A Synthetic Pulse Injection System for the CHIME/FRB Experiment

Marcus Merryfield

Master of Science

Physics

McGill University

Montréal, Québec

2020-12-15

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

ABSTRACT

The Canadian Hydrogen Intensity Mapping Experiment Fast Radio Burst project (CHIME/FRB) has begun detecting FRBs at an unprecedented rate. This allows for the first time the study of FRB properties in a large, coherent population. However, the CHIME/FRB detection pipeline is subject to many subtle selection effects. Thus, the detection sample from CHIME/FRB is not representative of the true FRB population. In order to correct for the biases introduced during CHIME/FRB event detection, a synthetic pulse injection system was developed which allows for the injection of a large population of simulated FRBs into the live telescope datastream. By injecting pulses drawn from a realistic FRB population, the detection signal-to-noise ratio (SNR) of synthetic pulses could be compared across pulse input parameters. Injected pulses were calibrated to physical energy units (Jy ms) in real time, and the pulse position in the telescope field-of-view was simulated, providing an authentic representation of detecting real FRBs on the sky. The final set of injections and corresponding detections will be reweighted such that the output distribution matches what has actually been observed by CHIME/FRB. This will begin to correct for the telescope's selection effects.

RÉSUMÉ

Le projet visant l'étude des sursauts radio rapides (FRBs, de l'anglais *Fast Radio Bursts*) avec le télescope *Canadian Hydrogen Intensity Mapping Experiment* (CHIME) a commencé à découvrir des FRBs à un rythme sans précédent. Ces nombreuses découvertes rendent possible une caractérisation des FRBs et de leurs propriétés via l'étude d'une population importante et cohérente pour la toute première fois. Cependant, le pipeline de détection des FRBs de CHIME (CHIME/FRB) est sujet à de nombreux biais de sélection. Par conséquent, l'échantillon de FRBs détectés par CHIME/FRB n'est pas représentatif de la véritable population. Pour corriger les biais affectant le pipeline, un système d'injection de FRBs synthétiques a été mis au point, et une grande population de FRBs simulés a été injectée en temps réel dans le flux de données du télescope. En injectant des sursauts tirés d'une population réaliste, le rapport signal sur bruit (SNR) a pu être comparé pour les différents paramètres d'entrée qui encodent les propriétés physiques des FRBs simulés. Les sursauts injectés ont été calibrés en une unité d'énergie physique (Jy ms) en temps réel. La position des sursauts dans le champ de vision du télescope a également été simulée, ce qui a permis l'obtention d'une représentation authentique de la détectabilité de véritables FRBs. L'ensemble final d'injections et des détections correspondantes seront pondérées de manière à ce que la distribution des paramètres de sorties corresponde ce qui a été effectivement observé par CHIME/FRB. Cela constitue un premier pas vers l'élimination des biais de sélection du pipeline de détection.

ACKNOWLEDGEMENTS

There are so many people I have to thank for getting me where I am now, but nobody has contributed to my development as a scientist more than my supervisor, Prof. Vicky Kaspi. Being a part of CHIME/FRB has been a wild ride, and having watched the system evolve from its first detection to the many we have now is inspirational beyond words. Thanks for letting me be a part of it!

I've met a ton of inspiring scientists as a part of the CHIME/FRB cohort at McGill who deserve thanks. Thanks to my officemates Ziggy, Pragya, Alex, and Émilie for being so supportive, answering my endless questions, and acting as mentors each in their own way. Thanks to Bridget for being a seriously encouraging and sympathetic friend, and a wickedly smart researcher to work alongside. Thanks to the rest of the MSI dungeon crew – Andrew Z., Rafael, Simon, Coni. The willingness to be distracted by me when I just needed a friend to talk to was always greatly appreciated! Thanks to Alice, who has helped motivate me to continue working hard. Thanks to Shiny, Shriharsh, Kiyo, and Dustin, who have all acted as important mentors to me throughout the development of the CHIME/FRB injection system. Thanks to everyone not already mentioned who is a part of CHIME in any way, the hard work of the entire collaboration continually astonishes me and pushes me to be a better researcher.

Thanks to James for being such a pillar of positivity for me, especially during the pandemic. For keeping me grounded, focused, inspired, and helping me whistle while I work.

Thanks to my parents, without whom I would never have made it to where I am today. Thanks for your countless support, endless wisdom, gentle guidance, and your love – particularly while writing this thesis.

CONTRIBUTIONS BREAKDOWN

The CHIME telescope and its three experiments (Cosmology, FRB, Pulsar) described in Chapter 2 are a result of countless effort from many individuals. A list of members in the CHIME/FRB collaboration, for which the work in this thesis was developed, can be found in the author list of The CHIME/FRB Collaboration et al. (2020a).

Chapter 3 is work done primarily by Ziggy Pleunis and Chitrang Patel (§3.1, §3.2), Emmanuel Fonseca (§3.3), Bridget Andersen (§3.4), and the author. Contributions of the author include significant development of the intensity airflow pipeline (§3.1) and the DM pipeline (§3.2).

The bulk of original work is presented in Chapter 4. In particular, §4.1 and §4.2 is largely original material. The data flow presented in §4.2 was developed by Dustin Lang, expert data plumber extraordinaire. Without his major contribution, the injection system would not exist. §4.2.1 describes material primarily developed by Shriharsh Tendulkar. §4.2.4 describes material developed with major contributions from Shriharsh Tendulkar, Deborah Good, and Adam Dong. Development of `simpulse` (besides the API) described in §4.3.1 is work done primarily by Kendrick Smith. The data-driven beam model (§4.3.2) created by Paul Scholz used data from Saurabh Singh and Dallas Wulf, alongside a fit to a model which was developed by Gary Hinshaw. The process of verifying the calibration of injections and comparing this calibration to the CHIME/FRB detections (§4.3.3) was done by the author in collaboration with Bridget Andersen and Kiyoshi Masui. §4.4 and §4.5 describes work done by Kiyoshi Masui, Moritz Munchmeyer, Matt Dobbs, and Kaitlyn Shin.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
Introduction to Fast Radio Bursts

Astrophysical radio transients provide meaningful insight towards many areas of astrophysics. The discovery of the first pulsar in July of 1967 by Jocelyn Bell (PSR J1921+2153, Hewish et al., 1968) launched the study of fast radio emission. Since their discovery, pulsars have played a critical role in astrophysics. Numerous surveys searching for these periodic sources have discovered over 2000 pulsars in the Galaxy, as well as in the Small & Large Magellanic clouds. Pulsars' bright radio pulses have frequency-dependent arrival times due to dispersion from intervening ionised plasma in the interstellar medium. This effect is characterised by a value known as the dispersion measure (DM, see Equation 1.2), which is a census of the number of electrons along the line of sight from the source to the observer, and thus has allowed pulsars to be used as a probe of the electron content in the Milky Way (Cordes & Lazio, 2002; Yao et al., 2017). Further, when travelling through magnetic fields, investigating the pulsar population's polarisation properties using Faraday rotation measure (RM, see Equation 1.13) allows inference of the large scale structures in the Milky Way's magnetic field (Han et al., 2006). Not only are pulsars an excellent probe of Galactic properties, they act as extreme interstellar laboratories, allowing the study of effects ranging from gravitational wave radiation (Hulse & Taylor, 1975) to the properties of nuclear matter at extreme densities (Lattimer & Prakash, 2007). However, pulsar luminosities constrain their detection to sources of Galactic origin.

Even the most energetic of pulsar pulses from the young pulsar B0531+21 (known as the Crab pulsar) are only detectable to ∼Mpc distance using some of the world's most sensitive telescopes (Cordes et al., 2004). To probe properties such as the electron content and magnetic field of the intergalactic medium (IGM), more luminous radio sources originating from cosmological distances must be detected. This description of luminous, distant radio transients is befitting of *Fast Radio Bursts*.

Fast radio bursts, or FRBs, were first discovered through searching archival data from a 1.4 GHz survey of the Magellanic Clouds using the multibeam receiver on the 64-m Parkes Radio Telescope (Lorimer et al., 2007). The first event, dubbed the "Lorimer burst," bore characteristics of the pulses emitted by pulsars. However, there were two important distinguishing factors: the burst was one-off with no subsequent pulses detected to date, and the signal's location and degree of dispersion meant it must have originated from far outside the Milky Way or the Magellanic Clouds. Subsequent such pulses were detected by Parkes upon further review of archival data (Keane et al., 2012) and in the High Time Resolution Survey (HTRU, Thornton et al., 2013), establishing an apparent extragalactic population of radio bursts. While there was some question as to whether these first Parkes detections were in reality astrophysical (Petroff et al., 2015a), the detection of FRB 121102 at the Arecibo observatory by Spitler et al. (2014) eliminated the possibility that FRBs were due only to instrumental effects or radio interference at Parkes. Compounding its importance, repeat pulses from FRB 121102 were later detected (Spitler et al., 2016), dividing FRBs into classes of repeating and thus far non-repeating sources, and ruling out cataclysmic progenitor models for the repeaters. The catalogue of verified FRBs[1] (Petroff et al., 2016) grew slowly in the early days of FRB detections, as no FRB survey yet covered a sufficiently large sky area to detect many sources. However, cutting edge surveys such as the Canadian Hydrogen Intensity Mapping Experiment FRB Project (CHIME/FRB, CHIME/FRB Collaboration et al., 2018) which cover large swathes of the sky and search for FRBs in real time, are increasing the detection rate of FRBs by orders of magnitude.

Large populations of detected FRBs will lead to a plethora of rich science. Akin to the use of pulsars in our Galaxy as tools to probe the Milky Way's electron content and magnetic field, FRBs will provide much deeper understanding of the electron content and magnetic field in the IGM. Of particular importance is the use of FRBs to potentially solve the so called "missing baryon problem." The question of missing baryons has arisen because current observations severely underestimate the baryonic mass in the universe expected from cosmology (Cen & Ostriker, 1999). FRB DMs, which have components due to electrons in the Milky Way, IGM, and the FRB host galaxy, are a highly sensitive probe of baryons in the universe. McQuinn (2014) has shown that a population of ∼hundreds of FRBs with ∼arcminute localization could provide interesting constraints on the IGM baryon distribution around massive halos.

Despite the relative youth of FRB science, powerful measurements using FRBs are already being made. For example, Ravi et al. (2016) used the RM of the bright and highly linearly polarised source FRB 150807 to constrain the magnetic field of

[1] http://frbcat.org/

the cosmic web in the direction of the burst to at most order ∼nG, consistent with some simulations of the large scale cosmological magnetic field structure (Marinacci et al., 2015). Prochaska et al. (2019) detected FRB 181112 using the Australian Square Kilometer Array Pathfinder telescope (ASKAP, see Bannister et al., 2017), which allowed for its ∼arcsecond localization. This event intersected the halo of a foreground galaxy (FG-181112), and the DM and scattering properties of the detection showed that the density of FG-181112's halo was lower than inferences of halo densities made from quasar absorption lines (e.g. Lan & Fukugita, 2017). Macquart et al. (2020) have also showed with a sample of five "gold standard" FRB localizations from ASKAP that FRBs already have the capability to account for the universe's missing baryons, albeit with considerable uncertainty.

A final major discovery in the field to note is the detection of a bright, millisecond-duration radio burst associated with Galactic magnetar SGR 1935+2154 (The CHIME/FRB Collaboration et al., 2020a; Bochenek et al., 2020). While this event helped bridge the large gap between the intrinsic luminosity of pulsar/magnetar radio bursts and FRBs, it was still an ∼order of magnitude less bright than the dimmest of FRBs seen so far.

This thesis will begin with an overview of the essential observable properties of FRBs, as each plays an important role in affirming or discrediting various progenitor models. In §1.1, the generic properties of dispersion delay (§1.1.1), intrinsic width (§1.1.2), scattering (§1.1.3), brightness (§1.1.4) intrinsic radio spectrum (§1.1.5), and polarisation (§1.1.6) of FRBs will be discussed. Such a discussion of each property is important for this thesis, as FRB survey sensitivity is a function of many of these parameters on top of intrinsic survey effects such as the telescope's sensitivity pattern on the sky (also known as the telescope 'beam'). §1.1.7 will present what is known about the repetition of FRBs and the frequency-time structure common to repeating sources, and §1.1.8 will discuss observations and constraints on multiwavelength FRB counterparts. §1.2 will present a brief summary of popular FRB progenitor models, and their viability in the context of current observations. Finally, §1.3 will present the layout of the rest of the thesis.

1.1 Characteristics of Fast Radio Bursts

1.1.1 Dispersion

The interstellar and intergalactic media Fast Radio Burst emission travels through is very diffuse, but it is not a vacuum. Rather, it is made up largely of ionised plasma (Ferrière, 2001; Davé et al., 2001). When solving Maxwell's equations in such a charged, low-density medium, it can be shown that radiation has a frequency-dependent group velocity, resulting in a corresponding frequency-dependent time delay in the detected emission of astrophysical sources. The time delay for emission in the frequency range $\nu_1 < \nu < \nu_2$ is given by:

$$\begin{aligned}\Delta t &= \frac{e^2 \int_0^d n_e dl}{2\pi m_e c}\left(\frac{1}{\nu_1^2} - \frac{1}{\nu_2^2}\right) \\ &\approx \frac{1}{2.41\times10^{-4}}\left(\frac{\mathrm{DM}}{\mathrm{pc\ cm^{-3}}}\right)\left(\left[\frac{\nu_1}{\mathrm{MHz}}\right]^{-2} - \left[\frac{\nu_2}{\mathrm{MHz}}\right]^{-2}\right)\ \mathrm{s}\end{aligned} \tag{1.1}$$

where e and m_e is the electron charge and mass respectively, c is the speed of light, d is the distance from source to observer, and n_e is the number density of electrons in the intervening medium. The DM is defined by the integral of n_e along the line

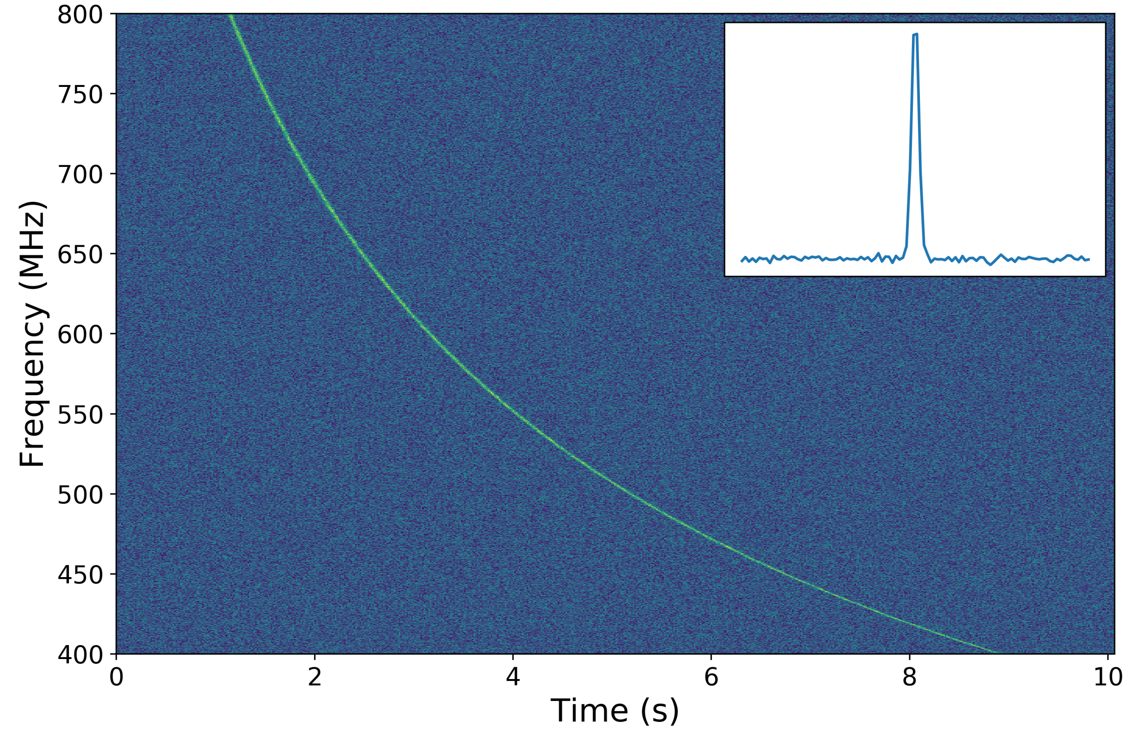

Figure 1–1: An example frequency vs. time (“waterfall”) plot of a dispersed burst generated using the `simpulse` package (see §4.3.1). The colour scale shows the intensity in each frequency-time bin. The pulse was generated with a DM of 400 pc/cm^3, which, using Equation 1.1, corresponds to a time delay across the CHIME frequency band (400-800 MHz, see Chapter 2) of ~ 7.77 seconds. The inset shows a cutout of the burst in the dispersion corrected, frequency-summed timeseries. The pulse was injected into white noise with fluence chosen to correspond to a nominal signal-to-noise (SNR) ratio of 50.

of sight:

$$\mathrm{DM} = \int_0^d n_e dl, \tag{1.2}$$

and can be derived from pulsar and FRB detections based on the emission's time delay as a function of frequency. Figure 1–1 shows an example of a dispersed pulse.

The DM is one of the most important characteristics of FRBs. Two popular electron density models of the Milky Way have been generated using pulsar DMs (NE2001[2] , Cordes & Lazio (2002); YMW16[3] , Yao et al. (2017)) and it is through comparing FRB DMs with the maximum Galactic DM along a given line of sight that their extragalactic nature can be inferred. Figure 1–2 shows clearly that, for the FRB population, the ratio of the observed DM to the maximum Galactic contribution is greater than one.

While the DM contribution from the Galaxy is one of the more constrained components contributing to an FRB's dispersion measure, it comes with its own uncertainties. Comparing the NE2001 and YMW16 models, for example, shows considerable discrepancies in predicted DM, particularly at low Galactic latitudes (Yao et al., 2017). Outside of the Galactic contribution, there are a number of other regions which contribute to the value of DM for a given source. The Milky Way halo, the IGM, and the host galaxy also contribute a fraction of the total electron density through which FRB radiation travels. This can be summarised in a simple equation:

$$\mathrm{DM_{obs}} = \mathrm{DM_{MW}} + \mathrm{DM_{halo}} + \mathrm{DM_{IGM}} + \frac{\mathrm{DM_{host}}}{1+z}, \tag{1.3}$$

where the observed DM of a given FRB ($\mathrm{DM_{obs}}$) is made up of contributions from the Milky Way ($\mathrm{DM_{MW}}$), the Milky Way halo ($\mathrm{DM_{halo}}$), the IGM ($\mathrm{DM_{IGM}}$), and the FRB host galaxy ($\mathrm{DM_{host}}$). $\mathrm{DM_{host}}$ is multiplied by a factor of $(1+z)^{-1}$, since in the

[2] https://www.nrl.navy.mil/rsd/RORF/ne2001/

[3] http://119.78.162.254/dmodel/index.php

observer's frame the degree of dispersion has been stretched due to the expansion of space time. Likewise, it should be noted that $\mathrm{DM}_{\mathrm{IGM}}$ is inherently a function of redshift since dispersion in the IGM happens over cosmological timescales for distant pulses. Often the excess DM is referred to, which is given by all contributions not coming from the Milky Way ($\mathrm{DM}_{\mathrm{E}} = \mathrm{DM}_{\mathrm{obs}} - \mathrm{DM}_{\mathrm{MW}} - \mathrm{DM}_{\mathrm{halo}} = \mathrm{DM}_{\mathrm{IGM}} + \frac{\mathrm{DM}_{\mathrm{host}}}{1+z}$).

The Milky Way halo contribution is not well constrained. X-ray observations of the hot, diffuse gas in the Milky Way halo can lead to estimates of DM from the gas's emission measure ($\mathrm{EM} = \int_0^d n_e^2 dl$). However, using the currently available EM data along with electron density estimates from cosmological hydrodynamic simulations, $\mathrm{DM}_{\mathrm{halo}}$ estimates still have ∼order of magnitude uncertainty (Keating & Pen, 2020). The estimated Milky Way halo contribution of $\mathrm{DM}_{\mathrm{halo}} \approx 30$ pc cm^{-3} from Dolag et al. (2015) is often used as a simple baseline. However, particularly for low excess DM FRBs, it's important to remember that this value is not well constrained as underestimating the Milky Way halo contribution could cause erroneous classification of a source as extragalactic.

