**Formal Review (Anonymous Referee #1, 31 Dec 2020):**

*General comments: The Behn et al. contribution tackles a persistent question in the glaciological community, namely whether the stress exponent of 3 typically used in ice flow constitutive laws has a robust physical justification, despite the lack of direct supporting experimental data (even in the classic Glen papers of the 1950s). The work is timely, of widespread interest, and with high potential impact. This study concludes that the combination of two deformation mechanisms – grain boundary sliding and dislocation creep – combine through a grain size dependency to produce a composite constitutive law with a stress exponent of approximately 3. I find the paper has a nice flow and clear macroscopic logic. At the same time, potentially ambiguous definitions and an incomplete treatment of uncertainty means that this paper as written is unlikely to be perceived in the community as close to the final word on the topic. Restructuring the analysis to more explicitly incorporate uncertainty will take some time but yield a contribution with stronger impact.*

Thanks for these suggestions. Below we describe specific changes that can be made to address the ambiguity in the definitions and the role of uncertainty in our model.

*Specific comments: In this section, I list a number of focused comments and questions that I feel the authors should address in a revised version of this manuscript.*
*Although we all have a mental picture of grain size, for the quantitative approach taken in this manuscript, a more exact definition of the term would be useful. Is grain size the diameter of the equivalent area of a circle, or some other measure? Tied to that, the grain circularity is a potentially critical concept yet barely mentioned. The comparisons with natural data evaluated size (and even that not clearly with the same definitions), but not circularity. Granted, circularity may not have a large practical effect, but some evaluation would assuage concerns.*

Grain size is a challenging material property to quantify because grains are irregular and are typically measured in a 2-D cross-section (e.g., thin section) through a 3-D sample. In practice grain size is measured in field and laboratory samples in a number of ways. Here we adapt the line intercept technique used by *Alley & Woods* (1996) in characterizing the GISP2 core. In this approach the average distance between grain boundaries along a series of lines through a sample is measured and then scaled by a correction factor of order 1 (=1.5 for circular grains; *Gifkens*, 1970) to account for the fact that when making a thin section many grains are cut near their edge as opposed to near their center (*Gow*, 1969). This approach was also used in the measurement of grain sizes in the derivation of the flow laws (*Goldsby & Kohlstedt*, 2001) and thus allows us to compare our calculated grain sizes to the GISP2 ice core data in a self-consistent manner. Thus, in both the wattmeter and our measurements we have implicitly assumed that grains are circular. This can be clarified in a revised manuscript.


Alley, R. B. and Woods, G. A., 1996, Impurity influence on normal grain growth in the GISP2 ice core, Greenland, J. Glaciol., 42(141), 255–260, doi:10.3189/S0022143000004111.

Gifkens, R.C., 1970, Optical Microscopy of Metals, American Elsevier, New York, 208 pp.

Goldsby, D. L. and Kohlstedt, D., 2001, Superplasticdeformation of ice: Experimental observations, J. Geophys. Res., 106(B6), 11,017–11,030, doi:10.1029/2000JB900336.

Gow, A.J., 1969. On the rate of growth of grains and crystals in south polar firn. Journal of Glaciology 8 (53), 241-252.

*Stress is used throughout the manuscript, and it too would benefit from clearer definitions and identification of the relationships among the various forms applied. For example, piezometers/wattmeters typically use differential stress, but Equations 17-19 use shear stress and Figures 3 and 4 plot effective stress. In earlier equations, such as Eq 5, the form of stress isn't specified. In addition, I did not see a definition for "effective stress", though I presume it represents the square root of the second invariant of the deviatoric tensor. By defining these terms and explaining how the equations use the appropriate formulations (e.g., shear stress cannot directly be used in the flow law), confidence will be higher in the calculation results.*

The reviewer is correct that we were somewhat sloppy in our definition of stress in the manuscript. The stress used in the definition of the wattmeter (defined as sigma in Eqs 5–14) and in the flow law (Eq. 16) is the Von Mises equivalent stress. However, in our simplified geometries the only non-zero term of the deviatoric stress tensor is the shear stress (defined as tau in Eqs 17–19). We relate the shear stress to the Von Mises stress through the square root of the second invariant of the stress tensor, which in this geometry simply reduces to $\tau = \sqrt{3}\sigma$ .

