Synchronizing ALBATROS Antenna Data Using The ORBCOMM Satellite Network

William Frost[a] **and Jonathan Sievers, PhD**[a]

[a]Department of Physics, McGill University, 3600 Rue Université, Montréeal, QC, H3A 2T8, Canada

This manuscript was compiled on April 24, 2020

This project aims to quantify the performance of the ORBCOMM satellite network intended for use in synchronizing ALBATROS antenna data. This synchronization requires nanosecond timing errors if an ALBATROS antenna array is to be capable of observing the sky at low radio frequencies (< 100 MHz) using radio interferometry. To achieve this, we first use a simple wire antenna hooked up to ALBRATOS front-end and back-end electronics and measure orthogonal polarizations of ORBCOMM satellite signals. Cross-correlating these two signals therefore gives insight into expected timing errors when using ORBCOMM, and constrains the integration time needed to obtain the desired nanosecond accuracy. Using data from a limited observation time of 2 minutes, we obtain time errors in the cross-cor between the orders of 10^0 ns and 10^{-1} ns, which are good results given the operating frequencies for an ALBATROS antenna array.

Radio Astronomy | 21-cm Cosmology | Antenna Synchronisation

The field of 21-cm cosmology aims to observe the universe using the faint, red-shifted, 21-cm emission line of neutral hydrogen (1)(2). Since hydrogen has remained the most abundant element throughout the history of the cosmos, this emission line could allow for an unprecedented mapping of matter distribution in our universe. It would allow for observations into epochs such as the Dark Ages, the period following the release of the Cosmic Microwave Background. For these early epochs, their corresponding red-shifted 21-cm hydrogen lines are observed on Earth at radio frequencies of less than 100 MHz and at less than 20 MHz for the Dark Ages (3).

Notable experiments that have delved into observing these low frequencies include Grote Reber's *Cosmic Static at 144 Meters Wavelength* (4) looking at 2.1 MHz and the RAE-B satellite, sent in orbit around the moon, which was able to observe at less than 10 MHz (5). However, these experiments were conducted in the 60's and 70's with angular resolutions in the data of no less than 5^o (6), a resolution 10 times greater than the angular diameter of the moon at 0.5^o.

Therefore, improved measurements in those frequency ranges are the focus of recent initiatives. One of these is ALBATROS, the Array of Long Baseline Antennas for Taking Radio Observations from the Sub-antarctic, operating out of remote sites with low radio frequency interference (RFI) like the McGill Arctic Research Station (MARS) and Marion Island.

One of the goals with ALBRATROS will be to observe at frequencies of $\sim$ 10 MHz. To achieve this, the cross-correlation of signals from an ALBATROS antenna array requires accurate time-keeping capabilities, which directly impacts the inferred sky location of celestial sources. Due to the remote nature of ALBATROS sites, the goal of this project is to characterize the timing error offered by commercial satellite signals from ORBCOMM, acting as a clock mechanism, as a function of integration time.

This is accomplished by getting data from an ORBCOMM signal observed using one dual-polarization dipole antenna. Since both polarizartions originate from the same source, they should effectively possess no time delay between them, allowing us to reconstruct expected timing errors due to noise and to constrain the integration time necessary for nanosecond timing.

Background and Methods

Polyphase Filter Banks. To acquire data, the ALBATROS project uses an ensemble of Long Wavelength Antennas (LWA) (7), each performing a Polyphase Filter Bank (PFB) front-end on incoming signals and then passing them through a Discrete Fourier Transform (DFT). A DFT is defined as follows:

$$X(k)=\sum_{n=0}^{N-1} x(n)e^{-i2\pi nk/N} \quad [1]$$

Applying a PFB on a signal corresponds to multiplying said signal with a window function, splitting it into equal chunks and summing over those chunks 1.

