Lab 2

Lukas Finkbeiner

March 6, 2020

Abstract

We investigate the properties of the signal from the hyperfine hydrogen transition with a collector dish connected to an ADC by several layers of mixers and filters, so that we are left with frequencies more amenable to currently-realizable sampling rates. We calibrate our instruments to the thermal noise of our system and by averaging over many blocks of consecutive data. We use humans, as studied blackbodies, to demonstrate the principle behind thermal calibration. We also include rotation matrices and our timestamps for some of the data that we collected, as an illustration of the Doppler corrections to be applied when converting frequencies to Doppler velocities on a power spectrum. We offer examples of all these processes with a measurement of the HI line at zenith and when pointing at Cassiopeia. In a second experiment, we investigate the speed of light by measuring minima in the magnitude-squared voltages of standing waves in a waveguide. Only able to get meaningful results for the using waveguides.

1 Introduction and Background

In the first of two experiments, we seek to observe the 21-cm HI line, at 1420.4 MHz. However, we want to look at frequencies better suited to our lab equipment. For example, our PicoScope has a maximum sampling rate of 62.5 MHz, so anything above 125 MHz would be aliased.

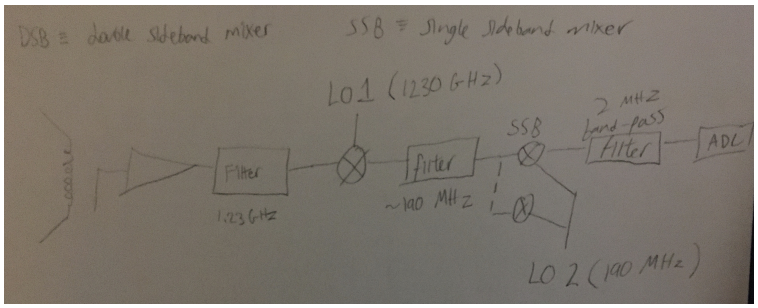

Figure 1: To reduce the frequencies of incoming signals to manageable levels, we employ a multi-stage combination of filters and mixers. The signal enters via transformer connected to the collector. The ADC represents the final step, when the PicoScope translates voltages over time into digital arrays.

The data that we sample from the Big Horn via the PicoScope will feature arbitrary units which depend on our setup. To calibrate the intensity of the spectrum, we first calculate the gain:

$$G = \frac{T_{\text{sys, cal}} - T_{\text{sys, cold}}}{\sum(s_{\text{cal}} - s_{\text{cold}})} \sum s_{\text{cold}} \approx \frac{300 \text{ K}}{\sum(s_{\text{cal}} - s_{\text{cold}})} \sum s_{\text{cold}} \tag{1}$$

Where $T_{\text{sys, cal}} \approx 98.6^\circ$ F ≈ 310 K represents the temperature of the human flesh that we use to cover most of the telescope. $T_{\text{sys, cold}}$ represents the temperature of the cold sky. We expect $T_{\text{sys, cold}} \ll T_{\text{sys, cal}}$, whereby we get the approximated form on the right. s_{cal} represents the spectrum for which we attempt to maximize thermal noise in the collector, and s_{cold} represents the spectrum for which we attempt to minimize.

To remove constant sources of interference in our data and to obtain the shape of the line, we divide an 'on' spectrum by an 'off' spectrum.

$$s_{\text{line}} = \frac{s_{\text{on}}}{s_{\text{off}}} \tag{2}$$

The fully-calibrated spectrum represents a scaling of this s_{line} by the thermal gain from equation 1:

$$T_{\text{line}} = s_{line} \times G \tag{3}$$

To plot our calibrated spectrum against Doppler velocity, we use the following relation:

$$v = -c\frac{\Delta\nu}{\nu_0} \tag{4}$$

Finally, to account for the times at which we observed, as well as to convert from galactic coordinates to topocentric coordinates (for which telescope pointing is more intuitive), we used the following rotation matrices:

Equatorial to galactic (radians):

$$\begin{bmatrix} -.054876 & -.873437 & -.483835 \\ .494109 & -.444830 & .74698 \\ -.867666 & -.198706 & .455984 \end{bmatrix} \tag{5}$$

Equatorial to hour-angle:

$$\begin{bmatrix} \cos(\text{LST}) & \sin(\text{LST}) & 0 \\ \sin(\text{LST}) & -\cos(\text{LST}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