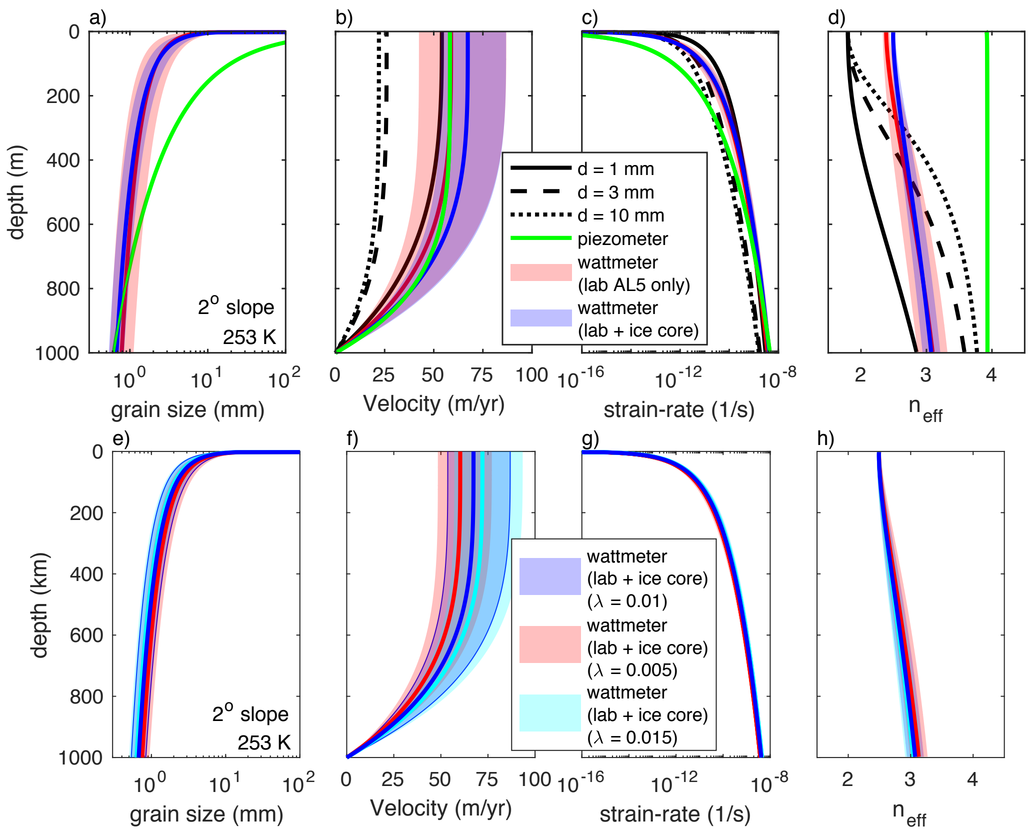
*This manuscript implies strongly that the grain size is a better wattmeter than piezometer. That may be, but which approach better matches geologic/glaciologic reality re- mains unresolved in the community: despite the availability of the wattmeter for over a decade, many studies still rely on a piezometric approach. As such, a more explicit comparison in this paper seems warranted. The partial approach presented here has some merit, but feels incomplete and, in places, inaccurate. For example, the authors present evidence for the wattmeter matching grain size data, and I agree in part. Taking Figure 5 for the moment, there is no illustrated scenario in which the model matches the experimental data for all three strain-rate cases. The highest strain rate (blue) never achieves a steady state, so we cannot really evaluate that case. The lowest strain rate case (black) and middle (red) are best fit with different lambda values, neither of which is what Austin and Evans use in their original study. Yet I don't see a treatment of this uncertainty in, for example, Fig 6. (At least, as far as I can understand from the text, the uncertainty shown in Fig 6 does not include a variation in lambda.)*

To better illustrate the influence of the assumed lambda value (the fraction of the total work responsible for increases in internal energy), we have added a new row of panels to Figure 6. These panels illustrate the changes in grain size, velocity, strain-rate, and effective stress exponent that result from varying lambda over the range of values used in the comparison to the laboratory data (Figure 5). We emphasize that the differences in the wattmeter predictions due to the uncertainty in lambda are significantly smaller than the variations associated with the uncertainty in the grain growth parameters. The revised version of Figure 6 (below) will be included in the revised manuscript. As pointed out by *Austin & Evans* (2007; 2009), lambda values from ~0 to 0.4 are reported and these are not well constrained; they emphasize that lambda is a scaling factor. Similar values to what we report for ice have been determined by applying the wattmeter to recrystallization of quartzite (*Tokle et al*., in revision) and olivine (*Holtzman* et al., 2018).

Austin, N. and Evans, B., 2009, The kinetics of microstructural evolution during deformation of calcite, J. Geophys. Res., 114(B9), B09402, doi:10.1029/2008JB006138.

