Response to reviewer 1

We thank the referee for their thorough and constructive review of the manuscript “Nonlinear dynamics of time-variable slope circulation.” The detailed comments and suggestions have been helpful for clarifying the presentation and strengthening the manuscript. Below, we respond point by point to the individual comments.

In addition to the comments reproduced here, we have gone through remarks in the annotated PDF and will revise the manuscript accordingly. We hope that these changes address the referee’s concerns and improve the clarity of the paper.

Comment:

I found the introduction confusing. It's not very clear what past studies found, what has been explained, what hasn't, and what the present authors' hypotheses are. The flow of logic is too circuitous and hard to follow. A more straightforward logic is desired.

For example, the paragraph on lines 59–72 forms a nice hypothesis (retrograde currents arresting topographic Rossby waves, resulting in prograde form stress and hence the asymmetry between the prograde and retrograde directions). So I thought that this is the hypothesis the paper is going to test and that this is the end of "the problem statement". But, then a discussion of Holloway's mechanism starts AGAIN . . .

. . . As I read this paragraph, it dawned on me that a discussion of another mechanism of generating prograde form stress HAS JUST STARTED.

The problem in this particular instance is that the authors have cast the problem into "steady forcing vs variable forcing". But this contrast (steady vs variable) is hardly essential. In Holloway's mechanism, you still need eddies, which form variability. The real issue is that there are two mechanisms.

So, the whole discussion of the authors' hypothesis should start from a statement that there are potentially two mechanisms to generate retrograde stress. One is XXX and the other is YYY. This should be the structure of the discussion . . . or at least this is one way to convey the idea smoothly to the reader.

. . . But then, at the very end of Section 5 (Discussion, lines 558–560) it becomes clear that the 2nd hypothesis isn't tested in the present study.

. . . As (I hope) you can see from this difficulty I had, the logic of the discussion becomes clearer, only when you review what you have read and re-organize the ideas in your head.

In general, ideas and arguments should be arranged to form a straightforward narrative, reducing redundancy as much as possible, so that the reader can understand the logic easily just by reading once from top to bottom.

I've inserted comments to the manuscript PDF file, which I'm attaching to this review report, as well as wrote other comments below. They, especially those in the manuscript, haven't been edited much. Some of the questions I ask in earlier comments are resolved later in the text, and also some of my comments include a lot of redundancy. I'm sorry about that but at the same time I hope this helps the authors see potential difficulty some of their future readers will have.

Response:
We agree that the Introduction section was confusing, especially since the reviewer got the impression that we were testing two contrasting hypotheses. We have therefore made significant modifications to this. The new structure goes roughly as follows:

1. Mention of the dynamics of gyres in the Arctic Ocean and a presentation of the findings of Sjur et al. and Johnston et al. - ending up with the observed discrepancies from existing linear theories.
2. A discussion of how the 'background' cyclonic/prograde offset seems to be consistent with 'Neptune' forcing – and vorticity fluxes across constant-H contours.
3. Limitations of such interpretation and an introduction to form stress, arising when analysis is done by integration along contours that contain depth variations. The asymmetric form stress response to prograde and retrograde flows, and how time-variable forcing is therefore needed for form stress to generate a net prograde flow.
4. Specification of the aim: to study the dynamics that may lie behind Sjur et al.'s results, focusing particularly on variable wind forcing – and how the non-linear response manifests when the dynamics are analyzed along integration paths without and with depth variations. Also motivate the choice behind studying two forcing time scales.
5. A listing of the structure of the rest of the paper.

Comment:
There are still things that I ended up not understanding:

- *Past studies are sometimes alluded for steady-state response, but it's not clear from the text what the past studies have shown about the steady-state response of the system the present authors deal with (barotropic flow along a corrugated continental slope on an f plane). Can the steady state response be regarded as the limit of slowly-varying forcing and the mechanism of arrested Rossby waves remain the same?*

Response:
The introduction has been considerably modified, addressing this point. Specifically, we state that the long forcing period can be regarded as a near-steady situation, where arrested Rossby wave theory applies.

