RC2
(edits are shown in red)

i) There should be a clearer distinction between the directly measured H and O isotope ratios and the deuterium excess, which is calculated from the two isotope ratios (Eq.2.) and therefore is to some extent redundant as far as analytical techniques are concerned.

Response:
It is important to include dxs in this paper, as dxs has long been an important proxy record in ice core science. However, we have made the following change:

"The dxs variable is not a direct isotopic measurement, but rather, it is calculated from the direct measurement of oxygen and hydrogen isotope ratios. The precise calculation of dxs (‰) has historically been challenging…"

ii) The last statement in 1.1 (p3, L21-24) is a result/conclusion and does not belong there.

Response:
We have removed the following sentences:
"Due to decreased density, firn ice has been difficult to measure on other CFA systems. Despite our best efforts, we recommend that low-density firn column samples be analyzed discretely. However, if the firn ice is allowed to sinter for an extended period during storage, this greatly aids in CFA measurement."

And, the following sentence has been changed:
"We then perform inter-lab and inter-system isotopic testing of firn ice and deep ice to determine how the CRDS-CFA system performs with ice of different densities."

iii) I do not understand "and normalized to Standard Light Antarctic Precipitation (SLAP)" (P4, L4). Eq. 1 is the standard definition of the isotope ratio relative to VSMOW and contains no reference to a normalization relative to SLAP.

Response:
We recognize that this is confusing. As we do normalize to VSMOW, we have removed mention of SLAP.

We have removed "and normalized to Standard Light Antarctic Precipitation (SLAP, -55.5‰ relative to VSMOW)" from the following sentence:

"The ratio of heavy to light water isotopes in a water sample is expressed in delta notation (Epstein, 1953; Mook, 2000) relative to Vienna Standard Mean Ocean Water (VSMOW, 0.00‰):"

In Section 2.4.1, we do make sure to state that:
"The protocol used for the correction of measured water isotope values is performed at the beginning and end of a melt day using laboratory isotopic standards that have been calibrated against the VSMOW2-SLAP2 scale (i.e. BSW, ASW, GSW, and PSW). By definition, SLAP2 has values for δD and $\delta^{18}O$ of -427.5 and -55.5 ‰, respectively."

The original primary standards for VSMOW and SLAP have been used up, thus we calibrate to the IAEA primary standards now in use: VSMOW2 and SLAP2. Table 2 shows the values for these primary standards, as well as the other standards in use in this study.

iv) The statement "The precise measurement of dxs (‰ has historically been challenging because IRMS methodology requires that oxygen and hydrogen water isotope ratios be analyzed on separate systems, which increases uncertainty. The CRDS-CFA technique removes multi-system uncertainty because both isotopic values are measured simultaneously on a single sample with the same system." (P4, L10-13) may be correct, but at this point it is unproven and speculative. It should either be backed up by references or phrased as a possibility that is evaluated later based on results. Furthermore, when looking at the measurement uncertainties in Table 5 it seems that the preparation creates (not surprisingly) some correlated errors for oxygen and hydrogen isotope ratio measurements (uncertainty for dxs is lower than expected from error propagation) while the errors for CRDS measurement seem to be independent (uncertainty for dxs exactly matches a Gaussian/ independent error propagation).

Response:
We have changed the wording to indicate that this is a possibility:

"The precise calculation of dxs (‰) has historically been challenging because IRMS methodology requires that oxygen and hydrogen water isotope ratios be analyzed on separate systems using separate samples, which could increase uncertainty. The CRDS-CFA technique removes the potential for multi-system uncertainty because both isotopic values are measured simultaneously on a single sample with the same system."

Furthermore, throughout the paper, we have changed mention of "system uncertainty" to "full-system precision". The full-system precision for dxs is given as 0.55 per mil.

And the caption for Table 6 now reads:
"Table 6: The isotopic measurement precision (1σ) of varying parts of the CRDS-CFA system (‰). The difference in quadrature of the full-system precision derived from GTI (System) and the precision derived from IHW (Vapor-Side) gives an estimate of the noise added on the preparation side of the system (Prep-Side). The Prep-Side values include noise added to the isotopic signal upstream of the Valco valve, while the Vapor-Side values include noise added to the isotopic signal downstream of the Valco valve."

v) The heading "Data and results" is not clear. Data would include results, but there is a separate section results and discussion. Should it be "Data processing (or evaluation) and method? There are indeed several results and to some extent a discussion of these results, which I find confusing given that there is also a results and discussion section.

Response: This is a good suggestion. We have elected to title this section "Data Processing Methodologies and Evaluation". We think that by including the word "evaluation", the reader will expect that some results related to data processing will be given in this section. We feel it is best to keep these data processing results included in this portion of the paper (Section 2), and separate the actual ice core measurement results into Section 3. We have thus retitled Section 3 from "Results and Discussion" to "Ice Core Measurement Results".

vi) P4, L21: "bolted with thermal paste" needs rephrasing

Response:

"To regulate temperature, the melt head is bolted to an aluminum heater block (51.0 mm x 51.0 mm x 45.8 mm) that is maintained at 14.6 ± 0.1 °C by internally circulating dilute propylene glycol from a 6-liter temperature-controlled bath. To increase the thermal conductivity between the melt head and heater block, a layer of polysynthetic silver thermal compound was applied to the interface prior to assembly. Traditionally, melt head temperatures have been controlled with electric resistive heaters that act on the mass of the heater block alone. Temperature control with the 6-liter circulating bath provides 19 times greater thermal mass, allowing for greater temperature stability and therefore more consistent melt rates."

vii) P4, L22-23:" This method is preferable to proportional-integral-derivative controlled resistive heaters used in most systems, which introduce small temperature fluctuations that can result in varying melt rates." is an evaluation of the method not based on presented facts and does not belong into the method section.

