An Analysis of Local Universe Fast Radio Bursts

Xander Jenkin

Advised by Dr. Dongzi Li

A Junior Paper

Submitted to the Department of Astrophysical Sciences

in Partial Fulfillment of the Requirements for

the Degree of Bachelor of Arts

Princeton University

Spring 2024

Abstract

Fast Radio Bursts (FRBs) are bursts of radio emission on the millisecond scale whose origins are currently unknown. Out of the possible theories, magnetars are often associated with FRB sources, but it is still unknown what overall fraction of FRBs are ultimately generated by magnetars. FRBs are similar to pulses from pulsars, but do not repeat so consistently—about 4% of FRB sources are known to repeat, but the fast majority have not been found to repeat at all. These discrepancies between FRB sources make FRBs a largely undiscovered frontier, with lots of progress made in recent years with more continuing to be made moving forward. In a previous paper, we successfully detected two FRBs from the same source using data from the Five-hundred-meter Aperture Space Telescope (FAST). Now, we continue our research by expanding our search to attempt to detect additional FRBs by broadening the search window while also performing deeper analysis of the two bursts we previously discovered. In this paper, we were unable to find additional bursts by broadening our detection timing window. We did successfully extract our bursts into Python, remove additional background noise and man-made Radio Frequency Interference (RFI), and analyze our FRBs' Stokes parameters to find the polarization and Rotation Measure (RM) of the bursts. We also estimated the burst rate ($R \approx (0.55 \pm 0.30)$ bursts hr^{-1}) for the previous target which we have detections for. These results allow us to better understand the FRBs we detected as well as the source they came from, as the RM nearly doubled between the two detections (-268.123 $\pm$ 2.534 rad m^{-2} for the first and -531.638 $\pm$ 1.186 rad m^{-2}) that were $\sim$ 3 months apart, indicating that the target is in a highly magneto-active environment. In the future, we could continue to search for bursts in additional sources with FAST then use our established process to analyze them as well.

This paper represents my work in accordance with University regulations.

/s/ Xander Jenkin

Contents

1 Introduction

1.1 What are Fast Radio Bursts (FRBs)?

Fast Radio Bursts (FRBs) are short, broadband radio emission on the millisecond scale which originate from unknown sources that are usually extragalactic (though recent discoveries have been made within our galaxy). FRBs currently have largely unknown origins, and while many FRB sources are known to repeat, the majority do not, making the mechanics behind how FRBs occur remain a mystery at the current time. Pulsar pulses are very similar to FRBs, though repeating FRB sources do not repeat nearly as consistently as pulsars, and FRBs have flux densities significantly larger ($\sim 10^{10}$ $\times$) on a much smaller millisecond window with vastly different spectra [Cordes & Chatterjee(2019)].

From 2019 to 2022, the number of confirmed unique FRB sources went from around 60 to over 600, with that number consistently increasing each year as more FRBs are detected [Petroff et al.(2022)Petroff, Hessels, & Lorimer]. Only about $\sim 4\%$ of these sources are known repeaters, but all of our current non-repeating FRB sources are only labelled such because we have not detected another FRB after an initial detection. In theory, all non-repeating FRB sources could be potentially repeating sources that are unconfirmed, further motivating research into the unknown origins of FRBs and the full nature of FRB sources.

1.2 FRB Origins

Many FRBs have been localized to areas associating them with magnetars (young, highly magnetized neutron stars that are X-ray active [Kaspi & Beloborodov(2017)]), but overall we still do not know exactly where they come from, and previous assumptions such as FRBs being only extragalactic have been recently disproven. There is still plenty of room for our ideas of their origins to change as we learn more and analyze more bursts from

repeating and non-repeating FRB sources.

2 Data Selection

2.1 FAST and CHIME Targets

The data we used was collected from the Five-hundred-meter Aperture Spherical Telescope (FAST) in Guizhou, China, which is the largest single-dish telescope ever made. FAST has unprecedented sensitivity which can resolve faint radio bursts, so it is a great instrument to use to check known FRB sources for new bursts (both to discover more bursts from repeating sources and to possibly confirm a non-repeating source as a repeater).

The data collected from FAST was guided by targets that were FRB sources detected from the Canadian Hydrogen Intensity Mapping Experiment (CHIME) radio telescope. This was a broader survey with much less sensitive instruments, so the goal of using FAST was to take known FRB Targets from CHIME and examine them in hopes of finding more FRBs that could not be previously detected.

Our FAST survey spanned across 12 targets, where targets 1-7 are known non-repeating FRB sources and targets 8-12 are know repeating FRB sources. More information and target parameters can be found in Appendix A.4.

3 Previous Burst Search

We will now be briefly summarizing the results of our previous paper (the Junior Paper submitted for Fall 2023)—where we successfully identified two new FRBs from one of our 12 FRB targets—before moving on to what was done during this new portion of the project.

3.1 Our Choice of DM and De-Dispersion

The Dispersion Measure (DM) is the column density (in pc cm^{-3}) of dispersion due to the interstellar medium integrated over all electrons along the distance d from the observer to the source [Cordes & Chatterjee(2019)].

$$\mathrm{DM} \equiv \int_0^d n_e dl \tag{1}$$

The time delay of FRBs as they arrive varies inversely with frequency, resulting in the signal becoming 'dispersed' across frequency over time. DM is the leading term in the time delay, followed by the Faraday Rotation Measure (RM), and the Emission Measure (EM), which each scale to an additional power of ν^{-n}, making each subsequent term less significant in affecting the time delay.

$$t(\nu) = 4.15\mathrm{ms}\left(\frac{\mathrm{DM}}{\nu^2}\right) \pm 28.6\mathrm{ps}\left(\frac{\mathrm{RM}}{\nu^3}\right) + 0.251\mathrm{ps}\left(\frac{\mathrm{EM}}{\nu^4}\right) \tag{2}$$

where

$$\mathrm{RM} \equiv \int n_e B_{||} ds \quad , \quad \mathrm{EM} \equiv \int n_e^2 ds \tag{3}$$

The RM has units of rad m^{-2} and the EM has units of pc cm^{-6}.

This means that incoming FRBs must be 'de-dispersed' to account for this time delay (where each frequency is subsequently shifted according to the FRB source's DM). Previously, we only accounted for the DM term in the time delay, which is good enough when initially searching for bursts. Later, we will discuss the RM in the next term when we more closely analyze our previously discovered bursts.

3.2 Radio Frequency Interference (RFI)

Radio Frequency Interference is often difficult to fully detect and even more difficult to remove (especially before we know where the bursts are and what they look like). The challenge of detecting new FRBs is largely distinguishing between real bursts and man-made RFI. We made various attempts to remove RFI in software using PRESTO (PulsaR Exploration and Search TOolkit) which has functions for FRBs in addition to Pulsar detection tools [Amiri et al.(2020)Amiri, Andersen, Bandura, & The CHIME/FRB Collaboration*].

3.3 Searching through the Bursts

Still, despite our efforts in software, the majority of single pulse detections (which are ideally bursts), will just be RFI. Our general strategy was to look through the most prominent bursts and look for prominent ($> 10\sigma$) Gaussian-shaped peaks with a vertical set of marks that follow the structure of the burst across frequency and a smooth diagonal 'ribbon' across DM. More details can be found in Appendix A.1; we are keeping these sections on previous steps brief, as they are much more detailed in the previous paper, and we would like to quickly move on to what was done for the new portion of this project.

3.4 Previous Results

Previously, we scanned 5 out of the 12 targets and identified two FRBs from the known repeater Target 11. Both exhibit a prominent Gaussian-like shape, with clear vertical marks along the burst in frequency and a bright 'ribbon' across DM just as we expect.

Burst #	Target	Date	Time (UT)	Signal/Noise (σ)	Peak Flux	Fluency
1	11	2021-10-20	03:12:10.824	22.38	0.034 Jy	0.27 Jy ms
2	11	2022-01-08	22:28:02.530	31.33	0.047 Jy	0.38 Jy ms

Table 1: List of found bursts from Target 11 with locations, times, and parameters. Peak Flux ($\sim 0.0015\sigma$ Jy) and Fluency ($\sim 0.012\sigma$ Jy ms) are statistical estimates given the average results of FAST (not directly measured) which can have up to an order of magnitude of undertainty).

We ended this project successfully having found FRBs, and our next steps were to continue searching for new bursts in the targets, constrain the rate of Target 11 given the two bursts we found, and to more closely analyze the structure of the bursts to examine their polarizations and RMs.