Comment:

- *The Neptune effect (barotropic flow induced by mesoscale eddies impinging on bottom slopes) is mentioned as a second potential mechanism to explain the prograde bias, but this hypothesis isn't tested as the barotropic model doesn't produce mesoscale eddies →*

→ Much later in the text (lines 558–561) this issue is taken up as a missing feature in the present study. That means, I was misled to think that this was a second hypothesis that the present paper would test. I hope the introduction will be revised so that this misunderstanding wouldn't happen.

Response:
The introduction has been considerably modified, addressing this point as well.

Comment:

- *How does the amplitude of corrugation enter the discussion? What determines the minimum amplitude for the arrested Rossby wave mechanism to work? How does the amplitude control nonlinearity? Does an increase in the amplitude monotonically increase form stress?*

Response:
Thank you for raising these points about the role of corrugation amplitude.

From the steady-state QG analysis in Bai et al (2021), we expect the amplitude of arrested waves to depend linearly on corrugation amplitude. So both form stress and nonlinear vorticity advection should depend on the square of the corrugation amplitude.
Near the arrest condition, the form stress is very sensitive to the mean velocity U. Here, small changes in U can balance relatively large changes in surface stress, which underlies the saturation behavior we find. Whether such saturation persists as the surface stress is increased further depends on whether the maximum attainable form stress at arrest is large enough to balance the wind input, and this in turn is expected to depend on the corrugation amplitude. For small amplitudes, the form stress may not be strong enough to sustain a pronounced saturation regime, whereas for larger amplitudes it can.
Thus, in principle, the basic Rossby wave arrest mechanism itself should not be sensitive to corrugation amplitude; but whether form stress can sustain a persistent saturation as the surface stress increases is expected to rely on the corrugation amplitude being sufficiently large.
A possible complicating factor is that there are different arrest conditions connected to different wave modes (as we see in Fig. 9, different modes have different arrest speeds). This implies that as the corrugation height changes for the same forcing strength, saturation might not happen at mode 2, but instead at mode 1.

We have done additional simulations with varying corrugation amplitudes and are currently working on incorporating these results into the manuscript.

Comment:

- *What happens in the limit of R → 0 when there is corrugation?*

Response:
This seemingly innocent question is actually quite complicated, and beyond the scope of the present paper. However, the short answer is that the kinetic energy of flows tends to increase with decreasing R; but depending on several factors it may be the mean flow energy or the wave/eddy energy that increases.

We outline some of the physics here. We consider the equation for the domain mean x-component velocity (say U) which in the limit of small depth variations (relative to the mean depth H0) and with R=0 (and no other form of friction/dissipation) can be written as

H0 dU/dt = T + F , (1)

where T the is the wind stress and F the domain-mean form stress. We consider a wind forcing that varies sinusoidally in time with frequency omega. Without corrugations F=0, and Eq. (1) predicts the velocity amplitude

|U| = T_max/(H0 omega) (2)

In the limit of steady forcing (omega = 0), |U| approaches infinity.

The question is what happens in the case with corrugations where F may be non zero. For a prograde flow F is zero when R=0 (see e.g Bai et al. 2021), and Eq. (2) gives an estimate on the amplitude of U.

For retrograde flow over corrugations with a single wave number k, F becomes large when the flow speed is near the phase speed of the topographic wave. If the flow is steady (omega =0) the phase speed is U_T= beta_T k^2, where beta_T is the topographic beta. (If the flow U oscillates the problem becomes considerably more challenging, but if omega << Uk the "arrest" speed is approximately given by U_T). If the retrograde flow is accelerated by the wind stress to a speed very close to U_T, topographic waves are excited resonantly and their amplitudes approximately grow linearly with time. Here, F is non-zero and nearly balances the wind stress, which limits the mean flow speed to around U_T.

However, the wind stress work (T x U_T) pumps energy into the wave field which eventually becomes unstable and turbulent. The studies by e.g., Constantinou and Young (2017) and Constantinou (2018) show that when the eddy kinetic energy grow beyond a threshold the flow jumps to a new equilibrium state where U > U_T. In this state the wave/eddy amplitude is small and the form drag would be zero for an inviscid flow. Here Eq. (2) would predict the flow amplitude.

We are currently debating how much of this to incorporate in a revised manuscript since, as said, it is really beyond the scope of the current study. Note, for example, that we are unable to reach the non-dissipative limit. We could let $R \rightarrow 0$ but would be unable to control numerical viscosity with the present code.