Hour-angle to topocentric:

$$\begin{bmatrix} -\sin(\text{latitude}) & 0 & \cos(\text{latitude}) \\ 0 & -1 & 0 \\ \cos(\text{latitude}) & 0 & \sin(\text{latitude}) \end{bmatrix} \tag{7}$$

To apply these matrices, we convert between angle pairs and rectangular vectors via

$$\vec{x} = \begin{bmatrix} \cos(\phi)\cos(\theta) \\ \cos(\phi)\sin(\theta) \\ \sin(\phi) \end{bmatrix} \tag{8}$$

and

$$(\theta, \phi) = \text{atan2}(y, x), \sin^{-1}(z)) \tag{9}$$

For our second experiment, we seek to calculate to calculate the wavelength λ_{sl} of a signal inside the waveguide based on the positions of the minima in the squared voltage, which we call nulls. We expect a linear dependence of null spacings on λ_{sl}. We use the following equation for this first experiment:

$$x_m = A + m\frac{\lambda_{\text{sl}}}{2} \tag{10}$$

m is the index of the null, x_m is the position of null m, and A represents a constant offset which can relate whether the far end of the waveguide is open or shorted.

The rectangular waveguide, to which we switch for the X-band phase of the experiment, abides a nonlinear relationship between the waveguide's width and the wavelengths of the input in the waveguide and in free space.

$$\lambda_g = \frac{\lambda_{\text{fs}}}{[1 - (\frac{\lambda_{\text{fs}}}{2a})^2]^{1/2}} \tag{11}$$

λ_g is the guide wavelength, which we directly measure. λ_{fs} is the free-space wavelength of the signal, which we will calculate using the known input frequency and the relation $c = f\lambda_{\text{fs}}$. a is the width of the waveguide.

Finally, to analyze our results we define the reduced chi-squared function as follows:

$$\chi_r^2 = \frac{1}{N - M}\sum_i \frac{|y_i - \hat{y_i}|^2}{\sigma_i^2} \tag{12}$$

Where N is the number of data and M is the number of fit parameters. For each measurement i, y_i is the observed value, $\hat{y_i}$ is the value predicted by the fit, and σ_i is the expected error on that measurement.

The reduced chi-squared function will provide something clear to minimize when fitting the data to a model. Specifically, we want to calculate the width a based on our observations of null spacings, so we will calculate the value of a which minimizes the reduced chi-squared.

2 Methods

After setting the two local oscillators to 1230 and 190 MHz, as labeled in figure 1, we set the power levels to 8 and 0 dBm, respectively. These numbers are difficult to theoretically justify; through trial and error, and particularly through consultation with other groups, we found that these numbers led to the greatest HI resolution for the Big Horn pointed directly upward (zenith).

However, these numbers changed when taking data for Cassiopeia–$(120°, 0°)$ in galactic coordinates, transformed to equatorial via the rotation matrices from the introduction–by which point our signal-to-noise ratio was too low. We switched to 10 dBm for both and proceeded with the pico-sampling at 62.5 / 6 MHz, the rate we used for all captures.

Furthermore, for each source, we took 2 blocks of 10,000 samples of 16,000 time-voltage tuples. We labeled one block as the 'on-spectrum' and held the first local oscillator at 1230 MHz. We then switched to 1231 MHz (to push the HI signal closer to the 2MHz band-pass filter cut-off) and labeled that second block as the 'off-spectrum.'

The results as they are shown here are the products of reducing these larger data blocks into single arrays of 16,000 time-voltage^2 tuples via unweighted averaging; first, all 10,000 samples in a block are Fourier transformed, then we average the square magnitudes of each Fourier transform. We chose to reduce via the mean rather than the median, because the aforementioned attenuation of the PicoScope was inconsistent and thus the mean was more effective at capturing the true signal.

Later on, we will see imperfections in the setup in the form of nearly symmetric 'bunny ear' spikes as well as an extreme spike at the center; we found the PicoScope to sometimes artifically attenuate and offset the signal going into the second port, because we would switch the cables going into A and B and we would see the same shape; if there were a problem in our mixer stages, we would expect the output to switch similarly.

To obtain the gain necessary for normalization of the spectrum (equation 3), we took 100 samples with the Big Horn pointing up at the cold sky (the reduction of this creates our cold spectrum) and 100 samples with three humans attempting to cover up the collector (the reduction of this creates our hot spectrum).

We measure λ_{sl} and its counterpart λ_g, the guide wavelengths, as the distance between the nulls in the guide output. We measure a with a set of calipers and attempt to compare this with our results from a least-squares fit to the data based on equation (ultimately I failed to produce results) 11.

