# Radio Transient Calibration Techniques for CHIME/FRB


Bridget Andersen


Master of Science


Department of Physics

McGill University

Montréal, Québec


2020-06-15

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

**ABSTRACT**

Fast radio bursts (FRBs) are bright radio transients of micro-to-millisecond duration and unknown extragalactic origin. Central to the mystery of FRBs are their extremely high characteristic energies, which surpass the typical energies of other radio transients of similar duration, like Galactic pulsar and magnetar bursts, by orders of magnitude. Calibrating FRB-detecting telescopes for burst flux and fluence determination is crucial for FRB science, as these measurements enable studies of the FRB energy and brightness distribution in comparison to progenitor theories.

The Canadian Hydrogen Intensity Mapping Experiment (CHIME) is a radio interferometer located near Penticton, B.C., that has detected an unprecedented number of FRBs. The efficiency with which CHIME detects these bursts is enabled by its novel design, consisting of four 20-m by 100-m cylindrical reflectors with 256 dual-polarization feeds lined along each axis that are sensitive to a wide bandwidth of 400−800 MHz. The CHIME/FRB project operates commensally on the CHIME data stream, continuously searching a grid of 1,024 formed beams over a ∼200 square degree field of view at 1-ms time resolution.

Such a novel design also produces novel challenges for CHIME/FRB flux calibration. In this thesis, we provide a comprehensive review of these challenges, as well as an automated flux calibration software pipeline that was developed to calibrate bursts detected in the first CHIME/FRB catalog, consisting of 535 events detected between July 25th, 2018 and July 1st, 2019.

# ABRÉGÉ

Les sursauts radio rapides (FRB pour *fast radio bursts* en anglais) sont des phénomènes transitoires radio lumineux d'une durée de l'ordre de quelques $\mu$s à quelques ms et d'origine extragalactique inconnue. Le mystère des sursauts radio rapides réside dans leurs énergies caractéristiques extrêmement élevées, qui dépassent de plusieurs ordres de grandeur les énergies typiques d'autres transitoires radio de durée similaire, tels que les sursauts des pulsars et des magnétars galactiques. La calibration des télescopes de détection des FRB pour la détermination du flux et de la fluence des sursauts est cruciale pour la science des FRB, car ces mesures permettent d'étudier l'énergie des FRB et la distribution de leur luminosité en comparaison avec les théories de progéniteurs.

L'expérience canadienne de cartographie de l'intensité de l'hydrogène (CHIME pour *Canadian Hydrogen Intensity Mapping Experiment* en anglais) est un interféromètre radio situé près de Penticton, en Colombie-Britannique, qui a détecté un nombre sans précédent de FRB. L'efficacité avec laquelle CHIME détecte ces rafales est rendue possible par sa conception novatrice, qui consiste en quatre réflecteurs cylindriques de 20-m sur 100-m avec 256 détecteurs à double polarisation alignés le long de chaque axe, sensibles à une large bande passante de $400-800\,$MHz. Le projet CHIME/FRB fonctionne conjointement avec le flux de données CHIME, en recherchant continuellement sur une grille de 1,024 faisceaux formés sur un champ de vision de $\sim$200 degrés carrés avec une résolution temporelle de 1-ms.

Une telle conception novatrice produit également de nouveaux défis pour la calibration du flux CHIME/FRB. Dans cette thèse, nous fournissons une revue complète de ces défis, ainsi qu'un pipeline logiciel de calibration de flux automatisé qui a été développé pour calibrer les sursauts détectés dans le premier catalogue CHIME/FRB, composé de 535 événements détectés entre le 25 juillet 2018 et le 1er juillet 2019.

# ACKNOWLEDGEMENTS

Working on this thesis has been an exhilarating experience, and I have so many colleagues and friends to thank. First, thank you to my supervisor, Vicky Kaspi, for her steadfast support, and for leading the team through uncharted waters. While orchestrating major discoveries in the FRB field and navigating through the COVID-19 pandemic, you've fostered an exciting and friendly environment for me to grow in as a researcher. Thank you.

Thank you to the humans of the McGill Space Institute, for the continuous spark of seminars and conversations, for the coffee and cookies and Moon puzzles. Thank you to Jojo for the support and professional development mentoring, and for providing perspective when I needed it. Thank you to Emily Petroff, for keeping me motivated to write. Thank you to Chitrang for your infectious enthusiasm and optimism and guidance in all things CHIME/FRB. Thank you to Shiny for your software advice, and for that one time we got pizza by the dishes on-site. Thank you to Pragya, Ziggy, and Alex for brainstorming and answering all of my questions when I got stuck. Thank you to Emmanuel, for your detailed write-ups and encouragement. You're going to be a great professor.

Thanks to Coni and Rafa for motivating me to put one foot in front of the other, and teaching me how to safely climb mountains. Thank you to the rest of the MSI dungeon dwellers — Alice, Andrew Z., and Simon — for friendship and motivation and everyday antics. Thank you to Marcus Merryfield for being such a generous friend and for navigating these wild few years with me. You're a crazy smart researcher and I've learned so much from you.

Thank you to Jacob, for your rock solid presence and level, keen, caring mind. For sharing the stars with me from the hills of LA to the forests of Maine.

And thank you to my family and my Mom and my Dad, for purple tomatoes and cinnamon scones, long car rides and brown bag lunches, desert adventures and a green home. For your burning sense of curiosity and constant encouragement. There's no justice to be done here with words. Thank you.

## CONTRIBUTION BREAKDOWN

The commissioning and operation of the CHIME telescope and its three commensal experiments (CHIME/Cosmology, CHIME/FRB, and CHIME/Pulsar) is a collective and monumental effort, encompassing years of work from dedicated engineers, software developers, project managers, professors, postdocs, and graduate students. A list of individuals that have contributed to the CHIME/FRB project can be found in the authors of CHIME/FRB Collaboration et al. (2021). Chapter 2 summarizes parts of this work related to the flux calibration pipeline.

The pipeline described in this thesis makes extensive use of the composite formed and primary beam model for CHIME/FRB (§2.3). The codebase for this beam model was developed by Paul Scholz and Alex Josephy. The initial primary beam model (without the 30 MHz ripple) was based on simulations by Meiling Deng. The primary beam response in the robust beam model (with the 30 MHz ripple) makes use of holography measurements by Dallas Wulf and Saurabh Singh, in combination with a coupling model by Gary Hinshaw. In addition, the `fitburst` morphology modeling software (§2.5.1) was developed by Emmanuel Fonseca and Kiyoshi Masui, and the various localization methods (§2.5.2) were developed by Alex Josephy (header localization), Paul Scholz (intensity localization), and Daniele Michilli (baseband localization).

The majority of work presented in Chapters 3 and 4 is original to the author of this thesis, in collaboration with other individuals in the CHIME and CHIME/FRB collaborations. Chitrang Patel, Juan Mena Parra, Seth Siegel, and

Matt Dobbs collaborated with the author to develop the initial method of FRB flux calibration and error estimation using steady source transits. Seth Siegel developed the software used to predict flux calibration source brightness in the CHIME band, as described in §3.3.2. Shiny Brar collaborated with the author on developing the software infrastructure for the flux calibration pipeline. The CHIME/FRB injections infrastructure was a massive effort by the following individuals: Marcus Merryfield, Shriharsh Tendulkar, Kiyoshi Masui, Dustin Lang, Adam Dong, Deborah Good, Kendrick Smith, Moritz Munchmeyer, and Kaitlyn Shin. Finally, the injections comparison described in §3.3.6 was completed by the author and Marcus Merryfield.

TABLE OF CONTENTS

LIST OF TABLES

Table | page

LIST OF FIGURES

# CHAPTER 1
## An Introduction to Fast Radio Bursts

Fast Radio Bursts (FRBs) are an enigmatic class of astrophysical transients currently posing one of the greatest mysteries in the field of radio astronomy. The name of these transients is apt, as it concisely encompasses the primary properties that define them. "Fast" – FRB durations are extremely short, lasting on the order of micro-to-milliseconds. "Radio" – FRBs have been detected in the radio part of the electromagnetic spectrum from 0.3 to 8 GHz (Chawla et al., 2020; Gajjar et al., 2018). "Burst" – FRBs are extremely energetic events with typical luminosities on the order of $\sim 10^{43}\,\mathrm{erg\,s^{-1}}$, briefly radiating as much energy as a typical radio galaxy (Katz, 2018; Fanaroff & Riley, 1974). Within the phase space of radio transients, these properties set FRBs in a unique position. In particular, with the exception of the recent detection of a bright burst from Galactic magnetar SGR 1935+2154 (see §1.4), radio transients of similar duration to FRBs, like pulsars and rotating radio transients (RRATs), are many orders of magnitude less energetic (e.g., Petroff et al., 2019; Keane, 2018).

Another primary characteristic of FRBs is their extragalactic origin, which is manifested in the large frequency-dependent time delays that FRBs take on as they propagate through the ionized plasma of the intergalactic medium (described in detail in §1.1.1). FRB dispersive delays are far in excess of what is expected from Milky Way plasma, which was the first piece of evidence toward their extragalactic nature. In the past few years, a handful of precise localizations have confirmed that FRBs originate from galaxies at mega-to-gigaparsec distances (e.g., Heintz et al., 2020; Marcote et al., 2020; Ravi et al., 2019).

Estimated rates predict that thousands of FRBs occur across the sky per day (e.g., Bhandari et al., 2018), but only a small fraction of these bursts are actually detected due the limited field-of-view and sensitivity of most radio telescopes. Still, search pipelines attuned to FRBs have been developed at telescopes worldwide over the past decade, resulting in the detection and publication of more than two hundred bursts to-date (Petroff et al., 2019). As outlined in §1.3, these detections have driven significant progress in the observational characterization of FRBs, and the sparse initial attributes described above have bloomed into a sprawling landscape of diverse properties. Although a long and evolving list of progenitor theories have been proposed to match these observations (Platts et al., 2019), as of yet, a comprehensive physical explanation of the FRB phenomenon remains elusive.

In this chapter I lay the groundwork for understanding the current state of the FRB field, especially with regards to burst energetics and the observational determination of burst brightnesses. In summary, I review fundamental FRB properties imparted by the intergalactic medium (§1.1), popular progenitor models (§1.2), particularly notable FRB detections (§1.3), and the science related to burst brightnesses (§1.4).

## 1.1 Characteristics Imparted by the Intervening Medium

As an FRB travels from extragalactic distances, interactions with ionized plasma in the intervening medium cause distinctive transformations to the intrinsic signal. In this section I highlight three of these transformations: dispersion, scattering, and scintillation, which are described in detail by Lorimer & Kramer (2004).

### 1.1.1 Dispersion

The cold diffuse plasma that makes up the intergalactic (IGM) and interstellar (ISM) medium has a frequency-dependent refractive index, so that the group velocity of a radio signal of frequency $\nu$ scales approximately as $1 - \nu^{-2}$. In practice, this means that the highest frequency of an FRB, $\nu_{\rm high}$, will arrive at the telescope before the lowest frequency, $\nu_{\rm low}$, with a time delay given by:

$$\Delta t = \frac{e^2 \int_0^d n_e dl}{2\pi m_e c}\left(\frac{1}{\nu_{\rm low}^2} - \frac{1}{\nu_{\rm high}^2}\right) \tag{1.1}$$

$$\approx 4.15 \cdot 10^6\,{\rm ms}\left[\frac{\rm DM}{\rm pc\,cm^{-3}}\right]\left[\left(\frac{\nu_{\rm low}}{\rm MHz}\right)^{-2} - \left(\frac{\nu_{\rm high}}{\rm MHz}\right)^{-2}\right] \tag{1.2}$$

where $e$ and $m_e$ are the electron charge and mass, $c$ is the speed of light, $d$ is the distance from the FRB source to the observer, and $n_e$ is the number density of electrons in the intervening medium, which is typically $n_e \sim 0.03\,{\rm cm}^{-3}$ for the ISM. The proportional value, DM, in the equation above is a quantity called the dispersion measure, which is the integrated electron density along the line of sight:

$$\mathrm{DM} = \int_0^d n_e dl \tag{1.3}$$

typically given in units of $\mathrm{pc\,cm^{-3}}$. The DMs of known FRBs range from 103.5 to $2596.1\,\mathrm{pc\,cm^{-3}}$ (Petroff et al., 2016). As an example, an FRB detected over a bandwidth of 400 to 800 MHz at a DM of $500\,\mathrm{pc\,cm^{-3}}$ will arrive at the telescope with a delay of about 10 seconds. An FRB that encounters more electrons along the way, with a DM of $1000\,\mathrm{pc\,cm^{-3}}$, will have a delay of 20 seconds.

This dispersion delay spreads an FRB signal over frequency and time, decreasing its SNR in the band-integrated time series. In most FRB detection algorithms, dispersion is corrected for incoherently in the intensity data by shifting the time bins in each frequency channel in accordance with the expected delay for a given DM. This essentially maximizes the SNR of the burst in the band-integrated time series, so that it can be robustly detected with a matched-filter (see Figure 1–1 for an example of a dispersed and de-dispersed burst). In blind searches the DM of a given burst is not known beforehand, requiring iteration over thousands of DM trials. Thus, dedispersion composes a significant portion of the computational currency of a standard FRB search.

As a census of the intervening medium, the DM can be a very useful quantity for estimating FRB distances and probing the structure of the IGM. As an FRB travels from extragalactic distances, it passes through material local to the source (such as a pulsar wind nebula or supernova remnant), the ISM of the host galaxy, the IGM, and the halo and ISM of the Milky Way. The final DM measured at the telescope encodes contributions from each of these locations:

$$\mathrm{DM_{obs}}(z) = \mathrm{DM_{host}}(z) + \mathrm{DM_{IGM}}(z) + \mathrm{DM_{MW,ISM}} + \mathrm{DM_{MW,Halo}} \tag{1.4}$$

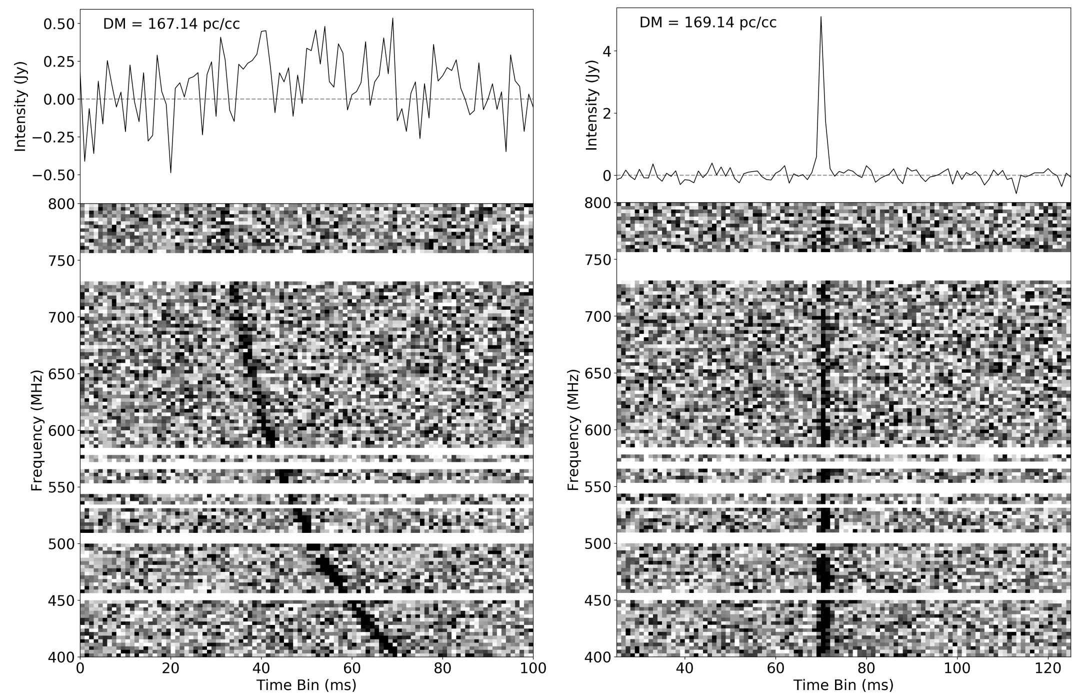


Figure 1–1: Frequency vs. time ("waterfall") plots for one of the first 13 FRBs detected by CHIME/FRB (CHIME/FRB Collaboration et al., 2019a). The top panel of each plot shows the band-averaged timeseries. The bottom panel of each plot shows the dynamic spectrum, where the colour scale shows the intensity (Jy) in each frequency-time bin. The white horizontal stripes indicate frequency channels where RFI has been excised. The left plot shows a burst that is sub-optimally de-dispersed. The right plot shows a burst that is de-dispersed to optimize the SNR in the band-averaged timeseries.

where $z$ is the redshift of the FRB. The free electron distribution in the Milky Way has been constrained using two different models calibrated to pulsars of known DM and distance: NE2001 (Cordes & Lazio, 2002) and YMW16 (Yao et al., 2017). Pre-CHIME/FRB, all known FRB DMs ranged from 1.4 to 200 times the Galactic contribution expected from these models (Petroff et al., 2016), indicating their extragalactic nature. The FRB host galaxy and Milky Way halo contributions are poorly constrained, but are typically taken to be $\mathrm{DM_{host}} \sim 50/(1+z)\,\mathrm{pc\,cm^{-3}}$ and $\mathrm{DM_{MW,Halo}} \sim 50\,\mathrm{pc\,cm^{-3}}$, respectively (Prochaska & Zheng, 2019; Deng & Zhang, 2014).

With estimates of all other contributions, the remaining DM from an FRB can be approximately attributed to the IGM. Additional assumptions about the baryon ionization and the electron-to-baryon fractions in the IGM in turn allow this DM to be used as a rough proxy for FRB redshift: $z \sim \mathrm{DM_{IGM}}/(1000\,\mathrm{pc\,cm^{-3}})$ (Ioka, 2003; Inoue, 2004; Zhang, 2018). If an FRB has already been localized to a host galaxy of known redshift, then this analysis can be reversed to constrain the composition and density of the IGM. A large enough sample of localized FRBs would allow us to investigate some of the most fundamental questions about the Universe, such as the "missing" baryon problem (Macquart et al., 2020) and the epoch of Helium reionization (Zheng et al., 2014).

### 1.1.2 Scattering and Scintillation

Inhomogeneities and turbulence in the intervening medium can induce multi-path propagation of an FRB signal. When a signal is scattered, parts that travel along longer path lengths arrive at the observer later, which temporally broadens the pulse. This effect is commonly modeled by placing the inhomogeneities on a thin screen of material in the path of propagation, which essentially convolves the pulse with a one-sided decaying exponential of characteristic scattering time $\tau$: $e^{-t/\tau}$ (Scheuer, 1968; Salpeter, 1969; Rickett, 1977). As shown in Figure 1–2, this effect is also highly frequency-dependent, with the scattering time scaling as $\tau \propto \nu^{-4.4}$ when modeling the intervening medium by Kolmogorov turbulence).

Scattering times for FRBs at 1 GHz are typically on the order of a few milliseconds (Petroff et al., 2016).

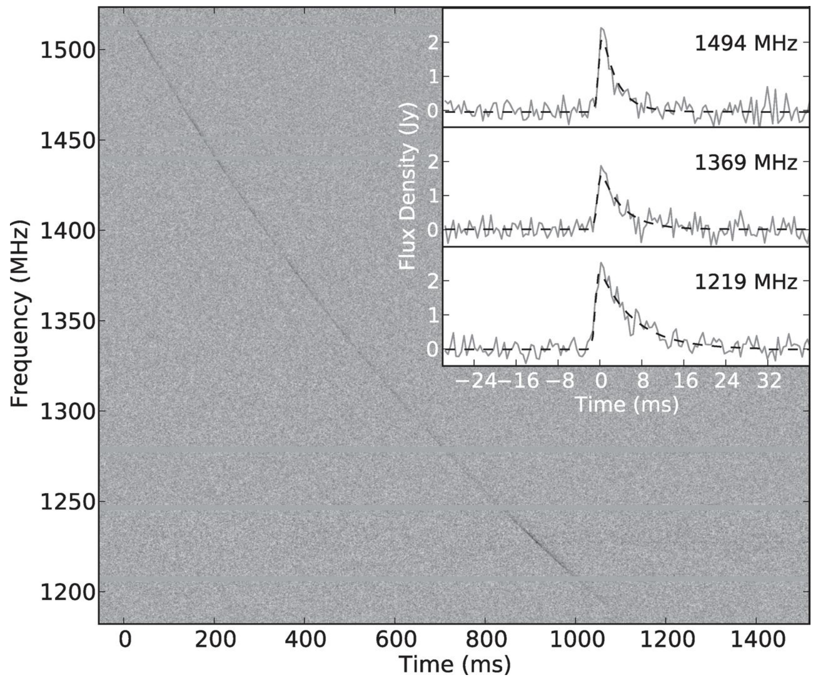


Figure 1–2: A frequency vs. time plot of a burst from Thornton et al. (2013) showing the uncorrected dispersive delay. The inset shows the de-dispersed pulse shape observed at different frequencies. The exponential tail characteristic of scattering is observed to get progressively wider at lower frequencies. This figure was taken directly from Thornton et al. (2013).

Diffraction and refraction by medium inhomogeneities can also cause differences in the phase of the FRB signal, which can interfere constructively or destructively. To the observer, this scintillation can result in a complex structure of enhanced and diminished intensity that changes on a range of timescales and bandwidths depending on the geometries involved (for a review, Lorimer & Kramer 2004 or Rickett 1977).

It is worth noting the existence of several other propagation effects impacting FRBs, including Faraday rotation, plasma lensing, and free-free absorption (see Cordes & Chatterjee 2019 for a review). Unwrapping the intricacies and degeneracies between each of these effects and the intrinsic FRB spectrum is vital not just for understanding FRB origins, but also for unlocking FRBs as powerful tools to probe the structure and magnetization of the IGM.

### 1.2 Progenitor Models

In the early days of FRB science, the extreme nature and sparse initial properties of FRBs formed an open-ended playground for theorists. The running joke repeated in presentations and papers was that there were more FRB theories than there were actual detected FRBs! Although this is not true anymore, the sentiment captures the breadth of initial FRB models, covering progenitors from pulsar giant pulses (e.g., Cordes & Wasserman, 2016) to active galactic nuclei interactions (e.g., Romero et al., 2016) to annihilating black holes (Keane et al., 2012) to superconducting cosmic strings (e.g., Brandenberger et al., 2017). Even now, each new observational discovery sprouts another branch of progenitor models. A comprehensive review of these models is out of the scope of this thesis (see Platts et al. 2019 for a catalog). Here, I summarize some of the most popular FRB theories and in §1.3 I mention a few additional models as they have been spurred by particular FRB discoveries.

Since the high brightnesses and short durations characteristic of the FRB population together imply coherent emission from a compact region, a particularly well-studied subset of FRB progenitor models are those involving compact objects like pulsars (e.g., Lyutikov et al., 2016), magnetars (e.g., Margalit et al., 2020), and black holes (e.g., Mingarelli et al., 2015). The connection with magnetars has especially been bolstered by some detections of high linear polarization (e.g., Ravi et al., 2016; Petroff et al., 2017; Michilli et al., 2018), localizations to star-forming regions (e.g., Bassa et al., 2017; Tendulkar et al., 2017; Marcote et al., 2020), repetition statistics similar to Galactic magnetar flares (e.g., Wadiasingh & Timokhin, 2019; Cheng et al., 2020), sufficiently high magnetar volumetric birth rates to plausibly explain the FRB rate (e.g., Nicholl et al., 2017), and the recent detection of a bright FRB-like burst from Galactic magnetar SGR 1935+2154 (more on this in section §1.4; CHIME/FRB Collaboration et al., 2020a; Bochenek et al., 2020). Generally, magnetar models can be divided based on whether the radio emission originates from within the neutron star magnetosphere (usually involving disturbances in the neutron star crust propagating outward resulting in curvature radiation; e.g., Lu et al. 2020; Kumar et al. 2017; Lyutikov 2020; Wadiasingh & Timokhin 2019) or at a much further distance through a magnetar flare interacting with the surrounding medium (e.g., the synchrotron maser blastwave model, Lyubarsky 2014; Metzger et al. 2019). For a clear and concise summary of magnetar models and predictions, see Margalit et al. (2020).

Although models that are non-cataclysmic have been slightly more favoured due to the detection of repetition in some FRBs (see §1.3.1), the possibility remains that repeaters and as-of-yet non-repeaters are from different progenitor populations. Thus there are some models that are cataclysmic, such as mergers of various combinations of compact objects (NS-NS, Totani 2013; BH-BH, Zhang 2016; WD-WD, Kashiyama et al. 2013; NS-BH, Mingarelli et al. 2015) or collapsing compact objects (e.g., Fuller & Ott, 2015; Shand et al., 2016).

### 1.3 History of Detection

FRB science is a relatively new field that has rapidly progressed over the past decade. The first FRB was published in 2007, after it was discovered in archival data from a Parkes telescope pulsar survey of the Magellanic Clouds (Lorimer et al., 2007). Deemed the "Lorimer Burst," this first FRB was detected with a DM of $375\,\mathrm{pc\,cm^{-3}}$, which is 15 times the NE2001 Galactic contribution along the same line of sight ($25\,\mathrm{pc\,cm^{-3}}$), and far beyond that of any Magellanic Cloud Pulsar at the time (the next largest was PSR J0131$-$7310 at $205\,\mathrm{pc\,cm^{-3}}$; Manchester et al. 2006). Although this single burst was met with a healthy dose of skepticism from the community, the 2013 detection of four more high-DM bursts from Parkes solidified FRBs as a new extragalactic population (Thornton et al., 2013).