The contribution of the IGM to the observed DM of FRBs is readily estimated based on estimates of the baryon density from ΛCDM cosmology fits by Planck Collaboration et al. (2018). Assuming the baryons derived from this fit are uniformly and isotropically distributed, as well as considering the fraction of baryons made up of ionised Hydrogen and Helium while applying appropriate redshift dependence leads to an integral over redshift making up the $\mathrm{DM}_{\mathrm{IGM}}$ contribution (Deng & Zhang, 2014). Overestimating $\mathrm{DM}_{\mathrm{IGM}}$ by assuming the universe is made up of only Hydrogen and that all baryons in the universe exist in the IGM leads to a rough DM-z relation

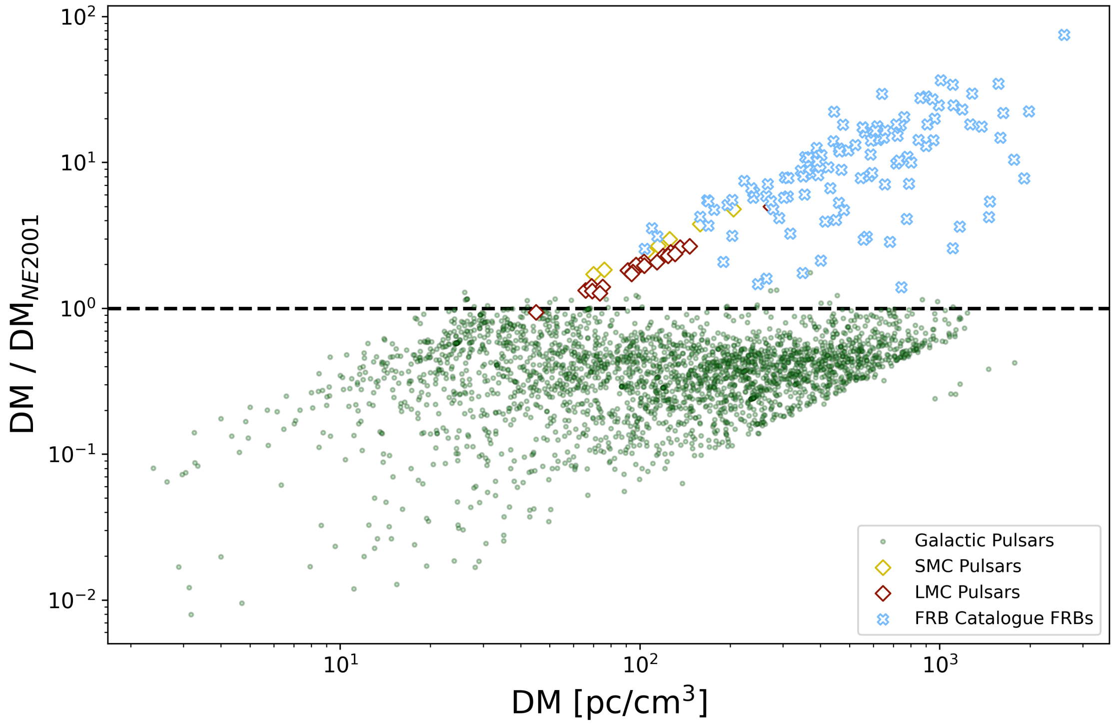

Figure 1–2: Comparing the dispersion measures of pulsars from the ATNF catalogue[4] (Hobbs et al., 2004) and FRBs from the FRB catalogue against the maximum Galactic DM contribution along the line of sight. The inferred maximum Galactic contribution is calculated using the NE2001 Galactic DM model.

(Zhang, 2018):

$$z \sim \mathrm{DM_{IGM}} / \left(855 \text{ pc cm}^{-3}\right). \tag{1.4}$$

Note that Equation 1.4 gives a redshift upper limit, is valid for $z \lesssim 3$, and is more uncertain at lower redshift (Zhang, 2018). While this DM-z relation is just an approximation, it still allows for meaningful science to be done in the era of detecting

[4] https://www.atnf.csiro.au/people/pulsar/psrcat/

many non-localized FRBs (e.g. Masui & Sigurdson, 2015). Once FRB surveys are able to localize many sources to their host galaxies, the true potential of FRBs as cosmic probes is unlocked. The ability to precisely determine each FRB's DM allows for a census of all the electrons along the FRB's line of sight. With enough distant sources this information can be used to determine the epoch of Helium reionisation (Zheng et al., 2014) as well as determine what fraction of baryons exist in the IGM (McQuinn, 2014).

The most difficult to determine of all DM components is $\mathrm{DM_{host}}$. Given the variety of galaxy morphologies, the potential for the FRB to lie anywhere within the host, the possibility of dense circumburst plasma, the random orientations of host galaxies, and the many conceivable FRB sight lines, this contribution is a challenge to estimate for a given source without localizing the host galaxy. Estimates of the distribution of host galaxy DMs have been made using a variety of methods. The probability distribution of host galaxy DM can be inferred from Markov Chain Monte-Carlo (MCMC) techniques if a set of bursts with known z are detected (Yang & Zhang, 2016), however the current dataset of FRBs does not yet have enough localized events for this to be used as a practical method. Another method has estimated the host galaxy DM distribution by assuming a narrow FRB luminosity distribution, using MCMC techniques with available FRB fluxes. The analysis was performed with a set of 21 bursts, and the results suggested large host galaxy DMs ($\sim 200 - 300$ pc cm^{-3}) are more likely (Yang et al., 2017). Directly modelling the electron densities in the thick disk, thin disk, spiral arms, and Galactic centre of the host has also been used as a method to estimate the host DM, assuming the event is viewed at random inclination angles (CHIME/FRB Collaboration et al., 2019a; Macquart et al., 2020). Estimation of FRB host DMs is still a relatively new problem however, and once many FRBs are localized, understanding of the host DM distribution will greatly improve.

It is important to note that the discrete nature of the frequency-time channels in typical FRB detection instruments means there will always be some amount of intrachannel smearing of the pulse within one frequency-time bin. A discussion of this effect, known as dispersion smearing, and on the methods used to search for FRBs at different DMs will be given in §2.2.1 and §2.2.2.

1.1.2 Width

FRBs have an associated intrinsic timescale, as emitted by the source. This timescale, known as the intrinsic width ($W_{\rm int}$) is derived from the expression:

$$W = \sqrt{W_{\rm int}^2 + t_{\rm samp}^2 + \Delta t_{\rm DM}^2 + t_{\rm DM_{err}}^2}, \tag{1.5}$$

where W is the observed width of the pulse, $t_{\rm samp}$ is the sampling time of the telescope which detected the pulse, $\Delta t_{\rm DM}$ is the dispersive delay across an individual frequency channel, and $\Delta t_{\rm DM_{err}}$ represents the contribution from dedispersing at a slightly incorrect DM. Note that the intrinsic width here represents what is detected at the radio telescope: a factor of $(1+z)$ must be multiplied to recover the intrinsic width at the source.

The intrinsic width of FRBs and its distribution offers some useful constraining information for FRB models, and it plays an important role in the detectability of bursts. It was from the intrinsic width of pulses from the first detected pulsar that

Hewish et al. (1968) were able to determine that the origin of the signal, if astrophysical, had to come from a region of space small enough to contain solely a white dwarf or neutron star (NS). Similar arguments can be made with FRB widths, where the region of emission in a model must be sufficiently small to explain the smallest widths observed (see §1.2). However, given the discretized time-sampling nature of radio data, detections of widths far below the sampling time of the instrument are not possible, and burst detectability quickly decreases below t_{samp} (Connor, 2019).

1.1.3 Scattering

The interstellar scattering of radio waves, first observed in the Crab pulsar (Hewish et al., 1968), manifests itself as the broadening in time at low frequency of a pulse (see Figure 1–3). The theory on the origin of scattering was first formalised by Scheuer (1968), who noted that a screen of inhomogenous media between a pulsating radio source and the observer could potentially explain amplitude variations between pulses. Salpeter (1969) was the first to point out that such a scattering screen also leads to temporal broadening, following the relation:

$$\tau \propto \nu^{-4}, \tag{1.6}$$

where τ is the scattering timescale at some frequency ν. τ can be fit by convolving an exponential tail to the pulse ($e^{-t/\tau}$, see Rickett 1977), with the resulting scattering timescale being used to understand the distribution of matter along the path of propagation.

The two important aspects of scattering models are the geometry of the scattering region, and the wavenumber spectrum of electron density fluctuations in the

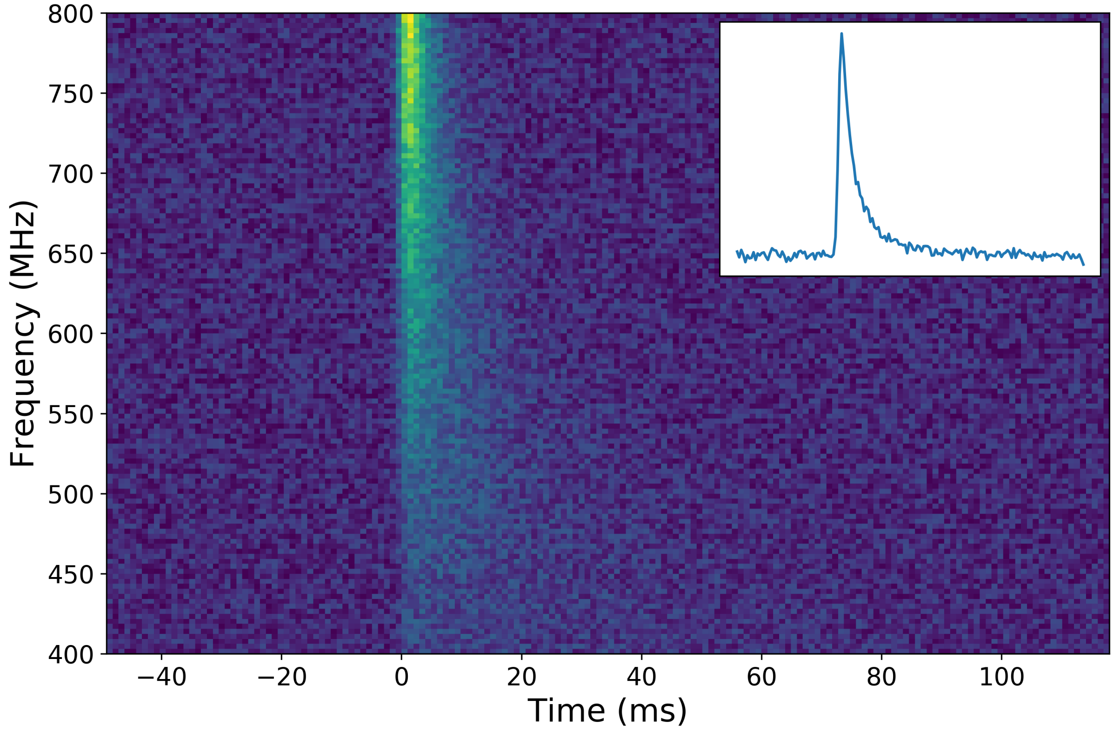

Figure 1–3: An example dedispersed waterfall plot of a scattered pulse generated using the `simpulse` package (see §4.3.1). The pulse was generated with an intrinsic width of 1 ms, and a scattering time at 600 MHz of 8.7 ms using the scattering index of Equation 1.6 (which corresponds to the pulse with the highest degree of scattering mentioned in CHIME/FRB Collaboration et al. 2019a). The inset shows the frequency-summed timeseries. The pulse was injected into white noise with fluence chosen to correspond to a nominal SNR ratio of 50.

region. For the former point, the most commonly considered geometry is either one or two thin scattering screens between the source and the observer. While the temporal broadening in Equation 1.6 assumed Gaussian inhomogeneities in the scattering medium, it is also common to assume Kolmogorov inhomogeneities which result in a scattering timescale $\tau \propto \nu^{-4.4}$ (Scheuer, 1968). Originally, the scattering timescale of pulses from the Crab pulsar showed its scattering to be most consistent with Kolmogorov density perturbations (Isaacman & Rankin, 1977). More recent observations of pulsar scattering show that large, corrugated sheets in the ionised interstellar medium with large density perturbations may best describe scattering phenomena (Simard & Pen, 2018).

Constraints on the scattering index α and DM-τ relation for FRBs are far less conclusive than for pulsars. Ravi (2019a) used FRBs which were scattered enough

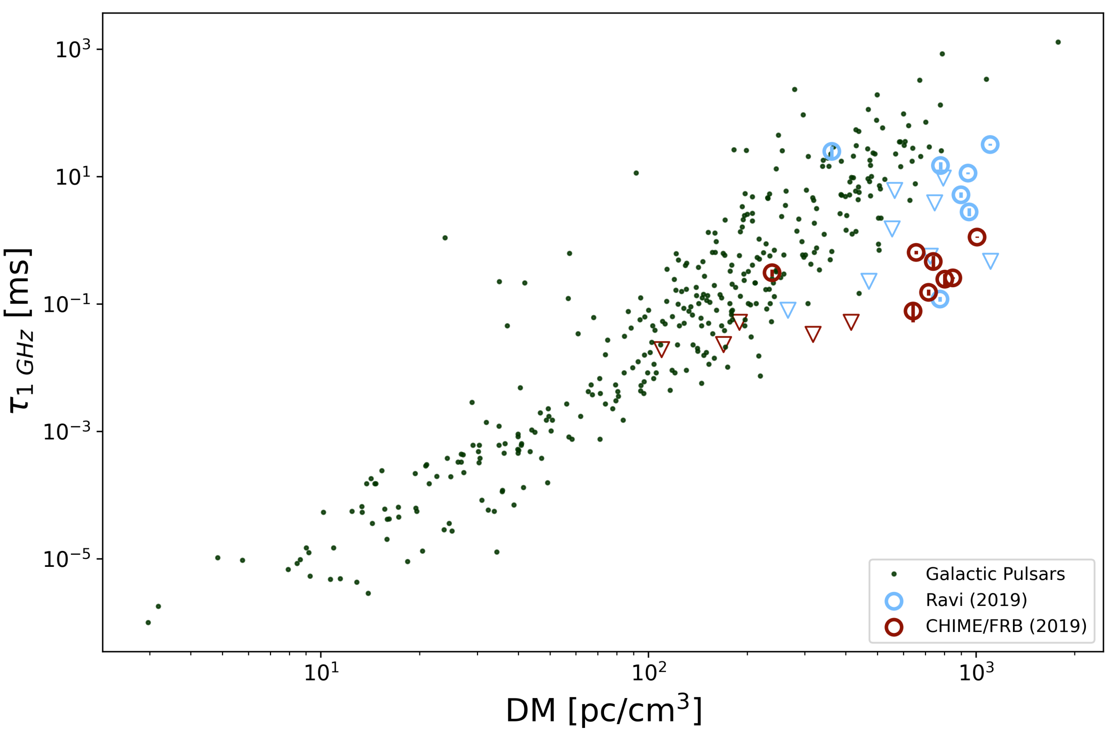

Figure 1–4: A comparison of the DM-τ relation between Galactic pulsars with measured scattering from the ATNF catalogue and FRBs with measured scattering in Ravi (2019a) and CHIME/FRB Collaboration et al. (2019a). FRBs which only have upper limits on scattering are denoted by a downward triangle, while the ones with measurements and errors are denoted by an open circle. All scattering times are referenced to 1 GHz, where necessary extrapolating using a scattering index $\alpha = -4$.

to fit for the scattering index α (where $\tau \propto \nu^{-\alpha}$) and recovered various α for FRB 010724 (Lorimer et al., 2007), FRB 090625 (Champion et al., 2016), FRB 110220 (Thornton et al., 2013), and FRB 131104 (Ravi et al., 2015). While the scattering index for FRB 010724 disagrees with both Gaussian and Kolmogorov scaling, the rest of the derived values agree with Gaussian scaling, and only FRB 110220 disagrees with Kolmogorov scaling. In general, it is common in the literature to assume a scattering index $\alpha = -4$ when fitting for scattering times.

One important point regarding the distribution of scattering times in FRBs is whether a correlation between DM and τ exists in the population. For pulsars with measured scattering, a correlation between scattering and DM exists across many orders of magnitude (see Figure 1–4 and Cordes et al., 2016). Comparing the DM-τ correlation of Galactic pulsars to FRBs, however, it is clear that FRBs are under-scattered for a given DM compared to the Galactic pulsar population, even after removing the Milky Way's DM contribution (Cordes et al., 2016). If the IGM is the main contributor to the excess DM of FRBs, then this is expected given the contribution from the IGM to scattering is expected to be small (Macquart & Koay, 2013). Further, if FRBs come from cosmological distances, any scattering occurring in the host galaxy is reduced by a factor of $(1 + z)^{-3}$ (Macquart & Koay, 2013). It is important to consider throughout the analysis of the DM-τ relationship for FRBs, however, that FRB surveys typically have a bias against detecting highly scattered events, particularly at low frequencies (CHIME/FRB Collaboration et al., 2019a).

Measurements of the FRB scattering can also be used to observe the density and turbulence of plasma in FRB host galaxies and the circumgalactic medium (CGM) of intervening galaxy halos (e.g. Prochaska et al., 2019). The existence of cool, ionised clumps in the CGM have been observed from intervening quasar absorption systems, and can also be observed using FRBs (Vedantham & Phinney, 2019). More useful analysis will be done when a distribution of many FRB scattering times that has been corrected for observational biases is available.

1.1.4 Brightness

One of the simplest imaginable observable properties of FRBs is how bright a radio telescope's detection appears against the radio sky background and system noise. In practice, however, determining this in physical units (i.e. in the flux density units of Jy) is far from trivial. First, a dedispersed timeseries must be created by choosing the DM which maximises the pulse SNR and averaging the signal over the receiver bandwidth. Then the flux density of the pulse $S(t)$ must be integrated over to obtain a fluence. This fluence is initially in arbitrary digital units determined by the voltage response of the receiver antenna. In order to convert the fluence to physical units, the radiometer equation is often used (Lorimer & Kramer, 2012):

$$\sigma_S = \frac{\mathrm{T}_{\mathrm{sys}}}{G\sqrt{2n_p \Delta\nu t_{\mathrm{samp}}}}, \tag{1.7}$$

where $\mathrm{T}_{\mathrm{sys}}$ is the system temperature in K (including contributions from the receiver temperature and sky temperature), G is the antenna gain in K/Jy (determined primarily from the area of the receiver dish), n_p is the number of polarisations combined to create the intensity data ($n_p = 2$ for most surveys), $\Delta\nu$ is the receiver bandwidth in Hz, and t_{samp} is the sampling time of the data in seconds. By taking σ_S as the root-mean-square (RMS) of the noise baseline (in Jy), the process of integrating over the pulse in the dedispersed timeseries thus gives a fluence.

Determining FRB fluences is complicated by the fact that it is difficult to determine the position of a given FRB in the beam. The beam pattern for a given radio survey will never have uniform sensitivity on the covered sky area, and for FRBs detected in a single beam it's possible the burst could have been located anywhere therein. This issue is slightly alleviated for surveys using phased array feeds, such as ASKAP. The frequency-dependent nature of beams can also aid in determining the location of a burst in the beam given the detected spectrum, though this is only truly viable by comparing spectra in multi-beam detections since it is difficult to disentangle the intrinsic FRB spectrum and that imparted by the beam model in single-beam detections. At the very least, the assumption can be made that the burst was detected in the most sensitive part of the beam, as was made by CHIME/FRB Collaboration et al. (2019a), giving a lower limit on the burst fluence.

The distribution of FRB fluences, commonly referred to as $\log N - \log S$, counts the number of sources which are expected to have a fluence above a given threshold, S_{obs}. This is often modelled as a power law with index γ:

$$N(> S_{\mathrm{obs}}) = S_{\mathrm{obs}}^{\gamma}. \tag{1.8}$$

In flat (i.e. Euclidean) space, for a population of FRBs whose rate doesn't evolve with cosmic time (i.e. has a constant number density $n(\vec{x}) = n_0$), an index $\gamma = 3/2$ is expected regardless of the luminosity function $f(L)$ of bursts. This comes from a combination of volume scaling as D_L^3 (where D_L is the luminosity distance) and flux scaling as $1/D_L^2$:

$$\begin{aligned} N(< D_L) &\propto D_L^3 \\ D_L &= \sqrt{\frac{f(L)}{4\pi S}} \propto S^{-1/2} \\ \implies N(> S) &\propto \left(S^{-1/2}\right)^3 \propto S^{-3/2}. \end{aligned} \tag{1.9}$$

Thus, deviations from $\gamma = 3/2$ contain valuable information about whether the FRB population comes from cosmological distances, and how the evolution of FRB event rate may change with z. A variety of values for γ have been presented in the literature, and its value is not well constrained. A number of authors have found evidence for γ consistent with the Euclidean expectation (Oppermann et al., 2016; Katz, 2016; Bhandari et al., 2018; Macquart & Ekers, 2018) while others have determined that a shallower slope ($\gamma < 3/2$) is more likely (Caleb et al., 2016; Vedantham et al., 2016). A larger sample of FRBs with well determined fluence, accounting for instrumental biases within the fluence distribution, will allow for the determination of properties such as whether the FRB density evolves with redshift (Caleb et al., 2016).

