<u>**Revision Notes to Reviewer's Comments**</u>

We have completed a minor revision of the earlier version of the manuscript by taking into consideration of the response of the reviewer-2. We hope this revised manuscript addresses all the reviewer's comments.

Reviewer's comments and suggestions greatly improved the structure, content, and the quality of this manuscript. We thank the reviewers for their valuable time spend in careful reading and constructive comments and suggestions. Their critiques gave us an opportunity to approach this research problem from various angles and expanded our understanding about the Northwest Pacific Ocean circulation.

The following revision notes explain details of what we have done to address specific comments from the each of the reviewers. *The reviewer's comments are in red text in italics*. Authors response is in black text. *Changes in the manuscript is in blue text in italics*.

<u>**Reply to Comments of Reviewer # 2**</u>

1 Specific comments

Caption to Figure 1, last sentence: change "assimilation region" to "region where oceanic observations where assimilated".

Thank you, we revised the Figure 1 caption as:

"Model domain and the major large-scale circulation features are shown in this figure. The shading marks the mean SSH in m, and thin green contours marks the SSH standard deviation at contour interval of 2 cm, computed from AVISO daily gridded SSH analysis for the period 1993 to 2018. The thick green contour marks the SSH standard deviation of 12 cm. The black box marks the region where oceanic observations were assimilated.".

Line 127: The grid-scale relaxation times corresponding to the diffusion coefficients when scaled by the model grid resolution: sounds tautological: "grid-scale... scaled by the model grid resolution". It would be better to rephrase these sentences, e.g., The corresponding relaxation times for the grid-scale features are 38 days for harmonic diffusion..., while for the features with the typical scale of the first baroclinic Rossby radius (60km, wavelengths 250km), the respective relaxation times are 1.1 and 3.9 years respectively.

Thank you, we revised this sentence as:

The corresponding relaxation times for the grid-scale features are 38 days for harmonic diffusion and 12 days for bi-harmonic diffusion, while for the features with the typical scale *of the first baroclinic Rossby radius (60 km), the respective relaxation times are 1.1 and 4.1 years respectively.*.

Line 191: "i.e., uncorrelated in space and time" does not imply that the matrices are diagonal, because different controls could be correlated between each other in coinciding space-time locations. Remove this comment.

Thank you, we revised this sentence as follows:

The formulation of $\mathbf{J}$ *assumes that the error covariance matrices* $\mathbf{Q}(t)$ *and* $\mathbf{R}(t)$ *are diagonal, i.e., uncorrelated.*.

Lines 200-229. This new piece of text requires more adjustments in terminology. I think the authors should try to switch from unclear wording (such as growth of nonlinear instabilities, incorrect? model state, non-smooth? cost function, linearized? adjoint, etc.) to simple mathematical concepts of linear algebra understandable to a general reader. The point is that the adjoint [of the tangent linear] model used to compute "sensitivities" is intrinsically unstable in non-linear regimes. Since this instability predominantly occurs on the grid scale (i.e., unstable modes are dominated by spatial variations on grid scales), the resulting gradient ("sensitivity map") could be severely contaminated by (biased to) these instable modes (eigenvectors) of the adjoint model if the integration time is long compared to the e-folding times of the fastest growing modes.

In saying "we are not making the modes stable and reducing their eigenvalues, but instead altering them to become smoother lower modes", the authors appear to have certain misunderstanding of the diffusive stabilization technique they use: if the eigenvalues of the unstable modes (eigenvectors) were not reduced, the magnitude of the "smoother lower(?) modes" would be as large as the magnitude of the fastest growing eigenvectors of the "unfiltered" adjoint propagator, i.e. of the adjoint model integration without inflated diffusion. I also strongly recommend discussing the choice of the diffusion magnitude. For example, what was the motivation of electing longer window at the expense of a larger diffusion (e.g., in the light of the existing theoretical estimates). Again, I recommend presenting the related material (including Appendix B) in the light of stabilizing the adjoint model via inflated diffusion. In the present form the presentation seems a bit clumsy and hardly understandable to the majority of the readers (see also my request to modify/update Appendix B, lines 774-783).