Response: Agreed. We have removed this line from the manuscript. Other researchers may test our method vs. other available technologies to determine what works best.

viii) P5, L4 and L8: Typical ranges of flow rate should be explicitly given, as well as the typical actual water levels in the 2mL vial. This is important to understand to which extent this reservoir may contribute to mixing occurring in the liquid phase of the CRDS-CFA system.

Response: We have modified the text as follows…

"The inner square catchment is pumped at a rate matching that of the ice stick melt. This liquid water is pushed through an 8-micrometer disposable filter backed by a 10-micrometer Peek frit 5 (IDEX A-411). With filters in place, the system is tolerant of ice with ash or dust layers.

The filtered water then enters a 2 ml glass, open-top vial where bubbles can escape (debubbler). The flow rate into the debubbler is 3.1 ml/min. An overflow tube, kept at a height of 6 mm from the bottom of the vial, regulates the height of the water. The vial has an inner diameter of 9.96 mm. A portion of the debubbled water is used as primary flow (i.e. for isotopic measurements), which is aspirated from the debubbler through a 1/16" OD 0.020" ID tubing via a selector valve (VICI Cheminert 10P-0392L). The primary flow rate is 0.1 ml/min. The primary water is then channeled through a Valco 6-port stream selection valve (Valco Instruments, Co. Inc.). High-pressure (80 psi) dry air is supplied to a downstream glass concentric nebulizer (Meinhard TL-HEN-150-A0.1), which creates a pressure difference. The nebulizer – first published by Emanuelsson et al. (2015) - converts the primary liquid water into a fine spray with a nominal droplet size of ~1.5 micrometers. The flow rate out of the nebulizer is 0.1 ml/min. Previous CFA isotopic systems have accomplished this step by pumping liquid water through a ~100 µm ID quartz capillary to a Swagelok Tee that acts as an open split (Gkinis et al. 2010; 2011). "

ix) P5,L11: The flow rate should be given. The suction effect does not depend on pressure, but on flow velocity.

See modified text above, comment viii. We have removed the following text for clarity "To induce suction," and "that pulls the primary water out of the debubbler, through the Valco selection valve, and out of the nebulizer." We also include the flow rate into the debubbler (3.13 ml/min), and for primary water out of the debubbler (0.1 ml/min), and the mixing volume of the debubbler (diameter 9.96 mm by 6 mm height).

We have also removed the following line from the manuscript:
"The auxiliary port water is 25 pulled in the same fashion as the primary flow by suction induced by the nebulizer."

x) P5,L29-30: I do not fully understand how 25 samples for 1m of ice core can give a 2.5 cm resolution. Is the resolution given calculated from mixing effects?

Response: This is fixed.

"It is programmed to partially fill approximately twenty-five 2 ml glass vials for every meter of ice melted, approximating a discrete ice sampling resolution of about 4.0 cm."

xi) The last paragraph in 2.2 is a discussion/conclusion and does not belong into the method description.

Response:
We have removed this paragraph:
"It should be noted that transforming the liquid water stream to vapor can be accomplished in different ways. Previous CFA isotopic systems have accomplished this by pumping liquid water through a ~100 μm ID quartz capillary to a Swagelok Tee that acts as an open split (Gkinis et al. 2010; 2011). This method has the advantage of simplicity, but relies heavily on precise placement of the capillary within the tee fitting. Alternatively, using a nebulizer as per Emanuelsson et al. (2015) provides stable water vapor concentrations within the CRDS vapor cavity. There are also fewer issues with system clogging. When clogs did occur in the nebulizer, they were easily identified and cleaned by back flushing the nebulizer using a vacuum."

However, to remain consistent with prior editorial comments, we have moved the citations for Emanuelsson et al. and Gkinis et al. to a preceding paragraph. "The nebulizer – first published by Emanuelsson et al. (2015) - converts the primary liquid water into a fine spray with a nominal droplet size of ~1.5 micrometers. The flow rate out of the nebulizer is 0.1 ml/min. Previous CFA isotopic systems have accomplished this step by pumping liquid water through a ~100 μm ID quartz capillary to a Swagelok Tee that acts as an open split (Gkinis et al. 2010; 2011)."

xii) P6, L16-17 : unclear meaning.

Response: We have removed this line from the manuscript for clarity.

"We found that a temperature of -30 °C produced the least amount of fractionation, but that fractionation during the freezing process prevented the mock ice from being used as strict isotopic standards. "

xiii) P7, L19: "can" or "was" or "routinely is" verified. What is (was) actually done?

Response:
We have made the following change:

"We tested and verified this depth assignment scheme using discrete samples of the same ice sections analyzed on the CRDS-CFA system."

xiv) P7, L20-30: This detailed explanation of the amount of data connected to computing resources is unnecessary. The authors simply should state the type of computer used and the time needed for processing. Any qualitative explanation that a more powerful computer will speed up processing is trivial. Statements about time savings and resolution of different methods should be moved to the end of the discussion into a sub chapter presenting a systematic comparison of different methods.

Response:
We have made a number of changes to address this comment.

We changed section titles or added titles, as follows:
Section 2.4 changed to "Post-Measurement Data Processing"
Section 2.4.1 added "Isotope Calibration"
Section 2.5 changed to Section 2.4.2 "System Performance"
Section 2.7 changed to Section 2.4.3 "Mixing Length Calculations"
Section 2.6 changed to Section 2.4.4 "Isotopic Step-Change Correction"
Section 2.5 added "Method Comparison"

pg. 6, lines 28-30 are changed to: "All data are recorded at the same data frequency as the isotopic data generated by the Picarro L2130-*i* (in this case, at approximately ~1.18 Hz intervals (0.85 second intervals. The data are exported as a raw text file generated by the Picarro and GUI."

