Response to Reviewer

By Guokun Lyu on behalf of all coauthors

We thank the reviewer for their careful reading of the manuscript and for their advice to improve the work. We have revised the manuscript according to the reviewer's recommendation. Below, we respond to the reviewer's questions and suggestions. Our responses are highlighted in blue.

GENERAL COMMENT:

The manuscript explores the impact of improved approximation of the adjoint code for the viscoplastic (VP) sea ice model on the overall quality of the optimized solution. The authors obtain quantitative estimates of the improvement and demonstrate that a better approximation of the adjoint velocity fields provides a more realistic distribution of the optimal corrections among other components of the sea ice state. Although the manuscript is of certain interest to ice modeling community, it has a number of deficiencies that need to be addressed before publication. In particular, the authors need to better articulate the position of their Arctic ice modeling system among other Arctic ice modeling systems with data assimilation capabilities. Presentation of the material also needs substantial improvements: I would suggest editing the manuscript by a person with a better command of English.

Response:

We thank the reviewer's suggestions. The complexity of this coupled ocean-sea ice model is similar to most of the current Arctic ocean-sea ice modeling and assimilation systems (see Table 1 in Uotila et al., 2019). The reconstructed sea ice and the mixed-layer sea water temperature and salinity in our previous Arctic reanalysis datasets are better than the TOPAZ4 and PIOMAS datasets (Lyu et al., 2021b). Section 2.1 explains more details about the thermodynamic-dynamic sea ice model. Besides, we edited the manuscript with native speakers.

SPECIFIC COMMENTS:

Introduction: the authors need to expand their overview of the data-driven ice modeling. In particular, they should discuss advantages of the adjoint/variational approach in forecasting seasonal changes of the sea ice state as compared to sequential methods (e.g., EnKF). Attention should also be paid to limited of differentiability of the VP rheology that imposes certain constraints on the utility of variational methods in ice forecasting.

Response:

We thank the reviewer's comments here. **We have introduced different data assimilation techniques in Paragraphs 3-4 in "1 Introduction".** In this part, we introduced the advantages/disadvantages of statistical-based and adjoint-based assimilation methods. Besides, we discussed the future developments and applications of the adjoint method in the coupled ocean and sea ice model and its potential difficulties in Section 6.

The TL/ADJ model is derived from the discrete coupled ocean and sea ice model. Besides "the limited differentiability of the VP rheology," other factors, including 1) nonlinearity vs approximated tangent linear approximation; 2) modifications on the adjoint model to ensure the stability of the adjoint model (L180-184 in the manuscript); 3) the “if”, “where” etc. statements in the discrete model, further hamper the usefulness/accuracy of the adjoint model. **We have examined how well the approximated TL/ADJ could predict error propagation in the nonlinear model (APPENDIX A in Lyu et al., 2021a, and our answer to Section 3.2 below).** Since we concentrate on approximating the adjoint of sea ice rheology in this study, we add paragraph 5 (L 59-66) to introduce the problems in adjoint of sea ice rheology in the current adjoint ocean and sea ice model, highlighting the necessity of this study.

Section 2.1:

a) The utilized approximation to the adjoint of the tangent linear (TL) ice model should be described in more detail. Instability of the TL/adjoint codes of the Hibler’s model is caused by violation of the positive semi-definite (PSD) property of the linearized C-grid discretization of the momentum equation, so it is instructive to give more details on what modifications of the code were made “to facilitate generation of the adjoint model”. The cited paper by Losch et al (2010) is insufficient.

Response:

We thank the reviewer’s comment.

The MITgcm ocean model was developed with automatic differentiation software (Transformation of Algorithms in FORTRAN, TAF, Giering and Kaminski, 1998), which could automatically generate the adjoint codes based on the discrete nonlinear models. Losch et al. (2010) reformulated the dynamic-thermodynamic sea ice model from climate models on an Arakawa C grid to match the MITgcm oceanic grid. Further, **they modified many codes for efficient and accurate automatic differentiation using TAF.** The modifications are mainly on the forward model and tedious; therefore, we didn’t explain more details. **We describe modifications of the adjoint codes in Section 2.3 (L180-184).**

b) A description of the “control run” (first mentioned in Table 2) is missing.

