Response to AMT 2015-400 Referee Comments

S. Garimella[1], T. B. Kristensen[2], K. Ignatius[2], A. Welti[2], J. Voigtländer[2], G. R. Kulkarni[3], F. Sagan[4], G. L. Kok[4], J. Dorsey[5], L. Nichman[5], D. Rothenberg[1], M. Rösch[1], A. Kirchgäßner[6], R. Ladkin[6], H. Wex[2], T. W. Wilson[7], L. A. Ladino[8], J. P. D. Abbatt[8], O. Stetzer[9], U. Lohmann[10], F. Stratmann[2], and D. J. Cziczo[1]

[1] {Massachusetts Institute of Technology, Cambridge, MA, United States}

[2] {Leibniz Institute for Tropospheric Research, Leipzig, Germany}

[3] {Pacific Northwest National Laboratory, Richland, WA, United States}

[4] {Droplet Measurement Technologies, Boulder, CO, United States}

[5] {University of Manchester, Manchester, United Kingdom}

[6] {British Antarctic Survey, Cambridge, United Kingdom}

[7] {University of Leeds, Leeds, United Kingdom}

[8] {University of Toronto, Toronto, Canada}

[9] {V-ZUG AG, Zurich, Switzerland}

[10] {Swiss Federal Institute of Technology, Zurich, Switzerland}

Correspondence to: D. J. Cziczo (djcziczo@mit.edu)

1. Response Interactive Comment by P. DeMott

The authors would like to thank Paul DeMott for the suggestions and insights provided in this review. After careful consideration of the information provided, we would like to offer the following responses.

1.1 General Comments

1. ...Can some statement be made upfront that this paper will include a standard description and describe methods developed by this group to utilize the instrument in more advanced ways?

The abstract has been updated to reflect this distinction. The following is added: "This study presents a standard description for using the SPIN instrument and also highlights methods to analyze measurements in more advanced ways" (2:5-7). "Also presented (as an advanced use), is a machine learning approach…" (2:12-13). Also, subsections for Section 3.1 and Section 3.2 have been added to specify standard and advanced data processing methods. A note has also been added (13:3-5) to state this explicitly.

2. …As a minimum, the way that AgI is generated in this study, its size mode (size dependence is clear in much past literature), and some assurance that this was directly comparable to previous studies is needed.

Section 4 now includes "Also, polydisperse, dry-generated NX illite and AgI (Sigma Aldrich >99 %..." (17:12-15). The AgI particles were found to have a narrow mobility size distribution between 200-300 nm, so additional size selection was not performed.

3. …Details are suggested for addition…

These are addressed in the specific comments.

4. …validation results using AgI versus RH, as well as the homogeneous freezing results not to take the more ideal form I would expect…

The literature for comparisons with AgI are largely those found in Stetzer et al., (2008), which shows a similar correspondence that is seen in this study. The homogeneous freezing figure shows data from dozens of freezing experiments. An earlier version of the figure included temperature and RH uncertainties, but these made the figure very difficult to read. Overall, this is intended to show a transition to homogeneous freezing in the expected regime, and this version of the figure is the best way we found to show this with the number of freezing experiments considered. Please refer to the specific comment response below.

5. …The figure showing a line denoting the reverse flow regime needs to state the flow rate and pressure for which calculations are made…

This information is now included in the figure caption (35:4-5).

… It was found at that time that slowing the flow by half was needed to observe homogeneous freezing ideally…

This was not required for detection of homogeneously frozen ice in the SPIN detector. More details are provided in the specific comments below.

1.2 Specific Comments

Page 3, Line 7 – The Koop et al. reference is completely about homogeneous freezing. It goes with first part of sentence, but not last, which refers to heterogeneous nucleation.

This now reads "Droplets freeze homogeneously below temperatures of ~-38ºC, including deliquesced haze droplets which do so below water saturation at such cold temperatures (Koop et al., 2000)" (3:5-7).

Page 4, line 2 - Are you certain that lower chamber weight has been achieved in parallel plate designs? My impression is that the SPIN instruments could be heavier than most CFDC's. Can you state a total weight?

The SPIN chamber is 300 lbs. This sentence has been removed.

Page 4, line 10: I suggest adding Jones et al. (2011) to the list. This CFDC development was actually the first to mimic the original device designed by Rogers (1988), followed later by Saito et al.

This reference has been added (4:8).

Page 4. line 21 - Have you measured this ice thickness as 1-2 mm? We have estimated and reported 1/10 of that thickness in the CSU CFDC's based on liquid meltwater measurements. If truly this thick, how much of the inter-plate distance is filled with ice, and do your calculations account for this? It would thus be useful to also report how wide the plate separation is in the SPIN (here, not only in the Figure caption). A horizontal distance scale on the bottom axis of Figure 1 would be useful, and I would presume that such thick ice layers should be shown (and again, accounted for in the calculations).

This estimate was based on measuring the volume of water that was depleted from the water reservoir during the icing process. Measurements of the melt water volume were consistently lower than this type of measurement (presumably due to the depletion of the ice layer after running experiments), so this is why we choose to report our ice thickness in this manner. This is now reflected in the text (4:19-20). We also investigated (qualitatively) the ice thickness with an endoscope camera and found that the thickness in the bottom (evaporation section) of the chamber to be much thicker than the in the rest of the chamber. Since direct observation of the ice layer did not yield quantitative estimates of the ice layer thickness, and melt water estimates suggested a much thinner ice layer than that from measuring the difference in the volume of water in the reservoir after icing, we choose to give this estimate of the average thickness. A mention of this qualitative inspection has been added as well (6:20-22). We also examined the sensitivity of the lamina calculations to ice layer thickness and found that assuming a negligibly thick ice layer (which is the case at the top of the chamber where the initial nucleation occurs) yielded results that best matched expected results (including water breakthrough and homogeneous freezing conditions). As the ice layer gets thicker towards the bottom of the chamber, the corresponding lamina RH values decrease, but the RH experienced by particles as they activate at the top of the chamber was found to be captured best by using a negligibly thick ice layer in the calculations. This is now specified (12:1-3).

Page 5, line 16 – How is the refrigeration controlled in the evaporation section of the chamber? Figure 3 only seems to show cooling each wall continuously at one temperature. What constant temperature are the evaporation section walls held at? That is, at the warm or cold wall temperature?

The evaporation temperature is regulated using the same refrigeration system as the warm wall, but with its own solenoid valve to control the amount of refrigeration. It also has dedicated PID heater control for each wall independently and one thermocouple measurement for each side. This setup allows for the evaporation section temperature to be set to the aerosol lamina temperature. This has been specified (10:10-12).

Page 6, line 17 - How is sheath air drying achieved? Can you comment on whether or not a one-way flow of dry air could replace the re-circulated sheath from in the system shown in Fig. 3? CFDC's have been known to be operated with one-way flow of with recirculation.

Are both practically possible with the SPIN? Just asking as a scientific point, not with regard to any perceived deficiency.

Sheath air drying can be achieved using any type of inline dryer, but molecular sieve is typically used. This has been specified (7:1). We have also often run the sheath air with one-way flow using both dry lab air and dry air cylinders. In general, both recirculating or one-way flow are possible.

Page 6, line 22: The knife edge simply places the sample flow initially in the center of the chamber. The limited range of temperature and supersaturation then depends on maintenance of laminar flow conditions, correct?

