Response to Referee # 1

**Simulation of humpback whale bubble-net feeding models**

*Bryngelson and Colonius*

We are grateful for the referee's efforts in improving the quality of this paper. We quote the comments of the referee and discuss changes made to the paper in response to these comments in the following.

*Referee #1*

The paper at hand discusses the very interesting topic of hump back whale bubble-net feeding and the modelling of the associated acoustic effects. Based on different scenarios, a detailed numerical model asses the resulting sound pressure levels within different bubble net configurations for various source, frequency and void fraction setups.

My bio-acoustic background is very limited as my field is more based in the physics of underwater acoustics. Therefore, I would like to apologize in advance if some of my questions are not appropriate, as the answer would be clear to any bio-acoustician.

In general, I found the paper and its investigations highly interesting. However, what I partly missed was the basis for the assumptions from measurements and observations, as well as the comparison with measurement data for the results or (if these are not available) a discussion what the results would mean for the feeding of the whales.

Regarding the input of the model, the following questions need to be answered in my view:

1. What dimensions of the bubble nets have been observed, why are the specific chosen parameters used here?

   The 10 m radius we use follows from the estimations of Devincent [10]. Some deviations in size have been observed, for example approximately 7.5 m radius by Hain [2]. The exact size depends on the size of the whale that creates it. Our choice of 10 m is thus not unique, but instead represents one such idealized bubble net. We have added this caviat to paragraph 1 of section II A.

2. What are the measured vocalization spectra observed during bubble net feeding?

   Whales are able to create complex and broadband waveforms [14]. We test a range of monochromatic signals as a proxy for considering these actual vocalizations. This is used to determine if specific acoustical phenomena can occur in a humpback whale generated bubble net, rather than represent the exact physical state of a specific bubble net. We now include this information in paragraph 2 of section II A.

3. What is the time scale of such a feeding operation and how will the bubble net change during that period?

   According to Devincent [10] and Hain [2], vocalizations were 45-58 s in duration, with the total process lasting less than a few minutes. Wiley [19] also observed a total dive duration of about 1 minute. The duration of these are significantly longer than the $< 1$ s required for the acoustical phenomena we observe to occur. Thus, we ignore changes in bubble-net shape over this period. We now include discussion of this in paragraph 1 of section II A.

4. What void fractions have been observed? If this could not be measured, what void fractions can realistically be produced by the whales and why are the specific values chosen here?

There are no field measurements for this and it's unclear what fractions whales can create. Our choices follow that of the previous studies that we compare to (e.g. [14,15,17,18]). However, we still investigated a range of void fractions to determine the sensitivity of our results to this parameter. We now introduce this information in section II A.

5. How do imperfections (such as the assumed full spiral vs. the single columns forming a spiral in figure 1a) effect the results?

   If the single columns are sufficiently close and of sufficiently uniform void fraction, then they should have little impact on the mechanisms we observe. Actual bubble nets have a broad range of sparsity of such columns (in space). Further, there are no measurements of the actual void fraction distributions. Thus, it would be challenging to determine the exact effect of this, even with significantly more simulations. We now discuss this idealization of our model in paragraph 1 of section II A.

Regarding the modelling results, I think it is important to address the following:

1. How do the found results relative to existing measurement data?

   Unfortunately there is no measurement data available.

2. If no measurement data is available, what would the results mean in theory? What effect would the observed levels have on the potential prey?

   The results are useful for determining the viability of the different mechanisms, revealing the motives of actual feeding humpback whales. They also has consequences for bubble curtains that are used to attenuate underwater sound for engineering purposes (e.g. pile-driving). If the sound levels in the bubble net are large, then the fish are more likely to prefer the quieter regions within the net, effectively corralling them. This is now discussed in paragraph 2 of section 1.

3. How big is the estimated acoustic effect on the prey in relation to the estimated effect from having "only" a bubble-net?