After FRB 121102 was detected a year later at Arecibo (Spitler et al., 2014), confirming that FRBs are not local to Parkes, the commissioning of instruments and surveys tailored to FRB detection began in earnest. In addition to continued surveys at Parkes (e.g., Champion et al. 2016), the Green Bank Telescope (GBT, e.g., Masui et al. 2015), and Arecibo (e.g., Spitler et al. 2016), new experiments came online at the upgraded Molonglo Observatory Synthesis Telescope (UTMOST, Bailes et al. 2017), the Australian Square Kilometre Array Pathfinder (ASKAP, Macquart et al. 2010), the Canadian Hydrogen Mapping Experiment (CHIME, CHIME/FRB Collaboration et al. 2018), the Deep Synoptic Array 10-dish prototype (DSA-10, Kocz et al. 2019), and the Westerbork Synthesis

Radio Telescope (WSRT/Apertif, Oostrum et al. 2020). A limited number of detections have also been made with the Five-hundred-meter Aperture Spherical radio Telescope (FAST, Zhu et al. 2020), Very Large Array (VLA, e.g., Law et al. 2020), and the European VLBI Network (EVN, Marcote et al. 2020). With combined efforts from all of these telescopes, the number of bursts detected has more than doubled since 2016, with 245 detected in total as of July 2020 (Petroff et al., 2016).[1]

Beyond adding to the sheer number of new detections, each telescope also brings new opportunities for FRB characterization. Together, these telescopes cover a large range of survey sensitivities, bandwidths, time resolutions, observing cadences, and localization capabilities, enabling wide-ranging studies of the FRB parameter space. Through these studies, the FRB population has revealed itself to be remarkably diverse. The past three years since this thesis began have been particularly exciting in this regard, with seemingly a new field-changing discovery every few months. Here I highlight two major subsets of discoveries that represent fundamental expansions of the FRB population parameter space, to give context for the current state of the field. I reserve a separate section (§1.4) for a discussion of burst brightnesses and energetics, which are the focus of this thesis.

### 1.3.1 Repetition

In 2016, 10 repeat detections were discovered at the same DM and sky position as FRB 121102 (Spitler et al., 2016). Up until this point, the apparent

[1] See http://frbcat.org/ for an up-to-date summary of detected bursts.

non-repeating nature of most FRBs had informed many prevailing origin theories, which invoked cataclysmic events like merging neutron stars (e.g., Totani, 2013) or collapsing supramassive neutron stars (e.g., Falcke & Rezzolla, 2014). This discovery solidified that the FRB phenomenon cannot be completely explained by cataclysmic progenitors, bringing renewed theoretical focus to non-cataclysmic origins.

Spitler et al. (2016) also first noted rapid spectral variations from burst to burst on the order of minutes or less that appeared to be intrinsic to the source. Further observations of FRB 121102 showed a strikingly complex class of bursts composed of sub-bursts that appear to drift towards lower frequencies at later times within the burst envelope, forming a spectral shape akin to a "sad trombone" sound (see, e.g., Figure 1 in Hessels et al. 2019). Observations at higher frequencies from $4 - 8\,\mathrm{GHz}$ revealed that FRB 121102 bursts exhibit extreme polarization properties implying an intensely magnetized environment (Michilli et al., 2018; Gajjar et al., 2018; Hilmarsson et al., 2020). These properties taken together have led to the development of a class of non-cataclysmic progenitor models involving a young highly-magnetized neutron star interacting with a dense and dynamic local environment such as a supernova remnant (e.g., Metzger et al., 2019) or around a massive black hole (e.g., Zhang, 2018), or producing bursts from different locations within its magnetosphere (Lyutikov, 2020).

The study of repeaters was blown wide open after the commissioning of the CHIME radio telescope. CHIME's large field of view and constant monitoring of the transiting sky means that it repeatedly observes a large portion of the northern hemisphere every day, making it uniquely primed for the detection of repeat bursts. As of the writing of this thesis, CHIME/FRB has detected a total of 18 new repeaters, many of which display the same complex spectro-temporal variations as FRB 121102 (CHIME/FRB Collaboration et al., 2019b; Fonseca et al., 2020). With this larger sample it was discovered that the temporal widths of repeaters are statistically larger (at the $4\sigma$ level) than as-of-yet non-repeaters, possibly indicating different emission mechanisms or local environments between the two populations. The discovery of new repeaters with CHIME also elucidated differences within the repeater population. Notably, new repeaters exhibit a wide range of polarization properties, with rotation measures typically orders of magnitude less extreme than FRB 121102 (CHIME/FRB Collaboration et al., 2019b; Fonseca et al., 2020; Michilli et al., 2018).

In 2020, CHIME/FRB also detected a 16.35 day periodicity in the clustering of burst arrival times from repeater FRB 180916, with a $\sim$4 day active window each cycle (CHIME/FRB Collaboration et al., 2020b). Before this discovery, all bursts from repeaters appeared to arrive sporadically, with some evidence of clustering, but no regular pattern (e.g., Opperman & Pen, 2017). This discovery has driven further scrutiny towards long-term trends in the arrival times of bursts from FRB 121102, unveiling a possible 157 day periodicity (Rajwade et al., 2020; Cruces et al., 2021a). Yet another set of progenitor theories have been developed to explain these apparent periodicities, including a neutron star interacting with strong winds from an orbital companion (e.g., Lyutikov et al., 2020; Ioka & Zhang,

2020) and a magnetar precessing on its axis freely (e.g., Levin et al., 2020; Zanazzi & Lai, 2020) or in a binary (e.g., Yang & Zou, 2020).

While it is now clear that repetition is a fundamental quality for some subset of FRBs, many key questions remain open-ended. What accounts for differences in properties between repeaters? Are repeaters and as-of-yet non-repeaters distinct populations? Or are all FRBs the same repeating population, but with different pulse morphologies and drastically different repeat rates (see, e.g., Caleb et al. 2019, Ravi 2019)?

### 1.3.2 Localizations

Localizing bursts to their host galaxies is another crucial step toward understanding FRB origins, as certain progenitors are more likely to exist in galaxies with different properties. Precise burst localization allows detailed characterization of the host galaxy and local FRB environment through multiwavelength follow-up. For example, FRB 121102 was localized to a low-metallicity dwarf galaxy in the same class as those known to host superluminous supernovae, suggesting a possible evolutionary link through the formation of a young magnetar (Bassa et al., 2017; Chatterjee et al., 2017; Marcote et al., 2017; Tendulkar et al., 2017; Metzger et al., 2017).

Most telescopes used throughout FRB detection history were single dishes that could only localize to arcminute-level precision, which is insufficient for unambiguous host identification. Direct FRB localization has only been accomplished within the past three years using a small number of interferometric telescopes with longer baselines, like the VLA, EVN, ASKAP, and DSA-10, yielding sub-arcsecond level precision. As a result, to-date, only 13 FRBs have been localized to host galaxies.[2] Within this small sample, the corresponding galaxies of both repeaters and non-repeaters display wildly varying properties, from star-forming spiral galaxies (Chittidi et al., 2020) to massive early-type spiral galaxies with negligible star formation (Bannister et al., 2019). These early results suggest that there are either multiple populations of burst progenitors, or that FRB progenitors can occur in diverse environments (Bhandari et al., 2020a). A larger number of localizations is required before more robust statistical conclusions can be made.

As we will see in section §3.1, burst localization is also crucial for accurately analyzing FRB brightnesses, as it provides an accurate distance for converting observed fluences to intrinsic source energies and enables correction of fluence measurements for telescope beam attenuation.

## 1.4 Energetics

The extreme energetics of FRBs remains one of the most integral properties that progenitor theories must contend with. The fundamental observables for FRB energetics are the peak specific flux density, the maximum energy per unit time per unit area per unit frequency that the burst reaches, and the fluence, the integrated flux density over the duration of the burst. FRB fluxes are measured in a unit common to radio astronomy called the Jansky (Jy; named after early radio astronomer Karl G. Jansky) where $1\,\mathrm{Jy} = 10^{-26}\,\mathrm{W\,m^{-2}\,Hz^{-1}}$, while fluence is measured in Jy ms. Observed peak fluxes range from $\mathcal{O}(10^{-2})$ to $\mathcal{O}(10^{2})$ Jy and

[2] See http://frbhosts.org for an up-to-date summary of hosts and properties.

fluences range from $\mathcal{O}(10^{-1})$ to $\mathcal{O}(10^{2})$ Jy ms (Petroff et al., 2016, with the caveat that most observed FRB flux and fluence values are lower limits, see §3.1).

With an estimate of the luminosity distance to an FRB ($D_L$), either derived approximately from the DM or obtained through a direct host galaxy localization, fluence ($F_\nu$) and flux ($S_\nu$) measured at the telescope can be converted into intrinsic isotropic-equivalent peak luminosities ($L_p$) and energies ($E$) at the source through cosmological versions of the inverse square law (e.g., Zhang, 2018):

$$L_p = 4\pi D_L^2 S_{\nu,p} \Delta\nu \approx \left(10^{40}\,\mathrm{erg\,s^{-1}}\right)\left(\frac{D_L}{100\,\mathrm{Mpc}}\right)^2 \frac{S_{\nu,p}}{\mathrm{Jy}}\frac{\Delta\nu}{\mathrm{GHz}} \tag{1.5}$$

$$E = \frac{4\pi D_L^2}{(1+z)} F_\nu \Delta\nu \approx \left(10^{37}\,\mathrm{erg}\right)\left(\frac{D_L}{100\,\mathrm{Mpc}}\right)^2 \frac{F_\nu}{\mathrm{Jy\,ms}}\frac{\Delta\nu}{\mathrm{GHz}} \tag{1.6}$$

where $\Delta\nu$ is the bandwidth of the receiving telescope and the $(1+z)$ term in the energy equation is the k-correction for the duration of the burst as it dilates between the source and the observer (e.g., Hogg, 1999; Macquart & Ekers, 2018a).

Both repeating and so-far non-repeating FRBs have been localized to cosmological distances ranging approximately from 150 Mpc to 4 Gpc which, paired with well-constrained fluence measurements, imply isotropic-equivalent burst energies spanning at least six orders of magnitude from $10^{36}$ to $10^{42}$ erg (Marcote et al. 2020; Ravi et al. 2019; energies here calculated assuming a fiducial burst bandwidth of 500 MHz). Even within the population of bursts detected from a single repeating source the energy range can be vast. For example, burst energies from FRB 121102 have ranged from $10^{37}$ to $10^{40}$ erg (Gourdji et al. 2019; Oostrum et al. 2017). The FRB energetics problem is thus multifaceted: models must account not only for such extreme energy outputs within short burst durations, but also for a large range of energy outputs both across the entire population and within a single repeating source.

Even the lowest energy of these bursts is orders of magnitude brighter than typical Galactic bursts. Giant radio pulses from the Crab pulsar have been observed with fluences up to $5{\cdot}10^3$ Jy ms at 1.3 GHz which, considering the distance of the Crab Nebula, yields an energy of $10^{31}$ erg (Bera & Chengalur, 2019), still 5 orders of magnitude lower than the faintest burst seen from FRB 180916 (Marcote et al., 2020). Until recently, the brightest burst from a Galactic magnetar was also $\sim 10^{31}$ erg at 6 GHz (magnetar 1E 1547.0−5408, with a fluence $> 200$ Jy ms; Burgay et al., 2018). This all changed in July 2020 with the simultaneous detection by the STARE-2 and CHIME/FRB experiments of a bright millisecond burst from magnetar SGR 1935+2154. The radio energy of this burst was $2.2 \cdot 10^{35}$ erg at 1.4 GHz (Bochenek et al., 2020) and $3 \cdot 10^{34}$ erg from $400 - 800$ MHz (CHIME/FRB Collaboration et al., 2020a), the former being just one order of magnitude fainter than the faintest burst seen from FRB 180916. This detection has significantly closed the gap between Galactic radio transients and FRBs, especially considering current FRB surveys are incomplete at low fluences corresponding to energies that could overlap with the SGR 1935+2154 burst. Altogether, this detection suggests that Galactic-analogue magnetars could explain part of the FRB population.

However, the analogue is not perfect. For one thing, SGR 1935+2154 is still orders of magnitude fainter than the brightest of FRBs. For example, FRB 180110 is at a DM-estimated luminosity distance of $\sim 3 - 5$ Gpc (Pol et al., 2019), but was detected with a fluence of $\sim 390$ Jy ms (Shannon et al., 2018). This implies an isotropic-equivalent energy of $\sim 10^{42} - 10^{43}$ erg, about eight to nine orders of magnitude more energetic than SGR 1935+2154. Additionally, no Galactic magnetar has matched the activity of FRB 121102, which has been observed to produce 18 bursts with energies $> 10^{37}$ erg within a span of just 30 minutes (Gajjar et al., 2018) and has been emitting bright bursts nearly continuously over the course of the 8 years since its discovery (Spitler et al., 2014, 2016). If particularly prolific repeating FRBs like FRB 121102 are produced by active magnetars, this suggests that they must differ in some way from the Galactic magnetar population.

On the theoretical side, some FRB models predict constraints on maximum and minimum FRB luminosities. For example, Lu & Kumar (2019) and Lu et al. (2020) derive a minimum and maximum luminosity range of $10^{36}$ to $10^{48}$ erg s$^{-1}$ assuming FRBs arise from coherent curvature emission in the magnetosphere of magnetars. Other models predict comparisons between energies emitted at different wavelengths, e.g., Metzger et al. (2019) describes a synchrotron maser model where FRB emission originates outside the magnetar magnetosphere, which predicts a coincident X-ray burst with energy $E_X/E_R \sim 10^{-5}$ times the emitted FRB (consistent with the energetics of an X-ray counterpart detected along with the recent SGR 1935+2154 burst; Mereghetti et al., 2020; Zhang et al., 2020). This is in contrast to magnetospheric models which either predict $E_X/E_R \sim 1$ or a wide and unconstrained range of ratios (e.g., Kumar et al., 2017; Lu & Kumar, 2019; Chen et al., 2020; Lyubarsky, 2020).

Beyond maximum and minimum brightness and multiwavelength limits, there are other crucial tests of FRB origins that depend on FRB brightness as a metric. Here I highlight a few:

- **Log N−Log S:** Since FRBs exist at cosmological distances, the fluence distribution across the entire population will be shaped by a combination of the intrinsic luminosity function and the evolution of progenitor prevalence with redshift. In the simplest case of a sample of sources of given intrinsic luminosity spread uniformly in flat (Euclidean) space, the number of FRBs above a given fluence threshold is expected to follow a power law $N(> F) \propto F^{\alpha}$, where $\alpha = -1.5$ (Petrosian, 1969). This case is expected to hold for FRBs originating from the local Universe. Deviations from this simple model can encode information about FRB progenitors at cosmological distances, for example if the rate of FRBs tracks the star formation history of the Universe (e.g., Niino, 2018). See Macquart & Ekers (2018c) and Macquart & Ekers (2018b) for an in-depth discussion of how different progenitors and luminosity functions are expected to manifest in the fluence distribution. While recent observational efforts undertaken to quantify the FRB fluence distribution using a simple power law model have yielded varying results from $\alpha = -0.83$ to $-2.2$ (e.g., Agarwal et al., 2020; James et al., 2019), these studies have been largely limited by small number statistics and uncorrected instrumental selection functions.
- **Energy Distribution:** The distribution of energies emitted by the FRB population as a whole as well as the population of bursts coming from a

single repeater can inform progenitor models. For example, the high-energy bursts from individual magnetars and giant radio pulses from pulsars have been well-described by power law distributions ($N(>E) \propto E^{\gamma}$) with indices of $\gamma = -0.6$ to $0.7$ (Göğüş et al., 1999) and $\gamma = -2$ to $-3$ (Popov & Stappers, 2007; Bera & Chengalur, 2019), respectively. Whereas sources with accretion disks like AGN exhibit a log-normal flux distribution (Kunjaya et al., 2011). Repeating FRB energy distributions showing similarity to any of the above would count as evidence toward their origins. As of yet, a consensus hasn't been reached on the power law index of FRB 121102, with values derived from different telescopes ranging from $\gamma = -0.6$ to $\gamma = -1.8$ (e.g., Law et al., 2017; Gourdji et al., 2019). Recently Cruces et al. (2021b) suggested that this discrepancy in values may mean that a single power law is a poor descriptor of the energy distribution over many orders of magnitude.

- **Dispersion−Brightness Relation:** An observational correlation between burst dispersion measure and fluence was first suggested by Shannon et al. (2018). A Kolmogorov–Smirnov comparison of the 20 ASKAP bursts and 26 Parkes bursts detected at the time showed that their dispersion measure distributions were statistically different, with the median DM of the ASKAP sample ($441\,\mathrm{pc\,cm^{-3}}$) being a factor of two smaller than the Parkes sample ($880\,\mathrm{pc\,cm^{-3}}$). Shannon et al. (2018) conducted a simple simulation to show that this discrepancy is likely not due to DM-detection biases from frequency and time resolution differences between the ASKAP and Parkes systems.

At the same time, the fluences detected by Parkes are on average lower those from ASKAP, suggesting a negative correlation in DM−brightness space. Using DM as a distance proxy, Shannon et al. (2018) showed that the ASKAP and Parkes samples have qualitatively overlapping energy distributions (see Figure 2 of Shannon et al. 2018), suggesting that high-fluence bursts in the ASKAP sample are nearby analogues of the more distant events from the Parkes sample. Determining whether this relation persists in a larger population of bursts detected with a uniform selection function could provide an alternate method for probing the cosmological evolution of the FRB energy distribution.

- **Luminosity−Width Comparison:** A potential comparison of interest is FRB intrinsic burst luminosity versus intrinsic temporal width. Hashimoto et al. (2019) found a weak but statistically significant positive correlation between luminosity and width for a subsample of 27 non-repeating bursts detected across multiple telescopes, but Hashimoto et al. (2020) failed to find a similar correlation for repeating FRBs, supporting the possibility of different physical origins for non-repeating and repeating FRBs. In particular, Hashimoto et al. (2019) highlights scenarios for non-repeating FRBs like AGN jets interacting with the surrounding medium (e.g., Romero et al., 2016), neutron star-asteroid collisions (e.g., Geng & Huang, 2015), and magnetars interacting with supernova remnants (e.g., Lyubarsky, 2014; Metzger et al., 2019), all of which predict a positive correlation between luminosity and width.

The revelatory potential of these tests and comparisons motivates precise physical measurements of FRB brightnesses, as well as the brightnesses of any Galactic analogues like SGR 1935+2154 that might be detected in FRB surveys. Plus, the FRB brightnesses are crucial components for FRB rate determinations and follow-up strategies.

Attempts to make general statements about the energetics of the FRB population (or any other intrinsic parameter distribution) have so far been restricted by the small number statistics associated with bursts detected at individual telescopes. Past efforts to combine FRB samples across several surveys have been limited by biases introduced in the data due to undefined and varying instrumental selection functions. CHIME is a novel radio telescope with a 200 square degree field of view, a large collecting area, and a powerful software backend, making it uniquely primed for detecting and characterizing a large number of FRBs from a constant selection function (CHIME/FRB Collaboration et al., 2018). In this thesis I will describe an automated pipeline that I have developed to flux calibrate the 535 bursts detected in the first CHIME FRB catalog, with the aim to constrain one of the most fundamental FRB qualities: their brightnesses.

The layout of the rest of this thesis is as follows: in Chapter 2 I provide an overview of the CHIME telescope and CHIME/FRB experiment, focusing on aspects of the project that are related to FRB flux calibration, such as the beam sensitivity pattern, burst localization, and post-detection burst characterization. In Chapter 3, I first review the basic principles of radio transient and FRB flux calibration, and then I provide a detailed description of the CHIME/FRB flux calibration pipeline that I have developed for this thesis. Finally, in Chapter 4, I give an overview of the preliminary scientific contributions of this pipeline and outline paths forward for the next pipeline iteration.

# CHAPTER 2
# Overview of CHIME/FRB

The Canadian Hydrogen Intensity Mapping Experiment (CHIME) is a radio telescope situated near Penticton in British Columbia, Canada, on the grounds of the Dominion Radio Astrophysical Observatory (DRAO). As the name suggests, the CHIME telescope was originally conceived for cosmological purposes, specifically for mapping baryon acoustic oscillation (BAO) features in redshifted neutral hydrogen to constrain the dark energy equation of state (this effort is known as CHIME/Cosmology; Newburgh et al. 2014). However, it was soon realized that many of the design features motivated by this initial goal — notably CHIME's large collecting area (8,000 $\mathrm{m}^2$), instantaneous field of view ($\sim$200 square degrees), and bandwidth ($400 - 800\,\mathrm{MHz}$) — also positioned CHIME as an unprecedentedly powerful tool for FRB detection. Strategic upgrades to the correlator and onsite computing architecture enabled the development of the CHIME Fast Radio Burst Project (CHIME/FRB): a real-time software pipeline for blind FRB detection operating commensally and continuously on the CHIME data stream.

Early estimates suggested that CHIME/FRB would be capable of detecting multiple FRBs per day, the highest detection rate among contemporary FRB surveys (CHIME/FRB Collaboration et al., 2018). Since commissioning, CHIME/FRB has largely lived up to these estimates, detecting hundreds of bursts within its first few years of operation. These detections have revolutionized the FRB field, leading to important results including confirmation of the existence of FRB emission down to 400 MHz (CHIME/FRB Collaboration et al., 2019a), the discovery of 18 new repeating FRB sources (CHIME/FRB Collaboration et al., 2019b; Fonseca et al., 2020), the discovery of periodicity in a repeating FRB (CHIME/FRB Collaboration et al., 2020b), the detection of a bright FRB-like burst from a Galactic magnetar (CHIME/FRB Collaboration et al., 2020a), and a catalog of 535 new FRBs detected from a uniform selection function (CHIME/FRB Collaboration et al., 2021).

In the following chapter I review the hardware and software of the CHIME/FRB system that made all of these discoveries possible, with a focus on aspects of the telescope related to the flux calibration pipeline. In particular, I summarize the physical configuration of the CHIME telescope (§2.1) as well as the path that an FRB takes through the CHIME system from first arrival at the telescope (§2.2) to detection by the CHIME/FRB backend (§2.4). I also include a description of the CHIME/FRB beamforming algorithm and primary beam (§2.3) as well as selected CHIME/FRB burst characterization analyses (§**??**), as they are integrated with the functionality of the flux calibration pipeline.

## 2.1 Reflecting Structure and Analog Signal Chain

The CHIME telescope consists of four 20-m by 100-m parabolic steel-mesh cylinders oriented parallel to each other with their axes aligned in the North-South direction (see Figure 2–1). Altogether, the four cylinders boast a total collecting area of 8,000 $m^2$, equivalent to over five hockey rinks or one and a half American football fields. This unusual design allows CHIME to be sensitive to a large field-of-view spanning $\sim$120° North-South along the local meridian and $\sim 1.3 - 2.5°$ East-West (depending on frequency). With no moving parts, CHIME operates as a transit telescope, scanning the entire sky north of declination $-11°$ every day as it passes overhead with the rotation of the Earth. Sources north of 70° declination are circumpolar and within the CHIME field of view in both "upper" and "lower" transit. These sources are observed twice per day on either side of the North Celestial Pole.

Each of the paraboloidal cylinders has a focal length of 5 m ($f/D = 0.25$). Along the focal line of each cylinder is a steel axis structure supporting an array of 256 dual-polarization feeds linearly-spaced 30.48 cm apart (or 1,024 feeds in total), forming a highly redundant grid pattern. The cloverleaf design of these feeds was developed particularly for CHIME, optimized for economical mass production and wide-band sensitivity in the $400 - 800$ MHz frequency range (Deng et al., 2014). This bandwidth was chosen to match the frequency of the 1.4 GHz neutral hydrogen line when redshifted between $z = 0.8$ (800 MHz) and $z = 2.5$ (400 MHz).

The light from any given FRB reflects off a CHIME cylinder and focuses onto the focal line, where its electric field registers as a voltage signal in the feeds. This signal is immediately amplified by first-stage low noise amplifiers (LNAs) attached directly to the feeds before traveling through a network of 50-m coaxial cables to two radio-frequency (RF) insulated shipping containers located between each pair of cylinders (the East and West receiver huts). With 2,048 inputs to transport (two for each polarization for each feed), this network consists of more than 100 km worth of cable. Once in the containers the signals pass through a second set of amplifiers and a $400 - 800\,\mathrm{MHz}$ bandpass filter. This entire signal chain introduces noise on the order of $\sim$20 K (see §3.1).

Figure 2–1: The CHIME telescope on September 15, 2016. Photo taken from CHIME/FRB Collaboration et al. (2018).

### 2.2 Correlator

The next step in the signal chain involves a system of custom electronics and compute nodes called the correlator. As the physical hardware of the telescope is fixed and stationary, CHIME relies on computationally-intensive interferometric methods to spatially sample the sky. The correlator is the beating heart of this effort, digitizing and channelizing analog voltages and completing a large volume of parallel data manipulations in real-time to facilitate downstream cosmological imaging and FRB detection.

Table 2–1: Salient properties of the CHIME telescope and CHIME/FRB backend

| **Parameter** | **Value** |
|---|---|
| Longitude | $119°37'03''.00$ West |
| Latitude | $49°19'13''.08$ North |
| Altitude | 547.9 m |
| Collecting area | 8,000 $m^2$ |
| Frequency range | $400 - 800$ MHz |
| Polarization | orthogonal linear |
| East-West field of view | $2.5° - 1.3°$ |
| North-South field of view | $\sim 120°$ |
| Focal ratio, $f/D$ | 0.25 |
| Receiver temperature | 50 K |
| Number of synthesized beams | 1,024 |
| Synthesized beam width (FWHM) | $40' - 20'$ |
| FRB search time resolution | 0.983 ms |
| FRB search frequency resolution | 24.4 kHz |
| Source transit duration | Celestial Equator: $10 - 5$ min<br>Declination = 45°: 14–7 min<br>North Celestial Pole: 24 hr |

**Note**: table is based on Table 1 in CHIME/FRB Collaboration et al. (2018). Where two numbers appear, they refer to the low and high frequency edges of the band, respectively. The quoted receiver temperature is a nominal value based on design specifications that includes ground spillover.