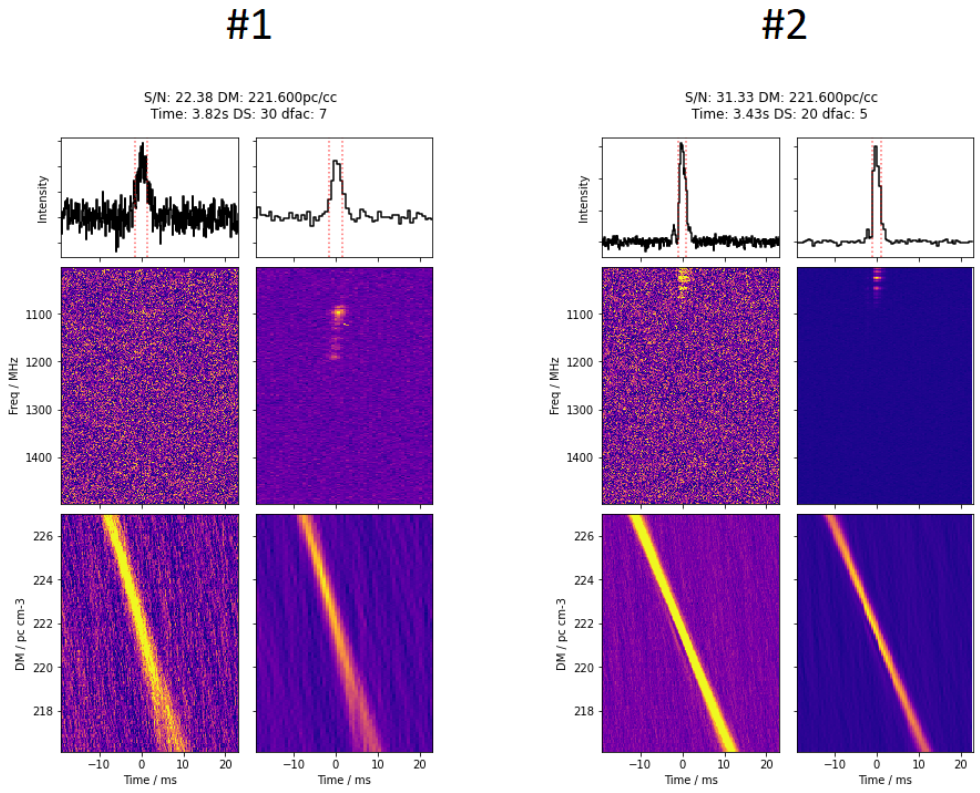

Figure 1: First and second bursts found from Target 11

Target Scanned	DM	Repeating?	Hours Scanned	Bursts Found
2	111.6	Not yet repeating	0.832	0
3	182.7	Not yet repeating	0.788	0
4	127.7	Not yet repeating	0.299	0
11	221.6	Repeater	5.503	2
12	287.1	Repeater	0.834	0

Table 2: Results for all the targets scanned for bursts.

4 Broadening the Search

Now, we will begin with what was newly done for this paper, beginning with our first step after the end of our previous paper, which is to broaden our search using a computationally inexpensive method to quickly check for more bursts.

4.1 Detection Timing Window

Previously, when we searched for single pulses in PRESTO we had only searched for peaks within a sub-millisecond window. We only found two FRBs across the targets, and so a quick way to check for more is to increase the maximum peak window size. This timing is not the full width at half maximum (FWHM) of the burst, which for FRBs is usually on the scale of a few milliseconds. If that were the case, our sub-millisecond window would have been far too short and neither of our FRBs would have been found. Instead, this is only the window of the very top of the burst and how long it maintains values around its peak.

Increasing this window will also likely increase the amount of RFI that gets marked as single pulses, so it's not the best thing to do for an initial FRB search; however, since we previously identified bursts already, this can be a method to try and find bursts that we may have otherwise missed.

After completing all of the previous steps in PRESTO, this can be done by using the `-m` or `--maxwidth` flag when running `single_pulse_search.py` and setting the flag to whatever

value you want (in milliseconds) instead of the default sub-millisecond window. Because this is only changing the maximum window, you will still see all of the previous detections at lower windows—thus, any bursts we have already found will still be there. We are still primarily looking for very significant single pulses with a $> 10\sigma$ Signal/Noise ratio, but when increasing the window size like this, we are only interested in any new single pulse detections that are comparable to previously found bursts and are much higher and distinct from previously detected single pulses that were discarded as RFI.

Searching the data for single pulses in PRESTO with variable window:

```
single_pulse_search.py -b Lband_rfifind.mask_DMXXX.XX.dat -m 10
```

where `10` should be replaced with your window (ms)[1].

This makes the process of manually looking through the detections easier, as the vast majority of new detections (which there will be many of) from the larger window won't be very significant and can be discarded as RFI.

4.2 Larger Window Results

We increased the maximum window flag `-m` 1ms, 10ms, 100ms, and 1000ms and re-ran `single_pulse_search.py` for our Target 11 surveys to see if we as well as the other targets. Realistically, we might not need to test it at 100ms and beyond, as anything larger than that likely won't resemble a single pulse or burst anyways. However, we did this anyway since it is quick and we wanted to test a wide range of windows to see if there are any significant new detections (full table of results in Appendix A.5).

After doing this, we did see an increase in total overall single pulse detections with moderate significance ($> 5\sigma$) for all sources scanned (as expected) when raising the

[1]Fortunately, this single pulse search for peaks occurs after the de-dispersion and RFI masking in PRESTO, so we can manually change the window timing for the last step without re-running all of the previous code. For Target 11 alone, those prior steps take over 30 hours to run, so this saves us a significant amount of time and makes testing different timing windows fast and easy.

maximum window to 1ms, but we did not see this change when raising it further to 10ms, 100ms, or 1000ms. This confirms the idea that raising the window will increase overall detections but only to the point where they still resemble pulses detectable by the software.

When filtering these by Signal/Noise ratios $> 10\sigma$, we found that there are no new single pulse detections after raising the maximum window in the majority of the sources. This again confirms our suspicion that most of the new detections from the larger window won't be significant enough to possibly be FRBs. Only two new single pulse detections $> 10\sigma$ were found after increasing the maximum window to 1ms: one from our Target 2 survey (17.53σ) and one from our first Target 11 survey (13.85σ). Both of these could be potential candidates, and they could possibly be evidence of Target 11 having a third burst during this observing duration and Target 2 (thought to be a non-repeating FRB source) repeating. Our next step is to visually inspect them to see if they exhibit the properties of FRBs (with the same process and criteria as done for the two Target 11 bursts we discovered previously).

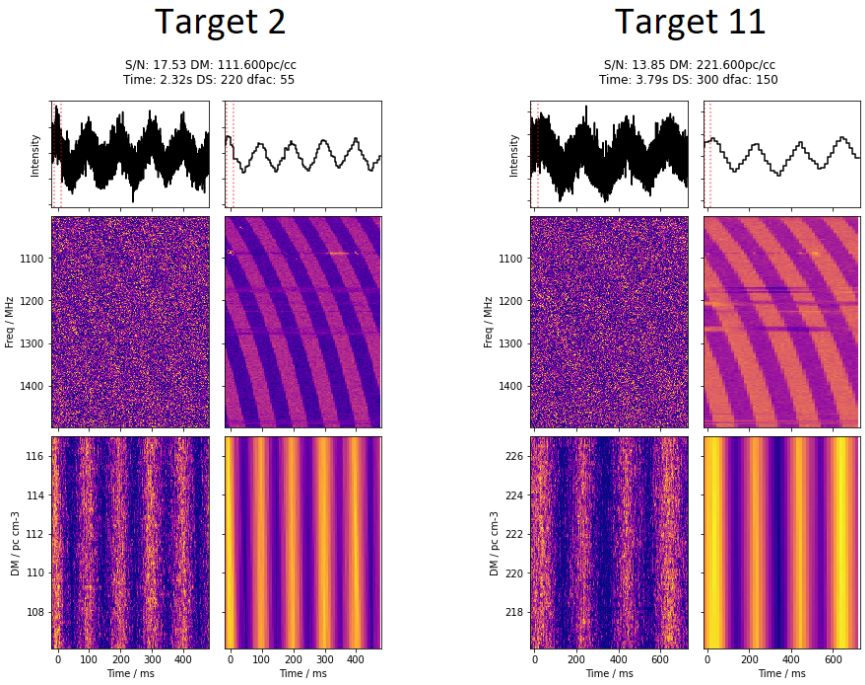

Figure 2: New Single Pulse detections found in Target 2 and Target 11

Plotting both of these detections (Fig. 2) just as we did before, neither appear to be single Gaussian-shaped pulses with vertical marks across frequency when de-dispersed and a single 'ribbon' across DM. Instead, they appear to be sinusoidal signals that exhibit odd, repeating bands that vary across time even after de-dispersion. These are clearly not FRBs and likely some sort of RFI, and were only detected because they have broader peaks that were picked up when we increased the width of the detection window.

