List of edits and disposition for V2 05/14/2020

The following edits are due to the new way of evaluating uncertainty that arises due to the unintended rotation of the ETMs.

1. Line 571: To estimate ϵ_{rot}, we consider the maximum conditions, when the Pcal beam offsets or center of force offset vector, $\vec{a}$, is aligned or anti-aligned with the interferometer beam displacement vector, $\vec{b}$. -----> talk about U distribution of $a b cos \theta$.
   Text updated as follows:
   The interferometer beam position offsets from center are 29\,mm for the X-end ETM and 22\,mm for the Y-end ETM. The uncertainty introduced by unintended rotation of the ETM, ϵ_{rot}, is proportional to the dot product of the Pcal and interferometer beam offset vectors, $\vec{a}$ and $\vec{b}$, i.e. $\epsilon_{rot} \propto |\vec{a}| |\vec{b}| \cos\phi$. Because ϕ, the angle between $\vec{a}$ and $\vec{b}$, is equally probable to be any value between -π and π, we use a sine wave distribution, i.e. a U-shaped probability density function\,\cite{bendat}, to estimate the variance in $ \cos\phi$ and form a Type\,B estimate for ϵ_{rot} using \eref{eq:pcaldisp}, $\epsilon_{rot} = Mab/(\sqrt{2}I)$. The values of the relative uncertainty estimates for both the X-end and Y-end stations at LHO are listed in \tref{tab:rot_values}. Because of these large interferometer beam position offsets, ϵ_{rot} is currently one of the largest sources of uncertainty for the LIGO Pcal systems.

2. Line 572 (later in the paragraph): ϵ_{rot} is currently one of the largest sources of uncertainty for the LIGO Pcal systems. -----> It is 0.40% for X end and 0.31% for Y end, making it the largest source of uncertainty in the displacement estimation of the ETM at the X end and almost equal to the NIST uncertainty for Y end.
   Left as is.
3. Uncertainty values updated in table 4 and 5.
4. Uncertainty in the displacement factor X increases: 0.52% for X end, 0.45% for Y end from 0.46% and 0.41%. Updated in Table 5.
5. Line 603: Values updated to 0.52% and 0.45% for X and Y end respectively.
6. Line 648 (1st sentence): Uncertainty not common to both the end stations have now increased to 0.41% and 0.31% for X and Y end stations. Updated.
7. Line 648 (Next sentence): Uncertainty relative to the two end stations is now 0.51%. Updated this value.
8. Line 648 (Next sentence): The 1.0046 is now at 0.9-sigma (within 1-sigma), ~36% chance of getting 0.46% off or more. Updated 1.1 sigma to 0.9 sigma
9. Line 650 (2nd sentence): The relative standard error on the weighted mean is 0.020% (should have been 0.20% in the first place). This value is updated to 0.25%.
10. Line 653: The 0.38 \,\% uncertainties in $X^{c}_{_X}$ and $X^{c}_{_Y}$ are slightly smaller than those for $X_{_X}$ (0.46\,\%) and $X_{_Y}$ (0.41\,\%). ------> The 0.41 \,\% uncertainties in $X^{c}_{_X}$ and $X^{c}_{_Y}$ are slightly smaller than those for $X_{_X}$ (0.52\,\%) and $X_{_Y}$ (0.45\,\%).
11. Uncertainty on the geometric mean, C_x and C_y, and X^{c} updated in table 6.
12. Double check if anything else has to change.

The following edits are due to increased Type B window for Pcal lab temperature.

1. Uncertainty on $\Delta T_{_{LN}}$ increased to 34%. Updated in Table 2.
2. Uncertainty on $\xi_{_{LN}}$ updated to 0.066% from 0.052% in table 2 and 6.
3. Uncertainty on $\xi_{_{EL}}$ updated to 0.40% from 0.20% in table 2.
4. Uncertainty on $\Delta T_{_{EL}}$ updated to 28% from 20% in table 2.
5. Check if this affects any other uncertainty.

The following edits are due to the removing error bars in the Fig 5.

1. Line 439: Ratio changed to 0.359% from 0.362%
2. Fig 5 will change. No error bars in the lower panel, 20x error bars on upper panel Sudarshan to do it
   Done.
3. Line 435 (caption): Need to talk about the systematics being larger than the statistical variations and hence the decision not to weight them by variance(rse^2). Sudarshan to do it

   Caption updated to add:
   The error bars have been omitted because the variations in the data are dominated by systematic variations between the measurement suites and not the statistical variations within each suite of between 25 and \textcolor{red}{ 1000} measurements. Weighting by the relative standard errors on the means would thus bias the estimate of the overall mean in favor

of suites with larger numbers of measurements.
What about the upper panel. Do we still use the error bars to estimate the step? I don't think so. So we need to note that they are shown, but not used for weighting.

Added to caption for upper panel:
Note that they are not used for weighting, as explained below.

Jeff's comments (05/07/2020)

%%%% Methodology %%%%
This section reads *so much* more clearly now. Excellent work!
- Equation (12) has its italics (or lack thereof) confused, with equations for X_{_X}^{c} and X_{_Y}^{c} differently italicized. No preference either way, but ever you chose, make sure equation 13 is consistent with your choice too.