We revised the paragraph as follows:

One of the known challenges of using adjoint based 4DVAR assimilation systems is the growth of nonlinear instabilities with integration time. Two approaches are normally sought to deal with this problem: 1) increasing diffusivity and viscosity coefficients in the adjoint of the background model [Hoteit et al.(2005), Kohl et al.(2007)], and 2) decreasing the length of the assimilation window. In this study, we chose an assimilation window of two months, which is the longest window that produced a skillful optimized state estimate using increased *diffusivity and viscosity coefficients in the adjoint model. Our hypothesis is that the adjoint of tangent linear model that compute the gradients or "sensitivities" become unstable in non-linear regimes that are most prominent at small-scales. This results in corruption of the sensitivities which grows with integration time and leads to less useful gradients that can slow or stop the iterative optimization over long assimilation windows. By increasing the diffusivity and viscosity in the adjoint of background model, we remove these small-scale features in the sensitivities that will allow us to use longer assimilation windows in our eddy-permitting model.*

In addition to increased diffusivity and viscosity, the KPP mixing parameterization is disabled in the adjoint model simulation to minimize its contribution to the nonlinearity of the adjoint in order to permit longer assimilation windows. The mixing parameters computed by KPP in the forward run were retained in the linearized adjoint, so it is only the effects of changes in the parameters due to changes in the model state that were not included. This limits the fidelity of the adjoint in the surface layer and leads to higher model representational error. Relative to the forward model integration, the horizontal second-order viscosity and diffusivity coefficients in the adjoint simulation were increased by a factor of 50 (50×10^2 m^2 s^{-1}) and the fourth-order coefficients were increased by a factor of 5 (5×10^{11} m^4 s^{-1}). The vertical second-order viscosity and diffusivity coefficients were also increased by a factor of 50 ($50 \times 2 \times 10^{-5}$ m^2 s^{-1}). These coefficients were selected to allow successful model fits to observations over all of the two-month assimilation windows. The sensitivity of the adjoint model simulation to changes in the viscosity and diffusivity coefficients is examined in Appendix B for the special case of a linear cost function. .

Table 1: adding the right column is informative. I would also recommend to add "x 60" after the numbers in the right column (e.g., change 74,091 to 74,091 x 60 if there are 60 daily control fields) for better clarity. Same with open boundary controls.

The values shown in the right most column marks the number of control points for each day of the assimilation window. We modified the Table including this detail.

Line 388-390: "non-linearities in the solution, which we smooth from our gradients, mean that control is lost and the cost descent is less efficient...." To clarify presentation, I would change this to "the small-scale structure of the optimal solution is not well reproduced, as it is removed from the gradients by the inflated diffusion in the adjoint model, causing certain loss of the control by the initial conditions and less efficient descent ..."

Thank you for your comment. We revised the sentence as below:

Finally, the small-scale structure of the optimal solution is not well reproduced, as it is removed from the gradients by the increased diffusion in the adjoint model, causing certain loss of the control by the initial conditions and results in less efficient cost descent for longer assimilation periods. .

Lines 774-783: This paragraph looks somewhat naive. Eigenvectors of AT are complex while those of D are real. As a consequence, AT and D would never commute, so "leaving the commutation of the operators as a research question" is a bit senseless. Instead, I would recommend a quick alternative research: assessment of a few largest eigenvalues of AT using, say, a standard ARPACK routine for an implicit matrix (whose action on vector is defined via a separate routine – adjoint model in your case). You can do that for, say, 5-day (or longer, if possible) integrations of the adjoint model with original and inflated diffusivities. As an extra option, you could use an ARPACK routine that computes the respective eigenvectors and see changes in their spatial structure caused by inflated diffusion. Results of these calculations would be instructive and interesting to potential readers.

Thank you for your comment. We revised this paragraph clarifying about the commutation as follows:

Increasing the viscosity and diffusivity in the adjoint model can be expressed as augmenting the adjoint propagator at each timestep, $\mathbf{A}^T$*, with a viscous operator:* $\mathbf{A}^T - K_h\mathbf{D}$*, where* K_h *is the additional horizontal viscosity (in this case a space- and time-invariant scalar) and* $\mathbf{D}$ *is the horizontal viscous operator. Since the two operators:* $\mathbf{A}^T$ *and* $\mathbf{D}$ *do not commute in general, the eigenvectors of the sum will not be the same as for the adjoint along, so we refer to the increased viscosity as a smoothing, not a change in the eigenvalues. The enhanced viscosity will certainly reduce the growth of sensitivities, but may also change the modes of evolution, not merely their eigenvalues depending on the commutation of the two operators.*

References

[*Hoteit et al.(2005)*] Hoteit, I., B. Cornuelle, A. Kohl, and D. Stammer, 2005: Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **131 (613)**, 3659–3682.

[*Kohl et al.(2007)*] Kohl, A., D. Stammer, and B. Cornuelle, 2007: Interannual to decadal changes in the ECCO global synthesis. *Journal of Physical Oceanography*, **37 (2)**, 313–337.