The main justification as to using a PFB front-end compared to simply performing the DFT on raw signals lies in improvements to power leakage in the Fourier Transform (FT). This happens due to an input signal ending up spread out in multiple frequency channels. Using a PFB also limits scalloping loss (loss in energy) between frequency bin centres due to the non-flat nature of the single-bin frequency response.

Significance Statement

21-cm cosmology hopes to ultimately map vast parts of the universe previously invisible to our instruments. Observing this redshifted 21-cm signal requires, among many challenges, minimal radio interference. This pushes observation sites to be located far away from civilisation in hopes of limiting RFI. With ALBATROS, we also desire to minimize antenna power consumption. On top of the base components, having on-site clocks would increase the power requirements for long observation periods. Therefore, determining if satellite signals (namely from ORBCOMM) can act as clocks and offer the desired synchronization capabilities for ALBATROS would improve its overall power consumption, since we require no additional power to observe these signals along with our actual signals of interest.

[1]To whom correspondence should be addressed. E-mail: william.frost2mail.mcgill.ca

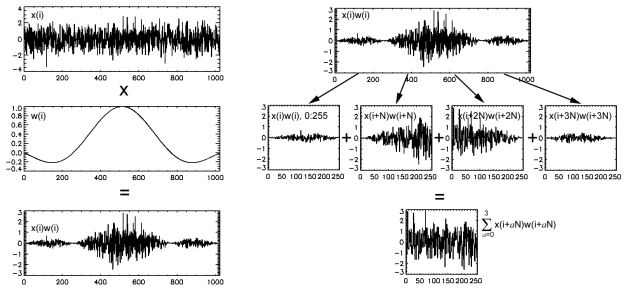


Fig. 1. Example of how a PFB front-end works. Here a part of a time stream is multiplied by a window function (sinc function) of the same length and split into 4 blocks. Then all equal indices of those blocks are summed together to obtain a new time stream that is 4 times shorter than the original. This is then input into a DFT. (8)

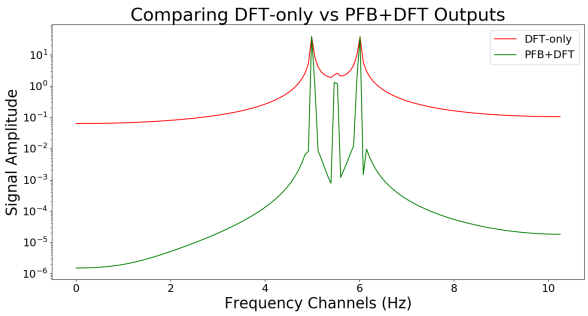


Fig. 2. Log plots of a PFB and an FFT (efficient DFT algorithm) performed on a simulated signal containing 2 large-amplitude signals at 5 and 6 Hz with a smaller signal at 5.5 Hz. The sampling rate was 20.5 Hz with an observing time of 400 s. Both methods are using the same number of channels. Using only a DFT does not allow the faint signal to be adequately detected, while applying a PFB front-end to the signal greatly reduces leakage and clearly shows the previously undetectable signal.

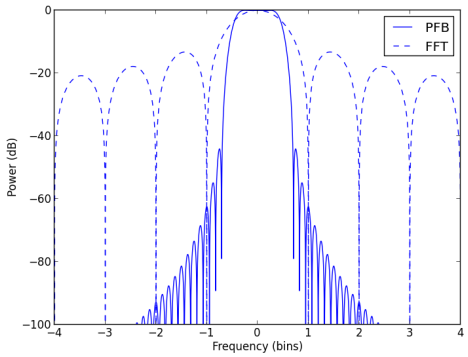


Fig. 3. The single-bin frequency response of a PFB compared with a FFT. Here, the length of the polyphase window is 8 times the length of the FFT. The narrow nature of the frequency response for a PFB shows why scalloping loss is less of an issue than with a simple FFT. (9)

One weakness of the PFB technique is that it is not an exactly invertable operation. However, an approximate inversion that has low residual difference with the original signal (apart from at the signal extremities) is possible and the code base used to implement it is taken from Richard Shaw's (UBC) pfb-inverse repository (10).

