, which does not make much sense for the L96 model, since it's quite cheap to run. If the same experiment is repeated multiple times, then I'd further suggest that the authors use a table to summarize the results (RMSE) in the form of mean +/- STD, which would make the performance comparison statistically more meaningful.

Following the reviewer's suggestion, we have repeated one standard experiment 10 times with different random seeds to derive statistically meaningful results.

We summarized the RMSE of the experiments in Table 2 in the form of mean and STD. It is clear the EMV is effective in reducing the analysis RMSE when the standard LETKF (CNTL) has large analysis errors and using the orthogonal component of EMV shows a better performance than EMV. However, adding the orthogonal component of IESV1 as the pseudomember results in the largest RMSE STD among the experiments. Even when the orthogonal IESV1 is sometimes less effective, it is still useful in improving the analysis accuracy (a smaller mean RMSE than that of CTRL). Please see the discussion at lines 130-141.

#### **Specific comments:**

1. L18, P1: The statement "The BEC is sampled by the background ensemble" does appear precise, since one cannot sample a covariance (perhaps you want to say "estimate" instead?)

Thank you. We have replaced “sampled” to “estimated” for this sentence.

2. L22, P1: In linear algebra, the statement "cause the ensemble space to be underestimated" also appears problematic. In which sense is a (linear) space underestimated? My guess is that here by "ensemble space" the authors possibly mean ensemble variance instead. If so, it seems that there are many similar instances later in which the terminology "ensemble space" is not correctly used (e.g., as in "apply SVD to the ensemble space" in L63, P3).

Thank you for your suggestion. We have modified this sentence to the use of finite ensemble size can cause underestimation of the background error variance and the background error correlation is less optimally represented due to the sampling error.

We have also revised the sentence at line 63 to emphasize that “SVD is applied to find the orthogonal vectors to represent the space spanned by the ensemble members.” We also clarified the sentences using the ensemble space (e.g. lines 68, 93, 104).

3. The authors stated that using Eq. 3 can preserve the ensemble mean and spread. It'd be good to provide a proof for this statement.

Appendix B is provided as a proof for the statement of Eq. 3.

4. L92, P3: Rephrase the text "as a demonstration of proof of concept".

This sentence is removed from the revised manuscript.

5. In Caption of Figure 4, "according" is duplicated.

Sorry for the typo. We have corrected it.