The CHIME correlator is a hybrid "FX" architecture, which is computationally efficient in the limit of many frequency channels (Denman et al., 2020). In this design, the data from each feed are first digitized and separated into frequency channels (in the "F-engine") before being spatially cross-correlated with other feeds to form beams on the sky (in the "X-engine").

Housed in the East and West receiver huts, the F-engine is composed of 128 custom "ICE" motherboards (Bandura et al., 2016b,a). These boards sample the analog feed voltages 800,000,000 times per second with 8-bit accuracy, leading to a total input data rate of 13.1 Tb per second for all 2,048 feeds combined. This information is then transferred to Field Programmable Gate Arrays (FPGAs) located on each motherboard, where computations are completed to channelize the data using a polyphase filterbank technique (PFB). The 400 MHz bandwidth for a each timestream is channelized into 1,024 frequency bins (each 390 kHz wide) with a sampling cadence of 2.56 $\mu$s. It is at this step that interferometric calibration occurs. Complex gain and phase values are applied to each data stream to correct for instrumental delays, enabling accurate beamforming (more details on this process in §2.3.2). The resulting data are rounded to 4 + 4-bit complex numbers, yielding a total data rate of 6.5 Tb per second leaving the F-engine.

The data are then transferred by high-speed fiber optic cable to the X-engine, which is housed in two more RF-insulated 40-ft shipping containers on the East side of the telescope (pictured on the right-hand side of Figure 2–1). The X-engine is a 256-node cluster with each node essentially consisting of four GPUs. Each GPU is responsible for processing a single frequency channel for all 2,048 feeds. The four frequency channels entering each node are separated by 100 MHz to avoid continuous gaps in the bandwidth due to node failures (Denman et al., 2020).

The X-engine is responsible for producing different data products for each commensal CHIME experiment. For CHIME/Cosmology, the X-engine correlates the digital signals between each pair of feeds to produce a set of $N^2$ time-averaged "visibilities" (where $N$ refers to the 2,048 feed inputs), which can then be manipulated offline to produce broad sky images of redshifted neutral hydrogen. For

CHIME/FRB, the X-engine combines signals from all the feeds to form instantaneous beams gridded across the CHIME field of view. This beamforming step, referred to as the "Level-0" or L0 process (see §2.4), runs in a series of `OpenCL` kernels on the GPUs. More details on this process are provided in the next section, see §2.3.

After beamforming, time resolution is traded for frequency resolution as the data are upchannelized by 128 times to 3 kHz frequency channels via an FFT. 8-bit Stokes-I intensities are then formed by squaring and summing the complex valued polarizations. Finally, the data are downsampled in time by a factor 3 and in frequency by a factor of 8, resulting in data with 0.983 ms time resolution and 16,384 frequency channels (24 kHz resolution). At a final rate of 131 Gbps, data are then sent to the CHIME/FRB detection pipeline (see §2.4).

### 2.3 CHIME/FRB Beamforming

The motivating tenet of the CHIME/FRB beamforming scheme is to spatially sample the CHIME field of view at the highest sensitivity and frequency resolution possible within the computational budget. To accomplish this goal, CHIME/FRB forms a closely packed grid of 1,024 static beams for each polarization and frequency. Arranged in four columns East-West and 256 rows North-South, the synthesized beams tile the $\sim$200 square degree span of CHIME's primary beam, allowing for high-sensitivity FRB detection and localization across a large instantaneous swath of sky.

Forming so many simultaneous beams is an incredibly computationally intensive process. To facilitate this process, CHIME/FRB uses a clever algorithm called Fast Fourier Transform (FFT; Ng et al. 2017) beamforming to form beams in the North-South direction. This algorithm leverages the regular grid-like layout of CHIME's feeds to relax the $\mathcal{O}(N^2)$ runtime of conventional beamforming to $\mathcal{O}(N \log N)$, where $N$ is the number of feeds (e.g., Tegmark & Zaldarriaga, 2009, 2010; Masui et al., 2019). This novel approach — which has only been implemented in a handful of other radio astronomy experiments to-date (e.g., the Waseda Radio Telescope, Otobe et al. 1994; the Medicina BEST-2 Array, Foster et al. 2014) — dramatically reduces the computational cost from beamforming, which is crucial for the feasibility of the CHIME/FRB real-time detection pipeline.

However, for all of its computational advantages, FFT beamforming also significantly complicates the CHIME/FRB bandpass response. In particular, the FFT beams exhibit complex structure as a function of frequency, which manifests as sharp discontinuous and periodic spectral features that change shape significantly over small displacements on-sky. As a further complication, these FFT beamforming structures are superimposed over bandpass ripples in the primary beam response, which arise due to standing waves in CHIME's cylindrical design. Untangling these spectral features from the intrinsic FRB spectra is a nontrivial task, which stands as one of the most fundamental challenges to CHIME/FRB flux calibration. In this section I review the synthesized beamforming process for CHIME/FRB (§2.3.1) and CHIME's primary beam response (§2.3.2) in some detail, to elucidate decisions made in the design of the flux calibration pipeline, the subject of this thesis.

### 2.3.1 Synthesized Beam Response

Like many interferometric experiments used for time-domain astronomy, CHIME/FRB operates in phased array mode, where localized beams are formed by summing feed signals with different time delays to cause constructive interference in particular directions. Mathematically, a composite beamformed signal at a particular time, $b(t)$, directed at steering angle $\theta_m$, can be written as

$$b(t, \theta_m) = \sum_{n=1}^{N} a_n \, x_n[t - \tau_n(\theta_m)] \tag{2.1}$$

where $N$ is the number of feeds, $x_n$ are the digitized feed signals, $a_n$ are the constant gains applied to correct for instrumental delays (described in §2.3.2), and $\tau_n(\theta_m)$ are the time delays required to point the beam to the specified direction (Mucci, 1984).

The above definition is formulated in the time domain, where $|b(t, \theta_m)|^2$ is taken as the intensity output of the beam. However, it is often advantageous to beamform in the frequency domain by taking the Fourier transform:

$$B(f, \theta_m) = \mathcal{F}\left\{b(t, \theta_m)\right\} = \sum_{n=1}^{N} a_n \, X_n[f] \, e^{-i2\pi f \tau_n(\theta_m)} \tag{2.2}$$

where $X_n[f] = \mathcal{F}\{x_n[t]\}$ and $f$ is the observation frequency. For CHIME/FRB beamforming, $X_n[f]$ are the channelized complex data from the F-engine. In this formulation, the intensity output of the beam is equivalently taken to be $|B(f, \theta_m)|^2$ (Maranda, 1989). Notice that forming $N$ beams using this method would take $\mathcal{O}(N^2)$ time.

The magic of FFT beamforming comes with the realization that the time delays required for beamforming an array of linearly spaced feeds with separation $d$ are given by: $\tau_n(\theta_m) = n\frac{d}{c}\sin(\theta_m)$, where $n$ is the feed index and $c$ is the speed of light. If we choose to form beams at steering angles

$$\theta_m = \sin^{-1}\left(\frac{c\,m}{f\,N\,d}\right) \tag{2.3}$$

then, by substituting $\tau_n(\theta_m)$, Equation 2.2 becomes:

$$B(f, \theta_m) = \sum_{n=1}^{N} a_n\, X_n[f]\, e^{-i2\pi mn/N} \tag{2.4}$$

This is just a discrete Fourier transform, mapping the spatial offsets in feed positions to angular beams on the sky. This expression can be evaluated to form $N$ beams using a fast Fourier transform in $\mathcal{O}(N \log N)$ time.

FRB surveys typically need to maximize broadband sensitivity to a single sky location to increase detection significance and allow for FRB spectral characterization. However, as dictated by Equation 2.3, the steering angles of FFT-formed beams are dependent on frequency as $\sin(\theta_m) \propto 1/f$, causing formed beams at higher frequencies to be closer together than those at lower frequencies. This effect chromatically smears the sensitivity pattern of a single beam (indexed by $m$) across the sky. To reduce this effect with CHIME/FRB, we use a method called "nearest-neighbour clamping" (Maranda, 1989; Ng et al., 2017). First, the 256 North-South feed inputs are zero-padded by a factor of two so that the FFT in Equation 2.4 forms 512 closely packed beams by Fourier interpolation. These beams are then subsampled (clamped) to form 256 beams at the desired pointings.

The most sensitive beam for each frequency is chosen for each pointing, forming a "Frankenstein" beam of combined components. This process is illustrated in the left panel of Figure 2–2 for a few CHIME/FRB beams near zenith, where the steering angle is North-South zenith angle.

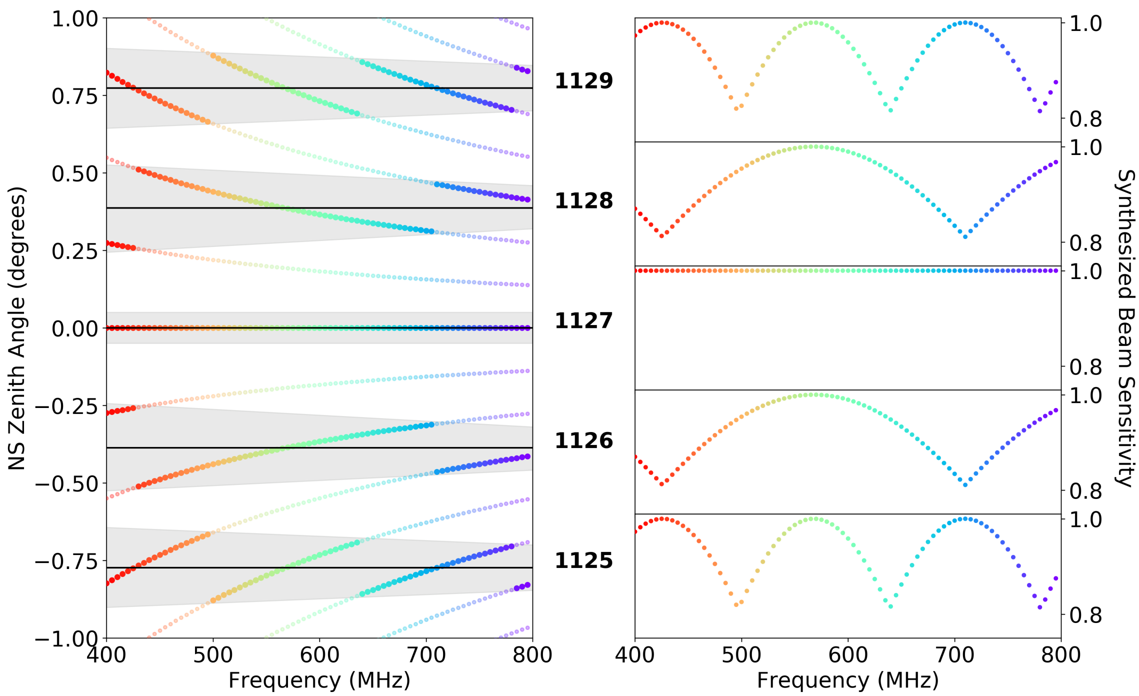


Figure 2–2: Illustration of the chromatic effect of FFT beamforming and clamping across the CHIME/FRB bandwidth. Figure adapted from Ng et al. (2017). (Left) The position of the dots represents the location of nominal peak sensitivity for each formed beam at a particular frequency. The coloured dots enclosed by the gray areas are the selected nearest-neighbour clamped beams, whereas the fainter coloured dots are the additional discarded beams. Horizontal black lines represent the nominal clamped beam centres, labeled by the corresponding beam number. (Right) The sensitivity versus frequency at the centre of each of the clamped beams in the left panel.

The full CHIME/FRB beamforming pipeline is a hybrid of the two techniques described above. FFT beamforming (Equation 2.4) with clamping is used to form the 256 rows of beams in the North-South direction, while brute-force phasing (Equation 2.2) is used to form the four columns of beams in the East-West direction. For the majority of the CHIME/FRB experiment duration, the configuration of the beam grid has been fixed. In the North-South direction, the beams are equally spaced in $\sin\theta$, where $\theta$ is the zenith angle. One beam is centred at zenith ($\theta = 0°$), while 127 beams tile to $\theta = -60°$ South and 128 beams tile to $\theta = 60°$ North. Beams get more elongated North-South at larger zenith angles, because the projected baselines between feeds shorten nearer to the horizon. In the East-West direction, the configuration is asymmetrical: one column is centred along the meridian while another column is formed 0.4° to the East and two other columns are formed 0.4° and 0.8° to the West (see the left panel in Figure 2–5). The resulting formed beams are labeled with integers according to their location in the grid: the rows South to North have an index ranging from 0 to 255 and the East to West columns add a factor of 0, 1000, 2000, or 3000 to the resulting index. For example, the beam at zenith in the meridian column corresponds to $127 + 1000 = 1127$.

The full implications of this hybrid beamforming scheme are complex and more effectively shown than told. Aided by figures, in the remainder of this section and the next I highlight some more detailed but central aspects of CHIME/FRB's beam response that play a role in the design of the flux calibration pipeline. These figures make use of both the synthesized and primary beam models (see

§2.3.3), which have been developed by other members of the CHIME/FRB and CHIME/Cosmology teams, respectively, in addition to steady source transit data from the flux calibration pipeline.

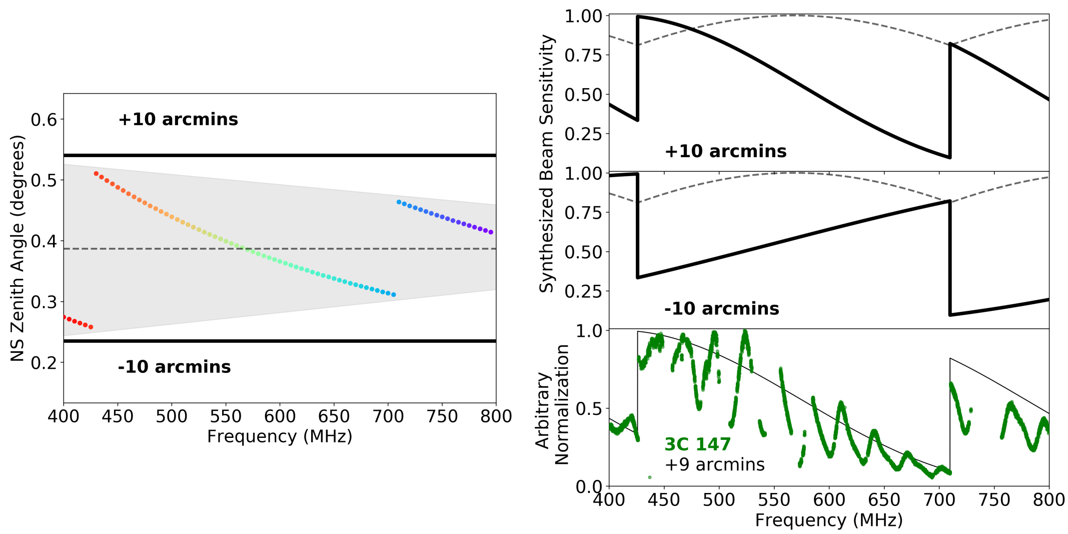


Figure 2–3: Variations in the response of synthesized beam 1128 as a function of frequency and location within the beam. (Left) A plot of the chromatic spreading of the FFT formed beams and clamping, similar to Figure 2–2 but focused on beam 1128. The thin grey dashed line indicates beam centre while the thick black lines mark 10 arcminute offsets on either side of beam centre. (Right) The thick black lines represent the sensitivity versus frequency 10 arcminutes above the centre of beam 1128 (Top) and 10 arcminutes below (Middle). The thin grey dashed line in each of these plots represents the response at beam centre. The bottom plot shows the synthesized beam response at the transit location of radio galaxy 3C 147, along with a background-subtracted and normalized observed spectrum of the source with CHIME/FRB.

**Severe Variations in North-South Spectral Structure:**

As illustrated in Figure 2–2, the clamping algorithm used by CHIME/FRB introduces periodic cusp-like structures in the resulting sensitivity versus frequency due to chromatic spatial smearing within the clamped beam extent. These cusp-like features, called “clamps,” change shape and severity depending on the beam being considered as well as the location on-sky. The right panel of Figure 2–2 shows the beam response versus frequency at the centre of five clamped beams surrounding zenith. Notably, the number of clamps in the CHIME/FRB bandwidth increases with the zenith angle of the clamped beam, which can cause a particularly complex response pattern for beams closer to the horizon. For example, beam 1128 at $0.4^\circ$ from zenith has only two clamps, while beam 1024 at $-44^\circ$ zenith angle has a total of 145 clamps (see the bottom panel of Figure 2–4).

The shape of the clamps also changes over spatial displacements within a single beam. In particular, as you move away from beam centre, the cusp-like structures turn into more severe discontinuous jumps. Figure 2–3 demonstrates this by showing the response of beam 1128 at the edges of its nominal $\sim 20'$ North-South full-width half-maximum (FWHM; at 600 MHz). The bottom right panel shows a CHIME/FRB observed spectrum from the transit of radio galaxy 3C 147 across beam 1128, which exhibits the sharp clamping discontinuities at $\sim$430 and $\sim$710 MHz. Note that the remaining oscillation in the spectrum is the 30 MHz ripple from the primary beam response (see §2.3.2).

Figure 2–4 shows a sampling of other CHIME/FRB synthesized beam responses corresponding with data from the transits of several bright supernova remnants and radio galaxies. This sampling covers a wide range of zenith angles, beam numbers, and offsets from the beam centre, which gives a qualitative demonstration of the complexity of the North-South synthesized beam response.

**Chromatic Sidelobes in East-West Beam Profile:**

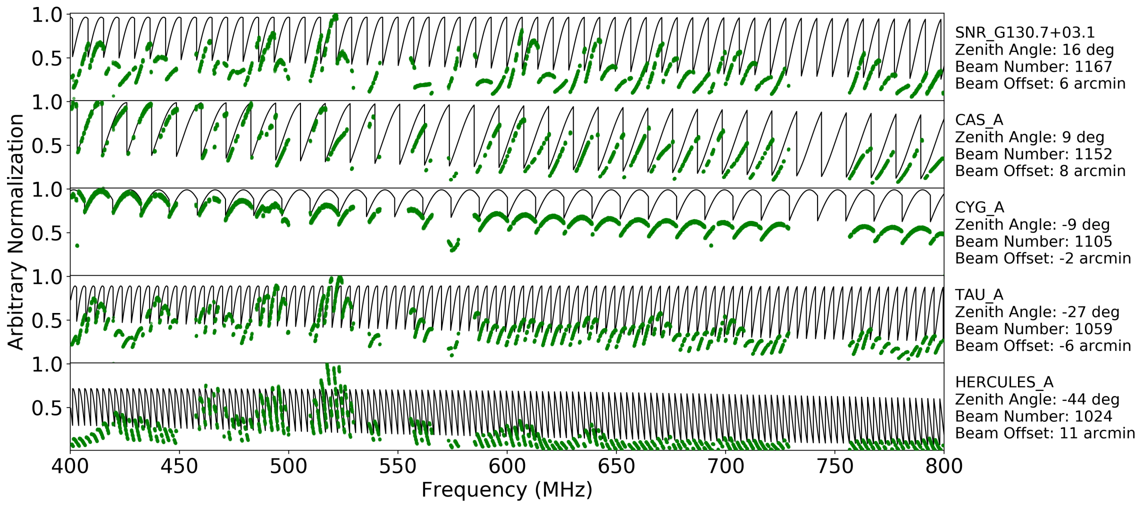


Figure 2–4: A sampling of CHIME/FRB synthesized beam responses at the locations of several bright supernova remnant and radio galaxy transits (black lines), along with background-subtracted and normalized observed spectra of each source (green dots). The source name, zenith angle, beam number, and offset from beam centre are labeled to the right of each plot.

The East-West profile of the CHIME/FRB formed beams consists of an intrinsic profile governed by exact phasing, which is then further attenuated by the primary beam response. The exact phasing profile has significant sidelobes of increased sensitivity due to the periodic nature of the Fourier transform (Equation 2.2). These sidelobes remain significant even with attenuation from the primary beam, as shown in the left panel of Figure 2–5. In particular, beams in the 3000 column have a sidelobe that is more sensitive than the main lobe in the middle of the CHIME/FRB band.

Another notable property of these sidelobes is their smeared frequency response. The sidelobes of a given beam spread chromatically East-West, as shown in the right panel of Figure 2–5, with lower frequencies spreading further than higher frequencies. This means that a burst detected in the East-West sidelobe of a given synthesized beam may be completely attenuated to the noise floor at some frequencies, making it appear band-limited. This is demonstrated in the top right panel of Figure 2–5.

### 2.3.2 Primary Beam Response

The full CHIME/FRB beam response is a combination of the synthesized beam pattern described in section §2.3.1 and the primary beam response, which is governed by the physical structure of the telescope itself. The fundamental primary beam response is that of a single feed over an ideal cylindrical reflector, which varies smoothly and spans $\sim 120°$ North-South along the local meridian and $\sim 2.5 - 1.3°$ East-West ($400 - 800$ MHz). However, in practice, the primary beam exhibits more complicated variations resulting from reflections within the telescope and cross-talk between neighbouring feeds on the focal line. The left panel of Figure 2–6 shows representative East-West and North-South cuts through the primary beam. One of the strongest components of these variations is a 30 MHz ripple in the primary beam response as a function of frequency, caused by cross-talk from a standing wave in the 5 m distance between the focal line and the receiver. The form of this ripple is dependent on zenith angle, as shown in the right panel of Figure 2–6.

The shape of these variations is also affected by the interferometric gain calibration process. As discussed in section §2.3.1, beamforming involves combining signals between feeds with geometric time delays to create peak sensitivity at

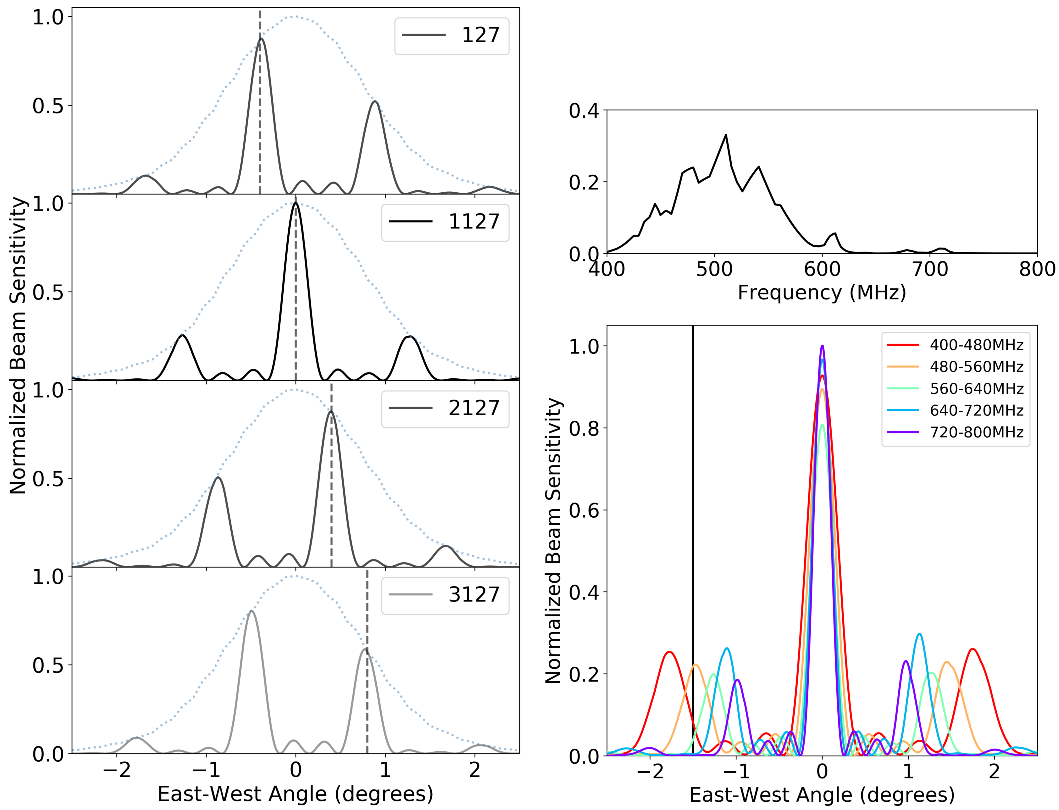


Figure 2–5: Representations of the East-West sensitivity profile of the CHIME/FRB formed beams. (Left) A plot of the East-West profile for the row of beams at zenith at 600 MHz (solid lines), which are attenuated by the primary beam envelope (dotted blue lines). The dashed vertical lines indicate the main lobe of each beam. (Right) The East-West profile of beam 1127. (Bottom) The profile is split into five subbands, with coloured lines representing each subband (red corresponding to low frequencies and purple corresponding to high frequencies). (Top) The frequency response of the beam at the sidelobe location marked by a vertical black line in the bottom plot.

specific locations on-sky. To do this successfully, additional instrumental delays must be accounted for. This is accomplished by solving for the delays empirically, using observations of a steady source of known position and brightness. Complex gain and phase values representing these delays are then applied to the digital data stream to calibrate the response at the location of the source. For CHIME/FRB, gain calibration is typically completed using daily meridian transits of Cygnus A (although not always, see section §3.2).