1.1.5 Frequency Spectrum

It is common to describe the frequency spectrum of FRBs as a power law:

$$S_\nu \propto \nu^\beta, \tag{1.10}$$

where β is the spectral index. Various radiative processes, particularly synchrotron emission, give rise to a power-law spectrum (Rybicki & Lightman, 1979), and the spectrum of pulsars is well described by a power law (Lorimer & Kramer, 2012), motivating this parameterization. The power-law model is simple, however, and there are many factors which make the detected spectra of FRBs far more complex than this simple description. Frequency-dependent beam effects, particularly from interferometers, are difficult to correct for fully, so recovering the intrinsic spectrum of the burst without instrumental effects already poses challenges. Further, interstellar scintillation or intrinsic radiation mechanisms can produce "knotty" FRB spectra, with patches of enhanced and reduced intensity with a characteristic frequency scale known as the *decorrelation bandwidth*.

While FRB spectral indices are important for the discrimination of models, the band-limited nature of radio telescopes makes it difficult to get a wide enough frequency coverage for a broad picture of the spectral energy distribution of FRBs. Further, FRB emission (particularly from repeating sources) is not always well described by a power law, as very band-limited patches of emission are often observed (see e.g. CHIME/FRB Collaboration et al., 2019b). This motivates instead the use of a "running" power law to describe FRB spectra, with an additional index ζ known as the spectral running:

$$S_\nu \propto \left(\frac{\nu}{\nu_0}\right)^{\beta+\zeta\ln(\nu/\nu_0)}, \tag{1.11}$$

where ν_0 is known as the "pivotal frequency" and is a chosen, fixed constant. This parameterization may be a better way to describe the more narrow bandwidth emission predicted in models such as FRBs resulting from plasma lensing events (Cordes et al., 2017).

Macquart et al. (2019) took a collection of 23 FRBs detected by ASKAP in order to try and grasp the spectral properties of FRBs in the ASKAP band (1152 – 1488

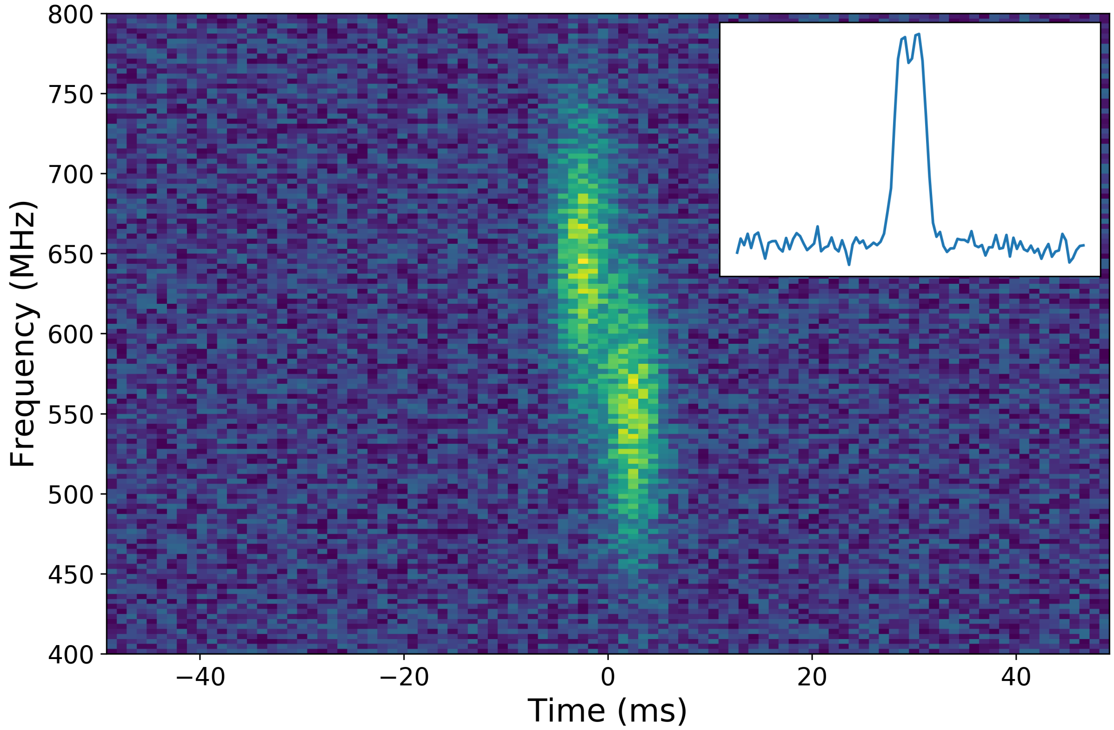

Figure 1–5: An example dedispersed waterfall plot of a pulse exhibiting downward drifting, band-limited structure generated using the `simpulse` package (see §4.3.1). Each sub-pulse was generated with an intrinsic width of 2 ms, with a drift rate of −20 MHz/ms (typical for bursts seen in the CHIME band, CHIME/FRB Collaboration et al., 2019b; Fonseca et al., 2020). The inset shows the frequency-summed timeseries. Each sub-pulse was injected into white noise with fluence chosen to correspond to a nominal SNR ratio of 15.

MHz). By taking the mean flux density in each spectral channel from their set of FRBs, they found a spectral index $\alpha = -1.5^{+0.2}_{-0.3}$. Notably, this is very similar to the weighted mean spectral index of pulsars ($\alpha = -1.6$, Jankowski et al., 2018), hinting at a potential connection between the emission of the two, though the sample is not large enough to draw any robust conclusions.

Notably, repeating FRBs tend to show complex time-frequency structure, including sub-bursts with finite bandwidth spaced closely in time (Hessels et al., 2019). These sub-bursts tend to drift downward in frequency, a characteristic aptly compared to a sad trombone (see Figure 1–5 for an example). While this structure appears primarily in repeating FRBs (Hessels et al., 2019; CHIME/FRB Collaboration et al., 2019b; Fonseca et al., 2020), there have also been (thus far observed) non-repeating FRBs exhibiting fine microstructure (Farah et al., 2018). Curiously, no FRB thus far observed has shown upward drifting sub-bursts, although some may argue (Wang et al., 2020) that a set of two bursts from FRB 180916.J0158+65 (191219A/B, Chawla et al., 2020), as well as the set of two bursts from the Galactic magnetar SGR 1935+2154 (The CHIME/FRB Collaboration et al., 2020a), are upward drifting sub-bursts with larger than usual separation.

1.1.6 Polarisation Properties

When the plasma through which radiation travels on its path from source to observer is magnetised, the right-circularly polarised and left-circularly polarised components of the radiation travel at different speeds. Which component travels faster than the other depends on the orientation of the magnetic field in the medium. The difference in effective index of refraction for the two circular polarisations results in a rotation of the position angle of linear polarisation as a function of frequency (Θ) imprinted on the radiation. This effect is known as Faraday rotation, and is described by Rybicki & Lightman (1979) in detail. The expression for Θ is:

$$\Theta = \mathrm{RM}\ \lambda^2, \tag{1.12}$$

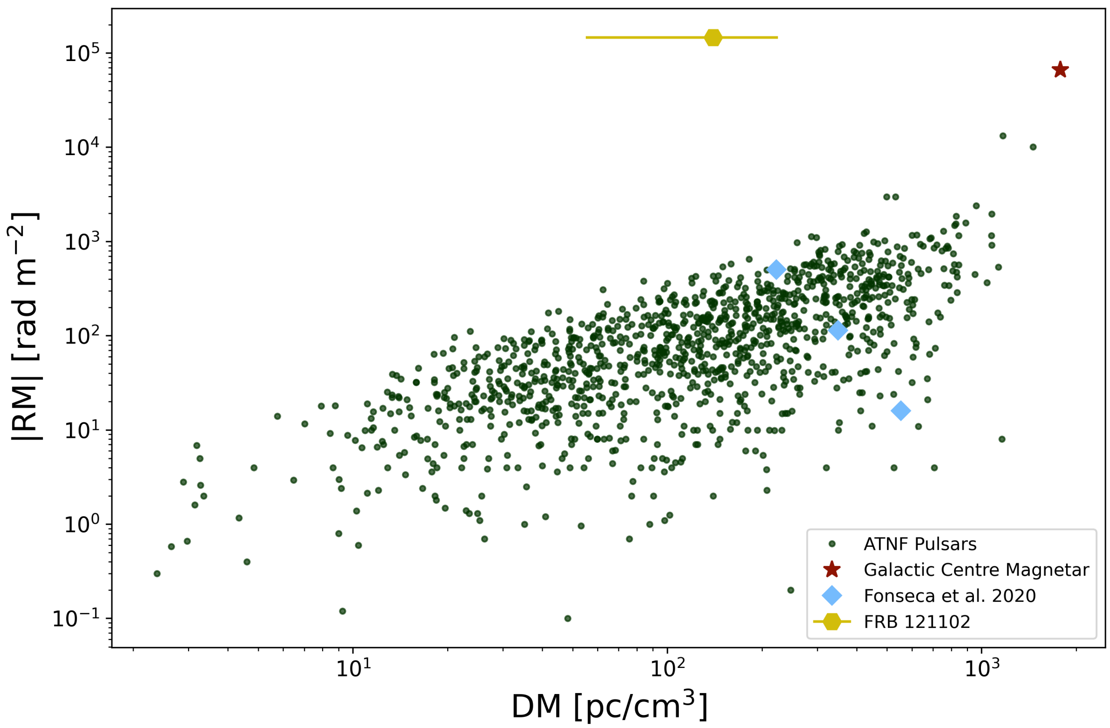

Figure 1–6: A comparison between the DM and RM of pulsars in the ATNF catalogue and some FRBs with measured RM. The Galactic centre magnetar, which has the highest value of RM measured among known pulsars, is marked with a red star, FRB 121102 is marked with a yellow hexagon, and bursts with measured RM from (Fonseca et al., 2020) are marked with blue diamonds. The DM of FRB 121102 plotted is the estimated host contribution, so that a comparison is being made in the context of contributions from the host galaxy and circumburst environment. The error bar shows the range of potential values for $\mathrm{DM}_{\mathrm{host}}$ (Tendulkar et al., 2017).

where RM is the rotation measure, and λ is the wavelength of light. RM is defined similarly to DM (see Equation 1.2), except including a contribution from the magnetic field parallel to the line-of-sight, $B_\parallel$:

$$\mathrm{RM} = \frac{e^3}{2\pi m_e^2 c^4} \int_0^d B_\parallel n_e dl. \tag{1.13}$$

RM acts as a powerful tool to investigate the magnetic properties of the FRB environment. Since the IGM is expected to contribute very little to FRB RMs due to its weak magnetic field (Marinacci et al., 2015), and the Galactic Faraday sky can be inferred from RM measurements towards Galactic objects (Oppermann et al., 2012), the contribution to the RM from the host galaxy and circumburst environment can be probed. If the parallel component of the magnetic field changes direction over the course of propagation, however, the RM contributions can effectively null themselves even in the presence of moderate or strong magnetic fields.

While the Galaxy's magnetic foreground limits the conclusions that can be drawn using RM measurements, some FRBs still have significant RM detections. FRB 121102, for example, has an extremely large value of RM ($\sim 10^5$ rad m^{-2}) that varies rapidly ($\sim 10^4$ units over 7 months), so an extreme and variable magneto-ionic environment associated with the source must be inferred (Michilli et al., 2018). The only known Galactic source with comparable RM is the Galactic centre magnetar SGR J1749$-$2900 (see Figure 1–6), which is very close ($\sim$0.1 pc) to the supermassive black hole at the centre of the Milky Way (Eatough et al., 2013). Fonseca et al. (2020), however, measured much less extreme RMs for three repeating sources. Further RM measurements of FRBs will reveal whether the extreme environment of FRB 121102 is peculiar to that particular source or if it is a more common characteristic of the (repeating) FRB population.

1.1.7 Repetition

Before the first detection of FRB repetition, many hours of dedicated follow-up time were given to various FRBs (see Petroff et al. 2015b; Rane & Lorimer 2017 for the observation time given to some sources), but no repeat bursts were detected. This put strong limits on models for FRBs which had expected repetition times on the order of hours, but the amount of dedicated telescope time required made excluding models with repetition time scales on the order of months to years difficult (Petroff et al., 2015b). Thus, the detection of multiple bursts from FRB 121102 (Spitler et al., 2016) using the Arecibo Observatory was an important milestone for FRB science.

FRB 121102's repetition meant cataclysmic models for its origin could be ruled out entirely, and follow-up observations could be done with more confidence. An interferometric observation with the Very Large Array (VLA) led to the localization of FRB 121102 to its host galaxy (Chatterjee et al., 2017), unambiguously showing that FRBs came from outside of the Milky Way. The host galaxy of the source along with a redshift were determined soon after (Tendulkar et al., 2017). FRB 121102's host galaxy is a low metallicity, star-forming dwarf galaxy at $z \sim 0.19$, which is particularly peculiar given the low fraction of baryonic mass contained in dwarf galaxies (Papastergis et al., 2012). FRB 121102's host, along with the existence of a co-located, compact, persistent radio counterpart (Marcote et al., 2017), have led to much speculation about FRB 121102's progenitor. As a source, it has arguably offered the most insight into the nature of Fast Radio Bursts, albeit through a biased lens.

With the discovery of FRB repetition, the potential for all FRBs to be repeating sources came into question. Caleb et al. (2019) simulated observing a universe of repeating FRBs with the Aercibo, MeerKAT, Parkes, and ASKAP observatories to try and determine if a single population of repeating bursts could be excluded using the available constraints at the time. Two cases for repetition statistics were assumed based on the observed wait-time distribution of FRB 121102 bursts (Oppermann et al., 2018; Zhang et al., 2018), and the choice of power-law indices for the energy distribution of pulses was in the range $-2 < \gamma < -1$ (similar to what has been found for FRB 121102, Law et al., 2017; Gourdji et al., 2019). Given these simulations, a population of only repeating FRBs could not be excluded given current observations, particularly in the case of steeper γ and longer repetition rates. Ravi (2019b) also noted after detection of the second repeating FRB by CHIME/FRB (FRB 180814.J0422+7, CHIME/FRB Collaboration et al., 2019c), that most observed FRBs likely come from repeating sources given the potential astrophysical scenarios of FRB production.

Since the discovery of the first two repeating FRBs, 17 further repeating sources have been discovered by CHIME/FRB (CHIME/FRB Collaboration et al., 2019b; Fonseca et al., 2020). One of these sources has been observed to repeat with a ~16.35 day periodicity (FRB 180916.J0158+65, The CHIME/FRB Collaboration et al., 2020b). This source is the only repeater with concrete periodic activity, though the nature of CHIME/FRB as a transit telescope makes it difficult to confirm whether the observed 16.35 day period is an alias of a shorter period. FRB 121102 has also been observed to have possible periodic activity on a longer timescale (P ~157 days, Rajwade et al., 2020), but the significance is hard to quantify due to the uneven observing cadence of the source. The further quantification of periodicity in FRB repeaters will provide clues as to the nature of these objects.

1.1.8 Multiwavelength Counterparts

Outside of FRB radio emission itself, the detection (or lack thereof) of emission at other wavelengths is of considerable interest to FRB theorists. Particularly, FRB host galaxies in the optical, high energy counterparts in the X-ray and gamma-ray, and gravitational wave (GW) counterparts detected by GW observatories may one day provide significant development in the understanding of FRB progenitors. The multi-wavelength follow-up of the NS-NS merger GW 170817 (Abbott et al., 2017) has proven what fruit can be borne from multi-messenger astrophysics, and multiwavelength detections of FRBs promise similarly exceptional results.

Host galaxy properties play an important role in understanding the origin of FRBs. Exploration of the host galaxy population and their mass, metallicity, and star-formation rate (SFR) can reveal whether FRBs form preferentially in environments conducive to certain formation channels such as supernovae, or NS-NS mergers. Further, the location of FRBs within host galaxies can confirm or rule-out FRB progenitors with active galactic nuclei (AGN). There are currently eight FRBs with localized host galaxies: FRB 121102 (Tendulkar et al., 2017), FRB 180924 (Bannister et al., 2019), FRB 181112 (Prochaska et al., 2019), FRB 190523 (Ravi et al., 2019), FRB 180916.J0158+65 (Marcote et al., 2020), FRB 190102, FRB 190608 (Bhandari et al., 2020), and FRB 191107 (Macquart et al., 2020).

Given the first known FRB host galaxy (that of 121102) was a peculiar low metallicity, star-forming dwarf galaxy, there was much speculation as to the connection between FRB progenitors and this particular environment. Metzger et al. (2017) noted that superluminous supernovae (SLSNe) and long gamma-ray bursts (LGRBs) are both preferentially hosted in 121102-like environments, giving the potential for a young magnetar to be the source of all three phenomena (SLSNe, LGRBs, and FRBs). However, the localization of sources besides FRB 121102 has shown that not all FRBs live in such peculiar environments (Bhandari et al., 2020). In fact, none of the other galaxies described in Bhandari et al. (2020) favour environments preferentially hosting SLSNe. Additional localizations and analysis of resulting host properties will provide strong constraints on the likely progenitors of FRBs.

Simultaneous or near simultaneous signals in the X-ray and gamma-ray wavelengths are predicted in a number of FRB progenitor models. The first near real-time follow-up of an FRB at high energy was performed by Petroff et al. (2015c) with FRB 140514, who did not detect any high energy counterpart in observations ∼hours after the FRB went off. This allowed any association with typical long GRBs and SLSNe to be excluded for this source.

The discovery of the first repeater meant simultaneous high energy and radio observations of the source location could be made, and real-time FRB detections checked for high energy counterparts. Scholz et al. (2017) performed simultaneous radio, X-ray, and gamma-ray observations of FRB 121102 during one of its active periods. Twelve radio bursts occurred during the X-ray and radio observations, but no high energy counterparts were observed. This put limits on the emission of X-rays and gamma-rays for these FRB 121102 events, however the flux limits were still an order of magnitude larger than the X-ray flux that would be detectable by an SGR 1806−20-like magnetar giant flare (Palmer et al., 2005) at the distance of FRB 121102.

While there have been no observed high energy counterparts of extragalactic FRBs, the Galactic radio burst from SGR 1935+2154 did have an X-ray counterpart. Not only did the radio burst coincide with the period of X-ray reactivation of the magnetar (Borghese et al., 2020), but there was an X-ray burst coincident with the radio burst and exhibiting similar temporal structure detected by the INTEGRAL (Mereghetti et al., 2020), Konus-wind (Ridnaia et al., 2020), AGILE (Tavani et al., 2020), and Insight-HXMT (Zhang et al., 2020) missions. Notably, the X-ray spectrum of the burst coincident with the radio burst was harder than typical SGR 1935+2154 X-ray bursts (Mereghetti et al., 2020). However, the high energy flux was still much less than that which would be detectable at extragalactic distances with current high energy missions.

1.2 Progenitor Models of Fast Radio Bursts

In the early days of FRB astronomy, measurable constraints on models were few. In fact, until recently, the number of FRB progenitor theories (collated by Platts et al. 2019)[5] outnumbered the number of detected FRBs. While it is tempting to come up with a single explanation for all FRBs, given the large variety of observed burst features (particularly the repetition seen in a subset of FRBs), it is entirely possible the population may be made up of more than one kind of progenitor.

There are a number of qualifications an FRB progenitor model should consider to viably explain the FRB phenomenon. For example:

[5] https://frbtheorycat.org

- **Energetics:** Typical FRB fluences are in the ∼Jy ms range at ∼1 GHz. The equivalent isotropic burst energy depends on the distance, which is not well constrained for unlocalized bursts. Taking FRB 121102 as an example, the range of burst energies spans over orders of magnitude (10^{37} erg $\sim 10^{40}$ erg, Law et al. 2017; Zhang et al. 2018; Gourdji et al. 2019). Thus FRB models must include radio emission mechanisms which provide such a large range of burst energies.
- **Duration:** Short (∼ms) widths are typical for FRBs. Often compact objects are invoked to produce emission, as the light travel time given the FRB duration is small. Regardless, progenitor models must account for the short intrinsic widths of FRBs.
- **Magnetic properties:** Sources such as FRB 110523 and FRB 121102 have been shown to live in extremely magnetic environments (Masui et al., 2015; Michilli et al., 2018), while sources such as FRB 150215 (alongside the aforementioned FRB 110523) are highly polarised (Petroff et al., 2017). Models should try to explain the highly magnetic environments of some FRBs, and the origin of the pulse polarisation.
- **Repeatability:** As discussed in §1.1.7, the existence of one or multiple populations of FRBs cannot yet be proven (see e.g. Caleb et al., 2019). If a model hopes to explain repetition in FRBs, it must explain the diversity of repetition rates ($\sim 10^{-1} - 10^{1}$ hr^{-1}, Fonseca et al. 2020), the periodicity of FRB 180916.J0158+65 observed by The CHIME/FRB Collaboration et al. (2020b), and the apparent periodicity of FRB 121102 observed by Rajwade et al. 2020.

