Response to referee #3:

Thank you for your comments. My manuscript has now been significantly improved as a result of your comments, the comments of the other referees, and also my own additional work.

I have made all the minor corrections that you mention to a forthcoming revised version. Detailed responses to your primary questions and comments are provided below:

1. *The relationship to Grimshaw and McHugh (2013) could be indicated more often. How do various equations compare: e.g. (42) here and (35) of GM13. At what stage does the derivation deviate from GM13? The "unknown functions" determined by GM13 - are they relevant here or consistent with the results?*

   The mean flow here and in Grimshaw and McHugh (2013) is indeed very similar, and found the same way. The other referees had the same comment and one insists that I reduce the derivation of the mean flow. As a result section 4 is much reduced, refering instead to Grimshaw and McHugh (2013) and other previous publications. Furthermore, I have added working to indicate when the results are different than GM13. The unknown functions are problematic when using the DJL equation, or other such integral result. Here the primitive equations are employed and it is clear that constants of integration must be zero so that the mean flow without waves is zero. Hence these unknown functions are not relevant here.

2. *It is noted that ϵ can be scaled away, into the time and height coordinates. Is there a precendent for this, e.g., GM13 or other weakly nonlinear internal wave studies? It is stated that there are 5 waves in the wave packet, and this sets ϵ. What is the definition of epsilon in terms of the number of waves, and what is the value of ϵ used in the examples?*

   Tabaei and Akylas (2007) also noted that the small parameter can be scaled away for a single layer case that is similar to here. I now mention this in the text. I have also included the definition of ϵ in terms of the number of waves in the text. Also I include now the value of ϵ for each example in the caption of the figures.

3. *What is the sign convention for k and n, starting in (16)?*

   All wavenumbers (k, n_1, n_2) are positive numbers, while negative signs are included in the phase functions where needed. I clarify this now in the text.

4. *(24)-(25) should be referenced to (16) in GM13.*

   I have now included this reference.

5. *In at least one of the examples, it would be useful to give the associated dimensional values relevant to the atmosphere: i.e., the horizontal and vertical wavelengths, the vertical velocity amplitude, and the maximum wave induced mean flow. Can these values be compared with the experimental data of McHugh et al. (2008a,b)?*

   Another referee suggested that I expand the conclusions section, which I have now done. I have taken the opportunity here to add a discussion with dimensional values of my results that attempts to compare with the previous observations. The theory here is much simpler than the observed flow of course, but still the comparison seems useful.