# Fast Radio Burst Morphology from Baseband Data with the CHIME Telescope

Daniela Breitman

Department of Physics
University of Toronto
Toronto, Ontario, Canada

August 21, 2021

A thesis submitted to University of Toronto in partial
fulfilment of the requirements of the Master of Science degree.

# Contents

# 1 Introduction

In 2007, a bright ($\sim$ 2 Jansky (Jy)) short ($\sim$ 8 ms) burst, today known as the Lorimer burst, was discovered in pulsar archival data from the Parkes Observatory (Lorimer et al., 2007). This burst is shown in Figure 1.1 with intensity plotted as a function of time and frequency. The top panels show the pulse profile at three different frequency channels. The burst had many characteristics different from pulsars. For starters, the data suggested that its origin was extragalactic, in contrast to pulsars. Moreover, the burst was unique - it was not observed to repeat. Today, these short extragalactic bursts are known as fast radio bursts (FRBs). Over 600 have been observed (Bannister et al., 2017, Burke-Spolaor & Bannister, 2014, Caleb et al., 2017, Champion et al., 2016, The CHIME/FRB Collaboration et al., 2021a) to date, and yet their physical origin remains elusive in spite of dozens of proposed models (Platts et al., 2019). To add to the mystery, a few years ago some FRBs have been discovered to repeat (CHIME/FRB Collaboration et al., 2020a, Connor et al., 2016, Spitler et al., 2016).

Although many models have been proposed (Platts et al., 2019), the favoured central engines for producing FRBs are pulsars and magnetars because the short burst duration of FRBs implies a relatively small emission region generated by coherent emission similar to that seen from pulsars and radio magnetars, but much brighter.

## 1.1 Pulsars

Radio pulsars are highly magnetized, rapidly rotating neutron stars with misaligned rotation and magnetic axes. The pulsar emits radiation beams from the magnetic poles, as it rotates, much like a lighthouse. An observer located in the path of the beam as the pulsar rotates detects a series of pulses with period corresponding to the pulsar rotation frequency. Thus, as the radio beam from B1919+21 - the first pulsar discovered - transits above a telescope, astronomers observe a series radio of radio pulses that correspond to rotations of the pulsar. Sometimes, when the magnetic and rotation axes are nearly orthogonal, beams from both poles can be observed so that two pulses correspond to a single pulsar rotation.

Pulsars are believed to emit in the entire electromagnetic spectrum and are sometimes observed to do so, but most of the known pulsars have been observed in the radio band (Kaspi & Kramer, 2016). Today, data for over 3000 pulsars has been published [1], with periods ranging from about 1 ms to about 23 s (Manchester et al., 2005). The majority of the pulsar population is confined to the Galactic plane, but is still very incomplete as the known pulsar population is still growing.

---

[1] https://www.atnf.csiro.au/research/pulsar/psrcat/
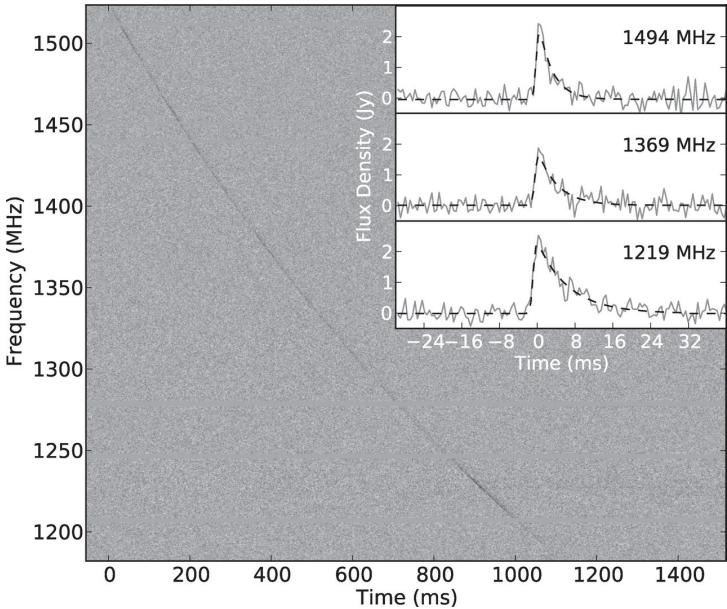
Figure 1.1: The first FRB discovered, known as the Lorimer burst. We see how the original signal is dispersed in frequency as a function of time. The three panels show the pulse flux density at three different frequency values: one at the bottom of the band, in the middle, and at the top of the band, with the peaks of the signal aligned at 0 ms. Figure 2 of (Thornton et al., 2013).

## 1.2 Magnetars

Magnetars are young neutron stars powered by the strongest magnetic fields in the Universe (Kaspi & Beloborodov, 2017). They are most often observed in X-rays and soft-$\gamma$ rays (Kaspi & Beloborodov, 2017). Magnetars have extremely diverse behaviours, from short bursts, to giant flares (Kaspi & Beloborodov, 2017). They are also active across a wide range of timescales, from milliseconds all the way to many months (Kaspi & Beloborodov, 2017).

Out of the thirty Galactic magnetars known (Kaspi & Beloborodov, 2017), radio pulsations have only been detected in five so far [2]. About a year ago, a bright radio burst from the Galactic magnetar SGR 1935+2154 was detected by both CHIME/FRB (with a band-averaged flux integrated over the burst duration aka fluence of 700 kJy ms) and STARE2 telescopes (CHIME/FRB Collaboration et al., 2020b, Margalit et al., 2020). This was the first radio burst detected from this magnetar, and more faint radio bursts have been detected from it since. This is an exceptional discovery, because, for a long time, it was thought that FRBs come from magnetars (Lyubarsky, 2014), but the only challenge with this theory was that FRB luminosities are orders of magnitude brighter than those of Galactic magnetars. This detection provides strong evidence that magnetars could be the origin of at least the faintest of FRBs (CHIME/FRB Collaboration et al., 2020b, Margalit et al., 2020).

---

[2]http://www.physics.mcgill.ca/~pulsar/magnetar/main.html

## 1.3    Dispersion

Dispersion occurs when the radio signal traverses clouds of free electrons which reduce the group velocity of the signal while affecting lower frequencies more significantly than higher ones. For an observer at a distance $d$ from the source, the dispersion delay time $t$ is: (Condon & Ransom, 2016)

$$t = \left( \frac{q_e^2}{2\pi m_e c} \right) \nu^{-2} \int_0^d n_e dl, \qquad (1.1)$$

where $n_e$ is the free electron column density along the line of sight, $\nu$ is the frequency, $c$ is the speed of light, and $q_e$ and $m_e$ are the charge and mass of the electron, respectively. We then define the dispersion measure (DM) as the integrated free electron density along the line of sight from an observer to a source at distance $d$:

$$\text{DM} \equiv \int_0^d n_e dl$$

Therefore, a signal with an extremely high DM could either come from a very dense region such as the Galactic centre or from extremely far away. FRBs are therefore thought to be extragalactic due to their unusually high DM compared to the expected Galactic DM contribution. The Lorimer burst, for example, had a DM of 375 pc cm$^{-3}$ (Thornton et al., 2013) in comparison with the 75 pc cm$^{-3}$ predicted by NE2001 (Cordes & Lazio, 2002) - a model of the Milky Way electron density. The 300 excess DM units are due to free electrons in the intergalactic medium (IGM), host halo, and host contribution, with the total observed DM calculated as follows: (Chawla et al., 2021)

$$\text{DM}_{\textbf{total}} = \text{DM}_{\text{MW}} + \text{DM}_{\text{WM,halo}} + \text{DM}_{\text{IGM}} + \frac{\text{DM}_{\text{host,halo}}}{1+z} + \frac{\text{DM}_{\text{host}}}{1+z}, \qquad (1.2)$$

where $z$ is the redshift of the FRB. The two leading models of the Milky Way free electron density are NE2001 (Cordes & Lazio, 2002) and YMW16 (Yao et al., 2017). These models predict a DM $\lesssim 30$ at the edges of the Milky Way, and DM $\gtrsim 1000$ toward the Galactic centre. Each of these models includes mathematical descriptions of the free electron density of various components of the Milky Way galaxy and are calibrated using pulsar data and independent distance measurements. The Milky Way halo contribution to DM is highly uncertain, but is usually expected to be less than 50 cm pc$^{-3}$ (Dolag et al., 2015, Keating & Pen, 2020, Prochaska & Zheng, 2019). The IGM contribution to the DM can be calculated using the Macquart DM-z relation (Macquart et al., 2020) among other methods. An estimate is $\text{DM}_{\text{IGM}} = 1000 \cdot z$. Like for the Milky Way halo, host galaxy halo estimates are also highly uncertain for all host types and are also usually under 50 pc cm$^{-3}$. There are models to describe the DM distributions for various galaxy types other than spiral such as dwarf (Xu & Han, 2015) and elliptical (Chawla et al.,

2021, Yao et al., 2017). In addition to dispersion, other important propagation effects are Faraday rotation (Brentjens & de Bruyn, 2005) and scintillation (e.g. Simard & Pen (2018)).

## 1.4   Fast Radio Bursts

In order to build new theories and rule out others, we must be aware of our limitations by carefully inspecting all the properties we can possibly measure for a given FRB source. In the following sections, we introduce leading theories of FRB origin and then discuss how various observable parameters can be used to test these models.

For more in depth reviews of the fast radio burst phenomenon, see (Petroff et al., 2019), (Rane & Lorimer, 2017), and (Katz, 2016).

There are many FRB emission and progenitor models that have been proposed. In fact, at some time, more proposed theories existed than FRBs detected! The two most popular categories of FRB progenitor models are interaction models and supernova remnant models. For an overview of FRB theories, see (Platts et al., 2019).

## 1.4.1    Emission Mechanisms
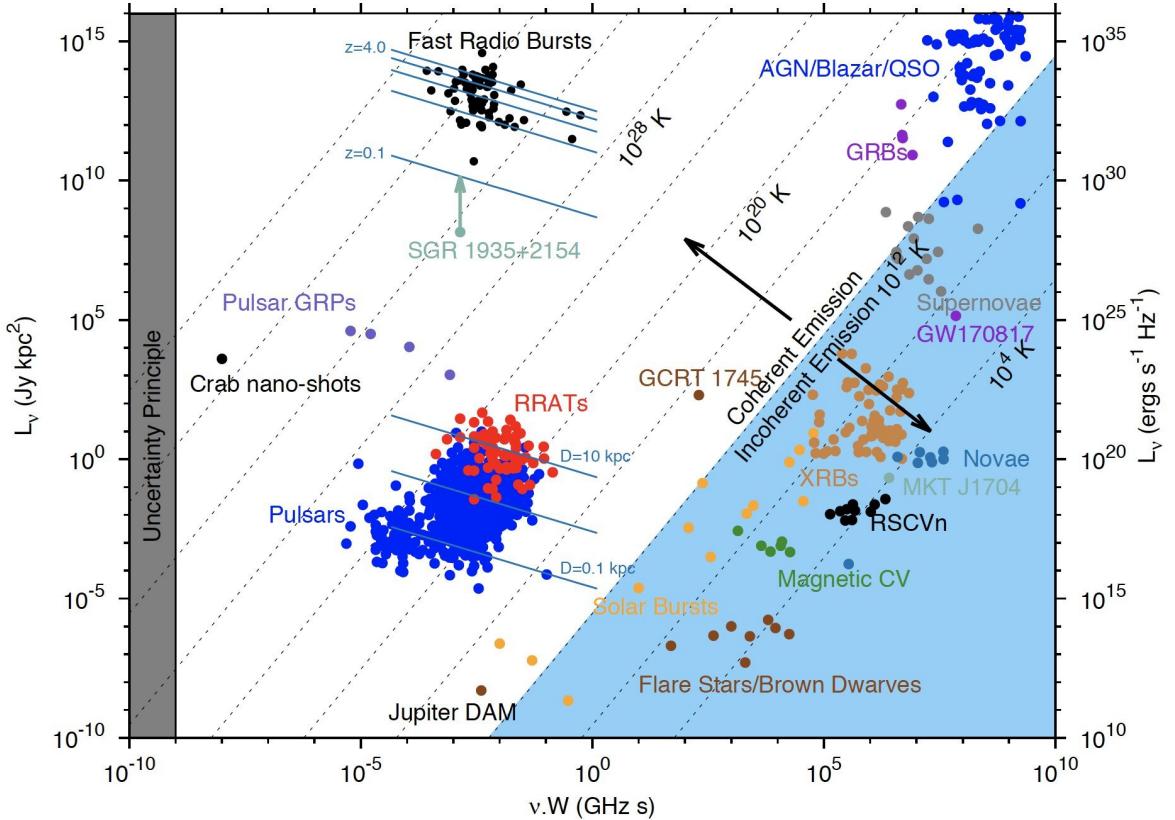


Figure 1.2: Pseudo-luminosity against emission frequency multiplied by the burst width of different classes of radio transients. Diagonal dashed lines have constant effective brightness temperature, with $T_B \sim 10^{12}$ K being the threshold between coherent and incoherent emission processes. Note that FRBs stand out among the radio transient because of their luminosity. Figure adapted from Keane (2018).