- **Host galaxy properties:** The properties of FRB host galaxies may be one of the simpler ways to discern what makes FRBs. Bhandari et al. (2020) showed how the host galaxy properties of localized FRBs compare to those of other energetic events, such as SLSNe, WD-WD mergers, double NS mergers, and core-collapse SNe. For repeating models invoking young NSs, hosts with young stellar populations will be preferred, whereas FRBs produced by mergers would be more likely to come from hosts with intermediate-to-old stellar populations. The relative dearth of localizations so far makes it hard to make robust conclusions using host galaxy properties, however.

The above list is by-no-means a definitive set of check-boxes an FRB progenitor model must satisfy to be considered viable, but rather it serves to summarize some of the most important clues available as to the nature of FRBs.

The abundance of FRB theories makes it impossible to give a concise overview of all viable models. However, the association of an energetic radio burst with the Galactic magnetar SGR 1935+2154 makes magnetar progenitor models for FRBs particularly popular. That being said, the SGR 1935+2154 burst was orders of magnitude less energetic than typical FRBs, meaning if it were due to the same process which generates typical FRBs it would have had to be on the very low-energy end of the luminosity function (The CHIME/FRB Collaboration et al., 2020a; Bochenek et al., 2020). It is also uncertain whether the rate of such magnetar bursts observed in our Galaxy can satisfy FRB rate constraints (The CHIME/FRB Collaboration et al., 2020a; Bochenek et al., 2020). So, the SGR 1935+2154 event does not rule out multiple FRB progenitors. It is, however, a smoking gun detection that cannot be ignored when theorizing the origins of FRBs. Here, a brief description of two popular FRB progenitor models is given:

- Metzger et al. (2019) present how synchrotron maser emission from ultra-relativistic magnetized shocks may produce FRB emission. While the shock may be driven by any sufficiently energetic central engine surrounded by layers of magneto-ionic matter, young magnetars conveniently provide both strong magnetic fields and a dense surrounding medium through which the shock can propagate. Among other predictions, the synchrotron maser shock model explains the unique downward sub-pulse drifting seen in repeating FRBs through the deceleration of the forward shock. It is also notable that this model predicts X-ray bursts coincident with FRBs, having a relative fluence $E_X/E_{\rm radio} \sim 10^{-5}$. Such an X-ray burst was observed in the case of the SGR 1935+2154 radio burst (Margalit et al., 2020).
- Lu et al. (2020) present an avenue for magnetospheric curvature radiation to produce FRBs. From a disturbance on the NS crust, Alfvén waves propagate through the NS magnetosphere. Further away from the NS surface, the charge density becomes small enough such that the plasma current associated with the wave can no longer be sustained. The lack of plasma current results in charge starvation and thus a strong electric field generated parallel to the magnetic field. The electric field accelerates charge clumps to high Lorentz factors, producing curvature radiation in the radio band. In this model, the heated NS surface at the site of the crustal disturbance produces soft X-rays, while the Alfvén waves propagating near the NS surface produce hard X-rays through

inverse Compton scattering. The coherent radiation in the magnetosphere can naturally explain downward drifting sub-pulses from *radius-to-frequency mapping* (Lyutikov, 2020). This model hopes to unify the relatively faint radio burst of SGR 1935+2154 with the extragalactic FRB population, given the volumetric rate at the faint end of the FRB luminosity function may match with the rate of SGR 1935+2154-like bursts.

1.3 Layout of the Thesis

Undoubtedly, the introduction of this thesis raises more questions about FRBs and their underlying population than it provides answers. However, the CHIME/FRB instrument has massive potential to elucidate many of these questions with its large, coherent sample of FRBs. Chapter 2 will provide details about the CHIME telescope and the CHIME/FRB experiment, including data processing in the live pipeline and the observational biases introduced as a result. Chapter 3 will then focus on the intensity data and subsequent data processing done by CHIME/FRB post-detection, in order to understand how the parameter values are obtained in the detected population. Chapter 4 will explain the injection system for CHIME/FRB, including the system architecture, the burst modelling process, and how the data will be used to correct the CHIME/FRB population for observational biases. Finally, Chapter 5 will present the conclusions of the work done in this thesis, and its impact on the results of CHIME/FRB.

CHAPTER 2
The Canadian Hydrogen Intensity Mapping Experiment

The Canadian Hydrogen Intensity Mapping Experiment (CHIME, see Newburgh et al., 2014) is a recently commissioned radio telescope near Penticton, British Columbia. CHIME was originally conceived solely as a cosmology experiment, with the goal of detecting a baryon acoustic oscillation (BAO) signal from neutral hydrogen at $z = 0.8 - 2.5$ (see Bassett & Hlozek, 2010, for a review on BAO). CHIME was thus designed to have an enormous field of view (~ 200 deg^2) and sufficient sensitivity in order to detect the BAO signal.

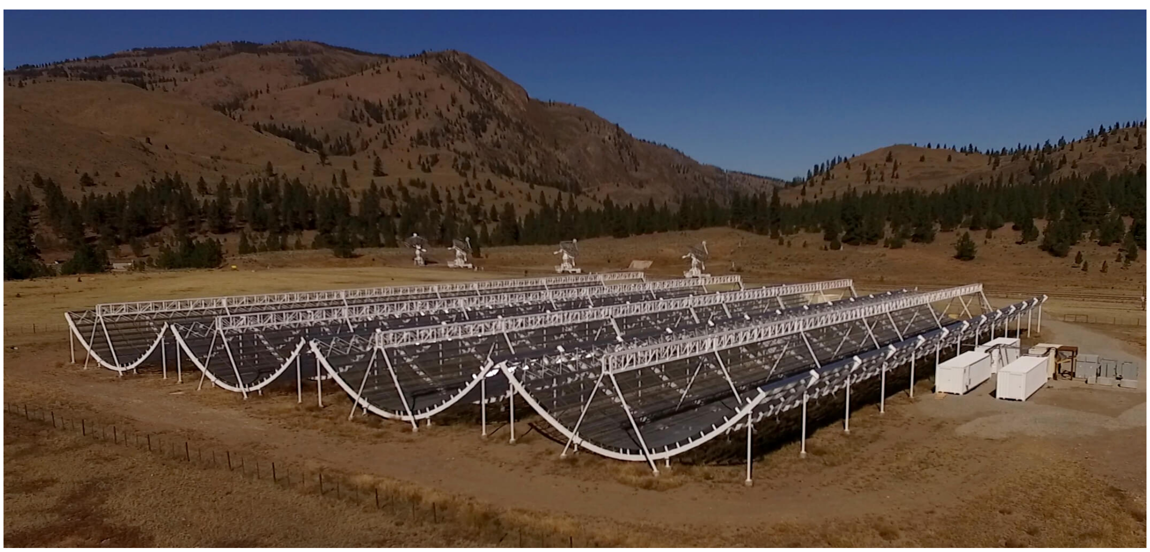

Figure 2–1: The CHIME telescope on 2016 September 15 at the Dominion Radio Astrophysical Observatory (DRAO) near Penticton, British Columbia. Figure from CHIME/FRB Collaboration et al. (2018).

The CHIME telescope is composed of four 20-m wide and 100-m long cylindrical paraboloidal reflectors with North-South alignment. Along the axis of each cylinder hang 256 dual-polarisation feeds, each spaced by 30 cm, along 80 m of the focal line (which is placed at a focal height of 5 m). CHIME has no moving parts: instead of pointing towards a location on the sky, it scans the sky passing overhead. The expense saved from this simplicity in design allowed investment in a very powerful FX correlator.

The 'F-engine' (named for the Fourier transform operations it performs) processes a total data rate of 13.1 Tb s^{-1} across 128 Field Programmable Gate Arrays (FPGAs) housed on specially designed 'ICE' motherboards (for information on the design of the ICE boards, see Bandura et al., 2016). The F-engine digitizes and channelizes the raw output data from the antennas into 1,024 evenly spaced frequency channels from $400-800$ MHz[6] . The F-engine is housed in specialized "receiver huts" which are shipping containers equipped with radio frequency interference (RFI) shielding and liquid cooling for the electronics.

The 'X-engine' (named for the cross correlation operations it performs) receives its inputs from the F-engine through 1,024 optical high-speed transceivers. The X-engine houses 256 nodes, and each node houses two dual-chip GPUs, meaning each node has effectively four GPUs and thus can handle the datastream for four of the

[6] Note that the CHIME frequency range is actually from $400.1953125-800.1953125$ MHz, but this will be abbreviated to $400-800$ MHz, for brevity.

Parameter	Value
Longitude	119°37′26″ W
Latitude	49°19′16″N
Altitude	545 m
Collecting area	8000 m^2
E-W field of view	2.5° (400 MHz) to 1.3° (800 MHz)
N-S field of view	$\sim 120°$
Frequency range	400−800 MHz
Frequency channels	1,024 (Baseband data), 16,384 (FRB intensity data)
Time resolution	2.56 μs (Baseband data), 0.98304 ms (FRB intensity data)

Table 2–1: Key values for the CHIME telescope and CHIME/FRB project.

1,024 frequency channels from the F-engine. Each X-engine node is liquid cooled with direct-to-chip cooling for the CPUs and GPUs, and the nodes are stored in two RFI shielded 40-ft shipping containers adjacent to the telescope (visible in Figure 2–1).

The large field of view offered by CHIME means it not only has potential as an excellent cosmology experiment, but may also search for fast radio transients over an unprecedented instantaneous sky coverage. CHIME's original design was based solely on its ability to map neutral hydrogen, but the realization that CHIME would make an excellent radio transient detector led to X-engine upgrades allowing for radio transient searches. An overview of the resulting FRB search project with CHIME (CHIME/FRB) will be given in §2.1. The real-time search and detection pipeline for CHIME/FRB will then be discussed in §2.2.

2.1 CHIME/FRB

The output digital data of the CHIME X-engine is sent to three separate projects: the CHIME/Cosmology project (Newburgh et al., 2014), the CHIME/FRB project

(CHIME/FRB Collaboration et al., 2018), and the CHIME/Pulsar project (CHIME/Pulsar Collaboration et al., 2020). The CHIME/FRB project takes advantage of massive computational power and CHIME's enormous field of view to search for FRBs in real time. CHIME/FRB can be described as a software-driven experiment, employing the powerful CHIME FX correlator along with a dedicated FRB search backend to detect FRBs.

CHIME/FRB consists of 132 compute nodes housed similarly to the X-engine, in RFI shielded shipping containers placed adjacent to the telescope. A combination of RFI mitigation and the CHIME/FRB dedispersion search algorithm `bonsai` (K. M. Smith et al., in prep.) runs on 128 of these nodes as the "Level 1" or L1 process (see §2.2.2). A set of two more nodes (one of which is a backup in case of node failure) is used to run multi-beam grouping, extragalactic event determination, and action determination algorithms, which are known as the L2 and L3 processes (see §2.2.3). Finally a set of two nodes (again with one being a backup) performs action implementation and hosts databases and web interfaces, known collectively as the L4 process (see §2.2.4).

2.2 The CHIME/FRB Real-time Pipeline

Here, a high-level overview of the CHIME/FRB real-time pipeline is given. Figure 2–2 provides a succinct overview of each level in the pipeline. It should be noted that though the L0 process prepares X-engine data for CHIME/FRB, the X-engine also sends data to the CHIME/Cosmology and CHIME/Pulsar experiments.

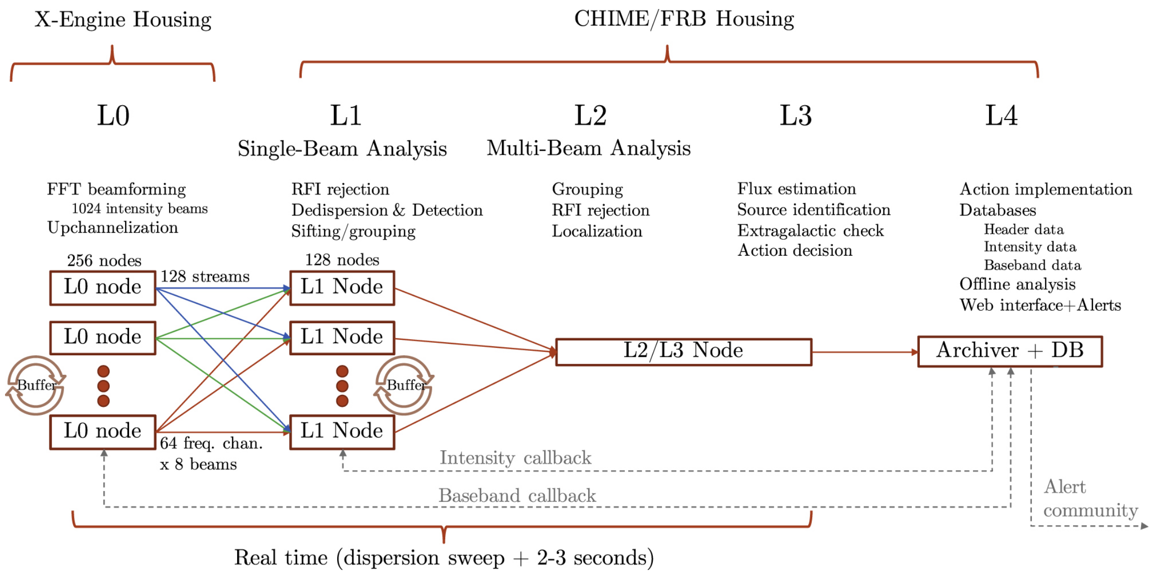

Figure 2–2: A block diagram of the CHIME/FRB software pipeline. The various "levels" of the pipeline (L0−L4) are described in §2.2.1−§2.2.4. The L1 buffer, available in each FRB node, is for intensity data callbacks, and buffers data at various levels of downsampling for ∼240 s. The L0 buffer is for baseband callbacks, and buffers data for ∼30 s. Figure from CHIME/FRB Collaboration et al. (2018).

2.2.1 L0

CHIME/FRB uses a hybrid beamforming technique described in detail in Ng et al. (2017). While typical radio telescopes form targeted beams by coherently summing antenna signals, it was noted by Tegmark & Zaldarriaga (2009) that a series of regularly spaced antennas can take advantage of the fast Fourier transform (FFT) and have the number of correlator operations scale as $O(N \log_2 N)$ (where N is the number of antennas). This is much more desirable computationally than the $O(N^2)$ scaling of a traditional interferometer.

256 beams are formed N−S using the FFT beamforming technique, evenly spaced in $\sin\theta$, where θ is the zenith angle. Note that there is a chromaticity in

the location of these formed beams, which left unaccounted for, would completely smear the beam pattern spatially across the CHIME band. To account for this, the most sensitive set of beams closest to the desired beam-location on the sky are chosen per-frequency. While this does not perfectly overlap each beam on the sky as a function of frequency, one can do better by zero-padding the input data to the FFT in order to arbitrarily increase the number of formed beams (albeit at a computational cost). For CHIME/FRB, input data to the FFT is zero-padded to give a factor of two more beams, or 512 total, which aids in reducing the effect of spatial chromaticity (also referred to as beam "clamping") in the FFT beams. In the E−W direction, 4 beams are made using exact phasing, meaning the final number of beams formed by CHIME/FRB is $256 \times 4 = 1024$. Daily telescope gain solutions are produced for each beam during the transits of steady radio sources Cygnus A, Taurus A, Cassiopeia A, and Virgo A. The gain solutions are always phase-referenced to Cygnus A.

The last part of the L0 pipeline is up-channelization. Though input baseband (i.e. raw antenna voltage) data have 1,024 frequency channels, for the purposes of FRB searching, it is better to have higher frequency resolution. This is because of dispersion smearing, which is the effect of a dedispersion transform leaving residual DM delay within one frequency channel. Three sets of 128, 2.56-μs baseband samples are collected, each individually Fourier transformed, and recombined to result in output data with 16,384 frequency channels at a time resolution of 0.98304 ms.

2.2.2 L1

Each L1 node receives the full 16k frequency data for 8 beams. Thus, eight instances of the L1 process are run on each node. The collective process of RFI mitigation, dedispersion, and detection is known as L1a, while the process of identifying pulse candidates and grouping significant peaks in the pulse detection phase space is known as L1b.

After initial packet assembly and buffer management, each intensity array is sent to the RFI mitigation algorithm. CHIME/FRB's initial RFI mitigation strategy consists of a number of "transforms" which iteratively operate on intensity data. The series of transforms currently used by CHIME/FRB consists of nine "clippers" and two "detrenders". The clippers apply clipping transforms, which remove statistical outliers in both frequency and time space. Five of the clippers apply a 3σ threshold, while four apply a 5σ threshold. The nine clippers are iterated six times before sending the output to two detrenders, which remove polynomial trends in both frequency and time. Finally, after applying the initial set of transforms, the entire process is repeated once more before the data advance to the dedispersion algorithm. The details of the RFI chain sequence and its thresholds were determined through a set of empirical tests using data from the CHIME Pathfinder telescope (Bandura et al., 2014), as well as data from CHIME/FRB once it was available.

The FRB burst search algorithm, `bonsai`, is the most computationally expensive part of the L1 process. At its core, `bonsai` is a tree dedispersion algorithm (outlined first by Taylor, 1974). However, there are a couple major differences. `bonsai` regrids the intensity data into ν^{-2}, time space, meaning dispersed pulses will appear as

DS factor	Widths searched (ms)	Max DM (pc cm^{-3})	ΔDM (pc cm^{-3})
1	$[1, 2, 3, 4] \times 0.98304$	1656.13	1.62
2	$[1, 2, 3, 4] \times 1.96608$	3312.27	3.23
4	$[1, 2, 3, 4] \times 3.93216$	6624.55	6.47
8	$[1, 2, 3, 4] \times 7.86432$	13249.11	12.94
16	$[1, 2, 3, 4] \times 15.72864$	13249.11	25.88

Table 2–2: An overview of the dedispersion trees searched in `bonsai`. The downsampling (DS) factor is shown for each tree, along with the widths searched in that tree, and finally the maximum DM searched and the tree's corresponding error (ΔDM) to two decimal places. Note the downsampling factors, the width of the DM search bins, and the maximum DM of each tree is configurable.

straight lines rather than curves. The tree dedispersion implementation thus sums over these straight-line tracks. Besides the base 0.98304 ms time resolution tree, `bonsai` has four more downsampling trees over which the dedispersion is performed, with each subsequent tree downsampling in time by a factor of two to a maximum time bin size of ∼16 ms. While for low DM pulses the dedispersion is performed in all trees, past certain threshold values of DM only the downsampled trees are searched due to memory restrictions.

`bonsai` searches over several other search parameters outside of DM: in all, the search is over a 4D parameter space consisting of (DM, β, t, W), where β, t, and W are spectral index, arrival time of the pulse, and width respectively. DMs are searched to a maximum of $\sim$ 13000 pc cm^{-3}. Widths are searched up to a maximum of $4 \times \Delta t_{tree}$ (where Δt_{tree} is the size of the time bin in the tree), which comes as a result of the `ab4` peak finding kernel for `bonsai` spanning anywhere from one to four time bins. Specifications of the widths and DMs searched in each dedispersion tree are shown in Table 2–2. Two values of β are searched, $\beta = \pm 3$. Though a flat β

isn't searched, the chosen values of β greatly improve SNR for pulses with emission located only in the top or bottom of the band, while maintaining near-optimal search conditions for flatter pulses.

Despite calculating SNRs for many pulse search parameters, writing out recovered SNRs from `bonsai` at its full resolution would be extremely computationally intensive. Instead, the array of SNR estimates in the 4D space are "coarse-grained" substantially in DM and arrival time, and the highest SNR candidate in these coarse-grained bins is selected. Coarse-grained events with SNRs greater than a tunable threshold (currently set to 8.5σ) are then vetted by L1b, which discriminates astrophysical candidates from RFI. L1b makes this classification by comparing the SNR of other nearby `bonsai` triggers in the DM-time plane, using a neural network trained on a set of RFI and astrophysical labelled data. A lightweight "L1 header" containing information about the coarse-grained DM and time, an RFI rating from L1b, and SNR slices around the maximum SNR bin in W and β.

Intensity data at L1 are buffered for ~ 240 s, which approximately corresponds to the dispersion sweep across the CHIME band of an FRB with the maximum search DM in `bonsai`. If an event qualifies as astrophysical at the end of the real-time pipeline (i.e. L4), then the intensity data are dumped to disk for further offline analysis. Raw voltage data from L0 are also buffered, but for a smaller amount of time due to memory constraints (~ 30 s). Events classified as astrophysical above a secondary tunable SNR threshold will also have their baseband data dumped to disk.

2.2.3 L2/L3

After processing at L1, each beam has its L1 headers sent to the L2/L3 node. L2/L3 takes L1 headers from every beam and first attempts to group them in DM, t, and sky position. A threshold is set in each of the parameters, and events whose parameters all match within the thresholds are grouped together. DM and time thresholds reflect the size of the coarse-grain L1 triggers, while spatial thresholds reflect the size of the CHIME/FRB beams. After grouping, the grouped set of L1 headers is brought together as an L2 header before being sent along to L4.