   This is currently an open question, though it is known that bubble curtains of sufficiently high void fraction can discourage fish from crossing [10]. We discuss this in paragraph 5 of section 1.

In addition to the general above points, please find some more specific remarks below.

1. Figure 1: picture (b) is hard to understand without further explanation. I suggest some more discussion is added, or it is omitted here

   We have added additional discussion to paragraph 1 of section I to describe figure 1 (b).

2. Section I: Whereas the "trapping effect" from a bubble wall intuitively makes sense to me, I do not fully get the mechanisms which are hinted to, where the whales are using vocalizations for feeding. I think this topic deserves a paragraph of detailed explanation for readers, which are not fully familiar with the topic

   We have added additional discussion to paragraph 2 of section I to describe this effect further.

3. Section II: "nor buoyant effects due to rising bubbles." Please explain why it is justified to neglect this effect, which generally is of high importance for bubble cloud dynamics. Please also relate this to the average duration of such a net feeding phenomenon. I understand from a later remark, that the operation takes minutes, which to me would mean that the bubbles rise and change shape significantly during this period.

   We agree that this should have been expounded upon. It is believed that the smaller bubbles leftover from such a feeding operation are the ones involved in the acoustical phenomena we analyze. For these bubbles, buoyant effects are of course less important. Further, the duration of a single whale call would take $< 1$ s to propagate through the net (and, as noted, less than a minute for the entire operation to occur). These times are much shorter than the velocities due to buoyancy of the small bubbles. We thus analyze this quasi-static process (with respect to the rising bubbles) as a proxy for actual bubble-net hunting. We have added a more full discussion of this to paragraph 1 of section II A.

4. Section II: "Whale vocalizations are represented as one-way waves emitted from a line-source of length $L_a = 0.2r_o$" Please explain what the source assumed source characteristic is based on.

   The line source length is a model parameter that does not have a direct physical meaning. One can draw a connection to the width of a humpback whale, which is close to the $L_a$ we use. However, our conclusions were insensitive to the length of this source (both doubling and halving) regardless. We now mention this in paragraph 2 of section II A.

5. Section II: Line 99 and after. Please explain what motivates the choice of the source locations.

   The location of the single line source with grazing angle $\theta = 50^\circ$ is unimportant, since this configuration was chosen just to determine if waveguiding could be observed. This is now added to the discussion of such insensitivities in paragraph 3 of section III A. The distance from the net center for the (b) cases that are equally-azimuthally-spaced is based on estimates of Wiley [19], who observed vocalization within one diameter of the net center. Of course other reasonable choices could be made, though our results are again relatively insensitive to this distance because little attenuation occurs from the vocalization location to the bubble net wall. Discussion of this is now included in paragraph 3 of section II A.

6. Section II. B: It is not clear to me how the effect the irregular distribution of bubble columns visible from Figure 1a and the irregular occurrence of bubble clouds in the columns are accounted for here or if they are omitted for simplicity. If the latter is the case, please discuss this in more detail.

   The irregularity of these bubble columns are not taken into account. As mentioned in a previous response (number 5 in model input questions), this could reduce the efficacy of the net, though in principal should not change the mechanisms by which the nets work. We have added discussion of this to paragraph 1 of section II A.

7. Figure 4: "Example $\sigma = 0.7$ polydisperse cases with larger $\alpha_o$ and smaller $f$" Please explain what is meant by later $\alpha_o$ here, it seems it is the same value ($10^{-4}$) as in the previous example

   Subplot (a) uses $\alpha_o = 10^{-3}$, which is larger than the $10^{-4}$ of the other cases. This figure caption has been revised to emphasize how these cases differ.

8. Section III. B / Figure 5: The figures seem to indicate that all whales/sources are emitting perfectly phase synchronized signals. As this might not be the most realistic case, how are the results effected, if the signals of the single sources are emitted with a random offset in time domain (or a random phase difference in frequency domain)?