We deleted the following sentence:
"However, to process the CRDS-CFA data requires multiple computer processors, and for the largest tasks, super-computing can reduce processing times from hours or days to minutes."

We moved the following to the new Section "2.5 Method Comparison"
"The Picarro L2130-*i* generates and records isotopic data at approximately ~0.85 second intervals. For a melt rate of 2.5 cm/min, this translates into sub-millimeter isotopic resolution, and roughly 2,800 data points per meter. At 12 meters of ice core per day, and recording 36 total quality control parameters, each melt day generates over 1.2 million data points. For a deep ice core of ~3,000 meters, this equates to ~300 million data points recovered in the course of 250 melt days. In comparison, traditional IRMS produces about ~60 ice core data points per day, equating to the analysis of ~2.0 meters of ice per day at 3.0 cm resolution. For a deep ice core of ~3,000 meters, this would generate a total of 99,000 data points for a single parameter and require about 1,500 analysis days to complete. The CRDS-CFA system presented here provides a ~6 fold time-savings and an order of magnitude improvement in data density compared to traditional IRMS methods."

xv) P8, L15-20: The definition of accuracy seems to differ from the typical definition of analytical accuracy. Analytical accuracy includes all sources of error and thus cannot be lower than the precision of a measurement. What is presented as accuracy seems to be the bias of the measurement. Bias can be corrected, but the overall accuracy needs to consider (in addition to all other sources of uncertainty) the uncertainty of the correction.

We now state "precision" and "bias" below.

"We also use the laboratory isotopic standards to define precision and bias of the CRDS-CFA system, occurring downstream of the Valco stream selection valve. Except for a small section of tubing between the Valco valve and the nebulizer, these values account for vapor-phase noise added to the isotopic signal. We use the IAEA/WMO definitions of precision and bias. Precision is defined as the degree of internal agreement among independent measurements made under specific conditions, while bias is defined as the difference between the expectation of the test results and an accepted reference value. The vapor-phase precision is determined by taking the average standard deviation of each of the last 5 minutes of each 20-minute laboratory standard run, while the vapor-phase bias is defined as the difference of the measured GSW value from its known value. These values can then be compared over the long-term to monitor the performance of the CRDS-CFA system. The average vapor-phase precision (1σ) and average vapor-phase bias for δD was 0.49 and 0.39 ‰, respectively. For δ18O, these values were 0.05 and 0.01 ‰, respectively. In Section 3, we define a precision value that takes into account all components of the CRDS-CFA system."

xvi) P8,L20-21: This statement is out of place and somewhat speculative since it does not present any

quantitative comparison of precision and accuracy. If correctly determined accuracy and precision of different methods can be compared directly.

We have decided to remove the following text:

"We note that there is no universally accepted method for defining internal precision and accuracy in water isotope CFA systems. For example, Gkinis (2010; 2011) used only 3 laboratory isotopic standards, whereas we use four. This should be taken into account when comparing internal precision and accuracy across different CFA papers."

See more discussion of precision below…

xvii) P9, L5-6: Table 3 and Figure 3 are redundant. Furthermore, they are results and should be presented and discussed in Chapter 3.

Although Figure 3 and Table 3 show results, we think they should be kept in Section 2. As stated in comment (v), we renamed Section 2 to "Data Processing Methodologies and Evaluation". We think that by including the word "evaluation", the reader will expect that some results related to data processing will be given in this section. We feel it is best to keep these data processing results included in this portion of the paper (Section 2), and separate the actual ice core measurement results into Section 3. We have thus retitled Section 3 from "Results and Discussion" to "Ice Core Measurement Results".

And, although Figure 3 and Table 3 are based on the same data, we felt it necessary to give exact values in table form for the data shown in graphical form. Just showing the graph makes visually determining Alan Deviation values difficult, especially with the log scale on the y-axis. However, just giving a few values in the Table makes it hard to see the form of the Alan Deviation graph, which is important as the trough indicates the optimal integration time – some researchers may find the Table helpful to compare a set of Alan Deviations from other CFA instruments to our instrument.

xviii) P9,L18: At this point there is no explanation how the vapor phase mixing is determined. Furthermore, this is a result.

Response: We have removed the following sentences from the manuscript.
"In laboratory tests, we note that removal of the 1st Picarro instrument filter can reduce mixing in the vapor phase by 40%. However, the removal of the filter also risks damaging the laser cavity, and we chose to leave the filter in place."

The fact that the Picarro instrument filter increases the mixing effect in still suggested as a possible mixing effect in the manuscript: "Possible contributors to the mixing effect include: … vapor interactions with two Picarro instrument filters (Mykrolis Wafergaurd) prior to entering the laser cavity." Since we don't extensively treat this topic, we feel it is best to remove the results shown in red above.

xix) P9, L20-P10, L6: The description of the mixing correction is not clear. It seems to be based on the deviation from the expected value. This works for a sample with known isotope ratios, but not for ice cores with a priori unknown isotope ratios (and it would not be necessary if the correct isotope ratios were known). The correct procedure would be to solve Eq. 5 for ct (as function of t) for tests with known "correct" isotope ratios and then use ct as function of t as described in Eq. 5. Is this what has been done? Clarification is needed. Furthermore, corrections have uncertainties which will impact the accuracy of the result. There is no information on the uncertainty of this correction, except that it "can only be applied at ct>0.65", which implies a (not explained) threshold for error

introduced by the correction. There is no information about magnitude or uncertainty of ct, let alone an evaluation how this impacts the final result.

Response:
In answering the first question, yes, we have performed tests with isotope ratios using mock ice with quantifiable isotope values. We then solve for Eq. 5 for c_n (previously c_t), and apply that same correction to the isotope data with the a priori unknown isotope ratios.