Comment:

- *The along-shore (x direction) structure of the form stress and Rossby waves is not discussed. If the form stress is perfectly periodic at the wavelength of the corrugation, the authors' re-entrant (periodic) domain is perfectly fine, but is this clean structure expected to hold for all realistic parameter range? I hope at least some words will be added about this issue.*

Response:

In our simulations the along-shore structure of both the Rossby wave and the associated form stress is indeed perfectly periodic at the wavelength of the bottom corrugation. This is evident in the residual perturbation field (show for selected x-values in Fig. 11), where the cross-shore profile of pressure repeats at the wavelength of the corrugations. We will clarify this explicitly in the text by noting that the pressure structure is periodic at the topographic wavelength.

Comment:

It's not clear to me what the role of the linear part of bottom form stress is: H ∇ϕ. Under the linear limit (weak forcing τ_0), is the form stress term always negligible even if R is very small? In a stratified ocean, generation of lee waves can exert net stress on the fluid and this is a linear mechanism, I think, but I'm not familiar with the barotropic situation .

Response:

What happens in the limit of small R is related to the earlier question regarding the limit R -> 0. For a topography of the form h0 sin(kx) and in the limit of small U far from resonance with the bathymetry, the topographic form stress is proportional to (h0k)^2 RU; see Bai et al. (2021). So in this regime the form stress is linear in U, and symmetric for prograde and retrograde flow. This does not mean that form stress is negligible but symmetric regarding forcing direction. Outside this far-from-resonance regime, the form stress is expected to become asymmetric.

In the revised manuscript, even though we have already included references to past studies, we will add some lines of explanation of how form stress is expected to behave.

Comment:

This is definitely a minor point. I'm glad to learn about the new package Oceananigans, but the text needs to tell the reader how the package contributes to the present study. If you know basic finite differencing, you can in principle construct the same numerical model. So . . .

- *Do the author need the package? For example, does the package use sophisticated numerical methods for higher accuracy, which would be a lot of work to implement?*
- *Or, do they use the package just for convenience? . . . Since there is already a well-written package, why not just use it? . . . Is this the idea?*

- *Or, does the package include tools for analysis and that's the main reason for utilizing this package?*

Response:
We agree that in principle, we could have implemented the shallow water model ourselves. However, we chose to use one of the existing and well-tested ocean models out there, in this case Oceananigans.jl, mainly for reproducibility and convenience.

Comment:
I'm not familiar with quasi-geostrophy for barotropic flow. For a 1.5-layer model, for example, PV is $q = (\zeta + f)/D$, where $D = H + h$, where H = const is the mean layer thickness and h is the deviation from it. In the QG limit, $q \approx (\zeta - f h/H)/H$. This is the so-called QG PV.

For the barotropic case, $D = H(x,y) + \eta$. Does that mean that the QG PV is $(\zeta - f\eta/H)/H$ and the rigid-lid approximation ($\eta \ll H$) is equivalent to QG (of course assuming $\omega \ll f$)? Or, with corrugated bottom, the nonlinear momentum term tends to be large enough to break geostrophy and so that $h(x,y) \ll H_0$ is a necessary condition for QG?

But then, why not just use $q = (\zeta + f)/D \approx (\zeta + f)/H$ for PV balance? At least for the const-H line integration, this q should work. Also, is the $h \ll H_0$ approximation really necessary in the const-y integration to relate PV with bottom form stress?

In any case, I think that a more careful discussion on the necessity of the approximation that $h(x,y) \ll H_0$ is necessary.

Response:
The reviewer is correct. The QGPV for the barotropic (1-layer) model is the same as for the 1.5-layer model. In both cases we assume $f = f_0$ *(constant)*, $|\zeta| \ll f$ (small Rossby number) and $|h| \ll H_0$. In addition, we make the rigid lid approximation.

The reason we focus on a flux of QGPV rather than a flux of total shallow-water PV is simply that two terms on the right-hand side of the depth-integrated momentum equations combine—under QG scaling—to look like a flux of QGPV. This formulation is standard in the Transformed Eularian Mean (TEM) form of a momentum equation integrated around a periodic domain (this will be mentioned in the revised manuscript). One cannot manipulate these depth-integrated equations similarly to produce a term which looks like a flux of exact shallow-water PV.