Response:

The control run is iteration 0. We have added a description of the control run in L215 (before Table 2).

c) Description of the assimilation window (line 100): It is unlikely that ice state in, say, December is controllable by the initial conditions on January 1 (Table 3). Provide more detail on the formulation of the assimilation window. Was it a sequence of twelve monthly windows (e.g., line 163)

Response:

We use **a 1-year assimilation window (the whole 2012 year)** in the two assimilation experiments. This large assimilation window is unique to the ECCO-like reanalysis. To achieve this large assimilation window (beyond the predictability of the nonlinear system), we have to modify the adjoint model by:

1) disable the K-profiles mixing parameterization scheme

2) Increase the Laplacian diffusivity of heat and salinity to 500 $m^2 \cdot s^{-1}$ and lateral eddy

viscosity to 10,000 $m^2 \cdot s^{-1}$;

3) apply a spatial filter to sensitivity variables calculated in the adjoint of the thermodynamic sea ice model (see APPENDIX in (Lyu et al., 2021b) for detail)

4) modified the adjoint of a VP sea ice dynamic (see section 2.3 of this manuscript).

We have added descriptions of these modifications explicitly in L180-184.

It’s true sea ice conditions on January 1 are unlikely impacts the sea ice state at the end of the year due to sea ice memory. However, the adjoint method also adjusted the atmospheric forcing throughout the assimilation window to reduce the model-data misfits.

Section 2.3:

a) Second term in eq, 3 is incorrect: check arguments of the viscosity coefficients (no commas between the indices), replace the deformation rate tensor by its trace

Response:

We thank the reviewer for pointing out the mistakes. We have revised these mistakes in the revised manuscript (L157).

b) Lines 153-156: instability of the TL/adjoint code is caused by violation of the PSD property by the system solved by the LSOR method, and could be eliminated by using weak formulation of the ice momentum equation (Mehlmann & Richter, 2017). Besides, stabilization of the linearized rheology term could be achieved by augmenting of the exact adjoint code by Newtonian damping (Panteleev et al, 2021) rather than eliminating dependence of the viscosities on the deformation rates. In that respect, I would suggest to use the term “approximate adjoint” rather than “stabilized adjoint” throughout the text and add some discussion of the issue either here or in Section 6. Stabilization of an [exact] adjoint model implies adding extra terms for the purpose of moving the eigenvalues of the TL/adjoint matrices inside the unit circle.

Response:

We thank the reviewer’s advice on using the terminology “an approximate adjoint of VP sea ice dynamic” and for suggesting the studies of Mehlmann & Richter (2017) and Panteleev et al (2021). In the revised manuscript, **we have modified the terminology “an exact adjoint of a VP sea ice rheology” to “an approximated adjoint of a VP sea ice rheology.”**

Besides the “violation of the PSD property by the system solved by the LSOR method”, we explain more about the instabilities of the adjoint model generated based on the discrete ocean and sea ice models.

In the atmosphere and ocean communities, previous studies have pointed out several reasons for the instabilities of the adjoint model:

1) conflicts with Courant-Friedrichs-Lewy (CFL) conditions (Zhu and Kamachi, 2000) due to the discrete schemes; In this case, we have to reduce the model time step.

2) linearizing the strong nonlinear processes such as KPP mixing parameterization, sea ice rheology, etc. Linearizing these processes will lead to a much faster exponential error growth (or eigenvalues) than the nonlinear model; We usually eliminate the dependence of the strong nonlinear parameterization on the model state (e.g., the viscosities on the deformation rates in our model).

3) integrate the adjoint beyond the predictability limits of the nonlinear system; This

problem can be solved by adding additional terms to move the eigenvalues of adjoint matrices inside the unit circle, as the reviewer suggested.

As mentioned in the manuscript, the ocean and sea ice model's adjoint has suffered from instability for a long time when we include the adjoint of sea ice rheology. Toyoda et al. (2019) suggested that "the dependence of the viscosities on the deformation rates" results in strong nonlinearities (the second cause mentioned above), and eliminating their dependence could remove the abnormal positive engine value of the adjoint model. Therefore, we examine whether it works in a VP rheology and test the impacts on estimating the Arctic ocean and sea ice state.

c) Line 161: does "global mean" imply spatial *and annual* average in this context?

Response:

"global mean" indicates the average over the model domain at the current model time step. We can only use the mean adjoint sensitivities of sea ice velocity when applying this filtering process at the present step.

d) Line 166, 172 etc: "along the SIEs" SIE should be explained. If it stands for "sea ice extent", it is better to use "along the sea ice margins" (SIMs). Also, SIT and SIC should also be explained after their first appearance in the text.

Response:

We thank the reviewer's comments. We have revised the related abbreviations in the context.