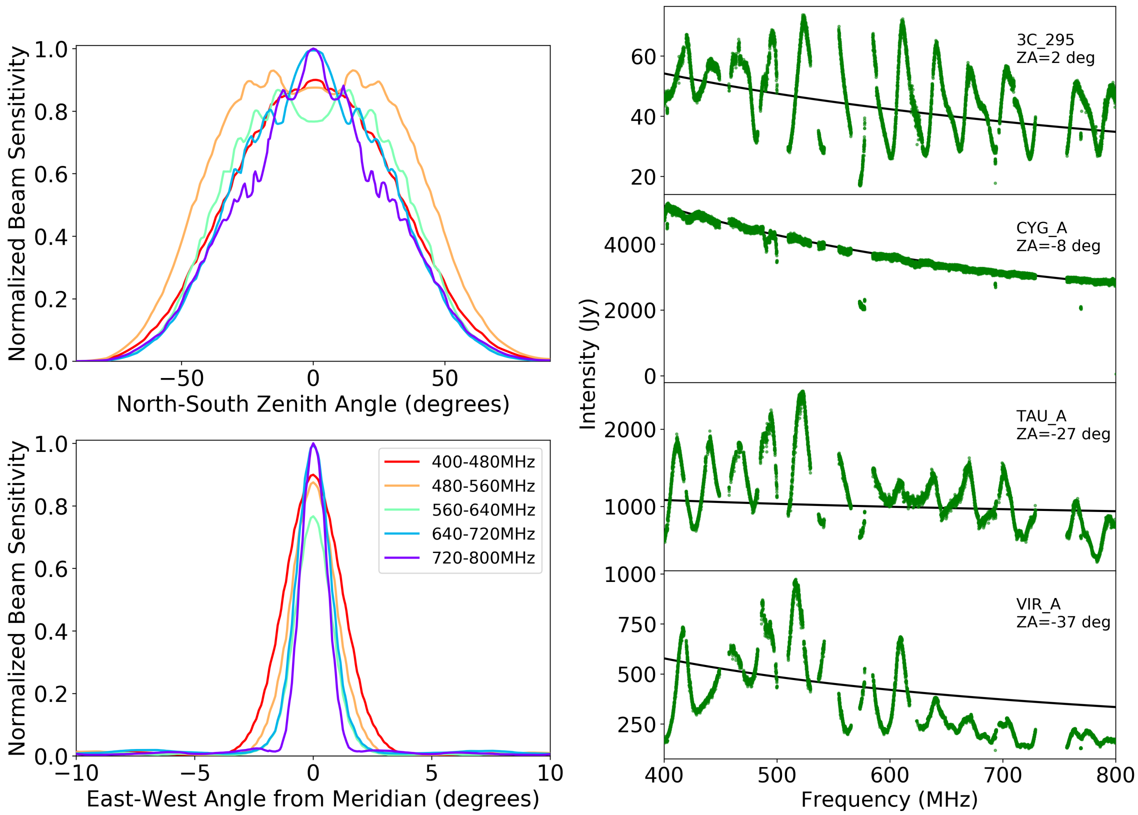


Figure 2–6: Representations of the sensitivity profile of the CHIME primary beam. (Left) Plots of the North-South profile along the meridian (Top) and the East-West profile at zenith (Bottom), split into five subbands, with coloured lines representing each subband (red corresponding to low frequencies and purple corresponding to high frequencies). (Right) Observed background-subtracted spectra for a number of steady sources at meridian transit, each corrected for the synthesized beam response (green dots). The solid black lines show the actual known spectrum of the steady source (see §3.3.2). Each plot is labeled with the source name and zenith angle.

This gain calibration process corrects for delays introduced in the analog chain between the feeds and the correlator, but it cannot completely correct for features arising from reflections and cross-talk within the reflecting cylinder, which vary as a function of location on sky. As a result, sources detected at or near the location of the complex gain calibrator (e.g., Cygnus A) will show their true spectrum, while away from the calibrator the detected spectrum will be modulated by the 30 MHz ripple, as demonstrated in the right panel of Figure 2–6.

### 2.3.3 Beam Model

The last two sections have demonstrated the well and truly Byzantine nature of the CHIME/FRB beam pattern, characterized by complex and rapidly varying sensitivity across both field of view and bandwidth. Improvements in our ability to predict the CHIME/FRB beam response will lead to improvements in our ability to separate beam attenuation from intrinsic FRB features. Beam modeling is therefore a crucial component of many CHIME/FRB analyses, including flux calibration. The CHIME/FRB beam model exists as its own code repository, developed by members of both the CHIME/FRB (FFT beamforming) and CHIME/Cosmology (primary beam) research teams.

The sensitivity patterns from FFT beamforming are purely digital and exactly deterministic, so modeling them is a relatively simple matter of re-calculating the beamforming process. CHIME/FRB beam model code simulates the formed beam positions and shapes analytically from angular delays, taking into account the $\sin\theta$ spacing in North-South zenith angle, the elongation of the beams closer to the horizon, and clamping of the frequency-dependent beam centres.

Characterizing the primary beam is much more difficult, as it involves accounting for complex physical reflections and cross-talk within the CHIME cylinders. This process is made even more challenging since CHIME is a drift-scan telescope that cannot be tilted in elevation and azimuth to probe the beam with bright steady sources. For the first year of the CHIME/FRB's run, the primary beam model was just a smooth cosine in the North-South direction and a Gaussian in the East-West direction, with no attempt to characterize the 30 MHz ripple.

This initial model has been rapidly improved thanks to the heroic efforts of the CHIME/Cosmology team, who have characterized the primary beam response using steady source and solar transits combined with analytical models of the North-South beam response. Concurrent observations by CHIME and the 26 m single-dish Galt telescope at DRAO are used to obtain high-SNR tracks of the transit of Tau A, which are averaged to form a representation of the East-West profile (for an example of this point-source holography technique applied to the CHIME Pathfinder, see Berger et al. 2016). The North-South profile is determined from fitting an analytical cross-talk model to observations of 37 bright radio point sources at different declinations. The resulting primary beam model is an outer product between the East-West and North-South profiles. Comparisons to steady source holography data and solar transits show that this current primary beam model is accurate to $\sim$10% within the main lobe of the primary beam, where most FRBs are detected. Papers describing the primary beam characterization effort in detail are in preparation by the CHIME/Cosmology team (Singh et al. 2021, in preparation; Wulf et al. 2021, in preparation).

## 2.4 Detection Pipeline Overview

After data are digitized, beamformed, and upchannelized in the correlator, they pass to a dedicated CHIME/FRB compute cluster consisting of 132 20-core CPU nodes housed in a separate RF-insulated shipping container from the X-engine. At this point, the data are a large volume of 8-bit integers representing the 1,024 total intensity beams sampled in 16,384 frequency channels and at 0.983 ms time resolution. This creates a total data rate of 142 Gbps or 1.5 PB per day entering the CHIME/FRB compute cluster. Saving all of this incoming data for post-processing is clearly infeasible, so the CHIME/FRB detection pipeline runs in real-time to determine which transient signals are worth saving. The latency of the CHIME/FRB real-time pipeline from arrival of the lowest frequency signal of an FRB to detection is just $\sim 2-3$ s.

The development of this pipeline is a monumental computational feat, requiring highly optimized algorithms and architecture. However, since the focus of this thesis is primarily post-processing of already saved FRB data, here I just provide an overview. For a much more detailed description of the pipeline, see CHIME/FRB Collaboration et al. (2018).

The pipeline is organized into five levels, denoted L0 to L4. L0 encompasses the FFT beamforming and upchannelization steps that take place in the X-engine, which have already been described in detail in sections §2.2 and §2.3. I summarize the rest of the levels here:

**L1 – Detecting Events:** Before any other data manipulation, the beamformed intensity data are first scrubbed of non-astrophysical sources of transient emission, or radio frequency interference (RFI). This is accomplished using a chain of data transforms that iteratively detrend (removing noisy wandering baselines or broad-band RFI) and clip (removing narrow signals above a certain threshold) in both the time and frequency axes. The current configuration of the chain is based on empirical training and adjustment using data from the CHIME Pathfinder (Bandura et al., 2014).

Next, a highly optimized assembly-language-kernelized algorithm called `bonsai` (Smith et al., in prep) is used for pulse dedispersion and detection. As described in section §1.1.1, dedispersion occupies a large portion of any blind FRB search's computation currency, and CHIME/FRB is no exception. L1 is by far the most computationally intensive part of the backend beyond the initial beamforming, and it takes up 128 of the 132 CHIME/FRB allotted compute nodes. `bonsai` tackles both dedispersion and matched filtering in what essentially amounts to a "dedispersion transform," which takes the input intensity data in frequency and time dimensions, and outputs at 4D array of detection SNRs with dimensions spanning DM, arrival time, spectral index, and pulse width. `bonsai`'s dedispersion step is based on the tree dedispersion algorithm (Taylor, 1974; Zackay & Ofek, 2017), but with several optmizations to overcome the normal pitfalls of the algorithm like memory bandwidth bottlenecks and suboptmized sensitivity to the $\nu^{-2}$ dispersive delay sweep (see CHIME/FRB Collaboration et al. (2018) for more details). The CHIME/FRB search covers a range of DM trials from 0 to 13,000 $\mathrm{pc\,cm^{-3}}$ and two spectral indices, $\beta = -3, 3$. Each of these trials is put through five different temporal downsampling trees (covering 1, 2, 4, 8, and 16 times the native 0.983 ms resolution), and each of these trees are matched-filter searched for burst widths up to four times the sampling time (reaching a maximum of $\sim$64 ms).

Given the massive search space, the 4D array of SNR values output from `bonsai` is too large to write out of the L1 nodes. Instead, the SNR matrix is downsampled by taking the maximum SNR within coarse-grained bins spanning a factor of 64 in both time and DM (leading to a resolution of approximately $\Delta\mathrm{DM} \approx 10\,\mathrm{pc\,cm^{-3}}$ and $\Delta t \approx 0.25\,\mathrm{s}$). The resulting coarse-grained triggers pass through a second round of RFI excision using a support vector machine classifier that analyzes the SNR behavior adjacent to each event detection. Remaining non-RFI candidate events from each beam, labeled with coarse metadata ("L1 headers") indicating their position in the detection phase space, are then forwarded to L2/L3.

**L2/L3 –** **Event Refinement and Classification:** The coarse-grained triggers from all 1,024 beams are then streamed to a single node for multi-beam collation and further event classification. First, events are grouped in time, DM, and sky position to identify multiple detections of a single incident FRB in the `bonsai` search space. Similarity of sky position is gauged by beam adjacency while similarity of time and DM are judged by dimension-specific thresholds that reflect the limits of the L1 trigger coarse-graining. After events are formed into groups, their coarse-grained information is combined into a single "L2 header", which passes through one final round of real-time RFI sifting. This round uses a stochastic gradient descent classifier to analyze the configuration of detection SNRs in the beams of each L2 grouping, with the premise that far-field astrophysical sources will have a more focused detection pattern than near-field RFI events. Astrophysical-deemed events then have their positions refined by comparing the detected SNRs from each beam to the CHIME/FRB beam model (this process is called "header localization", and is discussed more in §**??**).

With the improved position estimate, candidates are then compared to a database of known sources of radio transients combined from the ATNF pulsar catalog (Manchester et al., 2005), RRAT catalog[1] , and FRB catalog (Petroff et al., 2016). A likelihood ratio of association with nearby known sources is calculated based on both position and DM, and true associations are decided based on a ratio cutoff determined empirically from simulated events. For events that are not associated with known sources, the predicted maximum Galactic DM contribution is calculated using both NE2001 (Cordes & Lazio, 2002) and YMW16 (Yao et al., 2017). A given event is deemed "extragalactic" if its measured DM exceeds the maximum Galactic contribution by at least $5\sigma$. Events between $2\sigma$ and $5\sigma$ excess are labeled "ambiguous." Other events below $2\sigma$ are marked as potentially new "Galactic" candidates.

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

The final step encompassed in the L2/L3 level of the pipeline is deciding what actions to take for each detection. These decisions are made from set of mutable logical rules based on the classification and real-time properties of each event. Possible actions include sending the event header to the events database (which is done by default), triggering a callback dump of intensity data buffered in L1, triggering a callback dump of baseband (i.e. voltage) data buffered in L0, or sending out an alert to the community. The rules are configurable and can be set for entire groups of sources or for one particular source. For example, in the current configuration of CHIME/FRB, all unknown extragalactic events with detection SNRs above $8\sigma$ have intensity dumps triggered, and those with detection SNRs above $12\sigma$ also have baseband dumps triggered. In addition, if any event associated with Galactic magnetar SGR 1935+2154 (CHIME/FRB Collaboration et al., 2020a) is detected with an SNR above $10\sigma$ then both intensity and baseband callbacks are dumped.

**L4 – Organization and Action:** What L3 commands, L4 organizes and executes. L4 is host to both the events database and the CHIME/FRB archiver. Each event sent past the L1 level, regardless of whether it is classified as astrophysical or RFI in L2, has its header information stored in a relational database for use in pipeline debugging. To give an idea of the scale of this database, more than 100 million L1 headers have passed through the CHIME/FRB pipeline to-date. Events that continue onto the L4 level are each given a unique integer event number for identification and to track any associated data products. By querying the events database, any post-detection analysis can access the real-time detection headers, as well as any additional data products.

Data products are stored on the CHIME/FRB archiver, an on-site storage server offering ~750 TB of space. Both L0 and L1 have ring-buffers which save baseband and intensity data for a short period of time (~30 seconds for baseband and ~4 minutes for intensity). L4 triggers data transfers from these buffers to the archiver according to the actions determined by L3. When an intensity dump is triggered, Stokes-I intensity data are saved for all beams in which an event is detected, as well as immediately adjacent beams. When a baseband dump is triggered, 100 ms of complex voltage data surrounding the burst is saved for each frequency channel and polarization.

After storing data products, L4 also facilitates the final step in CHIME/FRB detection pipeline: human verification. Information and diagnostic plots of each new extragalactic or ambiguous event are displayed in a online portal hosted by L4. Aided by interactive waterfall plots, a diagram of the detected beams, a detection SNR versus DM curve, plots of nearby events in various phase spaces, and tools for known source association, two members of the CHIME/FRB collaboration must separately classify each event before it becomes a verified FRB candidate.

### 2.5 Post-Detection Analyses

The burst properties output to the events database at the end of the real-time detection pipeline are all approximate due to bonsai coarse-graining. To derive science-ready results, CHIME/FRB has a growing arsenal of post-detection software pipelines that operate on dumped intensity and baseband data to refine the parameters of each FRB. The flux calibration pipeline belongs to this group of analyses along with localization routines, morphological parameter fitting, exposure calculators, and fluence threshold calculators. In this section, I give a brief overview of CHIME/FRB's burst localization capabilities and morphological parameter fitting, as the flux calibration pipeline directly depends on the outputs of these two analyses.

### 2.5.1 Modeling Burst Morphology

`fitburst` is a least-squares fitting routine that models the two-dimensional dynamic spectrum of each burst in terms of fundamental burst parameters. The temporal shape of the burst is modeled by a Gaussian intrinsic profile convolved with an one-side decaying exponential function to encapsulate any scattering in the pulse profile (see section §1.1.2). The analytical functional form of this profile is given by Equation 4 of McKinnon (2014). This temporal fit produces values for the arrival time $t_{\mathrm{arr}}$, intrinsic width (the Gaussian $\sigma$), and the scattering time $\tau$ of the burst. The DM is also fit by delaying the temporal profile in each frequency channel according to the $\nu^{-2}$ dispersion sweep. In the frequency dimension, the spectral shape of the burst is fit by a continuous power-law function with a spectral index of $\alpha$ and an extra "spectral running" parameter $\beta$: $I(\nu) = (\nu/\nu_{\mathrm{ref}}) \exp(\alpha + \beta \ln(\nu/\nu_{\mathrm{ref}}))$, for some reference frequency $\nu_{\mathrm{ref}}$. This flexible function allows the spectral profile to vascillate between a regular, broadband power-law and a band-limited Gaussian, covering the wide variety of spectral shapes that have been observed in the FRB population. If a burst consists of multiple subbursts, each component is fit with a separate profile, but DM and scattering are taken to be global components across all fits.

Ideally `fitburst` runs as part of an automated pipeline on every burst that emerges from the real-time pipeline. For each burst, two models are automatically generated: one with zero assumed scattering ($\tau$=0 is fixed) and one where the scattering time is allowed to vary. Since not all detected bursts show multi-path scattering, an F-test is then calculated to determine which model is preferred. In practice, some bursts require additional manual refinements to the fitting, as determined by visual inspection of the fit and residuals. Once a fit has been verified, results are saved to the events database, so that they may be accessed by other analyses. For more information on `fitburst`, see CHIME/FRB Collaboration et al. (2019a), CHIME/FRB Collaboration et al. (2020a), or Fonseca et al. (2020).

### 2.5.2 Localization

The broadest level of localization available to CHIME/FRB is the centre of the formed beam in which a given FRB was detected with maximum SNR. The FWHM of a given formed beam ranges from $\sim 15' - 30'$ at zenith, which gives a rough idea of the likely location of an FRB. However, as discussed in section §2.3.1, CHIME/FRB formed beams have significant East-West sidelobes which spread chromatically out to $\sim 2°$ from the centre of the main lobe. The position of a burst detected in a given beam is therefore degenerate between the main lobe and the sidelobes.

The next narrowest level of localization can be obtained using multi-beam detection information in a given event's L2 header. Header localization works by fitting detected SNRs from each beam to predictions informed by the frequency-dependent beam model for a grid of locations. $\chi^2$ values for all locations are calculated and converted into confidence intervals according to empirical tests of the header localization method using pulses from pulsars of known position. Example localizations are shown in Figure 2–7. Note that multi-beam detections result in better localizations depending on the shape of the beam pattern, but generally the localization degeneracy between the main lobe and the sidelobes persists. Header localization is used in the real-time detection pipeline and it is the only method available in the first CHIME/FRB catalog.

As described in section 2.3.1, the FFT beamforming algorithm produces sharp drops in sensitivity in the CHIME/FRB beam pattern as a function of frequency called "clamps". The shape, amplitude, and frequency position of these clamps change rapidly depending on the observing beam and source declination. Due to the precision of this spatial dependence, modeling clamping structures in FRB intensity data from each of the detected beams is one method for determining burst positions to better accuracy than header localization. As of the writing of this thesis, automated intensity localization is still under development. Initial tests indicate that this method is able to break the degeneracy between the sidelobe and main lobe of a given beam, yielding localizations on the order of $\sim 2-3'$.

Finally, a localization method based on baseband data is available for the limited number of bursts detected with SNRs above $12\sigma$ (according to the rules applied in the L2/L3 level of the detection pipeline, see §2.4). In a similar method to header localization, baseband localization works by forming a tight grid of beams around the initial detection and comparing the map of resulting burst SNRs to that expected from the beam model. Fully automating this localization method is a work in progress. The localization capability of the baseband pipeline scales with the inverse of the burst SNR, where initial tests promise sub-arcminute localizations above an SNR of $\sim$60. A detailed description of baseband localization with CHIME/FRB, including possible systematics and remaining issues, is given in Michilli et al. (2020).

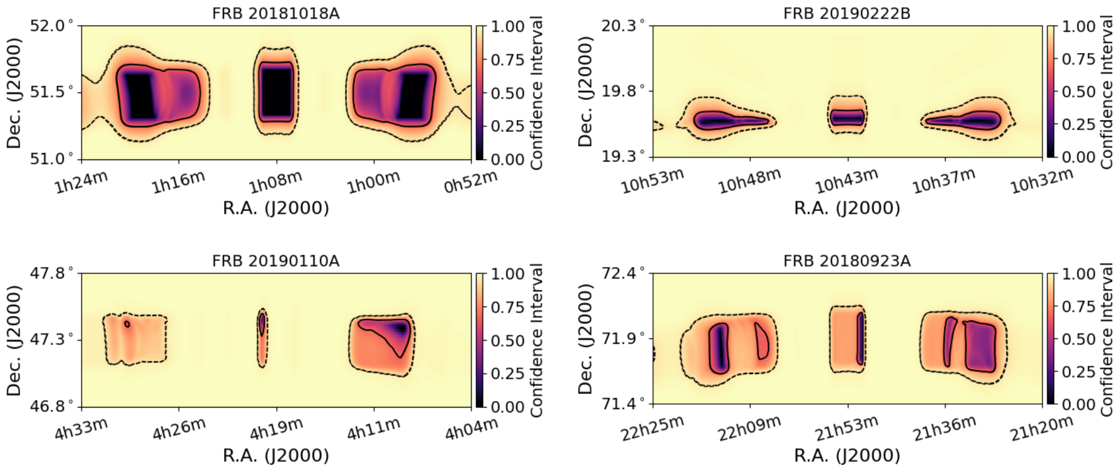


Figure 2–7: Example header localization confidence intervals for different beam detection patterns. Clockwise from the top left: single beam, two beams North-South, two beams East-West, and four beams in a square. In each example, the frame spans 5° in right ascension (scaled by cos(dec)), 1° in declination, and is centred at the beam with the highest SNR detection. Confidence intervals are obtained from contours of constant $\Delta\chi^2$. The colour scale encodes these intervals, such that the area enclosed by a given colour gives the $X\%$ confidence interval. The 68% and 95% intervals are shown with solid and dashed contours, respectively. Note that the common three-region pattern reflects the chromatically smeared side lobes of the formed beams (see also Figure 2–5). Figures and caption courtesy of Alex Josephy.

# CHAPTER 3
# CHIME/FRB Flux Calibration

When light from an FRB arrives at CHIME, it reflects off the cylindrical reflectors before being detected as voltages in an analog chain of feeds, filters, and amplifiers (§2.1). After digitization, the light passes through a series of data manipulations designed to tease out temporal, spectral, and spatial information (§2.2). Each of the steps in this processing chain imprints scalings and variations on the signal that are not intrinsic to the FRB. In particular, the interferometric beamforming scheme used by CHIME/FRB creates variations in sensitivity as a function of sky location that significantly modulate the FRB spectrum (§2.3). As a result, the Stokes-I intensity spectrum that emerges from the detection pipeline is a combination of the intrinsic FRB spectrum and the complicated transfer function of the CHIME/FRB instrument, all represented in some non-physical digitized units.

The fundamental goal of flux calibration is to reverse the effects of the instrumental transfer function and determine what a given FRB looked like before it reached the telescope. In particular, we would like to assign brightness values to the FRB intensity data in physically meaningful units that enable us to probe scientific questions related to FRB energetics (§1.4). The primary goal of this thesis was to implement this process for FRBs detected with CHIME.

I start this chapter with a review of flux calibration techniques commonly used in FRB astronomy (§3.1), followed by a summary of the particular challenges facing CHIME/FRB flux calibration during the development of this thesis (§3.2). Finally, I describe the implementation of the first-pass automated flux calibration pipeline for CHIME/FRB that I have developed (§3.3).

## 3.1 Radio Transient Calibration Techniques

The tried and true method for radio transient flux calibration is the radiometer equation, which appears in various forms throughout pulsar, RRAT, and FRB astronomy. Derived from physical first principles, the radiometer equation provides a process for conversion between instrumental units and physical flux density using just a few fundamental telescope parameters. For a more detailed discussion of the radiometer equation, see Appendix 1.4 of Lorimer & Kramer (2004) or Section 3.6 of Condon & Ransom (2016). In this section I provide just enough highlights to establish key terminology and context (§3.1.1), setting the stage for a discussion of how the radiometer equation is typically used to calibrate data from FRB experiments (§3.1.2).

### 3.1.1 Radiometer Equation Basics

For a radio transient to be detectable, its signal must significantly exceed noise fluctuations present in the telescope system. As a convention in radio astronomy, the amount of noise present in a system is quantified by a characteristic temperature value known as the system temperature. This is not the physical temperature of the system, but the temperature at which random thermal motions in an ideal resistor would produce the same power per unit frequency as the observed noise. Real observing systems have many independent sources of noise, and the system temperature is a sum of the contributions from each of these sources:

$$T_{\rm sys} = T_{\rm rec} + T_{\rm spill} + T_{\rm atm} + T_{\rm sky} \tag{3.1}$$

$T_{\rm rec}$ is the thermal noise produced in the receiver of the telescope, which encompasses the entire analog system including the reflector, feeds, amplifiers, transport cables, and bandpass filters that act on the signal before digitization. For CHIME/FRB, this value is about $T_{\rm rec} = 20\,{\rm K}$ (see §2.1 for a description of CHIME's analog system). The remaining terms are as follows: $T_{\rm spill}$ is the "spillover" noise from ground radiation entering the feeds (typically around 10 K or less), $T_{\rm atm}$ is the noise from Earth's atmospheric radiation (only important for high-frequency observations above 5 GHz), and $T_{\rm sky}$ is background radiation from astrophysical sources in the sky. $T_{\rm sky}$ is mostly composed of synchrotron radiation from electrons in the plane of the Milky Way. This component varies strongly based on frequency and sky position, ranging from tens to hundreds of Kelvin at high and low Galactic latitudes (Haslam et al., 1982; Remazeilles et al., 2015). Each of these independent sources contributes to random Gaussian noise in the voltages detected by the telescope receiver. The total system temperature is a measure of the power of these contributions, and it can be essentially thought of as the average power level that the telescope detects when pointed at a "blank" patch of sky (containing no strong radio sources).

The power per unit frequency detected from a given astrophysical source of flux density, $S_{\rm src}$, can also be described by a characteristic temperature value, $T_{\rm src}$, such that

$$S_{\rm src} = \frac{T_{\rm src}}{G} \tag{3.2}$$

where $G$ is the system gain, commonly given in units of $\mathrm{K\,Jy^{-1}}$. The gain is a measurement of antenna sensitivity which characterizes the instrumental response to a given increase in source brightness in Janskys. This response is largely determined by the effective collecting area of the telescope, $A_{\rm e}$, but also varies as a function of both frequency and sky location according to the telescope beam response, $B(f, \theta, \phi)$ (e.g., §2.3):

$$G = G_0 B(f, \theta, \phi) = \frac{A_{\rm e}(f)}{2k_{\rm B}} B(f, \theta, \phi) \tag{3.3}$$

where $k_{\rm B}$ is the Boltzmann constant and $G_0$ is the gain at peak beam sensitivity. The effective collecting area of a telescope is smaller than the geometric area of the reflectors by some aperture efficiency fraction: $A_{\rm e} = \eta A$, where $\eta \in [0, 1]$. This fraction encodes inefficiencies in the telescope that reduce the radiative power reaching the receiver, such as uneven aperture illumination and the finite reflectivity of the dish surface. For telescopes involved in radio transient astronomy, this fraction is typically on the order of $\eta \sim 0.6$ to 0.7 (e.g., 0.65 for the Parkes Ultra-Wide Bandwidth receiver, Hobbs et al. 2020; 0.72 for ASKAP, McConnell et al. 2016).

