Comments from Adam Beardsley (Arizona State University, moving to Winona State University in Minnesota) on July 15, 2020

Sorry again for taking so long to read this over and provide comments. I've just finished Sec 7, and want to send along what I have so far so you know I'm still working on it. I'm also in the process of packing up our house, so might be disconnected from science any day. I'll try to send the rest of my comments in the next couple days, but if I go silent, it's because I'm driving across the country.

It really is a great body of work! The writing is very clear, and I enjoy the way it's broken up into a collection of experiments and tests to verify the system. I've provided some comments/ thoughts below - let me know if anything is unclear.

More to follow, including my thoughts on 8.9.

Cheers,
-Adam

Comments

Bottom left par on p. 5: yellow dots in Fig 4?

Are H and V too confusing? e.g. H-plane (horizontal polarization vs magnetic plane of dipole), V-dipole (could be confused with stokes V?)

"A 1% uncertainty in the beam measurement may introduce a 4% systematic uncertainty in the final power spectrum." I'm not certain what this sentence is trying to say, and how this translation is done. But it sounds like an understatement to me. Depending on the nature of the beam error, 1% could completely wash out the signal (which is ~0.01% the magnitude of the foregrounds).

Fig 7: Sim seems to match the frequency of ripples, but not necessarily the phase (bottom panel). Is there a physical explanation for the ripple? It's also interesting that the measured FWHM for H-baseline is lower than expected. Not sure what to make of that, if anything.

Regarding redundant calibration, there is a technique being developed by a few groups to incorporate non-redundancy. Jon Sievers (now at McGill) proposed it a few years ago in an unpublished article (https://arxiv.org/pdf/1701.01860.pdf). Ruby Byrne (U-Washington) has a similar technique, I'm not exactly sure the difference, but the UW crew insists it's different (https://ui.adsabs.harvard.edu/abs/2020arXiv200408463B/abstract). I'm not completely convinced of the technique in general, but something to keep an eye on.

Fig 12: Do the tasks create new data containers? For example, does Task1 take Data container 1 as input, and create Data container 2 as output? Does this overwrite Data Container 1 or create a new file? Data sets are so far relatively small, but you may want to think ahead to ways to not multiply data volume (e.g. HERA never saves calibrated visibilities, only the calibration solutions separate from the raw data, then applies on the fly).

The expression for relative calibration, I'm not sure I understand why you divide by the amplitude of the noise source visibility. Maybe it's meant to be the amplitude of ON minus OFF, with the assumption that the noise source amplitude is constant?

tlpipe seems like a really useful toolset. I wonder about its wider utility - how might it interface with different file formats (from what I remember Tianlai's files are unique)?

Fig 13: Wow, higher amplitude in the "cross-pol" (H-V)? Is the CNS circularly polarized? Small note: missing) in y-axis label in top plots.

Sec 6.2 shows amplitude gain stability, and argues the variability seen from the CNS is due to it coming through the sidelobes. I agree this is probably the case, but was the stability of the CNS itself ever verified?

Sections 6.2 and 6.3 have minor typos (usually spurious s's)

Fig 18, that bimodal distribution is quite interesting! My best guess is that you had roughly equal number of warm and cold days? Or maybe something in the system changed during the 12 day run… have you tried breaking up the data and histograming subsets to isolate the two distributions?

Sec 6.4, first paragraph, missing table reference

Section 7 lacks some motivation. I'm all for making images - it's the most sure fire way of knowing your telescope is "looking up." But it's not clear to me why several imaging algorithms are used, and there isn't much analysis to compare between them. Specific comments below.

Fig 20: I'm not sure I understand the procedure here. It was divided into three sections, but it would appear that CasA is by far the brightest thing in the entire strip, so what was being cleaned in the other panels? Is it possible you ended up cleaning CasA's sidelobes in those other panels? The text says the gaussian beam was assumed - is this a gaussian synthesized beam? If so, I don't understand how the sidelobes could be supressed. In general, I'm not sure I agree the BFMTV has improved the map. There are clearly additional, non-physical features in the bottom panel, and importantly the reader doesn't get to see how these features vary in frequency - which could be a show-stopper for cosmology.

Fig 23: It's hard to know without numbers quoted, but the dynamic range here looks substantially lower than in Fig 21. If the algorithms were behaving similarly, I would expect the additional frequency synthesis would clean up side lobes some. But they appear bright and blurry in Fig 23 compared to 21. A natural question for a referee will be - why not use the same data and calibration for these two images for apples-to-apples comparison? What conclusion am I meant to draw here?