The FRB emission mechanism must explain the physics that occur in the progenitor to produce the required energy to emit the FRB. Incoherent emission with power $\propto n$, where $n$ is the number of particles, would require the blackbody to have a brightness temperature $\gtrsim 10^{35}$ K, which is not physical. Even in a relativistic scenario, the brightness temperature would be reduced by the Lorentz factor $\gamma$, which, even for $\gamma \sim 10^6$ still leaves us with a very high unphysical brightness temperature. The boundary between coherent and incoherent emission is at a brightness temperature $\sim 10^{12}$ K (which is where the boundary in Figure 1.2 is drawn). The FRB emission mechanism must therefore be coherent with power $\propto n^2$, with the electromagnetic waves emitted in phase and thus adding up.

Many FRB emission mechanisms have been proposed. Generally, they fall into two categories: near-field models and far-away models. Near-field models propose that the radiation is emitted inside the magnetosphere. Far-away models propose that relativistic beaming is converted partially into coherent emission. Since the jet is relativistic, the conversion into incoherent emission occurs far away from the source. Some coher-

ent emission mechanisms that have been proposed are synchrotron masers (far-away), bunched particles (also known as antennae, near-field). Particles accelerated by a magnetic field are bunched into groups with the same phase and thus are coherent. There are various astrophysical bunching mechanisms such as magnetic braking, and magnetic reconnection. Magnetic braking occurs when ionized material from a rotating object is captured and ejected into space by the magnetic field, thus spinning down the object. The bunched ejecta is accelerated radially along the magnetic field lines and emits high energy radiation.

As discussed in Section 1.2, the discovery of an FRB from Galactic magnetar SGR 1935+2154 is of particular importance as it tells us that objects that produce FRBs may share the same emission mechanism as magnetars. We can see that the gap in phase space between FRBs and giant pulses in Figure 1.2 has been closed thanks to this discovery. Moreover, this discovery tells us that magnetars are potential progenitors for FRBs.

### 1.4.2 Progenitor Theories

According to Lu & Kumar (2018), the progenitor theory must explain the FRB event rate, energetics, and timescale (i.e. burst width). Currently, the event rate estimated from the CHIME/FRB first catalog (The CHIME/FRB Collaboration et al., 2021a) is $\sim 1200$ /sky /day above fluence of 5 Jy ms, with energetics $\lesssim 10^{43}$ erg/s, and widths of order $\mu$s to ms.

Two main types of progenitor theories have been proposed so far: interaction models, and supernova remnants (Platts et al., 2019). Interaction models propose that the FRB signal is generated by interactions, such as merging, of two objects such as neutron stars. Such models may produce both repeating FRBs and non-repeating FRBs, depending on whether the interaction is catastrophic in nature. A strong support toward these models would be if we observed the FRB to be in a binary system, or observed a gravitational wave and / or X-ray counterpart.

The second type of progenitor theories is supernova remnants such as pulsars and magnetars. For a pulsar to have bursts as energetic as what is observed from FRBs, the pulsars would have to be rotating very rapidly. The Crab pulsar is notorious for its giant pulses. Models similar to those explaining Crab giant pulses suggest that FRBs could be generated by giant pulses from young rapidly rotating pulsars. The large Faraday rotation measures observed for FRBs are expected in this model and are due to the supernova remnant resulting from the creation of the young pulsar. A strong evidence for this model would be to observe a regular periodicity of FRB repeaters. Magnetars are strong candidates for FRB emission as they are extremely powerful objects which could produce flares on energy scales similar to those observed in FRBs. These young systems are surrounded by very dense plasma, and therefore one supporting argument for these models would be evidence of a large local DM contribution.

We now examine all available observables and discuss how they can be used to dis-

tinguish between some of these models.

### 1.4.3  Observable Quantities

- Localization: The sky localization of the FRB is important because potentially associating an FRB to its host galaxy can tell us the redshift (if the redshift of the galaxy is known), the galaxy type, and the position of the source within the galaxy which tells us if it is a younger or older object. All of these may help identify the progenitor of the FRB. $\sim 1$ arcsec localization is needed in order to successfully associate an FRB to its host galaxy. Localization is usually improved via the detection of the same signal by multiple dishes, either via interferometry, very long baseline interferometry (VLBI) (see Cassanelli et al. (2021)), or multi-wavelength counterpart detection (see Tohuvavohu (2020)). CHIME/FRB intensity data is usually localized to $\sim 1$ arcmin resolution, which is usually not enough to unambiguously associate a burst with a host galaxy. However, the CHIME project is currently expanding to two outrigger sites near CHIME which will allow for a baseline long enough to improve localization capacity to 50 milliarcsec resolution by 2023.

- Multi-wavelength counterpart: Observing counterparts to FRBs in other bands of the electromagnetic spectrum or in the form of gravitational waves is key to narrowing down potential FRB emission models and progenitor theories. Currently, the CHIME/FRB system works with a virtual observatory event (VOEvent) program which brings subscribers from other surveys such as Swift's Burst Alert Telescope (BAT) to receive alerts of FRB detection. This effort has already yielded a possible co-detections of FRBs in the radio and X-ray bands (Tohuvavohu, 2020).

- Redshift: Once an FRB has been unambiguously associated to a host galaxy, its redshift can be precisely determined via spectroscopy. If no host association is available, the redshift can be estimated from its DM using the Macquart relation. On the other hand, if the host galaxy and its redshift are known, a robust estimate of $DM_{IGM}$ can be used to obtain $DM_{host}$ (see Eq. 1.2), which tells us about the local environment of the FRB.

- DM: The original FRB signal is dispersed. Two main methods are used in order to obtain its DM and de-disperse the FRB. One is to maximize the signal-to-noise (S/N) of the burst, the other is to maximize the structure by aligning the voids in the signal. Currently, FRBs up to DM of 3000 pc cm$^{-3}$ have been detected. The DM distribution of the FRB population can rule out models that suggest FRBs are in young supernova remnants, for example, since such models require FRBs to exhibit a large local DM.

- Dispersion index: The dispersion index $\alpha$ is defined as $\Delta t \propto \nu^{\alpha}$, relaxing the assumption that $\Delta t \propto \nu^{-2}$ as shown in Equation 1.1.

- Waterfall: The intensity plotted as a function of frequency and time: also known as the waterfall or the spectrum, shows the spectral structure of the FRB as a function of time. FRB signal has been observed from 100 MHz all the way to 8 GHz. Some bursts exhibit a very complex waterfall with multiple components that change with frequency. For repeating bursts, for example, Pleunis et al. (2021) found that the components exhibit a downward drifting pattern which can be measured as well.

- Width: Fitting a model to the FRB spectrum allows us to extract the full width half maximum (FWHM) of the burst signal. If there are many components, each component then has its own width.

- Scattering: Scattering occurs due to the presence of high density clouds of plasma. This effect is frequency dependent and is inversely proportional to $\nu^4$, indicative that lower frequencies are much more affected than higher frequencies.

- Scattering index: The scattering index relaxes the assumption that scattering is proportional to $\nu^{-4}$.

- Polarization: Saving the polarization data, as opposed to just the total intensity, allows us to measure a few other properties of FRBs and their propagation. One is the intrinsic polarization angle of the source $\psi_0$, modulated by Faraday rotation which is described by the Faraday rotation measure (RM). We can also measure the polarization position angle (PPA), obtained by de-rotating the signal with the RM value. The PPA describes the geometry of the polarized signal before being altered by Faraday rotation. Measuring the PPA for many different FRBs is key in understanding more about whether there are multiple FRB populations and what are their emission mechanisms and progenitors.

- Flux and fluence: the measured flux and fluence (integrated flux) are extremely important as any FRB emission model must be able to explain these energy scales. For FRBs detected by CHIME/FRB, fluxes and therefore fluences are lower limits due to the fact that the burst localization is not known well enough in order to locate them within the instrument beam, and thus, we assume they were detected at peak sensitivity.

- Repeaters: For repeating FRBs, in addition to the above properties, we can also measure their repetition rate and their periodicity. In the case of an FRB repeating irregularly, it is also especially important to measure when and for how long it is active, since models will have to explain these patterns.

## 1.5 The Canadian Hydrogen Intensity Mapping Experiment

The Canadian Hydrogen Intensity Mapping Experiment (CHIME) is located near Penticton, British Columbia, Canada on the grounds of the Dominion Radio Astrophysical Observatory (DRAO). As seen in Figure 1.3, CHIME is located in a valley and is surrounded by mountains which help reduce the radio frequency interference (RFI). CHIME is a transit telescope, and therefore has no moving parts. It is composed of four 20 m wide 100 m long cylindrical paraboloidal reflectors.



Figure 1.3: Canadian Hydrogen Intensity Mapping Experiment in Penticton, British Columbia, Canada, in the South-East direction. The four cylindrical reflectors lay along the North-South direction and each have 256 dual polarization feeds at their focal line. The 26 m Galt telescope is visible on the left. Drone image courtesy of CHIME Collaboration.

The first motivation for the CHIME project was to map the neutral hydrogen content of the entire Northern hemisphere in order to measure the accelerating expansion of the Universe at redshift $0.8 \lesssim z \lesssim 2.5$ and providing an additional constrain on cosmological parameters. To measure the weak signal of the baryon acoustic oscillation (BAO), a large collection area and fast mapping is desired, motivating the choice of a telescope with hundreds of feeds. Since no sky position is favoured, a transit telescope is a much more economical alternative to the conventional pointing dish design. The key properties of the CHIME telescope are listed in Table 1.1.

The CHIME instrument has four cylindrical reflectors, each 20 m wide and 100 m long with 256 dual polarization feeds, for a total of 2048 signal paths. Coaxial cables carry the analog signal from the feeds to the CHIME correlator, which is made up of two parts, the F and X engines. The F engine uses custom Field Programmable Gate Array (FPGA) boards to digitize and channelize the analog signal from the feeds at a rate of 13.1 Tb/s for the 2048 timestreams. It then channelizes the 400 MHz bandwidth of each

| Parameter | Value |
|---|---|
| Collecting area | 8000 m$^2$ |
| Longitude | 119$^o$37'25'.25 W |
| Latitude | 49$^o$19'14".52 N |
| Frequency range | 400 - 800 MHz |
| Polarization | Orthogonal linear |
| E-W FoV | 2.5$^o$ (400 MHz) - 1.3$^o$ (800 MHz) |
| N-S FoV | $\sim 100^o$ |
| Number of feeds | 1024 |
| Intensity real-time time resolution | 0.98304 ms |
| Intensity real-time frequency resolution | 24.4 kHz |
| Source transit duration | North Celestial Pole: 24 hrs |
| | 45$^o$: 7 - 14 min |
| | Equator: 5 - 10 min |

Table 1.1: Properties of the CHIME telescope related to the CHIME/FRB project (CHIME/FRB Collaboration et al., 2018).

timestream into 1024 frequency bins, each 390 kHz wide, using a polyphase filter bank. The X engine consists of a GPU cluster that performs spatial correlation of the signal from all feeds in order to form a sky map (Ng et al., 2017).

### 1.5.1 CHIME/FRB Real-Time Radio Transient Search Pipeline

The CHIME/FRB real-time pipeline is composed of four stages, as shown in Figure 1.4. The purpose of the pipeline is to process data in real time and trigger an intensity data callback for astrophysical events (grey dashed arrow). Astrophysical events meeting certain criteria also trigger a baseband (i.e. raw voltage) data callback (grey dashed arrow). In this section, we summarise the role of every one of the four stages of the real-time pipeline.
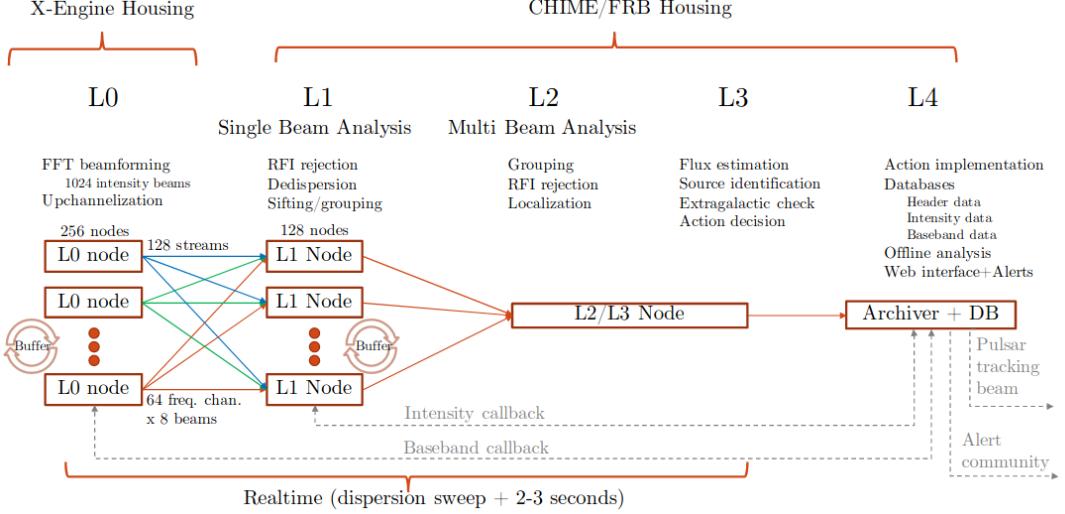
Figure 1.4: The four stages of the CHIME/FRB science pipeline. Stage 0 involves beam-forming the data. Stages L0 to L3 are done in real time. After stage L0, the real-time pipeline will de-disperse incoming data and remove RFI (L1), group multi-beam events (L2), classify the candidates, call back data to save, and choose actions (L3). The last stage, L4, which implements the actions chosen in L3, does not occur in real time. Figure from CHIME/FRB Collaboration et al. (2018).