The knife-edge inlet is a necessary but insufficient feature to maintain a limited range of thermodynamic conditions. The text (7:6-7) has been updated to this effect.

Page 6, line 23 – It might be useful to point out that the 4:1 growth to evaporation ratio here is much greater than the 2:1 ratio used by Rogers et al. (2001). But this does bring to mind that you should state all of the dimensions in writing here. Especially, it is important to know the cross-sectional dimension of the actual interior chamber region. It is shown as 3 cm in Fig. 2, but the plate separation is 1 cm correct? So with the ice added, the actual separation is 8mm?

This information is now included in Section 2.2 (6:20-22).

Page 6-7: Not addressed anywhere in this section is how walls have been treated to retain ice. I realize that this is found in publications written about chambers that the SPIN design followed, but it is another small detail that deserves discussion somewhere, as stability of ice surfaces is one of the critical issues for these types of instruments.

The wall surface is sandblasted aluminum, but it is otherwise untreated (unlike the ebonizing for other chambers). This lack of treatment does not seem to prevent the formation of the ice layer in the chamber, but the sandblasted aluminum has been specified (6:18-19).

Page 7 – Is the OPC quite similar to the DMT CASPOL (Glen and Brooks, 2013) or is it very different?

Though also manufactured by DMT, the CASPOL instrument is sufficiently different from the SPIN OPC (e.g. different backscatter angles, lack of forward scatter detection in SPIN, and different size range), that direct comparison was determined inappropriate by DMT.

Page 8, line 5 – Concerning the saturation of counting at 3900 per cc for a 1 LPM flow rate, is there any means to correct for "live time" as done for some multi-channel analyzers?

This functionality does not exist with the current design, unfortunately.

Page 8, lines 9-15 – The question here is what is present as actual functionality with a SPIN versus what is possible via a user's initiative, and has been demonstrated? . That is, should this say "potentially allowing. . ." and ". . .much of the chamber operation could conceivably be performed remotely"? The impression is given that full remote operation is already possible, but that is not demonstrated in this paper. These things are all possible for any group with some initiative and technical skills (and some have done so), so without thorough demonstration in this paper, I think that such statements should be avoided.

Technically speaking, full remote operation of the chamber is possible in the out-of-the-box configuration (after installing something like TeamViewer). The only real limiting step is that the water reservoir will eventually run out of water and cannot be re-filled remotely. In practice, we have avoided icing the chamber remotely to avoid flooding the detector in the unlikely event of a system failure, but this too is possible. Given a sufficient water supply, the entire chamber operation can indeed be performed remotely. A sentence has been added to clarify this (8:15-19).

Page 8, line 23 – Can you say what type of dewpoint sensor is used, its accuracy, and how well can it resolve a dewpoint temperature of -40◦C? This typically requires an advanced sensor.

The sensor information has been added (9:4). The sensor is rated to -80ºC with accuracy of ±2ºC.

Page 9, line 1 – When you say the chamber is cooled to icing temperature, you mean both walls and both wall sections (growth and evaporation) cooled uniformly I assume?

Yes, this is what we mean. This has been clarified in the text (9:6-7).

Page 9 – This might be another place to mention how the walls are treated to be "wettable," and what they are made of, materially.

This information was added to the instrument description section (6:18-19).

Page 9, lines 6-7 – Is this "dwell counter" time the total time of water in the chamber or the total time that that water remains prior to it being quickly removed? In other words, how long is water actually in the chamber? The time mentioned is quite short for other CFDC's, yet the ice thickness mentioned is much thicker, so just looking for details that users may wish to know.

The dwell counter is the amount of time that the chamber is full of water (as specified in the text 9:12-13). The total time any water is in the chamber is closer to 30 seconds. The fill sequence itself is very consistent, as measured by the difference in the amount of water in the reservoir before and after filling. This is clarified in the text (9:18-20).

Page 10, line 5: Doesn't this describe the same effect described above as the "back-ground"? It is not clear. Also, please add that this is uniformly more frequent at higher

supersaturations "for the SPIN." Such a generalization may not apply to all CFDCs for this process that is still rather poorly understood, and may vary depending on the method used to obtain wall wettability.

This has been clarified, and the supersaturation dependence has been specified to the SPIN chamber (10:14).

Page 10, lines 6-8 – The statement that frost background of the same order as INP number concentrations means that the INP data will include a significant artifact is understood, but this is likely quantifiable. It certainly does not require that the frost be only 0.1

This has been simplified now as an example "For example, if a laboratory experiment with 10 INP cm-3 were to report activated fractions at the 1 % level, it would require a background of no more than 100 counts L-1." (11:4-6). An exemplary background level is shown in Figure 11.

Page 10, lines 11-12 – It is useful and could be important to report what a typical ramp rate is in terms of cooling rate or d(RH)/dt.

This has been added (10:19).

Page 10, line 13: How long are the typical filtered air periods in order to provide a sufficient measure of background frost numbers?

The filtered air periods are typically 3-5 minutes. This has been specified (10:17).

Page 11, lines 13-15 - This seems a variation in the Rogers approach, which may be worth noting. Rogers did not consider along-wall differences, just average wall temperatures in locating the lamina and the flow profile. I am not sure how one uses this information, since the lamina position and velocity profile calculation usually requires assumption of a single value along the walls. What is actually reported then to associate with an INP measurement (for sample conditions and residence time)? Richardson (2009) did consider along wall

temperature differences, but he needed FLUENT simulations to resolve the importance of this factor in defining and conditions and affecting freezing.

This approach extends the Rogers approach using our ability to map along-wall differences. We simply perform the calculation at multiple locations, and include the variability in our reported values. This has been clarified as "extending the method from Rogers (1988)" (11:22). We also use FLUENT in a later section to look at the behavior of perfectly isothermal walls.

Page 13 and Figure 6 – While much of this comment is editorial, I place it here because I feel that this figure requires better explanation. Can the scales and labels be made larger?

This has been done.

Please explain the "Size" units or otherwise please explicitly remind of the form of the size parameter. If the standard definition, I could not get the magnitudes correct, or else the wavelength is different than listed for the laser. Also, if the lower size threshold is 0.4 micron diameter, these figures really show particles growing to two orders of magnitude larger in size within the SPIN? Perhaps better, why not show this figure with actual geometric particle size? Finally, has this exercise been done for specific size ranges of INPs, and what are the RH conditions associated with part b (should be stated)?

The units for each parameter are raw intensity counts, and this has been specified in the caption (36:3) and the text (14:19). For use in the machine learning algorithm, it is best not to pre-transform the data to different units, especially since the power law approximation between counts and diameter is not exact. Since the algorithm is run on the raw intensity counts, the figure is intended to reflect this. This exercise is performed on all of the SPIN data in this study to report activation results and corresponding uncertainties. The RH chosen for part b depends on the specific experiment, but in this example part b is at -40ºC and S_{ice}~1.4.

Page 14, Figure 7 – This figure may bear more discussing as well. It may be helpful here to reiterate that the evaporation region reduces "droplet" fractions. Otherwise, one may wonder why full droplet activation, or anything close to it, is not achieved in this experiment at higher supersaturation, One wonders anyway, since once droplet breakthrough occurs, it seems unusual not to see all droplets retaining sizes well above the OPC lower threshold size (e.g., DeMott et al. 2015). Is this feature due somehow to droplet evaporation after exiting the SPIN and entering the detector region (i.e., due to heat transfer)? The supersaturation seems too high to think that the AgI is not fully activating as droplets at some point.