Once the L2 header is made, event positions are refined. While one could simply take the centre of the highest SNR beam to be the event location for multi-beam events, instead the detected SNRs are compared to what is expected from the beam model for both values of β. This process allows for much better position refinement, and for very high SNR detections, L2 header localization can in principle localize events to within $\sim 0.3'$ (CHIME/FRB Collaboration et al., 2018).

L2 also performs its own RFI excision on L2 event headers. This method focuses on the SNR distribution of the L2 header in neighbouring beams. Astrophysical events are expected to be in the far-field of the telescope, and should have sharply focused SNR distributions within grouped beams. RFI events, on the other hand, are typically terrestrial in the near-field having much broader SNR distributions within grouped beams. A more detailed explanation of L2 RFI mitigation may be found in CHIME/FRB Collaboration et al. (2018).

After grouping, the *known source sifter* determines whether an event should be associated with a known source. The location of the L2 event in DM and sky location space is compared to a database of known sources. Events are checked against pulsars in the ATNF catalogue (Hobbs et al., 2004), the rotating radio transient catalogue (or "RRATologue"[7] , the FRB catalogue (Petroff et al., 2016), and CHIME/FRB's own previous detections. Events grouped with previously detected FRBs or known repeaters are classified as repeaters, and have a lower threshold for intensity callback (8σ) than FRB candidates as the association with a repeating source makes a strong argument for the astrophysical origin of the event.

If there is no connection to a known source, the pulse's DM is compared to Galactic DM models (Cordes & Lazio, 2002; Yao et al., 2017) to determine if the event is likely extragalactic. For events determined as unknown, Galactic sources, no intensity data are called back. This is because a large number of bright pulsars can be detected far off meridian, and such events are difficult to associate with a known source. Callbacks for unknown, Galactic sources would be too difficult for the system to manage given these false positives.

The last step in L2/L3 is to determine "actions" for each event. These are tunable actions the system will perform on an event based on the criteria it fulfills. For instance, intensity callbacks and baseband callbacks are queued for unknown extragalactic sources above a certain SNR threshold. These actions can (and have) been changed to allow for flexibility as improvements are made to the system. As an example, the threshold for intensity callback of unknown extragalactic candidates was originally 10σ, and has since been lowered to 8.5σ. Even more specific actions

[7] http://astro.phys.wvu.edu/rratalog/

can be requested, such as how intensity data dumps are determined from a less stringent range of location and DM for the magnetar SGR 1935+2154, given that the source has proven extremely interesting. More functionality will be added in the future, such as the ability to send out alerts to the FRB community for interesting events.

2.2.4 L4

Once L2/L3 headers are sent to the L4 node, L4 implements the actions requested from L2/L3. Specifically, its primary job is to request intensity data callback from all relevant beams (including non-detection beams surrounding those in which an event was detected). L4 also hosts the CHIME/FRB archive, which stores the more than 100 million L1 event headers which have passed through the system to date, regardless of their classification.

L4 also hosts a web display for candidate FRB events and browsing databases. Notably, events are not labelled as FRBs until they pass a human verification system, where two system monitors (or "tsars") must classify the event as a guaranteed FRB candidate. This is to avoid false-positives from RFI or off-meridian known sources. To aid in the classification process, there is an online dynamic spectrum plotting tool which can be used to inspect the intensity data. The products of the CHIME/FRB intensity data pipeline outlined in Chapter 3 are also available to browse and aid in decision making.

CHAPTER 3
Fitting CHIME/FRB Intensity Data

While Chapter 2 gave an overview of how intensity data dumps are triggered for promising FRB candidates passing through the live CHIME/FRB datastream, the output parameters given in the lightweight detection headers are not final. The real-time pipeline represents only the process that leads to a detection trigger. The intensity data pipeline, on the other hand, represents a set of processes that are run on dumped intensity data to recover the best-fit parameters for each FRB.

The intensity data pipeline is an automated pipeline which processes events that have been classified as FRB candidates from the real-time pipeline. The raw data from the telescope are stored on the CHIME/FRB archiver, which is a high powered storage server offering $\sim$ 750 TB of raw storage on 16 TB drives. Data are saved in compact msgpack[8] format, with 16,384 frequency channels and 1,024 time samples ($\sim$ 1 s) per data chunk. Raw data are then read from the archiver and processed by the intensity data pipeline for parameter determination. The pipeline itself is discussed in §3.1, while the individual analysis components of the pipeline are discussed in §3.2–§3.4. Major contributions from the author in this chapter include

[8] https://msgpack.org/index.html

helping develop the intensity airflow pipeline (§3.1) and being the primary developer of the DM pipeline (§3.2).

3.1 The CHIME/FRB Intensity Pipeline

The CHIME/FRB intensity data pipeline is a set of processes which read intensity data, clean them, reduce them to a more usable format, and fit them for burst parameters. An overview of the pipeline is shown in Figure 3–1. The intensity pipeline uses a framework called Apache Airflow[9] to schedule tasks and monitor the status of the pipeline.

Airflow is a powerful orchestration tool, and its flexibility allows for easy integration of tasks into a pipeline workflow. Airflow is scalable, meaning it's simple given enough computing power to queue up arbitrary numbers of workers. Work is organized using a structure known as a *Directed Acyclic Graph* (or DAG). In Airflow, DAGs are made as python scripts which represent tasks and task dependencies as code. The intensity pipeline is simple, as it only has four tasks:

1. Running a DM optimization script (described in §3.2) which processes, cleans, and reduces the raw intensity data into a more usable format (known hereafter as *cascade* files), then optimizes DM using an SNR optimizing algorithm and a burst structure optimizing algorithm. Two cascade files are made: one containing data for only the beam which had the highest SNR detection, and one containing data for all detection beams and non-detection beams on the detection beams' perimeter.

[9] `https://airflow.apache.org/docs/stable/`

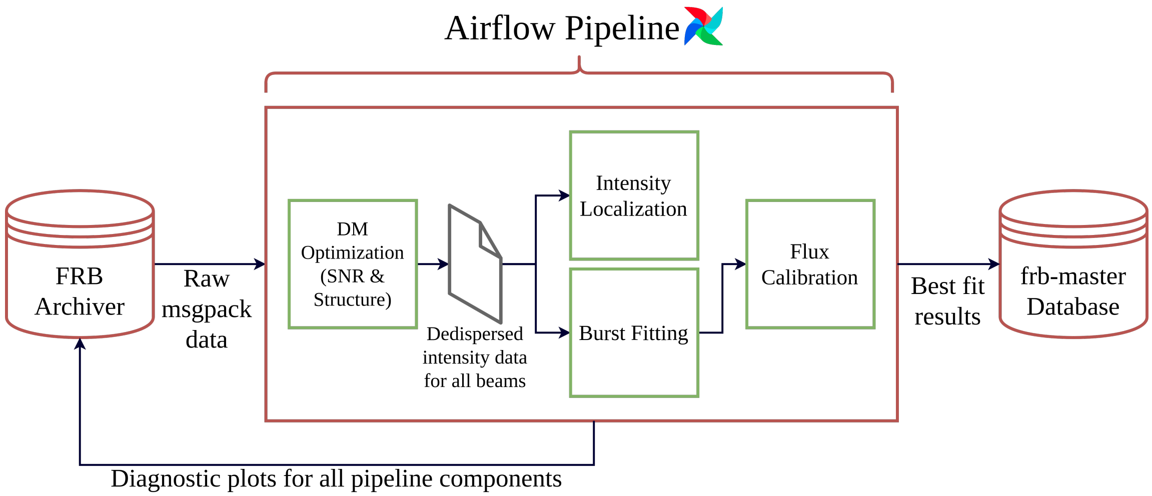

Figure 3–1: A block diagram of the CHIME/FRB intensity data analysis pipeline. Each of the main components in the Airflow pipeline (DM optimization, full burst fitting, and flux calibration) are described in detail in §3.2−§3.4.

2. Reading the all-beams cascade file and running a localization strategy which fits the CHIME/FRB beam model (described in §4.3) to each beam, and uses the best-fit beam location as an estimate of the event's sky position.
3. Running the CHIME/FRB burst fitting script, `fitburst` (described in §3.3), on data from the highest SNR beam. `fitburst` uses χ^2 optimization to obtain the best-fit DM, spectral index β, spectral running ζ, intrinsic width $W_{\rm int}$, and scattering time at 600 MHz $\tau_{600\ \rm MHz}$.
4. Calibrating the data from the highest SNR beam into physical energy units using CHIME/FRB calibration spectra of transiting steady radio sources and then determining the burst fluence and peak flux (described in §3.4).

As shown in Figure 3–1, tasks 2 and 3 (intensity localization and burst fitting) both depend only on the completion of task 1 (DM optimization), so it is possible

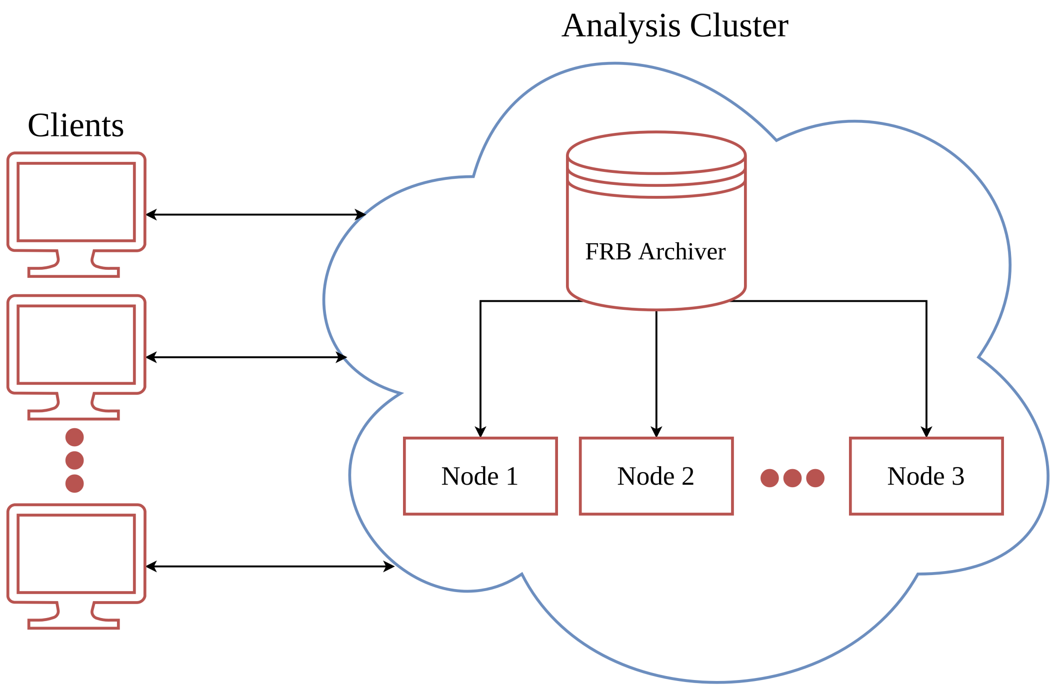

Figure 3–2: A schematic of the CHIME/FRB analysis cluster. Clients interact with the cluster through HTTP requests. Clients can request for analysis jobs to be spawned, as well as for the retrieval of information stored in CHIME/FRB databases.

to have them running in parallel. The ease of constructing parallel tasks in Airflow makes it a powerful tool for offline analysis. The intensity data pipeline is designed to run every five minutes, checking for any CHIME/FRB event IDs classified as unknown, extragalactic sources which have not yet had their intensity callback data analysed. Ten events may be processed simultaneously, with the flexibility to scale the pipeline up to multiprocess arbitrary numbers of events.

Each task in the intensity Airflow pipeline is run on the CHIME/FRB analysis cluster. The cluster consists of ten physical worker nodes, which are load balanced by a manager node. The manager node assigns jobs to the worker nodes, balancing

CPU and memory usage. Cluster jobs can be broken down into various software levels:

- **Containers:** A container is a standalone unit of software which has all code and dependencies wrapped into a software image which can be run reliably from any computer environment.
- **Tasks:** A task consists of a container image and commands to be run inside the container image when launched on the cluster. While containers provide an appropriate runtime environment for code to be run, they don't perform work on their own. Tasks, on the other hand, can be thought of as the atomic unit of work on the analysis cluster.
- **Stacks:** A stack is an orchestration of tasks which run concurrently on one or many of the cluster nodes. Stacks can be used to run the same service multiple times (e.g. to have multiple instances of a database in case one copy fails), or to run multiple tasks which might rely on one another (e.g. servers with different code dependencies running on separate containers). Stacks are self-healing, so will attempt to revive themselves on failure.

Containers, tasks, and stacks all exist within the Docker[10] framework. Docker is an OS-level virtualization service that is used to make CHIME/FRB container images. The intensity pipeline runs as a Docker stack, which requests the previously listed analysis tasks to be deployed as jobs on the cluster. All requests for tasks to be performed on the cluster are managed through the CHIME/FRB *Application*

[10] https://www.docker.com/

Programming Interface (or API) known as `frb-master`. `frb-master` accepts HTTP requests to perform actions such as retrieving information from databases hosted on the cluster or spawning tasks.

The web display page for candidate FRB events (see §2.2.4) contains all of the diagnostic information for intensity pipeline fits. Tsars validate and save pipeline output if there are no issues with the processing, but also have the ability to re-run each pipeline component with manual input parameters in cases where the processing is inadequate. Once tsars are satisfied with the quality of data from the intensity pipeline, they mark the output as finalized for science use.

3.2 Data Processing & Initial DM Fitting

The CHIME/FRB data processing and DM fitting pipeline (or simply the DM pipeline) is the first task run in the intensity pipeline. The first step in the DM pipeline is RFI excision and data cleaning. A brief overview is given here of the processing applied to the intensity data:

- The raw msgpack data are unpacked into a set of 2D arrays containing intensities and weights.
- Any samples with zero intensity are set to `nan` (not-a-number) values, since these samples likely arise from missed packets.
- A set of statistical measures is used to remove overly noisy frequency channels in the intensity data. First, the kurtosis, skew, and coefficient of variation are calculated for each frequency channel. Then, if any of the per-channel statistics are more than 3σ away from the median of the respective statistic calculated across all channels, then the channel is removed.

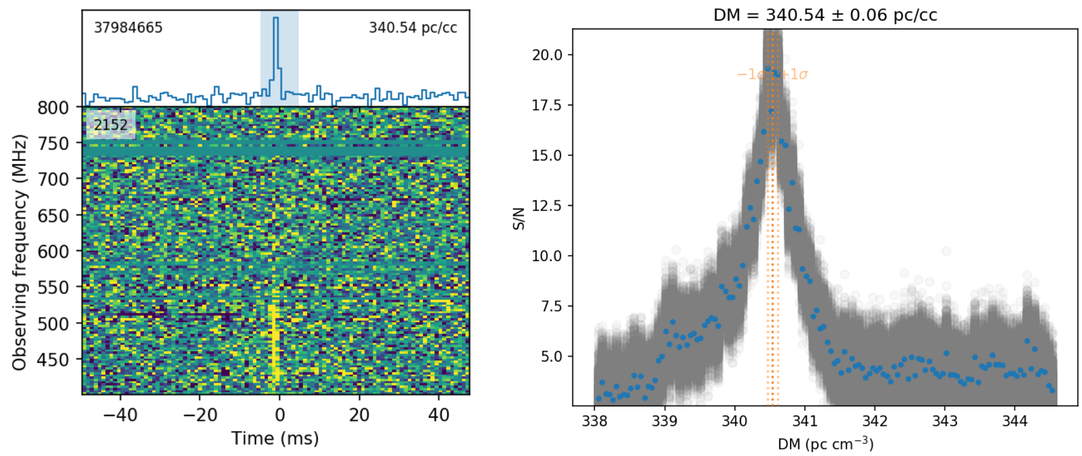

Figure 3–3: **Left:** An example of a burst waterfall from the intensity analysis pipeline. The burst waterfall shows the L1 event number in the top left of the frequency summed timeseries, the best-fit SNR optimizing DM in the top right of the timeseries, and the detected beam number in the top left of the dynamic spectrum. **Right:** A diagnostic plot for the SNR optimizing dedispersion algorithm from the intensity analysis pipeline. Each blue point shows the SNR for each DM trial, and the grey points show the 1000 perturbed SNRs for each DM trial. The dotted orange lines show the best-fit DM and the 1σ error region.

- Data are trimmed in time so that only a reasonable length snapshot around the dedispersed pulse remains.

After data processing, the best-fit DM is determined through two algorithms: an SNR optimizing algorithm, and a burst structure optimizing algorithm.

The SNR optimizing method uses a 1D boxcar convolution to determine DM. Boxcar convolution takes a boxcar kernel $g_W(t)$ (having width W in units of time samples) and convolves it with a 1D timeseries $f(t)$. In this case, the 1D timeseries is the frequency summed intensity data dedispersed to a given trial DM, $f_{\mathrm{DM}_i}(t)$. Each trial DM (DM_i) is determined by making a set of dedispersion transforms centred on the initial L1 pipeline DM within twice the L1 DM uncertainty and spaced equally by 0.025 pc cm^{-3}. This spacing is chosen because it corresponds to a DM sweep of $\sim$ 0.5 ms across the CHIME band, which is approximately the length of half of a time sample. Thus by Nyquist's theorem, all the available information in the time-series will be retrieved with this DM trial spacing. A set (T_W) of reasonable boxcar width trials W_j are also used for each DM trial. Succinctly, the SNR optimization algorithm for each DM trial can be summarized as:

$$\mathrm{SNR}_i = \max[f_{\mathrm{DM}_i}(t) * g_{W_j}(t),\ \forall\ W_j \in T_W], \tag{3.1}$$

where SNR_i is the SNR for DM trial DM_i, W_j is the trial width pulled from the set T_W, $f_{\mathrm{DM}_i}(t)$ is the frequency summed, dedispersed timeseries for the given DM trial, and $g_{W_j}(t)$ is a boxcar spanning the trial width. Then, instead of just picking the highest SNR_i and calling its corresponding DM_i the optimal DM, a slightly more complex method is used to find the optimal DM which allows for rough DM error estimation:

1. For each SNR_i, a number is drawn from a standard normal distribution, resulting in a perturbed SNR denoted SNR'_i
2. After determining each SNR'_i, the DM_i maximizing the perturbed SNR is added to the set T_{DM}
3. Steps 1 and 2 are repeated for a total of 1000 trials
4. The mean of T_{DM} is chosen as the optimal DM with a 1σ error given by the standard deviation of T_{DM}

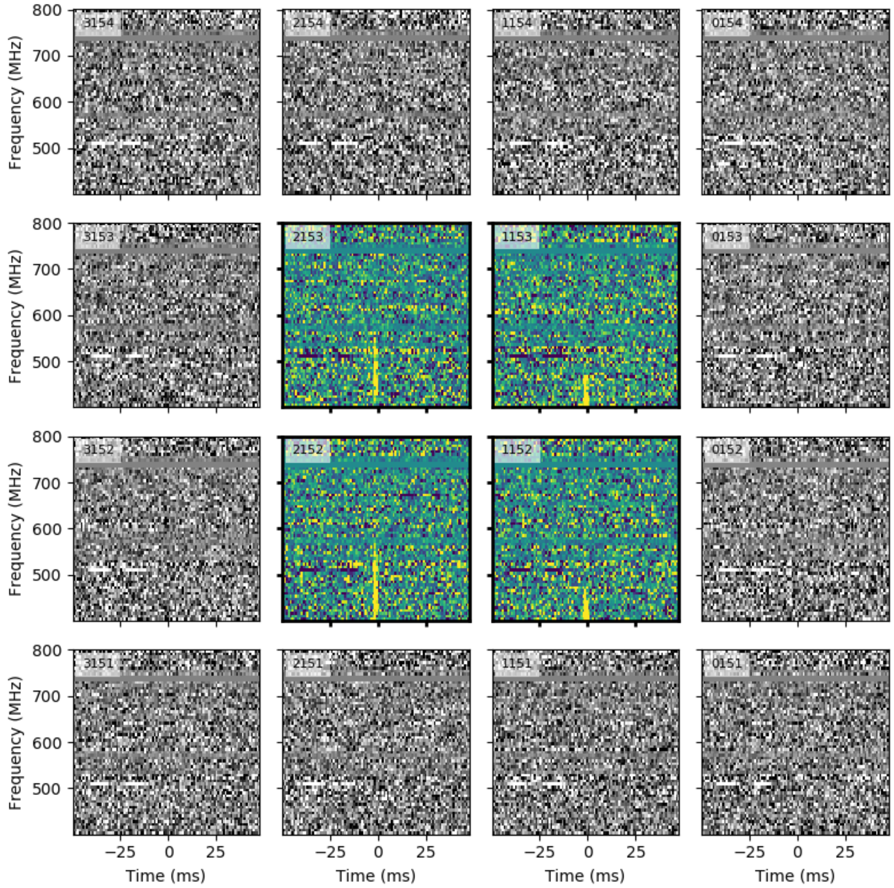

Figure 3–4: An example output "all-beams" waterfall from the intensity pipeline for the same event as in Figure 3–3. Coloured dynamic spectra represent beams wherein there was a detection, and greyscale dynamic spectra represent beams for which there are callback data but no detection. The spectral modulation effects from the different beam columns are clearly visible, as the burst appears more narrow-banded in the 1000 beam column than the 2000 beam column.