- **L0**: The L0 stage beamforms and upchannelizes the data. In the beamforming step, we synthesize a grid of 256 beams along the N-S direction and four beams in the E-W direction using a fast Fourier transform (FFT) algorithm (see Tegmark & Zaldarriaga (2009) and Masui et al. (2019) for details). This yields a total of 1024 formed static beams that are closely spaced. The spacing of the beams is a tunable parameter such that, along the N-S direction, the beams span $60^o$ above and below the zenith, along the E-W direction, the beams span $0.4^o$. The data stream is then upchannelized from a 2.56 $\mu$s time resolution and 1024 frequency channels to form the intensity data stream with a 0.983 ms time resolution and 16 384 frequency channels using a Fourier transform and downsampling in time.

- **L1**: The 128 L1 compute nodes each analyse 8 of the 1024 beams in parallel. Each node has a buffer which contains a total of 7 min of data. If triggered, the intensity data in the buffer can then be saved to disk. Every single beam is cleaned of RFI using both 1D and 2D excision methods. The most computationally expensive task of the real-time pipeline is to convert the time vs frequency data into time vs DM in order to efficiently detect dispersed impulse signals. We search over trial DM values up to DM of 13 000 pc cm$^{-3}$. In order to perform the search efficiently, we use a tree de-dispersion algorithm (CHIME/FRB Collaboration et al., 2018). The L1 stage outputs an L1 header for each detection, which includes the DM, arrival time, and sky coordinates of the trigger.

- **L2/L3**: A single node carries out the L2 and L3 stages. The L2 stage begins by scanning L1 headers and grouping them in time, DM, and sky position to form

multi-beam events. RFI events are sent straight to L4, while astrophysical candidates will be passed along to L3. The goal of the L2/L3 stage is to improve the sky localization, provide an initial event flux estimation, and decide which actions should be carried out for a given event such as save intensity data to disk, or save high-time-resolution (baseband) data to disk, for example. Following the detection of an event with the real-time pipeline, a baseband dump can be triggered within $\sim 14$ s of reaching the telescope dish. A buffer with 35.5 s of baseband data is maintained, therefore up to $\sim 20$ s of baseband can be saved for a given event, hence limiting the maximum DM of baseband sources to roughly 1000 pc cm$^{-3}$ at CHIME frequencies.

- **L4**: The last stage of the pipeline does not occur in real time, and involves carrying out the actions chosen in the L2/L3 stage and offline analysis.

For a full overview of CHIME/FRB and its systems, see CHIME/FRB Collaboration et al. (2018).

## 1.5.2 Baseband Data

The ultimate goal of FRB science is to understand the physics behind the FRB phenomenon and to identify the object that produces these signals. Studying the properties listed above for the FRB population is one way of bringing us closer to this goal.

Morphological analyses of FRBs are relatively unexplored territory when it comes to high time resolution, because not many surveys are able to record such data. However, it is a powerful way to constrain FRB emission mechanisms because the duration of sub-bursts of an FRB place a limit on the size of the emitting region. For example, a 1 $\mu$s component corresponds to an emitting region $\lesssim 300$ m (Nimmo et al., 2021).

Recent studies such as Nimmo et al. (2021), Pleunis et al. (2021) have shown morphological analyses to be a very promising avenue to understand the FRB phenomenon. Pleunis et al. (2021), for example, performed a morphological study of the entire CHIME/FRB Catalog 1 sample (The CHIME/FRB Collaboration et al., 2021a), which contains the intensity data of 536 FRBs. They were able to identify a significant difference between properties of repeaters and non-repeaters, thus suggesting that the two might come from separate populations and hence have different physical origins. One limitation of that study is the $\sim 1$ ms time resolution of the data, which makes it difficult to differentiate intrinsic burst structure from propagation effects (Gourdji et al., 2019). This leads to a difficulty to precisely determine the DM, pulse width, and scattering timescale.

In this work, some of these challenges can be mitigated thanks to our use of high-time-resolution data. However, some challenges remain. For example, it is much more difficult to model the data as it is generally more complex at higher resolution, with more structure resolved. The main goal of this work is thus to develop a processing pipeline to robustly model the complex temporal structure of high-time-resolution data and provide

a 1-D fit to both the pulse profile and spectrum. The second goal of this work is to apply this pipeline on a sample of 35 baseband events to assess its performance and draw conclusions about the physical properties measured. This work is thus the very first morphological study of a sample of this size and time resolution.

### 1.5.3 Thesis Outline

So far, we have been thoroughly introduced to the FRB phenomenon in Sections 1.1 - 1.4 and to the CHIME science mission and real-time detection system in Section 1.5. We have seen that the FRB phenomenon is very complex, and many questions remain to be answered. For instance: what is the FRB emission mechanism? In Chapter 2, we dive into the processing pipeline of the CHIME/FRB high resolution data product, known as baseband data. Chapter 3 forms the core of this thesis, where we describe the implementation of a processing pipeline to fit a complex model to baseband data thus enabling a systematic morphology study of FRBs at high resolution for the very first time. Finally, in Chapter 4, we apply the pipeline to a sample of bursts and draw conclusions in Chapter 5.

# 2 Baseband Pipeline

The real-time pipeline triggers baseband dumps for bright events (S/N threshold varies between 8-10) as described in the previous section. This typically yields 1-10 baseband dumps per day Michilli et al. (2021). The raw data of size $\sim 100$ GB per event is then processed with the baseband pipeline and ultimately produces a spectrum (intensity as a function of frequency and time) of size $\sim 3$ GB per event. During this processing, the event is de-dispersed and beamformed at multiple different sky locations in order to find its best localization. We begin by describing the beamforming and coherent de-dispersion processes which are called multiple times within the baseband pipeline. Lastly, we describe the localization process and characterise the CHIME/FRB localization capability.

Beamforming is the process of using the phase information of baseband data to phase-reference the signal in a specific direction within the field of view of the telescope prior to co-adding over feeds. This process maximizes the S/N of the burst, effectively pointing the telescope to that direction. To ensure that the instrument is optimally calibrated, we use the daily transit of a bright source to calibrate the gains of each of the receivers. The baseband pipeline uses the calibration closest to the source observation time as a phase reference. We then form beams using the real-time pipeline localization. Currently, beamforming is the most time consuming stage of the processing with one beam formed per hour with our current processing capabilities.

The detected radio signal is originally dispersed by free electrons along the path of the signal such that we observe high frequency signal that is less affected by dispersion before lower frequency signal that is more affected. In order to correct for this effect, we apply a de-dispersion algorithm to recover the original pulse. There are two main types of de-dispersion algorithms: coherent and incoherent. The most common way is to incoherently de-disperse simply by applying a time delay to each frequency channel with respect to a given reference frequency. However, this method is limited by the frequency resolution of the data, since each frequency channel has its own dispersion delay. This intrachannel residual smearing will smooth out features shorter than the smearing time and hence reduce the S/N. This effect becomes increasingly consequential with DM and is inversely proportional to frequency cubed. For example, for a source with DM 500 pc cm$^{-3}$ with 16 384 channels, the smearing at the bottom of the band is $\sim 1.6$ ms. For the same event but with the 1024 channels of native baseband resolution, the smearing at the bottom of the band is $\sim 25$ ms.
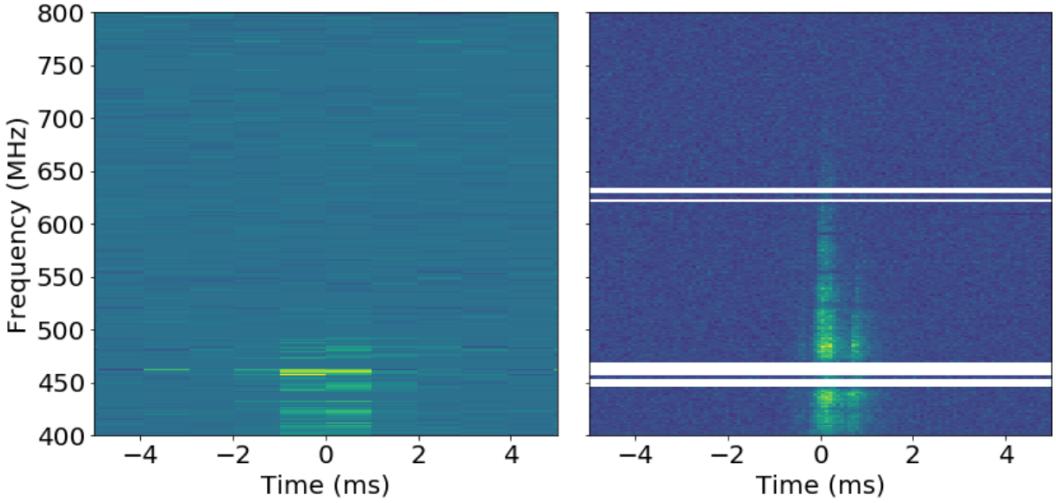
Figure 2.1: Burst from FRB 20190618A incoherently de-dispersed by the intensity pipeline (left) with $\sim 1$ ms resolution and coherently de-dispersed by the baseband pipeline (right) with 81.92 $\mu$s resolution (32 times the native resolution) both to the same value of 228.93 pc cm$^{-3}$. Note that the colour scale in both plots corresponds to the intensity but are independent between the two plots. Figure from (Michilli et al., 2021).

The second kind of de-dispersion involves using the phase information contained within the baseband data to model the intervening medium as a cold tenuous plasma, and thus mathematically correct for intrachannel smearing by rotating the phase with the inverse of the following transfer function:

$$H(\nu + \nu_0) = \exp\left(\frac{2\pi i \cdot \nu^2 \cdot k_{DM} \cdot DM}{\nu_0^2 \cdot (\nu + \nu_0)}\right), \qquad (2.1)$$

where $\nu$ is the frequency offset with respect to the reference frequency $\nu_0$, and $k_{DM}^{-1} = 2.41 \times 10^{-4}$ MHz$^2$ pc cm$^{-3}$ s$^{-1}$ is the dispersion constant (Michilli et al., 2021). Coherent de-dispersion thus allows us to resolve temporal features at higher resolution than incoherent de-dispersion. Figure 2.1 shows a comparison between the same event incoherently de-dispersed with the real-time pipeline and coherently de-dispersed with the baseband pipeline to the same value of 228.93 pc cm$^{-3}$. The plots show frequency as a function of time data of FRB 20190618A. In the intensity plot (left), the burst is barely resolved, while in the baseband data (right) we are able to distinguish three temporal components, spectral structure, and scattering. Note that the colour scales in both plots correspond to the intensity but are independent. The striking difference between the two plots is not only thanks to the higher resolution of baseband data, but also due to the beamforming of the data to the best source localization to maximise S/N.

In order to find the best localization of the burst, we map the intensity of the burst over a grid around the initial localization. We then fit a model of the CHIME/FRB beam response to the intensity to improve the accuracy of the localization. For a burst with signal-to-noise ratio S/N, this technique yields a localization precision $\sigma_\theta$ at 600 MHz

estimated by:

$$\sigma_\theta \sim \frac{8}{S/N} \text{arcmin.} \qquad (2.2)$$

Therefore, an event with S/N = 100 can be localized with precision $\sigma_\theta \sim 5$ arcsec (see Michilli et al. (2021) for more detail).

## 2.1 Baseband Pipeline
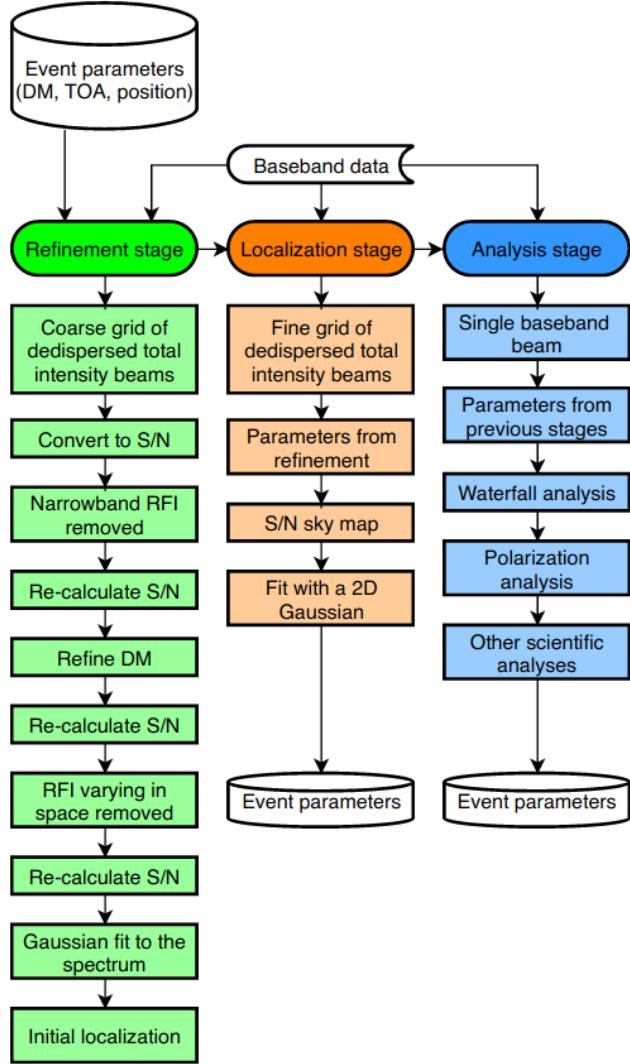


Figure 2.2: Schematic of the CHIME/FRB baseband pipeline. Figure from Michilli et al. (2021).