To make this clearer, we have modified the first paragraph of Section 2.4.4:
"During a typical analysis day, every 4th ice core stick is separated by isotopically distinct push ice. Ideally, the isotopic shift between push ice and the beginning of an ice core stick at the melt head would register instantaneously on the CRDS instrument. However, a substantial mixing effect is evident in the first ~5-7 cm of an ice core sample. This effect can be partially corrected for by separately determining mixing coefficients (c_n), quantified using an instantaneous shift in mock ice isotopic values at the melt head. The isotopic values of the mock ice sticks are approximately -240 and -120 ‰. As the water for the mock ice sticks moves through the CRDS-CFA system, the water will mix together, resulting in a smoothed transition over time (rather than an instantaneous transition) as measured by the CRDS instrument. …"

And:
"An example of an isotopic step-change correction, applied to an isotopic shift between push ice and the beginning of an ice core stick, is shown in Figure 6. For c_n ≥ 0.65, the correction saves an additional ~2.25-3.15 cm of data that would otherwise be discarded. For a typical deep ice core of 3000 m, there are 750 step-changes (every 4th ice core stick). This correction saves ~16.9 – 23.7 m of isotopic data. In the data output file we produce, the depths of isotopic step-change corrections are flagged so researchers can decide whether to include some or all of this data in their calculations. In Table 5, we provide standard error estimates for isotopic step-change corrections of mock ice sticks for different bins of data at $c_n \geq 0.5$.

We have also included a new Table 5:
"Table 5: Standard deviation of the mean for 50 determinations of isotopic step-change corrections using mock ice (δD step-change values of approximately -240 to -120 ‰). All values are given in per mil (‰). Each data bin includes 20 consecutive data points (spanning ~17 seconds of analysis time), increasing from c_n = 0.5. The data in each bin are averaged for each determination. Corrected values of dxs are calculated from the corrected δD and δ^{18}O values. From c_n = 0.5, about 100 seconds are needed to achieve greater than 98% of the expected mock ice value of -120 ‰. For comparison, the calculation of the standard deviation of the mean for 50 determinations of 17 second averages of isotopically homogenous water yields values for δD, δ^{18}O, and dxs of 0.01, 0.01, and 0.05 ‰, respectively. The isotopically homogenous water was allowed to run through the CRDS-CFA system for 1 hour before the 17-second averages were determined."

Bin (sec)	**δD**	**δ^{18}O**	dxs
1 (0-17)	1.06	0.06	1.16
2 (17-34)	0.54	0.06	0.46
3 (34-51)	0.39	0.05	0.19
4 (51-68)	0.31	0.04	0.1
5 (68-85)	0.26	0.04	0.09
6 (85-102)	0.23	0.04	0.08

xx) Subchapter 2.6 in general: This subchapter also contains a mixture of method description and results. Here a clear description of the procedure should be given, results should be presented and discussed in Chapter 3.

Note: Section 2.6 changed to Section 2.4.4 “Isotopic Step-Change Correction”.

As stated before, we have elected to have two sections that contain results. We renamed Section 2 to “Data Processing Methodologies and Evaluation”. We think that by including the word “evaluation”, the reader will expect that some results related to data processing will be given in this section. We feel it is best to keep these data processing results included in this portion of the paper (Section 2), and separate the actual ice core measurement results into Section 3. We have thus retitled Section 3 from “Results and Discussion” to “Ice Core Measurement Results”.

xxi) P9,L29: “at time t-1” needs rephrasing. Here t is the number of a given time interval, not really a time (an index needs to be a number). Probably better to use n

Response:
We have made the following changes:
“

$$\delta_{mc} = \frac{\delta_m - \delta_{mp}(1 - c_n)}{c_n} \quad (5)$$

where δ_m is the measured isotopic value at index n, δ_{mp} is the previously measured isotopic value at $n - 1$, c_n is the memory coefficient at n, and δ_{mc} is the corrected isotopic value at n.”

Subsequent mentions of c_t are changed to c_n as well.

xxii) P10, L20 “We modified”
Fixed.

xxiii) P10, last line: “to maximize the fit” is not clear.

Response:
We have changed “where μ_1, μ_2, σ_1, and σ_2 are predetermined to maximize the fit.” to “where μ_1, μ_2, σ_1, and σ_2 are determined by least squares optimization.”

xxiv) Chapter 2.7 in general: Again there is a mixture of method description and results, especially P11, L16 to P12 L6 contain mainly results and discussion/conclusions.

Note that Section 2.7 changed to Section 2.4.3 “Mixing Length Calculations”.

As stated before, we have elected to have two sections that contain results. We renamed Section 2 to “Data Processing Methodologies and Evaluation”. We think that by including the word “evaluation”, the reader will expect that some results related to data processing will be given in this section. We feel it is best to keep these data processing results included in this portion of the paper (Section 2), and separate the actual ice core measurement results into Section 3. We have thus retitled Section 3 from “Results and Discussion” to “Ice Core Measurement Results”.

xxv) Chapter 2.6 and 2.7: The isotope step correction and mixing length are tightly connected, here it seems that they are separate issues. However, the necessity for a correction is the consequence of mixing. The order in which the two subchapters are presented should be reversed and the correction be identified as procedure to correct for the impact of mixing.

Response:
We have reversed these sections.
Section 2.7 changed to Section 2.4.3 “Mixing Length Calculations”
Section 2.6 changed to Section 2.4.4 “Isotopic Step-Change Correction”

The beginning of section 2.4.3 has changed, and now reads:
Isotopic mixing effects occur in the CRDS-CFA system. Possible contributors to the mixing effect include: liquid mixing in tubing and the debubbler, liquid drag on tubing walls, vapor mixing downstream of the nebulizer, vapor interactions with two Picarro instrument filters (Mykrolis Wafergaurd) prior to entering the laser cavity, adsorption of water molecules onto the laser cavity walls, and diffusional effects (probably negligible) that can occur at any point in the CRDS-CFA system. To characterize the mixing, a transfer function and impulse response function of the system can be defined using mock ice or laboratory water standards.