The burst structure optimizing dedispersion algorithm[11] is more complicated in concept than the SNR optimizing dedispersion algorithm, and is only considered the

[11] https://github.com/danielemichilli/DM_phase

more useful expression of DM when FRBs show complex frequency-time structure (see §1.1.5). Since the use of the structure optimizing DM determined by the DM pipeline is still being tested and currently only gives reasonable output for very high SNR bursts, the algorithm will not be outlined here, as a detailed description is given elsewhere (Seymour et al., 2019).

After the optimal DMs are determined, the cascade cutout of the pulse at the SNR-optimizing DM is saved on the CHIME/FRB archiver. Diagnostic plots of each DM optimization method are saved as well for debugging purposes (see Figure 3–3 for an example). Then, all beams for which the event has intensity callback (which includes all detected beams and non-detection beams on the perimeter of the detection beams) have their data dedispersed and trimmed to the best-fit DM. The 3D array of intensity data over beam number, frequency, and time is then also saved on the archiver for easy access to intensity data in all available beams. Finally, a series of waterfalls showing the dedispersed snapshot surrounding the burst in all available beams is made (see Figure 3–4), which helps intuitively show the effect of the beam model on the FRB emission.

3.3 Full Burst Fitting

The full burst-fitting algorithm used by CHIME/FRB, `fitburst`, is automatically run on each event after the initial DM fitting. The algorithm uses least-squares

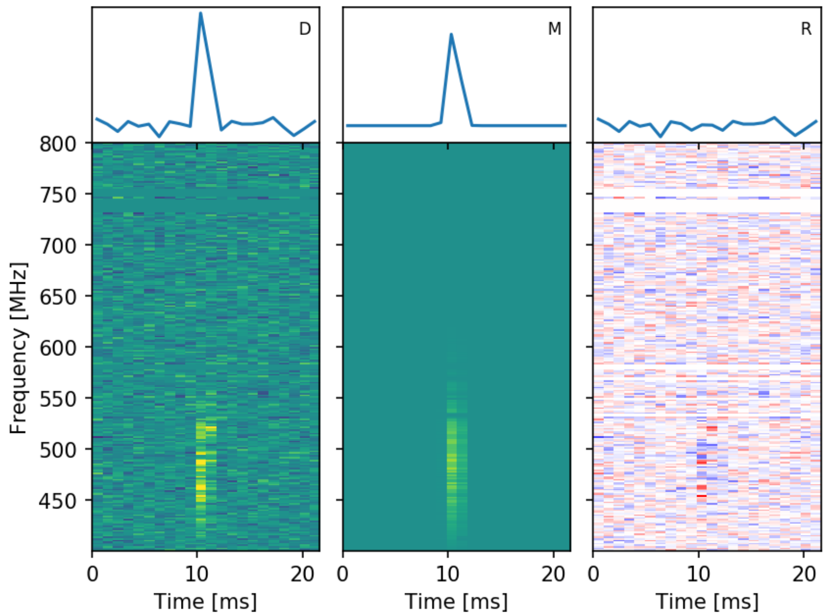

Figure 3–5: An example diagnostic plot from `fitburst` showing the burst data (D), the best-fit model (M), and the resulting residuals (R) for the same event as in Figure 3–3. Since `fitburst` doesn't fit the beam model and FRB spectra tend to be more complex than the simple running power law given in Equation 1.11, some structure in the residuals is expected.

fitting[12] with the Trust Region Reflective algorithm to determine the best-fit parameters of the burst. The algorithm has been described in detail elsewhere (e.g. CHIME/FRB Collaboration et al., 2019a; Josephy et al., 2019; CHIME/FRB Collaboration et al., 2019b) and continues to be refined, so will be described only briefly here.

[12] https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.least_squares.html

`fitburst` fits each event twice, once for a set of parameters assuming no scattering, and once for a set of parameters including scattering. The full set of parameters searched by fitburst are DM, spectral index β, spectral running ζ, intrinsic width W_{int}, scattering time referenced to 600 MHz $\tau_{600\ \mathrm{MHz}}$, and burst arrival time referenced to 600 MHz $t_{600\ \mathrm{MHz}}$. The scattering index α is fixed to -4, and pulse dispersion is assumed to be exactly proportional to ν^{-2}. An initial guess for DM and W_{int} is taken from the DM pipeline best-fit values, while default β and ζ assume a flat spectrum. Least-squares optimization is then used to calculate the best parameters by minimizing the residuals between the model function and the data. Reasonable bounds on each of the model parameters are given to restrict the algorithm to appropriate parts of parameter space. Once a fit has been determined, results are saved in a database, and a diagnostic plot showing the data, model, and residuals for each of the unscattered and scattered fits are shown (see Figure 3–5 for an example).

Since the unscattered and scattered models are nested, the scattered model will almost always give a better fit (i.e. higher SNR) given it has more parameters. So, in order to determine whether there really is significant scattering, the F-test statistic is calculated. The F-test is often used to determine whether the simpler of two nested linear models is statistically favoured. Taking the null hypothesis that the scattered model does not provide a significantly better fit than the unscattered model, the F-test statistic is calculated as:

$$f = \frac{\left(\chi^2_{\mathrm{us}} - \chi^2_{\mathrm{s}}\right) / \left(p_{\mathrm{us}} - p_{\mathrm{s}}\right)}{\chi^2_{\mathrm{s}} / \left(n - p_{\mathrm{s}}\right)}, \tag{3.2}$$

where χ^2_{us} and χ^2_{s} are the χ^2 values from the least-squares fitting for the unscattered and scattered models, p_{us} and p_{s} are the number of parameters in the unscattered and scattered models, and n is the number of points in the array for which the models are calculated. If the value of the F-test statistic is small ($f < 0.01$), then the null hypothesis is rejected and the scattered fit for the FRB is reported. In the case the unscattered fit is preferred, the scattering in the scattered fit is taken as an upper limit.

3.4 Burst Fluence Calibration

One of the most complex parts of the intensity pipeline is burst calibration. This is the process of taking the telescope's digital beamformer (BF) units and converting them to physical energy units (Jy). Because of the unique shape of CHIME's reflectors and antennas, the sensitivity pattern of the telescope on the sky is extremely complicated, making it non-trivial to determine FRB fluxes and fluences. To get around the difficulty of modelling the beam, CHIME/FRB uses a method of calibration which uses steady source calibrators (e.g. Cygnus A, Taurus A, Cassiopeia A) transiting in various locations in the beam to obtain calibration spectra.

Calibration is run after full burst fitting, so that the region wherein to calculate the fluence is well defined for bursts with different widths and degrees of scattering. The calibration method is detailed in CHIME/FRB Collaboration et al. (2019a), so here only a brief overview will be given. The process is as follows:

- Intensity data are dumped for transits of steady source calibrators through their most sensitive transit beam. These intensity dumps are done daily when possible, however system instability means that steady source data aren't necessarily available for each daily CHIME observation.

- The intensity data for each calibrator source are used to determine calibration spectra by comparing the on-source and off-source spectra. The calculation for each calibration spectrum is given by:

$$S_{\nu,\ \text{BF-to-Jy}} = \frac{(S_{\nu,\ \text{Calibrator, on}} - S_{\nu,\ \text{Calibrator, off}})\ [\text{BF}]}{(S_{\nu,\ \text{FFT}} \times F_{\nu,\ \text{Calibrator}})\ [\text{Jy}]}, \tag{3.3}$$

  where $S_{\nu,\ \text{BF-to-Jy}}$ is the resulting BF to Jy spectrum, $S_{\nu,\ \text{Calibrator, on/off}}$ are the (BF) intensities of the on/off spectrum in the calibrator transit intensity data, $S_{\nu,\ \text{FFT}}$ are the FFT sensitivities at the location of the calibrator source at peak transit (in relative units), and $F_{\nu,\ \text{Calibrator}}$ is the model spectrum in Jy of the steady source.

- The "best" calibration spectrum (determined first by a burst's spatial proximity to a calibrator, then temporal proximity) is divided out of the burst's intensity data to convert from BF units to Jy units. Note that, even if an FRB goes off in a different beam than its chosen calibrator, no attempt is made to correct the FRB spectrum for the differing FFT beam sensitivities. Due to burst localization uncertainties, this correction would likely introduce unphysical spikes from FFT beam clamping (see §2.2.1).

Calibration data for each steady source are only taken within the full-width-half-maximum (FWHM) of the most sensitive beam column, so any bursts detected outside of that region will have larger fluences than what is predicted with this method. Thus, all fluxes and fluences from CHIME/FRB are reported as lower limits. An example of a diagnostic plot produced from the fluence pipeline is shown in Figure 3–6.

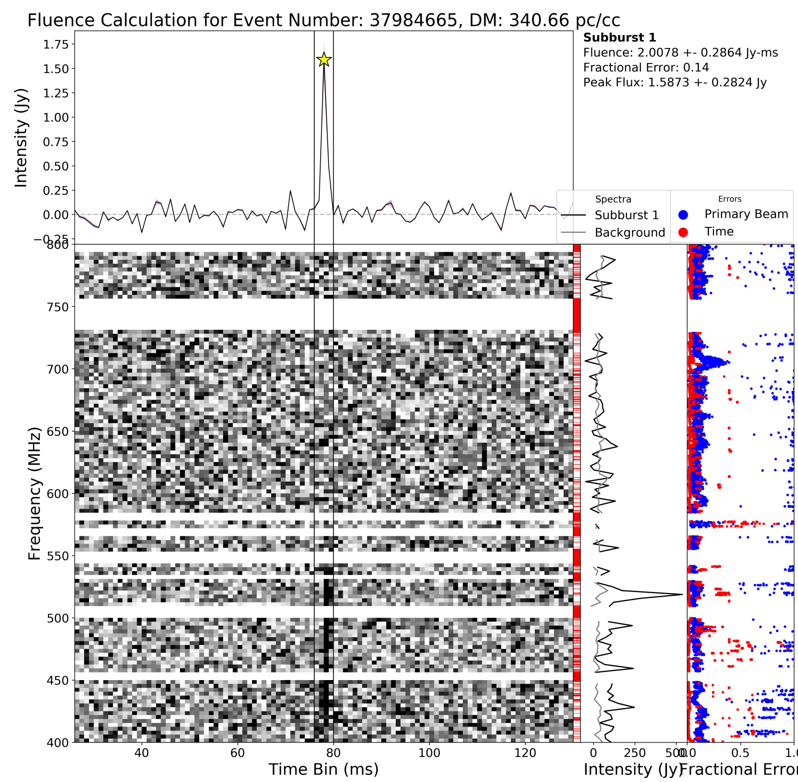

Figure 3–6: An example diagnostic plot from the calibration pipeline for the same event as in Figure 3–3. The region for which the fluence is calculated is shown with vertical black lines, and the peak of the time-series which is used to calculate the peak flux is marked with a yellow star. The red lines on the right side of the dynamic spectrum represent frequencies which have been masked for the fluence calculation at full resolution. The blue and red points on the right show the fractional error introduced in the flux and fluence calculation due to the position of the calibrator source in the primary beam ("Primary Beam", blue) and due to day-to-day variation in the calibration spectrum of the calibrator source ("Time", red).

CHAPTER 4
The Injection System

Thus far, Chapter 1 has given an overview of FRBs and what can be learned from their parameters, and Chapters 2 and 3 have given an overview of how CHIME/FRB triggers FRB detections and measures the burst properties. CHIME/FRB's sample of events is by far the largest coherent sample of FRBs thus far in the field. While CHIME/FRB's design allows for the highest FRB detection rate of all current surveys, like any telescope, it is subject to selection biases affecting our interpretation of the observed population of bursts. These biases come primarily from the fact that CHIME's beams are complicated, and from the RFI littering on-sky data.

In order to undo the biases which are introduced in the real-time pipeline, CHIME/FRB has developed a service which injects model populations of synthetic FRBs into the live detection pipeline. By injecting pulses over a broad set of parameters and tracking the (non-)detection status of injections, it is possible to construct a probability distribution of event detection, $P(\mathbf{X})$ (where $\mathbf{X}$ is the "true" set of FRB parameters), in a large multi-variable parameter space. This probability distribution, known as the *selection function*, is independent of the model FRB population. In principle, the selection function characterizes the CHIME/FRB search pipelines biases to all parameters in $\mathbf{X}$. While there are some examples of attempts at correcting observational biases with on-sky injection studies (Farah et al., 2019; Keane

& Walker, 2019), the full characterization of telescope selection functions with injections has not yet been discussed in the literature.

While the CHIME/FRB injection system is under continual development, it has reached a level of maturity capable of beginning to correct for some observational biases in the main set of FRB observables: dispersion measure (DM), fluence (S_ν), intrinsic width (W), scattering time (τ), and the spectral parameters β and ζ (see Equation 1.11). A preliminary, offline test of biases in the `bonsai` dedispersion algorithm for CHIME/FRB's first detections is given in §4.1, from which the live injection system grew. The injection system architecture is laid out in §4.2, while the methods of simulating bursts in the live pipeline are given in §4.3. The process of generating populations of FRBs for injection is given in §4.4, and finally an introduction to the formalism of using injections to correct for biases in the CHIME/FRB sample is outlined in §4.5. Major contributions from the author in this chapter include developing and running the initial offline injection analysis (§4.1), developing the injection API and `simpulse` workers (described in §4.2), verifying the calibrations of injections (§4.3.3), and organizing the delivery of a large scale injection sample ($N \sim 85,000$) used in selection function characterization.

4.1 Offline Injections

When formalising the results of CHIME/FRB's first 13 FRBs, it was apparent that the majority of detections were scattered (CHIME/FRB Collaboration et al., 2019a). The degree of scattering observed was so strong that, through population synthesis analysis focused on detailing FRB scattering properties in spiral host galaxies, putting FRBs in special, electron dense locations within the host was necessary

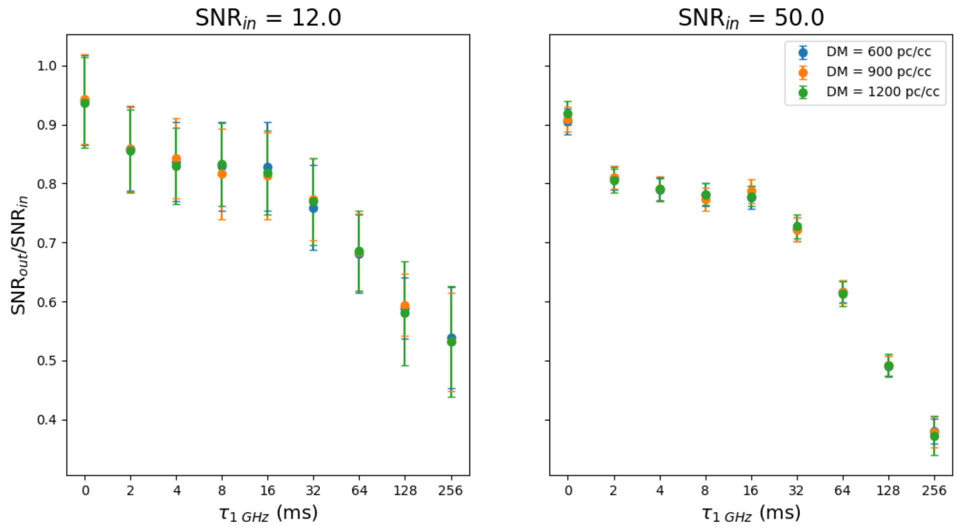

Figure 4–1: Summary of the offline injection simulations for various scattering times and DMs. The mean recovered SNR fraction ($\mathrm{SNR_{out}/SNR_{in}}$) is plotted for each trial, with the error bars showing the 1σ deviation for each set of 100 bursts. The degree of scattering heavily impacted the recovered SNR for bursts of a constant fluence, while various DMs in the range of CHIME/FRB's first 13 detections do not significantly change the recovered SNR for the same degree of scattering.

to explain the observed scattering properties. Scattering broadens pulses temporally (particularly at CHIME's low observing frequencies) which results in lower detection sensitivity from CHIME/FRB because `bonsai` only searches for FRBs with simple, boxcar-like kernels (see §2.2.2). Thus the amount of scattering in CHIME/FRB's first 13 detections was surprising, as the search pipeline should have, if anything, been biased against detecting scattered bursts.

To clarify the nature of scattering in these detections, an initial, offline series of simulations was run by the author to study potential biases in CHIME/FRB's dedispersion search. This analysis would ensure that the statistical scattering properties of the observed FRB population was not due to any potential bias *towards* detecting scattered events in `bonsai` through comparing the recovered L1 SNR of simulated bursts with various degrees of scattering. For the simulations, sets of bursts with known signal-to-noise ratios (assuming a Gaussian noise background), dispersion, and scattering were injected into a Gaussian noise background and run through an offline version of `bonsai`. Pulses were modelled using `simpulse` (see §4.3.1), and had 1σ Gaussian widths of 1 ms. Two input SNRs were used for each scattering and DM injection trial: one representing a population of "bright" bursts ($\mathrm{SNR_{in}} = 50$) and one representing a population of bursts near CHIME/FRB's detection threshold ($\mathrm{SNR_{in}} = 12$). Three trial DM values were used for injected bursts in order to probe any differences in detectability for pulses with DMs in the range of CHIME/FRB's first sample. Finally, a set of logarithmically spaced scattering times was used, spaced from $\tau_{1\ \mathrm{GHz}} = 2$ ms to $\tau_{1\ \mathrm{GHz}} = 256$ ms. A set of unscattered pulses was also injected as a reference.

For each set of parameters, 100 pulses were injected, and each of the detection SNRs recorded and saved. Figure 4–1 shows the results of the analysis. Evidently, CHIME/FRB's detection pipeline detects highly scattered FRBs with less efficiency than unscattered FRBs having identical intrinsic signal-to-noise. Thus, any strong scattering properties observed in the FRB population detected by CHIME are intrinsic to the population as opposed to being caused by a selection effect. This strengthened the conclusion in CHIME/FRB Collaboration et al. (2019a) that, because CHIME detected many FRBs with large degrees of scattering within the first 13 detections, they likely occupy special locations in their host galaxy.

While the offline injections allowed for the simple observation that `bonsai` alone is biased against detecting scattered pulses, it by no means characterizes all of the biases present in CHIME/FRB's detection pipeline. There are still complex effects not accounted for in these simulations, such as RFI (and its subsequent excision) in on-sky data. The development of an injection system attached to the real-time pipeline was necessary to encompass all of these effects.

4.2 System Architecture

Injections are orchestrated by an injection API, known as `mimic`. At its core, `mimic` manages injections in the real-time pipeline by interacting with L1 nodes via remote procedure calls (RPCs). Each L1 node is outfitted with an RPC server, which has two crucial functions for the purposes of injections. First, a request can be made to start (or stop) the duplication of on-sky data for one set of four intensity beams that are being processed by a given node. One of the 128 L1 nodes has been outfitted as a "receiver node" (L1′) which processes duplicated beams from the "sender node" using the same software as the other L1 nodes. Second, a request can be made to the L1′ node to add a synthetic FRB to one of the duplicated beams. The data from the receiver node are tagged with an arbitrary offset in their labelled beam numbers before being sent to L2/L3, which prevents the misclassification of injections as astrophysical signals. A summary of `mimic` as it interacts with the real-time pipeline is given in Figure 4–2.

