Response to the comments of Referee #1

We are grateful to the reviewer for his careful reading of the paper once again. The reviewer's comments helped us greatly to further improve the presentation of our ideas and results. The changes made in the revised manuscript, following the reviewer's suggestion, are marked red (new addition) and gray (deleted parts). In what follows, we repeat each of the reviewer's comments in italic, and provide our response in roman.

The revised paper has added material that attempts to support the (new) sentence (Lines 414-416) The central argument of this paper is that applying a single localization function for the localization of covariances between multiple state variables in an EnKF scheme may lead to a rank deficient estimate of the background covariance matrix. As I describe in the Matters which require attention portion of my review, much of the supporting material for this new sentence is either unsubstantiated or incorrect. As such, despite improvements that have been made to other parts of the paper and despite the fact that I still think that their suggested methods for inter-variable localization and their description of the generalized multi-variate Askey localization function would be of interest to readers of non-linear processes in geophysics, I do not think the paper should be published in its current form. My overall recommendation is to accept the paper after revisions that address the concerns outlined below.

Matters which require attention

1. Lines 85-87 suggest that the Schur product between a semi-positive definite symmetric matrix (like that of the Gaspari-Cohn localization matrix) and a pre-existing positive definite matrix may result in a matrix that has zero eigenvalues. If understand localization correctly, it is incorrect to suggest that univariate localization will reduce the rank of the original covariance matrix. In most practical examples, it massively increases the rank of the covariance matrix (without creating negative eigenvalues) and this is one of the main reasons it has proven to be so useful. If the authors have a relevant example where univariate localization reduces the rank of the matrix to which the localization is applied, please include it in an Appendix. Otherwise drop this sentence.

We agree with the reviewer about univariate localization. Our point there was about the localization, in the system with multiple state variables, with the same localization function (with the same localization length for each variable), and when the cross-covariance between multiple state variables are not negligible. In that case, localization may not work as efficiently as it would for the univariate case, due to the problem of rank deficiency of the localization matrix.

We modifed the sentence to avoid confusion.

2. Lines 87-88 suggest that the symmetry of a matrix implies that its eigenvalues are non-negative. This is also incorrect. For example, the negative of the identity matrix is symmetric but both of its eigenvalues are negative. Symmetry does not imply positive eigenvalues.

We did not mean that symmetric matrices have non-negative eigenvalues (it is clearly false). Rather, we meant that $\mathbf{P}^b$, which is positive semi-definite in its nature, is also symmetric and has non-negative eigenvalues. We agree that our comment can be confusing and we modified the sentence.

3. Lines 99-101. Do the authors know that Kang et al. (2011) was not also motivated by the need to achieve a semi-positive definite background error covariance matrix? Zeroing out the inter-variable covariances does not produce negative eigenvalues. Many researchers have been trying to find localization functions that filter spurious correlations AND do not produce negative eigenvalues. I think you need to justify your claim that Kang was not interested in preserving semi-positive definiteness or drop the statement.

It appears to us that the primary goal of Kang et al. (2011) is to introduce the localization method that filters the spurious correlation between distinct variables, not about prserving positive-definiteness or rank of the matrix. We checked the reference again to make sure. Words such as "positive-definiteness", "rank", or "eigenvalues" were never mentioned.

We did not mean to say that they were not interested in preserving positive definiteness, or they were ignoring it. We just stated that it was not their motivation. We slightly modified the sentence to make this clear.

4. Lines 182-184. The statement "although C is rank-deficient and thus so is the localized covariance matrix" is incorrect. To see this, take the element wise-product of the localization matrix defined by (5) with the identity matrix one then will recover the identity matrix which has full rank and is not rank deficient.

The reviewer is correct. Whether the localized covariance matrix $\tilde{\mathbf{P}}^b$ is rank-deficient or not depends on the structure of $\mathbf{P}^b$ (your example of $\mathbf{P}^b = \mathbf{I}$ is an extreme case that $\mathbf{P}^b$ is of full rank with zero off-diagonal block matrices. However, usually $\mathbf{P}^b$'s are not in this "nice" shape.). We modified the statement to avoid confusion.