3 Observations

We collected data over many days of troubleshooting. The thermal data, as pictured in figure 1, have a straightforward gap between them, but our gain calculation (of about 30) was well below that of other groups.

Other results surely point to some misunderstanding regarding the setup of the filters and mixer stages. Consider the rippling spikes which run along figure 2; the noise detracts from our results, and seems to vanish (at least, as a clear problem) when we raise the local oscillator power levels for the Cassiopeia measurements (the comparative smoothness will become obvious with the fully calibrated spectra in the analysis section). However, when we raised power levels for this initial HI line capture, the signal got drowned out somewhere.

For the sake of brevity, we will plot the data for the 3GHz section later, in the analysis, along with the least-squares fit model.

All measurements come with an intrinsic reading error ($\pm$.025 cm) and, more importantly, poor resolution of the null's location. This poor resolution was partially due to variation in human judgment, but principally to the low signal-to-noise ratio. Near the null, the

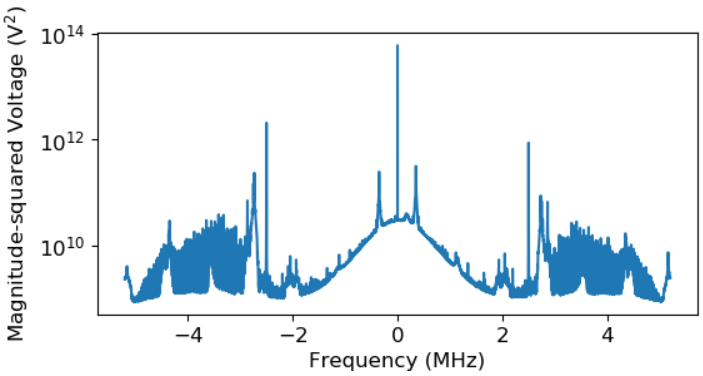

Figure 2: A semi-log plot over the range of frequencies sampled. We assume that our 2 MHz low-pass filter works, so we will dismiss the signals farther than 2 MHz from the center. Additionally, the large central spike and smaller 'bunny ear' spikes (±.5 MHz) appear on all data sets as persistent interference; we partially ignore these by limiting the y-axis.

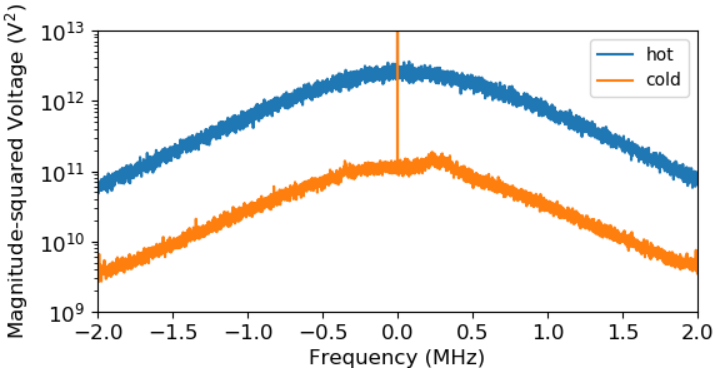

Figure 3: The 'hot' data correspond to three humans standing in front of the collector. The 'cold' data correspond to the collector pointing up at the cold sky. The noise is not an issue, because it is regular: observe the even discrepancy between the curves over the domain of interest.

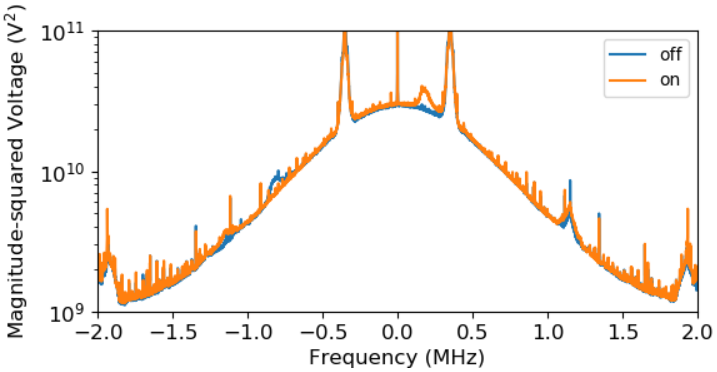

Figure 4: Combined plot of the 'on' and 'off' (LO1 at 1230 and 1231 MHz, respectively) power spectra. As we expect, the HI signal shifts by about 1 MHz between the two plots. This also supports our interpretation of the other patterns as interference: these patterns do not move between spectra.

