Response to referee #1:

Thank you for your comments. I have now made many changes and will submit a revised manuscript shortly to address your comments and the comments of the other two referees.

Please keep in mind that a group of referees often have different levels of expertise and conflicting opinions, sometimes forcing an author to compromise, as happened with the previous editor (I explain below).

Detailed responses to your primary comments are provided below:

1. *Three pages (pp 278-280) are needlessly devoted to the derivation of the wave-induced mean flow. ... and the derivation is contorted besides. Equation (26) confusingly gives* $\bar{u}$ *as if it is order* α^0*, but is then said to scale as* α_0 *in (32) and finally it is revealed that* $\bar{u}$ *is in fact order* α^2 *in (37). This should have been self-evident from the start, if the author has started with the horizontal momentum equation as done, for example by Dosser and Sutherland (2001), which had previously been cited by the author but is not referenced here.*

   In a previous submission I had not included a derivation of the mean flow, however another reviewer insisted that I had to demonstrate convincingly that the mean flow was order α^2, as he was convinced it was not. As a result I added a complete derivation of the mean flow, which did indeed satisfy this previous reveiwer. Furthermore, the derivation had to connect with my choice of variables, accounting for the differences from previous derivations. I never claimed that the mean flow derivation was fundamentally different. I have now removed much of the derivation and refer to previous work.

2. *He says he follows Acheson (1976) and Scinocca and Shepherd (1992), but does nothing of the kind. the former gives a contorted derivation based on wave action and the latter gives a derivation based on Hamiltonian fluid dynamics. As I stated before, It's derivation is well-established (Section 3.4.5 of the textbook "Internal Gravity Waves" by Sutherland (2010)) ...*

   Acheson (1976) seems to be the first to find an explicit expression for the wave-induced mean flow for internal waves, and my previous submission merely intended to credit him for that. Scinocca and Shepherd (1992) generalized the idea using a Hamiltonian approach and again I was merely mentioned this significant advance. The specific approach I used was in Grimshaw and McHugh (2013), which in turn credits Sutherland (2006). I have now reworded this paragraph and removed the reference to Acheson (1976) and Scinocca and Shepherd (1992).

3. *Rather than starting with horizontal momentum to derive the mean horizontal flow, he starts with vorticity (27-28) and in (29-30) erroneously switches to a hybrid co-ordinate system involving z and $\zeta = \epsilon z$. If he is performing a multi-scale analysis he does not say so much. He then converts from vorticity to buoyancy and then gives results in (40-41) below in terms of I and R, which work with vertical velocity, not buoyancy. (Incidentally, it is only implied that R in (18) represents vertical velocity).*

   As I mentioned above, I had added a complete derivation of the mean flow to satisfy a different reviewer. You however complained that this derivation was totally unnecessary, as you have repeated again. However I could not remove it completely and still satisfy this other reviewer, so I reduced the derivation to the compromise in my previous submission, which I felt was just sufficient to demonstrate that the mean flow was order α^2. This section is now much reduced.

4. *...it is evident that there is a theoretical error in the author's paper. In (15) the author defines a translating horizontal co-ordinate, $\xi = x - c_p t$, that moves horizontally at the horizontal phase speed. And yet, the author arrives at (38), which is the mean flow found by others in a stationary frame. If the author's theory was rigorous, the two results should have differed by a factor $-c_p$.*

   My results are correct and completely rigorous. In general, the inclusion of a uniform background flow U would not change the results, as a coordinate system moving with U would delete U from the system. The waves merely behave as if there was no background flow. U would reappear in the end only if the absolute mean velocity in the non-moving system was desired. In my case U is merely the phase speed c_p. Yes, to obtain absolute velocities in a non-moving coordinate system I would have to add c_p to the total horizontal velocity. But this does not change the wave-induced mean flow, which is correct.

5. *...But the major theoretical flaw, which I mentioned in my previous two*

reviews and which persists in this paper, is that the interface conditions are incorrect. The conditions (61-62) have been truncated at order α^2 and so neglect the weakly nonlinear effect of the wave-induced mean flow acting back upon the waves. But the NLS equations (50), (53), and (54) are order α^3 . . . and so do include this effect.

I will first state unequivocally that the interface conditions are *correct*. I have truncated terms only when they do not contribute to the final amplitude equations. When the contribution from the mean flow terms in the interfacial conditions are formally included, these terms are an order-of-magnitude smaller than other nonlinear terms. Hence the final set of equations can only have the linear terms in the interfacial conditions. I had explained this in my original submission to NPG, and now I have expanded upon it in the forthcoming revision. This means that the variation from linear theory in the reflection and transmission coefficients only happens for large amplitude waves. It cannot be meaningfully included in the theory at this order. This is a fact of this theory, not a flaw. This was also the case in Grimshaw and McHugh (2013).

The reason this feature does not appear at this order can be traced to the fact that the waves in question are not interfacial waves travelling along this interface, which if they were would cause to first set of terms to balance automatically and drop out. Since this is not the case, then the nonlinear interfacial terms that are two orders away do not survive.