### 2.1.1 Refinement

The purpose of the refinement stage is to use the data to improve burst parameters provided by the real-time intensity pipeline (DM, TOA, and position). The first step of the refinement consists in forming a widespread grid of beams on the sky, since the source sky location is not known very precisely, especially in the E-W direction. The

grid consists of a small grid of 5 × 3 beams and a row of 19 beams on each side of the grid in the E-W direction as shown in Figure 2.3 for a bright pulse from the Crab pulsar B0531+21.
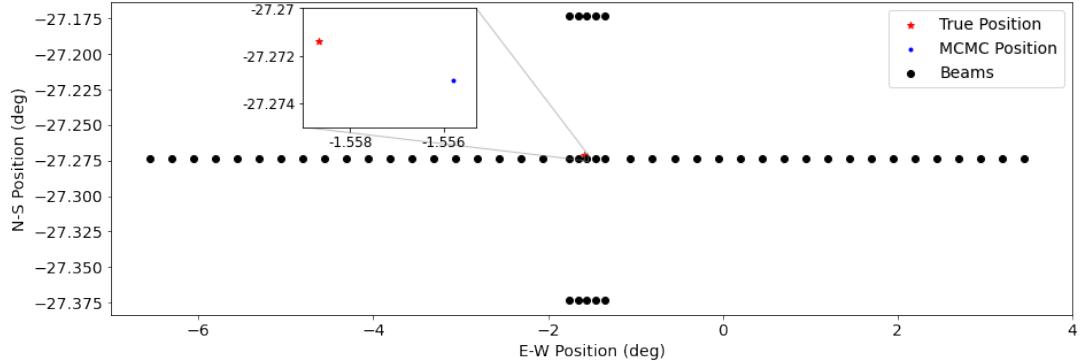


Figure 2.3: Formed beams at the refinement stage (black) for a bright pulse (S/N ∼ 360) from the Crab pulsar B0531+21 with its known position (red star) and MCMC localization (blue) (Astropy Collaboration et al., 2013, McNamara, 1971). The error bars on the MCMC fit are omitted.

Next, we may downsample the data by some factor, which, for a factor of 2, for example, would involve summing every two bins together and then normalizing the result by $\sqrt{2}$. We then clean the data from narrow-band and spatial RFI and cut the spectrum in frequency to only keep the frequency range which contains the burst signal.
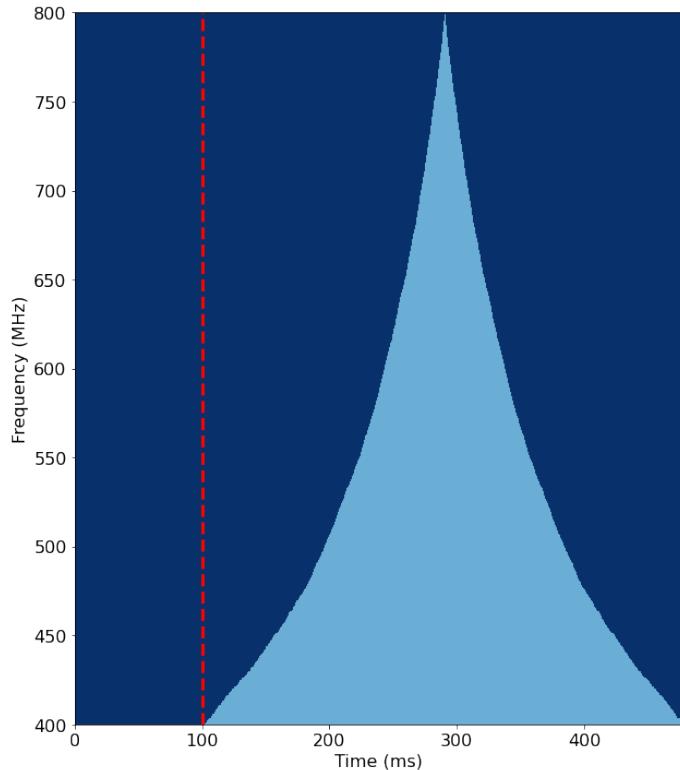


Figure 2.4: An illustration of the artefact occurring due to the uncertainty in DM. The artefact is the pale blue triangle. All around it is valid data. The data is usually cut off at the dashed red line at the base of the triangle. If the pulse signal happens to be on the right of the red line, then it is cut out and only noise is left.

17

Finally, we convert every beam to S/N by normalizing the data such that the noise floor has a median of 0 and a standard deviation of 1. Due to the uncertainty in the DM assigned by the real-time pipeline, the de-dispersion during the beamforming step requires adding empty bins in order to keep the data array rectangular. These empty bins must be zero-padded when taking an FFT. After the FFT, these bins end up with some negative value and thus create an artefact in the data that is shown in pale blue in Figure 2.4. This artefact is usually cut off (dashed red line in Figure 2.4). However, in certain cases when the data dump is too short, cutting off the artefact can also remove the burst signal. In Figure 2.4, this would occur when the data is on the right of the dashed red line.

To avoid accidentally removing the pulse, we first perform the normalization after cutting at the edge of the artefact as before, and then, we check if the pulse is present. If the pulse is present, no additional steps are taken. However, if no pulse is present or if the pulse is located at the very edge, then we identify the artefact and replace it with white noise, thus recovering the pulse.
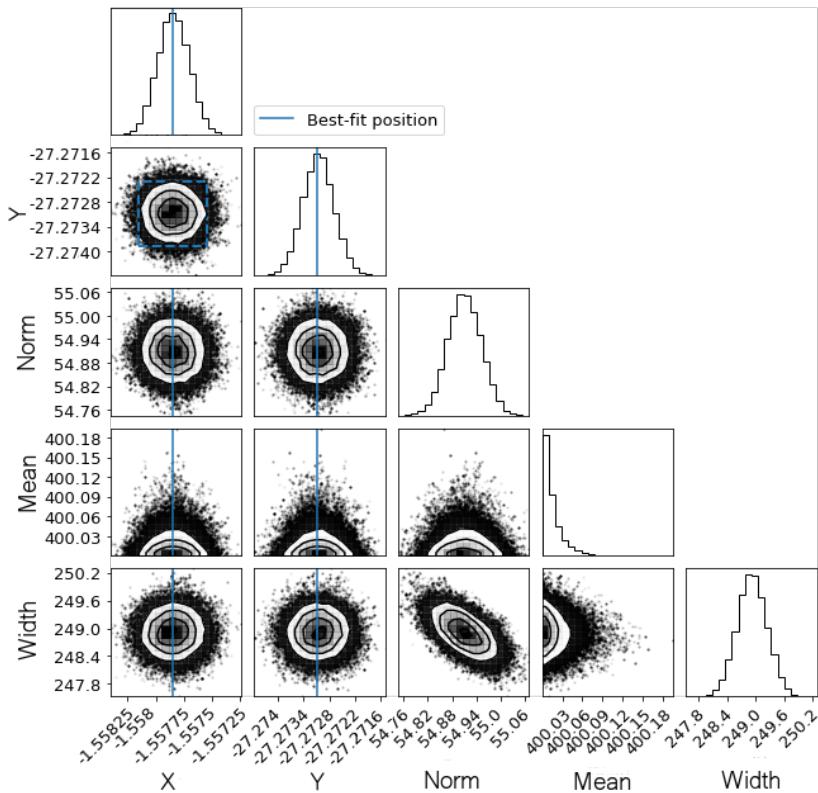


Figure 2.5: Contour plots for the MCMC position fitting of a Crab pulse with S/N $\sim$ 360. Best position indicated by blue lines. The offset from the true position is -0.029 and 0.0017 in X and Y position, respectively. The model fit to the spectrum is a Gaussian, so the other three parameters are the normalization constant, the mean (MHz), and the width of the Gaussian (MHz).

Once the burst signal has been cleaned of all RFI and artefacts, we use Markov Chain Monte Carlo (MCMC) sampling on the refinement beam configuration to improve the initial burst localization as input to the localization stage. Figure 2.5 shows the MCMC

contours of the localization of the same Crab burst as in Figure 2.3. The localization best fit, shown in blue, is found by fitting a Gaussian or power-law model to the burst spectrum using MCMC sampling. In this example, a Gaussian was fit to the spectrum, although both models would have been appropriate. This burst is brightest at low frequencies, hence a mean around 400 MHz. This burst is also broadband, hence a Gaussian width of 249 MHz. The MCMC error is not used in the subsequent steps and hence is omitted from the plot.

The MCMC relies on an instance of the CHIME/FRB beam model, described in Section 2.1 of The CHIME/FRB Collaboration et al. (2021a). The performance of this MCMC fitting can be improved mostly by better understanding the CHIME beam and improving its model. The MCMC sampling is performed using likelihood clustering following Hou et al. (2012). The configuration of the refinement beam pattern shown in Figure 2.3 is very important for the MCMC because it determines the quality of the MCMC result, especially for sidelobe events. A sidelobe event occurs when a burst is detected outside of the three main lobes of the CHIME/FRB formed beams. This often happens when the instrument detects a particularly bright burst long before or after the source transit. This may result in the telescope assigning it an offset E-W position.

If the refinement beam grid is too tight, there is not enough information for the MCMC to determine a better position. However, if the grid is too wide, there may not be enough beams to precisely determine the best localization. If more beams are created accordingly, then the beamforming phase will be excessively long. Currently, the beamforming takes roughly 3 hours for 30 beams. On the other hand, if a wide grid with too few beams is chosen, then the localization returned by the MCMC will not be very precise. To minimize the number of beams while maximising the amount of information they contain without spreading the beams too far apart, we decided on a composite pattern made of a small grid and two wings of beams in the E-W direction on each side of the grid, which help better localize sidelobe events. Once the MCMC runs on the refinement beam configuration, it returns an initial guess for the burst localization that is input into the localization stage.

### 2.1.2 Localization

In the localization stage, a tight square grid of $5 \times 5$ overlapping beams spanning roughly 0.3 deg$^2$ is created around the initial localization provided by the MCMC at the end of the refinement step. To find the best localization, the brightest time bin of the profile of every beam is selected, which yields a map of the burst brightness as a function of position. The beam is modelled with a simple 2-D Gaussian is then fit to this map and yields the best localization for the burst. After the best localization is found, a beam is formed at this best location thus maximizing the S/N of the source. This single beam is then saved and is used as input in many scientific pipelines such as the morphology pipeline that is the focus of this work and is described in the following chapter.

# 3 Fast Radio Burst Morphology

## 3.1 Pulse Morphology Pipeline

In this section, we will use a pulse from FRB 20180916 (The CHIME/FRB Collaboration et al., 2021a) as an example while we go over every step of the pulse morphology pipeline. A graph of the key steps of the morphology pipeline is shown in Figure 3.1. The analysis consists in three main steps. First, the data is loaded, cleaned of RFI, and the DM is corrected using the methods described in the refinement step of the baseband pipeline (see Section 2.1.1). The data are then used to estimate parameter initial conditions. Second, these initial conditions are fed into a non-linear least squares (LS) algorithm whose resulting fit is the initial condition for the MCMC fit. At the end of this step, we have a pulse profile fit of $N$ components. The last step is to perform a fit on the spectrum of every component. All of the fit parameters are then passed on to `fitburst`, an algorithm that fits the entire 2D waterfall via least-squares that is described in detail in The CHIME/FRB Collaboration et al. (2021a) (see Section 3.3 and Appendix B).
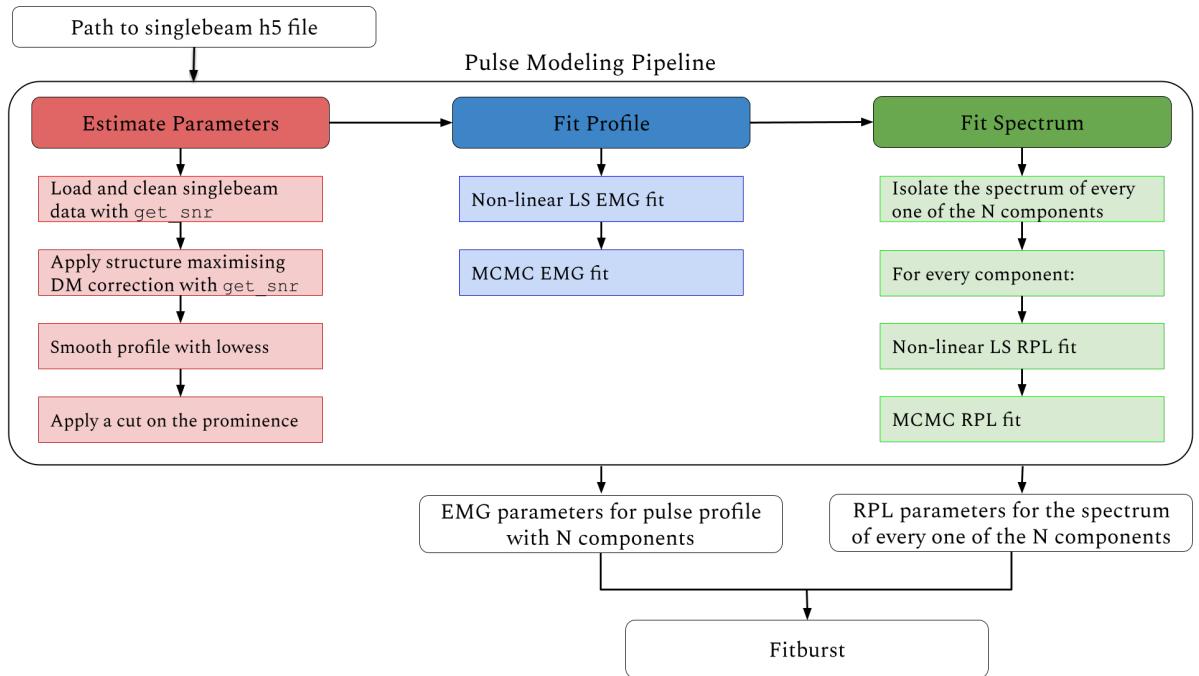


Figure 3.1: The three steps of the pulse modeling pipeline. The first step consists in reading the data, cleaning it, applying a DM correction, and estimating initial conditions from it. The second step uses those initial conditions to fit a sum of $N$ exponentially modified Gaussians (EMGs) to the pulse profile with $N$ components, first using least squares (LS) and then using Markov Chain Monte Carlo (MCMC). Next, we fit the spectrum of each of the $N$ components with a running power-law (RPL) model in the same way. Finally, the fit parameters are passed on to `fitburst` that performs a least squares fit on the entire 2D waterfall.

