Dear Professor Christopher Gerbi,

We appreciate your helpful comments. Here we present our responses to the comments. Our responses are in black; your comments are in blue. Revisions of the manuscript are in italics.

1. Many places in the manuscript and in the conclusions (e.g., lines 198, 209, 230, 231, 286, 294, 472, 473) n is described as varying. Yet Line 232 (similar to line 364) claims that "n can be regarded as relatively independent of both temperature and concentration". I find those statements incompatible, but perhaps I am missing something. In my limited time spent with these data, I do not see a systematic relationship, but I have difficulty seeing n as independent, particularly given how far outside of uncertainty the values lie according to Table 1. Possible resolutions are that the uncertainties are higher than reported or that the data aren't able to capture some additional complexity.

Reply: Thank you for the suggestion. We found no systematic changes in values of n with changes in solute concentration or temperature. However, the values of *n* of some samples showed significant uncertainty, and we cannot rule out the impact of this uncertainty on the value of *n*, although our fitting method strives to minimize the uncertainty of the value of *n*. As you suggested, we don't use "independent" to describe it.

Uncertainties in *n* are also brought by bubbles. The amount and size of air bubbles trapped along grain boundaries can strongly affect the value of *n* (Azuma et al., 2012). In the study of Azuma et al. (2012), the value of *n* of bubble ice varied between 6 and 20, with the majority lied between 6 and 11. With bubbles inducing such large uncertainties, we agree that a potential trend, if existed, would be overwritten by the uncertainties. But for this study, we have to work with the current data and acknowledge the weakness.

Now the sentences are changed to the following:

Line 232: "*While no systematic changes were observed in the value of n as the solute concentration or temperature changed, we cannot rule out that the uncertainty in the value of n may be higher than the systematic changes due to solute concentration and temperature.*"

We also modified the first paragraph in Section 5.4:

Line 362: "*The presence of air bubbles could induce large uncertainties in the fitting of n (Azuma et al., 2012), which may overshadow the effect of different impurities contents on n, if existed. However, in this study, we have to make inferences based on the available data set, while acknowledging the uncertainties. Given the same experimental conditions, we assume the effect of air bubbles should apply equally to pure and doped ice samples, and thus, we can investigate the effect of soluble impurities on grain growth by comparing samples with different impurity concentrations. As elucidated in the previous subsections, the values of n suggest that the grain growth*