Section 3.1:

a) Line 189: provide details on "could not be further reduced": does M1QN3 stall at a certain step magnitude during the line search? What gradient reduction was achieved in both the FD and VP cases at the convergence?

Response:

We have provided more details on the optimization "In adjoint-FD and adjoint-VP, the optimizations stall at iterations 13 and 32, and the further cost function reductions at the last two successive iterations are 0.7% and 0.2% of the total cost, respectively. After the optimizations, the total cost and norms of the gradients are reduced by 32.3% and 59.2% in adjoint-FD and 40.2% and (89.3%) in adjoint-VP." (L215-218)

b) Lines 196-197: replacing "adjoint of the full sea ice dynamics" by "approximate adjoint of the sea ice rheology" would be more appropriate.

Response:

We thank the reviewer's advice. We have revised "adjoint of the full sea ice dynamics" to "approximate adjoint of the sea ice rheology" throughout the manuscript.

Section 3.2:

a) Line 222: why sea ice in the Atlantic sector exhibits stronger nonlinearity? Explain.

Response:

We thank the reviewer's question. This statement, "where sea-ice shows strong nonlinearity and the tangent linear model could only capture part of the sea-ice changes (Appendix B in Lyu et al., 2021a).” comes from our previous research.

To evaluate how well the tangent linear model could predict the nonlinear errors, we performed three model simulations: 1) $F(C_{atm}(t))$ with default atmosphere forcing; 2) add perturbations $+\Delta C_{atm}(t)$ to atmosphere forcing $F(C_{atm}(t)+\Delta C_{atm}(t))$; 3) add perturbations $-\Delta C_{atm}(t)$ to atmosphere forcing $F(C_{atm}(t)-\Delta C_{atm}(t))$. The three model simulations are integrated for one year, and we evaluated the linear and higher-order components of the responses following the Taylor expansions:

$$\Delta F_1 = F(C_{atm}(t)+\Delta C_{atm}(t)) - F(C_{atm}(t)) = \frac{\partial F}{\partial C}\cdot \Delta C_{atm}(t) + o(\Delta C_{atm}(t)) \quad \text{(A1)}$$

$$\Delta F_2 = F(C_{atm}(t)-\Delta C_{atm}(t)) - F(C_{atm}(t)) = -\frac{\partial F}{\partial C}\cdot \Delta C_{atm}(t) + o(-\Delta C_{atm}(t)) \quad \text{(A2)}$$

where the linear and nonlinear error components are obtained by $\frac{1}{2}(\Delta F_1 - \Delta F_2)$ and $\frac{1}{2}(\Delta F_1 + \Delta F_2)$, respectively. By integrating the tangent linear model with the same atmospheric forcing perturbations $\Delta C_{atm}(t)$, we get the error evolution for the model variables.

Figure R1 reveals that much stronger high-order signals exist in the Atlantic sector in both the melting season (Figure R1b) and the freezing season (Figure R1e). The tangent linear model could represent the linear signal well in the melting season (Figure R1a, c) but its accuracy is much lower in the freezing season (Figure R1d,f). However, the causes for these differences need to be further investigated.