   This is an excellent point that should have been emphasized. Indeed, this is not the most realistic case, though it does result in maximum constructive interference between the line sources and thus the loudest possible net interior. We now mention these points in paragraph 2 of section III B.

Response to Referee # 2
**Simulation of humpback whale bubble-net feeding models**
*Bryngelson and Colonius*

We are grateful for the referee's efforts in improving the quality of this paper. We quote the comments of the referee and discuss changes made to the paper in response to these comments in the following.

*Referee #2*

This paper is a major revision of an original paper submitted to JASA-EL that took an approach for coupling acoustics with multiphase flow and used it to simulate sound propagation in the bubble nets produced by humpback whales, work that followed on from previously published work by Leighton. The original paper was rejected with much of the concern being centred around the fact there were too few new results to report to significantly add to Leighton's work, that the work presented took an overly simplistic approach with monochromatic insonifications of simple annular bubble nets containing bubbles of a single size, and that the new work failed to fully grasp the latest conclusions in Leighton's work and the reservations he had about some of the simplifications which the authors took on face value. The attempt to further develop the original work of Leighton is an interesting problem and is worth pursuing. Though the criticisms of the original paper were extensive, the authors have made genuine and serious attempt at responding by addressing each of the issues in turn. Although some of the original simple approaches (such as the annular ring geometry and monodisperse populations) are repeated again, the authors justify these as instructive means of introducing the theory before moving on to the more sophisticated approaches (spiral net geometry, polydisperse populations, etc). The authors have substantially improved the quality of the paper, and I believe that it will add to the scientific literature in this field and, after revision to correct some of the ambiguous acoustic terminology used, its publication would now be justified. However, there are a one or two aspects of the paper where the authors let themselves down. There are some sloppy uses of acoustic quantities and units which are ambiguous and which need clarification. It may be the case that these can be cleared up but simple re-wording of some of some statements, or it may require a little more effort to address them, I am not sure because the meanings are not clear to me. I would suggest that the authors consult an acoustics text book, or better still the international standard ISO 18405:2017 which provides definitions of some of the terms they use. If they are adopting these definitions (I suspect not), then they need to state this. If they are using alternative definitions of their own, then they need to make these clear to avoid ambiguity. It may be the case that the errors and ambiguities do not necessarily alter the substantive conclusions of the paper which relate to the relative level of insonification of different spatial regions in the different bubble nets, but the figures presented and the values within them may well change. Subject to clearing up these issues, I would support publication.

1. Page 7, line 87: I am not quite sure why the authors use the density of pure water (998 kg/m$^3$) rather than that of sea water. Can they explain? The density of surface seawater ranges from about 1020 to 1029 kg/m$^3$, depending on the temperature and salinity.

   We appreciate this clarification. We use the usual pure water density as opposed to approximating the density of surface seawater (or density variations in space with salinity concentrations, for that matter!) Of course, our conclusions are insensitive to this small change in density. We have added discussion of this to Page 7.

2. Page 8, line 95: The statement "The sound pressure of the vocalizations is about $A = 180$ dB re 1 $\mu$Pa" is not clear. Do the authors mean that this is the value of the Source Level

of the vocalisations? This would be a reasonable value to assume for the Source Level of a humpback whale. If so, the units are incorrect, and should be "dB re 1 $\mu$Pa·m" (sometimes written as "dB re 1 $\mu$Pa at 1 m" in the literature). If the authors really mean that this is the "sound pressure", then the units of this are pascals (Pa). If they mean that this is the Sound Pressure Level (SPL) then, as a level of a quantity, this is expressed in decibels relative to 1 $\mu$Pa as stated in the manuscript. However, the sound pressure (and the SPL) vary with position in the acoustic field (diminishing with range from the source), so the position in the acoustic field must also then be specified. In which case, is this position the entry point of the wave to the bubble net?