Signal Acquisition. To first acquire ORBCOMM signals, a dual-polarisation, wire dipole antenna hooked up to a LWA front-end and SNAP FPGA board back-end (11) is used to obtain many PFB channels of satellite data. This transformed data is generated using 2048 frequency channels, time stamped and then correlated off-site. It is important to note that we do not necessarily store all 2048 channels, but only keep those of interest to us since not all of the observed bandwidth is needed.

Even though the advantages of the PFB technique are not necessarily needed to extract the ORBCOMM signals at the locations where ALBATROS will operate, given that the hardware used to acquire data for ALBATROS does implement it to observe cosmological signals, we need to make sure that we can get adequate timing data from that PFB hardware. Additionally, when observing ORBCOMM signals in potentially RFI-heavy places, such as what was done for this report, the PFB can come in handy to mitigate leakage of nearby radio sources into our channels of interest.

Additionally, the ORBCOMM signals observed operate within 137-138 MHz. The SNAP board ADC samples at 250 MHz, meaning it adequately represents signals up to 125 MHz. Due to ALBRATROS being deployed in remote regions with low RFI, the fact that the clock signals will be aliased should not be a factor in the field as there will most likely be no 112-113 MHz sources that could blend in with ORBCOMM observations.

Offline Correlation. Since ALBATROS data is correlated separately from the observation process, computationally heavy calculations can be run on better hardware and thus decrease the power consumption of an antenna over an observation period. Once data is collected from the SNAP board, the process becomes:

1) Auto-correlate the ALBRATROS data to pick out satellite signal regions in Fourier space.
2) Invert back these satellite signals to obtain a time-stream for both polarizations of the signal
3) Cross-correlate those satellite time-streams and quantify the timing error we can have with ORBCOMM satellites.

The cross-correlation h of 2 time-domain signals $f(t)$ and $g(t)$ as a function of time-delay dt is defined by

$$h(dt) = \ < f(t)g(t+dt) > \ . \qquad [2]$$

In Fourier space, one can use the convolution theorem to express the FT of h as

$$H(k) = \overline{F(k)} * G(k) \ , \qquad [3]$$

where $H(k)$, $F(k)$, $G(k)$ are the fourier transforms of $h(dt)$, $f(t)$, $g(t)$ respectively and $\overline{F(k)}$ represents the conjugate of $F(k)$. Using this property, $h(dt)$ can be determined as

$$h(dt) = \mathcal{F}^{-1}\left(\overline{F(k)} * G(k)\right) \qquad [4]$$

where $\mathcal{F}^{-1}$ is the inverse FT. Furthermore, in Fourier space, auto-correlation can be thought of as the cross-correlation of identical signals such that

$$H(k) = \overline{F(k)} * F(k) = |F(k)|^2 \ . \qquad [5]$$

The Fourier operations shown in Eq. 4 can be performed with or without a PFB front-end. However, once the data is saved, there are no limitations set by hardware to the amount of baseband data you can funnel into a FFT. Because the PFB front-end becomes useful when constrained to small sample sizes (as is the case with the SNAP board), it becomes unnecessary to use it in the case where we do not have this limitation. An FFT is thus preferred for the offline cross-correlation. Any DFT, as well as Eqs. 5, 4, can be performed using relevant python numpy functions.

Additionally, one can use techniques to improve smoothness in the cross-correlation. When performing a real FFT (i.e. np.fft.rfft()) on real input, one can append zero-values to the end of the FT and invert back to obtain a finer interpolation between the points in real space. Since the goal of our cross-correlation is to find its peak and evaluate how precisely we can infer its location, the zero-padding approach can become useful as it retains roughly the same Signal to Noise Ratio (SNR) while allowing for the peak of the cross-correlation to be resolved with more points.