Regardless, this was not a long process compared to all of the combined steps when we previously searched for bursts, and it could have potentially yielded additional FRBs, so it was a valuable step to take that made sense when trying to expand this project beyond the results of my previous paper.

5 Rate Estimation

5.1 Poisson Statistics

The Poisson distribution is a probability distribution that can be used to model the probability of events independent of the time elapsed since previous events. The probability of k events occurring assuming Poisson statistics is:

$$P(\text{of } k \text{ events}) \equiv \frac{\lambda^k e^{-\lambda}}{k!} \tag{4}$$

where λ is the expected number of events [Yates(2005)].

If you want to express the probability of k events occurring in a time interval t assuming a rate of r events per t, you can rewrite this as:

$$\lambda \equiv rt, \qquad P(k \text{ events in time } t) = \frac{(rt)^k e^{-rt}}{k!} \tag{5}$$

5.2 Why Use Poisson Statistics?

Poisson statistics assume independence between events, and FRB occurrences are generally thought to be not independent. Similar to their origins, the actual probability distribution representative of FRBs is unknown, but we know that they likely do not exactly follow Poisson statistics [Lawrence et al.(2017)Lawrence, Vander Wiel, Law, Burke Spolaor, & Bower]. This method of using Poisson statistics for FRB detections does NOT predict when future events will occur and how many we expect to occur. Instead, this process is used to compare the rates between different observations to guide the efficiency of the instruments for calibration.

It is informative for those doing instrumentation to analyze the rates of various observations conducted using their instrument, and it is also informative for those looking for bursts to compare their rates with others using the same instrument. Thus, we will now compute the rate based on our previous results and compare it to results from others searching for FRBs with FAST.

5.3 General Poisson Rate Calculation

Based on Eq. 5, we can very easily determine the probability of seeing k events if we know the rate r and time t. However, reversing this and finding the rate and time for a given k and probability is very difficult. Analytically, we cannot solve for a single solution. There are, in fact, multiple solutions for $\lambda = rt$, as there exist multiple possible Poisson distributions which can have the same probability for a particular value of k.

This makes things challenging, so our practical strategy for calculating this rate is to compute many different Poisson distributions and find which R result in desired probabilities for a given k.

5.4 How to Calculate Rates for our FRBs

If we want to know the rate R of FRBs coming from a repeating FRB source, we can use our total observing time T, number of FRBs detected during that time k, and the target probability to have seen k events. In this case, it isn't useful for us to use the probability of observing *exactly* k events; instead, we want to know the probabilities of seeing at most or at least k events, which involves summing the area under the Poisson distribution before or after k.

To estimate upper and lower bounds for our rate R, we should have upper and lower bounds for how likely it is to see at least or at most k events. If these bounds are symmetric (i.e. upper bound + lower bound = 1) we can say that the probability to see at least the upper bound is the same as the probability to see at most the lower bound (and vice versa), so the choices of if we assume at most or at least k events (whether we integrate from the left or the right) are equivalent.

As previously mentioned, to estimate R for a given k, we have to test many different possible Poisson distributions. Now that we are trying to integrate to k and sum to target upper and lower bound probabilities, we need to generate and integrate across many, many Poisson distributions proportional to how precise we want our R estimates—which will have upper and lower estimate values. We choose to simulate this in Python by generating many distributions and checking for which R values have total probabilities integrated around k which are closest to our bounds.

5.5 Numerical Estimation

Since we need to find the area under the Poisson distribution about k, we need to be able to continuously integrate across it. The Poisson distribution is an inherently *discrete* probability distribution that is specifically defined for an integer number of events k. This makes sense in terms of not being able to see 2.5 FRB detections, but is not convenient

when trying to integrate accurately between k events.

This discreteness restriction lies in the factorial of Eq. 5, but fortunately we can use the 'gamma function' $\Gamma(k)$,

$$\Gamma(n) = (n+1)!, \quad n \in \mathbb{Z} > 0, \qquad\qquad \Gamma(x) \equiv \int_0^\infty t^{x-1} e^{-t} dt, \quad x \in \mathbb{R} \tag{6}$$

where we can equate it to the factorial while also defining it for non-integer values of k,

$$P(k \text{ events in time } t) = \frac{(rt)^k e^{-rt}}{\Gamma(k-1)} \tag{7}$$

making Eq. 5 into a continuous function which we can now integrate over numerically.

In Python, we can use `scipy.special.factorial`, which uses the gamma function for non-integer values but not values < 0, which is perfect for estimating a continuous Poisson distribution. This does partially contribute to the erratic behavior of our Poisson estimation near 0 later on, but for the rest of the integration, this works incredibly well to create a smooth-appearing Poisson distribution estimate.

5.5.1 Trapezoidal Integration

Even though we now have a continuous probability distribution to integrate over, because we cannot analytically solve for R given k, the most practical way to do this for our purposes is to numerically integrate to our desired precision for R. The fact that the upper and lower R we find is not critical for any of the other parts of our results but is instead a valuable metric to compute alongside them further motivates this choice.

The continuous integral over $P(x)$ can be trapezoidally approximated as

$$\int P(k)dk \approx \sum_i^N \frac{P(k_{i-1}) + P(k_i)}{2} \Delta k_i \equiv \text{trapz}[P(k)] \tag{8}$$

where k_i is the set of k values we integrate over and Δk_i is proportional to the width of how finely we integrate.

We can take this function and integrate over it many times testing different values of R for a given k in Python by looping over `np.trapz()` and outputting the R values whose distributions integrated closest to the inputted lower and upper (1− lower) percentages (full final Python script in Appendix A.2).

5.5.2 Problems

The continuous Poisson distribution (even with a continuous factorial) does not behave well for negative values and no longer predicts probability past 0. This is good because it reduces our integration to

$$\int_{-\infty}^{\infty} P(k)dk \quad => \quad \int_{0}^{\infty} P(k)dk = 1 \tag{9}$$

which seems great for numerical computer estimation, since we can just integrate from 0 to however many N points of precision we need to where the probabilities become small and the total amount integrated becomes close enough to 1 (the Poisson distribution approaches 0 quite quickly, so in practice we don't have to sample out too far to get $> 99\%$).

However, a problem emerges in precision when we have estimate values of k between [0, 1], as for many Poisson distributions (including the ones that closely match my results), the slope near 0 becomes incredibly steep—probabilities near 0 do not always reach 0 smoothly, so even numeric trapezoidal integration becomes inaccurate.

When my Python script to numerically estimate the rate was initially behaving oddly, I tested it by trapezoidally integrating my resulting Poisson distributions out to large N (where $P(N) < 10^{-10}$), and the entire distribution only summed to $< 83\%$. This is far lower than our goal of being almost ≈ 1, and this resulted in completely inaccurate rate

estimations with Poisson distributions that did not give the true targeted probabilities.

5.5.3 Solutions

We remedied this problem by taking advantage of the simple integration fact that

$$\int_0^k P(k)dk = 1 - \int_k^\infty P(k)dk \tag{10}$$

where we can instead just numerically compute the integral from k out to a large number N such that $P(N)$ becomes very small (which our trapezoidal integration is very good at).

Doing this yielded much better results when attempting to re-compute the rate results of other FRB papers.

5.6 Target 11 Rate Results

My Python script (Appendix A.2) when given our parameters for Target 11 with $k = 2$ bursts gave a final rate estimate of $R \approx (0.55 \pm 0.30)$ bursts hr^{-1}, with full results below.

