Ensemble Riemannian Data Assimilation over the Wasserstein Space

By: Sagar K. Tamang, Ardeshir Ebtehaj, Peter J. van Leeuwen, Dongmian Zou, and Gilad Lerman

Responses to review comments (Reviewer 2)

Comment: This paper introduces an optimal transport framework for updating discrete representations of posterior probability density functions during ensemble data assimilation. This work further provides proof-of-concept assimilation experiments comparing the algorithm the authors introduce (and call EnRDA) to standard DA methods (3D-var, particle filters, ensemble Kalman filters). The authors mention two serious issues that will need to be overcome for EnRDA to become a viable strategy for users in the DA community: 1) the high computational expense (that scales super-cubicly with the ensemble size) associated with computing optimal transport maps and 2) the need for the observation operator to be bijective, e.g. all state dimensions must be observable.

Reply: We very much appreciate the thoughtful review and comments on the manuscript. We have included a track-change color coded version in which red colored text is the updated text in response to the reviewer's feedback. Some of the changed text in the manuscript are also copied in the replies and shown with red color for convenience. Please also find item-by-item replies to your comments as follows.

Some minor comments:

1. The reader would benefit from more discussion of the Sinkhorn algorithm and how/why gamma is chosen as well as how eta is chosen in practice. On line 290, you mention that these parameters are set by "trial and error". Can you offer the reader more guidance on how to make these choices in practice, even if ad-hoc? Do these or should these parameters vary at different assimilation time steps? On line 290 you have gamma=3 and in the caption of Figure 5 you have gamma=0.003. Are both correct? Are these for the same DA experiment at different times? Or different experiments? This should be clarified in the text. In Figure 3 you demonstrate that one order of magnitude change in gamma results in quite different joint distributions with the same prior and observation pdfs. What does a 3 order of magnitude change in gamma do?

   Reply: Thank you for this comment. The displacement parameter η can be tuned offline through cross-validation by minimizing the mean squared error or any other error metric of interest. The error shall be defined with respect to a reference point such as ground-based observations. Please see the updated text in lines 196–199, where we have addressed this comment. However, the value of the regularization parameter is highly dependent on the transportation cost matrix. In practice, one can begin with γ set as the largest element of the transportation cost matrix and gradually reduce it to find the minimum value of γ that provides a stable solution of Sinkhorn's algorithm. Please see the updated text in lines 260–264. Both the displacement and regularization parameters are static in our implementation, however, future research may come up with new ideas for dynamic updating. Please see line 417–426, where we addressed this issue.

Thanks for your attention to the details. We double checked and the reported values of γ are correct. The choice of γ is different for different experiments and dependent on the transportation cost matrix. The value of $\gamma = 0.003$ is set for two arbitrary Gaussian mixture models defined in Figure 3 whereas $\gamma = 3$ is set for our experimental setting in the 1-D advection-diffusion model. Please see the updated text in lines 260–264, where we clarified that the value of the regularization parameter is different for different experiments.

2. The barriers to wide usage of this approach are quite high, yet if overcome the EnRDA could prove a very powerful DA method. As such I believe these barriers and the research advances needed to overcome them warrant a longer discussion than you offer in section 5.

   Reply: Thank you for your perspective. We agree and did our best to be upfront about the barriers of the presented approach. Please see the expanded discussion about the limitation in Section 5, 432–434, and 449–452 (also copied below for convenience).

   "Another promising area is to utilize EnRDA only over the observed dimensions of the state-space and similar to the EnKF, use the ensemble covariance to update the unobserved part of state-space through a hybrid approach."

   "Furthermore, recent advances in approximation of the Wasserstein distance using a combination of 1-D Radon projections and dimensionality reduction (Meng et al., 2019), can significantly reduce the computational cost to make EnRDA a viable methodology for tackling high-dimensional geophysical DA problems."

3. Typographical mistake — you have two periods ending a sentence on line 398.

   Reply: Thank you, fixed.

Once again we would like to take the opportunity and thank you for the insights and feedback that helped us to improve the manuscript. We hope that the replies and changes we made in the manuscript meet your expectation.

References

Meng, C., Ke, Y., Zhang, J., Zhang, M., Zhong, W., and Ma, P.: Large-scale optimal transport map estimation using projection pursuit, in: Advances in Neural Information Processing Systems, edited by Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., and Garnett, R., vol. 32, Curran Associates, Inc., 2019.