### 3.1.1 Parameter Estimation

The first step of the analysis is to load and clean the data using the methods described in the refinement stage of the baseband pipeline. The baseband pipeline applies a S/N optimizing DM correction. This step can be detrimental in certain cases because when the pulse morphology is complex, e.g. exhibits downward drifting structure, the S/N optimizing DM algorithm may perform poorly. Figure 3.2 shows the S/N as a function of DM. The left panel shows the waterfall when the DM is too low, as we see the entire signal stretch out in an arc. The middle panel shows the best DM for this pulse the signal of every component lines up. The third panel shows the S/N maximising DM that is higher than the correct DM because the downward drifting structure of the components allows to increase the overall S/N by stacking up the various components. To avoid this issue, we may skip the S/N optimizing DM correction. We next apply a structure maximising DM correction using `DM phase` (Seymour et al., 2019). `DM phase` maximizes the coherent power of the radio signal rather than its S/N to calculate the best DM, and is thus robust to the issue shown in Figure 3.2.
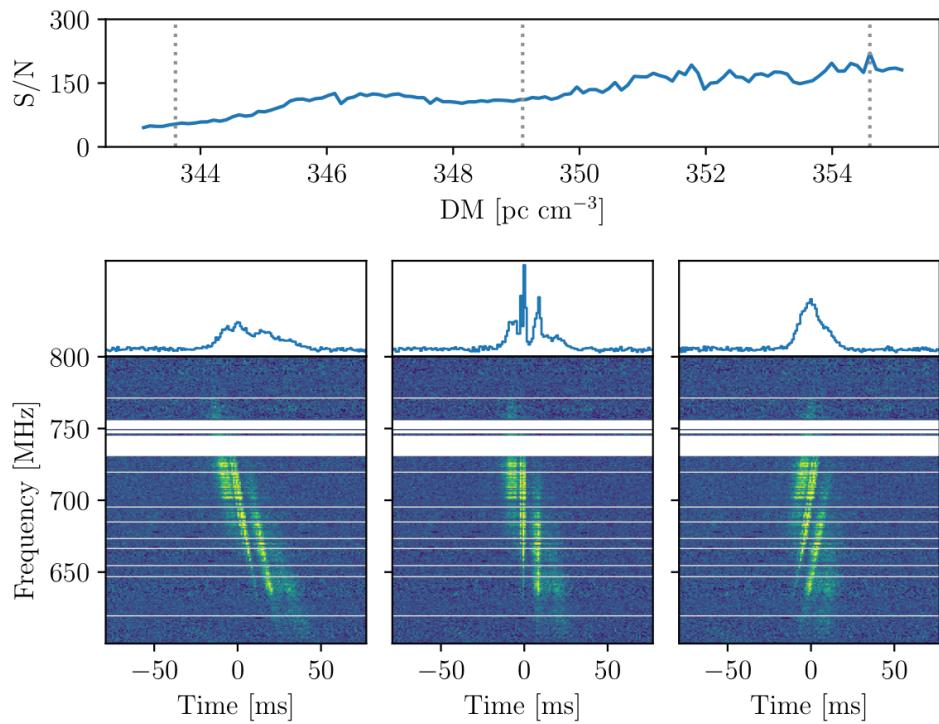


Figure 3.2: The top panel shows S/N as a function of DM. The three dotted lines correspond to the DM of each of the three plots below. The left plot shows the waterfall when the DM is underestimated. The middle plot shows the waterfall de-dispersed at the correct DM, where the signal in each of the components is aligned. The right plot shows the S/N maximising DM, which stacks the downward drifting components. This is a burst of FRB 20180916 detected by CHIME/FRB on December 22, 2018. Image courtesy of Ziggy Pleunis (Pleunis, 2021).
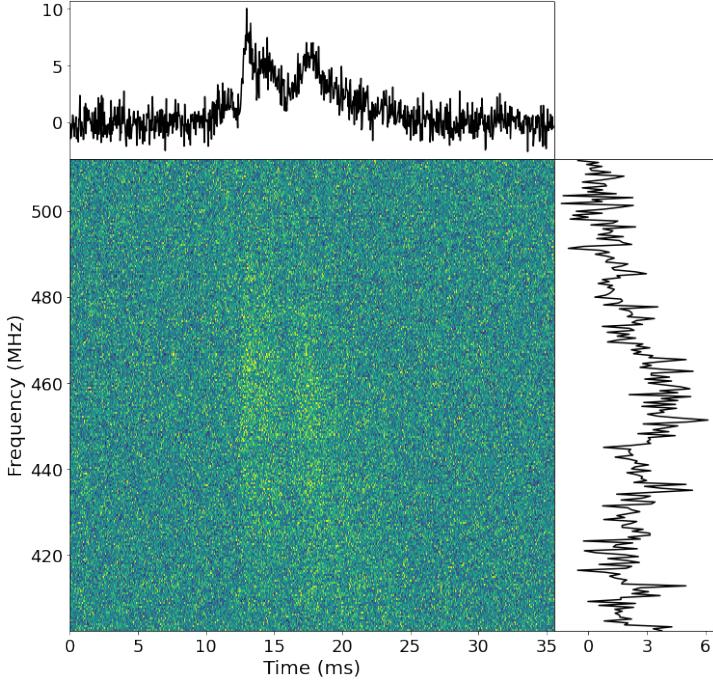
Figure 3.3: Spectrum after RFI cleaning and DM correction. This is a burst of FRB 20180916 detected by CHIME/FRB on December 25, 2018 (The CHIME/FRB Collaboration et al., 2021a).

This data will now be used for the rest of the analysis. The next step is to find the number of components in the pulse profile. To do so, we apply locally weighted scatterplot smoothing (lowess) implemented by `statsmodels` (Seabold & Perktold, 2010). Lowess is a non-parametric regression method which weights a neighbourhood of data with some kernel. Thus, the only parameter it requires is the neighbourhood size to use for the smoothing. A small neighbourhood will smooth very little, while a large neighbourhood will smooth a lot. The pipeline typically uses a neighbourhood size of 1.5% - 20% of the profile, depending on burst properties such as S/N, time resolution, and pulse width. Figure 3.4 shows the profile for the same burst as Figure 3.3 but at half the resolution for visual purposes. In Figure 3.4, the pulse profile is plotted in grey, and the cyan star around 5 ms indicates the data point we wish to smooth. The blue oversized dots around it indicate a neighbourhood size of 3%. The tricube kernel plotted above the data in blue weighs the points in the neighbourhood. A local weighted linear regression is then performed in order to calculate the smooth point marked by a red diamond. The red line is the resulting smooth pulse profile.
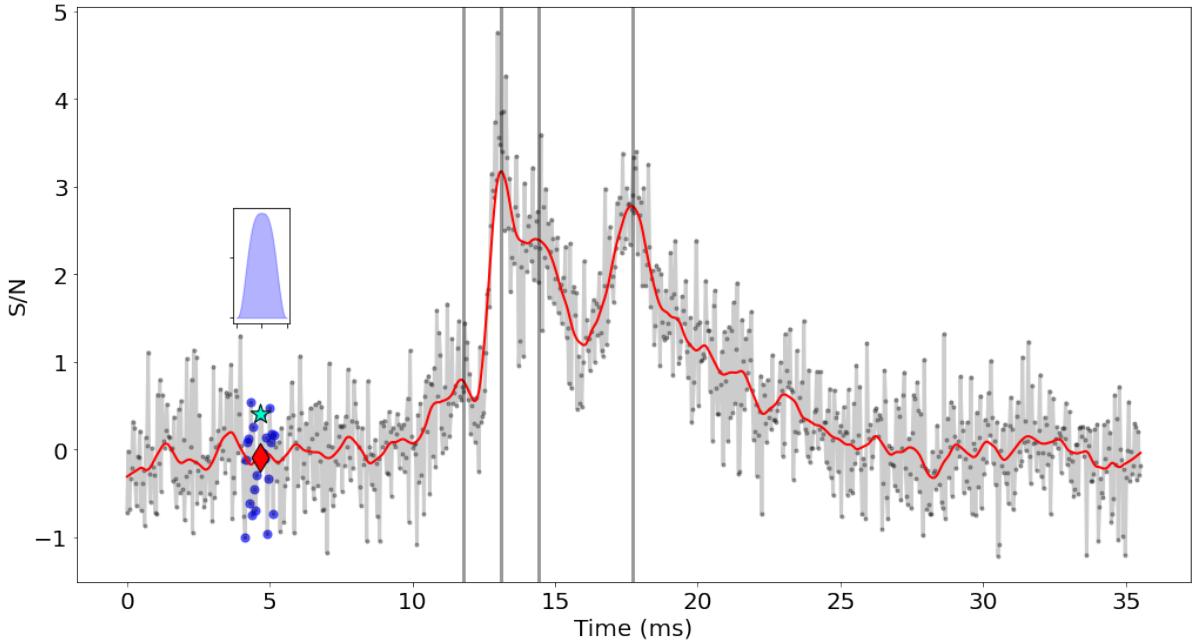
Figure 3.4: Locally weighted scatterplot smoothing (lowess) with a 3% smoothing fraction. The pulse profile is in grey. The cyan star marks the data point we wish to smooth. The oversized blue dots around it form a neighbourhood determined by the smoothing fraction. The data in the neighbourhood is weighted with a tricube kernel shown in blue above the pulse profile. Regression is then applied in order to obtain the resulting red diamond. The red line is the lowess smoothing for the entire pulse profile. The vertical lines mark peaks found using the smooth profile. This is a burst of FRB 20180916 detected by CHIME/FRB on December 25, 2018 (The CHIME/FRB Collaboration et al., 2021a).

The smooth profile is then used to find the pulse components using `find_peaks` implemented by `SciPy` (Virtanen et al., 2020) .[1] The pulse components are marked by vertical grey lines in Figure 3.4.

### 3.1.2 Fitting the Pulse Profile

From the previous step, we have the number of temporal components and their peak location. We now can fit a sum of exponentially modified Gaussians (EMGs) to the pulse profile, one EMG per burst component, with the scattering timescale set to be the same for all components. For a burst with $N$ components, the EMG model is:

$$\text{EMG}(x; A_i, \mu_i, \sigma_i, \tau) = \sum_{i=1}^{N} A_i \frac{1}{2\tau_i} \exp\left(\frac{1}{2\tau_i}(2\mu_i + \frac{\sigma_i^2}{\tau_i} - 2x)\right) \cdot \text{erfc}\left(\frac{\mu_i^2 + \frac{\sigma_i^2}{\tau_i} - x}{2\sqrt{\sigma_i}}\right),$$
(3.1)

where $A$ is the amplitude, $\mu$ is the Gaussian mean, $\sigma$ is the variance, and $\tau$ is the scattering timescale for burst components $i = 1, \ldots, N$. We first fit this model using a non-linear least squares algorithm `curve_fit` implemented by `SciPy`. The results from this fit are then used as the initial condition for the walkers of the Markov chain. The

---

[1]https://scipy.org/

MCMC sampling algorithm is implemented by `emcee`(Foreman-Mackey et al., 2013). We use wide uniform priors for all parameters. We use the square of the residuals as the log likelihood function:

$$\ln \mathcal{L}(x; y, \theta) = -\sum_{i=1}^{T}(y - \text{EMG}(x;\theta))^2, \qquad (3.2)$$

where $x$ is the independent variable, $y$ is the data, and $\theta$ is the set of EMG parameters.
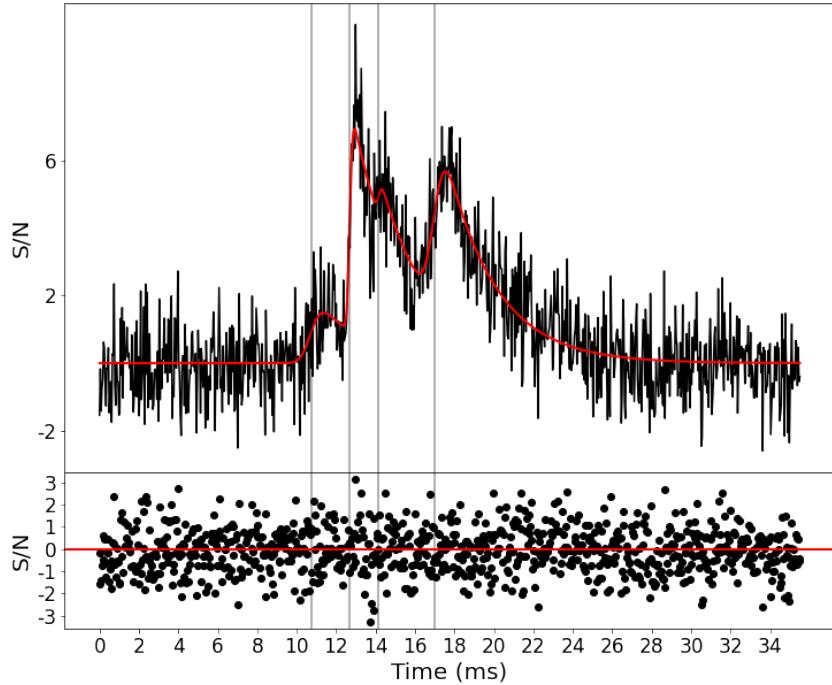


Figure 3.5: The top panel shows the data in black with a time resolution of $4.096 \cdot 10^{-2}$ ms and the model fit in red. The model is a sum of four exponentially modified Gaussians fit to the pulse profile with MCMC sampling. The vertical lines correspond to the parameter $\mu$ of each component as defined in Eq. 3.1. The bottom panel shows the residuals. The reduced chi squared of this fit is 1.16. This is the same data as in Figure 3.4.