Austin, N. J. and Evans, B., 2007, Paleowattmeters: A scaling relation for dynamically recrystallized grain size, Geol., 35(4), 343, doi:10.1130/G23244A.1.
Holtzman, B.K., Chrysochoos, A., and Daridon, L., 2018. A thermodynamical framework for analysis of microstructural evolution: Application to olivine rocks at high temperature. Journal of Geophysical Research: Solid Earth 123, 8474-8507. doi:10.1029/2018JB015613.
Tokle, L. and Hirth, G., 2020. Assessment of quartz grain growth and the application of the wattmeter to predict recrystallized grain sizes (submitted to JGR, available on EarthArXiv: https://eartharxiv.org/repository/view/1865/)

*In addition, on line 122, the text implies that all internal energy goes to grain boundary area. That may be sufficiently accurate, but should be justified by exploring the potential of dislocations and elasticity to store energy. Austin and Evans (2009), for example, mention dislocation-driven energy variations after eq 19. The questions and notes I raise in this paragraph all lead to a concern that uncertainty in parameter values and applied processes preclude a robust conclusion about the stress exponent derived from the presented data.*




By setting $\lambda_{GBS} = \lambda_{disl}$ in all our models, the effects of grain boundary energy, grain geometry, and lambda, can all be grouped into a single "scaling factor" $= {}^{c\gamma}/_{\lambda}$ in Eq. (14). Thus, the variations in lambda shown in the new version of Figure 6 elucidate the behavior of the model with respect to variations in any of these parameters. We also note that while we use a lower value of lambda compared to *Austin & Evans* (2007; 2009), the ability of the wattmeter to fit grain size data across multiple orders of magnitude in stress and strain-rate from the lab (Fig. 5) to the GISP2 ice core data (Fig. 7) provides strong motivation that this approach is capturing the first-order physics of grain size evolution.

It is important to emphasize that the wattmeter formulation models the rate of *change* in the internal energy (with the lambda scaling factor), and relates this to the grain size reduction rate (and thus increase in internal energy owing to increase in grain boundary area). This is a bit confusing, because the driving force for the grain size reduction is indeed the internal energy associated with dislocations. The key assumption here, outlined by *Austin and Evans* (2007; 2009), is that the rate of change in grain size is greater than the rate of change in stress – thus the dislocation density can be considered constant for a given stress.

*Another significant discussion component that would lead towards a greater impact of this paper is a comparison of the effect of grain size against other rheological controls. The principle factor to address is anisotropy due to fabric development. The relationship will change with depth and affect the stress-strain-rate environment, and may affect grain size evolution.*

This is an excellent point. The development of fabric in the shear plane will weaken the ice, resulting in a larger strain rate for the same stress. Because the stress is fixed by the surface slope, this will result in a positive feedback in which enhanced fabric development will drive further grain size reduction (due to the enhanced work rate). Future studies should investigate these feedbacks in order to quantify the magnitude of this effect.

*Technical Comments Equations 17 and 18: I do not understand the source nor assigned value for viscosity; it doesn't appear in Table 1 and is somewhat at odds with the formulation of Eq. 16. I imagine I am missing something here, so an explanation would help.*

The viscosity is simply reformulated from the flow law (Eq. 16) in terms of the stress, e.g.,

$$\eta(\sigma) = A^{-1}\sigma^{1-n}d^{m}\exp\left(\frac{Q}{RT}\right)$$

*Line 224: I do not follow how the iteration between Eqs 14 and 18 works in practice.*

To calculate grain size, we first make a guess at the initial shear stress and grain size. Using these values, we calculate viscosity and use Eq (18) to make a new estimate of the shear stress. Based on our new estimate of shear stress and the corresponding strain-rate (calculated from the flow law), we use the wattmeter to calculate an updated steady-state grain size (Eq. 14). These new estimates for stress and grain size are then used to recalculate viscosity, which is in turn fed back in Eq (18) for the next iteration. We continue to iterate in this manner until the shear stress varies by less than 0.1%.

*Figure 4: This is a relatively small comment, but can sow uncertainty. As I read it, the right panel is derived from the slope of the left panel. However, the right panel seems to have the same number of teal and blue dots. I am not clear how the authors calculate slope at the termini of the series of discrete points.*

We actually made calculations at a much finer resolution in stress than is shown on the figure, thus while the dots are indeed slightly offset (by half the spacing in our sampling of stress) it is hard to see this in the figure. For clarity, we will adjust the figure to plot the dots in panel (b) halfway between the dots in panel (a).

*Lines 353-355: "In this scenario…" These two sentences feel to me to be a circular argument. I concur with the comment provided by PD Bons that the manuscript should recognize that natural data do not necessarily require n=3, and in fact n=4 may be a more accurate representation.*

Please see our response to Dr. Bons below, in which we address these concerns.