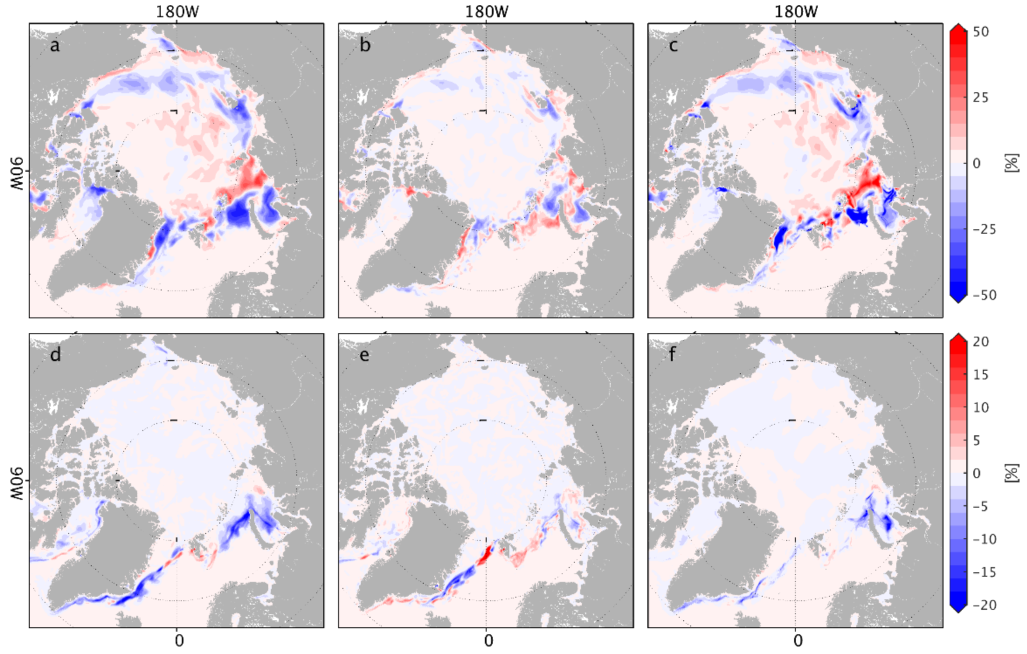

Figure R1. Linear (a) and nonlinear (b) components of the SIC changes averaged over May-June computed based on equations (A1)-(A2). Panel (c) is SIC changes predicted by the approximated tangent linear model in May-June. Panels (d)-(e) are the same as panels (a)-(c), except that they are averaged over November-December. Adopted from Lyu et al. (2021a).

b) Figure 4d: why there is a gap from May to October? No satellite data?

Response:
From April 15 to October 15, the satellite cannot observe sea ice thickness.

Section 6:
I would suggest to add a few computational quantities related to the improvement in the approximation of the adjoint model, such as the comparison of the cost function gradient decline curves with iterations, extra computational time (needed to execute 32 instead of 13 iterations), and extra CPU time/memory requirements related to the necessity of memorizing background viscosities and computing extra terms in the updated adjoint model. I would also add more discussion on the issues of instability and differentiability of the linearized VP rheology, and on the utility of the variational methods in optimizing ice conditions in general, especially in view of alternative rheologies (e.g. Mohr-Coulomb, elasto-brittle) based on more sophisticated models parameterizing sub-grid ice dynamics.

Response:

We thank the reviewer's comments here. **We have added the cost and gradients reductions in adjoint-FD and adjoint-VP in L215-218. Including the adjoint of sea ice rheology, we note that the adjoint model requires ~1.2 times of using the adjoint of a free-drift sea ice model (L187-188).** Further, we need to memorize ice velocity during the iterative solving processes, and the other variables are recomputed. For our configurations (480×416 horizontal grids ×50 levels) with a one-year assimilation window, one iteration (forward+adjoint run) requires 24 hour*208 CPU times.

In this study and our previous study (Lyu et al.,2021a, b), we have tested the performance of the adjoint method in optimizing the initial and atmosphere conditions. In Section 6, we further discuss the development of optimizing model states and uncertain parameters based on the adjoint model. Once the forward sea ice model is developed with other more sophisticated sea ice rheologies, we believe the major problem is how to approximate (stabilize) the adjoint of sea ice rheology. However, since we are not sure about the nonlinearity and computational requirements of more sophisticated sea ice rheologies, we refrain from further discussion on the choice of alternative sea ice rheologies.

TECHNICAL CORRECTIONS:
There are numerous grammar and stylistic errors that should be corrected: To name a few:
Response: We thank the reviewer's comments and have corrected the manuscript's grammar and stylistic errors.
L 47 and after: change hereafter to hereinafter.
Response: we have changed "hereafter" to "hereinafter" throughout the manuscript (e.g., L187-

188).

L 61: “may overestimate”, also change “towards” to “with respect to”
Response: We have rephrased these words.

L174: change “contradicts the impacts” to “resist [or oppose] the impact”
Response: we have changed the word “contradicts” to “oppose” (L207).

L214: add “in Fig. 3d” after RMSEs
Response: We added “in Fig. 3d” after RMSEs (L243).

L222: change “exist” to “persist”
Response: we have changed “exist” to “persist” (L252).

L226: unclear which “three simulations”, change “may related” to “may be related”
Response:
The three simulations represent “the control run” and “the two assimilation runs”. We have changed the words “three simulations” to “the control run, and the two assimilation runs”. We also changed “may related” to “may be related” (L256-257).

L227-228: change to “threshold thickness… which triggers sea ice formation in open water”
Response: We thank the reviewer’s comments. We have changed “demarcation” to “threshold” (L257).