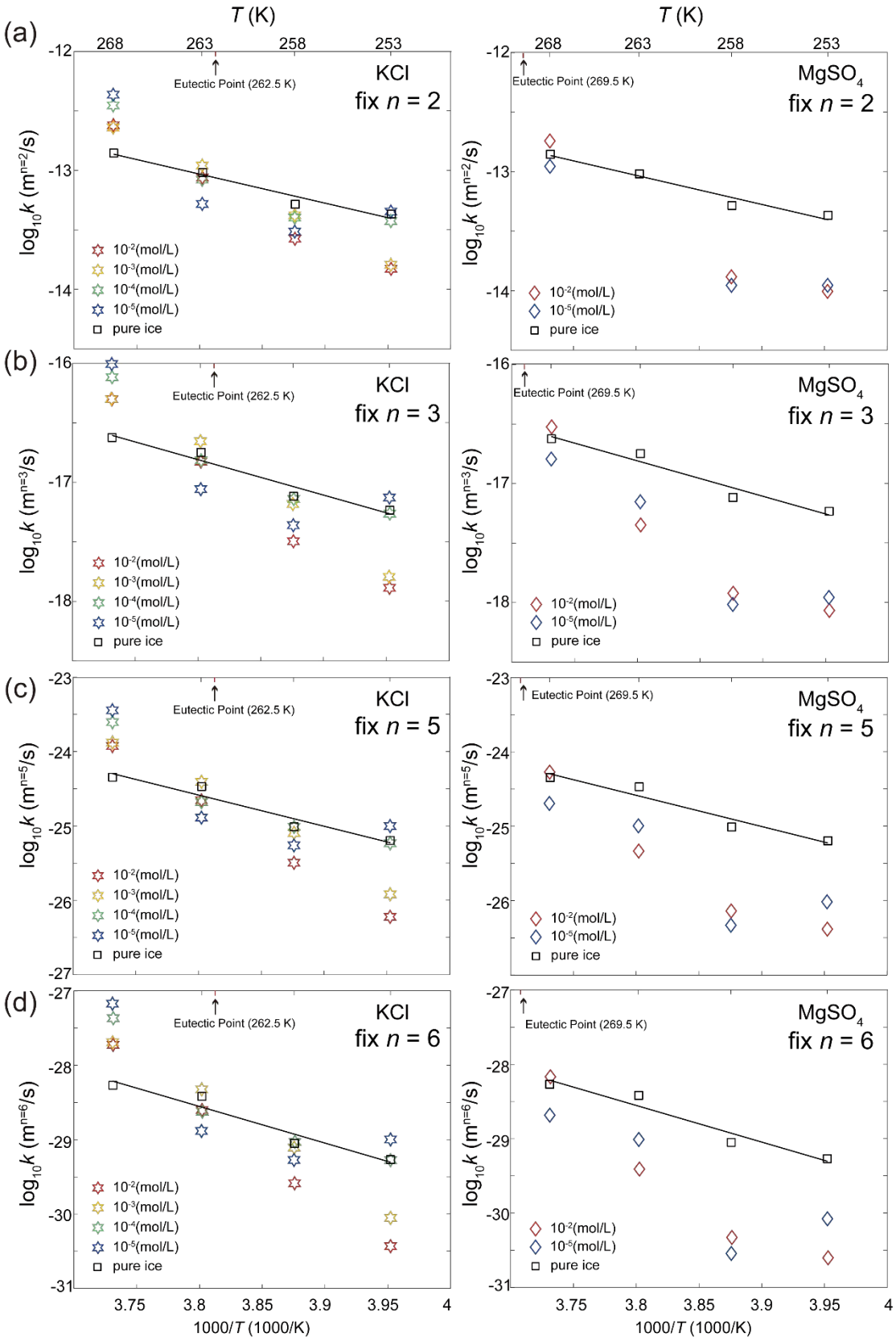
occurred in a multi-phase system. We have not observed any evident systematic trends in n with changes in temperature or solute concentration. We tentatively assume that the grain growth in all doped ice samples can be considered as controlled by the same mechanism. As illustrated in Appendix Figure A8 (a), the majority of fitted values of n lie around 4 and 5. And, as illustrated in Table A1, n = 4 provides good fits for all samples. We take n = 4 as a representative value for all samples in the following discussions."

We also expanded Appendix Table A1. Values of k for n = 2 to 6 are all presented.

2. Throughout the manuscript, including in the conclusions, n is characterized as having an average value. I do not have a sense of how that average is calculated, and how much the experimental condition choices affect the determination of that average. Meaning, if 5 experiments had been run above the eutectic but still in the same temperature range, I expect the average n would be different. So it is hard for me to see how reliable the average is.