The gain is central to flux calibration as it acts as a bridge between the physical brightness of a source and the instrumental response. Peak gain values cover a wide range across different telescope systems. For example, the Parkes 21-cm Multibeam receiver had a gain of $G_0 \sim 0.6\,\mathrm{K\,Jy^{-1}}$ (Manchester et al., 2001), while the recently commissioned FAST telescope has a gain of $G_0 \sim$

$16\,\mathrm{K\,Jy^{-1}}$ (Jiang et al., 2019). The CHIME gain can be very roughly estimated by assuming that the effective area is just the illuminated area of the reflectors: $A_\mathrm{e} \sim 80\,\mathrm{m} \times 80\,\mathrm{m} = 6400\,\mathrm{m}^2$. If we conservatively assume an aperture efficiency of $\eta \sim 0.5$, this corresponds to a gain of $G_0 \sim 1.15\,\mathrm{K\,Jy^{-1}}$.

The gain can also be used convert the system temperature into a representative flux value known as the system equivalent flux density (SEFD), given by $\mathrm{SEFD} = T_\mathrm{sys}/G$. As the ratio between two fundamental system parameters, the SEFD is another useful characterization of the sensitivity of a given telescope, with lower values indicating more sensitive systems. For example, the SEFD of ASKAP is $\sim$2000 Jy (McConnell et al., 2016), the SEFD of CHIME/FRB is roughly $\sim$50 Jy (confirmed using steady source transits from the flux calibration pipeline), and the SEFD of the Arecibo L-Band Feed Array was $\sim$3.5 Jy[1] .

All of the terms defined above combine together very elegantly in the radiometer equation. A "radiometer" is a device that measures the average power of a noise-like signal in a well-defined frequency range. It is an umbrella term that includes radio telescope receivers under the assumption that the measured signal resembles thermal Gaussian noise. To detect faint radio signals, the amount of noise in the power measured by a radiometer can be beat down by integrating over multiple samples. When a radio telescope is pointed at a blank patch of sky, the

[1] http://www.naic.edu/alfa/performance/

root mean square fluctuation in the measured power is given by

$$\Delta T_{\rm sys} = \frac{T_{\rm sys}}{\sqrt{N}} = \frac{T_{\rm sys}}{\sqrt{n_{\rm p} t_{\rm s} \Delta f}} \tag{3.4}$$

where $N = n_{\rm p} t_{\rm s} \Delta f$ is the number of samples integrated over a $\Delta f$ frequency bandwidth in $t_{\rm s}$ sampling time with $n_{\rm p}$ summed polarizations.[2] Substituting $\Delta T_{\rm sys} = G \Delta S_{\rm sys}$, we arrive at

$$\Delta S_{\rm sys} = \frac{T_{\rm sys}}{G\sqrt{n_{\rm p} t_{\rm s} \Delta f}} = \frac{\rm SEFD}{\sqrt{n_{\rm p} t_{\rm s} \Delta f}} \tag{3.5}$$

This is the radiometer equation. The usefulness of this equation for the flux calibration becomes more apparent when it is re-framed in terms of the signal to noise ratio of the detected power ($S_{\rm det}$). By definition:

$$S/N = \frac{S_{\rm det}}{\sigma_{S_{\rm sys}}} \tag{3.6}$$

where $\sigma_{S_{\rm sys}}$ is the standard deviation of the radiometer fluctuations. If we subtract the off-pulse mean from the time series of detected power, then $\Delta S_{\rm sys} = \sigma_{S_{\rm sys}}$. In this case, we can rearrange Equation 3.6 and plug in the expression for the rms variation from radiometer equation ($\Delta S_{\rm sys} \rightarrow \sigma_{S_{\rm sys}}$), arriving at:

$$S_{\rm det} = \frac{(S/N)\, T_{\rm sys}}{G\sqrt{n_{\rm p} t_{\rm s} \Delta f}} = \frac{(S/N)\, \rm SEFD}{\sqrt{n_{\rm p} t_{\rm s} \Delta f}} \tag{3.7}$$

[2] This derivation based on $N{=}n_{\rm p} t_{\rm s} \Delta f$ independent data points is technically a simplification. For the full derivation involving Nyquist sampling and calculations of square-law detector noise, see Section 3.6.2 of Condon & Ransom (2016).

A given FRB profile in instrumental units can therefore be calibrated into Jy units by simply subtracting from the time series its off-pulse mean, dividing by the off-pulse standard deviation to derive a time series of signal to noise ratios, then multiplying by the right-hand side of Equation 3.5. In this case, $t_{\mathrm{s}}$ is the sampling time of the telescope and $\Delta f$ is the bandwidth of each frequency channel. The power of this calibration method is in its simplicity, since it requires knowledge of just a few system parameters: either $T_{\mathrm{sys}}$ and $G$ or just the SEFD.

### 3.1.2 FRB Flux Calibration

After decades of verification in the field of pulsar astronomy, the radiometer equation is considered a very robust method for the flux calibration of radio transients. As a result, nearly all FRB experiments and surveys in operation today use some variation of the radiometer equation to calibrate their data (e.g., Arecibo Pulsar ALFA Survey, Spitler et al. 2014; Parkes High Time Resolution Universe Survey, Champion et al. 2016; Green Bank Northern Celestial Cap Pulsar Survey, Parent et al. 2020; Commensal Radio Astronomy FAST Survey, Zhu et al. 2020; UTMOST, Farah et al. 2019). Its use in the field is so routine that the flux calibration descriptions in many FRB papers consist of just a short paragraph or table listing the relevant system parameters.

Key to the use of the radiometer equation is the accurate determination of parameters like the system temperature, gain, or SEFD. These values are typically determined using observations of sources of known intensity, such as a bright astrophysical source of radio emission (e.g., a radio galaxy or supernova remnant), a previously calibrated noise diode, or a pair of hot and cold "loads" of calibrated temperature (e.g., a radio-absorbing material in an oven or liquid nitrogen bath). For example, FAST alternately injected a $T \sim 11\,\mathrm{K}$ noise diode into an observation of a blank patch of sky and used the difference in detected power to calibrate the system temperature $T_{\mathrm{sys}} \sim 23\,\mathrm{K}$ (Jiang et al., 2020). Alternatively, ASKAP does not have a noise diode, and instead determines the SEFD using observations of a Seyfert galaxy of well-known flux, PKS B1934-638 (McConnell et al., 2016). The new Parkes Ultra-Wide Bandwidth Receiver, on the other hand, was calibrated by comparing the spectrum observed with an ambient-temperature absorber covering the feed (the "hot load") versus the spectrum observed in a blank patch of sky (the "cold load"; Hobbs et al. 2020). For a review of each of these methods, see O'Neil (2002).

Once these values are measured, they can change over time due to variations in the telescope structure, receiver electronics, environmental conditions, or calibrator source intensity. FRB experiments make different assumptions about the temporal stability of the measured system parameter values based on their system design. For example, the gain values quoted in current FRB papers for single-dish telescopes like the Parkes Multibeam receiver and the GBT were measured years ago, in 1996 (notably the birth year of the writer of this thesis; Staveley-Smith et al., 1996) and 2009[3] , respectively. These systems have cryogenically cooled receivers, which tend to be stable over long periods of time. In contrast, newer

[3] https://www.gb.nrao.edu/~rmaddale/GBT/ReceiverPerformance/PlaningObservations.htm

interferometric experiments like ASKAP and UTMOST rely on calibrations from within several months to within several hours of an observed FRB (e.g., for ASKAP, fly's eye mode observations are calibrated within months to days: Bannister et al. 2017, Shannon et al. 2018, whereas interferometric observations are calibrated within hours: Bannister et al. 2019; for UTMOST, the SEFD is calibrated daily: Venkatraman Krishnan et al. 2020). This is partially an outcome of the additional phase calibration requirement for interferometric observing, but also due to the lack of cooling on the receivers of these instruments, which causes more variability.

Beyond the determination of system parameters, the most central obstacle to accurate FRB flux calibration is burst localization. Unlike pulsars, most FRBs observationally appear to burst only once (it is still up for debate whether these FRBs are intrinsically one-off events or if there is simply a very long wait time between repeat bursts). This presents a particularly difficult challenge, as FRB-detecting telescopes must often rely on just a single burst for localization purposes. As described in §1.3.2, most telescopes in FRB detection history have been limited in their ability to spatially sample their field of view, leading to large uncertainties in detected burst positions within the telescope beam pattern. Without precise knowledge the location of the FRB in the beam, measured fluences cannot be properly corrected for attenuation due to the beam response ($B(f, \theta, \phi)$). As a result, the vast majority of fluences published to-date are, in fact, lower limits, calculated under the simplifying assumption that the FRBs were detected at beam boresight (using $G_0$).

For many of the existing single-dish telescopes used in nascent FRB surveys, localizations are basically unconstrained within the beam pattern. For example, at the GBT, FRB positions are typically reported as the centre of the telescope pointing, with an uncertainty corresponding to the FWHM of the beam (e.g., Parent et al., 2020; Scholz et al., 2016). For the Parkes Multibeam receiver, which tiles the Parkes field of view with 13 independent beams, most localizations are similarly taken to be the FWHM of the highest SNR beam in which the burst was detected (e.g., Thornton et al., 2013; Champion et al., 2016). The fluences measured from these observations could be biased low by a factor of two or more.

In principle, this calibration challenge posed by localization can be solved using interferometeric telescopes with longer baselines, like the VLA, EVN, and DSA-10, which have the ability to localize on the sub-arcsecond level. These telescopes can beamform in the precise direction of the FRB and use nearby phase calibration sources to obtain accurate flux and fluence measurements. However, by design, these telescopes typically have a small field of view, which significantly decreases the chances of a one-off FRB detection. For example, project `realfast`[4] has been completing both commensal and targeted FRB searches at the VLA (field of view $\sim 30'$ at L-band) since late 2015 (e.g., Law et al., 2015, 2019) and has detected only a single one-off FRB in hundreds of hours of observation (Law et al., 2020). Despite this limitation, long-baseline interferometers have proven absolutely crucial for following up and precisely localizing repeating FRBs that

[4] `http://realfast.io/`

are first broadly localized at other observatories (as described in §1.3.2). Once these precise localizations are determined, other telescopes like the GBT and Arecibo can then obtain accurate fluences during follow-up observations centred on the repeating burst positions (e.g., Gourdji et al., 2019; Chawla et al., 2020). In this way, a symbiosis has formed between collaborating telescopes of different localization capabilities.

As the FRB field has matured, new experiments have developed that balance the design tradeoff between a large field of view and localization capability. The Commensal Real-Time ASKAP Fast Transients (CRAFT) survey is one such experiment that is particularly well-poised for the accurate determination of one-off burst fluences. Each 12 m antenna in the ASKAP array is equipped with a phased array feed (PAF) which tiles 36 synthesized beams across a 30 square degree field of view (Bannister et al., 2017). The PAF beams are separated from each other by just $0.9^{\circ}$, forming a closely packed square grid. With baselines up to 6 km, ASKAP is capable of obtaining sub-arcsecond localizations when antenna signals are combined interferometrically. This localization capability, in addition to the relatively large field of view, has allowed ASKAP to obtain sub-arcsecond positions and accurate fluences for eight one-off FRBs (Bannister et al., 2019; Prochaska et al., 2019; Bhandari et al., 2020b; Heintz et al., 2020; Macquart et al., 2020). The array can also operate in a fly's eye mode, with 5 to 12 antennae pointed in different directions to maximize sky coverage, providing an instantaneous field of view of up to 360 square degrees. Even in this observing mode, ASKAP produces accurate fluence measurements using a method similar to

CHIME/FRB header localization (§2.5.2), which compares the pattern of measured fluxes in adjacent beams to predictions from a beam model (see Section 4.1 of Bannister et al. 2017 for a detailed description). Leveraging the dense sampling of the focal plane and a relatively simple Gaussian model for the PAF beams, this method results in localizations within $10'$ (90% confidence) and fluence measurements within 20% accuracy (Bannister et al., 2017). To-date, ASKAP has detected a total of 23 FRBs in this observing mode, which constitutes the majority of one-off FRBs with accurate fluence measurements (Bannister et al., 2017; Shannon et al., 2018; Macquart et al., 2019).

## 3.2 CHIME/FRB Calibration Challenges

As described in the previous section, radio transient flux calibration ultimately involves quantifying properties related to the fundamental performance of a given observing system. The requirements for accurate FRB flux calibration, in particular, distill into three main components: determination of the time-variable system sensitivity, precise FRB localization, and characterization of the telescope beam response as a function of frequency and on-sky location.

For the first bursts detected from CHIME/FRB, this flux calibration process became an especially elaborate challenge requiring the characterization of a novel and still-developing system within an accelerated timeframe. By July 20th, 2018, less than a year after first light for the CHIME telescope on September 7th, 2017, a limited version of the CHIME/FRB beamforming and detection pipeline was up and running semi-stably. On July 25th, less than a week later, the first FRB was detected (CHIME/FRB Collaboration et al., 2019a). At this point, the system was still in a pre-commissioning period, where only a small and variable number of beams were being searched at any given time and the complex gain calibration strategy was under development and constantly changing. Despite the instability of the system during this time, 14 FRBs were detected before pre-commissioning ended and the full 1,024 beam system was implemented on August 27th, 2018. Since then, CHIME/FRB has steadily detected FRBs, and 535 FRBs were accumulated within the first year of operation (CHIME/FRB Collaboration et al., 2021).

These frequent detections early in the commissioning of the experiment, and ensuing scientific results, necessitated the development of the CHIME/FRB flux calibration pipeline before many aspects of the system were fully operational, let alone thoroughly quantified or understood. Here I highlight the state of our knowledge of the system during the time period encompassing the first catalog (July 25th, 2018 to July 1st, 2019) in order to give context for the design and limitations of the automated pipeline presented in the next section:

- **System Parameters:** Robust measurements of key system parameters for CHIME/FRB had not yet been completed at the start of flux calibration pipeline development, although a few nominal values existed. Based on design specifications, the receiver temperature for the full analog chain of CHIME was expected to be approximately 50 K (including spillover noise from the ground; Bandura et al., 2014). The gain at peak sensitivity could also be very roughly determined by estimating the illuminated area of the reflectors, yielding a value of $G_0 \sim 1.15\,\mathrm{K\,Jy^{-1}}$ (see §3.1.1 for this derivation).

Later, in May and June of 2019, CHIME/Cosmology used a pair of hot and cold loads to measure the receiver ($\sim$20 K) and system ($\sim$50$-$60 K) temperatures from two feeds at different locations along the CHIME focal line, confirming the design specifications.[5] These measurements also constrained the effective area of a single feed to be $\sim$3.5 m$^2$ (band-averaged), which, multiplied by 1,024 feeds, roughly yields a total effective area of 3,600 m$^2$, or a gain of $G_0 \sim 1.3\,\mathrm{K\,Jy^{-1}}$. After the development of the flux calibration pipeline, I used steady source transit observations to sample the CHIME/FRB beamformed SEFD at a number of zenith angles along the meridian, with values near zenith reaching $\sim$50 Jy. However, to this day, a detailed accounting of the system parameters (considering any beamforming inefficiencies, quantization biases, feed coupling, and variations as a function of location on-sky) remains a work in progress.

- **Beam Model:** As described in §2.3.3, CHIME/FRB's synthesized beam pattern is purely digital and exactly deterministic. Since modeling it simply requires re-calculating the beamforming process, an accurate model of the FFT beamformed beams was available early in the commissioning of the flux calibration pipeline, by July 2018. Modeling the primary beam, however, is much more difficult, as it involves accounting for complex physical reflections and cross-talk within the CHIME cylinders. As a result, for the first year

[5] Details can be found in documents #0836, #0865, and #1216 in the internal CHIME doclib.

of CHIME/FRB's run, the primary beam model was just a smooth cosine in the North-South direction and a Gaussian in the East-West direction. This early model made no attempt to characterize the 30 MHz ripple, an oscillating sensitivity pattern across the CHIME bandwidth on the order of $\sim$40% in amplitude that varies in shape depending on North-South zenith angle (see §2.3.2). Over the course of the first year of operation, CHIME/Cosmology was able to better characterize the 30 MHz ripple using a combination of analytical models and steady source transits, resulting in a new beam model accurate to within $\sim$10% in the main lobe of the primary beam. This data-driven primary beam model was incorporated into the CHIME/FRB codebase in late 2019, after the first-pass fluence pipeline was already developed.

- **Localization:** The only thoroughly tested and automated form of localization developed for CHIME/FRB is header localization, which works by fitting detected SNRs from each beam to predictions informed by the frequency-dependent beam model (intensity and baseband localization are still in development, see §2.5.2). Owing to CHIME/FRB's significant and highly chromatic formed-beam sidelobes, header localization regions generally consist of three "islands" representing the main lobe and sidelobes of the highest SNR formed beam in which the burst was detected. Multi-beam detections result in better localizations depending on the pattern of detected beams, but generally the degeneracy between the main lobes and sidelobes persists, especially at higher confidence levels (see Figure 2–7). Given this degeneracy, the spectral structure of CHIME/FRB's beam pattern and overall beam response can change significantly over the extent of the header localization region obtained for each burst, making it difficult to reliably correct fluence measurements for beam attenuation. I have created two figures to illustrate this point. Figure 3–1 shows the beam response at several locations within the 68% confidence localization region for a single-beam detection, demonstrating how the clamping behaviour and overall beam response can change rapidly both within the main lobe localization and in the sidelobes. Figure 3–2 shows a histogram of 10,000 band-averaged beam sensitivities sampled according to the probability density of the header localization region shown in Figure 3–1. The overall beam response varies by roughly a factor of ∼10 between the main lobe and the sidelobes, and the sidelobes dominate the sampled responses.

## 3.3 CHIME/FRB Calibration Pipeline

The motivating question for the design of the flux calibration pipeline is the following: given the beam modeling and localization limitations of CHIME/FRB outlined in the previous section, how do we leverage our existing resources to obtain meaningful constraints on the flux and fluence of each detected FRB? Answering this question prompted two salient design decisions:

1) Since we did not have an accurate model of the primary beam when the calibration pipeline was first developed, we characterize and correct for the 30 MHz ripple empirically using daily transit observations of steady sources with known spectral properties. By comparing the known flux of a source to

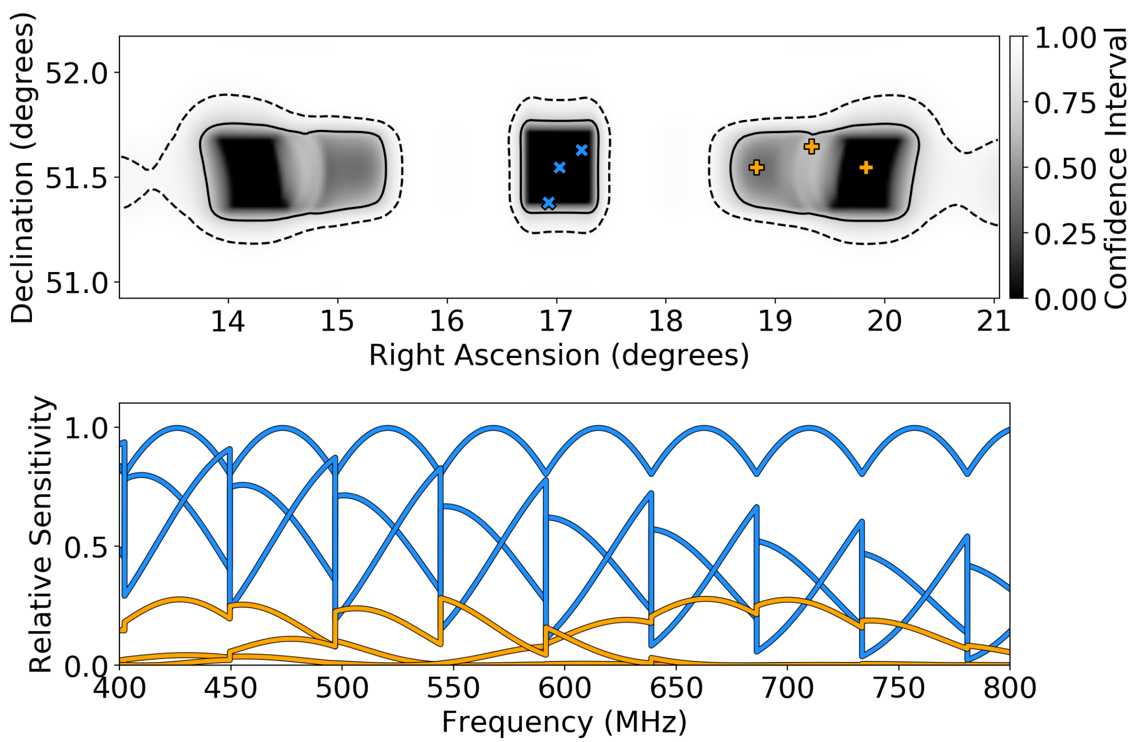


Figure 3–1: (Top) An example header localization confidence interval for a single beam detection. The confidence interval colouring is explained in the caption of Figure 2–7. The 68% and 95% intervals are labelled by the solid and dashed contours, respectively. The blue crosses represent sampling locations in the main lobe of the localization, while the orange plus signs represent sampling locations in the sidelobes. (Bottom) The CHIME/FRB beam response as a function of frequency at the samplings marked in the top figure. Note that these beam responses were generated using a smooth approximation of the primary beam model (see §2.3.2). Blue lines correspond to the main lobe samplings, orange lines correspond to the sidelobe samplings.

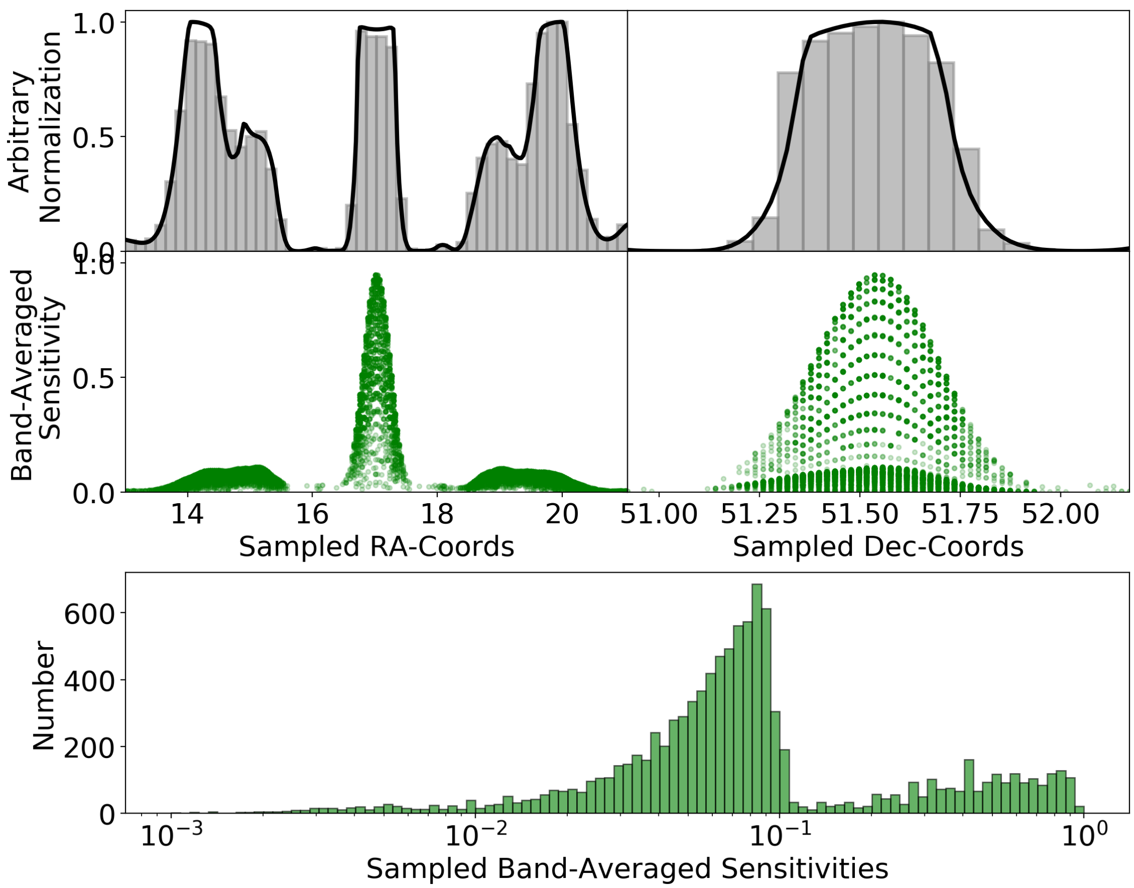


Figure 3–2: A visualization of 10,000 band-averaged CHIME/FRB beam sensitivities sampled from the header localization region shown in Figure 3–1. (Top) The histograms show the normalized number of sampled points as a function of right ascension (Left) and declination (Right). The black lines show normalized cross-sections in declination (Left) and right ascension (Right) of the 2D header localization probability density shown in Figure 3–1. (Middle) The sampled band-averaged sensitivities as a function of right ascension (Left) and declination (Right). (Bottom) A histogram of sampled band-averaged sensitivities.

the observed total-intensity units output by the beamformer, we solve for the beamformer to Jansky conversion across the primary beam directly rather than relying on measurements of the system temperature, gain, or SEFD in combination with an approximate beam model. Determining this conversion for each frequency channel creates a "calibration spectrum" that encodes the 30 MHz ripple in the direction of a given calibration source. This spectrum is then applied to the total-intensity data of FRBs nearby in zenith angle to derive a burst dynamic spectrum in physical units roughly corrected for attenuation from the North-South pattern of the primary beam.