L236 remove “larger”; change “constant” to “spatially uniform”

Response: we have removed “larger”; In this study, we used spatial-varying ice thickness errors. For accuracy, we have deleted these words.

L293 change “which” to “but
Response: We thank the reviewer’s comment. We have changed “which” to “but” (L330).

Figure 8: Color bar units? Both m/s and degrees Celsius? Comment in the caption.
Response: Units in the color bars are dimensionless since the adjustments are normalized by their prior uncertainties. We have labeled the color bar with “dimensionless” and commented in the Figure 8 caption.

L315: “By replacing [long list of variables] with their values and estimate [] contributions to cost reductions.” Incomplete sentence. Split into several, clarify what you meant.
Response:
We thank the reviewer’s suggestion. We have rewritten these words to “By replacing the adjusted initial temperature and salinity, wind vectors, 2-m air temperature, and the remaining control variables with NCEP/RA1 datasets, we integrate the model and estimate their contributions to the total cost reductions and individual components” (L380-383). We test the relative contributions of individual control variables to cost function reduction.

L319: variables

Response:

We thank the reviewer for pointing out the mistakes and have deleted the duplicated "variables" (L385).

L323 "relies", "the wind vectors", "the temperature and salinity".

Response: We have corrected these mistakes here (L388).

L326-330: Appears repetitive of the previous paragraph. Reformulate more clearly.

Response:

We have rewritten the paragraph to "Overall, adjoint-FD attributes model-data misfits more on 2-m air temperature than adjoint-VP, resulting in over-adjustments of 2-m air temperature. By using an approximated adjoint of the sea ice rheology (adjoint-VP), we reduce the model-data misfits by slightly adjusting the control variables. This leads to the conclusion that the large 2-m air temperature adjustments in adjoint-FD are likely an overcompensation for wind errors that could not be corrected appropriately because of large errors in the respective cost function gradients." (L391-395).

L362: move "thinner" after "sea ice"
L363" "while sea ice in the Eurasian Basin and central Arctic is thicker by 0.4m"
L367: remove "thickened"

Responses: we thank the reviewer's corrections and have rewritten the paragraphs and corrected the related errors (L408-420).

L420: change "seems" to "is evident"

Response: we have changed the words "seems" to "is evident" (L484).

L421: "is less efficient to reduce" to "are less efficient in reducing"

Response: We thank the reviewer's comment and have changed the words to "less efficiently reduce " (L484).

L429-430: change "which is melted" to "which destroys sea ice by"

Response: We thank the reviewer's correction, and we have changed "which is melted" to "which destroys sea ice by" (L489).

L431: This statement looks trivial (it would be surprising if the effect were opposite)

Response:

We thank the reviewer's comments. **Including the adjoint of sea ice rheology is deemed to improve the accuracy of the adjoint ocean-sea ice model if we have no further modification on the adjoint model.** However, we further enhanced viscosity and diffusivity (or added additional viscosity and diffusivity) in the adjoint model to stabilize the adjoint model beyond the predictability limit. It is unclear whether the enhanced viscosity and diffusion in the adjoint model could damp out the signals introduced by the adjoint of sea ice rheology. This study confirms that including this approximated sea ice rheology improves the ocean and sea ice estimation by adjusting the control variable slightly.

Reference:

[1] Lyu, G., Koehl, A., Serra, N., and Stammer, D.: Assessing the current and future Arctic Ocean observing system with observing system simulating experiments, Quarterly Journal of the Royal Meteorological Society, 147, 2670-2690, https://doi.org/10.1002/qj.4044, 2021a.

Lyu, G., Koehl, A., Serra, N., Stammer, D., and Xie, J.: Arctic ocean–sea ice reanalysis for the period 2007–2016 using the adjoint method, Quarterly Journal of the Royal Meteorological Society, 147, 1908-1929, https://doi.org/10.1002/qj.4002, 2021b.

Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthélemy, A., Bricaud, C., Carton, J., Fučkar, N., Garric, G., Iovino, D., Kauker, F., Korhonen, M., Lien, V. S., Marnela, M., Massonnet, F., Mignac, D., Peterson, K. A., Sadikni, R., Shi, L., Tietsche, S., Toyoda, T., Xie, J., and Zhang, Z.: An assessment of ten ocean reanalyses in the polar regions, Climate Dynamics, 52, 1613-1650, 10.1007/s00382-018-4242-z, 2019.

Zhu, J. and Kamachi, M.: The role of time step size in numerical stability of tangent linear models, Monthly Weather Review, 128, 1562-1572, 2000.