Comment:
Figure 3 compares a nonlinear solution with the prediction from the linear theory. Before this, we need a comparison of a "linear" numerical solution with the theory. We can reduce the amplitude of the wind stress by a factor of 1/1000 or anything very small for the numerical model. This numerical solution should be very "linear". In this situation, the numerical simulation doesn't develop eddies (correct?) and so we don't have to worry about signal-to-noise ratio even with very week winds (because there is little "noise" in the numerical solution). This comparison would give us confidence in the linear theory, in the authors' calculation of various terms (for example, how accurately is the contour length C calculated for the numerical model?), and in the numerical solution itself (for example, because of the finite resolution, there may be some artificial form stress due to the zig-zag nature of the slope), and at the same time it would be a way to assess the impacts of the assumptions that went into the theory (rigid lid approximation, what else?) and the numerical model's deviations from the idealized situation (for example, what's the impact of the artificial boundary at y = L_y in the numerical model?).

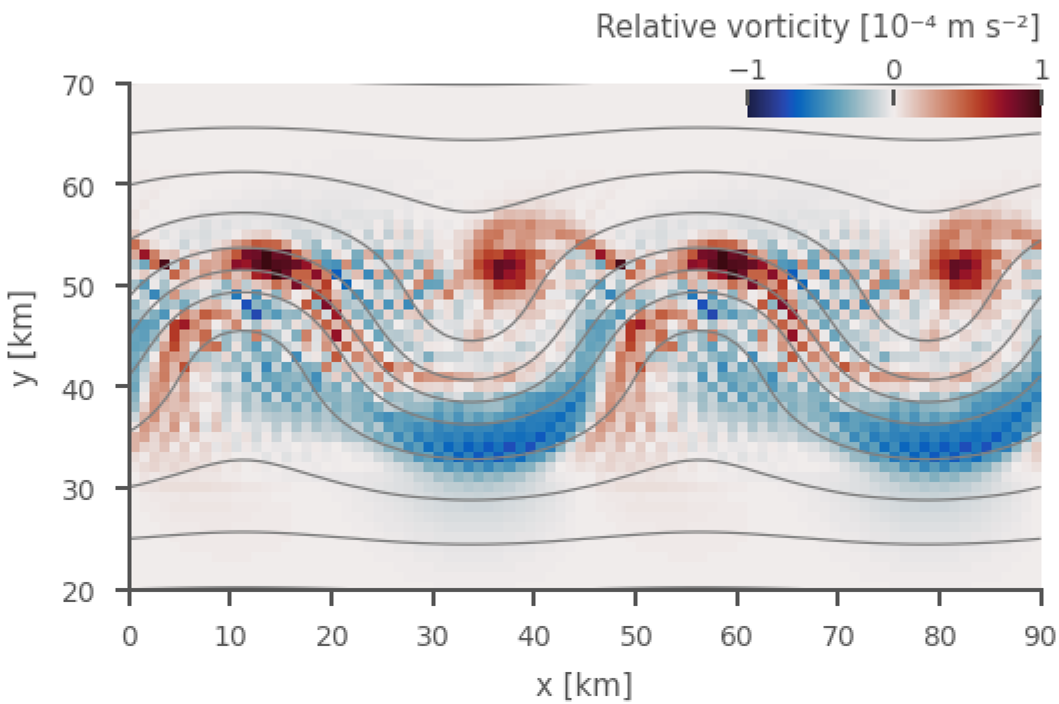
After this, we would be more confident that the deviation of the nonlinear solution from the linear theory is due to nonlinearity.

Response:
To address this comment, we perform an initial evaluation of the linear theory against a nearly linear numerical simulation. We reduce the wind-stress amplitude to 10%, which we find to produce a weak, laminar response without eddies, and we find excellent agreement between the simulated circulation and the linear estimates. This is added to the appendix, which is referred to when the linear model is introduced in the theory section.

Comment:
The spatial structure of the nonlinear term should be discussed. If it is exactly periodic in the x direction at the periodicity of the bottom corrugation, the width L_x of the domain shouldn't matter as long as it is an integral multiple of the corrugation wavelength.

In this regard, I'm wondering whether the x-width of the domain might be too restrictive. It includes only two bumps (Figure 2). Nonlinearity might "want" to develop a larger structure?

