Referee comment 2:

General comments:

Johnson et al. present a novel technique for collecting in situ measurements of the near ice-water interface of glacier ice using a frame that can be equipped with a suite of instruments and attached to the ice. Understanding heat transport through the near ice-water interface of glacier ice is key to accurately estimating submarine melting, which can be underestimated using current theory according to some observational studies. The technique of Johnson et al. seeks address a current lack of direct observations with scope to improve glacial melt estimates in the future. The authors tested their apparatus, fitted with a hydrophone, underwater camera and thermistor string, on growlers in Hornsund Fjord, Svalbard. I think this is an exciting technique and a nice, if very brief, showcase of some measurement capabilities. I think that, with some adjustments, a paper is a nice way to present this research as a proof-of-concept study.

Specific comments:

I agree with Referee #1's comment that the deployment description is quite lengthy, for example, details about boating safety and frame redeployment after unsuccessful attempts could be omitted. I also wonder if further data exploration might be feasible/possible. Temperature measurements spanning 20 minutes are mentioned in section 3.4, but only 2.5 minutes worth of data are shown. If images and audio were also collected for the 20 minutes, could you present coincident data and look for, e.g., images of bubbles being released (learning more about this seems to be a major part of the justification) corresponding to audio pulses and maybe being seen somehow in the temperature data? Some kind of quantified results from what I've mentioned above or even something like bubble release rate from the audio data that you could compare with other literature to show that your apparatus is providing novel observations would really help the discussion which is mostly rationale and outlook at the moment. The outlook is good though and proposes some interesting future applications for your system.

We thank the referee, Matthew Corkill, for the thoughtful comments.

We will go through the deployment description carefully and make an effort to trim down any superfluous information.

One important clarification, which we can and should make clearer in the manuscript, is that with the iteration of the frame described here we were able to collect acoustic and video data simultaneously, or temperature data alone. We did not manage to make all three types of measurements simultaneously.

The point about including more quantitative data analysis is nevertheless well-taken. While we do not have coincident acoustic or video data to go with the temperature data, we would be happy to incorporate some basic quantitative analysis of the temperature data, such as the average thermal gradient observed. We did not display the full 20 minutes worth of temperature data graphically because we did not feel that doing so would provide any additional insight, as the other 18 minutes are qualitatively similar to the 2 minutes that are shown. In terms of the acoustic and video data, we can absolutely identify some individual bubbles in the video data, estimate their size and rise rate through the water, and also find the corresponding pulse generated by the release in the acoustic data, and compare its primary frequency to the natural frequency expected base on the video data size estimate.

Briefly on the frame design, could the pair of flotation spheres collide and interfere with audio data? Perhaps something like a large, sealed PVC pipe could be attached parallel to and above the crossbeam instead of the spheres (or the aluminium crossbeam could even be replaced by a larger floating one).

The concern about noise from the flotation spheres interfering with the acoustic data is valid, and in the future a solid float fixed to the frame would indeed be a good way to eliminate this issue. For the present data, the fact that the separation between the floats and the hydrophone is several meters, whereas the hydrophone is only a few centimeters from the ice face, means that spherical spreading of the sound, which results in a 1/r dependence for the amplitude of the signal, should reduce the impact of noise from the floats. Additionally, noise generated by the floats colliding should be at a lower frequency than the bubble pulses, making it fairly easy to differentiate and filter out if necessary.

Technical corrections:

Line 15: Consider moving details of salt and gas to the third sentence where they're mentioned and then combining the first two sentences.

Agreed, we will do this.

Line 19: Replace "this interface" with "the proximal boundary".

Agreed.

Line 35: Poorly how?

Specifically, McChonnochie and Kerr 2017 point out that the common form of the parameterizations assumes that the boundary layer is governed by a shear-driven instability, which they suggest is only valid at water velocities in excess of about ~5 cm/s, and at velocities lower than this the thickness should instead be controlled by a convective instability. They suggest that this results in an underestimate of melting in the low flow speed regime. We will make this more clear in the text.

Line 51: Please check style guide regarding capitalising figure # and appendix #.

Understood, we will correct this throughout the manuscript.

Line 107: Remove “then” after “crossbeam was”.

Agreed.

Line 121: How was the size range 3-5 m determined? Also, spaces should be removed either side of dashes in this paragraph or added elsewhere.

This is the rough ranges of sizes of the growlers that we successfully deployed the frame on. As discussed around lines 84-92, growlers smaller than this generally weren't large enough to accomodate the frame, whereas larger pieces of floating ice pose more danger to the operators. Section 3.1 is intended to be a summary of the performance capabilities of the frame; if it would be useful, we can refer back to the previous text where the reasons for these limitations are explained in more detail.
We will fix the inconsistency with the dashes.

Line 133: Quantify? “Enabling calculation of a rise velocity of …”.

We will add some calculated rise velocities of bubbles.

Line 141: “figure 3e-f”.

We will fix this.

Line 156: “…ice face of a floating growler about 30 cm below the sea surface…”.

We will make this change.

Line 183: What was the wall thickness of the tube?

The wall thickness was 0.065 inches; we will specify this in the text.

Line 185: What material was the threaded rod?

The threaded rod was stainless steel; we will specify this in the text.

Line 209: “remain”

We will fix this.

Figure 2: It would be nice to show the thermistor array and the optional standoff beam. It would also be nice to label the instruments with what they actually are.

This is a good suggestion, and we can easily add a picture of the thermistor array. The reason for the vague label "instruments" was to emphasize the fact that other types of instruments could be deployed, but we agree that it would be better to label them clearly.

Figure 3: As mentioned above, it would be nice if more data could be included here. If you have images of a bubble being released and rising, I think it might be very nice to include multiple images showing this. If the ice-face images could be linked to the audio, I would be great to timestamp the images so that they could be linked to a longer time series of audio data. I understand that this may not be possible due to distance between instruments, but that distance may be an important consideration for future deployments if the goal is to link data from all the different instruments.

As discussed in our response to the earlier comment, we will add some more quantitative analysis of the acoustic and video data. Specifically, we could present several frames of video data showing a rising bubble, as suggested, as well as the synchronous acoustic data of the bubble-release pulse.

Figure 3a: This may be easier to relate to panels c-f if time were put on the x axis (though I also understand the logic of having distance on the x axis). It might be interesting to include a black line at some specific temperature contour to better see patterns.

As mentioned earlier, the temperature data was not collected at the same time as the acoustic and video data. Ideally it would be a separate figure, but we are constrained to three figures by the brief communication format.

Figure 3d: In the caption, “A blown-up view of a bubble-release pulse recorded from the ice frame. e…”

We will fix this.