

Answers to reviewer comments 2

We thank the reviewer for his constructive comments. We address them below and refer to sections in the manuscript we changed accordingly.

The topic is important and under-researched, so the paper is timely. The authors perform their experiments on laboratory snow, but make a credible effort to mimic different types of weak layers. Microstructure characterization is thorough and lends substance to the results. I have many positive things to say about the paper, which I will be happy to see ultimately published. However, I leave the niceties for future comments since I am seriously stuck on one essential question regarding the mechanical analysis. Let me come straight to the point.

The authors analyze the stress strain behavior of their samples until 'failure' which is, in Figure 4, marked by a little red dot at which the record abruptly terminates. Since no other information is given, I have to presume that, at this point, some catastrophic process occurs that makes further recording futile. (I note that it would have been tremendously helpful for this referee to be able to access the unprocessed data, so I would not have to resort to such wild assumptions). However, I have hardly ever in my professional life as a materials scientist seen a material that would fail with an abrupt and complete loss of cohesion, at least not if plasticity was involved before. There is almost always some cohesion post 'failure' in quasi-brittle materials. Also, there is residual dissipation.

As to snow, it is well known to exhibit strain softening at high and strain hardening behavior at low strain rates. The authors discuss this but add to my confusion by arguing that, prior to failure, their samples are 'perfectly plastic'. This leaves me utterly confused. A perfectly plastic material by definition does not undergo brittle failure. So what exactly happens at the point of failure?

We agree that a clear description of the different stages of deformation and failure in the experiments is important. We observed a linear stress increase which later leveled off indicating plastic deformation and subsequently failure occurred. We did not observe brittle failure (changed wording in line 192 to be clearer). In fact, in section 3.1, we aimed at quantifying the amount of non-recoverable deformation.

Section 2.5 now includes a description of how failure was characterized and determined in the experiments. Failure was defined based on the particular situation and observed in the high-speed images when the slab started moving horizontally. To illustrate that, we added to the case of sample 2 (tested on 07/01/18) a force-displacement graph (now Figure 4a), the slab displacement and shear strain of the weak layer with time (before, at, and after failure (Figure 5a and 5b), as well as near infrared image of the weak layer zone before, at, and after failure overlain with DIC shear strain analysis (Figure 5c-d).

The definition of 'failure' by the authors is as follows: "As failure we consider the abrupt decrease during a few milliseconds in the shear stress with respect to the global shear strain measured on the sample. Hence, we refer to failure as the catastrophic fracture of the before connected ice structure.". This definition is near-meaningless: In a disordered material like snow under increasing imposed displacement, the stress undergoes up-and-down fluctuations all the time. That load fluctuates with respect to the global shear strain is self-evident since shear displacement (viz global shear strain) is what is controlled in the experiment. But how are such fluctuations, which may well imply a decrease of stress over a few milliseconds, distinguished from 'failure'? Please answer me the following questions:

- 1) What is 'abrupt' ?
- 2) What is a 'decrease' (as opposed to a mere fluctuation)?
- 3) Over how many milliseconds does the decrease need to occur?
- 4) Are you referring in your definition of failure to the magnitude of the stress decrease, or the rate of stress decrease, or some combination of both, or to something altogether different?
- 5) What is the reason for not showing, in Figure 4, any data after the red dot?

A detailed description of how failure was determined and a graph (Figure 4a) illustrating the concept is provided in Section 2.5, as described in the above answer. Failure was defined based on the particular situation and was chosen before rigid body motion of the slab began. The force-displacement graph in Figure 4a shows that failure cannot be mistaken by up-and-down fluctuations of the force signal, but is characterized by a prominent force decrease of different magnitude than mere force fluctuations. In addition, Figure 5a shows the onset of rapid horizontal motion of the slab, which occurred right after failure. This acceleration of the slab occurred within about 18 ms.

The displacement measurements (Figure 4a) were taken at the load actuator and refer to the deformation of the entire sample which does obviously not capture strain concentrations. Values of the shear strain in the weak layer were determined with digital image correlation. These data are available until failure. After failure the upper part of the snow sample (slab) slides with respect to the base of the snow sample which does not allow for meaningful DIC-

derived strain measurements from the region of the weak layer in the images. The above considerations are now better described in the paper in Section 2.5.

The subsequent statement that failure is the 'catastrophic fracture of the before connected ice structure', i.e., that it implies a loss of global connectivity, would be warranted only if the authors had data to substantiate it. This would require an in situ characterization of the bond microstructure which, given the scale and image rate of available Xray tomography equipment, the authors clearly cannot do. So the statement is irrelevant, because unsupported by data.

By 'catastrophic failure of the before connected ice structure' we meant that after failure the slab is disconnected from the base of the sample. Due to this horizontal discontinuity in the structure along the weak layer we observed a sliding motion that was very much apparent in videos and high-speed images of the experiments – but is, however, not supported by measurements. To prevent confusion, we avoid the term.

I am now at a loss how to proceed. The authors do not give a tangible definition of the key concept of 'failure'. Failure is when their data record ends (but why??). Without a clear criterion for 'failure', the experiments cannot be reproduced by other scientists since we would simply not know where to put the red dot in Figure 4 and why.

I am leaving this review incomplete at this point because the interactive format, I hope, provides the authors with the possibility to clarify the point where I am stuck so that, once unstuck, I can proceed to give a meaningful and constructive opinion on the rest of the paper.

We hope the revision clarifies the raised issues and thank the reviewer for his comments.