The role of the evaporation section in reducing droplet fractions is now noted (16:8-9). Though a detailed discussion is beyond the scope of this paper, the result reported in DeMott et al., 2015 of only 1/3 of the particles experiencing the reported supersaturation is consistent with this figure. If normalized to this correction factor, the activated fractions are perhaps closer to what one might expect (i.e. ~2/3 of particles actually in the lamina are activating into ice or droplets). Since previous AgI investigations have not probed this area of the parameter space (where a depolarization detector is required to distinguish phase), it is difficult to know what fractions to expect. Overall, this figure is just intended to show that such a discrimination is possible and what results might look like.

On another matter, when you say validation accuracies, what is the basis for validation?

Validation as it pertains to these techniques was intended to refer to statistical cross validation or out-of-bag classification certainty. For the sake of clarity, "validation" has been changed to "classification" in this context (16:15).

One also would like to know if there is any particular lower ice number concentration value for which this machine learning procedure is easily used. For example, data are shown at INP number concentrations as low as 1 per liter in Fig. 7. Are these values already corrected for background frost? The method seems to be most reliable for INP number concentrations of several 10s per liter.

In the example experiment shown in Figure 7., the error bars that are off the scale show where ice / droplet concentrations are statistically indistinguishable from zero. In this example, this corresponds to 10's per liter, and the values are already corrected for background frost by quantifying how many frost counts are classified as ice crystals by the algorithm (16:9-11).

For INP concentrations in much of the mixed-phase cloud regime, how long would one need to integrate ambient data to reliably use this method if INP number concentration is say even 4 per liter? Would this become problematic when droplet breakthrough occurs and nucleated ice concentrations are quite low? Any statements on the useful range of this otherwise elegant method would be appreciated. Finally, is it possible to speak about the use of a constant particle size that could distinguish ice versus aerosol and drops, perhaps as a function of temperature and RH? Just looking at the data in Figures 6 and 7, it seems possible to use this information for many instruments with similar residence times.

This method was chiefly developed as a way to process laboratory CFDC data as opposed to field data. Though it could be extended to do so, this is beyond the scope of the present study. An advantage of this approach is the ability to quantify the probability that the signal one would see is real vs. frost. For example, this could be first approached by constructing a machine learning classifier aimed at distinguishing frost from real ice: such a classifier and the corresponding estimate of classification accuracy could then be used to reduce the uncertainty associated with

INP number concentration estimates. The concept of integration time as it applies to something like Poisson statistics would be less important, since this approach would provide the probability that each detected crystal is either frost or real data.

This general ML approach could certainly be used in for other instruments with size-only data. Figure 6 shows a Gaussian kernel support vector machine approach, but a linear SVM that uses size only would find the optimal size (by maximizing the margin between ice and aerosol) to use to distinguish the two classes. Text motivating this future work has been added (16:16-21).

Page 14: How likely is it that AgI is in exactly the same form and size as for other studies referenced in the following results? It surprises me, considering past cloud seeding related literature regarding AgI activation, and how even minor contamination can lead to disparate results.

We are unable to provide a quantitative estimate of the contamination level in previous studies (though the sample used in this study is specified at >99% purity). We do recognize, however, that this is a real issue in the ability to use a material as a calibrant, which is why we chose a commercially available sample (with traceable batch #, etc.). Similar issues of differing contamination levels have been known to occur with dust calibrants such as ATD and illite, so we chose to use AgI because it was also used to characterize the ZINC chamber (Stetzer et al., 2008). A note to this effect has also been included (17:13-15).

Page 14, line 19 – Is this validation or calibration?

This change has been made to "calibrations" (17:3).

Page 15, lines 4-6 – The use of ammonium sulfate here is a classic calibration study for existing CFDCs, but it does lead me to ask about the expectation that sulfate will act only for homogeneous freezing in the lower temperature regime. Heterogeneous ice nucleation

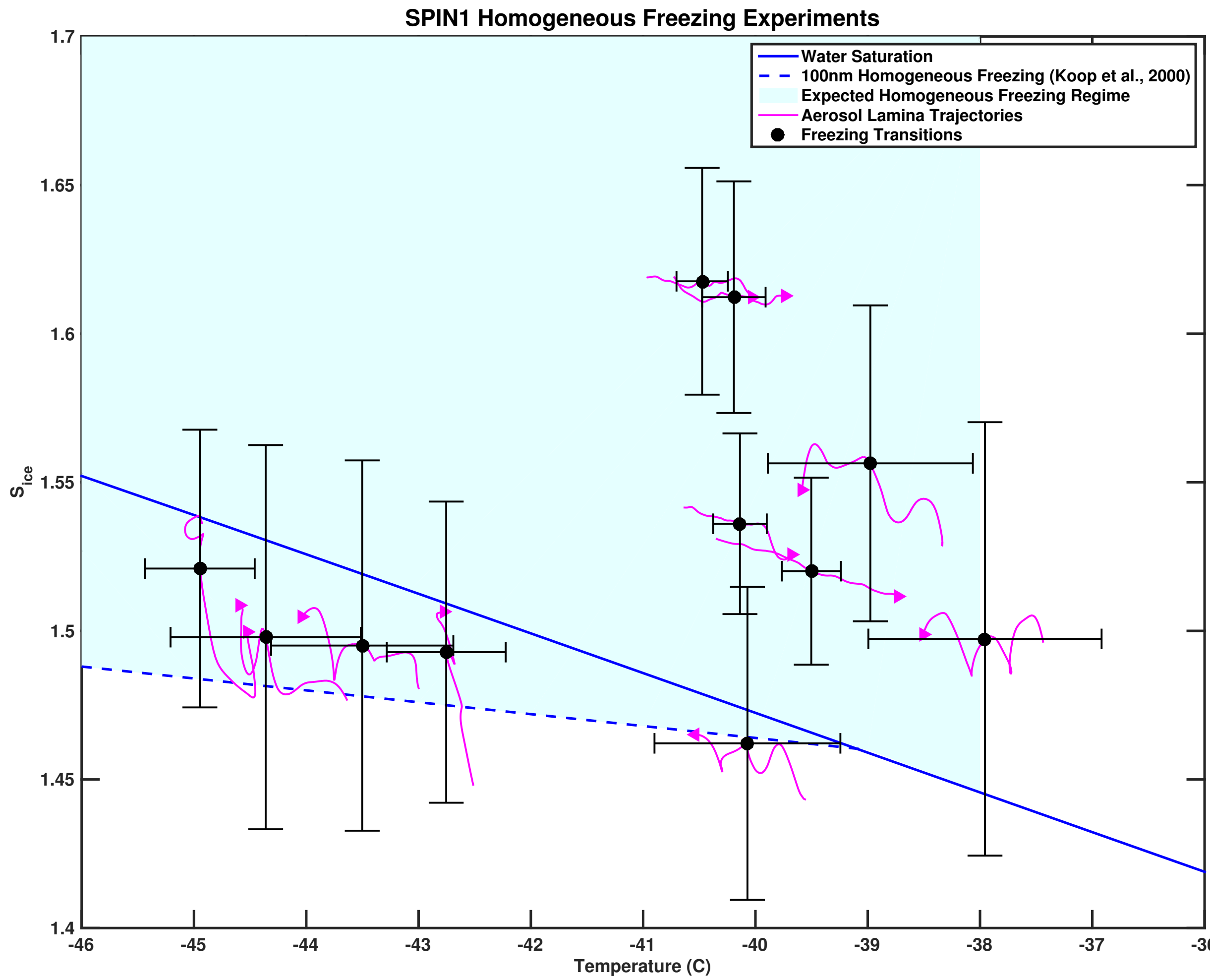
activity has been reported for crystalline ammonium sulfate aerosols (Abbatt et al., 2006). It might be important to state why this is not expected in this case. Also, for size selection, can you say the proportion of multiply charged particles? This could affect the homogeneous freezing transition, which I suggest next requires more detailed explanation and discussion.