Target	# FRBs (k)	Obs. Time	Center Probability	Upper R	Lower R
11	2	5.503 hr	68%	0.8405 hr^{-1}	0.2499 hr^{-1}

Table 3: Full Results for rate R estimation for the two bursts found in Target 11

where the Center Probability is the 'goal' probability that is in the middle of the lower bounded:

$$\text{Center Probability} \equiv 1 - 2(\text{lower percentage}) \tag{11}$$

In my case, we wanted my rate to be within 1σ on either side of the distribution. We took the percentage of $\sim 68\%$ corresponding to the middle area within 1σ and computed distributions corresponding to this area by looking for areas on either side of the middle $\sim 68\%$: 16% (lower percentage) and 84% ($1 - 16\%$, higher percentage). Those lower and higher percentages are what is inputted into my script to produce the upper and lower R

estimates.

When comparing with results from a completely different FAST survey found by [Sheikh et al.(2024)Sheikh, Farah, Pollak, Siemion, Chamma, Cruz, Davis, DeBoer, Gajjar, Karn, Kittling, Lu, Masters, Premnath, Schoultz, Shumaker, Singh, & Snodgrass]:

{ 35 bursts, 541 hours, avg. burst rate [0.0776, 0.0538] },

my script outputted an upper and lower range of [0.0776, 0.0555], which is not quite exact but is close enough to indicate that our numbers are somewhat representative and comparable with other FAST surveys.

6 Extraction of Bursts

Previously, we took the intensity at the location of the burst produced by PRESTO and plotted it directly from the `.fits` file to get an immediate idea of if we have found an FRB. To more closely examine its structure, we should now extract the portion of data containing the burst and turn it into a format we can fully examine and plot (here I used `matplotlib` in Python).

The process of extracting the burst involves knowing the DM of the target and the MJD (time, in seconds) of the burst. Once we have both of these, we de-disperse that portion of the `.fits` file into a `numpy` array. Most of the logistics of this come from having to de-disperse the data again and properly transfer over the frequency and time data properly aligned with the observed values, but this can be done using built-in functions with the PSRCHIVE pulsar analysis software (details of PSRCHIVE and burst extraction scripts in Appendix A.3) [Straten et al.(2010)Straten, Manchester, Johnston, & Reynolds].

Once we have 'cut out' the FRB from its containing `.fits` file and have it in loaded into Python (I loaded an `.npz` file, but anything that can store nested `NumPy` arrays of data will

work), we need to know how this FAST radio data is structured so that we can examine the burst itself.

6.1 Stokes Parameters

The electromagnetic radiation detected by FAST that contains our FRB arrives with polarization, which is stored by the detector and encoded into the data. The intensity of the different components of the incoming radiation is expressed as the Stokes vector $\vec{S}$.

6.1.1 I, Q, U, V

The components of $\vec{S}$ are named the 'Stokes parameters', and all four together describe the polarization of the incoming radiation:

$$\vec{S} = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} \equiv \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} \tag{12}$$

which are referred to as I, Q, U, and V, and we will need to obtain these components from our data to determine and analyze our FRBs' polarization.

6.1.2 XX, YY, $(YX)_{real}$, $(YX)_{imaginary}$

Unfortunately, our FAST data is not split up into the Stokes parameters I, Q, U, and V, which is ultimately what we want so that we can analyze the FRBs' polarization. This data is instead stored with the parameters XX, YY, $(YX)_{real}$, and $(YX)_{imaginary}$. These four parameters can easily be transformed into I, Q, U, and V with the following conversions [Straten et al.(2010)Straten, Manchester, Johnston, & Reynolds]:

$$I = XX + YY \tag{13}$$

$$Q = XX - YY \tag{14}$$

$$U = 2 \cdot (YX)_{real} \tag{15}$$

$$V = 2 \cdot (YX)_{imaginary} \tag{16}$$

Later on, we will need to be careful and do some processing on the data (Subtraction & Division) with these initial parameters first and only convert to our final I, Q, U, and V parameters afterwards—though this will be done before we re-mask for RFI and 'bin' the time and frequency resolutions into larger chunks when plotting.

6.1.3 Why Care About Polarization?

The rationale behind computing and analyzing the polarization of our FRBs primarily comes from the motivation to find the Rotation Measure (RM) of the burst which adds the next leading order term of the arrival time delay of our FRBs (Eq. 2). Previously, we neglected these two terms for the purposes of large-scale searching. Now, (much like we will do when we later more accurately remove the RFI and background noise) we can compute this RM and use it to inform future searches and learn more about the FRB source and how the bursts that come from it operate.

6.2 Fully Identifying the Burst

6.2.1 Simple Self-Calibration

To remove a large amount of background noise from around our burst, we can perform a simple self-calibration process that involves subtracting and dividing the data by the 'off burst' off(f), which is the average value of the intensity across time per frequency outside of the FWHM of the burst.

$$\text{off}_{XX}(f) \equiv \text{mean}[XX(t)](f) \tag{17}$$

and our new calibrated parameters become

$$XX(t,f) \equiv \frac{XX(t,f) - \text{off}_{XX}(f)}{\text{off}_{XX}(f)}, \quad YY(t,f) \equiv \frac{YY(t,f) - \text{off}_{YY}(f)}{\text{off}_{YY}(f)} \tag{18}$$

$$(YX)_r(t,f) \equiv \frac{(YX)_r(t,f) - \text{off}_{(YX)_r}(f)}{\sqrt{\text{off}_{XX}(f) \cdot \text{off}_{YY}(f)}}, \quad (YX)_i(t,f) \equiv \frac{(YX)_i(t,f) - \text{off}_{(YX)_i}(f)}{\sqrt{\text{off}_{XX}(f) \cdot \text{off}_{YY}(f)}} \tag{19}$$

where the subtraction and division is done differently for each of the four parameters, as the polarization is expressed differently between I, Q, U, and V (which we will later transform these parameters into).

It is also done on these parameters *before* converting into I, Q, U, and V because we want to calculate the relative amplitudes of XX and YY separately before combining them into I and Q—the noise levels may not be uniform between the two, so we want to remove them such that the differences fully contribute to I and Q. We also do this before re-masking RFI because this actually can remove low-level, noisy RFI that is uniform across frequencies, allowing us to mask less RFI in the next step, preserving more of the burst.

This method is a simple way to greatly reduce the intensity of the background by assuming that the background is largely unpolarized (whereas FRBs are highly polarized). This does not remove everything but maintains the burst while removing a vast majority of the background.

6.2.2 Masking RFI (more accurately)

Previously, we attempted to mask out the RFI before searching for bursts in PRESTO to varying degrees of success—we were able to find our two FRBs, but amongst many false positives that made the search and identification significantly time consuming.

Now that we have identified and confirmed our FRBs, masking out the RFI present during the burst becomes significantly easier (and we can do so much more precisely). We will

first take the intensity over time of our burst and fit a Gaussian to it using `scipy.optimize`. This allows us to estimate the FWHM of the burst (by computing $2\sqrt{2\ln 2} \cdot$ (standard deviation) of the fitted Gaussian), which we will use to only look for RFI outside of the burst's FWHM as shown below[2]. This allows us to easily discard any extrema without worry of masking out our FRBs.

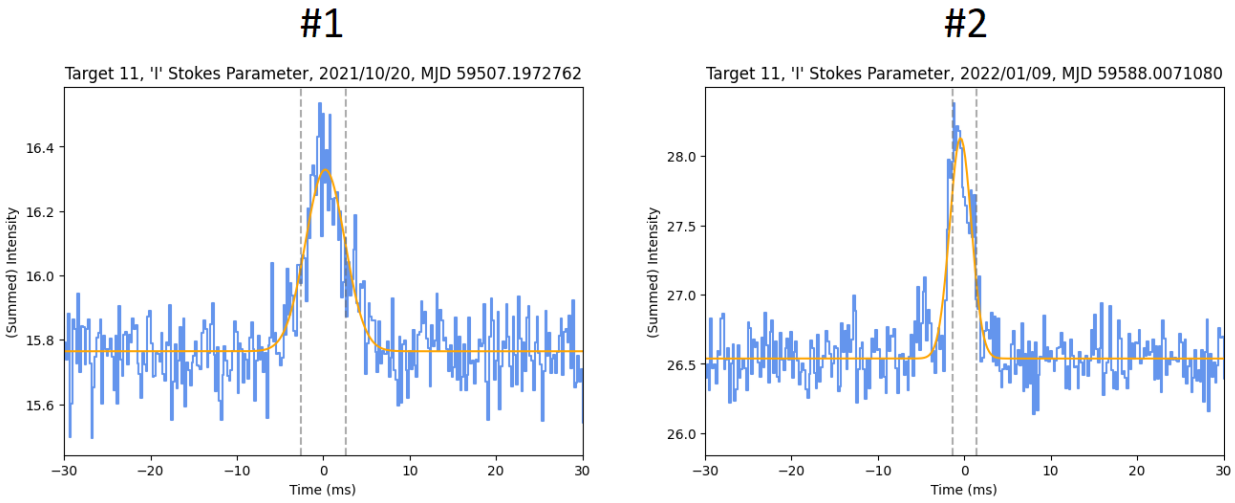

Figure 3: Visual Examples of Gaussian fits with FWHM estimates on bursts from Target 11

The RFI is now detected by taking the intensity across time outside of the FRB's FWHM for each frequency and checking to see if its variance is greater than the median (to be more resistant to extrema) variance across all frequencies times a threshold (which you can tune so that it removes RFI but doesn't block out large parts of the burst, I used values between 10 and 15). If that frequency's variance outside of the FWHM is greater, the entire frequency is masked out as RFI.