We appreciate the attempt to clarify our nomenclature. Whale vocalization sound levels for feeding calls range from 160 to 190 dB re 1 $\mu$Pa at 1 m (Thompson et al. JASA 1985). $A = 180$ dB re 1 $\mu$Pa s the peak sound pressure at the source location (not at 1 meter, nor at the entry point into the bubble net). However, the attenuation that occurs between the source and net is small and our conclusions are insensitive to it. We have appended the discussion of paragraph 4 of section II A to clarify these points.

3. Page 11, lines 127-138: A number of acronyms are used for the first time in this section, and they need to be defined at first use, or a glossary should be provided. Examples include: MFC, WENO, HLLC, TVD, and CFL.

   We regret omitting the definitions of these acronyms, which may be new to many readers. They have been added to the manuscript at their first use. They are also defined below for completeness:
   MFC: Multi-component Flow Code
   WENO: Weighted Essentially Non-Oscillatory
   HLLC: Harten, Lax, and van Leer Contact
   TVD: Total Variation Diminishing
   CFL: Courant-Friedrichs-Lewy.

4. Page 12, Figure 3 caption: I am not sure what the authors mean by "maximum deviations from ambient pressure". Do they mean simply the sound pressure (which is commonly defined as the perturbations on the quiescent (ambient) static pressure)? Or do they mean the maximum variations in the steady-state sound pressure value? This has implications for the terminology used elsewhere in the paper.

   We regret the confusing language used in this caption. This is simply the sound pressure. We have clarified related language throughout the manuscript.

5. Page 12, Figure 3 caption: When the authors say "a 180 dB pulse", I assume they mean the 180 dB re 1 $\mu$Pa specified on page 8 (though I am unclear if this was actually a Source Level). In any case, in the figure caption they should specify what the quantity is. If it is the SPL, then it would by convention be the level (in decibels) of the RMS sound pressure. Is the "pulse" a burst of constant acoustic frequency?

   We again regret the ambiguous language. The referee is correct that it is a monochromatic sine wave of frequency $f$ and peak sound pressure 180 dB re 1 $\mu$Pa pulse. We have modified the figure captions and appended paragraph 2 of section II A to clarify this.

6. Page 14, line 168: Is the word "pollution" the correct word in this context?

   We agree that a more appropriate word should be chosen. To be more precise, we have replaced "pollution" with "acoustic excitation".

7. Page 15, Figure 5 caption: What is meant by "Long-time pressure"? Do the authors mean by "long-time" that there is sufficient time for steady-state conditions to have been achieved? I assume that the "pressure" is really the sound pressure? The plot shows the ratio of pressures - is this the ratio of the absolute pressure to the ambient pressure - which is really just a measure of the sound pressure? Or is it the fractional change in sound pressure?

   Long-time is indeed vague in this context, since the flow itself is never steady, but rather becomes periodic with the acoustic excitation frequency $f$. The referee is correct: long-time in this context means that an essentially stationary state has been reached. We have modified the caption to define our use of "long-time". The referee also correctly ascertained that the figures show a measure of the sound pressure (the ratio of absolute to ambient pressures), which we have also added to the caption.

8. Page 16, Figure 6 (and surrounding text in Section III. B): How do the authors define $\mathrm{SPL}_{\max}$? The SPL is conventionally defined as the level of the RMS sound pressure expressed in decibels (the RMS being calculated over a specified time interval). The $\mathrm{SPL}_{\max}$ could be the maximum value of SPL (in dB) in a series of SPL values (calculated over time or space). Or it could be the maximum value of the sound pressure expressed in decibels (more typically termed the "level of the peak sound pressure").

   We showed $\mathrm{SPL}_{\max}$ as the maximum sound pressure level expressed in decibels (in a series of sound pressure levels). We have relabeled this quantity as "peak sound pressure level" or "peak SPL" throughout to clarify this.