Response:
The spatial structure of the nonlinear term is indeed periodic in the x-direction at the corrugation wavelength. The supplementary animations of relative vorticity (which directly connects to the nonlinear relative-vorticity flux) show that the vorticity field repeats at the corrugation wavelength. In the highly non-linear runs, eddies are forming on scales smaller than the corrugation wavelength and form a pattern that is periodic in x. We will clarify this observation in the text when introducing the animations.
As an illustration, we include below a snapshot from the vorticity-field animation (long forcing period, t=32, which is part of the highly nonlinear phase), showing the developed nonlinear eddy field and its periodicity at the corrugation wavelength. Here, gray lines indicate isobaths:



Comment:
For color plots, please consider using color-blind friendly pallets. For example, the orange-green contrast in Figure 4 doesn't seem to be particularly friendly to some type of color-blindness. Even though one is dashed and the other is solid, the legend uses only colors. There are resources on the Internet: for example, https://davidmathlogic.com/colorblind/ . Modern scientific plotting tools usually include a selection of color-blind-friendly schemes.

Having said that, it's always a good practice to increase the distinctiveness between the lines by using symbols and various dash-types in addition to colors.

Response:
We appreciate that the reviewer highlights the importance of color-blind-friendly palletes. As it turns out, the figures in the original submission already use a color-blind friendly palette, corresponding to the colormap “Wong” in the resource cited by the reviewer.

Comment:
Please use an expression such as "friction" and "frictional stress" instead of "bottom stress" throughout the paper. The latter expression is confusing because bottom form stress is also a "bottom stress".

Response:
We will change “bottom stress” to “bottom frictional stress” throughout the text to avoid confusion.

Comment:
I found it hard to follow the discussion of momentum flux and acceleration near the end of Section 4.2. You need a schematic diagram to follow it. If the discussion were simply about relation between a single spot of flux convergence and acceleration, it would be fine, but the text talks about multiple convergence zones occurring on the "flanks" of peaks of another quantity . . . In an oral presentation, the presenter would point to specific points in the line graphs in Figure 6 . . . In reading the paper, the reader needs a similar visual guide.

Also, I'm not familiar with the relation between Rossby wave generation and form stress. I guess Rossby waves carry prograde momentum and that momentum is imparted by the solid Earth and therefore the form stress must be prograde . . . Is this what's going on?

Response:
Yes, this is, in essence, what goes on. The general relationship between Rossby wave generation and momentum fluxes is well-established, and schematically outlined in section 15.1 of Vallis (2017). We will include a reference to this in the text. We will also add a reference to Held (1983) with respect to topographic generation of Rossby waves.

Comment:

I failed to see the value of the PV discussion. The approximate equivalence of PV flux with bottom form stress is mathematically derived and the numerical solution demonstrated the accuracy of the mathematical derivation (Figure 7) . . . What have we learned from this exercise?

Perhaps there has been a debate or confusion about the form-stress view and PV-flux view and the authors want to resolve it?

Response:

Thank you for pointing this out. This was an under-communicated point which we have now emphasized quite a bit more. Existing literature has mostly worked along two seemingly separate branches, focusing on i) vorticity fluxes when integrating the dynamics along const- H contours or ii) form stress when integrating along contours with topographic variations. Bridging those two perspectives—under QG scaling—seems quite useful, and we are, frankly, quite excited by this result. As mentioned in the Conclusions, this brings form stress (that is the thickness contribution to the PV flux) also into the context of the minimum potential enstrophy arguments traditionally used to explain prograde offsets (e.g. Bretherton and Haidvogel, 1976).

Comment:

It's not clear what the comparison between 16-day and 128-day forcings shows. By design (equation 16) higher forcing frequency leads to smaller amplitude. Is the weakness of nonlinear impacts in the 16-day solution, then, all due to the smaller amplitude? What would happen to a 128-day solution whose τ_0 is reduced so that u_max is exactly the same as that of the 16-day solution at the mid depth? It would still show the weaker depth dependence [1 + (ωH/R)^2]^{−1/2} . . . what else?

→ This question will be resolved later in the text, but I suggest including some guide earlier so that the reader doesn't have to carry this question to the very end of the discussion of the main results.