derivative of the magnitude-squared voltage with respect to waveguide slider position is at its most shallow. Consequently, there is a sort of range of positions over which the same null can be argued to appear. We decided that this uncertainty probably corresponds to a universal uncertainty of $\pm.2$ cm. We measured the width a of the XBand waveguide to be 22.29 mm to, nominally, within ten microns.

f (GHz)	Null Positions (cm)	f (GHz)	Null Positions (cm)
7.0	8.8, 15.15	9.5	8.9, 11.15, 13.35, 15.45, 17.7
7.2	11.95, 17.15	10.0	8.45, 10.45, 12.3, 14.35, 16.5
7.5	10.05, 14.4	10.5	9.95, 11.75, 13.5, 15.45, 17.3
8.0	10.15, 13.35, 16.7	11.0	9.35, 11.4, 12.65, 14.45, 16.2, 17.85
8.5	9.85, 12.7, 15.25	11.5	8.95, 10.45, 12.0, 13.8, 15.15, 16.8
9.0	9.3, 12.3, 14.25, 16.7	12.0	8.4, 9.85, 11.35, 12.85, 14.4, 15.9, 17.45

4 Analysis

Since the Horn was pointing straight up, we may say that the declination is equal to the latitude of the telescope. In our case, that is $37.873199° \approx 0.661012$ radians. In the directly upward case, we also may say that the right ascension is equal to the local sidereal time (LST) of data-capture.

We labeled all of our samples with unix time. Included below is a table including examples of date conversions to Doppler corrections for the events of interest. For the sake of brevity, we omit the counterparts for the Cassiopeia measurements.

Label	Unix Time	Julian Date	LST
On-line start	1582937209.9447305	2458908.5325225084	0.8334479646673989
On-line stop	1582937704.4486437	2458908.5382459336	0.8695077621535888
Off-line start	1582937837.5090616	2458908.5397859844	0.8792106794915676
Off-line stop	1582938387.1761112	2458908.5461478718	0.9192930359644231

Label	PST	Doppler Correction (m / s)
On-line start	2/28/20 @ 16:46:49	-27648.19266498681
On-line stop	2/28/20 @ 16:55:04	-27897.548197380296
Off-line start	2/28/20 @ 16:57:17	-27960.009947759394
Off-line stop	2/28/20 @ 17:06:27	-28197.027228226034

The Doppler correction exhibits, at first glance, a critical time sensitivity. Observe that, just for one source (directly overhead), the final Doppler correction differed from the initial by about 550 m / s. However, as we shall soon see, the velocity axis will span many megameters, so taking the average of the Doppler corrections should suffice for a global correction.

Now, our on-off method is somewhat arbitrary, and the frequency at which a single HI line appears will change as we change the frequency of the first local oscillator. Furthermore, the very formula for producing the velocity axis (equation 4) stipulates a primary frequency against which other frequencies are compared.

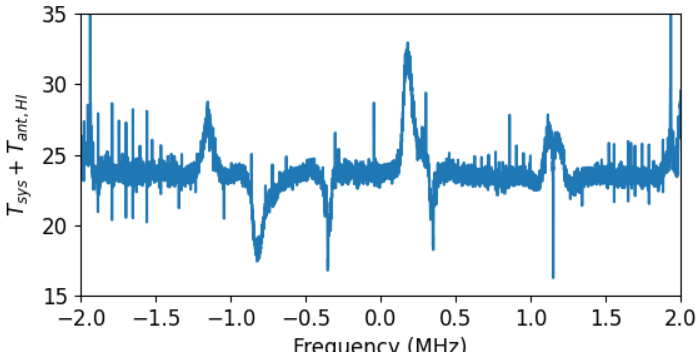

Figure 5: Here is a fully calibrated line. As mentioned before, there is plenty of noise with which to contend. Discounting the spikes at the cutoff of the low-pass filter, we can see that the HI line ($\approx$ .15 MHz) is the most prominent peak, and we can also see the impression left on the left side of the spectrum, by dividing out the 'off' signal (wherein the HI line was simply shifted, and not eliminated, from our data)