Besides the API itself, the injection system consists of several components, explained in §4.2.1-§4.2.3: the *injection driver*, `simpulse` workers, the injections database, and the *injection snatcher*. A subset of these components are brought

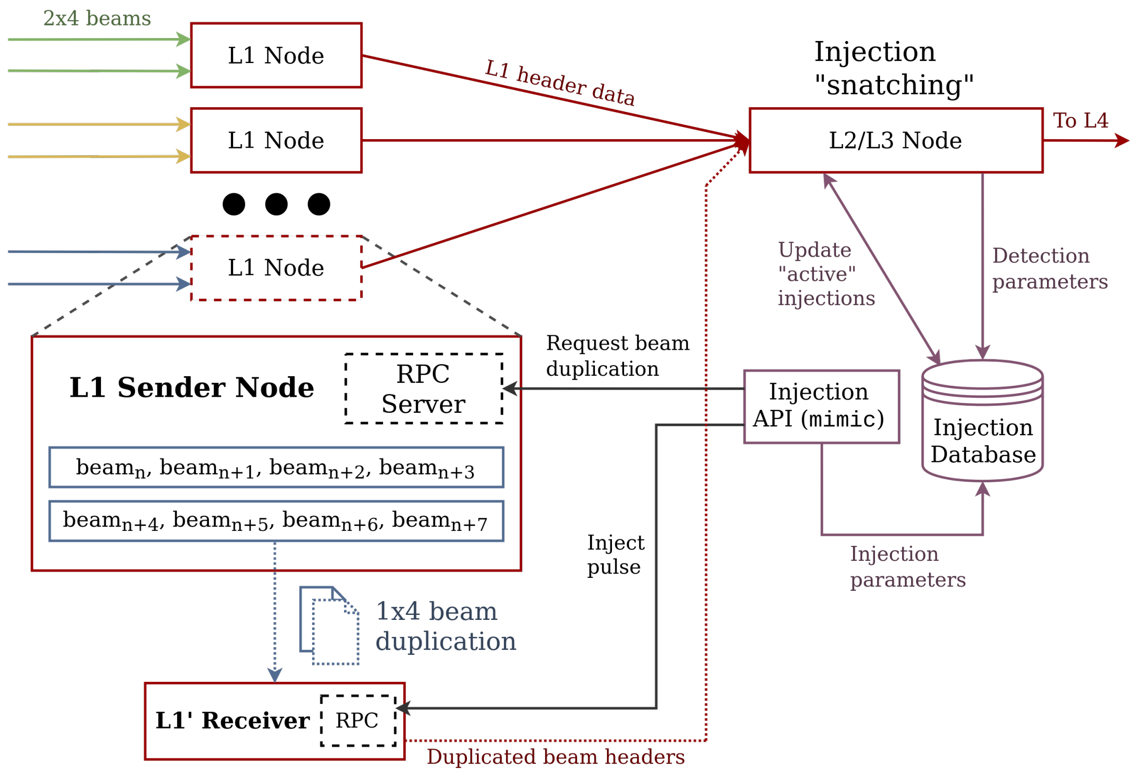

Figure 4–2: Schematic of the CHIME/FRB injection system. Each L1 node handles two datastreams containing full resolution intensity data for four beams. The injection API (`mimic`) interacts with the L1 nodes through RPC, and its capabilities are outlined in §4.2. The full injection system injects a population of FRBs into the on-sky datastreams of four beams from a given L1 sender node at a time. The injections' detection properties are then measured in the live pipeline, storing the injection and (non-)detection parameters in a database.

together as a Docker stack (see §3.1), which run persistently on the CHIME/FRB analysis cluster (see Figure 4–3).

4.2.1 The Injection Driver

While the injection API handles requests to start/stop beam duplication and inject pulses into duplicated beams, it does not make these requests on its own.

Once a population of FRBs has been generated for injection (see §4.4), a separate module known as the *injection driver* reads the resulting `hdf5`[13] *schedule file*, orchestrating the injections therein. The management of these injections is more-or-less straightforward, but with some important technical considerations.

The outgoing L1 sender network can only handle duplicating one set of four beams at a given time, and `bonsai` requires a moderate amount of time (∼several minutes) for its variance estimation of incoming data to reach a steady state. Given these two restrictions, the driver schedules injections in bundles of four beams. Each bundle's injections are run to completion before moving on to the next. This approach is not as ideal as fully randomizing the injections across all beams, because intermittent RFI could plague the detection statistics of a given set of beams. However, it is far more time efficient since each time a new set of beams is streamed to L1′ from a new L1 sender node, `bonsai`'s running variance estimation needs to stabilize. Thus, going through four sequential sender beams at a time minimizes this waiting time.

The injection driver handles four parallel threads at a time, known as *injection handlers*. Each injection handler is responsible for making requests to the injections API to generate pulses for one beam. The API, written in the `sanic`[14] framework, is naturally asynchronous and can handle simultaneous requests from the four injection handlers. Each handler keeps a running tally of the injections it has completed. This

[13] https://www.hdfgroup.org/solutions/hdf5/

[14] https://github.com/huge-success/sanic

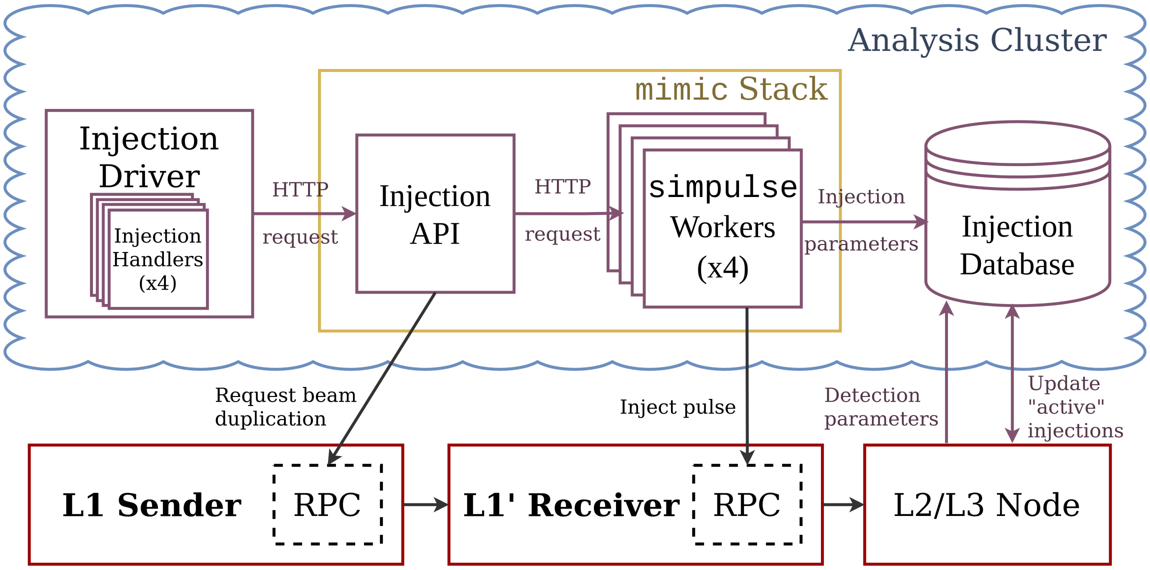

Figure 4–3: A detailed schematic of the individual components of the injection system and the `mimic` stack. HTTP requests are always handled with the python `requests` package.

tally is relayed to the injection driver, which in turn keeps track of each completed injection for a given schedule file. Such bookkeeping is important because networking issues or unplanned system shutdowns can otherwise interrupt injections.

4.2.2 `simpulse` Workers

Once given a request to generate an injection, the injection API cannot perform this action on its own. This is because the pulse modelling software used by CHIME/FRB, `simpulse`, runs only in `python2`, and the asynchronous `sanic` API framework used for injections only runs in `python3`. To get around this, a `flask`[15]

[15] `flask` is a lightweight web application framework: https://flask.palletsprojects.com/en/1.1.x/

API was created for simpulse which runs in python2. The simpulse API's sole purpose is to generate the data from injection requests to send to the L1′ receiver.

To keep up with the asynchronous requests from the injections API, four identical copies of the simpulse API are spawned in the mimic stack. Docker natively load-balances the requests made to each of the workers, allowing the simpulse workers to keep up with the incoming pulse generation requests. Besides the pulse generation functionality, there is an optional flag in the request to the simpulse API to allow for a dedispersed waterfall plot of the generated pulse to be made and stored on the CHIME/FRB archiver. These snapshots greatly aid in debugging the injection system, and help visualize the combined effects of the intrinsic pulse parameters and the beam modulation. An example pulse snapshot is shown in Figure 4–4.

4.2.3 The Injection Database

Injection and detection parameters for injections are stored in a RethinkDB[16] database hosted on the CHIME/FRB archiver. The database itself is not a part of the mimic stack, but rather is built as a part of frb-master (see §3.1), which already functioned to interact with databases at the beginning of the injection system's development. Each injected population is labelled with a unique program name, which makes the full results of any given run easily retrievable.

The injection database's role is to store both the injection parameters and header detection parameters, which are given in Table 4–1. RethinkDB is a good choice of architecture for the injections database because it is malleable – if the capability of

[16] https://rethinkdb.com/

Injection Parameters	Detection Parameters
Expected Arrival Time (μs precision)	Detection Time (μs precision)
Dispersion Measure (pc cm^{-3})	`bonsai` DM (pc cm^{-3})
Fluence (Jy ms)	`bonsai` SNR
Width (Gaussian FWHM, ms)	`bonsai` Width (bins)
Spectral index (β)	`bonsai` spectral index ($\beta' = \pm 3$)
Spectral running (ζ)	L2 RFI Grade (0-10)
$\tau_{1\text{ GHz}}$ (ms)	
Beam number	
Beam coordinates	

Table 4–1: A list of the injection parameters currently available as input to `mimic`, and relevant detection parameters for injections detected in the realtime pipeline.

injections is expanded with new input parameters (e.g. scattering index α), then previous database entries will remain unaffected. Within RethinkDB it is also possible to tag database entries with an expiration time. This is useful for injections because it allows a list of active injections to be retrieved, with injections expiring if undetected by the injection snatcher within the expiration time. Undetected injections are removed from the list of active injections, and labelled as non-detections.

4.2.4 The Injection Snatcher

One of the most important aspects of the injection pipeline is properly flagging injections as they come through the real-time pipeline, labelling them as either detected or not detected. The injection snatcher, which acts as a module in the L2/L3 pipeline, fulfills this purpose.

Largely built from the framework of the known source sifter (see §2.2.3), the injection snatcher inspects all incoming L2 headers and compares their expected arrival times and DMs to the list of active injections. If the arrival time of an L2 header matches within half a second of an active injection and if the DM of that same L2 header also matches within twice the DM error of an active injection, then the detection is taken as a match to the injection. The detection parameters for the event are then sent to the injection database and matched with the corresponding injection parameters.

During the construction of the injection system, the injection snatcher was tested for a set of several thousand simple, bright bursts to ensure its efficacy. The snatcher successfully retrieved these bursts with nearly 100% detection efficiency, assuring its ability to match active injections with the corresponding detections.

4.3 Simulating FRBs

A crucial part of making sure the injection system simulates realistic on-sky FRBs is the process of properly modelling both the FRB itself and the effects of the system. In the case of CHIME/FRB, injections are input after beamforming, but before RFI mitigation. Thus, beamforming effects must be forward modelled, while RFI mitigation effects occur naturally in the live system. The software used to model FRBs in the injection system is described in §4.3.1, while the beam model and the methods used to calibrate from input Jy units to telescope BF units are described in §4.3.2 and §4.3.3.

4.3.1 `simpulse`

The `simpulse` library generates dispersed, scattered FRBs using the frequency-time relationships given in Equation 1.2 and Equation 1.6, with the intrinsic burst widths assumed to be Gaussian. While the process of generating a dispersed pulse with a Gaussian width then convolving with a one-sided scattering exponential is simple, `simpulse` also models dispersion smearing (see §2.2.1) and other such channelization effects, ensuring a realistic FRB signal.

`simpulse` makes all the same assumptions about the FRB frequency and time spectra that `fitburst` makes. This includes assuming a scattering index of -4, a pulse dispersion proportionality of exactly ν^{-2}, and the use of a running power-law frequency spectrum (Equation 1.11). This allows bursts modelled by `fitburst` and bursts generated by `simpulse` to be directly compared.

4.3.2 Modelling the Beam

Outside of just an FRB's intrinsic properties, the injection system models the combined spectral signatures of the CHIME telescope's primary beam and the FFT synthesized beams (see §2.2.1). While the effect of the FFT synthesized beams is known precisely and can be calculated, this is not true of the telescope's primary beam, which arises as a result of the telescope shape and antenna design. To model the telescope primary beam, the CHIME/Cosmology team developed a method which fits a coupled antenna model using a cylindrical reflector to holography data of steady source and solar transits. The results of the fit were saved to a file, which has been interpolated into a beam model module used for sensitivity calculations with CHIME/FRB (Hinshaw et al., 2020).

The beam model uses an (x, y) coordinate system, where x represents hour angle and y represents degrees North from zenith (where zenith is the origin of the coordinate system). So, an (x, y) coordinate can be seen as a transformation of $(\mathrm{ra}, \mathrm{dec})$ for a specific date and time. CHIME/Cosmology references its calibration to the position of CygA in the beam at meridian, $(x_{\mathrm{CygA}}, y_{\mathrm{CygA}})$, meaning the 16,384

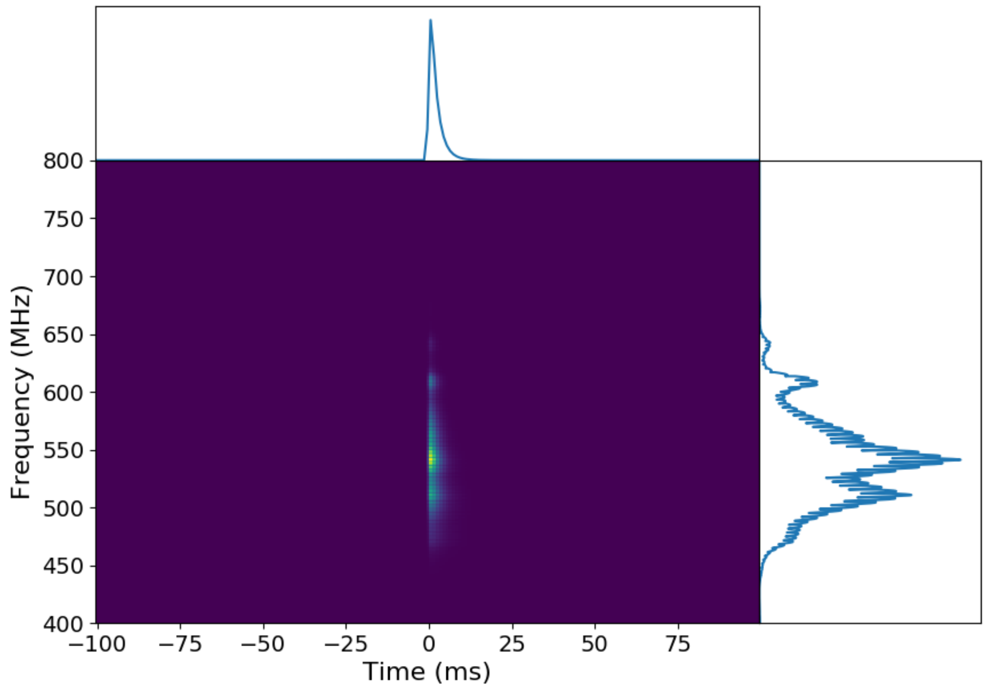

Figure 4–4: An example waterfall plot from the `simpulse` API for an arbitrary injection. Summed time-series and frequency spectra are shown above and to the right, respectively. The frequency spectrum shows the complex modulations induced by the beam model, in particular the periodic spikes induced by FFT beam clamping (see §2.2.1).

beam sensitivities as a function of frequency is an array of ones at that location. From there, the sensitivity arrays are in units relative to the sensitivity at that location, with frequency averaged sensitivities at the centre of each of the 1,024 beams ranging from $\sim 0.05 - 1.1$. The full array of sensitivities from the beam model are passed as input to the `simpulse` workers, and modulate the pulse before it is sent to duplicated beams for injection.

4.3.3 Calibrating Injections

A critical step in the process of simulating injections is ensuring that the input fluence for an injected burst is correct, such that bursts are being injected in meaningful, physical units. Since the duplicated beams have datastreams in BF units by definition, this means taking the input fluence for an injection and converting it from Jy ms to BF ms. Because CHIME/Cosmology calibrates daily on steady sources, intensity data for CHIME/FRB are calibrated up to a constant multiplied by a factor taking into account the number of antennas labelled as bad input to the beamformer. The total conversion factor is (Masui, 2020):

$$C = \frac{(1024 f_{\rm good})^2 \times 128}{16 \times 0.806745 \times 400} \ \text{BF units} \ \ \text{Jy}^{-1}. \tag{4.1}$$

Each of the numerical factors in Equation 4.1 comes from careful bookkeeping in the CHIME/Cosmology calibration process. The factor $f_{\rm good}$ is the fraction of good antenna feeds, and typically ranges from 0.9 to 1.0.

Given the calibration factor C, data can be converted between CHIME/FRB BF units and Jy units. When CHIME/FRB began its pre-commissioning phase, the input data from L0 to L1 did not account for the factor $f_{\rm good}^2$. Since the value $f_{\rm good}$ is time-varying, the L1 intensity data could not be calibrated to a static factor in real-time. However, in April 2020, a correction for the factor $f_{\rm good}^2$ was applied to the script which manages gain conversions, and thus CHIME/FRB intensity data became calibrated in real-time up to a static factor and beam model. The use of this calibration constant was tested against CygA transit data and compared to CygA's expected Jy spectrum. The calibrated flux was consistently within $\sim$5% of

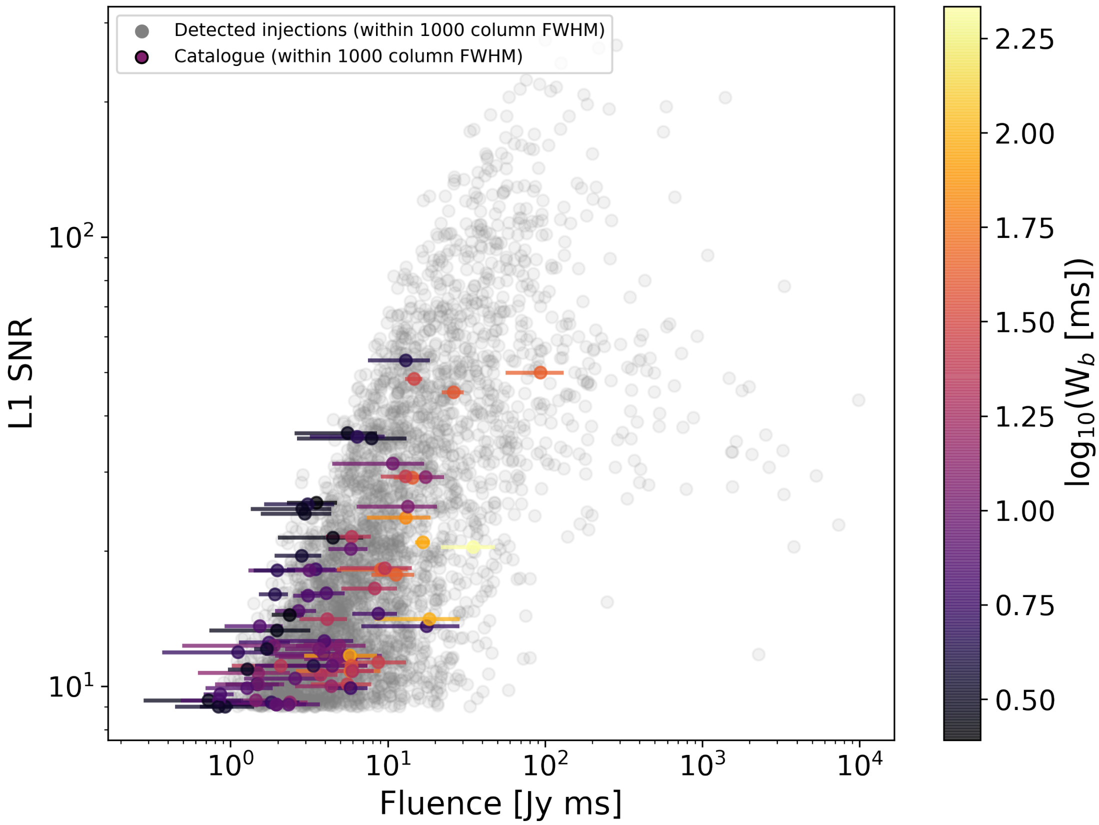

Figure 4–5: A comparison of the fluence and L1 SNR relationship for injected pulses (grey) and catalogue FRBs detected within the FWHM of the most sensitive beam column (colour). Catalogue bursts are coloured by their broadened width. Error bars on the catalogue bursts come from the fluence calibration, and contain contributions from the spatial and temporal offset of the calibrator from the burst position and time of arrival.

the expectation (Masui, 2020), and on average did an excellent job describing the expected spectrum. While the testing of the calibration factor against steady source transits was successful, the verification for the purposes of FRB detections was more complicated.

Because injections are performed after the L1 intensity data buffer, it was not possible to request callback data of injected FRBs and test their calibration using either CHIME/FRB's calibration method (see §3.4) or the calibration constant. So, instead the fluence and `bonsai` SNR relationship between bursts in the first CHIME/FRB catalogue (CHIME/FRB Collaboration, in prep.) and a population of $N \sim 85,000$ injected bursts (drawn from distributions explained in §4.4) was compared. Noting that the fluence calibration method derives lower limits on the fluence of detected bursts (see §3.4), a cut was made on both the CHIME/FRB catalogue bursts and the injections such that all were detected within the 600 MHz FWHM of beams in the most sensitive beam column. This gave a subset of bursts in the catalogue with the most accurate fluences alongside a comparable set of injections.