2) Due to the limitations of header localization, we follow suit from other early FRB surveys and calculate our fluences assuming that each burst was detected at beam boresight. For our purposes we take "boresight" to mean along the meridian of the primary beam (at the peak sensitivity of the burst declination arc). We do not correct our fluence measurements for a burst's unknown location in the synthesized beam pattern. Thus, our fluence measurements are biased low, as bursts off-meridian will experience beam attenuation from both the primary and synthesized beam pattern that we are not correcting for. The measurements produced from the pipeline are therefore most appropriately interpreted as lower limits, with an uncertainty on the limiting value.

In this section I present the implementation details of the first-pass CHIME/FRB flux calibration pipeline that I have developed for this thesis. The steps in the pipeline are outlined as flowcharts in Figures 3–3 and 3–10. First, I briefly describe the infrastructure of the pipeline (§3.3.1). Then, I describe the four broad sections of the pipeline in sequence: the automated calibration source transit observations (§3.3.2), the beamformer to Jansky (BF/Jy) conversion spectrum calculation (§3.3.3), calibrating the FRB intensity data (§3.3.4), and calculating the flux and fluence and corresponding errors (§3.3.5). Finally, in section 3.3.6, I describe an initial test of the pipeline conducted on simulated bursts injected into CHIME/FRB intensity data.

### 3.3.1 Pipeline Infrastructure

Since CHIME/FRB detects a large volume of bursts, the flux calibration pipeline is automated as much as possible to save human work hours. The BF/Jy calculation, intensity calibration, and flux calculation stages are configured to run in jobs distributed on a separate 10-node on-site cluster known as the analysis cluster. These nodes are located in the same seacan as the detection pipeline cluster, but are reserved for post-detection analyses. Using a platform called Docker,[6] the flux calibration pipeline code is organized into a software package called a container, which includes all the infrastructure and dependencies needed to run the code in any environment. This container can be launched onto the cluster to run any stage of the flux calibration pipeline. The cluster is managed by a custom built load balancing service using Docker Swarm.[7] Each calibration pipeline job is allocated a single core and 16 GB of RAM. This setup allows

[6] `https://www.docker.com/`

[7] `https://docs.docker.com/get-started/orchestration/`

multiple jobs to be run on the cluster in parallel, which is conducive for analyzing large batches of bursts.

Pipeline data products — such as the downsampled calibration source intensity data, beamformer to Jansky conversion spectra, and calibrated FRB dynamic spectra — are stored on the on-site CHIME/FRB archiver. Output fluence measurements are stored in the events database while metadata for calibration data products are stored in a separate calibration database. These databases can be queried from anywhere through authenticated HTTP requests using a custom Python API called `frb-master` (which is accessed through a wrapper called `frb-api` that handles authentications). I have developed the part of this infrastructure that handles metadata queries related to calibration spectra using the RethinkDB query language.[8] The querying is designed to be flexible, allowing the user to obtain the available calibration spectra on a given date, in a date range, from a given calibration source, or nearest to a spatial location on-sky. This functionality is used extensively in the intensity calibration and fluence calculation stages of the pipeline.

### 3.3.2 Steady Source Acquisition Stage

The first part of the fluence pipeline (the "Steady Source Acquisition Stage" in Figure 3–3) deals with scheduling, processing, and storing daily total-intensity transit observations of steady calibration sources.

[8] `https://rethinkdb.com/`

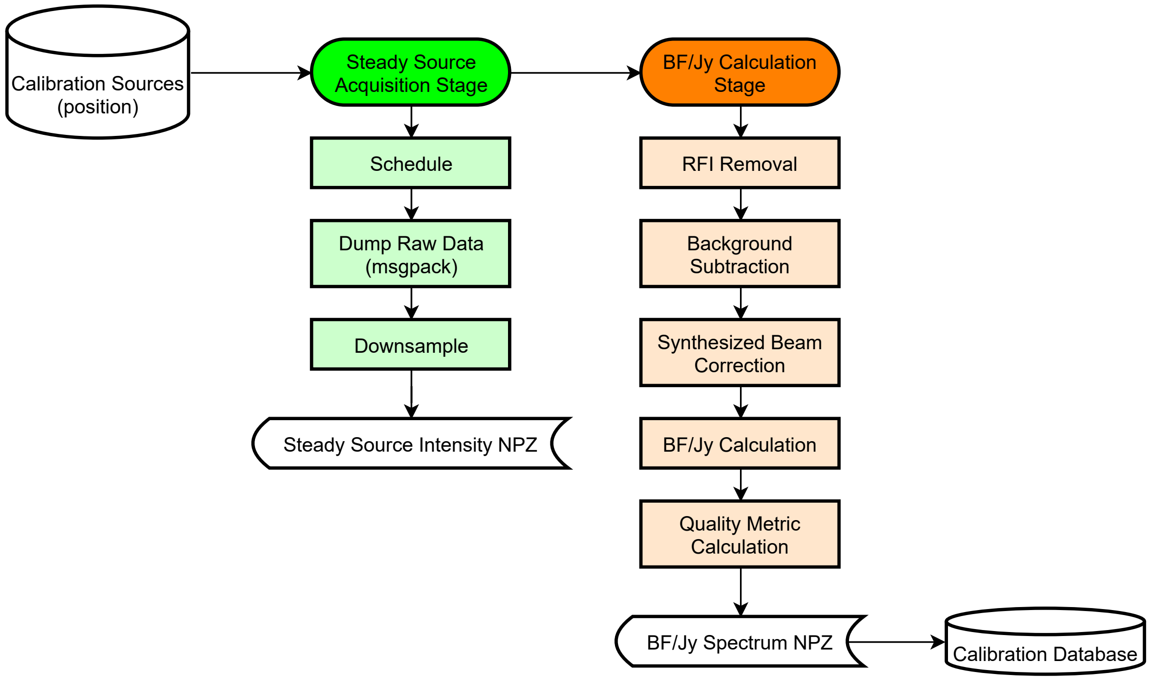


Figure 3–3: A flowchart of the first half of the CHIME/FRB flux calibration pipeline, which encompasses steady source acquisitions and determination of beamformer unit to Jansky conversion spectra.

Calibration sources were selected from two different catalogs, Perley & Butler (2016) and Vollmer et al. (2010), which provide measurements of the source flux across the CHIME band. Perley & Butler (2016) presents fluxes from the VLA measured in 2014 and 2016 at frequencies ranging from 220 MHz to 48.1 GHz. Combining these measurements with 73.8 MHz legacy observations from 1998, Perley & Butler (2016) model the frequency-dependence of the flux density of each calibrator by fitting a polynomial to the log transformed measurements. The resulting uncertainty on the flux of these sources in the CHIME band is $\sim$2$-$4%. Vollmer et al. (2010) presents a compilation of flux densities at frequencies ranging from 159 MHz to 8.4 GHz, derived from cross-identifying sources in different radio catalogs. Since these measurements come from a wide range of telescopes observing at different epochs, the uncertainties on the flux densities in the CHIME band are larger than those from Perley & Butler (2016), on the order of ∼5-15%.

Sources were selected from these catalogs according to the following three criteria: (1) the source is within the CHIME field of view (above $-11^\circ$ declination), (2) the flux density of the source at 600 MHz is greater than 10 Jy (below which confusion noise becomes significant), and (3) the catalogs provide flux density measurements for at least three different radio frequencies spanning the CHIME band. A total of 35 sources matched these criteria (14 from Perley & Butler 2016 and 21 from Vollmer et al. 2010). We use the same methodology as Perley & Butler (2016) to interpolate the spectrum of each source and determine the flux in the CHIME band (see section 4 of Perley & Butler 2016). The 35 selected sources are listed in Table 3–1, and plots of their observed average transit time series and spectra are shown in Figure 3–4.

Note that there are a few sources in our calibrator sample, such as TauA, CasA, and 3C138, that are known to be variable on the order of a few percent per year. For example, TauA's flux density is declining by approximately 0.25% every year and 3C138 is declining on the order of $\sim 1-3\%$ per year (Baars et al., 1977; Perley & Butler, 2016). This decline is much smaller than the errors on the flux expected from beam attenuation and the time-variable system sensitivity, and so we keep these sources in our sample to achieve more coverage of the primary beam (see the discussion of error calculation in §3.3.5).

| Source | Name | Type | Nominal RA (J2000) [hh:mm:ss.s] | Nominal Dec (J2000) [dd:mm:ss.s] |
|---|---|---|---|---|
| 1 | 3C_61.1 | Seyfert 2 Galaxy | 02:22:46.992 | +86:18:47.88 |
| 2 | 3C_27 | Seyfert 2 Galaxy | 00:55:58.944 | +68:22:22.44 |
| 3 | SNR_G130.7+03.1 | Supernova Remnant | 02:05:35.844 | +64:49:40.08 |
| 4 | NVSS_J235054+644018 | Radio Galaxy | 23:50:53.4 | +64:40:19.56 |
| 5 | SN_1572A | Supernova Remnant | 00:25:23.94 | +64:08:26.88 |
| 6 | 3C_430 | Quasar | 21:18:18.504 | +60:48:12.42 |
| 7 | CAS_A | Supernova Remnant | 23:23:27.9408 | +58:48:42.408 |
| 8 | 4C_55.06 | Radio Galaxy | 03:27:19.248 | +55:20:31.92 |
| 9 | 3C_295 | Seyfert 2 Galaxy | 14:11:20.52 | +52:12:9.972 |
| 10 | 3C_147 | Seyfert 1 Galaxy | 05:42:36.1368 | +49:51:07.236 |
| 11 | 3C_380 | Seyfert 1 Galaxy | 18:29:31.7808 | +48:44:46.176 |
| 12 | 3C_196 | Seyfert 1 Galaxy | 08:13:36.0336 | +48:13:2.568 |
| 13 | 3C_129 | Radio Galaxy | 04:49:01.752 | +45:01:35.76 |
| 14 | GB6_B2040+4242 | Radio Source | 20:42:34.992 | +42:57:03.959 |
| 15 | NGC_1265 | Radio Galaxy | 03:18:15.888 | +41:53:44.88 |
| 16 | 2C_396 | Quasar | 04:32:36.408 | +41:38:28.68 |
| 17 | CYG_A | Seyfert 2 Galaxy | 19:59:28.356 | +40:44:02.112 |
| 18 | 3C_452 | Seyfert 2 Galaxy | 22:45:48.312 | +39:41:08.88 |
| 19 | 3C_134 | Radio Galaxy | 05:04:42.096 | +38:06:02.88 |
| 20 | 2MASX_J21555232+3800285 | Seyfert Galaxy | 21:55:52.296 | +38:00:33.84 |
| 21 | 3C_48 | Seyfert Galaxy | 01:37:41.2992 | +33:09:35.136 |
| 22 | 3C_286 | Seyfert 1 Galaxy | 13:31:08.2872 | +30:30:32.94 |
| 23 | 3C_123 | Active Galactic Nucleus | 04:37:04.3752 | +29:40:13.8 |
| 24 | NGC_7720 | Seyfert 1 Galaxy | 23:38:27.948 | +27:01:15.96 |
| 25 | 3C_133 | Active Galactic Nucleus | 05:02:58.704 | +25:16:22.8 |
| 26 | 3C_409 | Radio Source | 20:14:27.408 | +23:34:46.92 |
| 27 | TAU_A | Supernova Remnant | 05:34:31.9704 | +22:00:52.056 |
| 28 | 3C_138 | Seyfert 1 Galaxy | 05:21:09.8856 | +16:38:22.056 |
| 29 | 3C_454.3 | Quasar | 22:53:57.72 | +16:08:53.52 |
| 30 | PKS_J1923+1429 | Radio Source | 19:23:41.904 | +14:30:33.12 |
| 31 | VIR_A | Brightest Cluster Galaxy | 12:30:49.42 | +12:23:28.04 |
| 32 | 2MASX_J03585442+1026033 | Seyfert 2 Galaxy | 03:58:54.696 | +10:26:2.04 |
| 33 | HERCULES_A | Brightest Cluster Galaxy | 16:51:08.148 | +04:59:33.324 |
| 34 | 3C_353 | Active Galactic Nucleus | 17:20:28.1592 | −00:58:46.632 |
| 35 | 3C_161 | Quasar | 06:27:10.0960 | −05:53:04.768 |

Table 3–1: Information about the CHIME/FRB flux calibration sources.

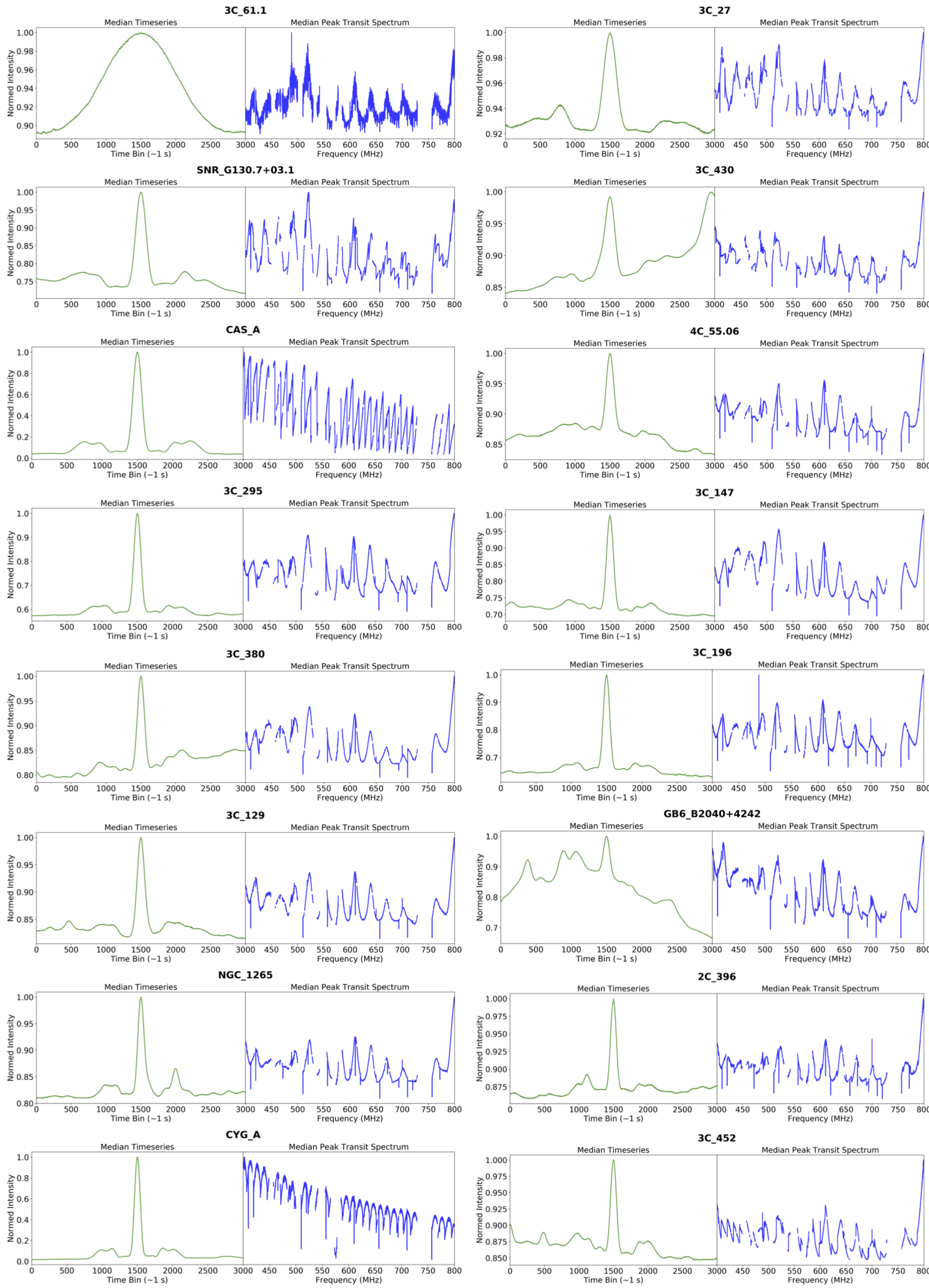


Figure 3–4: CHIME/FRB data showing the average normalized transit time series and spectra for each of the flux calibration sources listed in Table 3–1.

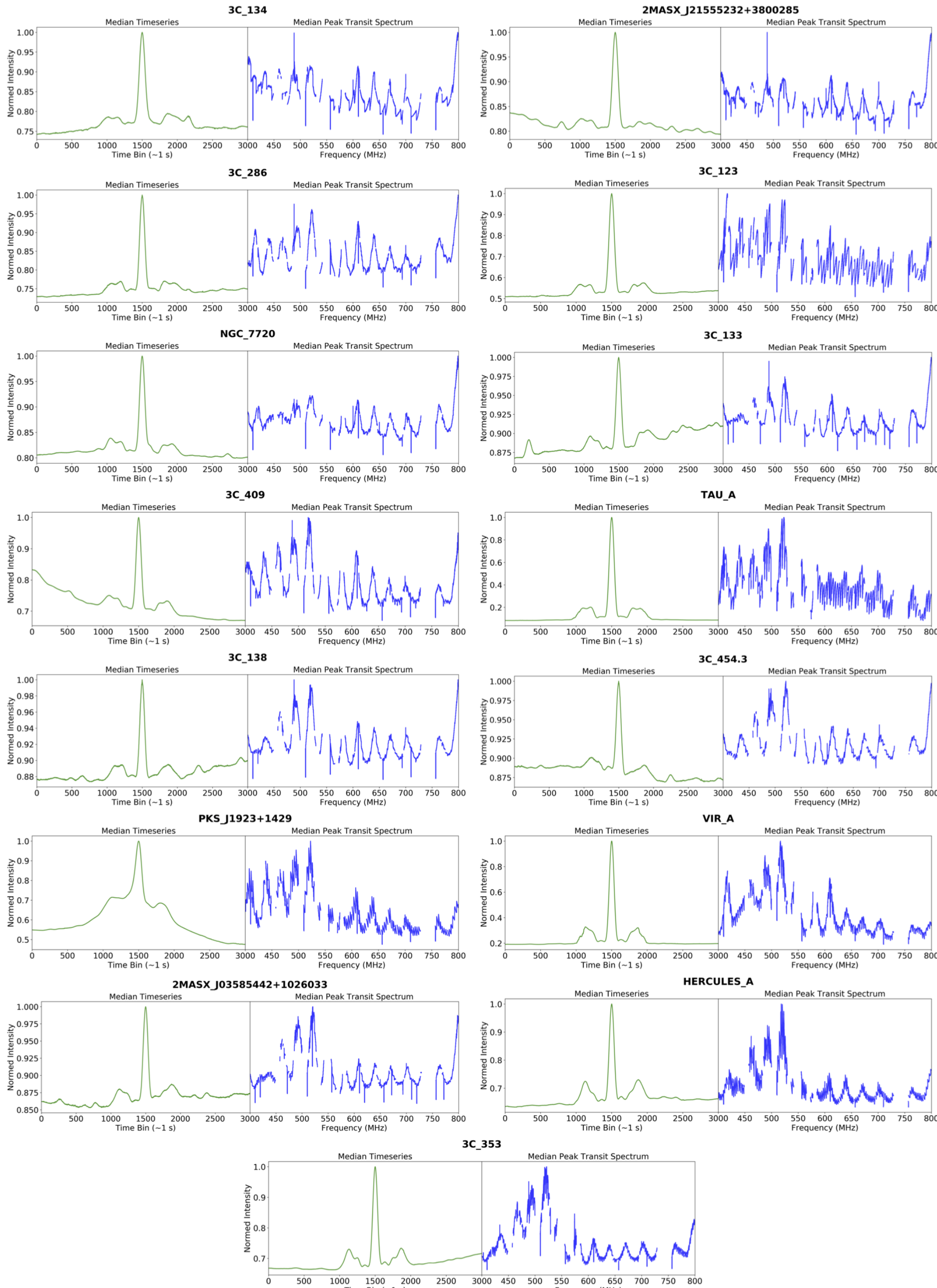


Figure 3–4: Continued, CHIME/FRB data showing the average normalized transit time series and spectra for each of the flux calibration sources listed in Table 3–1.

The steady source acquisition stage runs in a simple python script on the compute node that hosts L4. Taking the calibrator locations and observation duration as input, this script uses `pyephem`[9] in combination with the FFT formed beam model to predict when a given source will transit and which beam in the 1000 column it will cross. When a source is close to transiting, the script creates a folder on the CHIME/FRB archiver, triggers an intensity dump in the transit beam, waits the observation duration, and then stops the dump. Although it only takes a source $\sim$5$-$14 minutes to pass through the main lobe of the primary beam, each calibrator is observed for one hour centred around transit in order to obtain a measurement of the adjacent sky background. Note that, for circumpolar calibrators, we only take observations of upper transits.

The raw intensity data from these observations are output in msgpack[10] format with 16,384 frequency channels (24 kHz resolution) and 0.983 ms time resolution, where each msgpack stores one second of data. A full hour-long observation at this resolution takes up 60 GB of memory on the archiver. To save space, the raw data from each observation are unpacked, downsampled in time to 1-s resolution (taking the median over 1,024 1-ms samples for each new 1-s sample), and organized into a 2D array. The final result is a 200 MB NPZ file[11]

---

[9] https://rhodesmill.org/pyephem/

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

[11] https://imageio.readthedocs.io/en/stable/format_npz.html

representing the dynamic spectrum of the observation. After the NPZ has been saved, the raw data are deleted.

The steady source acquisition stage came online on November 16th, 2018. Ideally this part of the pipeline runs continuously, saving 35 observations every day that CHIME/FRB is operational. In reality there are gaps in coverage due to delays in restarting the pipeline after power outages and system upgrades, as shown in Figure 3–5. Notably, from December 3rd, 2018 to February 8th, 2019, steady source acquisitions were paused to debug storage issues on the archiver, although FRB data were still being saved during this time. As of August 2020, we had taken a total of 7063 calibrator observations, totaling 1.7 TB of data.

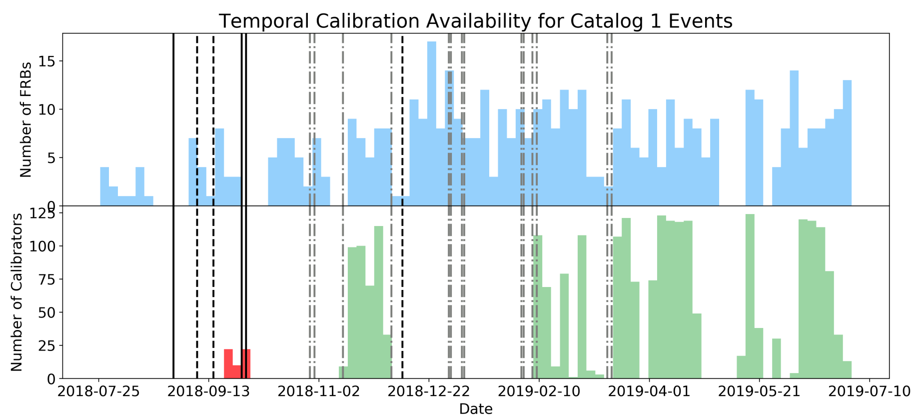


Figure 3–5: Histograms in time of the number of FRBs detected (Top) and the number of calibrator observations obtained (Bottom). Vertical lines denote dates where changes were made to the system: solid lines indicate changes in the beam configuration, dashed lines indicate changes in how the data are scaled at L0 (see §3.3.4), and dashed-dotted lines indicate changes in the complex gain calibrator. Note the period of time from the beginning of the Catalog to February 2019, where calibrator observations are sparse.

### 3.3.3 BF/Jy Calculation Stage

The second stage of the fluence pipeline (the "BF/Jy Calculation Stage" in Figure 3–3) deals with extracting a beamformer unit to Jansky (BF/Jy) conversion spectrum from each steady source observation, as well as additional metrics.

The first step in this process involves removing radio frequency interference (RFI) from the steady source spectra. Although CHIME is mostly sheltered from RFI due to its location in a government-mandated radio quiet zone, CHIME observations are still plagued by RFI mainly from the LTE band ($\sim 700-800\,\mathrm{MHz}$) and transiting airplanes. Some frequency channels are already masked during the upstream gain calibration process (both due to flagged persistent RFI and failures in the calculation of gain calibrations for certain frequency channels), but this is a rough first pass that leaves behind significant residual RFI (see the left panel of Figure 3–6). The data run through three RFI filters in the flux calibration pipeline. One is a static mask that removes channels that were empirically determined to be consistently bad. The second stage detects any additional bad frequency channels using a sliding window algorithm that calculates median values of various statistics over a given frequency channel and time window (median absolute deviation and kurtosis), and then flags outlier channels based on empirical thresholds. The final RFI removal stage searches for bad time bins by taking the gradient of the observation time series and flagging significant spikes. Altogether this RFI method is rather aggressive. It produces clean data but removes a significant fraction of CHIME's bandwidth, typically leaving $\sim 260\,\mathrm{MHz}$ of usable bandwidth (see the right panel of Figure 3–6).