Made equation 12 and 13 consistent.

%%%% Conclusion %%%%
These are more "for your consideration" than concrete suggestions.

- Order of paragraphs:
    First paragraph = Improvements, highlighting NIST improvement in power calibration is key.
    Second paragraph = Unintended rotation as second worst problem.
    Third paragraph = Discussion of how NIST and community want to continue to improve power calibration.
I wonder if changing the order to First, Third, Second, might flow a bit better? I understand your current flow: First & Second are "discuss the results in the paper, and explain the biggest limits to the results presented," then Third & Fourth are "talk about these results in the context of the

community, and talk about what folks are doing to improve the situation elsewhere / differently than us."
After re-reading, slightly editing, and re-considering, we have decided to leave the order as-is.

- Since you mention need for 1% *detector* calibration to be relevant in the Hubble “game" in the introduction, I felt a craving for a phrase or an extra sentence after you quote the 2% value from the O3A calibration paper, to remind the reader that some of that 2% comes from the absolute fiducial reference, but not all. Especially since the start of the paragraph goes in to a bit as to why *detector* calibration is hard. At the moment, the connection between the second-to-last sentence “In some frequency bands […] are below 2% [39]” and the last sentence “As the SNR of GW detections […]..." And then maybe an extra phrase / sentence bringing back the statement about Hubble?
New paragraph has been added that ties the 2% calibration accuracy achieved to what will be required to “continue to extend the scientific reach of GW detections.” We think this adequately addresses this suggestion.

Niko’s comments (05/18/2020)

> Line 374: Could the sentence be shortened? The GS digital multimeter could be a footnote.
>
> Text updated as suggested.
>
> Line 376: … Also, responsivity ratio measurements between one or more elements of this array of sensors and the GS → responsivity ratio measurements between th WSs and GS
>
> Text updated to remove “elements of array,” etc.

Line 519: Alternative sentence → In order to mitigate the impact of point absorbers in the mirror coatings, both LIGO observatories position their ifo beam off-center on the ETM. The rotation induced by this beam positioning can increase or decrease the sensed displacement.

Corrected an error in the text, made minor changes, but left structure as is, to introduce topic by beam offsets first, not point absorbers.

Rick's comments (05/18/2020)

Line 101: ORCID IDs for authors

Dripta and Niko got an ID. Sudarshan?

Line 179: Are we calculating vec{b} correctly?

Noted that we have not gotten to the bottom of this issue. It's something that could be investigated in the future. Rick recalls that Evan didn't recall if he had considered this "way back when" when Rick asked him about it.

Evan's comments (05/18/2020)

a) I like the interesting method of using the two end station measurements jointly, but I suggest some attention be paid to description of the end station comparisons and how \chi_{XY} is used in section 2. The text doesn't give me a clear sense of what is going on "underneath the hood". Perhaps writing out the formulas for geometric mean and combining measurements would help and slightly expanding the description with a few more sentences (or point the reader ahead to Section 3 if some more description is given there based on real measurements)

Added a new equation (eq. 13) and a few more sentences in section 2.2 clarifying the steps involved in calculating C factors.

b) there's not a strong argument of the choice to use a geometric mean. Perhaps that could also be fleshed out a bit with a few more words.

Text has been expanded to give more detail and motivation.

c) I don't see in table 6 where the unintended rotation is accounted for \epsilon_{rot}. I guess it's included in C…? This could be called out more explicitly because it's the second largest effect and I don't easily see where it enters.

Table 6 is a new addition that shows all the factors contributing to the calculation of C factors.

1) Confused about footnote labeling. Example on page 3 shows "8" and "9" but I thought I would see a symbol or something. Maybe something to fix with the journal editor. Also footnote "10" on page 10

The journal wants the footnote numbered in continuation to the affiliations of the authors.

2) Footnote "8" on page 3 - should join references [21][22] --> [21,22]

Corrected.

3) Page 3 units of R(\omega) is hard to tell if the "m" is supposed to be in the denominator or numerator. Maybe try something like 1/(N m) or similar

Corrected.

4) Suggest change "The remainder of this document is organized as follows." --> "This article is organized as follows."

Changed as suggested.

5) This sentence is a little bit cryptic: “Sensor calibration factors are incorporated into displacement factors that convert the power sensor outputs to estimates of ETM displacement.” Does it need to be here?

Agreed. Sentence removed as suggested.

6) Page 5: the word “chamber” is unclear to a new reader as it’s the first time used. Consider a small rewording?

Changed “chamber” to “vacuum chamber” and used “vacuum chamber” later in the sentence too.

7) Page 10: suggest changing “But errors induced by uncertainties in factors that are not common to both end stations (…)…” to “Errors induced by uncertainties in factors that are not common to both end stations (…), however, …”

Changed “But,” to “However,”

8) In tables, the variable labels are sometimes offset in the column. Fine to proceed, but will need to be corrected for the journal.

This was intentional. Factors that contribute to the main factors are indented to indicate that they are contributing factors. We will try to find a way to emphasize this aspect.