5. Lines 414-416. Here we here that the central argument of the revised paper is that applying a single localization function for the localization of covariances between multiple state variables in an EnKF scheme may lead to a rank deficient estimate of the background covariance matrix. I do not believe that rank-deficiency is the primary problem with univariate localization so making this the central argument of the revised paper has decreased the papers appeal for me. Suppose I have many more ensemble members than variable types; e.g. I might have 5 model variables at each grid point and 80 ensemble members. In this case, I could use a univariate localization matrix of the form given by equation (5), where each of the C_0 *matrices was the identity matrix and then apply this to my 80 member ensemble covariance matrix. The resulting covariance matrix would almost certainly have full rank even if the number of model grid points was in the tens of millions because the localization has zeroed out all inter-grid-point covariances leaving 80 ensemble members to describe the covariance between the 5 model variables on each grid point. Without doubt the rank of the matrix obtained using (5) would be very much higher than the rank 79 of the unlocalized ensemble covariance matrix. Note that it is easy to construct dynamical data assimilation systems where the rank of the true forecast error covariance matrix is rank deficient (e.g. Bishop et al., 2003, J.Atmos. Sci.). Hence, the best localization strategies may be those that stop short of making the localized ensemble covariance matrix full rank. I think there is broad agreement that the localization matrix should attenuate spurious correlations and also ensure that the localized covariance matrix has no negative eigenvalues. (To ensure this, one only needs to ensure that the localization matrix is semi-positive definite). The value of your study is that you have presented new and effective methods for generating multi-variate semi-positive definite localization matrices. In my view, thats the only argument you have to make.*

We completely agree with the reviewer. As we mentioned in our previous response, whether or not the rank of localized covariance matrix increases or not depends on the structure (or rank) of the original matrix $\mathbf{P}^b$ and as the reviewer points out, if you have large enough ensemble members, then $\mathbf{P}^b$ may be close to be of full rank.

The line numbers the reviewer gave about this point do not seem correct (lines 414-416 are part of the acknowledgment). We guess that the reviewer commented on the first sentence of our Discussion section. We modified this following the reviewer's suggestion.

Minor comments:

1. Line 16. Abstract: In order to be just a bit clearer, change multiple state variables to multiple state variables that exist at the same location

Done.

2. Line 165: Replace of d by of the separation distance d

Done.

3. Eqs (9)-(11) please add more discussion of the meaning of the terms μ_{12}, μ_{11} *and* μ_{22} *in equations (9) through (11). Perhaps mention the values you ended up choosing to use for these parameters in your experiments? In equation (9) and (11) can one think of the factors beta and B as simply being normalization factors that ensure that the correlation between the variables becomes equal to unity when the distance between the variables is equal to zero?*

The values of μ_{ij}'s used in the experiments have been mentioned at the end of Section 3.c. We are not sure if we could give physical meanings to the parameters μ_{ij}'s. Obviously allowing these terms to vary (rather than fixing them at a value) makes the functions in (9) and (11) more flexible. Part of the role of β_{ij}s and the function B is, as the reviewer mentions, normalization so that ρ_{ij} is a proper correlation function. However, a more important role for β_{ij}'s and B is to introduce cross-correlation structure and to ensure positive-definiteness of the resulting localized marix.

4. Figure 4. Please give more explanation of Figure 4 within the figure caption. What is the difference between the circles and the squares? Perhaps reiterate the percentile significance of the box and whiskers plots. Clarify that the reason that there are no grey lines for the S4, beta=1 case is because the Askey function is not defined in this case, etc. Consider adding more explanation of what the figure is showing to the other figures as well.

Done.

5. Figure 3. Consider rephrasing the caption to For the partially observed case, locations of observations of X and Y are indicated by the black dots and grey circles, respectively.

Done.

Remark

Please also note that the year of the paper cited (Porcu et al.) has been corrected from 2012 to 2013.