Abbatt et al., (2006) report that depositional freezing of solid ammonium sulfate becomes increasingly important (at the 0.1% - 1% activation level) at temperatures colder than ~ -45ºC, which is coldest temperature reported in this study. Though this could cause early ice formation at colder temperatures, it is not an important factor for the larger activated fractions considered in this study (17:23-18:2). Koop et al., (2000) suggest that the difference in supersaturation for homogeneous freezing for solution droplet formed on 200 nm particles and 500 nm particles is ~1%. The resulting uncertainty is much smaller than the other measurement uncertainties of these experiments (typical uncertainties in supersaturation in the SPIN chamber are ±~5%), so this effect is not considered (18:8-11).

Page 15, lines 8-9 – The broad sweep of Figure 9, which is nice as a first inspection based on a large suite of measurements, nevertheless suggests to me that the instrument is not reproducing conditions of homogeneous freezing as clearly as observed by some other CFDCs. This transition and the definitive statements made in this section bear some quantitative inspection. What would a line drawn vertically in the figure (e.g., a single experiment) indicate for the active fraction of particles as a function of RH? How would it compare to predictions based on water activity theory for freezing? This is a standard on which to compare, such as shown in DeMott et al. (2009), Richardson et al. (2010) and some other past literature. Stating clear agreement with Koop et al. (2000) suggests the need for

such a presentation. Furthermore, does the figure show false detection of ice at temperatures warmer than homogeneous freezing?

The figure below is an early version of Figure 9 with fewer experiments shown. The magenta arrows show the trajectory of the aerosol lamina. In general, freezing of droplets at the 1% level occurs between -38ºC and -40ºC for temperature ramps, and it occurs at levels similar to those predicted by Koop et al., (2000) for humidity ramps. When increasing temperature, the disappearance of ice crystals occurs at colder temperatures than those at which ice appears when decreasing temperature. This behavior is similar to the general hysteresis observed when comparing humidity up ramps and down ramps. We chose to show current version of Figure 9 because it shows data from more experiments, but we have provided the version below to clarify the uncertainties associated with such measurements and to show that homogenous freezing occurs at expected levels within this uncertainty. Also these uncertainties have been added to the figure caption (39:7-8).



Page 15, lines 15-21 - I must say that it is not sufficient simply to say that the homogeneous freezing behavior is "captured" by the SPIN, versus requiring much more detailed inspection in future studies. As I said above, it does not look ideal (activated fractions do not increase orders of magnitude as shown in other CFDC studies of homogeneous freezing), but this is difficult to tell in this mapped figure. A detailed comparison will be needed, and this suggests an additional figure. What are flow rates and residence time in the chamber? What is efficiency of ice detection below water saturation (i.e., are ice crystals predicted to grow to sizes that will be easily distinguishable via size or polarization)? Noting the discussion in Richardson (2009) and Richardson et al. (2010),

especially the need to reduce flow rates in shorter-length CFDCs such as SPIN in order to detect (at least by size) the presence of ice crystals growing above about 1 micron at temperatures below -40C, I wonder if this complication is present in any of the data shown here (i.e., explaining the need for high supersaturations to realize high activated fraction)?

Please see the above figure and corresponding explanation. The flow rates for these (and all other experiments in this study) are 10 Lpm sheath flow + 1 Lpm sample flow, which leads to ~ 10s residence times in the chamber. The homogenous freezing transition causes the fraction of ice crystals to increase from < 0.1% to > 60% in the expected region of phase space.

Page 16, line 3 – Editorially, I suggest removing "significantly" here. A few percent above water saturation is not much in order to detect immersion freezing (DeMott et al. 2015 and references therein). It seems that the depolarization detector will be required in many cases. Hence my earlier question about the lower level of sensitivity for ice amongst a field of many liquid particles using this method. It would seem useful to design the instrument with a longer evaporation region.

This change has been made (18:17).

Page 16, line 7 – AgI aerosols only provide a benchmark if they are carefully produced in exactly the same manner based on a great deal of past literature. Is agreement expected or fortuitous?

Please refer to the previous response regarding differing levels of contamination between calibrants and the choice of AgI to compare to Stetzer et al., (2008).

Page 16, lines 16-20 – I think that this statement should be omitted. It is not true. Other portable INP chambers have simply switched to two-stage refrigeration systems to achieve lower temperature conditions in the past. See, for example, DeMott et al. (2003) (already

referenced here) and Prenni et al. (2007), ground-based and aircraft examples of using a refrigeration system to extend to lower temperatures. It really is not a difficult change for instruments that can be flexible with the configurations used.

The second half of this sentence referring to other chambers has been removed (19:8-9).

Page 18, lines 2 to 5 – Here I suggest referring to agreement of your results with Richardson et al. (2009), who quantified uncertainties in a similar manner.

This reference has been added (21:7).

Page 18, lines 7 to 9 – Considering comments above, I consider this a point that is as yet unproven.

Please see previous responses regarding homogenous freezing.

Page 18, line 14 – demonstrate, not "validate", in my opinion.

This change has been made throughout.

Page 18, lines 16 to 19 – This is a judgement statement that I consider inappropriate for a scientific article. The introduction of this instrument commercially is an important milestone. Nevertheless, the lowered barrier involves a financial investment that is about 4-5 times the cost of a CCN instrument. An individual building an instrument could do so cheaper, and in the process obtain the invaluable experience and investment that my colleagues here have demonstrated is needed to understand and effectively interpret such measurements. It is certainly true that the availability of a commercial instrument could "potentially" lead to an increase in temporal and spatial coverage of INP measurements. This remains to be seen.

This has been revised only to draw similarity to introduction of the SPIN and CCN as *commercially available* chambers. "Potentially" has been added as well (21:15-18).

Page 18, lines 21 to 23 – The full performance of the depolarization detector would seem to be worthy of additional study. A comparison of ice detection versus other instruments and a standard such as a cloud chamber simulation would be useful for the future.

A statement to this effect has been added. (21:23-22:2)

Page 19, paragraph 1 - The discussion here should reflect any responses to my questions and comments above. Again, utility has been demonstrated, not validated quantitatively.

The language has been modified to remove "validation/validate" throughout.

"Page 19, lines 15 to 17 – I suggest striking the statement regarding "SPIN's availability. . ." in preference to something to the effect that "The commercial availability of such a device may allow for increased coverage of INP measurements that will help constrain. . ."

This change has been made (22:17-19).

In preference to some of this material overall, I wonder if you could say what future needs for demonstration and validation remain (I would say that quantitative intercomparison is needed), and what challenges may remain?