And the beginning of Section 2.4.4 now reads:
During a typical analysis day, every 4th ice core stick is separated by isotopically distinct push ice. Ideally, the isotopic shift between push ice and the beginning of an ice core stick at the melt head would register instantaneously on the CRDS instrument. However, a substantial mixing effect is evident in the first ~5-7 cm of an ice core sample. This effect can be partially corrected for by separately determining mixing coefficients (c_n), quantified using an instantaneous shift in mock ice isotopic values at the melt head. The isotopic values of the mock ice sticks are approximately -240 and -120 ‰. As the water for the mock ice sticks moves through the CRDS-CFA system, the water will mix together, resulting in a smoothed transition over time (rather than an instantaneous transition) as measured by the CRDS instrument.

xxvi) Chapter 3 contains not only results and discussion; it includes method description such as the origin and handling of ice cores analyzed as well as description of tests. The chapter presents three evaluations by comparing results for “real samples” using different methods (including the method presented here). This is excellent work. Unfortunately the value of this is difficult to see and understand for the reader due to the “parallel presentation and discussion”. I find it very difficult to compare performance data (e.g. precision and accuracy, mixing length etc.) for different types of samples (artificial mixtures, water, ice cores and firn etc). To do this the reader has to go back and forth not only between the three subchapters of Chapter 3, but also between chapter 2 and 3. The authors should present a clear comparison between performance parameters under different conditions, for example as Table (or bar graph). This would provide the reader with the possibility to understand differences in performance between artificial samples, water, different ice or firn cores. This would also provide a solid basis for a straightforward comparison of the newly developed method with existing methods, which should be combined in one subchapter.

Response:
As stated before, we have elected to have two sections that contain results. We renamed Section 2 to “Data Processing Methodologies and Evaluation”. We think that by including the word “evaluation”, the reader will expect that some results related to data processing will be given in this section. We feel it is best to keep these data processing results included in this portion of the paper (Section 2), and separate the actual ice core measurement results into Section 3. We have thus retitled Section 3 from “Results and Discussion” to “Ice Core Measurement Results”.

In Section 2.4.3, we present mixing length values that are valid for measurements of solid ice from any location (Greenland ice, Antarctic ice, etc.), assuming a melt rate on the CRDS-CFA system of 2.5 cm/min, and a total mixing volume that is similar to our system. This mixing length value is referred to as $\sigma_{mock\ ice}$. We have now made note that the $\sigma_{mock\ ice}$ values are not valid for firn ice, as the density of ice in the firn is considerably lower than solid ice. It is also important to note that a direct comparison of $\sigma_{mock\ ice}$ to other studies is not possible, as other studies utilize different flow rates and have different total mixing volumes. See changes below:

"The $\sigma_{mock\ ice}$ values are not valid for firn ice, which has much lower densities than deeper ice due to the bubbles and vapor pathways that are present in the firn. The remainder of this paper, except for Section 3.3, relates to tests of solid ice beneath the firn column. To illustrate the importance of $\sigma_{mock\ ice}$, consider a series of solid ice core samples that have an average layer thickness of ~23.0 cm/year (a typical value at ~200 m depth in the WAIS Divide ice core). A $\sigma_{mock\ ice}$ value of ~1 cm would have almost no affect on the attenuation of the natural signal. However, if the average layer thickness was ~2.0 cm/year (a typical value for very low accumulation sites in East Antarctica), then $\sigma_{mock\ ice}$ = ~1 cm would have a larger attenuation affect on the isotopic signal. For low accumulation sites, it may be necessary to adjust the melt rate or decrease the total mixing volume of a CRDS-CFA system when measuring ice cores, as this could decrease the $\sigma_{mock\ ice}$ values. Figure 5 shows a comparison of isotopic signal corrections using $\sigma_{mock\ ice}$ for the CRDS-CFA system vs. a typical diffusion length occurring naturally from diffusional processes in the Antarctic ice sheet at the location of the WAIS Divide ice core. In our data output file, values of $\sigma_{mock\ ice}$ and melt rate are included for all depths to allow for CRDS-CFA mixing length corrections (i.e. deconvolution corrections). It is at the discretion of the output file user whether to make these corrections. "

We also note that we have not yet devised a way to test $\sigma_{mock\ ice}$ for low-density firn- ice. Thus, we cannot define a mixing length reliably for lower density ice. Instead, we test the ability to measure firn ice on CRDS-CFA by comparing measurements to discrete samples. Ultimately, we recommended that researchers use discrete sampling methods for firn ice, as the uncertainty of CRDS-CFA firn measurements is too high due to the lower densities.

In Section 3.3, we changed:
"Until this point, our analyses have focused on solid ice; we now shift focus to firn ice, which involves additional complexities. Due to the increased porosity and lower density of firn ice, liquid water at the melt head can wick upwards by capillary action through the pore spaces, artificially smoothing the isotopic signal and decreasing signal amplitude. This can, for example, alter frequency analysis calculations or cause the misinterpretation of the size of summer-winter signals. The melt rate is also difficult to control, as density variations over a few centimeters can cause the melt rate to increase or decrease unexpectedly. Furthermore, we have not devised a reliable way to test $\sigma_{mock\ ice}$ for firn, as the interconnected vapor pathways and trapped air bubbles are difficult to replicate in the lab. For this reason, we can only test the reliability of CRDS-CFA firn measurements by comparing the CFA measurements to discrete samples over the same intervals of ice. To test the repeatability of CRDS-CFA firn measurements, two identical 1-meter sections from 4, 9, and 29 meters depth in a South Pole firn core are analyzed under the same conditions of measurement (hereafter referred to as South Pole Firn; SPF). The reproducibility of these CRDS-CFA firn measurements is then compared with a third identical 1-meter SPF section discretely sampled at 1 cm increments using a Picarro L2130-*i* (Figure 10)."