Response:

Thank you for this helpful comment. We have revised the introduction to state more explicitly that we focus on two contrasting regimes: a short forcing period comparable to the frictional damping timescale, and a long forcing period that allows the system to approach a quasi–steady state.

When the results from the two focus cases (16 days and 128 days period) are first introduced, we now explicitly state that part of the difference between the two experiments is indeed the difference in circulation strength. We also add a forward reference to the section where we vary the forcing strength, where this is examined more closely.

However, we also emphasize already in the discussion of the two scatterplots that amplitude is not the only relevant distinction. Because the 16-day forcing period is comparable to the damping timescale, the circulation does not fully adjust before the forcing reverses, and the response behaves more like a low-pass filter. In contrast, in the 128-day case the forcing varies slowly compared with the damping, allowing the flow to reach distinct prograde and retrograde states.

Comment:
The construction of Figure 9 is explained solely in words (lines 450–460) and is hard to follow. Also, the reason why this procedure will tease out Rossby waves isn't clear. Perhaps write the procedure and the assumptions both in math.

Response:
We have added equations describing how the pressure perturbations are calculated. We have also elaborated on why this procedure reveals the Rossby waves.

Comment:
This is just a suggestion: Since most of the important conclusions are derived by analyzing the ω = 2π/(128 d) solution, it's a bit distracting that the text keeps comparing it with the ω = 2π/(16 d) solution. The meaning of this comparison becomes clearer only when the systematic amplitude sensitivity is presented in Section 4.5.

So, I think the logical structure of the Results would be simpler if the numerical experiments are ordered this way:

1. *ω = 2π/(128 d) only*

1.1) "linear" . . . no corrugation (or extremely small τ_0 ?);

1.2) "weakly nonlinear" . . . smallish τ_0 (so that this is similar to (2.1) below);

1.3) "fully nonlinear" . . . τ_0 = standard size.

2. *Sensitivity solutions*

2.1) same as (1.3) but $\omega = 2\pi/(16\ d)$.

2.2) sensitivity to τ_0

2.3) impacts of corrugation.

Response:
We appreciate the suggestion and agree that clarifying the logical role of the two forcing periods is important. In the revised manuscript we have not reorganized the results section exactly as proposed, but we have made several changes that we hope address the underlying concern. First, we have revised the introduction to more clearly motivate the use of two contrasting forcing periods. Second, we now include a linear weak-forcing experiment (new B1). Finally, we have moved the impact of corrugations to follow the section on the PV perspective.

Comment:
Figure 8 is very informative. And this makes me wonder whether resonance is possible or not. Because the computational domain is re-entrant (cyclic in x) or the actual Arctic Ocean has closed-f/H basins, wind forcing should include frequencies and wavenumbers that match the cycle travel times and wavenumbers of topographic Rossby waves.

Response:
If we understand the reviewer correctly, the reviewer asks whether the resonance between arrested Rossby waves and bathymetry we observe in idealized simulations is also possible in the actual Arctic Ocean. As we touch upon in the discussion, real topography contains a wide spectrum of wavelengths. On this basis, we consider it possible that similar arrest mechanisms could operate in the real Arctic. We will also stress this point more in the conclusion of the revised manuscript.

Comment:
It's not clear to me how the prograde-retrograde asymmetry in Figure 3b is explained, where the currents aren't fast enough to arrest Rossby waves. Is it still due to bottom form stress?

I suppose that this lack of discussion on this point is because the mechanism has already been explained by past studies and it's entirely possible that I missed such discussion in the present manuscript. But it would still be nice if the mechanism is briefly explained in *the discussion of Figures 3 and 4 before entering the discussion of "saturation" and also in the discussion of Figure 11a, where corrugation is removed. (I suppose the prograde-retrograde asymmetry is gone when there is no corrugation.)*

Response:
The reviewer is correct that the deviations from linear prediction in Fig. 3b is due to a net topographic form stress (diagnosed on y-contours) in the prograde direction. While the arrest of topographic Rossby waves resonating with the bathymetry produces strong form stress and a pronounced asymmetry, there is also an asymmetry in the magnitude of topographic form stress even for more moderate flow strength. Only for very weak flow does the form stress become approximately symmetric with respect to flow direction. This behavior has been documented in previous studies, and we now emphasize this more clearly in the introduction and theory section.