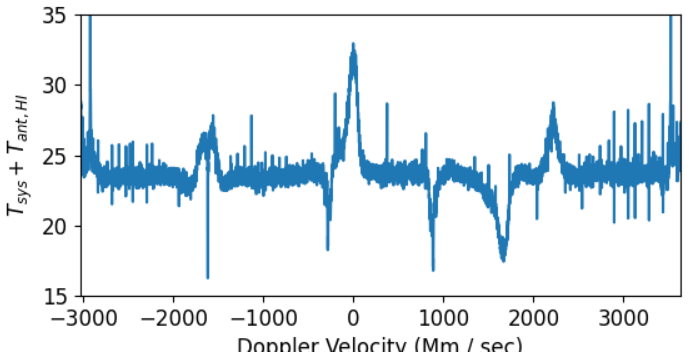

Figure 6: Here is that same line, now with a velocity axis. The width of the peak can tell us about the temperature of the transitioning hydrogen. However, the only other spikes correspond to instrumental noise, so there are no meaningful relative velocities to compute here.

Consequently, the single Gaussian does not tell us much, except perhaps the rough temperature of the cloud by the width of the Gaussian–as the spread in velocities increases, so too must the average kinetic energy of the gas particles. The intensity of the line from the cloud (i.e. the amplitude of the Gaussian) woud require additional information for us to learn something intrinsic about the cloud, like luminosity. Figure 6 is our example of this.

By contrast, the distances between average values of *separate* Gaussians on the same graph would give us direct information on relative velocities between the two corresponding clouds of hydrogen.

Figure 10 represents a potentially more obvious flaw in experimental setup. waveguide represented a We estimated there to be a range of about .4 cm over which the waveguide output could be said to be minimized.

The least-squares model did not largely improve on a simple averaging of differences, but it may be coincidence that the data came out in this fashion. A simple average of the null

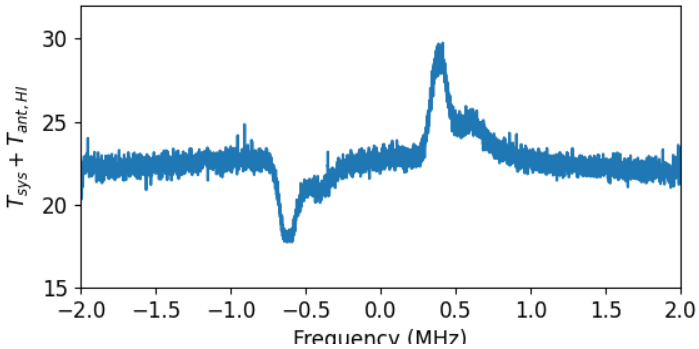

Figure 7: This is the fully calibrated spectrum for our Cassiopeia data. Notice how much smoother this curve is than its counterpart from when we pointed the Big Horn directly up. We also see that, although we implemented identical frequency switching tactics for both sets of data, we see the line land much more closely to the anticipated value of .4 MHz away from 0 (which, taking back through the mixer stages, should correspond to 1230.4 MHz)

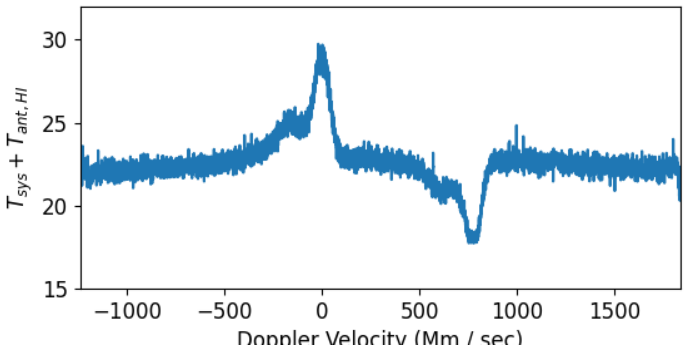

Figure 8: Just as before, we can modify the calibrated power spectrum to use an x-axis of velocities. Because we have two separate real (as opposed to instrumental) signals, we can procure information on relative velocities of the sources, if we apply a Gaussian fit.