Finally, the reviewer's comment regarding "compar[ing] performance data (e.g. precision and accuracy, mixing length etc.) for different types of samples (artificial mixtures, water, ice cores and firn etc)" deserves some clarification. We essentially tested two types of ice: 1) Solid ice beneath the firn column, and 2) Firn ice. Most of the paper focuses on the solid ice. Section 3.3 treats firn ice. In our initial tests of firn ice, we determined that CRDS-CFA data quality was secondary to discrete sampling methods, therefore, no further tests were done for firn ice. The rest of the paper relates to solid ice, and all performance data up until Section 3.3 is in relation to this.

To make this clear, we include the following changes:
In the beginning of Section 3.1, we write:
"To test CRDS-CFA precision for the entire system, nine sticks of solid ice were cut from a single meter of Greenland test ice (GTI) at about 400 meters depth. The ice sticks were obtained from the National Science Foundation National Ice Core Laboratory (Lakewood, Colorado) and were a by-product of the initial tests of the U.S. Deep Ice Sheet Coring Drill performed in Greenland in 2006 (Johnson et al., 2007)."

In the beginning of Section 3.3, we write:

"Until this point, our analyses have focused on solid ice; we now shift focus to firn ice, which involves additional complexities."

In the conclusion, we write:
"We have tested two types of ice cores on the CRDS-CFA system: firn ice and solid ice beneath the bubble close-off depth in the firn. Using South Pole firn samples, we find that the CRDS-CFA measurements are not of high-enough quality for most scientific analyses, due mainly to difficulties in controlling the melt rate with low-density firn ice. We suggest that researchers instead utilize high-resolution discrete samples for firn column measurements. For solid ice, we used identical Greenland ice core sticks to quantify the precision of the CRDS-CFA system. Additionally, we measured West Antarctic ice core sticks to test the reproducibility of two sets of measurements: 1) Traditional magnetic sector isotope ratio mass spectrometry (IRMS) vs. CRDS-CFA system measurements, and 2) Inter-lab comparisons of CFA and discrete measurements on the same type of CRDS instrument. The entirety of these solid ice tests validates the measurement of water isotopes on the CRDS-CFA system presented in this paper.

…

One exception to these solid ice results is found in the deuterium excess (dxs) measurements. At the highest frequencies (multi-year to decadal), we found discrepancies in comparisons of both inter-lab CRDS measurements of dxs and in CRDS and IRMS measurements of dxs. Additional studies at high resolution and between different labs are necessary to better understand these high-frequency discrepancies in dxs. Contrary to this, however, are the dxs results measured solely on the CRDS-CFA system presented in this paper. We found that dxs can be replicated at high frequencies, demonstrated by tests of identical Greenland ice core sticks. Furthermore, despite slightly different impulse response functions for δD and δ^{18}O, the dxs measured from replicate ice sticks on the CRDS-CFA system is correlated ($R^2 = 0.88$) with discrete CRDS dxs values measured on the same replicate ice. This indicates a negligible artifact in dxs calculations arising from the differing impulse response functions, and shows that CRDS-CFA system measurements of dxs can be used reliably as a climate proxy.

The overall results of this study for CRDS-CFA relative to IRMS show an improvement in δD precision (0.55 vs. 1.0 ‰, respectively), similar precision for δ^{18}O (0.09 vs. 0.10 ‰, respectively), and a substantial improvement in the precision of dxs (0.55 vs. 1.3 ‰, respectively). These CRDS-CFA precision results are also in line with prior ice core CFA studies. Gkinis et al. (2011) cites combined measurement uncertainty values of 0.2, 0.06, and 0.5 ‰ for δD, δ^{18}O, and dxs, respectively. Maselli et al. (2012) and Emanuelsson et al. (2015) report similar results. However, unlike these prior studies, the precision values cited in this paper were determined using replicate ice core sticks. We suggest that future CFA studies utilize this approach, as it is most representative of the added noise introduced by CRDS-CFA systems into water isotope measurements. In terms of CRDS-CFA system mixing lengths, these values depend on both flow rate and total mixing volume in the system. Thus, mixing lengths (sometimes referred to as diffusion lengths) published in other CRDS-CFA studies should be carefully reviewed in comparison to this study. Finally, the stated precision values in this paper result in high signal-to-noise ratios for most ice core measurements. For example, at ~215 meters depth in the WAIS Divide ice core the average annual signal (i.e. the difference in maximum summer and minimum winter values) for δD, δ^{18}O, and dxs is about 20.0, 2.5, and 4.0 ‰, respectively. This corresponds to a CRDS-CFA signal-to-noise ratio for δD, δ^{18}O, and dxs of about 36, 28, and 7, respectively. In most instances, researchers look for a signal-to-noise ratio of at least 5 for paleoclimate data. Researchers should consider this hypothetical threshold when interpreting future ice core data measured on the CRDS-CFA system presented in this study, or on other systems.

xxvii) P12, L9-10: I find the use of the terms "external precision" or repeatability as specific goals of this test somewhat confusing. The authors should stick to clearly defined criteria such as accuracy, reproducibility, bias, mixing length etc. If the performance criteria are different for different types of samples, then this is something that needs discussion and (ideally) explanation.

Response:

Throughout the paper, we have replaced “system uncertainty” and “external precision” with “full-system precision”. This is done to prevent the reader from assuming that we have fully quantified all known uncertainties related to the isotopic measurements presented in the paper, and to remove the confusing word “external”. We have removed mention of repeatability. Bias (instead of accuracy) is now used in Section 2.4.2. And, in any case where we define a precision, we fully explain how that precision value was determined. We mention 3 types of precision measurements in the paper: vapor-phase precision downstream of the Valvo valve, a similar precision measurement using isotopically homogenous waters, and a full-system precision determined from replicate ice sticks.