The ability to search outside of the burst's FWHM and the manual tuning of the masking threshold on a per-burst basis makes this post-discovery RFI masking more accurate and lets us get a clearer view of the burst.

[2]Note that these are NOT the same as what will be later shown in the results of Section 6.3, as those will be calibrated, RFI masked, and binned within the time window. This is just shown for a visual idea of what will happen later, with the grey lines continuing to be an indicator of the result.

6.2.3 Intensity 'Waterfall' Plot, Profile, and Spectra

Ultimately, we want to plot three subsequent items for each of the four Stokes parameters: a 2-D colormap ('waterfall') of the FRB intensity that is calibrated and RFI masked, the 'profile', which is a sum along all frequencies measured which produces a plot of the overall burst intensity over time, and the spectra that is summed over *only* the time window inside the estimated FWHM of the burst.

Along with the spectra, it will be helpful for us later to calculate the estimated noise levels dI, dQ, dU, and dV which are $\pm$ the standard deviations of each of their stokes parameter counterparts.

6.3 Target 11 Stokes Plots

For each of the two bursts (and for each of their four Stokes parameters), we have the profile and spectra to the top and right of the 'waterfall'—with the grey dashed lines indicating the estimated FWHM which the spectra is summed over, and the transparent green lines around the spectra are the estimated noise levels for each of the parameters' spectra. Note that the intensities on top are all from I, since the alternating nature of the other three parameters means that they will not necessarily sum well into the clear Gaussian shapes we saw previously.

#1

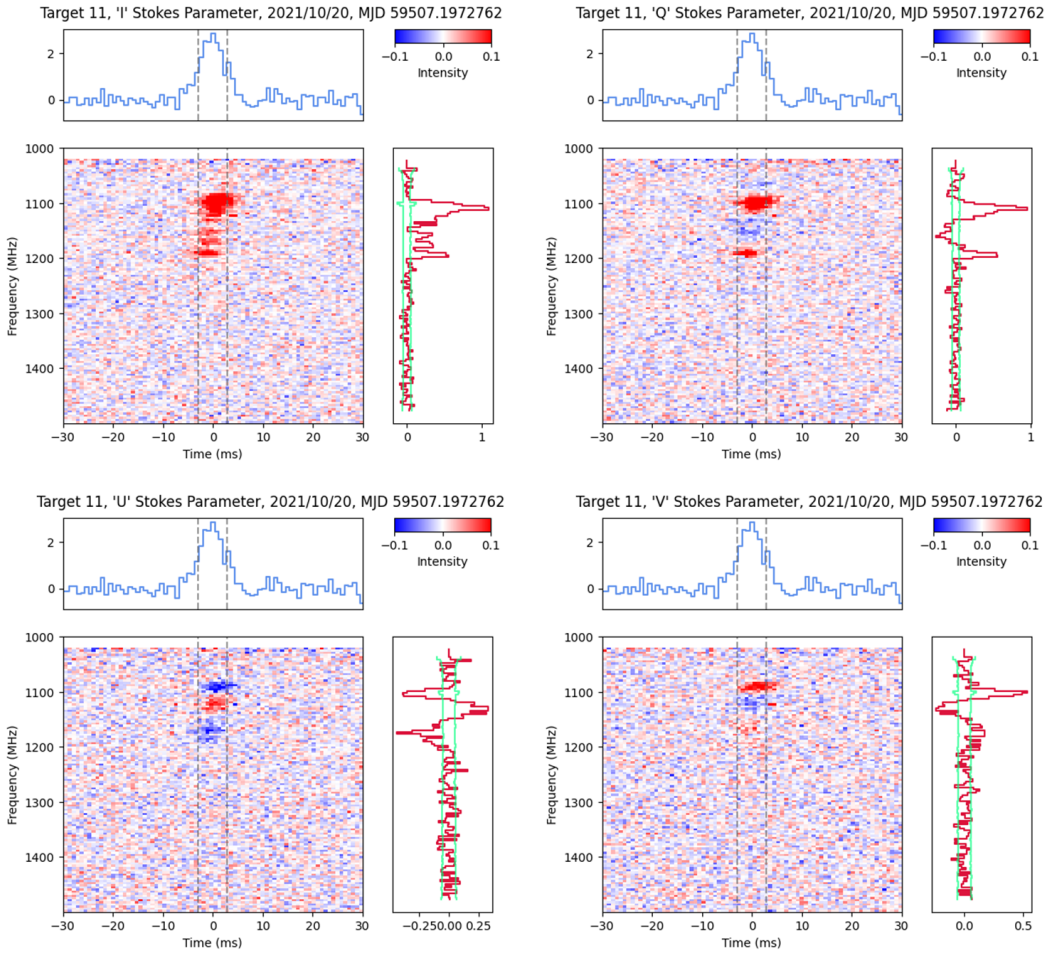

Figure 4: Profile, Waterfall, and Spectra of I, Q, U, V of the first FRB from Target 11

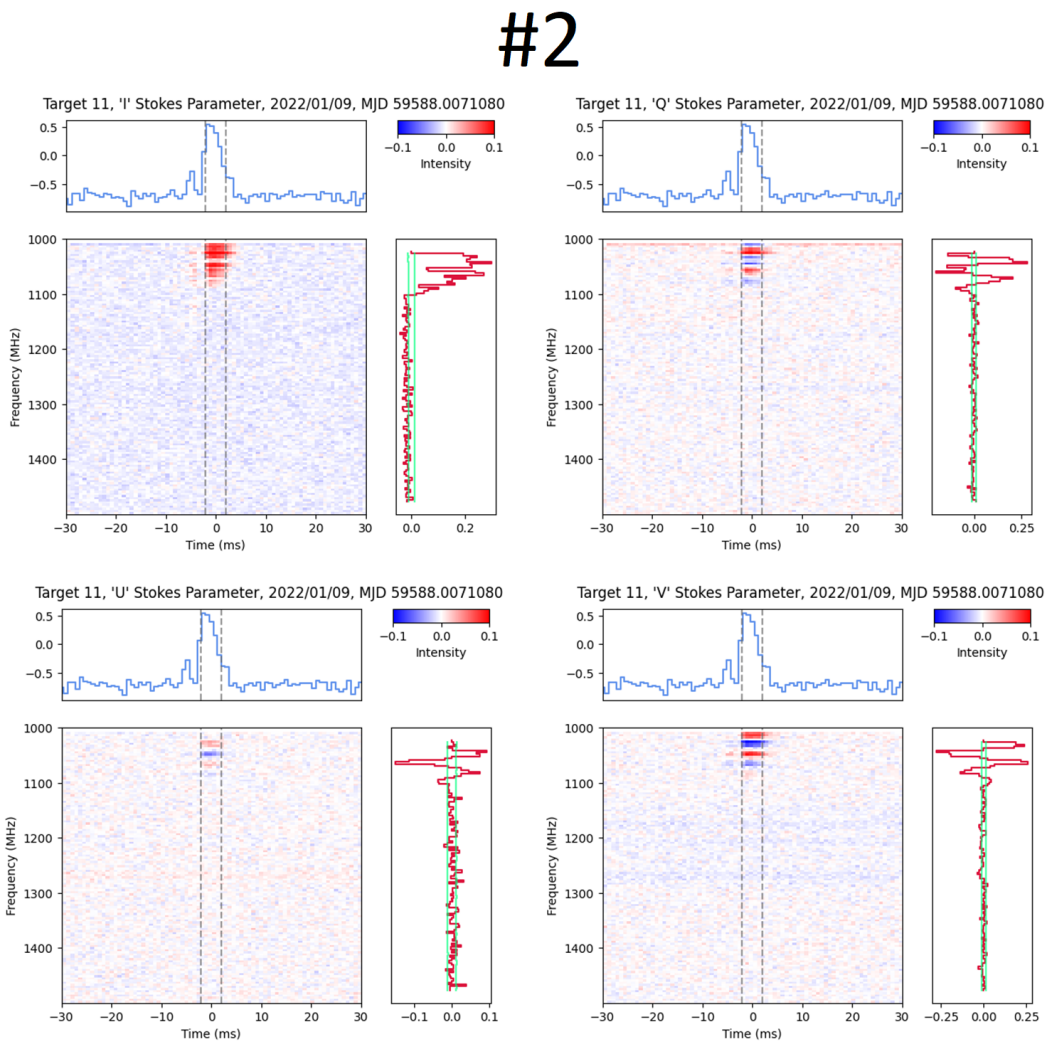

Figure 5: Profile, Waterfall, and Spectra of I, Q, U, V of the second FRB from Target 11