Figure 3.5 shows the best fit pulse profile in red against the data in black with the residuals in the bottom panel. The vertical lines correspond to the parameter $\mu$ of each component as defined in Eq. 3.1. Note that it is exactly the same data as in Figure 3.4. The quantile-quantile (Q-Q) (Wilk & Gnanadesikan, 1968) plot of the residuals is shown in Figure 3.7a. The red $45^o$ line is a normal distribution. The residuals shown in blue follow it closely and are therefore normal as well. The reduced chi squared of this fit is 1.16. We can therefore conclude that that this model describes the data extremely well. We next split the waterfall component-wise at the minima between the peaks, obtain the spectrum for every component, and proceed to fit each one.

### 3.1.3 Fitting the Spectrum

We fit a running power-law (RPL) model to the 1-D spectrum of every $i$th component of a burst with $N$ components:

$$\mathrm{RPL}_i(f; A_i, r_i, \gamma_i) = A_i(f/f_0)^{-\gamma_i + r_i \ln(f/f_0)}, \tag{3.3}$$

where $f_0$ is an arbitrary reference frequency, usually chosen to be the middle of the band at 600 MHz, $r$ is the spectral running and $\gamma$ is the spectral index. The RPL model is appropriate for the spectrum because the spectrum can be very versatile. Broadly speaking, spectra can be flat, Gaussian, or follow a power-law. The RPL model can fit all of these case scenarios: a flat spectrum would have $\gamma, r = 0$, a power-law-like spectrum would have $\gamma, r > 0$, and a Gaussian spectrum would have $r < 0$. Like the pulse profile, the spectrum also has complex structure, but it is often due to propagation effects such as scintillation, which are beyond the scope of this work.
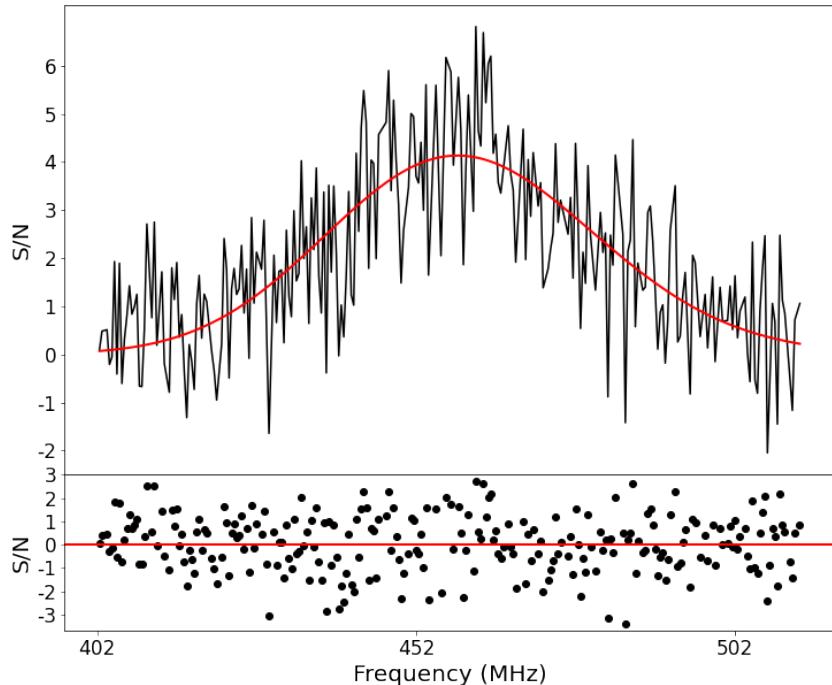


Figure 3.6: The top panel shows the spectrum of the brightest component, i.e. the second component, in black and the model fit in red. The model is a running power-law fit with MCMC sampling. The bottom panel shows the residuals. The reduced chi squared of this fit is 2.107. This is the same data as in Figure 3.3.

Like in the previous step, we first fit the model with a non-linear least squares, and then use the result as initial condition for the MCMC fit. The result of the MCMC fit for the brightest component of the burst is shown in Figure 3.6 in red against the spectrum in black. The residuals of this fit are shown in the bottom panel. The reduced chi squared of this fit is 2.107.

(a) Pulse profile fit residual Q-Q plot.       (b) Pulse spectrum fit residual Q-Q plot.
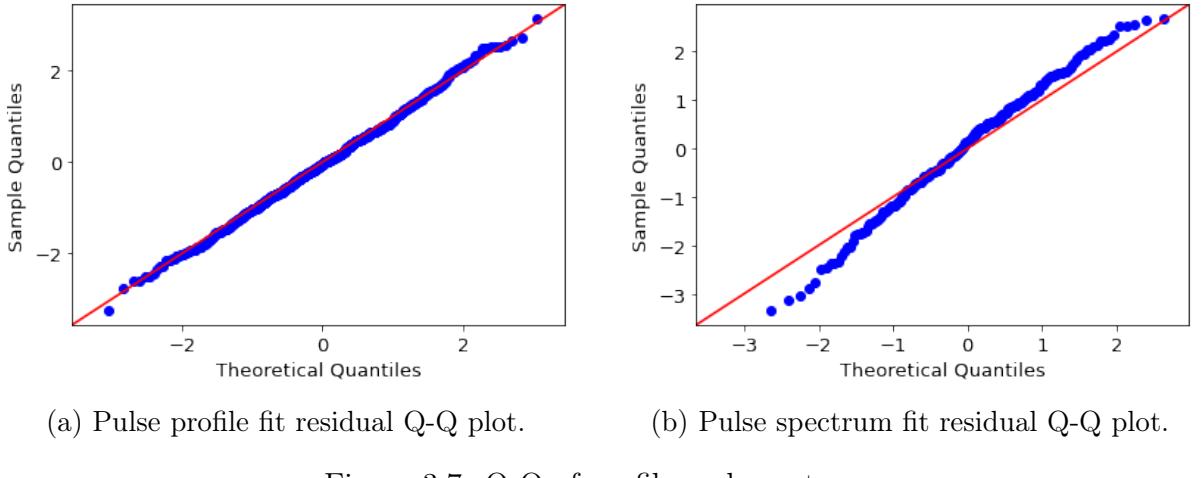
Figure 3.7: Q-Q of profile and spectrum

The Q-Q (Wilk & Gnanadesikan, 1968) plot of the residuals of the spectrum fit is shown in blue in Figure 3.7b, which confirms that the RPL model describes the data reasonably well, since it exhibits only a slight offset from the red line.

Finally, the fit parameters for the profile and the spectrum are passed on to `fitburst` that performs a 2D fit of the waterfall. A detailed description of `fitburst` can be found in Section 3.3 and Appendix B of The CHIME/FRB Collaboration et al. (2021a). In short, `fitburst` uses least-squares optimization implemented by `optimize.least_squares` provided the `SciPy` (Virtanen et al., 2020) software library. The least-squares optimization is performed using the noise-weighted fit residuals, which are defined as follows:

$$h_{t,f}(\theta) = \frac{y_{t,f}}{\sigma_f} - S_{t,f}(\theta), \tag{3.4}$$

where $y_{t,f}$ is the total intensity data as a function of discrete time $t$ and discrete frequency $f$ (the dynamic spectrum / waterfall), $S_{t,f}$ is the model defined as a function of the input parameters $\theta$, and $\sigma_f$ is the noise standard deviation measured in each spectral channel. The least-squares optimization thus minimizes $\chi^2(\theta) = \sum_{t,f}[h_{t,f}(\theta)]^2$ with respect to the fit parameters.
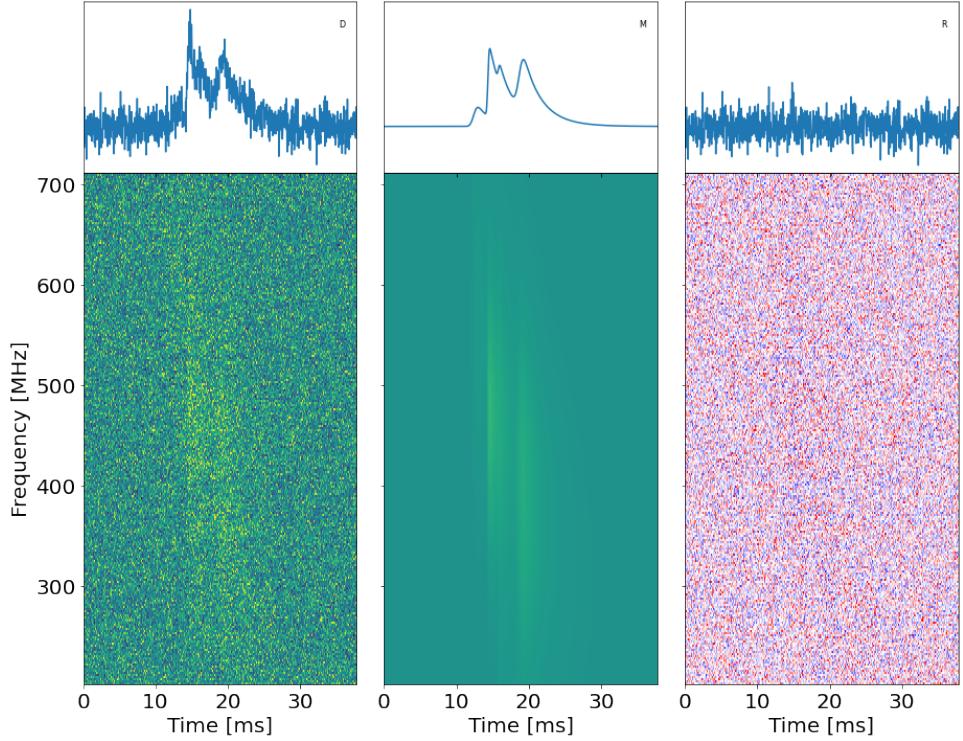
Figure 3.8: Two dimensional model fit to the FRB 20180916 pulse using `fitburst` where the initial conditions are given using all the one-dimensional fits from the pulse morphology pipeline. The left panel shows the data. The middle panel shows the model. The right panel shows the residuals. This is the same data as in Figure 3.3.

Figure 3.8 shows the `fitburst` two-dimensional fit to the data. In order to obtain such a complex fit, `fitburst` requires a very good initial condition. The initial condition for the `fitburst` parameters for this fit comes from the one-dimensional pulse profile and spectrum fits (Figures 3.5 and 3.6)from the pulse morphology pipeline. The fit is successful since there are no outstanding residuals in the right plot.

The next step is to apply this powerful tool to various data sets.

# 4 Results & Discussion

## 4.1 Sub-Second Periodicity

The first application of the pulse morphology pipeline has been on a sample of FRBs with extremely complex structure. The profile model has been used to test for periodicity in the components of those FRBs. The profile fit for FRB 20191221A, the longest ($\sim 3$ s) and most complex FRB of the sample, is shown in Figure 4.1.
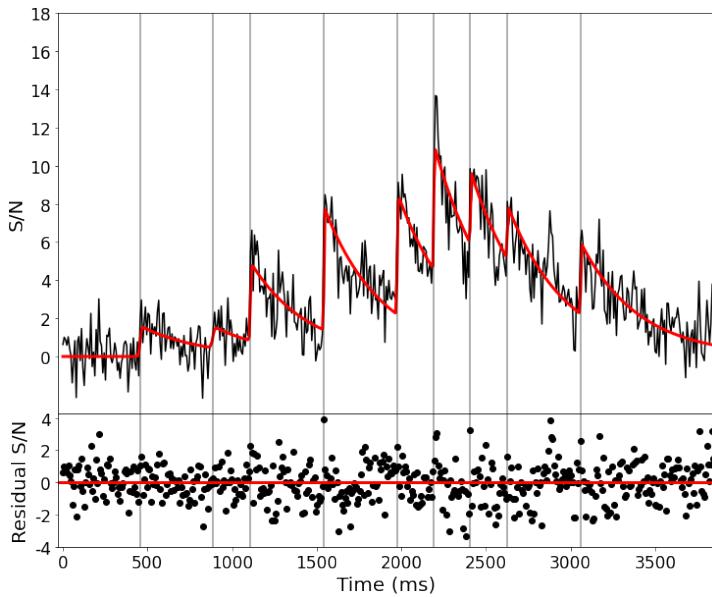


Figure 4.1: Modeling of the pulse profile of FRB 20191221A. The measured pulse profile is plotted in black with a time resolution of 7.86432 ms for visual purposes (the full resolution data at 0.98304 ms have been used in the fitting process). The exponentially modified Gaussian function (see Eq. 3.1) used to model nine peaks in the profile is plotted in red, and peak locations are highlighted by vertical lines. Fit residuals are shown in the bottom panel with a red horizontal line at zero residual.