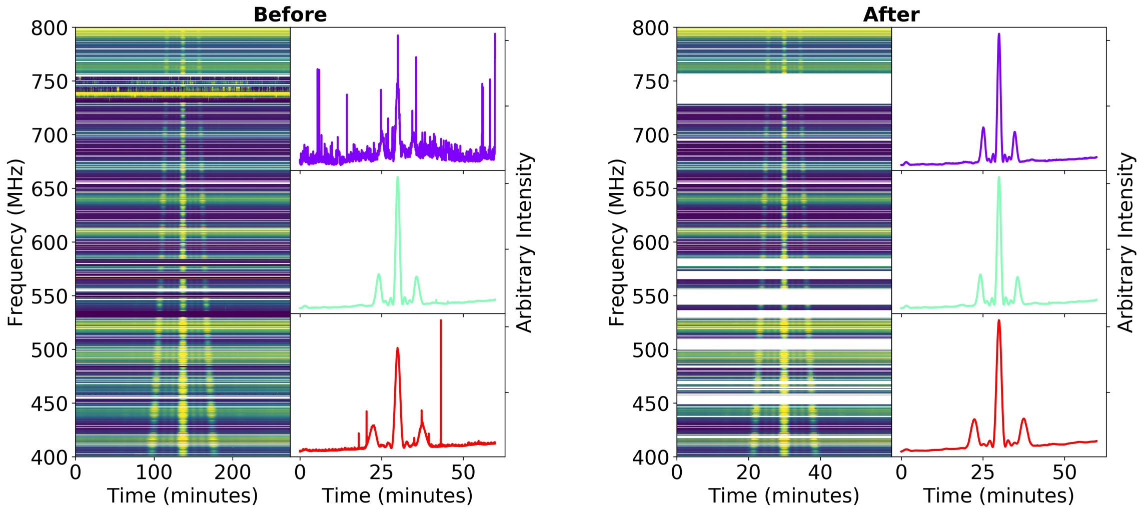


Figure 3–6: A plot showing the flux calibration pipeline RFI removal process for a February 10th, 2019 observation of 3C123. The left panel shows the dynamic spectrum of the observation before RFI removal (Left) and the time series averaged over three equal subbands (Right, with subbands corresponding to: Top, 666−800 MHz; Middle, 533−666 MHz; Bottom, 400−533 MHz). The right panel shows the same plots, but after RFI removal. The white parts of the dynamic spectrum indicate frequency channels that have been zapped due to RFI.

After RFI removal, the BF/Jy spectrum is calculated. This multistep process, outlined in Figure 3–7, can ultimately be broken down into the single equation:

$$C_{\nu,\mathrm{BF/Jy}} = \frac{(S_{\nu,\mathrm{cal,on}} - S_{\nu,\mathrm{cal,off}})\,[\mathrm{BF}]}{(B_{\nu,\mathrm{FFT}}(\theta,\phi) \times S_{\nu,\mathrm{cal,known}})\,[\mathrm{Jy}]} \tag{3.8}$$

where $C_{\nu,\mathrm{BF/Jy}}$ is the resulting calibration spectrum, $S_{\nu,\mathrm{cal,on}}$ is the intensity spectrum detected in beamformer units at the peak of the calibrator transit, $S_{\nu,\mathrm{cal,off}}$ is the background intensity spectrum detected in beamformer units before or after the calibrator has transited, $B_{\nu,\mathrm{FFT}}(\theta,\phi)$ is the synthesized beam sensitivity at the location that the calibrator reaches during peak transit (derived from the formed beam model), and $S_{\nu,\mathrm{cal,known}}$ is the known flux spectrum of the source modeled as described in §3.3.2.

In words, the calibration pipeline selects time bins to represent the on-source spectrum and off-source spectrum in beamformer units (see the top panel of Figure 3–7), and takes the median values in time over these bins for each spectrum. The resulting on-source and off-source spectra for an observation of source 3C295 are shown in the second panel from the top in Figure 3–7. The off-source is subtracted from the on-source and then divided by the FFT beam sensitivity as a function of frequency using the formed beam model (see §2.3). The result is the spectrum of the source in beamformer units, still encoding attenuation due to the primary beam model. This spectrum is shown in the second panel from the bottom in Figure 3–7, where the 30 MHz ripple is clear. Finally, this FFT-corrected spectrum is divided by the known spectrum of the source to produce the final BF/Jy spectrum shown in the bottom panel of Figure 3–7.

The BF/Jy spectrum is saved in an NPZ file on the archiver, and metadata about the spectrum are sent to the calibration database. This metadata includes descriptive information (like the name of the calibrator, the date of the observation, the beam that the source transits through, the path of the BF/Jy spectrum NPZ on the archiver, etc) as well as a series of metrics about the data (like the band-integrated SNR of the source transit, the FWHM of the transited beam, the effective bandwidth left over after RFI removal, etc). A "data quality" metric was also developed in order to automatically detect transit observations that are significantly disrupted by solar transits, strong near field RFI, or data scaling

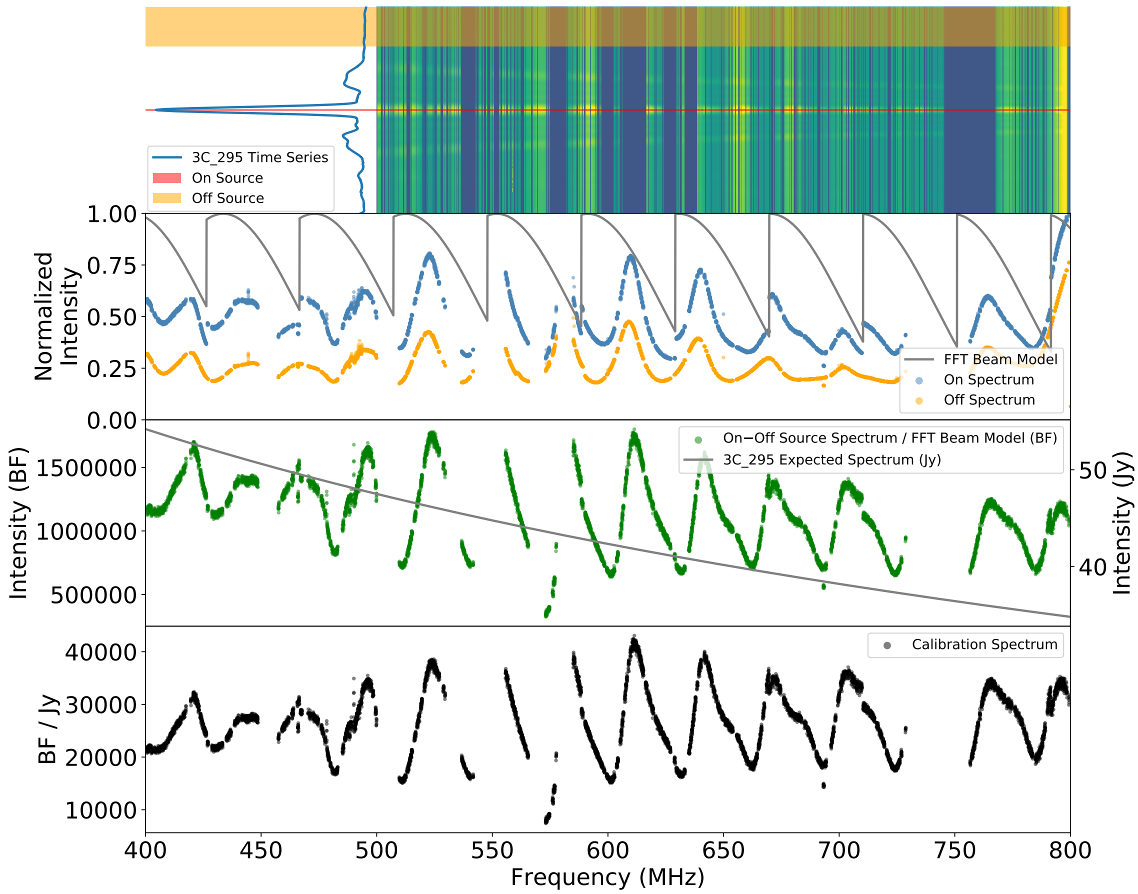


Figure 3–7: A series of plots showing different stages in the process of calculating a BF/Jy spectrum. The data in this figure come from a March 22nd, 2019 observation of Seyfert 2 galaxy 3C295 transiting through beam number 1134. The top panel shows the dynamic spectrum of the observation (Right), along with the frequency-summed time series (Left). For the dynamic spectrum, the x-axis corresponds to frequency ranging from 400 to 800 MHz, left to right, while the y-axis corresponds to time. The yellow region denotes the time bins chosen to represent the off-source spectrum, while the red region denotes the time bins chosen to represent the on-source spectrum. The panel second from the top shows the resulting on-source spectrum (blue dots), off-source spectrum (orange dots), and sensitivity of the formed beam at the location of the source during transit (grey solid line). The panel the second from the bottom shows the on−off source spectrum divided by the formed beam sensitivities in beamformer units (green dots) as well as the known spectrum of the source in Jy (grey solid line). The left-hand axis shows the source spectrum intensity in beamformer units, while the right-hand axis shows the known source spectrum intensity in Jy. The bottom panel shows the final resulting BF/Jy spectrum.

issues, rendering them unusable for flux calibration. Examples of these sorts of observations are shown in Figure 3–8. The data quality metric is calculated by comparing a given transit time series to an average template time series using the Pearson coefficient. The Pearson coefficient is a measure of linear correlation between two datasets, and it ranges from $-1$ to 1, where 1 is perfectly positively correlated. Through empirical testing, we found that a Pearson coefficient of greater than 0.95 indicates undisrupted or "good" data. Figure 3–9 shows the data quality of different calibration observations as a function of right ascension and time, demonstrating examples of the data quality metric catching solar transits and near field RFI.

### 3.3.4 Intensity Calibration Stage

The second half of the flux calibration pipeline, outlined in Figure 3–10, switches focus from processing calibrator observation data to processing FRB data. The "Intensity Calibration Stage," in particular, deals with applying a BF/Jy spectrum to FRB intensity data to obtain a calibrated waterfall.

Similarly to the steady source transit observations, raw FRB intensity data are saved on the archiver in msgpack format with 16,384 frequency channels (24 kHz resolution) and 0.983 ms time resolution. FRB data preprocessing is completed using `iautils`, a software package designed specifically for reading in and manipulating CHIME/FRB intensity data. This package was developed independently from the flux calibration pipeline by other members of the CHIME/FRB team. For each FRB that enters the flux calibration pipeline, msgpack data

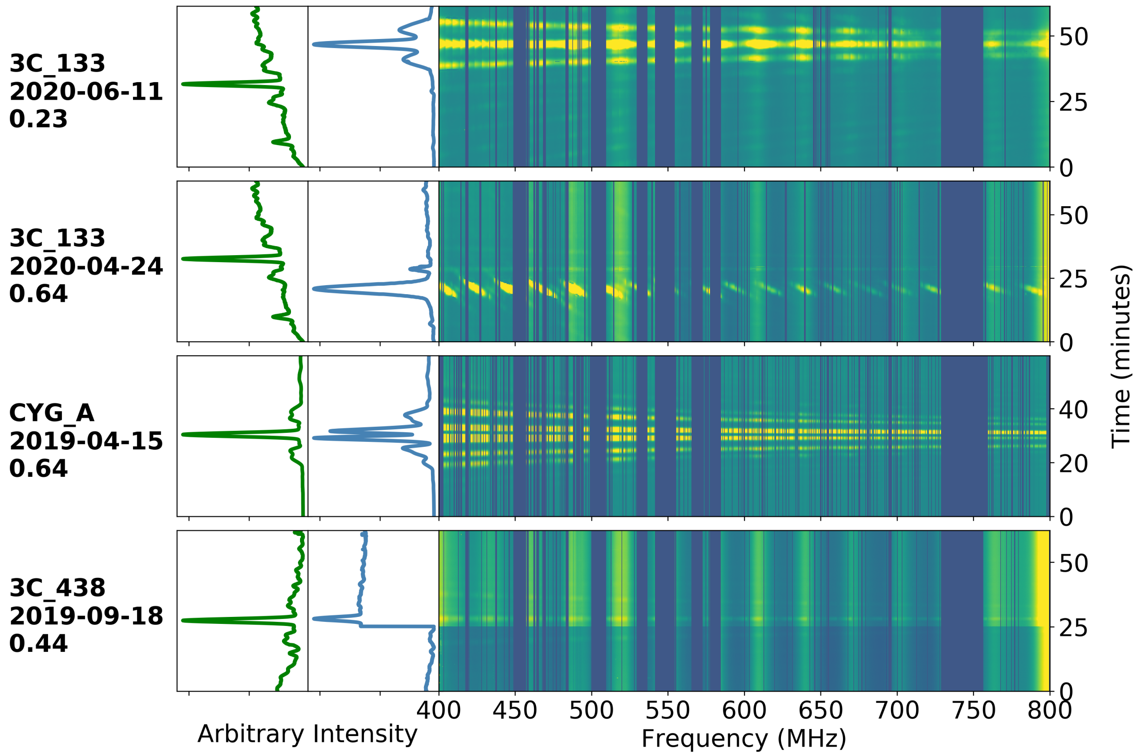


Figure 3–8: Examples of transit observations with bad data quality. The right column shows the "bad" transit dynamic spectrum, the middle column shows the corresponding "bad" transit time series, and the left column shows how the transit time series appears normally. Each observation is labeled on the left side of the plot by the calibration source that was transiting, the date, and the resulting data quality metric. Top to bottom, the plot shows examples of a solar transit in the main lobe, a solar transit in a far sidelobe of the primary beam, an unknown near field RFI source, and a data scaling jump.

are unpacked, organized into a 2D dynamic spectrum, zapped for RFI, and de-dispersed to the best-fit DM derived from `fitburst` (see §2.5.1).

After preprocessing, the FRB data are then calibrated. Each burst is paired with the calibration spectrum of the nearest steady source transit, closest first in zenith angle, then in time. We assume North-South beam symmetry, so

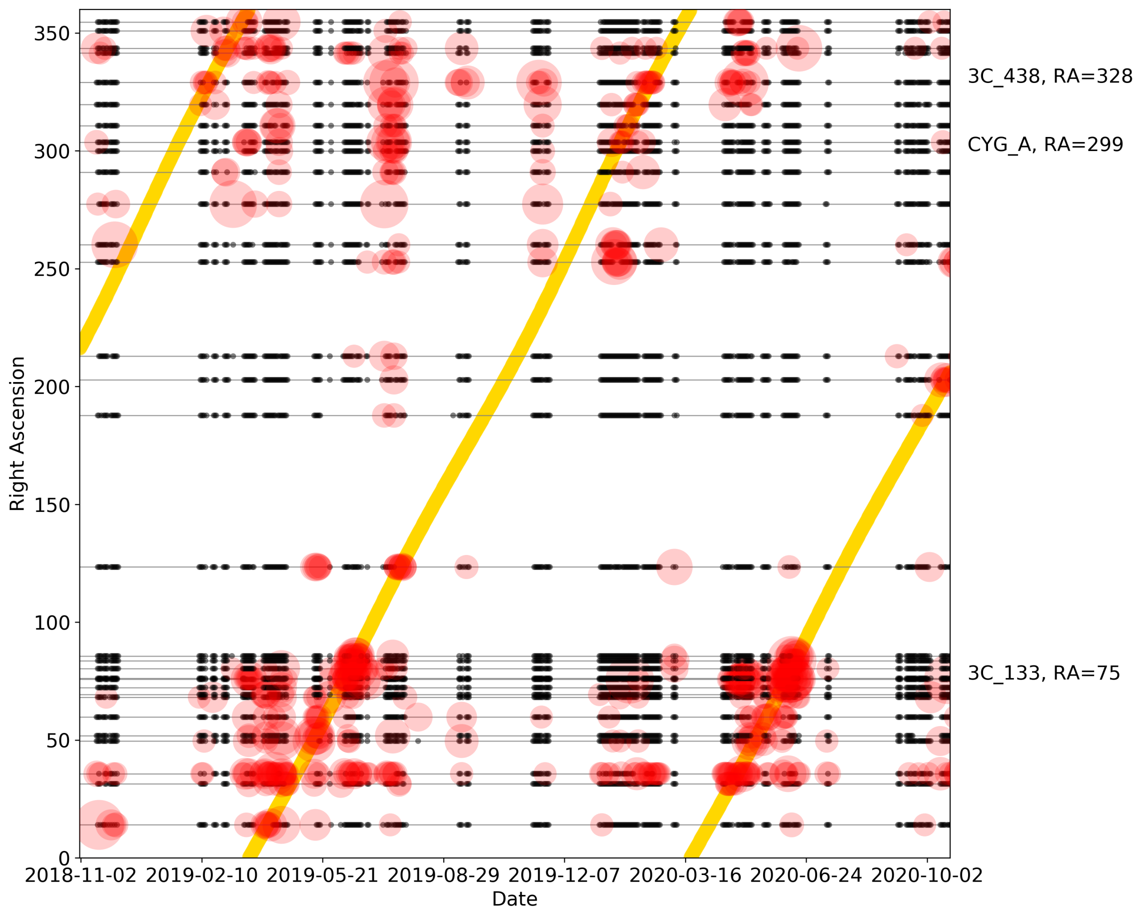


Figure 3–9: The data quality of the transit observations as a function of both time and right ascension. Each horizontal line represents a calibration source at a given right ascension, each black dot represents a “good” quality observation (data quality metric greater than 0.95). Red dots represent “bad” observations with the size of the dot correlating with the inverse of the quality metric. Yellow lines show the path of the Sun. Where the Sun crosses the calibrators, there tend to be clusters of “bad” observations (corresponding examples of main lobe and side lobe solar transits for calibration source 3C133 are shown in Figure 3–8). The vertical lines of bad observations on April 15th, 2019 and July 20th, 2019 show where the near field RFI was affecting observations (an example of this is shown in Figure 3–8 for a CygA observation).

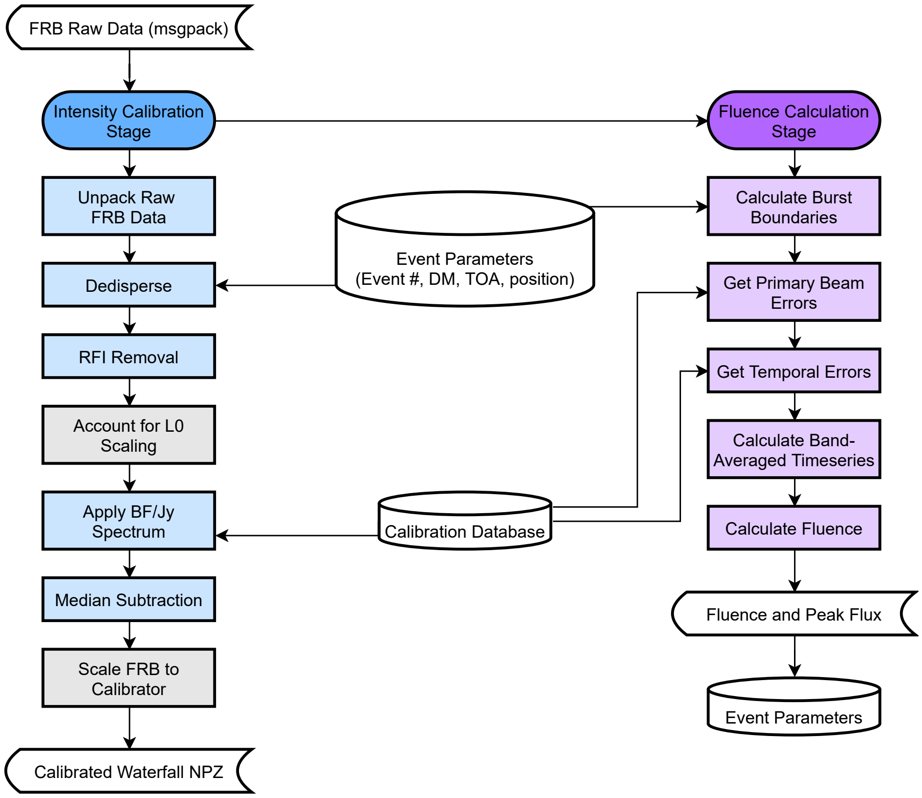


Figure 3–10: A flowchart of the second half of the CHIME/FRB flux calibration pipeline, which encompasses the intensity calibration and flux calibration calculations. Optional steps, that are not always executed, are greyed out.

that sources on both sides of zenith can be used to calibrate each event. Only calibration spectra with data quality metrics higher than 0.95 are used. If the automatic pipeline pairs a calibrator with an FRB more than two weeks apart, then the calibration proceeds but the results are flagged for inspection. If needed, the calibration can be manually re-run with a more appropriate calibrator. By dividing each frequency channel in the total-intensity FRB data (in beamformer units) by its corresponding BF/Jy conversion, we derive a dynamic spectrum in physical units (Jy) roughly corrected for North-South beam variations. As a final step, each frequency channel is subtracted by its median value. The resulting dynamic spectrum is saved in an NPZ on the archiver.

Figure 3–11 shows a spatial representation of the calibrator-FRB pairings for the bursts in the first catalog. 98% of FRBs are associated with a calibrator within 5 degrees in zenith angle (either on the same side of zenith or at North-South symmetric locations on opposite sides of zenith). Most of the remaining 2% of FRBs were detected at the edges of CHIME's North-South field of view (very high or low zenith angles), where 3C353 is the only available calibrator. Especially at high northern zenith angles, there is a paucity of calibrators since we currently do not dump observations for the lower transits of circumpolar sources. Note that these calibrator-FRB pairings are a function of spatial and temporal proximity as well as data quality. As a result, not all FRBs are associated with their nearest calibrator, as the data for that calibrator could be disrupted by solar transits for a couple of weeks at a time. This is the case for two FRBs detected $6-7$ degrees away from NGC7720, that are still paired with the source for calibration.

The temporal proximity of calibrators and FRBs was complicated for this first catalog, since the flux calibration pipeline was still in development for roughly half the period over which FRBs were being detected (see §3.3.2). As a result, before February 8th, 2019, FRBs could be paired with calibration observations taken two to three months before or after the burst arrival time. A total of 253 FRBs fall within this time period. The uncertainty in the fluence measurements due to these large separations is estimated in the "Fluence Calculation Stage" of the pipeline (§3.3.5). For FRBs detected after February 8th, 2019, temporal calibrator coverage is more regular, and 95% of bursts are associated with a calibrator within 2 weeks of the event time. The remaining 5% are cases where the calibrator data were significantly disrupted by solar transits.

As discussed in the introduction of §3.3, in general, we do not correct FRBs for attenuation due to their uncertain locations within the formed beam. However, after a more accurate model of the primary beam was developed in late 2019, we added an option into the pipeline that produces a per-frequency scaling between the location of the calibrator and the location of the FRB using the composite primary and synthesized beam model. This is useful functionality for when a burst has a precise localization determined using CHIME/FRB baseband or another long-baseline telescope. In these instances, we can apply this scaling to the dynamic spectrum to obtain accurate fluence and flux measurements corrected for synthesized beam attenuation, rather than lower bounds. This method was used to analyze the fluence distribution of the first repeating FRB with periodic activity, FRB 180916, which had a precise localization from the EVN (see §4.1; CHIME/FRB Collaboration et al., 2020b). For a test of this method, see §3.3.6.

Before the end of 2018, the system was still in a slightly chaotic state of commissioning. The Steady Source Acquisition Stage of the flux calibration pipeline was under development, so calibration observations were sparse. At the same time, the upstream CHIME complex gain calibration process was undergoing significant testing and changes. Some of these changes affected the way that the data were scaled during the complex gain calibration process, which in turn

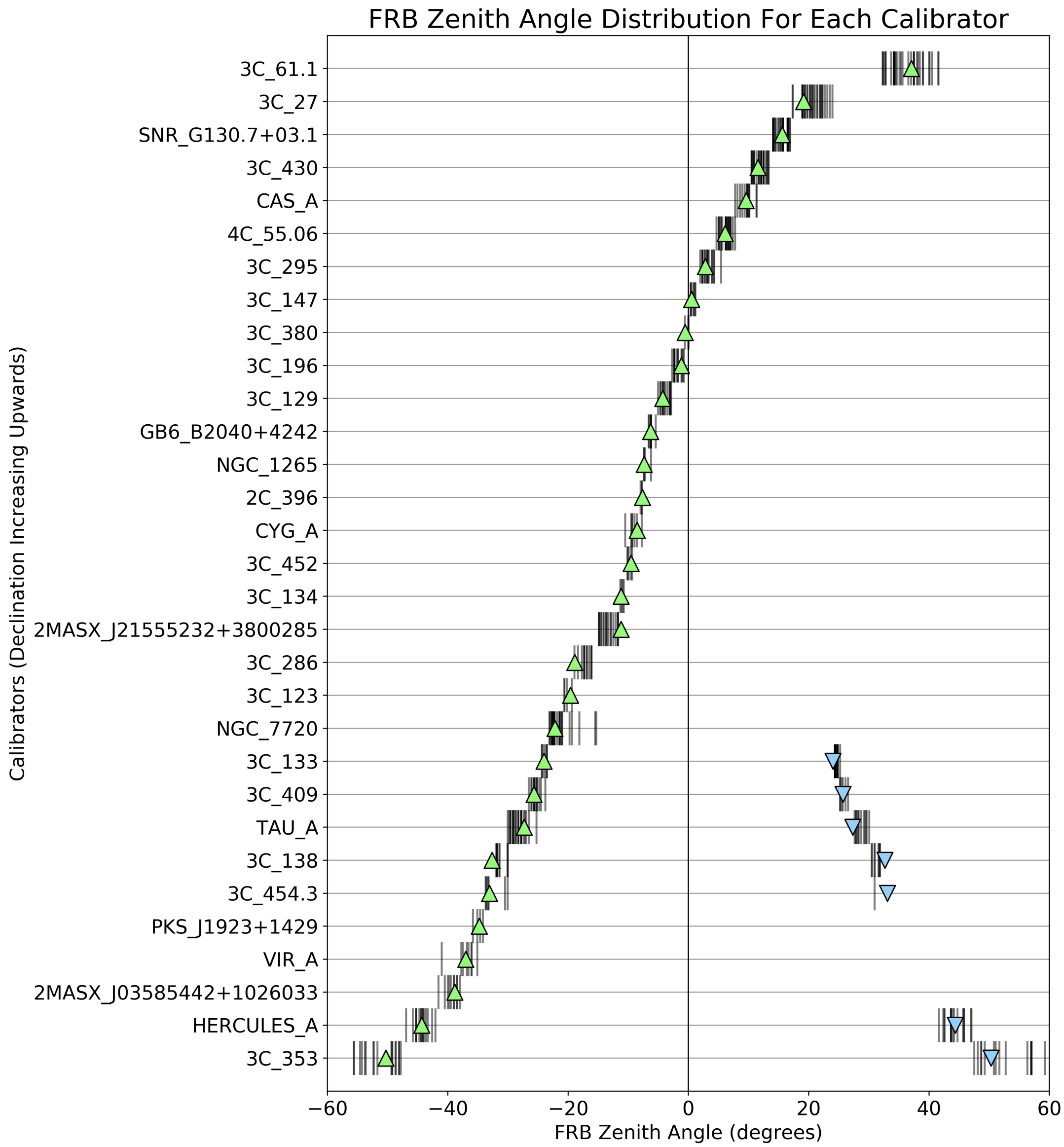


Figure 3–11: The association of FRBs with calibrators as a function of zenith angle for the first CHIME/FRB catalog (CHIME/FRB Collaboration et al., 2021). Each calibrator is labelled on the y-axis and represented by a horizontal line, with the y-axis sorted by increasing calibrator declination. The position of each calibrator is marked with an upward pointing green triangle. The symmetric position on the opposite side of zenith is marked with a downward pointing blue triangle. An FRB paired with a given calibrator is marked on the horizontal line of the calibrator by a small vertical line at the zenith angle of the FRB.

affected the overall scaling of CHIME/FRB intensity data downstream. These changes (marked in Figure 3–5) include:

- On September 7th, 2018, CHIME/FRB started pre-scaling the complex gains so that the standard deviation of the gain amplitude was 1. This resulted in an overall scaling of $\sigma_G^2$ in the resulting intensity data, where $\sigma_G$ is the standard deviation of the complex gain amplitude.
- On September 15th, 2018, the gain calculation procedure was updated to improve treatment of the two polarizations when determining the gains, introducing an empirically determined overall scaling factor of $\sim$3.246.
- On December 10th, 2018, we stopped pre-scaling the gains by the standard deviation, removing the $\sigma_G^2$ scaling.