This section now motivates such future work (21:23-22:2).

Figure 5 - By "degree of flow reversal" do you mean the fraction of the flow that is reversed? You need to state a total flow and pressure for these calculations right? What would it be at 300 mb, or something more typical of cirrus sampling conditions? Also, it could be useful to superimpose the lamina position on these flow figures.

This figure is to demonstrate the concept of flow reversal as opposed to show an extensive characterization of how it depends on all chamber variables. To this effect, the flow conditions pertaining to this plot have been added and "degree" has been changed to "fraction" and explained more thoroughly (35:1-11).

1.3 Editorial notes

Page 3, Lines 15-19 – I understand the flow here in saying that CFDC's have been used for both laboratory and field studies, but to be clear, such instruments were designed not to simply detect the conditions for ice nucleation in the laboratory, but to quantify the number concentrations active under a variety of conditions far beyond onset conditions or for specific nucleants. And they were designed for field measurements specifically, from the start. So the statements here are not historically accurate. I suggest, "Several types of instruments have been developed to measure the efficiency of heterogeneous nucleation of cloud droplets and ice crystals. Many of these have applicability for measurements in the laboratory, as well as intended application for field observations."

This change has been made (3:16-18).

Page 9, lines 21-22 – Just to advise that use of the term "time dependence" could be confusing, considering stochastic freezing time dependence that is not easily assessed with this type of instrument, and is not what is being referred to here. How about "temporal changes" of INP concentrations?

This change has been made (10:7).

Page 11. Line 17 – The Kulkarni and Kok reference seems incomplete. Is it a report? Is it freely available?

This is a freely available PNNL report. The location of PNNL has been added for completeness, according to the Copernicus citation specification for reports (27:11).

Page 17, lines 9-10: Are the top 13 thermocouple temperatures then averaged for reported SPIN conditions?

Yes, this has been clarified (20:6-7).

Figure 1 - Re: see above comment on plate separation, in that it could be useful to add a horizontal distance scale under temperature.

The width has been specified on the figure and in the figure caption (31:3).

Figure 9 – can the typical number concentrations of ammonium sulfate be stated in the figure caption?

This has been added (39:7).

2. Response Interactive Comment by Anonymous Referee #2

The authors would like to thank Anonymous Referee #2 for the suggestions and insights provided in this review. After careful consideration of the information provided, we would like to offer the following responses.

Itemized points:

page 3, line 2-5: Aerosol particles are not required for nucleation, rather because they are ubiquitous they assist nucleation. In a system with few enough particles homogeneous nucleation could proceed. The way in which particles facilitate the freezing is by changing the free energy barrier to phase transformation. Classical nucleation theory uses bulk thermodynamics to describe the free energy change as a surface/volume competition that is modified by a kinetic pre-factor. The described "mechanisms" are simply phenomenological descriptions of how this proceeds. Thus some thought should be put into accurately phrasing the first two sentences.

The accuracy of the phrasing has been improved by stating "Aerosol particles facilitate the nucleation…" (3:2).

page 3, line 7: The Koop et al. (2000) reference is used throughout the manuscript and also within Figs. 5, 9-10. It should be clear that the Koop theory and curve refers to homogeneous freezing in liquid solutions. Thus, the applicability of that theory to freezing in the atmosphere is not always clear – in their manuscript Koop et al. (2000) ignore geometry and consider only large droplets. A very different curve represents homogeneous freezing from the vapor phase – for examples see Fig. 5 in Murray and Jensen (2010) and Fig. 2 in Thomson et al. (2015). See also the discussion of the relevant figures below.

Homogenous freezing, as it is tested in the SPIN chamber for the purpose of this study, is on deliquesced ammonium sulfate haze droplets and supercooled ammonium sulfate cloud droplets, which is why the comparison to Koop et al., (2000) is made. The distinction that it applies to the freezing of solution droplets has been made throughout.

page 3, line 8-12: Although the IPCC reports are highly valuable and useful references, it seems as though citing the recent reports 2x in the papers primary motivating paragraph (Boucher et al., 2013; Stocker et al., 2014) may ignore primary source literature. Perhaps the points being made by the authors would be better emphasized or bolstered if they were slightly expanded and also referred to primary source material, where INP, mixed phase clouds, and/or cloud microphysics are concerned.

Additional motivating references have been added (3:9-13).

page 3, line 21: “RH’s" is introduced here but not defined until later on page 4. The abbreviation should be defined at first use and used thereafter.

The acronym is now defined here instead (3:22).

page 4, line 1-4: The sentence, “Development of parallel plate chamber geometry has simplified several technical aspects of the chamber design (e.g., lower chamber weight, less complex machining, and simpler refrigeration plumbing than for the cylindrical geometry) at the expense of edge effects (i.e., deviations from ideality at the chamber edges).", is confusing and relegates seemingly important details to the parenthetical. I suggest rephrasing, removing the parentheses, and adding detail to construct a more holistic picture of what is meant.

Combining multiple referee recommendations, this section has been simplified to indicate that the ZINC and PINC are CFDC chambers with parallel plate geometry (4:3).

page 4: The references to the ZINC and PINC systems and papers raise the question of, how much of SPIN can be understood by referring to the literature that describes these earlier systems? Are there things that can be taken out of those papers that are valid for SPIN? I think it is important to make explicit what is SPIN specific versus universal for parallel plate (vertical) CFDCs and/or CFDCs generally.

The SPIN design, including geometry inlet design etc., is adapted from these chambers (4:9). What is intended to be taken from this comparison is that laboratory results from the SPIN chamber are expected to be able to repeat those from previous chambers (as discussed in results section).

page 4, §“Theoretical principles": This section reads more like operational principles/procedures. In its current form it would be better suited to come after the design description. If the preference is to keep it in its place then it needs to be re-written in a more general format. However, I would suggest that the authors both outline the general operating principle and describe a suggested user protocol for typical experiments – not necessarily in the same section.

This section is intended to introduce the operating CFDC principles and highlight how these pertain to the SPIN chamber in particular. It introduces important concepts for the CFDC chambers in general (such ash how supersaturation is created, how droplets are evaporated, what droplet breakthrough is, etc.) so we feel it should precede the design section. However, per the referee’s request to ensure this section is general, the aspects that pertain only to the SPIN chamber have been relegated to parenthetical selections. The rest of the text is now general to all CFDC-style chambers. The section title has also been changed to “Operating principles” for clarity (4:16).

Throughout this section and the following two the manuscript lacks a clear trajectory and often uses imprecise language. Some effort should be made to more clearly delineate between the theoretical working principle, what could possibly be done in SPIN given the intended engineering, and what in fact is done in (with) SPIN.

For example,

page 4, line 19-20: "Controlling the temperature and RH is accomplished ...." It seems to me that temperature control has nothing to do with icing the chamber walls. The temperature is controlled by refrigeration and heaters, the saturation condition is established by icing the walls and controlling the temperature gradients. Likewise (line 23), "The two walls are held at different temperatures ...." That is a statement of operational protocol, because within some range it appears you could set the temperatures to whatever you want, they could even be equal. It would be better to say something like, 'In order to establish the necessary vapor supersaturations for nucleation the ice coated chamber walls are held at different temperatures ...'

The language in this section has been made more precise and less procedural (4:17-6:2).

page 5, line 6: How well are the particles constrained within the sample flow lamina? Garimella had a presentation at the 2015 AGU Fall meeting suggesting that this is a source of uncertainty (Garimella et al., 2015). That presentation or a manuscript in preparation could be referred to, to say something general about this.