The arrival times (vertical lines in Figure 4.1) have been used to identify a periodic separation of 216.8(1) ms between its components (The CHIME/FRB Collaboration et al., 2021b).

The long ($\sim 3$ s) duration and nine or more components forming the pulse profile make this source an outlier in the FRB population. The multi-component structures of the other two FRBs in the sample, 20210206A and 20210213A, show some indication of periodic separations of 2.8(1) and 10.7(1) ms, respectively, suggesting the possible existence of a group of FRBs with complex and periodic pulse profiles. Such short periodicities provide strong evidence for a neutron-star origin of these events. One hypothesis is that such periodic bursts are due to magnetic reconnection. A magnetic field is induced by the two inspiralling neutron stars. As the neutron stars approach each other, the induced

field builds up against the intrinsic magnetic field of each of the stars, which results in magnetic reconnection (see Section 1.4.1 and Platts et al. (2019) for a review of FRB emission mechanisms). If the signal is emitted during many orbital periods of the binary system, then it may yield a periodic multi-peaked signal like the ones we have observed. Moreover, these detections favour emission arising from the neutron-star magnetosphere (see near-field emission in Section 1.4.1),(Lu et al., 2020, Popov & Postnov, 2010) as opposed to emission regions located further away from the star (see far-away emission in Section 1.4.1), as predicted by some models (Metzger et al., 2019).

## 4.2   Bursts from Repeating FRB 20180916B

FRB 20180916B is one of the most prolific repeaters known. 38 of its bursts were used to demonstrate that it had a periodicity of 16.35 ±0.15 days (CHIME/FRB Collaboration et al., 2020b). In the following section, we consider a sample of 35 bursts from this source. We apply the pulse morphology pipeline and run `fitburst` on the sample. Note that in the previous chapter (Chapter 3), all of the plots and examples come from a burst of this source.

After performing the profile and spectrum fits on the entire sample of 35 bursts exactly as described in Chapter 3, we look at the time evolution and distributions of S/N, burst width, scattering timescale, spectral running and index. All of the results shown below are from `fitburst` after providing it with the fit results of the pulse morphology pipeline as described in the previous chapter.

Out of the sample of 35 events, the S/N maximizing DM correction step was skipped for eight events that were either too faint or had a complex morphology such that the S/N maximizing DM correction introduced the issue shown in Figure 3.2. For seven out of these eight, the structure maximizing DM correction was applied. Five of these seven used an initial DM guess of 348.772 pc / cm$^3$ (Nimmo et al., 2021) before this correction. The other two used a manual DM correction. The remaining one was so faint that only a manual DM could be applied, without the structure maximizing DM correction step. For these seven of these eight FRBs, the final DM used was higher than 348.772 pc / cm$^3$ from Nimmo et al. (2021), and looked just like Figure 7 in CHIME/FRB Collaboration et al. (2020b) where they study a sample of 17 bursts from this same FRB source.
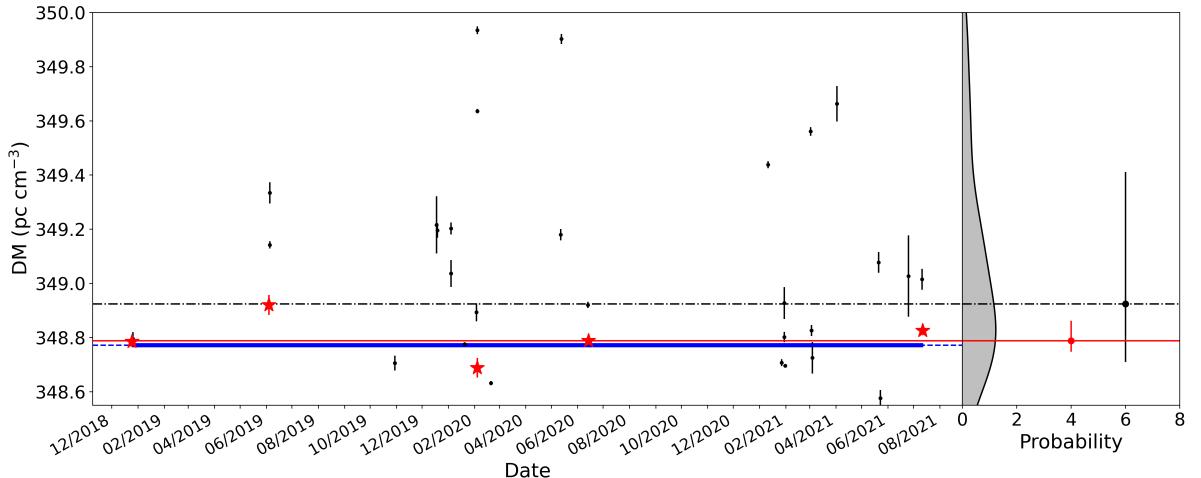
Figure 4.2: The left plot shows the evolution of DM as a function of time. The dashed blue line with the thin blue shaded area represents a DM of $348.772 \pm 0.006$ pc cm$^{-3}$ from Nimmo et al. (2021). The five brightest bursts are marked with red stars. The right plot shows the kernel density estimate of the entire sample. The dots represent the median of each sample, and the vertical lines are the 68% confidence intervals. The median of the red sample is at a DM of $348.79^{+0.07}_{-0.04}$ pc cm$^{-3}$, and that of the entire sample in black is of $348.9^{+0.5}_{-0.2}$ pc cm$^{-3}$.

Figure 4.2 shows the DM of FRB 20180916B bursts as a function of time. Based on this sample, there does not appear to be any DM evolution. The FRBs with S/N > 10 and time resolution under 0.08192 ms constitute the subsample of the five brightest FRBs in the sample. They will be marked in red in all plots in this section, while the black markers and dot-dashed lines pertain to the entire sample. In Figure 4.2, the subsample of bright bursts is shown to have a median DM of $348.79^{+0.07}_{-0.04}$ pc cm$^{-3}$ with 68 % confidence interval (CI).

Although both the entire population (in black) and the subsample of bright FRBs (in red) agree at $1\sigma$ with the value measured by Nimmo et al. (2021) of $348.772 \pm 0.006$ pc cm$^{-3}$ [1], shown in blue, the median value of the red sample is closer to their value.

This is expected because the DM of fainter and therefore more downsampled bursts is biased high due to the lower time resolution.

---

[1] Note that this value is from a S/N maximizing DM correction applied on the brightest burst in their sample. After applying a structure maximizing DM correction, they obtain a DM of 348.775 pc cm$^{-3}$, which is also consistent with our findings.
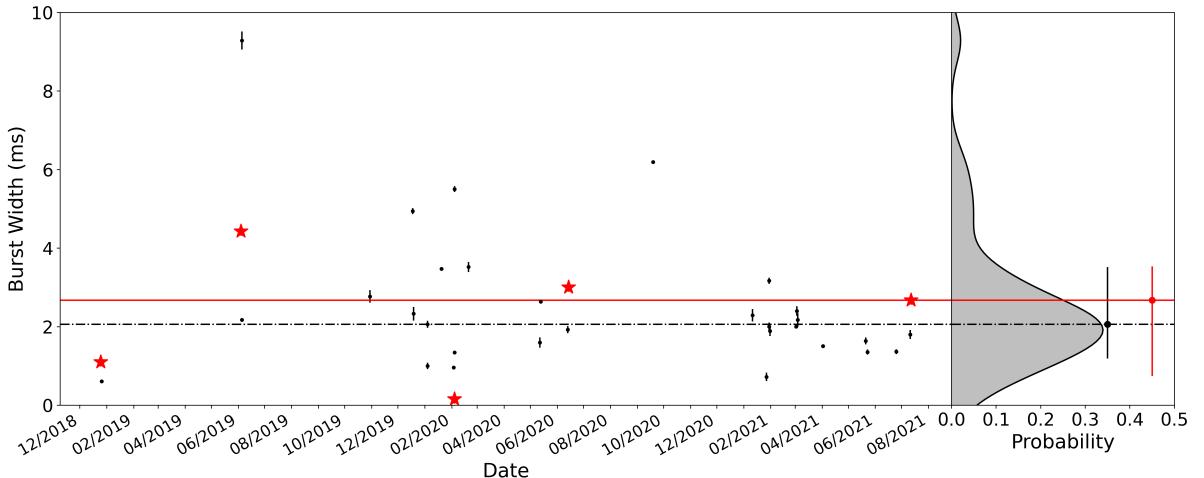
Figure 4.3: The left plot shows burst width as a function of time. The red stars mark the five brightest bursts in the sample. The right plot shows the kernel density estimate of each sample. The median of the entire sample, marked by the dot-dashed black line, is $2.1^{+1.4}_{-0.9}$ ms, and the median of the red sample, marked by the full red line, is of $2.7^{+0.8}_{-2.0}$ ms, both with 68% CI.

Figure 4.3 shows the width of the burst over time. If a burst has multiple components, we report the combined width as the sum of the width of each component. We also adjust the errors accordingly. Looking at the right plot of Figure 4.3, we see that the burst widths are all fairly consistent over time with no obvious pattern. The sample median is of $2.1^{+1.5}_{-0.9}$ ms, which agrees with other high-time-resolution measurements such as Nimmo et al. (2021) and Marcote et al. (2020). These widths are consistent with the bursts being produced near the source (i.e. see near-field emission mechanism in Section 1.4.1).
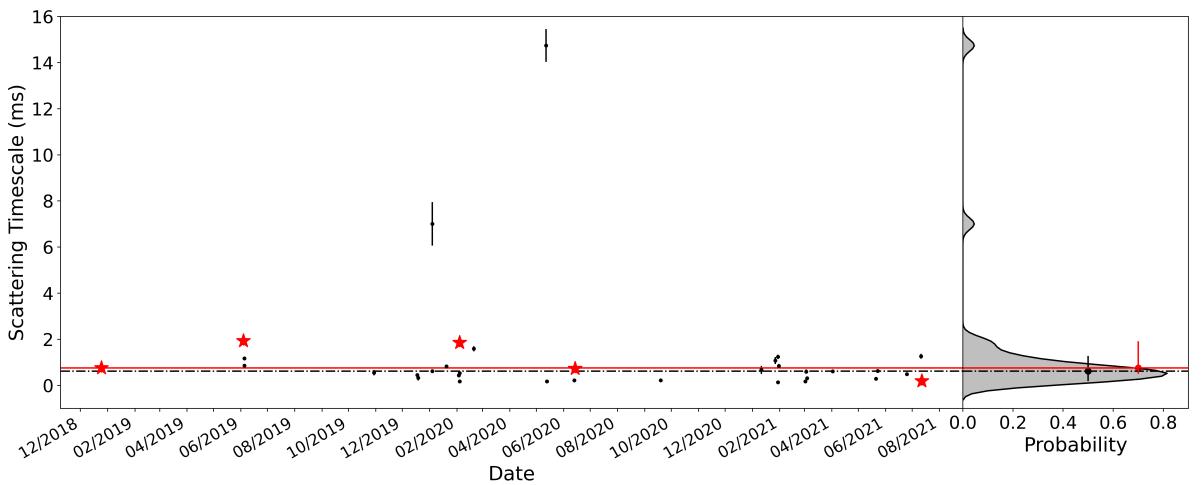


Figure 4.4: The left plot shows the scattering timescale as a function of time. The red stars mark the five brightest bursts. The right plot shows the kernel density estimate of each sample. The median of the entire sample, marked by the dot-dashed black line, is $0.0006^{+0.0006}_{-0.0004}$ s, and the median of the red sample, marked by the full red line, is of $0.0008^{+0.0011}_{-0.0002}$ s, both with 68% CI.

Figure 4.4 shows the scattering timescale over time. In the left plot, we see that there

is very little variation in scattering over time.

Next, we look at the correlations between some of the fit parameters.
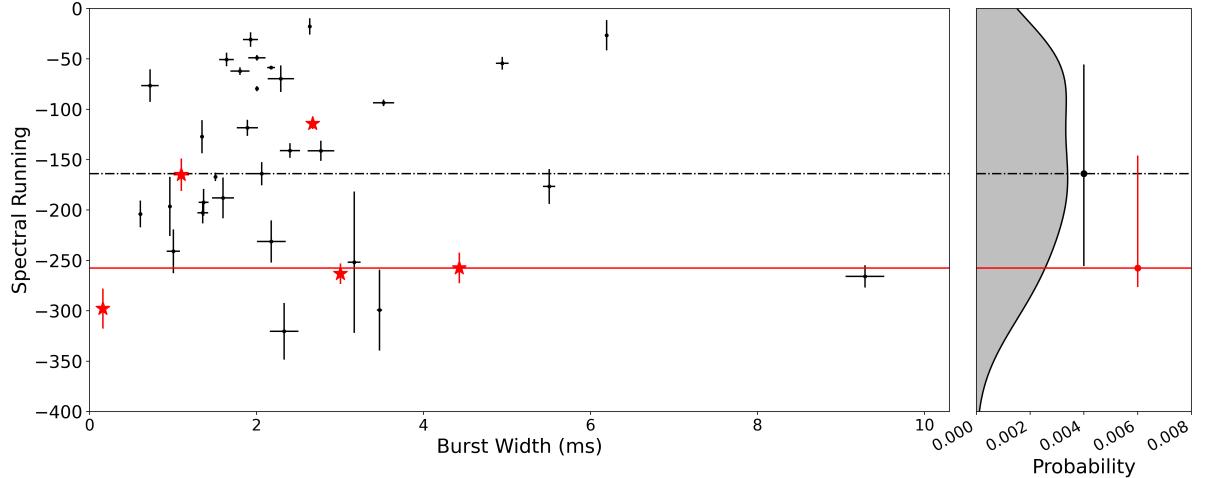


Figure 4.5: Spectral running ($r$) as a function of burst width in ms. The red stars mark the five brightest bursts. The right plot shows the kernel density estimate of each sample.