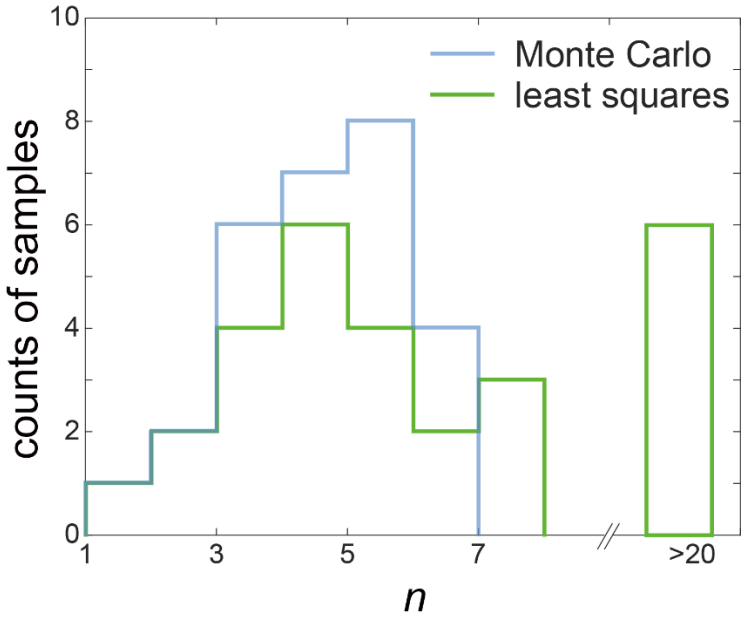
Reply: We apologize for the misleading usage of the word "average". In fact, 4 was used as a representative value for n. It is inevitable to acknowledge the presence of systematic errors in many stages of the experiment, and ultimately, this will also be reflected in the values of n. We agree with you that the average value does not hold significant meaning across different samples. Therefore, we have decided to modify the value of n to a range in the conclusion, rather than an exact average. However, we are still considering retaining $n = 4$ in our discussions, specifically for comparing the crystal growth constants (values of k) of samples with different compositions. As illustrated in Table A1 and newly added Figure A7, we have tested the goodness of fitting and calculated k values for different values of n from 2 to 6. We found that the variation in k values between different samples increased with increasing values of n. However, with different values of n, the order of k between different samples exhibits a consistent pattern. This also indicates that the choice of n values does not significantly affect the observed pattern of grain growth rates among different samples. We find that n = 4 gives reliable fits for all temperatures and impurity concentrations. As illustrated in the newly added Figure A8, the majority of fitted values of n lie between 3 to 6. Since the comparison for grain growth rate requires a common value of n, 4 is the best choice for further interpretation based on the available data set. We have revised the first paragraph of section 5.4, as described in the reply of comment 1.

We have added the following figures and captions to the Appendix.

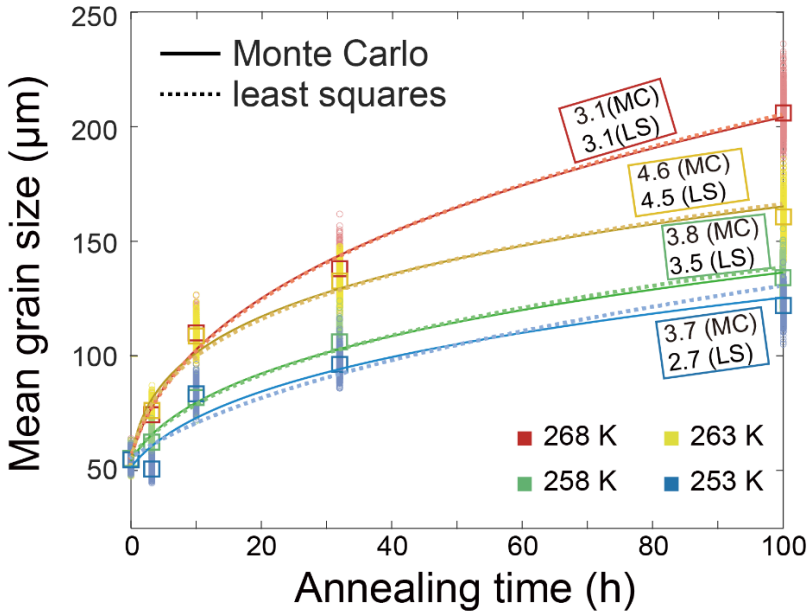


Appendix A7: Temperature dependence of the grain-growth rate constant k of KCl-doped ice (left column) and $MgSO_4$-doped ice (right column) calculated by fixing different values of n. (a) $n = 2$, (b) $n = 3$, (c) $n = 5$, (d) $n = 6$. Data for pure ice were plotted as black squares in both panels as a reference. Different colors represent

different impurity concentrations. The black line is the fit of pure ice.