Furthermore, to improve clarity of the stated precision values in the paper, we add the following:

In the first paragraph of Section 2.4.2:
“We also use the laboratory isotopic standards to define precision and bias of the CRDS-CFA system, occurring downstream of the Valco stream selection valve. Except for a small section of tubing between the Valco valve and the nebulizer, these values account for vapor-phase noise added to the isotopic signal. We use the following definitions of precision and bias. Precision is defined as the degree of internal agreement among independent measurements made under specific conditions, while bias is defined as the difference between the expectation of the test results and an accepted reference value. The vapor-phase precision is determined by taking the average standard deviation of each of the last 5 minutes of each 20-minute laboratory standard run, while the vapor-phase bias is defined as the difference of the measured GSW value from its known value. These values can then be compared over the long-term to monitor the performance of the CRDS-CFA system. The average vapor-phase precision (1σ) and average vapor-phase bias for δD was 0.49 and 0.39 ‰, respectively. For δ18O, these values were 0.05 and 0.01 ‰, respectively. In Section 3, we define a precision value that takes into account all components of the CRDS-CFA system.”

In the second paragraph of Section 3.1:
“The isotopic data for the seven 1-meter long GTI sticks analyzed on the CRDS-CFA system was measured at sub-millimeter resolution. These data were then averaged to 1 cm successive values (GTI-1cm) for each stick, resulting in seven values at each 1 cm increment, from which a standard deviation was determined. This resulted in a total of 100 standard deviation values, and the mean of these values is used to estimate the full-system precision (Figure 7). We find full-system precision values (1σ) for δD, δ^{18}O, and dxs of 0.55, 0.09, and 0.55 ‰, respectively. For comparison, traditional IRMS measurements using water equilibration methods for δD, δ^{18}O, and dxs are commonly reported with a precision of about 1.0 0.1, and 1.3 ‰, respectively.”

In the third paragraph of Section 3.1:
“The full-system precision can be subdivided into two parts: noise added to the isotopic signal on the preparation side of the system (i.e. noise occurring prior to the Valco stream selection valve, including depth registration) and noise added to the isotopic signal on the vapor side of the system (i.e. noise occurring downstream of the Valco stream selection valve, including the nebulizer and CRDS laser cavity). We can isolate the noise added on the vapor side of the system by analyzing a continuous stream of isotopically homogenous water (IHW) inputted directly into the nebulizer from the Valco valve and subsequently analyzed on the CRDS instrument. We take averages of the IHW over consecutive 24-second intervals (supplement), which correspond to the time needed to melt 1 cm of ice at a melt rate of 2.5 cm/min. The standard deviation of 100 consecutive 24-second averages of the IHW (the equivalent of a meter long section of ice) yields an IHW precision value, which is an estimate of the noise added to the isotopic signal on the vapor side of the system (similar to that determined in Section 2.4.2). The difference in quadrature of this IHW precision value and the full-system precision value (determined from GTI) gives an estimate of the noise added to the isotopic signal on the preparation side of the system upstream of the Valvo valve (Table 6). We find that the added preparation side noise is significantly larger than the noise added on the vapor side of the system. This highlights the importance of limiting mixing in the system and achieving high-quality depth registration measurements.”

The full-system precision can be subdivided into two parts: noise added to the isotopic signal on the preparation side of the system (i.e. noise occurring prior to the Valco stream selection valve, including depth registration) and noise added to the isotopic signal on the vapor side of the system (i.e. noise occurring downstream of the Valco stream selection valve, including the nebulizer and CRDS laser cavity). We can isolate the noise added on the vapor side of the system by analyzing a continuous stream of isotopically homogenous water (IHW) inputted directly into the nebulizer from the Valco valve and subsequently analyzed on the CRDS instrument. We take averages of the IHW over consecutive 24-second intervals (supplement), which correspond to the time needed to melt 1 cm of ice at a melt rate of 2.5 cm/min. The standard deviation of 100 consecutive 24-second averages of the IHW (the equivalent of a meter long section of ice) yields an IHW precision value, which is an estimate of the noise added to the isotopic signal on the vapor side of the system (similar to that determined in Section 2.4.2). The difference in quadrature of this IHW precision value and the full-system precision value (determined from GTI) gives an estimate of the noise added to the isotopic signal on the preparation side of the system upstream of the Valco valve (Table 6). We find that the added preparation side noise is significantly larger than the noise added on the vapor side of the system. This highlights the importance of limiting mixing in the system and achieving high-quality depth registration measurements.

In the last paragraph of Section 3.1:
"As a final analysis using GTI, the reproducibility of GTI-1cm CRDS-CFA data and 1 cm discrete CRDS data are compared. We define reproducibility as the closeness of the agreement between results of measurements of the same measure and carried out under changed conditions of measurement. In this case, the same water samples are analyzed using a Picarro L2130-*i,* but the exact conditions of measurement are different, in that some samples are analyzed on the CFA-CRDS system, while others are analyzed by discrete sample injection into a CRDS instrument. A scatter plot and linear regression of the resulting data for the two measurement types gives a slope, y-intercept, and R^2 value (Figure 8). The coefficient of determination values (R^2) for δD and $\delta^{18}O$ are 0.99 and 1.00, respectively, while dxs remains inherently more difficult to measure with an R^2 value of 0.88. Because this test shows the comparison of seven GTI sticks analyzed on CRDS-CFA compared with a single ice stick measured discretely on CRDS, the R^2 values represents the reproducibility of the average of seven ice sticks vs. a single ice stick carried out under changed conditions of measurement. We explore the reproducibility of sets of single ice sticks for IRMS and CRDS in the following section."