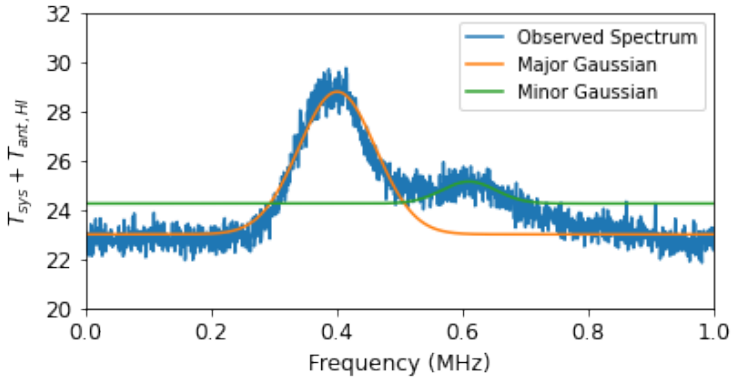

Figure 9: A close-up of figure 8, where we have used the ugradiocode.gauss module to calculate best fits, assuming that each lump here is a Gaussian. The 'major' Gaussian is a very strong fit: it hugs the data closely and over a long interval. By contrast, we may doubt the statistical significance of the 'minor' Gaussian: the interval, over which the fit and the data align, is comparatively thin. Furthermore, the ratio of the Gaussian's amplitude to the noise along the curve is quite low.

spacings for the open 3GHz waveguide gave a wavelength of 5.00556 cm ($c \approx 15.017 \times 10^5$ cm/s), while a least-squares calculation gave 5.00758 cm ($c \approx 15.023 \times 10^5$ cm/s).

We calculated the fit for the closed waveguide to be $x_m = 5.00167m + .43611$ ($c \approx 15.005 \times 10^5$ cm/s) and $x_m = 5.00758m - .57667$ ($c \approx 15.023 \times 10^5$ cm/s). I lack specific uncertainties on those parameters. The discrepancy in the y-intercept is the parameter of interest to the comparison of the two types of standing waves. Reflections from open ends will invert as they travel back, whereas reflections from shorted ends will maintain prior sign. Consequently, we observe that the voltage-squared pattern shows a simple translation in the positions of the nulls.

We calculated the standing wave ratio on the XBand to be 2.56. The standing wave with the waveguide open at the far end featured a higher amplitude because the impedance was matched far less well than it was for the shorted data. Therefore, more of the wave was reflected.

I was unable to correct my chi-squared minimization implementation for the XBand section.

5 Conclusions

In conclusion, we obtained results strongly in favor of expected values for the HI line observation. Despite hiccups in the experimental design manifesting as noise spikes in our first pair of blocks, we picked out a likely candidate for the HI line there. The Cassiopeia observations were even more promising. We considered how fitting individual Gaussians to the lumps on fully calibrated power spectra can provide insight into the temperatures and relative velocities of different clouds of hydrogen, when we exchange the frequency axis for doppler velocities via equation 4. We briefly probed some least-squares fits in a simple linear scenario (measuring the speed of light), questioning our assessments of error in the data, and

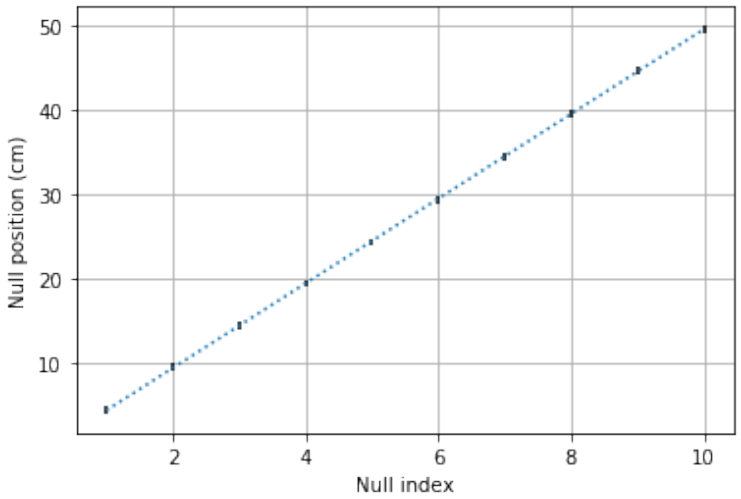

Figure 10: Our null displacements for the 3GHz waveguide are plotted alongside the line of best fit. Despite the appearance of this graph, it *does* have error bars on it. If there is an issue with error estimation, it is here visible and thus guides our meditation on the results from the XBand section.

failing to affirm the superiority of the least-squares approach. We highlighted the relevance of voltage standing wave ratio to impedance matching, and reiterated the theory behind the XBand data-fitting without success with actual data.

6 Acknowledgments

I collected and re-collected data for the zenith and Cassiopeia spectra. Mehdi calculated the on and off frequencies for the local oscillator and helped me debug my data-taking code. Mehdi and Rebecca collected the XBand waveguide data, and I helped with the 3GHz data. Rebecca also helped to lay out data-collection schedules.

Theory and background provided by Aaron Parsons. "LAB 2: Astronomy with the 21-cm Line and Waveguides." Updated February 2020.