Quantifying Timing Error. To obtain a timing error from the cross-correlated ORBCOMM signals, a Monte Carlo approach is implemented using random shifts in the (correlated) noise of the signals to introduce uncertainty in the data and fit a parabola about the cross-correlation peak. The standard deviation in the time-values obtained for this approach will be defined as the timing uncertainty provided by the ORBCOMM signal, given a certain observation time.

Results/Discussion

Because the SNAP board has a high sampling rate, even a few seconds of observation time takes up considerable storage. Therefore, fewer channels (26 to be exact) than the 2048 used by the SNAP board were sent to W. Frost allowing for $\sim$ 2 minutes of total observation time to be analysed with reasonable storage requirements. Initially, only 5 seconds of this total will be analysed and discussed, with the whole 2 minutes being analyzed further into the report.

Due to the data lacking resolution in only 26 channels as seen in Fig. 4, inverting back to a time stream allows for a better re-channeling of the original data and better visualization of the satellite channels. Fig. 5 shows the re-channelized data (using a PFB) and the better resolved frequency ranges corresponding to satellite signals.

We can manually select regions of interest (where we think a satellite signal is located) and create a new spectrum for each satellite we want to isolate, zeroing out everything outside a certain satellite's channels. We can take, for example, the channel interval housing satellite #2's signal, invert both polarizations to obtain time-streams of each and perform the cross-correlation on the polarizations. Fig. 6 shows this process performed on satellite #2 in Fig. 5.

To quantify the noise in Fig. 6, an overestimate of it is obtained by cross-correlating different polarizations of 2 different satellites. Fig. 7 shows what this (correlated) noise looks like.

With a cross-correlation and its respective noise in hand, the correlation peak can now be fit using a Monte Carlo approach. A least-squares, quadratic fit is performed many times to the points about this peak while randomly shifting the noise of Fig. 7 and adding it in to Fig. 6. By adopting this approach, we are able to conserve the correlated nature of our noise in the Monte Carlo process. The standard deviation in the peak's time-coordinate is then extracted, corresponding to our timing error estimate. In the previous case where

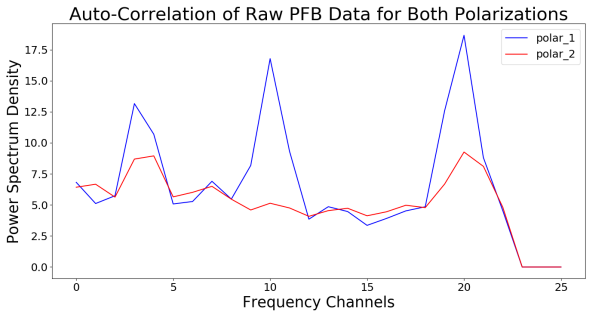


Fig. 4. This shows the unpacked, time-averaged and auto-correlated LWA data with 26 channels. It is obtained from $\sim$ 5 seconds of observation. We can observe three major peaks exhibited by the first polarization (blue), corresponding to 3 different ORBCOMM satellites. The left and right prominent blue spikes are shown to clearly shadow their opposite polarizations (red), lending credence to the assertion that we are indeed seeing the same satellite signal in those channel intervals.

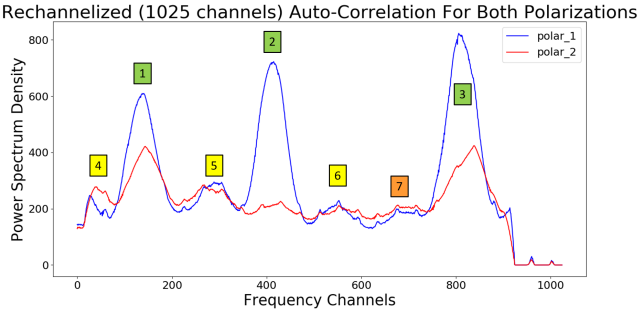


Fig. 5. This shows the time-averaged auto-correlation of the re-channelized time stream recovered by performing an inverse PFB on the data used to make Fig 4. We can distinguish 3 well-resolved satellites (green), 3 smaller signals (yellow) as well as an uncertain bump that could correspond to another satellite (orange).