6. *Despite raising this concern in past reviews, most of the figures continue to show shapshots at three times instead of time series and contour plots. I think the numerical results of the author are incorrect and time series might convince me otherwise. If the author made contour plots and time series of the control simulations with uniform stratification (Figs 3,7,13), as best as I can tell they would not look like the corresponding plots given by Sutherland (2006) regarding modulational stability and instability.*

   My numerical results are correct. In fact I show analytically that the case with constant N and no dispersion must have a wave packet shape that does not change, and then show that the numerical approach gives this result perfectly. Furthermore the numerical results are insensitive to changes in time step and grid spacing. There is no basis to suggest that the numerical results are incorrect.

I have now added several contour plots as you suggest. For the constant N case they do indeed match the plots of Sutherland (2006).

7. *Despite comments in previous reviews, the author continues to cite Whitham (1974), who said nothing specifically about modulational stability of internal waves as the text implies.*

   Whitham has a general discussion concerning modulations of waves which does include modulational stability, although he does not discuss internal waves specifically. I have now removed the reference to Whitham and changed this paragraph.

8. *In my earlier reviews I suggested the author refer to Sutherland (1996) which examined transmission and reflection of small and moderately large wavepackets from strong to weak stratification. ... The transmission and reflection coefficients were clearly shown to be a function of amplitude through the fully nonlinear simulations presented in that study.*

   While my present work has revealed several important issues, (particularly the development of increasingly smaller scale components in the mean flow), it is clear that *nonlinear* wave reflection does not occur at this order, hence that is not the main theme of the results, which is why I do not compare my results with those of Sutherland (1996).

9. *the wavepacket in Fig 13a doesn't look like it has a truncated cosine amplitude envelope at the outset.*

   The envelope is created at the bottom of the computational domain, and is a raised cosine. For the parameter ranges of figure 13a, the envelope evolves rapidly into the shape shown in figure 13a. I have added text to explain further how I create the wave packet, as mentioned below.

10. *The author insists that ϵ and α (the nondimensional amplitude) must be equal. Even if the assumption constitutes part of the rigorous derivation of the NLS equation, the equation itself describes small amplitude dispersive waves simply by initializing with small amplitude waves, making the weakly nonlinear terms, $|I|^2 I$ etc, negligibly small. If the author's*

results truly require the wavepacket width to be set by the amplitude instead of being independent quantities, then the application of the results to realistic circumstances are drawn into question.

The two parameters are indeed independent, however if they are of different orders then the overall results become less important. Keeping them different but of the same order does not reveal any different results, while setting them equal does allow a simpler form of the NLS equation. I have added some comments in section 5 to clarify this.

11. *It is incorrect to say n_1 is the vertical wavenumber: as it is given in (16), $-n_1$ is the vertical wavenumber.*

    I have chosen to have wavenumbers and frequencies be positive quantities, as is common in wave theory, and then insert the negative sign in the phase function where necessary. I now made this clear in section 3 to clarify.

12. *What is the practical reason for moving in a frame with the horizontal phase speed? It makes all interpretation of the wave-induced mean flow in text and figures confusing.*

    Phase functions, such as $\xi = x - c_p t$, are widely used in wave theory, such as Tabaei and Akylas (2007) and Shrira (1981). In slowly-varying theory as here the waves get filtered out, leaving only the dynamics of the amplitude function. Phase functions allow the filtering over the phase rather than just time or space. A similar theory can be developed without using ξ and the basic results are the same, but here the analysis is simpler with the inclusion of this phase function, particularly since there are multiple wave packets present. The wave-induced mean flow is unchanged with the use of ξ, as I have explained above.

13. *The origin of the rescaling given by U in (68) is unclear, if it is not incorrect. For example, the textbook of Sutherland (2010) suggests the wave-induced mean flow should increase as* $\sec\Theta = \sqrt{k^2 + n_1^2}/k$.

    Defining U as I have done is a matter of convenience only, really not a 'rescaling'. Using U allows me to show the mean flow on the same scale for different parameter values. I have reworded section 6 to clarify this. My definition of U does agree with result of Sutherland, but I have made U dimensionless accounting for the different form.

14. *It makes no sense to have $\mathcal{R} = -1$ in equation (73).*

    My wording here was poorly chosen, and I have now changed it.

15. *Also, I have great difficulty backing out physical variables from the control variables given by the author. The wavepacket width is supposed to be represented by ϵ, but its definition in (67) uses Q in of of ϵ for a proxy of wavepacket width.*

    I have deleted Q now and use ϵ directly in (67).

16. *Then, when the author presents results in Figure 3, etc, no mention is given of the value of ϵ or Q, or when the simulation began for that matter.*

    I will include the value of ϵ in the caption of each figure in the forthcoming revision. I have also added a comment in the text that explains further the initiation of the wave packet.