Figure 4.5 shows the spectral running, which reflects the bandwidth of the burst, as a function of the burst width. Like Pleunis et al. (2021) (see Figure 4), we observe that the vast majority of the repeater bursts occupy the upper triangular half of the plot. We observe a slight tendency for narrower bursts to have a higher spectral running, and thus this sample is characteristic of a repeating FRB. Overall, our observations agree with Pleunis et al. (2021).
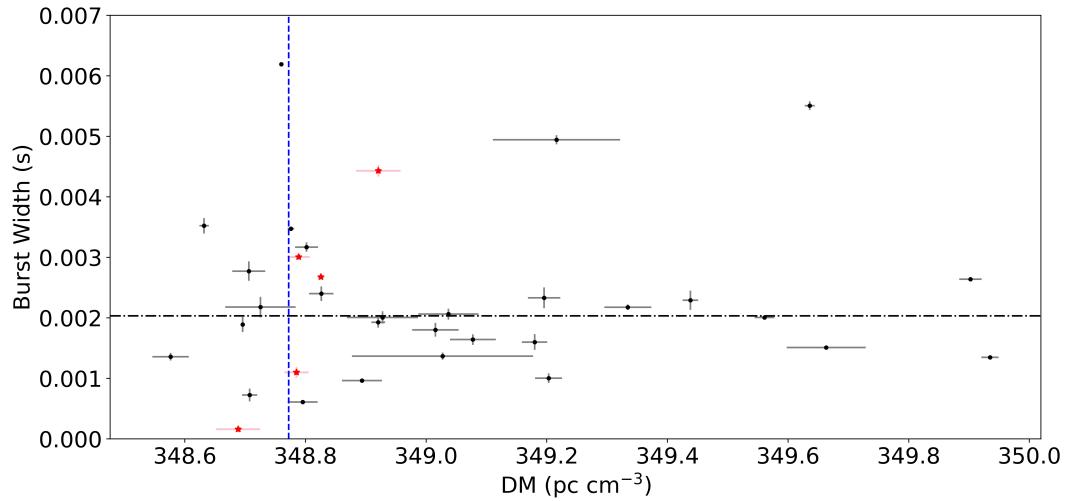


Figure 4.6: Burst width as a function of DM. The red stars mark the five brightest bursts. The vertical dashed blue line is at DM of 348.772 pm cm$^{-3}$. The dot-dashed horizontal black line is at the median burst width excluding an outlier with the highest burst width (see Figure 4.3, a burst width of almost 10 ms around July 2019).

In Figure 4.6, we show the burst width as a function of DM. The subsample of red stars (the same one as in the previous few plots) shows a faint trend that we tend to fit a

higher DM for wider bursts. This is caused by the same issue as explained in Chapter 3 in Figure 3.2. We tend to fit a higher DM to bursts with downward drifting morphology, which is especially a problem when the burst is very faint so that structure cannot be maximized successfully in the DM correction step and individual components can be difficult to distinguish.

## 4.3   Future Work

In the coming months, we hope to apply the methods and analysis described in the previous section to a large population of over 150 repeating and non-repeating bursts. The goal will be to study individual repeating FRBs as shown in the previous section, but also compare the properties of repeating and non-repeating FRBs to extract conclusions on whether they may or may not belong to a single population in a similar manner that Pleunis et al. (2021) have done but with much higher time resolution. Moreover, we also aim to make further conclusions about the FRB emission mechanism of these populations.

# 5    Conclusion

Fast radio burst morphology is an unexplored avenue that shows to be very promising for constraining the FRB emission mechanism. More specifically, studying the FRB component burst width at high time resolution is key to inferring the size of the burst emission region at the source and hence to understanding how the FRB signal is produced. The CHIME/FRB project has collected what is most likely the largest sample of high-time-resolution FRB data, which makes it an ideal survey to perform morphological analyses. In this work, we present a pipeline to process high-time-resolution data and a morphology pipeline to fit the resulting data with a complex model. The morphology pipeline uses MCMC sampling to fit a sum of exponentially modified Gaussians to the pulse profile and a running power-law to the pulse spectrum. The results from these 1-D fits are then input into `fitburst`, a 2-D least squares fitting algorithm. We then apply the pipeline on a sample of fast radio bursts. We use it to demonstrate the periodicity of the components of a sample of FRBs, one of which has nine or more components. We also use it on a sample of high-time-resolution bursts from the repeating FRB 20180916B. We look at the time evolution of its properties and discuss the correlations between various parameters. Our results agree with previous studies suggesting that the emission of this FRB occurs most likely close to the source (i.e. near-field). This work is the very first time that anyone has studied the morphology of this FRB using a sample of this size at such high time resolution.

# Bibliography

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: 10.1051/0004-6361/201322068

Bannister, K. W., Shannon, R. M., Macquart, J.-P., et al. 2017, ApJ, 841, L12. https://arxiv.org/abs/1705.07581

Brentjens, M. A., & de Bruyn, A. G. 2005, A&A, 441, 1217

Burke-Spolaor, S., & Bannister, K. W. 2014, ApJ, 792, 19, doi: 10.1088/0004-637X/792/1/19

Caleb, M., Flynn, C., Bailes, M., et al. 2017, MNRAS, 468, 3746. https://arxiv.org/abs/1703.10173

Cassanelli, T., Leung, C., Rahman, M., et al. 2021, arXiv e-prints, arXiv:2107.05659. https://arxiv.org/abs/2107.05659

Champion, D. J., Petroff, E., Kramer, M., et al. 2016, MNRAS, 460, L30, doi: 10.1093/mnrasl/slw069

Chawla, P., Kaspi, V. M., Ransom, S. M., et al. 2021, arXiv e-prints, arXiv:2107.10858. https://arxiv.org/abs/2107.10858

CHIME/FRB Collaboration, Amiri, M., Bandura, K., et al. 2018, ApJ, 863, 48, doi: 10.3847/1538-4357/aad188

CHIME/FRB Collaboration, Amiri, M., Andersen, B. C., et al. 2020a, Nature, 582, 351, doi: 10.1038/s41586-020-2398-2

CHIME/FRB Collaboration, Andersen, B. C., Bandura, K. M., et al. 2020b, Nature, 587, 54, doi: 10.1038/s41586-020-2863-y

Condon, J. J., & Ransom, S. M. 2016, Essential Radio Astronomy (Princeton University Press)

Connor, L., Pen, U.-L., & Oppermann, N. 2016, MNRAS, 458, L89, doi: 10.1093/mnrasl/slw026

Cordes, J. M., & Lazio, T. J. W. 2002, arXiv preprint astro-ph/0207156

Dolag, K., Gaensler, B. M., Beck, A. M., & Beck, M. C. 2015, MNRAS, 451, 4277, doi: 10.1093/mnras/stv1190

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

Gourdji, K., Michilli, D., Spitler, L. G., et al. 2019, ApJ, 877, L19, doi: 10.3847/2041-8213/ab1f8a

Hou, F., Goodman, J., Hogg, D. W., Weare, J., & Schwab, C. 2012, ApJ, 745, 198, doi: 10.1088/0004-637X/745/2/198

Kaspi, V. M., & Beloborodov, A. M. 2017, Annual Review of Astronomy and Astrophysics, 55, 261

Kaspi, V. M., & Kramer, M. 2016, arXiv e-prints, arXiv:1602.07738. https://arxiv.org/abs/1602.07738

Katz, J. I. 2016, Modern Physics Letters A, 31, 1630013, doi: 10.1142/S0217732316300135

Keane, E. F. 2018, Nature Astronomy, 2, 865, doi: 10.1038/s41550-018-0603-0

Keating, L. C., & Pen, U.-L. 2020, MNRAS, 496, L106, doi: 10.1093/mnrasl/slaa095

Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, Science, 318, 777. https://arxiv.org/abs/0709.4301

Lu, W., & Kumar, P. 2018, Monthly Notices of the Royal Astronomical Society, 477, 2470

Lu, W., Kumar, P., & Zhang, B. 2020, MNRAS, 498, 1397, doi: 10.1093/mnras/staa2450

Lyubarsky, Y. 2014, MNRAS, 442, L9, doi: 10.1093/mnrasl/slu046

Macquart, J. P., Prochaska, J. X., McQuinn, M., et al. 2020, Nature, 581, 391, doi: 10.1038/s41586-020-2300-2

Manchester, R. N., Hobbs, G. B., Teoh, A., & Hobbs, M. 2005, AJ, 129, 1993, doi: 10.1086/428488

Marcote, B., Nimmo, K., Hessels, J. W. T., et al. 2020, Nature, 577, 190, doi: 10.1038/s41586-019-1866-z

Margalit, B., Beniamini, P., Sridhar, N., & Metzger, B. D. 2020, ApJ, 899, L27, doi: 10.3847/2041-8213/abac57

Masui, K. W., Shaw, J. R., Ng, C., et al. 2019, ApJ, 879, 16, doi: 10.3847/1538-4357/ab229e

McNamara, B. J. 1971, PASP, 83, 491, doi: 10.1086/129160

Metzger, B. D., Margalit, B., & Sironi, L. 2019, MNRAS, 485, 4091, doi: 10.1093/mnras/stz700

Michilli, D., Masui, K. W., Mckinven, R., et al. 2021, ApJ, 910, 147, doi: 10.3847/1538-4357/abe626

Ng, C., Vanderlinde, K., Paradise, A., et al. 2017, arXiv e-prints, arXiv:1702.04728. https://arxiv.org/abs/1702.04728

Nimmo, K., Hessels, J. W. T., Keimpema, A., et al. 2021, Nature Astronomy, 5, 594, doi: 10.1038/s41550-021-01321-3

Petroff, E., Hessels, J. W. T., & Lorimer, D. R. 2019, A&A Rev., 27, 4, doi: 10.1007/s00159-019-0116-6

Platts, E., Weltman, A., Walters, A., et al. 2019, Phys. Rep., 821, 1, doi: 10.1016/j.physrep.2019.06.003

Pleunis, Z. 2021, Fast radio burst detection and morphology with the CHIME telescope (PhD Thesis). https://escholarship.mcgill.ca/concern/theses/f4752n54q

Pleunis, Z., Good, D. C., Kaspi, V. M., et al. 2021, arXiv e-prints, arXiv:2106.04356. https://arxiv.org/abs/2106.04356

Popov, S. B., & Postnov, K. A. 2010, in Evolution of Cosmic Objects through their Physical Activity, ed. H. A. Harutyunian, A. M. Mickaelian, & Y. Terzian, 129–132. https://arxiv.org/abs/0710.2006

Prochaska, J. X., & Zheng, Y. 2019, MNRAS, 485, 648, doi: 10.1093/mnras/stz261

Rane, A., & Lorimer, D. 2017, Journal of Astrophysics and Astronomy, 38, 55

Seabold, S., & Perktold, J. 2010, in 9th Python in Science Conference

Seymour, A., Michilli, D., & Pleunis, Z. 2019, DM_phase: Algorithm for correcting dispersion of radio signals. http://ascl.net/1910.004

Simard, D., & Pen, U.-L. 2018, MNRAS, 478, 983, doi: 10.1093/mnras/sty1140

Spitler, L. G., Scholz, P., Hessels, J. W. T., et al. 2016, Nature, 531, 202, doi: 10.1038/nature17168

Tegmark, M., & Zaldarriaga, M. 2009, Phys. Rev. D, 79, 083530, doi: 10.1103/PhysRevD.79.083530

The CHIME/FRB Collaboration, :, Amiri, M., et al. 2021a, ApJS, arXiv:2106.04352. https://arxiv.org/abs/2106.04352

The CHIME/FRB Collaboration, Andersen, B. C., Bandura, K., et al. 2021b, arXiv e-prints, arXiv:2107.08463. https://arxiv.org/abs/2107.08463

Thornton, D., Stappers, B., Bailes, M., et al. 2013, Science, 341, 53. https://arxiv.org/abs/1307.1628

Tohuvavohu, A. 2020, The Astronomer's Telegram, 14075, 1

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: 10.1038/s41592-019-0686-2

Wilk, M. B., & Gnanadesikan, R. 1968, Biometrika, 55, 1. http://www.jstor.org/stable/2334448

Xu, J., & Han, J. L. 2015, Research in Astronomy and Astrophysics, 15, 1629, doi: 10.1088/1674-4527/15/10/002

Yao, J. M., Manchester, R. N., & Wang, N. 2017, ApJ, 835, 29, doi: 10.3847/1538-4357/835/1/29