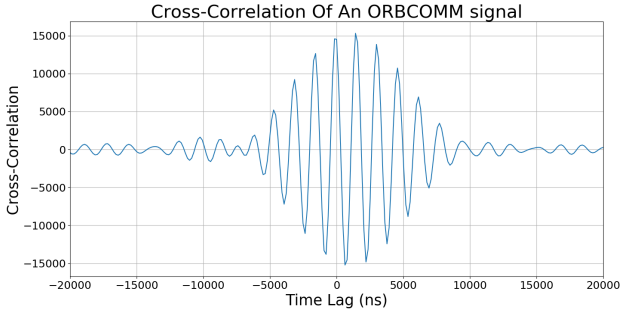


Fig. 6. Shows a zoomed-in view centered around zero of the cross-correlation of satellite #2 from Fig. 5. The peak of the cross-cor corresponds to the time-lag between the 2 signals used. To obtain this particular plot, the Fourier transforms used in Eq. 4 are zero-padded with an amount equal to the original FT length (100% more points). The fact that it is sufficiently centered around zero when it peaks leads to believe that there was not much difference between both observed polarizations of the satellite, as expected.

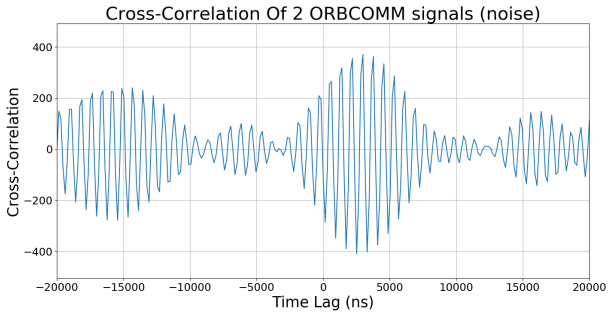


Fig. 7. An overestimate of the noise in an ORBCOMM signal. To obtain this, satellite #2's polarization #1 was cross-correlated with satellite #3's polarization #2. Again, Fig. 5 is used to number and distinguish these different signals. The noise obtained is said to be an overestimate as in actuality, the signal power is expected to correlate, meaning only the system noise would contribute to the total noise.

satellite #2 was observed for 5 seconds and interpolated with 100% more points, a timing error of 1.01 ns is obtained.

Depending on the zero-padding/interpolation performed before correlation, various timing errors are obtained as a function of observation time and displayed in Fig. 8. As before all these results are obtained by the cross-correlation of satellite #2 from Fig. 5.

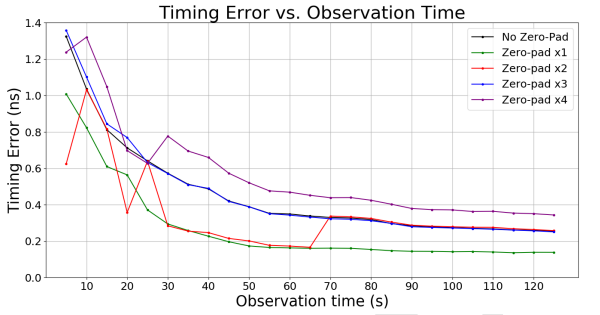
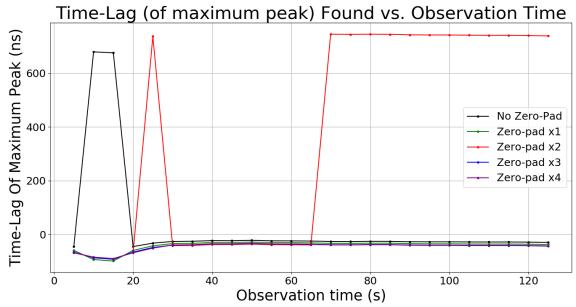