An adequate answer to this question is beyond the scope of this study, but a sentence with this reference and the need for future work has been added to Section 5 (20:13-14).

page 5, line 12-end: The effectiveness of the evaporation section is largely dependent on the relevant residence times, perhaps this could be addressed? Isn't the critical droplet size

related to the time a droplet might have to evaporate? Is it important droplets completely evaporate? Or can they simply decrease in size to below a threshold?

This has been clarified: "Depending on their residence time in the evaporation section, droplets over a critical size will not evaporate completely and are large enough to be detected by the OPC" (5:19-21).

page 7, lines 6-10: Both the aspherical morphology and depoloraization associated with ice crystals come from the anisotropy of the ice as a material. Ice crystals have fast and slow growth directions (that lead to their macroscopic morphology, e.g. Cahoon et al., 2006; Wettlaufer et al., 1999) and are also optically birefringent (e.g., Thomson et al., 2009; Lekner, 1991, 1999, and references therein). Both aspects are important.

This language is now more precise and references have been added (7:16).

page 7, line 16: "LabView software is used for instrument control and data acquisition. The SPIN software program consists of several different loops and sub-programs and allows for significant automation during operation. The Control Program starts and stops the other modules, updates the displays, controls the instrument set points, watches the alarms, and otherwise supervises the operation of the entire system." Given the lack of specificity much of this text seems superfluous, it would suffice to say: The SPIN instrument is operated and controlled via a LabView master control program. (or alternatively add pertinent information)

This section has been simplified and condensed (8:1-13).

page 8, line 5: What happens when particle counts approach this level? What is ignored? Does this bias results? Is it recommended that a dilution flow is added to keep raw particle counts below this threshold?

Though the instrument is never operated in these conditions due to data loss, there is no preferential bias for detecting certain particles. After the first ~64000 particles have been detected in a given second, no more data can be buffered, so this effect simply caps the number of detectable particles per second in processing (8:9-13).

page 8, lines 6-15: Is the entire icing scheme fully automated as one could believe from this paragraph?

As long as there is power to the system, water in the reservoir, and a stable internet connection, the system technically can be run fully remotely. We have generally avoided doing so, however, out of lack of necessity and not wanting to take the risk of something like a power failure occurring. What is possible remotely has been clarified (8:15-19).

page 9, line 4: "icing sequence" Quotation marks are generally used to set off material that represents quoted or spoken language. At times they are used to show sarcasm. Here neither appears to be the case – they seem to be indicating LabView program components. If these terms/phrases are being emphasized it would be better to use a different type-setting tool (bold, italics, etc.). The same mistake is made throughout the manuscript, eg., "ice dwell counter," etc.

Instances of erroneous quotation have been removed. Italics are now used instead throughout.

page 9, line 23 - page 10 line 2: "For the former, diverging the wall temperatures increases the chamber supersaturation,... and ramping both walls..." The uses of 'diverging' and 'converging' are awkward in this sentence, I suggest rephrasing (for example, 'increasing the temperature gradient between the walls'). They are correctly used a bit further in the text. Also, the walls are not 'ramped'. The temperatures are increased and decreased or temperatures are ramped.

This paragraph has been clarified with the suggested language (10:8-10).

page 11, line 12: extra comma after “Rogers"

Fixed (11:11).

§3.2 Data processing and methods: Given the audience of AMT this section lacks significant detail and some concentrated effort needs to go into a major re-writing. While machine learning and other data analysis techniques may be known to some, I do not think one can assume they are part of the atmospheric measurement community vocabulary. Detail and references must be added to the section, and if deemed excessive for direct inclusion I suggest that fundamental information now not included could become part of a supplement. For example in Figure 6 data from a GMM-KDE is presented – could a supplement include a step-by-step explanation of the analysis process, perhaps including some idealized system (where the separation is more clearly bi-modal)? These techniques seem to be a powerful tool, but it is hard to appreciate in the limit of the single example and poor explanation.

This section has been revised and separated to specify advanced use. We agree that the referencing in this section was lacking, especially for the intended audience. A more transparent explanation of the exact procedure used and the (now referenced) sub-steps are now provided (14:5-15:15). In particular, a reference to the basic concepts of supervised machine learning (including support vector machines) is provided (Mohri et al., 2012), the original description of KDE’s is now referenced (Rosenblatt, 1956; Parzen, 1962), GMM’s are now referenced (McLachlan and Peel, 2000), the specific KDE algorithm used in this study is referenced (Kristan et al., 2011), and bootstrap aggregation is referenced (Breiman, 1996). Furthermore, the pre-existing MATLAB functions used to perform this processing are also referenced (14:1-2) (please note that web URLs are provided based on how MATLAB code is referenced in other Copernicus papers).

page 12, line 1: "supervised machine learning" is introduced without a reference. Include a reference so that those unfamiliar with this have a standard text to which to refer.

A reference has been added. Please see above comment.

page 12, lines 3-4: "The also require fewer assumptions to be ...." Machine learning requires fewer assumptions than what?

This now reads "It also requires fewer assumptions to be made about particle classification and allow more flexibility in experimental design" (14:3-4).

page 12, line 6: I suggest rephrasing to say, "... historically been analyzed using post-evaporation section particle size as...."

This change has been made (13:6-7).

page 12, line 10: I suggest rephrasing to say, "... than the ice size and that droplets above that size do not survive..."

This has been revised (13:10-11).

page 12, line 15: By "efficiently" do you mean completely?

This change has been made (13:16).

page 13, line 4: The introduction of Kernel Density Estimation should include a reference.

This has been added. Please see the above comment.

page 13, line 7: The introduction of the 4-d Gaussian mixture model (GMM) should include a reference and likely Mixture and Model should begin with capital letters.

This has been added (please see the above comment), but the original capitalization is correct.

page 13, line 10: “aerosol only" Again a misuse of quotation marks. However, this time the phrase is also left undefined. If this is to be used to indicate a defined procedure/measurement it should be explicitly defined. The same is true of “aerosol + ice" and “ice only."

The quotations have been changed to italics and parenthetical descriptions have been provided (15:2-7).

page 13, line 18: Gaussian kernel support vector machine needs a reference.

This has been added. Please see the above comment.

page 13, lines 20 –: The final sentence of the paragraph, “Since a condensation ... " is long, awkwardly appended and does not follow from the previous material. Make this an independent paragraph that clearly explains how activated fractions are calculated.

This has been moved to its own paragraph and clarified. (15:16-19)

page 14, lines 4-6: “3-class supervised machine learning (bootstrap aggregated decision trees....)" At least two missing references here, 3-class supervised machine learning and bootstrap aggregated decision trees? Also missing is any explanation of why these methods are used versus what is described before and how are they different.