Appendix A8: Comparison of distributions of *n* between Monte Carlo fitting (blue line) and least squares fitting (green line). The x-axis represents fitted values of *n* and the y-axis represents number of samples. Note that the x-axis is not continuous, jumping to greater than 20 after position of $n = 8$.



Appendix A9: A comparison of fitting curves obtained by Monte Carlo method (colored solid line) and the least squares method (colored dashed line), using pure water ice samples as an example. The fitted values of *n* are presented next to each curve, with results from Monte Carlo method marked by MC and least squares method marked by LS. Different colors represent different annealing temperatures. The squares are the measured average grain size and the small circles are the grain size distribution used

for fitting by the Monte Carlo method.

We have also revised the conclusion.

Line 471: "*The values of the best-fit grain growth exponent n lie between 2.6 to 6.2.*"

Line 473: "*The grain growth exponent obtained from pure ice samples ranges from 3.1 to 4.6.*"

3. The manuscript spends significant time discussing the grain growth constant, *k*, yet no conclusion point relates to this value.

Reply: We have modified the descriptions in the Conclusions.

Line 472: *"The grain growth exponent for pure ice samples ranged from 3.1 to 4.6. Compared to previous studies, we observed that the grain growth rate in the laboratory is lower than that of artificial pure ice; and the activation energy for grain growth is close to that of glacial ice and artificial ice containing smaller bubbles. This may be because some bubbles still exit in our samples, although they are compressed by the confining pressure to sizes below the detection limit of our optical microscope."*

Line 475: *"Above the eutectic point, grains doped with soluble impurities exhibit a higher grain growth constant, compared to pure water ice, i.e., faster grain growth rate. Below the eutectic point, ice doped with a specific concentration of soluble impurities manifests a smaller grain growth constant, compared to pure water ice, i.e., slower grain growth rate."*

4. In conclusion point #4, I suggest separating the observations from the inferred cause (i.e., that soluble impurities enhance grain growth above the eutectic and the presence of a melt phase; similar comment for below the eutectic).

Reply: Thank you for the suggestions. We have separated them in different points. The third conclusion point is the same as in our response to question #3. The inferred cause was added as a new conclusion point.

"The enhancement in grain growth at temperatures above the eutectic point could be attributed to the formation of a molten phase by the doped salt. The inhibition in grain growth at temperatures below the eutectic point may be attributed to the formation of hydrates at grain boundaries."

5. In Figures 6, 7, and 9, please describe how the best fits are calculated.

In the Methods Section we introduce how to perform the best fit. Here, we provide a more in-depth explanation of the Monte Carlo fitting process and comparisons with the outcomes from the least squares fitting. For each sample, we introduced 500 random variations to the average grain size and generated 500 grain-size data. These new grain sizes conform to a log-normal distribution with a standard deviation of 0.02. This distribution is an analogue of the actual grain-size distribution of the sample. Such that, instead of using just one average value for the fitting, a grain-size distribution was used. Examples of the generated grain-size distributions are presented in Figure A10 (a). We