Additional edits are made throughout the paper, especially related to changing "system uncertainty" and "external precision" to "full-system precision".

For example:
"Table 6: The isotopic measurement precision (1σ) of varying parts of the CRDS-CFA system (‰). The difference in quadrature of the full-system precision derived from GTI (System) and the precision derived from IHW (Vapor-Side) gives an estimate of the noise added on the preparation side of the system (Prep-Side). The Prep-Side values include noise added to the isotopic signal upstream of the Valco valve, while the Vapor-Side values include noise added to the isotopic signal downstream of the Valco valve."

"Table 2: Water isotope standards in units of per mil (‰). The laboratory standards (BSW, ASW, GSW, and PSW) were calibrated to primary standards (VSMOW2, SLAP2, and GISP) in March 2010. The primary standards are reported with errors given by the IAEA. The laboratory standards are reported as a precision value given by the average standard deviation (1σ) of multiple isotopic determinations across multiple analysis platforms (IRMS and CRDS). In parentheses, the combined uncertainty of the laboratory standard and each primary standard are given, added in quadrature."

xxviii) The conclusion contains a significant amount of detailed repetitions from chapters 2 and 3 as well as performance comparison with other methods. This results in a rather long (especially for a measurement technic paper) conclusion. This can be avoided by a better structure (see above) which will also give the reader a much better chance to form an own opinion.

Response: We have reduced the length of the conclusion, and made substantial changes.

“We have presented a high-resolution continuous flow analysis (CFA) system based on cavity ring down spectroscopy technology (CRDS) that is specifically designed for water isotope analysis of ice cores. The CFA system converts ~1 m ice sticks into a continuous liquid water stream, which is then vaporized and analyzed on a CRDS instrument. The full system builds from previous water isotope CFA studies (Gkinis et al., 2010; 2011; Maselli et al., 2012; Emanuelsson et al., 2015) and includes novel improvements to the ice delivery mechanism and the melt head.

We have tested two types of ice cores on the CRDS-CFA system: firn ice and solid ice beneath the bubble close-off depth in the firn. Using South Pole firn samples, we find that the CRDS-CFA measurements are not of high-enough quality for most scientific analyses, due mainly to difficulties in controlling the melt rate with low-density firn ice. We suggest that researchers instead utilize high-resolution discrete samples for firn column measurements. For solid ice, we used identical Greenland ice core sticks to quantify the precision of the CRDS-CFA system. Additionally, we measured West Antarctic ice core sticks to test the reproducibility of two sets of measurements: 1) Traditional magnetic sector isotope ratio mass spectrometry (IRMS) vs. CRDS-CFA system measurements, and 2) Inter-lab comparisons of CFA and discrete measurements on the same type of CRDS instrument. The entirety of these solid ice tests validates the measurement of water isotopes on the CRDS-CFA system presented in this paper.

One exception to these solid ice results is found in the deuterium excess (dxs) measurements. At the highest frequencies (multi-year to decadal), we found discrepancies in comparisons of both inter-lab CRDS measurements of dxs and in CRDS and IRMS measurements of dxs. Additional studies at high resolution and between different labs are necessary to better understand these high-frequency discrepancies in dxs. Contrary to this, however, are the dxs results measured solely on the CRDS-CFA system presented in this paper. We found that dxs can be replicated at high frequencies, demonstrated by tests of identical Greenland ice core sticks. Furthermore, despite slightly different impulse response functions for δD and δ^{18}O, the dxs measured from replicate ice sticks on the CRDS-CFA system is correlated ($R^2 = 0.88$) with discrete CRDS dxs values measured on the same replicate ice. This indicates a negligible artifact in dxs calculations arising from the differing impulse response functions, and shows that CRDS-CFA system measurements of dxs can be used reliably as a climate proxy.

The overall results of this study for CRDS-CFA relative to IRMS show an improvement in δD precision (0.55 vs. 1.0 ‰, respectively), similar precision for δ^{18}O (0.09 vs. 0.10 ‰, respectively), and a substantial improvement in the precision of dxs (0.55 vs. 1.3 ‰, respectively). These CRDS-CFA precision results are also in line with prior ice core CFA studies. Gkinis et al. (2011) cites combined measurement uncertainty values of 0.2, 0.06, and 0.5 ‰ for δD, δ^{18}O, and dxs, respectively. Maselli et al. (2012) and Emanuelsson et al. (2015) report similar results. However, unlike these prior studies, the precision values cited in this paper were determined using replicate ice core sticks. We suggest that future CFA studies utilize this approach, as it is most representative of the added noise introduced by CRDS-CFA systems into water isotope measurements. In terms of CRDS-CFA system mixing lengths, these values depend on both flow rate and total mixing volume in the system. Thus, mixing lengths (sometimes referred to as diffusion lengths) published in other CRDS-CFA studies should be carefully reviewed in comparison to this study. Finally, the stated precision values in this paper result in high signal-to-noise ratios for most ice core measurements. For example, at ~215 meters depth in the WAIS Divide ice core the average annual signal (i.e. the difference in maximum summer and minimum winter values) for δD, δ^{18}O, and dxs is about 20.0, 2.5, and 4.0 ‰, respectively. This corresponds to a CRDS-CFA signal-to-noise ratio for δD, δ^{18}O, and dxs of about 36, 28, and 7, respectively. In most instances, researchers look for a signal-to-noise ratio of at least 5 for paleoclimate data. Researchers should consider this hypothetical threshold when interpreting future ice core data measured on the CRDS-CFA system presented in this study, or on other systems. “

Furthermore, we removed the last sentence in Section 3.3, and placed similar text in the conclusion above: “In any case, we find that our CRDS-CFA measurements of firn samples are not of high-enough quality to scientific analyses, and suggest that discrete measurements are a better option throughout the firn column.”