All of these plots show the bursts much clearer than previously (Fig. 1) with better RFI masking and background noise removal. All four of these parameters for both previously detected bursts show clear indicators of being from FRBs just as before, and the alternating intensity of Q and U in particular indicates that this is strongly polarized and that we will be able to compute a clear polarization and RM from these spectra.

7 Analysis of Bursts

7.1 Polarization

FRBs are thought to be highly polarized, and their polarization comes from the highly magnetic environments of their sources. The polarization of our incoming FRBs can be deduced from the differences between the four stokes parameters (for detections we are mostly concerned with I, but for analysis we now examine the others).

Specifically, the Linear Polarization is dependent on the Stokes parameters Q, U, and I

$$P_L \equiv \frac{\sqrt{Q+U}}{I} \tag{20}$$

We can also define the Faraday Dispersion Function (FDF) $F(\phi)$ as the Fourier Transform of the Polarization with respect to λ^2

$$F(\phi) \equiv \frac{1}{\pi}\int_{-\infty}^{\infty} P(\lambda^2)e^{2i\phi\lambda^2}d\lambda^2 \tag{21}$$

where ϕ is the Faraday depth, which has units of Rotation Measure (which will be convenient later). For our FRBs, we only care about λ^2 where our burst is present, so rather than integrating over the entire electromagnetic spectrum, we can introduce a window function $W(\lambda^2)$ which is only non-zero inside of our frequency/wavelength range:

$$\tilde{F}(\phi) \equiv \frac{1}{\pi}\int_{-\infty}^{\infty} W(\lambda^2)P(\lambda^2)e^{2i\phi\lambda^2}d\lambda^2 \tag{22}$$

this is called the "Dirty Faraday Dispersion Function" [Takahashi(2023)]. This is what we are concerned with for finding the rotation measure of our FRBs.

We can then produce $P(\lambda^2)$ (and thus $\tilde{F}(\phi)$) with our spectra $Q(f)$, $U(f)$, and $I(f)$ (using Eq. 20), which can easily be converted from being functions of frequency f to λ^2 in

software.

7.2 Rotation Measure (RM)

The Rotation Measure (RM) of the burst is proportional to the slope of the Polarization Spectrum $P(\lambda^2)$,

$$\mathrm{RM}(\lambda^2) \equiv \frac{1}{2}\frac{d}{d\lambda^2}[arg(P(\lambda^2))] \tag{23}$$

so for a constant RM value, we are looking for a consistent slope across the wavelengths of our burst. Fortunately, the RM value is also expressed as a peak ϕ_0 in the Dirty FDF $\tilde{F}(\phi)$.

For analysis we should visually look for both slopes in $P(\lambda^2)$ and a clear peak in $\tilde{F}(\phi)$ to confirm that the value for RM that our `RM-Tools` software[3] gives us is accurate.

Once again, the RM adds to the next leading order term after DM for the arrival time delay of the burst (Eq. 2), so finding the RM and analyzing its consistency tells us more about the details of the delays we should expect from an FRB source (which is more well informed as we have more bursts to analyze).

7.3 Target 11 Polarization and RM Results

The Spectra of Q, U, and I trimmed around the bursts' active frequencies (activity clearly beyond the noise threshold seen in the previous plots) fed into `RM-Tools` (Appendix A.3) with the plotting `-p` option. We also fed in the estimated noise levels dQ, dU, and dI, which gives us the error bars on the plots as well as outputting the estimated RM error.

[3]RM-Tools can be found at [Purcell et al.(2020)Purcell, Van Eck, West, Sun, & Gaensler], and commands used can be found in Appendix A.1.2

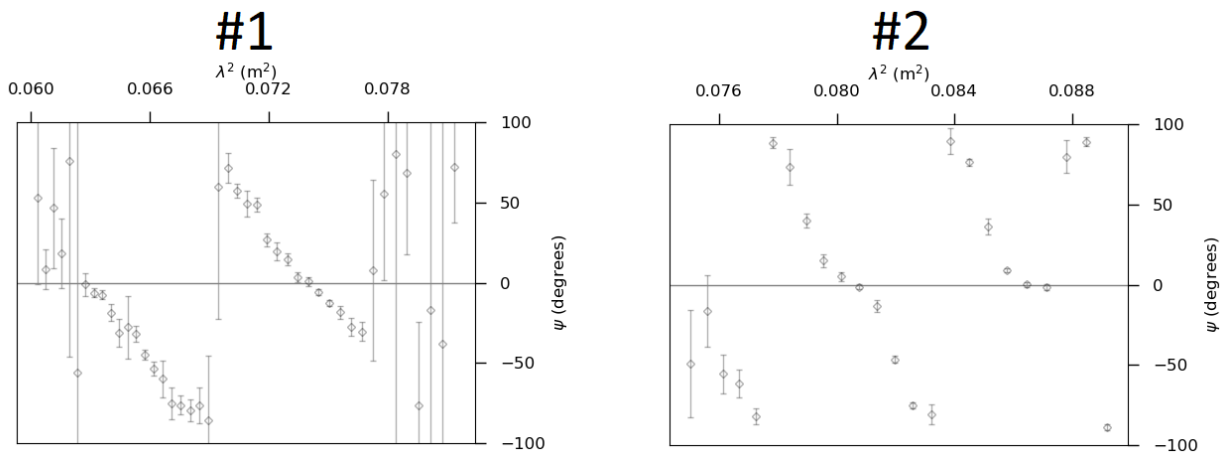

Figure 6: Plots of Polarization Angle v. Wavelength^2 for both bursts

For the first burst from Target 11, we see the negative slope is very clear within the wavelengths where the burst is present—we do not need the "downward lines" of points to be continuous (as it is split into two here), the important thing is that the slope is consistent. The first burst looks very linearly sloped downwards, and the second burst is is less clear (possibly due to it being more narrow-banded compared to the first burst, giving it fewer data points) but is still prominently sloped downwards in two sections similar to the first.