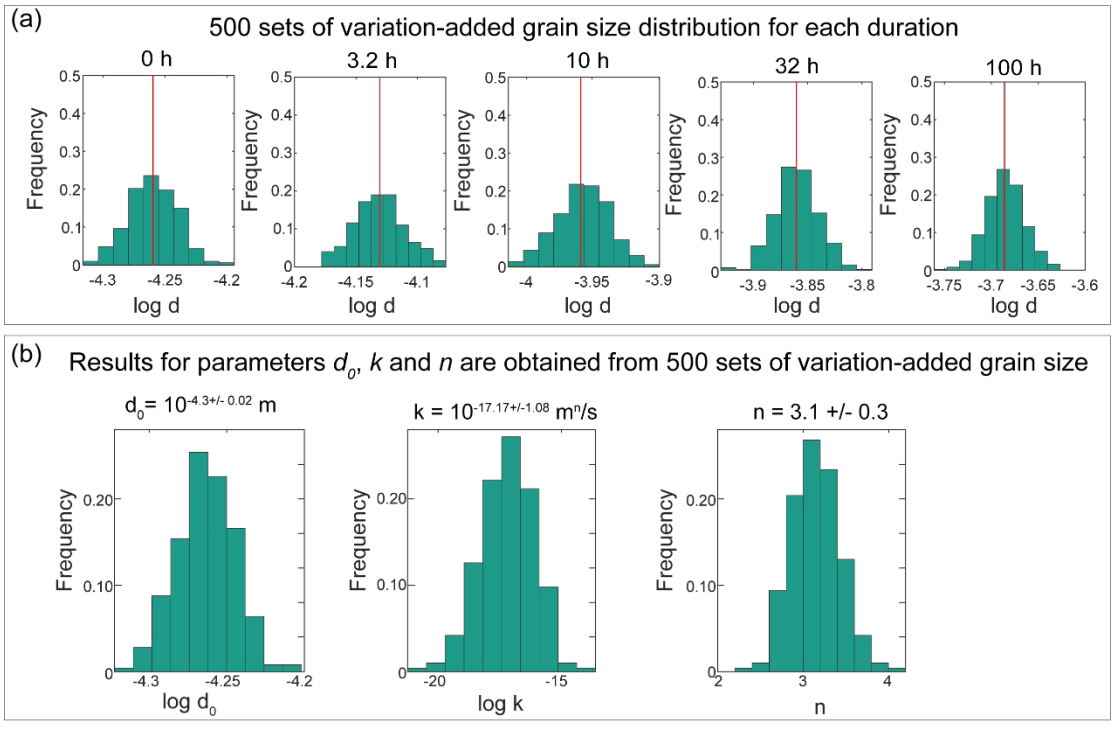
conducted 500 fittings based on variation-added grain sizes, resulting in 500 sets of outcomes. The averages of these fittings were then calculated to obtain the final fitting results. Figure A10 (b) show histograms illustrating the results of 500 fits for parameters d_0, k, and n. The title of each histogram presents the mean value of the corresponding parameter. This method reduces the effect of sample variations in fitting.

Figure A8 present the n values obtained through both methods, while Figure A10 illustrates a comparison between the two fitting curves for pure ice samples at four annealing temperatures. For most samples, the disparity in values of n obtained from the two methods' fitting is within 0.5. However, for samples with a value of n exceeding 6 fitted by the Monte Carlo method, there is a considerable difference in the outcomes of the two methods. The least squares method yields a larger n value, with the difference being at least greater than 1. Moreover, in certain samples where the 3.2-hour grain size data is lower than the initial grain size, the Monte Carlo fitting results range between 1.1 and 5.7, while the least squares fitting outcomes are typically much greater than 20.

However, these samples also exhibit substantial grain growth over longer annealing times, indicating that such large n values are clearly inconsistent with the grain size data. This underscores that, relative to traditional least squares method fitting, the Monte Carlo method can minimize the impact of the sample variations. Compared to the conventional least squares fitting method that relies solely on the average grain size, the Monte Carlo method does not change the original data, but incorporates the grain-size distribution into the fitting process, enhancing the reliability of the fit.

The above discussion is added as an Appendix section.

We have added the following figure and caption to the appendix.



Appendix A10: *An example illustrating the Monte Carlo method. Samples of pure ice annealed at 268K are used as examples. 500 sets of grain size data with variation were used for fitting, and corresponding fitting results were obtained. (a) For each duration, 500 sets of variation-added grain size data were generated by applying a range of random noise to the measured average grain size. In each histogram, the red vertical line represents the measured average grain size. (b) Results for parameters d_0, k and n are obtained from 500 sets of variation-added grain size. From left to right, three histograms of the parameters, d_0, k and n, were illustrated. The title of each histogram presents the mean value of the corresponding parameter.*