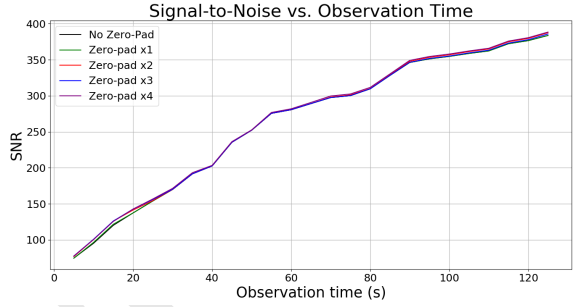
Fig. 8. A plot of timing error as a function of observation/integration time when cross-correlating the polarizations of satellite #2 of Fig. 5. The labeling indicates by how much the FTs were zero-padded before cross-correlation. For example, 'Zero-Pad x1' refers to the FT's size being increased by 100% by introducing zeros into it. In this case, it results in double the points being used to plot the cross-cor.

As observation time increases, it strictly decreases the timing error in the phase of our signals, given that no zero-padding is utilized. As zero-padding is incorporated in our cross-correlation calculations, there is no guarantee this behaviour will be respected by the Monte Carlo process. From Fig. 8, we observe that interpolating with 100% more points in the cross-correlation outputs the best timing errors. This leads to believe that less is more when it comes to zero-padding and should not be used profusely.

Additionally, Fig. 9a shows it is possible for the Monte Carlo least-squares fitting to change the peak it fits around depending on observation time. When looking at irregular jumps in timing errors of Fig. 8, some of these seem to coincide with a switch in peaks detected by the Monte Carlo process in Fig 9a. This switch in peaks is due to the way the cross-correlation is performed as a function of observation time. It is done by taking equal-length chunks of the satellite signals,



(a)



(b)

Fig. 9. Plots of the time-lag values found for the peaks in the cross-correlation as a function of observation time, as well as SNR. We observe that there seems to be no guarantee that the Monte Carlo process sticks to evaluating a unique cross-cor peak given different zero-pad and observation time values. However, we can seem to get a guarantee that the SNR in the cross-correlations will stay within the same ballpark, regardless of zero-padding.

cross-correlating those chunks and summing the chunks as we observe for longer periods of time. Since the summations can introduce slight differences in the shape of the cross-cor as time passes, it explains why sometimes the cross-cor peak changes location according to the Monte Carlo algorithm.

Fig 9b shows SNR as a function of observation time is similar for any zero-padding approach.

We see from Fig. 8 that timing errors of less than 1 ns are possible with a minimum of 10 seconds of integration, with a lower asymptote seemingly being reached at about 0.1 ns in the best case shown here. We can speak on the usefulness of this timing error by stating the limitations it imposes on the positional errors of celestial sources.

Since ALBRATROS hopes to observe at around 10 MHz, a timing error in the signal phase of 0.1 ns corresponds to $(10MHz) * (2\pi) * (0.1ns) \approx 0.006\ radians$. At the Marion Island location, a baseline of approximately 20 km allows for a resolution of $\sim \frac{3*10^8(m/s)/10^7(Hz)}{20000(m)} = 0.0015\ radians$. Therefore, this imposes a positional error in our celestial sources observed from Marion of $0.006 * 0.0015 = 9 * 10^{-6}\ radians$, or roughly 2 arcseconds.

Another way of interpreting the timing errors obtained is by using them to estimate the expected SNR in the phase. With 0.006 *radians* of error, we'll be able to resolve a radian with $SNR = \frac{1}{0.006} \approx 167$, which caps the SNR if the noise in the system is smaller than the noise in the phase.