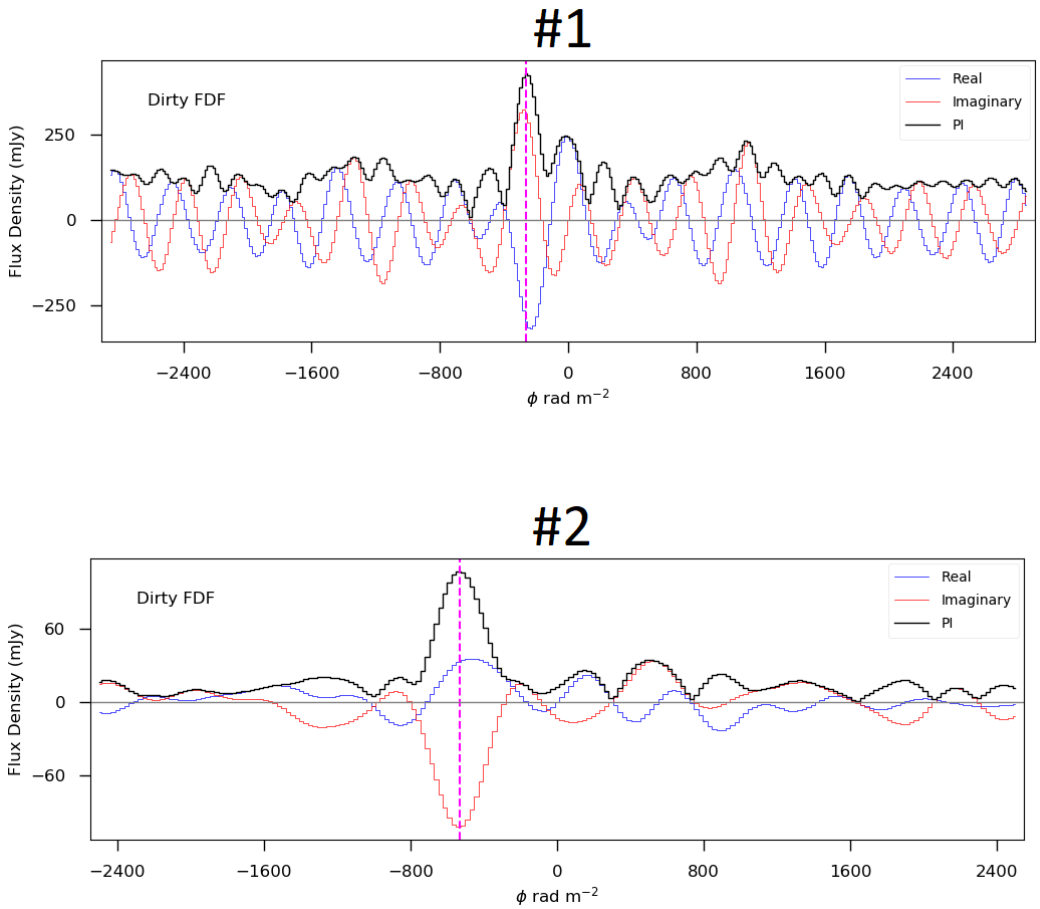

Figure 7: Plots of Dirty Faraday Dispersion Function for both bursts

In both Dirty FDF plots, we see clear peaks near the values proportional to the slopes of the previous two plots (rather than the highest peak at 0, which is indicative of unpolarized noise). This time, the second burst's Dirty FDF has a much clearer peak than the first; however, both are significant enough that—combined with the previous plots of $P(\lambda^2)$—we can be reasonably sure that our RM values are representative of these bursts.

Below are the full RM outputs as we calculated with `RM-Tools` (Appendix A.1.2) in accordance with the plots and conclusions above.

Burst #	Target	Date	Time (UT)	Fitted RM	RM Error
1	11	2021-10-20	03:12:10.824	-268.123	2.534
2	11	2022-01-08	22:28:02.530	-531.638	1.186

Table 4: List of found bursts from Target 11 with locations, times, and fitted RM parameters

Additionally, in both of the Dirty FDF plots, the Real and Imaginary components are prominently different under the peak at the RM value, indicating that these FRBs are highly polarized and that our assessment of RM is legitimate for these bursts.

We might initially be surprised that (as seen in Table 4) our RM for the second burst (-531.638) is nearly double that of the first burst (-268.123). As an FRB source, Target 11 is known to have a fast-varying RM because it is in a very magneto-active environment which may cause it to change quickly.

8 Conclusion

Since the previous project, we have expanded our FRB search to broader detection timing windows, estimated Target 11's rate when observed with FAST, showcased our bursts in all four Stokes parameters with improved background noise and RFI reduction, and determined the polarization of our bursts leading to an estimate of their Rotation Measures. While some of these tasks were unsuccessful, the prominence of these bursts' polarization provided a very strong result for the RM and supported the idea that Target 11 is a highly magneto-active environment causing not only strongly polarized bursts but also can have a largely variant RM between bursts.

8.1 Future Steps

In the future, the next steps would be to continue to search for more FRBs with PRESTO (which we now have a full pipeline to process and analyze) for the remaining targets we did

not initially scan, as well as further compare our rate estimates with other papers to confirm my method for estimating the Poisson statistics. Additional steps for these two bursts might be to estimate the EM for the fourth-order term of the time-delay 2 or to try and remove any low-level RFI that may be still remnant in these bursts to get more accurate measurements—some of the Stokes parameter plots (in section 6.3) still appear to have elevated noise that is uniform across frequency bands, which is indicative of RFI.

A Appendix

A.1 Code Commands Used

Previous Burst Search commands with PRESTO [Amiri et al.(2020)Amiri, Andersen, Bandura, & The CHIME/FRB Collaboration*]

Search your `.fits` file for prominent RFI with `rfifind`

```
rfifind -time 2.0 -o Lband /fitsfiles/Target_XX_tracking-M01_XXXX.fits
```

which will create a `.mask` file that tries to mask over any detected RFI

Next, prepare the data for de-dispersion by creating a topocentric, DM = 0 timeseries with `prepdata`

```
prepdata -psrfits -noscales -nooffsets -nobary -zerodm -downsamp 1 -dm XXX
   .XX -o Lband_rfifind.mask /fitsfiles/Target_XX_tracking-M01_XXXX.fits
```

which will output as a `.dat` file

Create a Fast Fourier Transform of the timeseries

```
realfft Lband_rfifind.mask.dat
```

whose output can inform what parameters you use in the next step.

De-disperse the data with `prepsubband`

```
prepsubband -nsub 64 -dmstep 1.00 -lodm XXX.X -numdms 1 -downsamp 1 -
   nobary -psrfits -o Lband_rfifind.mask /fitsfiles/Target_XX_tracking-
   M01_XXXX.fits
```

which will output a new de-dispersed `.dat` file

Search the outputted `.dat` file for single pulses

```
single_pulse_search.py -b Lband_rfifind.mask_DMXXX.XX.dat
```

which can be filtered through to find possible FRBs.

A.1.1 Variable Window Search

Searching the data for single pulses in PRESTO with variable window by adding the `-m` flag:

```
single_pulse_search.py -b Lband_rfifind.mask_DMXXX.XX.dat -m 10
```

where `10` should be replaced with your window (ms).

A.1.2 RM-Tools [Purcell et al.(2020)Purcell, Van Eck, West, Sun, & Gaensler]

To perform RM Synthesis, you should have a `.dat` file which contains `ASCII` formatted columns corresponding to: { f (in Hz), I, Q, U, dI, dQ, dU }

Then you execute the `do_RMsynth_1D.py` script (currently it only works in Python 2) with your `.dat` file as its argument

```
python2 do_RMsynth_1D.py list.dat -p
```

where the `-p` flag produces plots after the output.

To find the RM and dRM estimates, open the text of the newly outputted `list_RMsynth.dat` file

```
cat list_RMsynth.dat
```

and look for the values of `phiPeakPIchan_rm2` (RM) and `dPhiPeakPIchan_rm2` (dRM).

A.2 Rate Estimation Script

```
#!/usr/bin/env python3

resolution_mult = 100

import sys
import numpy as np
import matplotlib.pyplot as plt
from scipy.special import factorial


k = int(sys.argv[1])
percent_goal = float(sys.argv[2])
sample_time = float(sys.argv[3]) # (in hours)

def new_integrate_poission_to_k(rt, k, n):
```

```
    return np.sum(poisson(rt, np.linspace(0, n, n + 1))[:k+1])

def poisson(rt, k):
    exp = np.exp(-1 * rt)
    powr = np.power(rt, k)
    fact = factorial(k)
    value = (exp * powr) / fact
    return value

def find_closest_rt(target_percent, integral_v_rt, rt_list):
    index = np.argmin(np.abs(integral_v_rt - target_percent))
    closest_percent = integral_v_rt[index]
    rt = rt_list[index]
    return rt, closest_percent

n = 50

rt_res_mult = 2 * 100
rt_max = 50
rt_list = np.linspace(0, rt_max, rt_max * rt_res_mult)


print('Calculating possible rates with about', 100 * percent_goal, 'percent before k =', str(k), 'with a
    resolution of', 1/(10*resolution_mult))
print('Over', str(sample_time), 'hours')
print('')

print('...')
integral_up_to_k_v_rt = []
for rt_value in rt_list:
    integral_up_to_k_v_rt.append(new_integrate_poission_to_k(rt_value, k, n))
integral_up_to_k_v_rt = np.array(integral_up_to_k_v_rt)
print('...')

print('')
rt_result, percent_result = find_closest_rt(percent_goal, integral_up_to_k_v_rt, rt_list)
print('Closest (Rate * time):', str(rt_result))

print('Which gives a percentage of: ', str(percent_result))

print('')
rate_result = rt_result / sample_time
print('Final Rate: ', str(rate_result), 'bursts / hour (same units as time was entered)')
```

which is ran as

```
python estimate_rate_poisson_k.py 2 0.16 5.503
```

where the three arguments are k, the target percentage up to k (which is $= 1-$ percentage after), and the total observing time (in hours).