These two references have been added as well as “Bootstrap aggregated decision trees (Breiman, 1996) are used for the classification instead of SVM in this case, because this classification algorithm outperforms SVM in terms of classification error in the 3-class case: both algorithms are operationally interchangeable, so the better performing one with respect to classification error was chosen.” (15:21-16:2)

As a general comment to supplement the specific comments regarding the data processing section. Given this is not a computer science journal then one cannot expect that everyone

is well versed in machine learning, KDEs, GMMs, etc. As a reader it is very disturbing that much of this seems to be introduced as little more than buzz-words with acronyms. Some effort needs to be made to indicate what in practice is being done at each step. As I understand machine learning is used to predict outcomes – what exactly are the outcomes being predicted? How are the training vectors selected? Are KDEs used to make PDFs of size or of all OPC variables, it is not clear? A 2D GMM-KDE is used for visualization but a 4D used for the actual calculation? Does that mean that the 4D does a better job than we can understand by looking at the figures? Otherwise, what is the difference?

This section has been revised to help answer these questions and make the machine learning process more transparent (14:5-16:15). In brief, the machine learning is used to assign the class (aerosol, droplet, or ice) that data correspond to, and provide an uncertainty estimate of this assignment. The KDE process is what is used to select the training data via statistical sampling, which is now explained in the text. The KDE is 4D because it uses each of the four SPIN OPC variables. The 2D version is a transformed version of the 4D case, which is only presented for visualization. The 4D case is used in calculations because it does not collapse depolarization into the S1/P1 ratio, and does indeed have better performance.

Finally, should the community expect that all SPIN data should be reported in this manner (i.e. As a result of this type of analysis?)? Is this a computationally time consuming process, or is it quick and completely automated? Can (is) such an analysis be implemented in real time, such that one observes the results while running an experiment? What part of the analysis (training data, etc.) needs to be implemented separately for individual detectors on different units, or are all detectors identical to within the uncertainty?

Combining multiple referee recommendations, we have changed the manuscript to specify that the machine learning analysis is presented as a possible advanced use of SPIN, as opposed to an out-of-the box capability (13:3-5). This has been specified in the abstract as well.

page 18, line 14: What is meant by “validate"? Validate the performance as what? Perhaps this is too strong?

Combining multiple referee recommendations, the use of the word “validation” has been reconsidered throughout and replaced with less weighty vocabulary.

page 19, line 2: Here given what has been presented within the manuscript I take “homogeneous INP" to refer to homogeneous freezing of solution droplets. It is important to make the distinction between homogeneous freezing from solution and homogeneous freezing from the vapor.

This has been revised to say, “The experiments presented in this study illustrate SPIN’s measurements of freezing behavior of both the heterogeneous and homogeneous regimes and demonstrate that the SPIN chamber reproduces freezing data measured in previous studies” (22:3-5).

page 23, line 20: Kulkarni et al. ref. missing doi (doi missing from many refs)

This reference is a report without a doi and has been referenced as completely as possible based on the Copernicus guidelines (27:11). Missing doi’s for other references have been added.

page 23, line 21: Kulkarni and Kok reference seems incomplete. Is this a report? In a series? Is this document available publicly?

This is a freely available PNNL report. The location of PNNL has been added for completeness, according to the Copernicus citation specification for reports (27:11).

page 26, line 4: The Stocker IPCC reference is missing co-authors or an et. al.

The rest of the list has been added (30:1-5).

page 31, Figure 5: What are the units of the color bar? "The color scale shows the degree of flow reversal..." What is meant by "degree"? The small subplots can be better connected to the intensity plot. I assume for example in the flow reversal subplot -30 represents the warm wall and -45 the cold wall? This could be indicated in the figure and/or caption. Also, I take the color intensity to represent the maximum flow reversal velocity – is that true? Or is it a mean flow reversal velocity?

This has been clarified, combining multiple reviewer suggestions. The color scale shows the ratio of (upward) reverse flow to (downward) normal flow in the chamber, which is now specified in the caption and in the figure. The insets are now explained more thoroughly to answer these questions. "Two flow profiles are shown as insets: the coldest temperature in each corresponds to the cold wall temperature and the warmest to the warm wall temperature. Flow reversal occurs along the warm wall in one case (left, red circle) and not in the other (right, red square)" (35:9-11).

Also it must be made explicit that the Koop curve represents the onset of homogeneous freezing of liquid solutions (droplets).

This is now explicitly stated in the caption (35:8).

Given the dashed flow reversal line the phase space that can be explored without flow reversal within the chamber is quite limited. Is this an issue, does it affect the utility of this (and similar) instruments? It would be beneficial to discuss this issue in the text where the flow reversal and figure appear.

The results from the solution droplet homogenous freezing experiments show that flow reversal is adequately handled by the analytical model for CFDC chambers. This and qualitative comparison to CFD simulations is discussed in Section 5 (21:1-9).

Finally, is the plan for this figure to appear in a single column within the 2 column journal? If so the subplots and legend will become very difficult to make out. Perhaps the figure should be redrawn to optimize it for the 1 (or 2) column size?

In the past, such figures in Copernicus journals have occupied 1 column, even with 2 columns of text. All figures have been optimized for the expected size in the final paper.

page 32, Figure 6: Again think about the figure size in 1 or 2 columns. The panels will be very small, and it is difficult to read any of the numbers/axes labels and legends in the current full page format. I had to zoom into the *.pdf many times to recognize that the support vector points were plotted as open circles enclosing data points.

The fonts have all been increased, but this figure is to have its own page in the journal and be rotated to landscape view for readability.

What are the log10(size) units? I have not been able to convince myself of the meaning of these axes. In the upper 3 panels time 2 - time 1 does not look like 0 as it appears in (c). This could be due to rescaling because the intensity shading in (a)-(c) changes. I suggest that an absolute color scheme is chosen (perhaps including a wider color spectrum), such that colors in all panels represent the same probability density.

The units of intensity counts are now specified in the caption (36:3). The units of the figure are $\log_{10}$(intensity counts). Since the top three panels show probability density functions (that must be positive and integrate to 1), there is a rescaling of density that occurs during the deconvolution of c from b to preserve this integration rule. This rescaling is why the two look visually different as opposed to rescaling of the color bar, which is now consistent between panels.

page 33, Figure 7: Again think about the figure size in 1 or 2 columns.

This figure is also to have its own page in the journal and be rotated to landscape view for readability.

What does classification accuracy of 99% mean?

Cross validation of the training data lead to 1% missed classification in this case, which has been clarified (36:7).

It seems that some points would be much more certain than others. Does this mean that the total number of each is identified to a 99% certainty level? In all 3 dimensions? The 3D plot shows an area where the 3 colors (particles, ice, droplets) merge. I would assume in this area uncertainty would be high, while far from this uncertainty would be small?

The fraction of correctly classified particles using all four dimensions in the cross validation is 99%. Yes, the uncertainty is concentrated at the boundaries between classes, but the overall probability of classifying a data point correctly is 99%. Using the KDE approach to sample training data factors in the relative likelihood that a given class of particle (aerosol, droplet, or ice) will appear in an area of the parameter space: the resulting uncertainty estimate quantifies the probability that a class assignment is correct taking into account the shape of the underlying PDF of that class (14:13-18).

In the right- hand panels why does ice fraction equal water fraction at conditions sub-saturated with respect to water? Is this an area of phase space where the droplets and ice cannot be distinguished beyond the level of uncertainty? If so is this a result of small particle sizes? Is there a critical ice (droplet) size to distinguish solid from liquid?