The resulting fluence and L1 SNR relationship is shown in Figure 4–5. While it's possible to statistically compare the 2D distributions of the catalogue detections and the injections with tests such as the Peacock test (Peacock, 1983), it is clear visually that the two populations follow a similar trend. This evidence, combined with the successful calibration tests against steady sources, gives confidence that the injection system properly calibrates its bursts in real time.

4.4 Generating Populations of FRBs

When generating a population of FRBs to be injected, it's difficult to surmise what the best input distributions are. The distribution of parameters in injected bursts should give a reasonable fraction of detected events while still spanning the interesting parts of parameter space. Ideally, one would want to inject distributions which are based on physical models, as then the parameters of these models in the true population of bursts could be better understood when comparing the results of injections to the detected FRB population. However, in the case of FRBs, there are not yet physically well-motivated distributions on which to model populations.

In order to ensure FRBs span a range of interesting parameters while remaining detectable, CHIME/FRB currently uses a combination of uniform, log-normal, kernal density estimate (KDE), and power-law distributions, parts of which come from fits to the detection distributions in CHIME/FRB's first catalogue ("catalogue distributions"). The details of the distributions for the relevant injection parameters in Table 4–1 are described here briefly:

- For the DM distribution, 90% of the population is drawn from a log-normal fit to the catalogue distribution, while 10% of the population is drawn from a uniform distribution between 0 and 5000 pc cm^{-3}
- For the fluence distribution, the population is drawn from a power-law distribution with index γ. For the injections shown in Figure 4–5, $\gamma = -1.0$
- For the width distribution, 90% of the population is drawn from a log-normal fit to the catalogue distribution, while 10% of the population is drawn from a uniform distribution between 0 and 100 ms
- The distributions for the spectral indices (β and ζ) are drawn from a kernel density estimator fit to the catalogue data
- For the scattering distribution, 90% of the population is drawn from a log-normal fit to the catalogue distribution, while 10% of the population is drawn from a uniform distribution between 0 and 100 ms

To determine the sky distribution of an injection population, a uniform set of FRBs are first generated on sky, with parameters drawn as above. Then, the expected SNR for each FRB is calculated using the radiometer equation (Equation 1.7) and a fiducial system equivalent flux density derived from CygA transit data. After modulating the radiometer SNR by the band-averaged beam sensitivity, bursts below a certain threshold are cut from the sample to avoid injecting too many faint, undetectable events. The result is a population whose sky distribution traces the beam sensitivity as the sky transits across the beam. This allows not only for injections in the primary lobe of the CHIME/FRB beam, but also for the possibility of the brightest injections being detected in the sidelobes.

4.5 CHIME/FRB Population Bias Correction

The primary use of the CHIME/FRB injection sample is to "correct" the catalogue of CHIME/FRB detections for observational biases. In principle the selection function can be determined by injecting a sufficiently large number of simulated events to cover the entirety of the observable, multi-dimensional FRB parameter space. With a large enough sample, the selection function would be given by the detection fraction across many histogram bins in the parameter space. In practice, however, the parameter space is large enough that performing so many injections is not tractable. There are also further complications presented by the fact that CHIME/FRB measures fluences in a way that derives lower limits (see §3.4). Instead, parameter inference methods must be used by making some assumptions about the FRB population. The methods to correct observational biases are still heavily under development, and will be described in detail in a future publication (Munchmeyer et al., in prep.). As such, only a brief introduction will be given here.

Consider a model FRB population, $r(\mathbf{X})$, where $\mathbf{X}$ is a set of observables measured by the experiment. For CHIME/FRB, these observables currently consist of $\mathbf{X} = (\mathrm{DM}, S_\nu, W, \tau, \beta, \zeta)$. A simple analytical assumption is that the true FRB population has parameter distributions which are uncorrelated, except for DM and S_ν, since DM should track redshift and bursts further away will appear fainter to the telescope. Noting that the spectral parameters β and ζ come from an empirical model and are very covariant, the FRB population can be broken down as:

$$r(\mathbf{X}) = r_1(\mathrm{DM}, S_\nu) r_2(W) r_3(\tau) r_4(\beta, \zeta). \tag{4.2}$$

Equation 4.2 describes a “true” model population of FRBs, not what is detected by the telescope. Of the measured observables (denoted $\bar{\mathbf{X}}$), we assume that some are measured well enough to be considered exact (DM, W, τ) because the parameter measurement process cannot yet be fully forward modelled for injected bursts. However, since `fitburst` accurately measures these parameters, this can be considered a valid assumption. The only parameters that must be forward modelled are β and ζ, because `fitburst` does not attempt to correct for the beam model in its fit which drastically changes the observed frequency spectrum of pulses. The forward modelling here is performed by taking a least-squares fit of the injection spectra (attenuated by the beam model) upon population generation using a running power law.

After injecting FRBs, a fiducial population model must be constructed from the detected injections. Since the parameter distribution of detected injections will never perfectly match the catalog, each of the parameters in $\mathbf{X}$ must be simultaneously re-weighted to match the catalog distributions. The exact process of generating this fiducial model is still being refined, but it will result in a set of weights $w(\mathbf{X})$ which matches injections to the catalog when applied to the detected injections distribution.

The selection function can be constructed from the fiducial population model. Finding the full multi-dimensional selection function simultaneously in all parameters is difficult, so for now only a subset of parameters $\mathbf{Q}$ are used in any given analysis, with $\mathbf{X}$ now representing the parameters over which the selection function is marginalized. As an example, when analyzing the brightness distribution, the interesting parameters are DM and S_ν. The selection function is then described as $P(\mathbf{N}|\mathbf{Q}, \mathbf{X})$: the probability of $\mathbf{N}$ events being detected as a function of the parameters of interest $\mathbf{Q}$, where $\mathbf{X}$ are the assumptions made about the distributions of the other parameters. P is thus constructed as the fraction of re-weighted injection events detected in each $\mathbf{Q}$ bin. The corrected parameter distributions for the catalogue can then be inferred using maximum likelihood methods alongside the selection function and the CHIME/FRB catalogue parameter distribution. The exact formalism of the bias correction is still a work in progress, and the details and results will be found in a forthcoming paper (CHIME/FRB Collaboration, in prep.).

CHAPTER 5
Closing Remarks

With the CHIME/FRB experiment well under way, Fast Radio Burst science is being revolutionized seemingly every several months. From the discovery of a second repeating FRB followed by a host of more repeaters (CHIME/FRB Collaboration et al., 2019c,b; Fonseca et al., 2020), to the discovery of the first periodically repeating FRB (The CHIME/FRB Collaboration et al., 2020b), and finally the association of an FRB-like event with a Galactic magnetar (The CHIME/FRB Collaboration et al., 2020a), CHIME/FRB is at the forefront of discoveries in the field. The next great step for the experiment is to release a large-scale catalogue of events, and do various scientific analyses with the resulting parameter distributions – such as determining the volumetric event rate of FRBs.

In order to efficiently classify the parameters of CHIME/FRB's plentiful detections, the intensity pipeline described in Chapter 3 was built. Through a series of automated fits on intensity data, this pipeline now does an excellent job of automatically classifying events tagged as likely FRB candidates by the real-time pipeline. Its infrastructure allows for efficient classification, as it is naturally parallelized and scalable, and its modularity will allow for easy implementation of additional functionality. In the case of events with unsatisfactory fits, a stage of human vetting allows for pipeline reruns, ensuring quality output for science use.

The problem of properly reporting on the "true" FRB population's parameter distributions has been long complicated by observational biases resulting from RFI, telescope design, and detection algorithms. One method of accounting for these biases includes using an injected population of synthetic FRBs into the live telescope datastream across some range of parameter space, and correcting the telescope's detected population based on the recovery fraction of injected events. This has been lightly explored so far (Farah et al., 2019; Keane & Walker, 2019), but no large-scale FRB survey has yet implemented a robust bias correction strategy with injections. The injection system and bias correction methods presented in Chapter 4 act as a novel solution to this longstanding problem. CHIME/FRB has already begun injecting fiducial populations of FRBs, and continues to refine the methods which use the detection statistics of injections to correct for the telescope selection function.

While the CHIME/FRB injection system has reached a level of sophistication appropriate for science use, there is still work to be done. Currently, injections into only four duplicated beams at a time may be performed. Once the robustness of the injection snatcher has been more thoroughly tested, injections may be performed in non-duplicated beams, which will increase the efficiency of injections substantially. While populations of FRBs to inject are currently generated prior to the beginning of an injections run, eventually this will not be the case. A machine learning addition to the injection driver will be developed in the future, which hopes to use past injection statistics as well as current sky conditions to generate a running list of injections with near optimal parameters for gaining more information about biases. These additions will make the injection system more efficient at generating data for the purposes of bias correction, bolstering its current capability.

REFERENCES

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, PhRvL, 119, 161101

Bandura, K., Addison, G. E., Amiri, M., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9145, Ground-based and Airborne Telescopes V, ed. L. M. Stepp, R. Gilmozzi, & H. J. Hall, 914522

Bandura, K., Bender, A. N., Cliche, J. F., et al. 2016, Journal of Astronomical Instrumentation, 5, 1641005

Bannister, K. W., Shannon, R. M., Macquart, J. P., et al. 2017, ApJL, 841, L12

Bannister, K. W., Deller, A. T., Phillips, C., et al. 2019, Science, 365, 565

Bassett, B., & Hlozek, R. 2010, Baryon acoustic oscillations, ed. P. Ruiz-Lapuente, 246

Bhandari, S., Keane, E. F., Barr, E. D., et al. 2018, MNRAS, 475, 1427

Bhandari, S., Sadler, E. M., Prochaska, J. X., et al. 2020, ApJL, 895, L37

Bochenek, C. D., Ravi, V., Belov, K. V., et al. 2020, arXiv e-prints, arXiv:2005.10828

Borghese, A., Coti Zelati, F., Rea, N., et al. 2020, arXiv e-prints, arXiv:2006.00215

Caleb, M., Flynn, C., Bailes, M., et al. 2016, MNRAS, 458, 708

Caleb, M., Stappers, B. W., Rajwade, K., & Flynn, C. 2019, MNRAS, 484, 5500

Cen, R., & Ostriker, J. P. 1999, ApJ, 514, 1

Champion, D. J., Petroff, E., Kramer, M., et al. 2016, MNRAS, 460, L30

Chatterjee, S., Law, C. J., Wharton, R. S., et al. 2017, Nature, 541, 58

Chawla, P., Andersen, B. C., Bhardwaj, M., et al. 2020, arXiv e-prints, arXiv:2004.02862

CHIME/FRB Collaboration, Amiri, M., Bandura, K., et al. 2018, ApJ, 863, 48

—. 2019a, Nature, 566, 230

CHIME/FRB Collaboration, Andersen, B. C., Bandura, K., et al. 2019b, ApJL, 885, L24

CHIME/FRB Collaboration, Amiri, M., Bandura, K., et al. 2019c, Nature, 566, 235

CHIME/Pulsar Collaboration, Amiri, M., Bandura, K. M., et al. 2020, arXiv e-prints, arXiv:2008.05681

Connor, L. 2019, MNRAS, 487, 5753

Cordes, J. M., Bhat, N. D. R., Hankins, T. H., McLaughlin, M. A., & Kern, J. 2004, ApJ, 612, 375

Cordes, J. M., & Lazio, T. J. W. 2002, arXiv e-prints, astro

Cordes, J. M., Wasserman, I., Hessels, J. W. T., et al. 2017, ApJ, 842, 35

Cordes, J. M., Wharton, R. S., Spitler, L. G., Chatterjee, S., & Wasserman, I. 2016, arXiv e-prints, arXiv:1605.05890

Davé, R., Cen, R., Ostriker, J. P., et al. 2001, ApJ, 552, 473

Deng, W., & Zhang, B. 2014, ApJL, 783, L35

Dolag, K., Gaensler, B. M., Beck, A. M., & Beck, M. C. 2015, MNRAS, 451, 4277

Eatough, R. P., Falcke, H., Karuppusamy, R., et al. 2013, Nature, 501, 391

Farah, W., Flynn, C., Bailes, M., et al. 2018, MNRAS, 478, 1209

—. 2019, MNRAS, 488, 2989

Ferrière, K. M. 2001, Reviews of Modern Physics, 73, 1031

Fonseca, E., Andersen, B. C., Bhardwaj, M., et al. 2020, arXiv e-prints, arXiv:2001.03595

Gourdji, K., Michilli, D., Spitler, L. G., et al. 2019, ApJL, 877, L19

Han, J. L., Manchester, R. N., Lyne, A. G., Qiao, G. J., & van Straten, W. 2006, ApJ, 642, 868

Hessels, J. W. T., Spitler, L. G., Seymour, A. D., et al. 2019, ApJL, 876, L23

Hewish, A., Bell, S. J., Pilkington, J. D. H., Scott, P. F., & Collins, R. A. 1968, Nature, 217, 709

Hinshaw, G., Singh, S., & Wulf, D. 2020, CHIME Internal Documentation, 1183

Hobbs, G., Manchester, R., Teoh, A., & Hobbs, M. 2004, in IAU Symposium, Vol. 218, Young Neutron Stars and Their Environments, ed. F. Camilo & B. M. Gaensler, 139

Hulse, R. A., & Taylor, J. H. 1975, ApJL, 195, L51

Isaacman, R., & Rankin, J. M. 1977, ApJ, 214, 214

Jankowski, F., van Straten, W., Keane, E. F., et al. 2018, MNRAS, 473, 4436

Josephy, A., Chawla, P., Fonseca, E., et al. 2019, ApJL, 882, L18

Katz, J. I. 2016, ApJ, 818, 19

Keane, E. F., Stappers, B. W., Kramer, M., & Lyne, A. G. 2012, MNRAS, 425, L71

Keane, E. F., & Walker, C. R. H. 2019, arXiv e-prints, arXiv:1907.07948

Keating, L. C., & Pen, U.-L. 2020, arXiv e-prints, arXiv:2001.11105

Lan, T.-W., & Fukugita, M. 2017, ApJ, 850, 156

Lattimer, J. M., & Prakash, M. 2007, PhR, 442, 109

Law, C. J., Abruzzo, M. W., Bassa, C. G., et al. 2017, ApJ, 850, 76

Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, Science, 318, 777

Lorimer, D. R., & Kramer, M. 2012, Handbook of Pulsar Astronomy

Lu, W., Kumar, P., & Zhang, B. 2020, arXiv e-prints, arXiv:2005.06736

Lyutikov, M. 2020, ApJ, 889, 135

Macquart, J. P., & Ekers, R. D. 2018, MNRAS, 474, 1900

Macquart, J.-P., & Koay, J. Y. 2013, ApJ, 776, 125

Macquart, J. P., Shannon, R. M., Bannister, K. W., et al. 2019, ApJL, 872, L19

Macquart, J. P., Prochaska, J. X., McQuinn, M., et al. 2020, Nature, 581, 391

Marcote, B., Paragi, Z., Hessels, J. W. T., et al. 2017, ApJL, 834, L8

Marcote, B., Nimmo, K., Hessels, J. W. T., et al. 2020, Nature, 577, 190

Margalit, B., Beniamini, P., Sridhar, N., & Metzger, B. D. 2020, ApJL, 899, L27

Marinacci, F., Vogelsberger, M., Mocz, P., & Pakmor, R. 2015, MNRAS, 453, 3999

Masui, K. 2020, CHIME Internal Documentation, 1182

Masui, K., Lin, H.-H., Sievers, J., et al. 2015, Nature, 528, 523

Masui, K. W., & Sigurdson, K. 2015, PhRvL, 115, 121301

McQuinn, M. 2014, ApJL, 780, L33

Mereghetti, S., Savchenko, V., Ferrigno, C., et al. 2020, arXiv e-prints, arXiv:2005.06335

Metzger, B. D., Berger, E., & Margalit, B. 2017, ApJ, 841, 14

Metzger, B. D., Margalit, B., & Sironi, L. 2019, MNRAS, 485, 4091

Michilli, D., Seymour, A., Hessels, J. W. T., et al. 2018, Nature, 553, 182

Newburgh, L. B., Addison, G. E., Amiri, M., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9145, Proc. SPIE, 91454V

Ng, C., Vanderlinde, K., Paradise, A., et al. 2017, arXiv e-prints, arXiv:1702.04728

Oppermann, N., Connor, L. D., & Pen, U.-L. 2016, MNRAS, 461, 984

Oppermann, N., Yu, H.-R., & Pen, U.-L. 2018, MNRAS, 475, 5109

Oppermann, N., Junklewitz, H., Robbers, G., et al. 2012, A&A, 542, A93

Palmer, D. M., Barthelmy, S., Gehrels, N., et al. 2005, Nature, 434, 1107

Papastergis, E., Cattaneo, A., Huang, S., Giovanelli, R., & Haynes, M. P. 2012, ApJ, 759, 138

Peacock, J. A. 1983, MNRAS, 202, 615

Petroff, E., Keane, E. F., Barr, E. D., et al. 2015a, MNRAS, 451, 3933

Petroff, E., Johnston, S., Keane, E. F., et al. 2015b, MNRAS, 454, 457

Petroff, E., Bailes, M., Barr, E. D., et al. 2015c, MNRAS, 447, 246

Petroff, E., Barr, E. D., Jameson, A., et al. 2016, PASA, 33, e045

Petroff, E., Burke-Spolaor, S., Keane, E. F., et al. 2017, MNRAS, 469, 4465

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv e-prints, arXiv:1807.06209

Platts, E., Weltman, A., Walters, A., et al. 2019, PhR, 821, 1

Prochaska, J. X., Macquart, J.-P., McQuinn, M., et al. 2019, Science, 366, 231

Rajwade, K. M., Mickaliger, M. B., Stappers, B. W., et al. 2020, MNRAS, arXiv:2003.03596

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

Ravi, V. 2019a, MNRAS, 482, 1966

—. 2019b, Nature Astronomy, 3, 928

Ravi, V., Shannon, R. M., & Jameson, A. 2015, ApJL, 799, L5

Ravi, V., Shannon, R. M., Bailes, M., et al. 2016, Science, 354, 1249

Ravi, V., Catha, M., D'Addario, L., et al. 2019, Nature, 572, 352

Rickett, B. J. 1977, ARA&A, 15, 479

Ridnaia, A., Golenetskii, S., Aptekar, R., et al. 2020, The Astronomer's Telegram, 13688, 1

Rybicki, G. B., & Lightman, A. P. 1979, Radiative processes in astrophysics

Salpeter, E. E. 1969, Nature, 221, 31

Scheuer, P. A. G. 1968, Nature, 218, 920

Scholz, P., Bogdanov, S., Hessels, J. W. T., et al. 2017, ApJ, 846, 80

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

Simard, D., & Pen, U.-L. 2018, MNRAS, 478, 983

Spitler, L. G., Cordes, J. M., Hessels, J. W. T., et al. 2014, ApJ, 790, 101

Spitler, L. G., Scholz, P., Hessels, J. W. T., et al. 2016, Nature, 531, 202

Tavani, M., Ursi, A., Verrecchia, F., et al. 2020, The Astronomer's Telegram, 13686, 1

Taylor, J. H. 1974, A&AS, 15, 367

Tegmark, M., & Zaldarriaga, M. 2009, PhRvD, 79, 083530

Tendulkar, S. P., Bassa, C. G., Cordes, J. M., et al. 2017, ApJL, 834, L7

The CHIME/FRB Collaboration, :, Andersen, B. C., et al. 2020a, arXiv e-prints, arXiv:2005.10324

The CHIME/FRB Collaboration, Amiri, M., Andersen, B. C., et al. 2020b, arXiv e-prints, arXiv:2001.10275

Thornton, D., Stappers, B., Bailes, M., et al. 2013, Science, 341, 53

Vedantham, H. K., & Phinney, E. S. 2019, MNRAS, 483, 971

Vedantham, H. K., Ravi, V., Hallinan, G., & Shannon, R. M. 2016, ApJ, 830, 75

Wang, W.-Y., Xu, R., Zheng, X., & Chen, X. 2020, arXiv e-prints, arXiv:2005.02100

Yang, Y.-P., Luo, R., Li, Z., & Zhang, B. 2017, ApJL, 839, L25

Yang, Y.-P., & Zhang, B. 2016, ApJL, 830, L31

Yao, J. M., Manchester, R. N., & Wang, N. 2017, ApJ, 835, 29

Zhang, B. 2018, ApJL, 867, L21

Zhang, S. N., Tuo, Y. L., Xiong, S. L., et al. 2020, The Astronomer's Telegram, 13687, 1

Zhang, Y. G., Gajjar, V., Foster, G., et al. 2018, ApJ, 866, 149

Zheng, Z., Ofek, E. O., Kulkarni, S. R., Neill, J. D., & Juric, M. 2014, ApJ, 797, 71