A.3 Burst Extraction Scripts

The burst extraction is split into three scripts (two of them are ran directly, and one uses functions that depends on the the third), and are ran as follows:

Slice out the burst using `slice_dd_v5_single.py`

```
python slice_dd_v5_single.py --info new_directory/burstmjd.txt fitsfiles/
   FAST/Target_11/202XXXXX/Target_XX_tracking-M01_XXXX.fits
```

which takes in a text file `burstmjd.txt that contains the MJD and DM of the burst`

Convert the `.ddar` file produced by the previous script into a Python-readable format using `get_int_data_py3.py`

```
python get_int_data_py3.py
```

which will output an `.npz` file that can be easily read into Python as `NumPy` arrays. Note that `get_int_data_py3.py` does NOT take in any arguments, you must instead edit the file manually to change the folder containing the `.ddar` and other parameters.

More information and links to download the scripts can be found at

`https://dzliprojects.wiki/index.php/User:Xanderjenkin#Burst_Cutting_Scripts`

A.4 Target Lists

Source_Name	Observing_Mode	Int_Time(s)	RA/Start_DEC/ TargetNumber	DEC/End_DEC
Target_1	Tracking	14400	23:14:51.74	+48:20:24.7
Target_2	Tracking	14400	12:03:43.34	+27:33:08.6
Target_3	Tracking	14400	04:23:16.46	+16:04:01.9
Target_4	Tracking	14400	17:02:41.61	+21:34:35.0
Target_5	Tracking	14400	07:01:02.37	+51:15:55.0
Target_6	Tracking	14400	08:50:01.87	+09:46:46.5
Target_7	Tracking	14400	08:28:21.31	+29:05:09.9
Target_8	Tracking	14400	18:08:04.05	+22:13:15.6
Target_9	Tracking	14400	00:33:04.34	+28:49:51.2
Target_10	Tracking	14400	16:37:18.91	+41:27:03.9
Target_11	Tracking	21600	13:52:01.10	+48:07:12.7
Target_12	Tracking	14400	04:17:39.67	+07:55:10.9

Table 5: List of FRB sources requested for this FAST survey (pulled from our proposal). Not all of the hours requested were observed, so the total time of our survey is less than the sum of the times listed above, and not all of the data surveyed was analyzed (see Table 2 for hours and sources which were scanned for bursts).

Target	R.A deg.	$\sigma_{R.A}$ deg.	Dec deg.	σ_{Dec} deg.	S/N	DM pc cm^{-3}	DM_{ex} pc cm^{-3}	R_{FAST} /h		
1	348.72	0.01	48.340	0.006	44	208.6	45.8	0.3		
2	180.931	0.007	27.552	0.007	28	111.6	51.6	0.4		
3	65.818	0.007	16.067	0.009	27	182.7	56.8	0.5		
4	255.673	0.003	21.576	0.003	64	127.7	37.8	0.5		
5	105.259	0.006	51.265	0.004	57	147.1	32.3	0.3		
6	132.51	0.01	9.77	0.01	19	150.4	57.5	0.5		
7	127.088	0.009	29.086	0.006	35	141.9	43.9	0.4		
Repeaters										
Target	R.A deg.	$\sigma_{R.A}$ deg.	Dec deg.	σ_{Dec} deg.	DM pc cm^{-3}	DM_{ex} pc cm^{-3}	N_{bursts}	RM	active	R_{FAST} /h
8	272.0169	0.0075	22.221	0.0082	218.1	103.1	2	L		0.9
9	8.2681	0.0038	28.8309	0.0034	201.7	125.9	5		Y	2.3
10	249.3288	0.0104	41.4511	0.0086	222.2	156.4	2	vary		0.7
11	208.0046	0.0037	48.1202	0.0024	221.6	161.7	24	vary/L	Y	8.2
12	64.4153	0.0127	7.9197	0.0174	287.1	200.7	12		Y	6.1

Table 6: Full list of parameters for the proposed FRB sources. In this project, Targets 2, 3, 4, 11, and 12 were scanned for FRBs (more details in Table 2 The $DM_{ex} = DM - DM_{ISM} - DM_{MWhalo}$, while assuming that the Milky Way halo contribution is $DM_{MWhalo} = 30$ pc cm^{-3}. If the FRB from the source has $||RM|| > 400$ rad m^{-2}, the RM (Faraday rotation measure) column is labelled with 'L'. Definitions for RM and EM (emission measure) are included later on with the full time-delay $t(\nu)$ (Eq. 2).

A.5 Variable Window Search Table

Target #	Significance	default window	1ms	10ms	100ms	1000ms
2	$> 13\sigma$	6	32	32	32	32
2	$> 15\sigma$	0	5	5	5	5
4	$> 9\sigma$	4	4	4	4	4
11	$> 9\sigma$	24	55	55	55	55
11	$> 11\sigma$	3	11	11	11	11
12	$> 6\sigma$	8	13	13	13	13
12	$> 8\sigma$	1	1	1	1	1

Table 7: List of detected single pulses per target above specified Signal/Noise ratio significances using variable detection windows. Our goal was to see new detections $> 10\sigma$ (which most of the sources did not have) and visually inspect the highest peaks for FRBs.

References

[Amiri et al.(2020)Amiri, Andersen, Bandura, & The CHIME/FRB Collaboration*] Amiri, M., Andersen, B., Bandura, K., & The CHIME/FRB Collaboration*. 2020, Nature, 582, 351, doi: 10.1038/s41586-020-2398-2

[Cordes & Chatterjee(2019)] Cordes, J. M., & Chatterjee, S. 2019, Annual Review of Astronomy and Astrophysics, 57, 417, doi: 10.1146/annurev-astro-091918-104501

[Kaspi & Beloborodov(2017)] Kaspi, V. M., & Beloborodov, A. M. 2017, Annual Review of Astronomy and Astrophysics, 55, 261, doi: 10.1146/annurev-astro-081915-023329

[Lawrence et al.(2017)Lawrence, Vander Wiel, Law, Burke Spolaor, & Bower] Lawrence, E., Vander Wiel, S., Law, C., Burke Spolaor, S., & Bower, G. C. 2017, The Astronomical Journal, 154, 117, doi: 10.3847/1538-3881/aa844e

[Petroff et al.(2022)Petroff, Hessels, & Lorimer] Petroff, E., Hessels, J. W. T., & Lorimer, D. R. 2022, The Astronomy and Astrophysics Review, 30, 2, doi: 10.1007/s00159-022-00139-w

[Purcell et al.(2020)Purcell, Van Eck, West, Sun, & Gaensler] Purcell, C. R., Van Eck, C. L., West, J., Sun, X. H., & Gaensler, B. M. 2020, Astrophysics Source Code Library, ascl:2005.003. https://ui.adsabs.harvard.edu/abs/2020ascl.soft05003P

[Sheikh et al.(2024)Sheikh, Farah, Pollak, Siemion, Chamma, Cruz, Davis, DeBoer, Gajjar, Karn, Kittling
Sheikh, S. Z., Farah, W., Pollak, A. W., et al. 2024, Monthly Notices of the Royal Astronomical Society, 527, 10425, doi: 10.1093/mnras/stad3630

[Straten et al.(2010)Straten, Manchester, Johnston, & Reynolds] Straten, W. v., Manchester, R. N., Johnston, S., & Reynolds, J. E. 2010, Publications of the Astronomical Society of Australia, 27, 104, doi: 10.1071/AS09084

[Takahashi(2023)] Takahashi, K. 2023, Publications of the Astronomical Society of Japan, 75, S50, doi: 10.1093/pasj/psac111

[Yates(2005)] Yates, R. D. 2005, Probability and stochastic processes : a friendly introduction for electrical and computer engineers, 2nd edn., ed. D. J. Goodman (Hoboken, NJ: John Wiley & Sons)

Acknowledgements

Special Thanks to my wonderful advisor of two junior papers Dr. Dongzi Li (currently Princeton University), Suryarao Bethapudi (Max Planck Institute) for his immense help with the software, my previous advisors who first introduced me to Radio Astronomy Professors Steve Eales and Matt Smith (Cardiff University), and all of the faculty, staff, and other students in the Princeton University Astrophysics Department for all of your support!