Conclusions / Further Work

From a limited observation time of 2 minutes, a timing error of $\sim$ 0.1 ns can be achieved using ORBCOMM satellites for synchronization of an ALBATROS antenna array. It is important to keep in mind that for this research, much of what has been measured pertaining to noise is likely an overestimate. Since the observations were held in downtown Montreal in a metal-rich lab, things like signal time-lag and noise are likely to have been exaggerated compared to what will be seen at MARS and Marion.

Nevertheless, with what has been done so far, the ground work has been laid to properly estimate the timing accuracy offered by an ORBCOMM signal. Due to this research being performed during the COVID-19 pandemic, some of its aspects were not able to be fully explored before submission. Here the author will list what can be done by future researchers to better constrain the performance of ORBCOMM for usage by ALBRATROS.

- Perform the ORBCOMM observations with a better antenna setup. Not only was the antenna design crude (simple wires placed in an orthogonal fashion), but the lab in which it was placed to observe in was not conducive to polarizations from a same signal having zero time-lag. A better way to take data in future timing error measurements would be to observe with a better antenna (ideally ALBATROS) and be in an open-air location with less RFI than downtown Montreal. This will undoubtedly increase SNR in satellite data as well as allow for more satellites to be observed at once.

- Cross-correlate over longer observation periods than the 2 minutes explored in this report. Roughly 2 days of data was kept in the lab but has remained untouched as of writing this paper. The code-base used in the writing of this report can serve as a good starting point for someone picking the project back up in the near future, and can be found in this github repo.

- Investigate the appearance of unnatural power spikes observed in the instance where Fourier Space data was inverted with a pfb-inverse (10) and then re-channelized. Since these unnatural power spikes appear even from taking the inverse of random complex Gaussian noise, it was decided that simply removing them from auto-correlation plots by using a smoothing algorithm was acceptable. These power spikes did not seem to have an effect on cross-correlation data, but nevertheless, a better understanding of this artificial construct would not be a bad idea.

ACKNOWLEDGMENTS. Many thanks to Dr. Jonathan Sievers who's patience with this undergraduate student was appreciated. Thank you for helping me understand key concepts in radio astronomy and signal processing and nudging me in the right direction every now and then when I strayed from the path. I would also like to thank Nivek Ghazi for helping to clarify certain concepts regarding the interpretation and analysis of data coming from the SNAP FPGA back-end. Additionally, thanks to Dr. Sabrina Leslie for sharing valuable knowledge and resources pertaining to the writing and presentation of this report.

References

1. JR Pritchard, A Loeb, 21-cm cosmology in the 21st century (2012).
2. SR Furlanetto, et. al., Fundamental cosmology in the dark ages with 21-cm line fluctuations (Astro 2020 Science White Paper) (2019).
3. J Pritchard, A Loeb, Hydrogen was not ionized abruptly. *Nature* **486**, 772 (2010).
4. G Reber, Cosmic static at 144 meters wavelength. *Frankl. Inst.* **285**, p 1–10 (1968).
5. JK Alexander, ML Kaiser, JC Novaco, FR Grena, RR Weber, Scientific instrumentation of the radio-astronomy-explorer-2 satellite. *Astron. Astrophys.* **40**, 365–371 (1975).
6. HC Chiang, Observing the <100 mhz radio sky from the sub-antartic and arctic (Global 21cm Workshop Slides) (2019).
7. BC Hicks, et al., A wide-band, active antenna system for long wavelength radio astronomy (arXiv:1210.0506 [astro-ph.IM]) (2012).
8. DE Gary, Radio astronomy lecture 8: Digital cross-correlators (Physics 728 @ NJIT) (2019).
9. J Chennamangalam, The polyphase filter bank technique (https://casper.ssl.berkeley.edu/wiki/The_Polyphase_Filter_Bank_Technique) (2016).
10. R Shaw, Pfb inverse code repository (https://github.com/jrs65/pfb-inverse) (2015).
11. M Krasteva, Improved back end electronics enclosure for the albatros interferometer on marion island (Concordia BSc Honours Physics Thesis) (2019).