Anywhere that the lower error bar of either the ice or droplet fractions is not above zero, this fraction is statistically indistinguishable from zero (16:11-13). This is indeed an area of phase space that of where the droplets and ice cannot be distinguished beyond the level of uncertainty, but this is considering all four parameters and not just size. Also, since the algorithm is trying to distinguish between all three output classes, this area is also where droplet and ice are more difficult to distinguish from aerosol particles as well (in part due to size). In essence, in this transition region, all three classes begin to occupy similar regions of parameter space before a clearer separation emerges. This is also expected from physical considerations, as hygroscopic growth factors are large very near water saturation.

page 34, Figure 8: These panels are probably better arranged vertically for the 2 column format.

This change has been made to Figure 8.

Is the sigmoid fit used as a scaling factor for results? Is sigmoid the appropriate fit – perhaps the efficiency is unity to a size cutoff, below which the relationship is linear? Without one or more further data points the choice of fit should be supported based on some reasoning. Was the previously presented data (Figs. 6,7) processed using these relationships? I would suggest discussing these issues in the text where Figure 8 appears.

The fits in this are not used as a scaling factor for the results, nor do they affect the ML processing, which is done on the raw intensity counts. This figure is intended to show there is near unit counting efficiency for supermicron particles, which correspond to the expected size range of droplets and ice. Particle concentrations are derived from CPC concentrations, since the OPC is less sensitive to smaller sizes. In this figure the fit is simply to guide the eye, since as the reviewer points out, the OPC has unit efficiency down to a certain size. Below this size, the efficiency drops off rapidly (faster than linearly). For this reason, a sigmoid rather than a linear fit was chosen. The text has also been modified to read "rapidly decreasing detection efficiency with decreasing size is observed for sub-micrometer particles." (17:6-7)

page 35, Figure 9: Increase the size of the legend and labels.

This has been done.

Also, again specify that the "expected homogeneous freezing" refers to aqueous phase homogeneous freezing. See my previous references to the Koop article. In such a plot it is difficult to also incorporate uncertainty, but perhaps some indication of the uncertainty range of activated fraction can be given?

The Koop et al., (2000) has been specified throughout to refer to solution droplets. Temperature, saturation, and activation uncertainties have also been specified (39:7-8).

page 36, Figure 10: See early comments with respect to formatting, font etc.

Font sizes have been increased throughout for clarity.

page 37, Figure 11: See early comments with respect to formatting, font etc.

This figure had been shrunk to fit both it and the caption on a single page. A larger size is now included (42).

I wonder if this figure could also be incorporated into the earlier discussion of flow reversal etc. Comparing with Figure 5 and given that Sice=1.3, then flow reversal should be present in (e). However, I see no evidence of this and the scale incorporates only positive velocities. The two plots are not self-consistent, please explain.

The overall qualitative behavior of the low temperature, high-RH flow reversal is consistent between the fluent simulations and the calculations from Rogers (1988). There is a smaller amount of flow reversal in the model than in the calculation, but the results from the chamber experiments at these conditions (i.e. with the solution droplet homogenous freezing experiments) suggest that the model and analytical calculation provide reasonable estimates of the actual conditions in the chamber. Also this comparison suggests that the flow reversal conditions, as represented in the analytical model, capture the general behavior even at what is predicted to be the worst flow reversal conditions. These points are discussed in Section 5 (21:1-9) and the claims are supported by similar findings in a reference suggested by another reviewer.

Let me reiterate where I see a need for major revisions.

Please refer to the specific comments above for details.

• The data processing section is difficult to understand and incomplete in its current form.

This has been clarified and referenced more thoroughly.

• Operating procedures and analysis etc. specific to SPIN should be clearly delineated from general discussions of the CFDC principle (eg., How the saturations gradient is established.). Furthermore, I encourage the authors to highlight what they consider to be the best use practices of SPIN given the current state of understanding. A simple example would be the utilization of an icing temperature of -25◦C and time of 5 s – were these observed to somehow lead to optimal experimental conditions, or simply chosen at random?

The general vs. specific principles are now more explicitly delineated. The experimental methods are now presented as a standard best practices description of SPIN operation.

• Clearly delineate what part of this manuscript represents out-of-the-box SPIN operation/ measurements versus user implemented protocols and analysis. For example, is the water reservoir cooling to 2◦C standard for SPIN?

This distinction has been made, and specific points like the one mentioned are in the standard best practices section.

• If necessary re-write the abstract to reflect the scope of the standard versus user enabled capacity of SPIN.

The abstract has been clarified to differentiate these two use cases.

One more point of importance with regard to using the homogeneous freezing line of Koop. Typically a nucleation rate must be specified when plotting this curve. For example, the value J = 5 × 108 cm−3s−1 from Hoose and Möhler (2012) is often used.

With the 200nm dry ammonium sulfate particles we used for the experiments, the growth factor after deliquescence at the freezing RH's considered is ~1.5, so we use a nucleation rate $J \approx 10^{11} cm^{-3} s^{-1}$ to compare with the Koop et al., (2000) assumption of $\omega = 1/min$. This has now been specified in the text and captions (39:4).

3. Response Interactive Comment by Anonymous Referee #3

The authors would like to thank Anonymous Referee #3 for the suggestions and insights provided in this review. After careful consideration of the information provided, we would like to offer the following responses.

What I would like to see is some preliminary results of ambient measurements (e.g. at -30◦C at a defined supersaturation). From the statement on page 19, line 17, I am assuming that it is also meant for field measurements, but from what has been shown in the manuscript, I would tend to think that the SPIN is a laboratory CFDC.

A new Figure 11 has been added to address this point and also to show SPIN field performance in the cirrus regime (19:10-18, 41:1-9).

Page 4 – line 21 : how did the authors measure the thickness of the ice layers ? Is the thickness consistent from an icing to another ? Please clarify.

This estimate was based on measuring the volume of water that was depleted from the water reservoir during the icing process. Measurements of the melt water volume were consistently lower than this type of measurement (presumably due to the depletion of the ice layer after running experiments), so this is why we choose to report our ice thickness in this manner. This is now reflected in the text (4:19-20). We also investigated (qualitatively) the ice thickness with an endoscope camera and found that the thickness in the bottom (evaporation section) of the chamber to be much thicker than the in the rest of the chamber. Since direct observation of the ice layer did not yield quantitative estimates of the ice layer thickness, and melt water estimates suggested a much thinner ice layer than that from measuring the difference in the volume of water in the reservoir after icing, we choose to give this estimate of the average thickness. A mention of this qualitative inspection has been added as well (6:20-22).

Page 9 – lines 16,17 : Can the authors show a plot of the background concentration they are considering as ideal before an INP measurement ? It would also be good to include one showing when it is not ideal to keep on sampling.

A suitable background concentration depends entirely on the INP concentration desired to be measured. In the field this may be as low as a few particles per liter, but in the laboratory in can be up to a few hundred particles cm^{-3}. This is why the periodic background checks are to ensure that frost concentrations are much smaller than the INP concentration to be measured. Examples of required laboratory (11:4-6) background levels and actual field background levels (Figure 11) have been provided.

Page 10 – line 18 : The operation time before stopping the measurements is quite broad (2-5 hours). Could the authors clarify this part ? Which specific conditions (supersaturation, temperature) lead to shorter or longer operating time.

The stopping time depends on too many factors (such as ambient RH, instrument operating temperature, operating temperature history, ramp rates, and desired INP concentration to be measured, etc.) to specify a single time. For this reason, the periodic checks are very important to ensure a usable background (10:15-18).

Figures Please increase the font of the figures (from figure 6 to 11).

This has been done throughout.