9. Page 16, Figure 6: The vertical axis of Figure 6 is labelled as the quantity "SPLmax / A". As far as I can tell from the text, both of these quantities are expressed as levels in decibels. Dividing one level in decibels by another to determine a fractional difference has no real meaning. If the authors wish to express a fractional difference of levels in decibels then they should subtract the values rather than divide them (subtracting two levels in decibels is equivalent to taking the ratio of the relevant linear quantities - a 0.1 dB change in level is equivalent to about a 1% change in the linear quantity). This comment applies to other parts of the paper (and Figures 8 and 9).

   This is correct, and we agree that dividing two logarithmic quantities is confusing. We have changed figures 6, 7, and 8 to remove these ratios (labeling $A$ separately).

10. Page 16, lines 196-201: The term "loudness" has a specific meaning in acoustics, and I do not think that the authors mean that here. Instead I believe the authors mean the amplitude (or magnitude) of the acoustic field. Could they just not use the SPL as a measure of this?

    This is correct, we have misused the term loudness here. We have clarified our language to use peak sound pressure throughout (consistent with the results we plot in figures 6, 7, and 8.

11. Page 17, lines 205-212: The quantity "$\mathrm{SPL}_{\max}/A$" is again used as a metric without proper definition. See the comment above regarding Figure 6 for discussion of the proper use of level quantities expressed in decibels. In the statement "$A = 180$ dB vocalisations", the quantity $A$ must be specified.

    Following the reply to comment number 9, we have removed the use of this quantity entirely, instead referencing the sound pressure level itself.

12. Page 17, lines 205: Is the $\mathrm{SPL}_{\max} = 0.5A$ criterion completely arbitrary or based on evidence? How is it justified?

    This criterion is indeed arbitrary. We have adjusted our discussion throughout to make this clear, instead using it as a simple example of how quiet the net interior can remain.

13. Page 17, line 208: Do the authors really mean "damp" in this context? Or would "attenuate" be a more general word for the process?

    We agree that "damp" is incorrect here, and have switched it to "attenuate" as suggested.

14. Page 19, lines 246-247: What do the authors mean by "averaged sound pressure $\mathrm{SPL}_{\mathrm{avg}}$"? Over what time duration was the SPL calculated and or over what spatial domain was the spatial average taken? The SPL is by convention already an average of the sound pressure over time (the root-mean-square sound pressure expressed as a level on decibels). How was the $\mathrm{SPL}_{\mathrm{avg}}$ calculated? If the calculation was the result of averaging the SPL decibel values, then this is the geometric mean of the sound pressure and not the arithmetic mean.

    The spatial domain of the averaging is a $0.2r_o$-thick patch that occupies the bubble-free region of the spiral in, centered at $(-3r_o, 0)$, $(0, 2.5r_o)$, $(2r_o, 0)$, and $(0, -1.75r_o)$ for patch 1 through 4, respectively. The absolute sound pressure is averaged over the patch and the sound pressure level is computed via this average, so this is indeed an arithmetic mean of the sound pressure. We average over the time duration $t \in (15, 25)r_o/c_l$. These details are now fully included in section III C paragraph 3.

15. Page 20, Figure 8: See comment on Figure 6 regarding the quantity "$\mathrm{SPL}_{\max}/A$" and the meaning of this term which needs explaining.

    Following the replies to comments 9 and 11, we have removed the use of this quantity entirely, instead referencing the sound pressure level itself.

16. Page 21, Figure 9: See comment on Figure 6 regarding the quantity "$\mathrm{SPL}_{\max}/A$" and the meaning of this term which needs explaining.

    Following the replies to comments 9, 11, and 15 we have removed the use of this quantity entirely, instead referencing the sound pressure level itself.

17. References: The citations to the references are numbered in the main body of the text, yet the references at the end are not listed in numeric order but alphabetical order. Personally, I find this confusing and I think it would be better to list the references in numeric order when using number citations.

    We agree and will request that JASA adopt this styling in the printed paper.