The sparse coverage of calibration observations meant that some FRBs in the catalog needed to be calibrated using BF/Jy spectra determined across these date boundaries. In these cases, the scaling between the FRB and calibrator data is determined using records of the complex gains saved on the archiver. These scalings are applied to the BF/Jy spectrum before it is applied to the dynamic spectrum of the FRB. This process is encompassed in the "Account for L0 Scaling" step of Figure 3–10.

### 3.3.5 Fluence Calculation Stage

The final stage of the calibration pipeline deals with the calculation of FRB fluence and flux values from the calibrated dynamic spectra (the "Fluence Calculation Stage" in Figure 3–10).

The first step in the fluence calculation involves determining the boundaries of the burst extent in the dynamic spectrum. The fluence and flux values calculated from the pipeline are averaged over the entire 400−800 MHz CHIME band. For narrow-band bursts, this averages noise into our fluence and flux values. However, we choose to quote values from the same frequency range for consistency. In terms of determining the burst boundaries, this means that we only need to localize the burst along the temporal axis of the dynamic spectrum. This is accomplished using results from the `fitburst` routine, which models the burst morphology in time and frequency space and outputs fundamental burst parameters like the arrival time $t_{\mathrm{arr}}$, intrinsic width (the Gaussian $\sigma$), and the scattering time $\tau$ at 400 MHz (see §2.5.1 for a review). The start and end times encompassing the burst are defined by the $3\sigma$ Gaussian width, with an optional extra term added to the end time to account for a scattering tail, if present:

$$t_{\mathrm{start}} = t_{\mathrm{arr}} - 3\sigma \tag{3.9}$$

$$t_{\mathrm{end}} = t_{\mathrm{arr}} + 3\sigma + 5\tau \cdot (p < 0.001) \tag{3.10}$$

where $p$ is the p-value from an F-test comparison between scattered and un-scattered models fit to the burst. A p-value significance of 0.1% was used to declare the presence of significant scattering. In this case, we add five times the scattering time to the time denoting the end of the burst, corresponding to the time where the scattered emission would have decreased by a factor of $e^{-5} \approx 0.008$ (less than 1%).

Next, the band-averaged time series is derived by averaging the de-dispersed calibrated dynamic spectrum over the bandwidth remaining after RFI removal, and subtracting the resulting time series by the median of the off-pulse. The fluence is then calculated by integrating over the burst extent in the band-averaged time series, while the peak flux is the maximum value within the burst extent (at 0.983 ms resolution). Figure 3–12 shows an example band-averaged time series, with the burst extent and peak flux bin labeled.

The uncertainties on the fluence and flux values are estimated using steady source transit data. There are two main sources of error that are incorporated into the pipeline uncertainties: (1) the error due to differences in the primary beam response between the calibrator and the assumed FRB location along the meridian, or the "primary beam error," and (2) the error due to the temporal separation between the FRB and the calibrator transit, or the "time error" (this error encompasses temporal variations in the system sensitivity, as well as calibrator source variability). Each of these errors is first calculated as a relative or fractional uncertainty in each frequency channel (see the rightmost panel of Figure 3–12).

The primary beam error is estimated by using steady source observations from a single day to calibrate each other and measuring the average fractional error as a function of frequency compared to known flux values (the fractional error is given by: $(S_{\nu,\mathrm{meas}} - S_{\nu,\mathrm{exp}})/S_{\nu,\mathrm{exp}}$ where $S_{\nu,\mathrm{exp}}$ is the expected spectrum and $S_{\nu,\mathrm{meas}}$ is the measured spectrum). The calibrator pairs used to estimate this error are selected to match the spatial separation in zenith angle between the FRB and its paired calibrator. For example, if an FRB is less than 1° from its calibrator (in

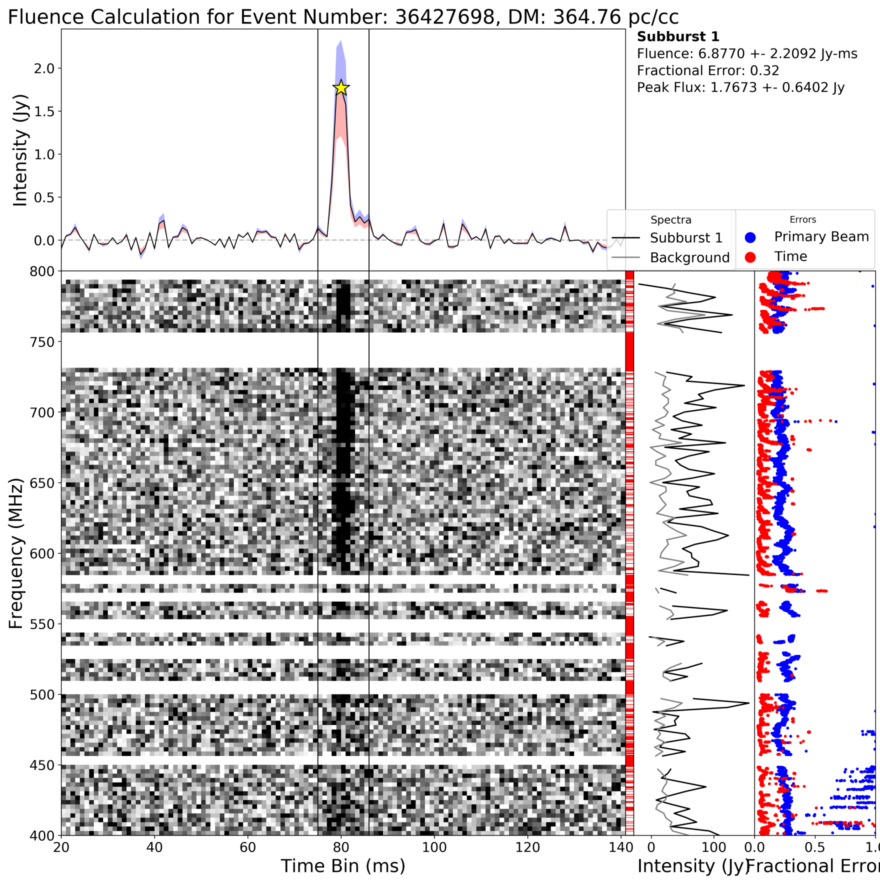


Figure 3–12: A diagnostic plot for the "Fluence Calculation Stage" of the flux calibration pipeline. The central panel is the dynamic spectrum of the burst, where the colour scale shows the intensity (Jy) in each frequency-time bin. Note that this dynamic spectrum is subbanded to 128 frequency channels for plotting clarity, but this is not included as a step in the fluence calculation. The white horizontal sections of the dynamic spectrum indicate subbanded channels that have been completely masked due to RFI. The red lines to the right of the dynamic spectrum indicate the frequencies that have been masked at full 16,834 resolution. The top panel shows the band-averaged time series of the burst (black line) as well as the upper bound (blue shading) and lower bound (red shading) uncertainty for each time bin. The yellow star indicates the bin from where the peak flux is taken. The vertical lines indicate the extent of the burst integrated over to obtain the fluence. The first panel to the right of the dynamic spectrum shows the subbanded average spectrum of both the burst (black) and the background (grey). The far right panel shows the fractional error in each frequency bin determined due to variations in time (red) and across the primary beam (blue).

zenith angle), then the primary beam errors are estimated using steady source pairs that are within $1^\circ$ of each other. If an FRB is $1-5^\circ$ from its calibrator, then steady source pairs are selected that are within $5^\circ$ of each other. If an FRB is $5-10^\circ$ from its calibrator, then steady source pairs are selected that are within $10^\circ$ of each other. If an FRB/calibrator pair are from opposite sides of zenith, then steady source pairs are selected from a similar distance on opposite sides of zenith. This primary beam error is typically on the order of $20-30\%$ (band-averaged), depending on the distance between the FRB and the calibrator. The left panel of Figure 3–13 shows the average fractional primary beam errors derived for different spatial separations.

For a given FRB, the time error is determined by measuring the rms variation in the BF/Jy spectra of the paired calibrator over a period of roughly two weeks surrounding the burst arrival (the fractional error is given by: $\Delta C_{\nu,\mathrm{BF/Jy}}/\overline{C}_{\nu,\mathrm{BF/Jy}}$ where $\Delta C_{\nu,\mathrm{BF/Jy}}$ is the rms of the BF/Jy spectra and $\overline{C}_{\nu,\mathrm{BF/Jy}}$ is the average). The time error is typically on the order of $10-20\%$, depending on the calibrator used (see the right panel of Figure 3–13). As previously mentioned, from the start of operation up until February 8th, 2019, the flux calibration pipeline was still being commissioned and steady source observations were sparse, resulting in FRB/calibrator pairings with temporal separations beyond two weeks. We conservatively estimate the time error for bursts detected during this time by taking the fractional rms variation in the calibration spectrum over the entire period, yielding errors typically on the order of $20-30\%$.

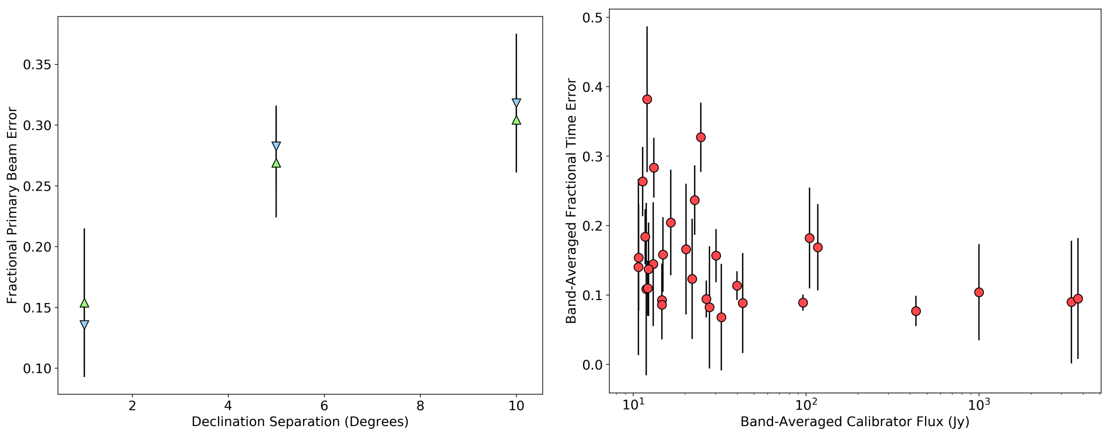


Figure 3–13: The fractional errors averaged over both bandwidth and bursts in the first CHIME/FRB catalog (CHIME/FRB Collaboration et al., 2021). (Left) The average fractional error due to the primary beam, split by separation between the FRB and calibrator (points from left to right correspond to: $<1°$, $1-5°$, $5-10°$). Upward pointing green triangles indicate FRB/calibrator pairs on the same side of zenith from each other, downward pointing blue triangles indicate pairs on the opposite side of zenith from each other. Error bars show the standard deviation of the averaged values. (Right) The average fractional error due to temporal variations, as a function of calibrator flux. Again, error bars show the standard deviation of the averaged values.

The next step involves applying the determined fractional errors to the FRB data and deriving the final errors on the fluence and flux values. First, the time and primary beam fractional errors are combined together to obtain an upper limit on the fractional uncertainty in each frequency channel. Since these errors are more systematic than purely random, with non-Gaussian underlying distributions, we choose to be conservative and sum them directly rather than add them in quadrature. Multiplying the dynamic spectrum by the combined fractional error along the frequency axis yields an uncertainty on the intensity in each frequency-time bin. We then band-average over the uncertainty dynamic spectrum, and add and subtract the resulting time series to the previously determined band-averaged time series for the burst. This essentially results in upper bound and lower bound burst profiles, which are indicated by the blue and red shading in the top panel of Figure 3–12. The error on the fluence is given by the average area between the upper and lower profiles, and the error on the peak flux is grabbed from the location of the maximum bin. Both of these errors are summed with the rms fluence and flux from the off-pulse region in the band-averaged time series to form the final, upper limit, error.

At the end of the pipeline, the resulting fluence, flux, and error values from the pipeline are pushed to the events database. We note that the estimated errors do not encapsulate the bias due to our assumption that each burst is detected along the meridian of the primary beam, which causes our fluence measurements to be biased low. As previously mentioned, the measurements produced from the pipeline are most appropriately interpreted as lower limits, with an uncertainty on the limiting value.

### 3.3.6 Injections Comparison

CHIME/FRB's high detection rate presents the opportunity for characterizing a large number of FRBs from a constant selection function, which is conducive for population studies. However, inferring the properties of the intrinsic FRB population requires a detailed accounting of the telescope selection function, i.e., which FRBs are being detected by the CHIME/FRB system, which are not, and how measured FRB properties compare to their intrinsic values. To probe this selection function, significant effort has been put towards developing a robust injections system for analyses completed in the first CHIME/FRB catalog. This system generates synthetic FRBs with user-generated properties, injects them into the real-time total-intensity data stream to be searched with the CHIME/FRB backend, and keeps track of the resulting properties measured from the real-time pipeline (including whether the FRB was detected or not). A detailed description of the injections infrastructure is outside the scope of this thesis (there is a description in the catalog paper; CHIME/FRB Collaboration et al. 2021). Here I describe an initial test of the flux calibration pipeline that I completed alongside members of the CHIME/FRB injections team.

Early in the commissioning of the CHIME/FRB experiment, the upstream complex gain calibration process was still in development. Over time, the procedure matured from phase-only calibration (during the pre-commissioning period until September 4th, 2018) to phase and amplitude calibration normalized to the daily transit of CygA (as of early 2019). Applying these phase and amplitude gains to the raw upstream baseband data essentially flux calibrates it up to a static primary beam model. Then, the process of FFT beamforming and upchannelization introduces a series of additional scalings, resulting in the observed "beamformer units" (BF) of the CHIME/FRB total-intensity data. In early April of 2020, after the system calibration and beamforming processes had been stable for a long time, these scaling factors were systematically tracked down,[12] and a conversion

[12] A valiant effort by collaboration member Kiyo Masui, see document #1182 in the internal CHIME doclib.

between the original upstream calibration and CHIME/FRB BF was determined:

$$1\,\mathrm{Jy} = \frac{(1{,}024 f_{\mathrm{good}})^2 \cdot 2 \cdot 128}{32 \cdot 0.806745 \cdot 400}\,\mathrm{BF} \approx 26{,}000 f_{\mathrm{good}}^2\,\mathrm{BF} \tag{3.11}$$

where $f_{\mathrm{good}}$ is the fraction of good feed inputs. This number is the only time-variable scaling factor introduced in the beamforming process (it generally varies between 70% and 95% based on environmental conditions around the telescope).

On April 20th, 2020, the L0 beamforming code was updated to account for the $f_{\mathrm{good}}^2$ scaling, so that CHIME/FRB intensity data taken after this date are calibrated in real-time up to the beam model and the static conversion factor given in Equation 3.11. Comparing CygA transit data calibrated using this conversion factor to the expected spectrum attenuated by the beam model shows that the calibrated flux is accurate to within $\sim$5%.

This new calibration method provided an avenue for testing the first-pass calibration pipeline using functionality from the injections infrastructure. This test involves simulating a series of synthetic FRB datasets, running them through the flux calibration pipeline, and comparing the resulting measured fluences to the nominal simulated values. Synthetic FRBs of given fluences and morphologies (temporal width, bandwidth, scattering, spectral index, etc) are generated with a piece of injections code called `simpulse`, producing dynamic spectra of the intrinsic FRBs in Jy units. Using the scaling factor given in Equation 3.11, the dynamic spectra are converted from Jy to BF units. Then the FRBs are scaled to a given location on-sky by applying the corresponding beam attenuation from the CHIME/FRB beam model (including the accurate, data-driven primary).

Finally, each pulse is added into an empty total-intensity dataset representing CHIME/FRB background noise.

We simulated a series of simple bursts with no scattering, intrinsic widths of 1 ms, flat spectra, and fluences of 0.5, 1, 3, 5, 10, 100, and 1,000 Jy ms. The bursts are injected at the centre of formed beam 1070. Each burst is run through the calibration pipeline twice: first without beam model scaling and then scaling between the location of the calibrator (in this case NGC 7720, which transits through beam 1071) and the location of the FRB, as described in §3.3.4. The results are displayed in Figure 3–14.

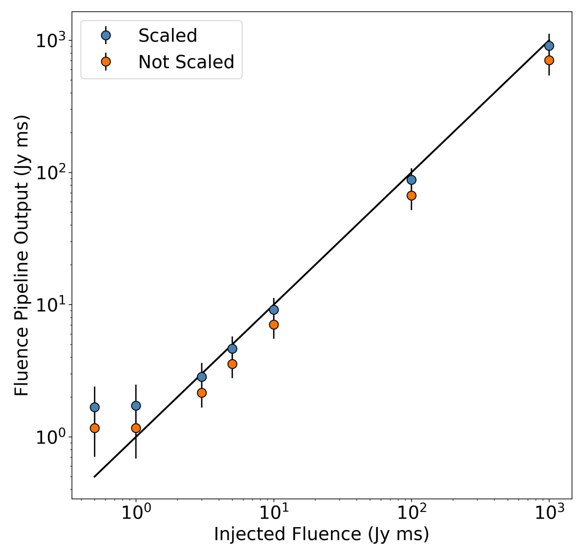


Figure 3–14: The nominal simulated burst fluence (before beam attenuation) versus the burst fluence derived from the flux calibration pipeline, for a series of bursts injected at the centre of formed beam 1070. The orange points indicate pipeline fluences derived without beam model scaling, and the blue points indicate pipeline fluences derived by scaling between the location of the calibrator and the location of the FRB. The error bars are the errors derived from the calibration pipeline. The black solid line denotes 1:1.

The injected fluence and flux pipeline fluence are in agreement within the errors output by the flux calibration pipeline. As the bursts are injected at the centre of a formed beam in the 1000 beam column, these injections represent a "best case scenario" for the fluence pipeline, which explains how even the lower-bound fluences are close to the injected values.

# CHAPTER 4
# Results and Future

The first-pass flux calibration pipeline described in this thesis was developed rapidly to provide burst characterization for early CHIME/FRB detections, extending from the first detection on July 25th, 2018 to the end of the first CHIME/FRB catalog nearly a year later on July 1st, 2019. As discussed in §3.2 and §3.3, the fluences calculated from this pipeline are lower limits, which could be underestimated by a factor of $\sim$10 or more for sidelobe detections (see, e.g., Figure 3–2). In addition, the flux calibration pipeline has not yet been integrated with the injections system, meaning that the instrumental selection function for pipeline fluence values is not yet characterized. As a result, the scientific output gleaned from these fluences has been very limited.

The limitations of the current flux calibration pipeline are symptomatic of the lag between the immediate scientific results afforded by early FRB detections, and our developing understanding of the technical intricacies involved with CHIME's novel design. Groundbreaking discoveries, like the existence of FRB emission down to 400 MHz (CHIME/FRB Collaboration et al., 2019a) and the identification of 18 new repeating FRB sources (CHIME/FRB Collaboration et al., 2019b; Fonseca et al., 2020), necessitated fluence constraints before subarcminute localization methods, a model of the primary beam, and integration with the injections system had been fully developed. However, once these technologies are available, more realistic fluences will be obtainable. Looking to the future, the next iteration of the CHIME/FRB flux calibration pipeline still holds immense potential for robustly probing fundamental questions related to FRB brightnesses.

In the final chapter of this thesis, we briefly review the preliminary scientific contributions of this early pipeline (§4.1), outline paths forward for the next fluence pipeline iteration (§4.2), and then finish with a few closing remarks (§4.3).

## 4.1 CHIME/FRB Science Results

The general procedure for the flux calibration pipeline was developed and manually applied to early bursts in order to obtain fluence and flux measurements for the first three scientific publications from CHIME/FRB: the existence of FRB emission down to 400 MHz (CHIME/FRB Collaboration et al., 2019a), the discovery of a second source of repeating FRBs (CHIME/FRB Collaboration et al., 2019c), and a CHIME/FRB detection of the first source of repeating FRBs, FRB 121102 (Josephy et al., 2019). The rest of the CHIME/FRB publications up to the present have used the automatic flux calibration pipeline that was developed for this thesis (a notable exception being the SGR 1935+2154 detection paper, see §4.2 and CHIME/FRB Collaboration et al. 2020a for more details). Here I highlight a few preliminary scientific analyses that I have completed using the automatic flux pipeline.

- **Repeater Peak Flux and Intrinsic Width Comparison:** In addition to the second source of repeating FRBs published in CHIME/FRB Collaboration et al. (2019c), CHIME/FRB has detected 17 new repeaters which have been published in two separate papers: eight in CHIME/FRB Collaboration et al. (2019b) (hereafter RN1) and nine in Fonseca et al. (2020) (hereafter RN2). Lower limit flux values derived from the automated flux pipeline are shown in Table 2 of each of these publications. A subset of the bursts in these papers are composed of multiple components in temporal succession, separated by timescales on the order of just a few milliseconds (see, for example, FRB 181222 in Figure 2 of RN1). In these cases, fluence and flux constraints were derived for each sub-burst.

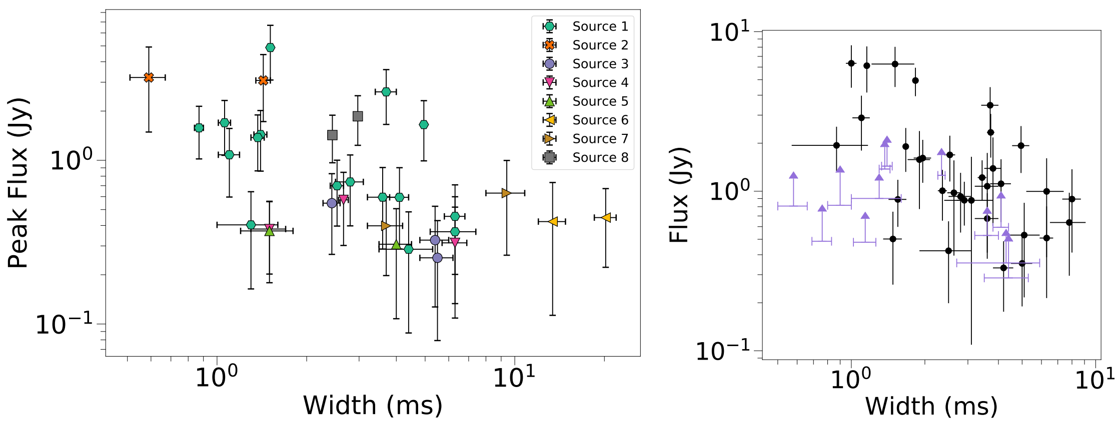


Figure 4–1: Plots showing the potential correlation between peak flux and intrinsic width for repeating FRBs detected with CHIME/FRB. (Left) The initial correlation for sub-bursts from all eight of the repeating sources detected in the RN1 paper (CHIME/FRB Collaboration et al., 2019b). The marker style and colour of each data point indicates which repeating source it came from. The bursts associated with each source are summarized in Table 2 of RN1. For example, Source 1 is FRB 180916. The peak fluxes in this panel are lower bounds. (Right) The potential correlation for the larger collection of sub-bursts from FRB 180916 (CHIME/FRB Collaboration et al., 2020b). Purple arrows indicate lower bound fluxes, while black dots indicate non-lower-bounds.

In RN1, I identified a potential correlation between the peak fluxes and intrinsic temporal widths of the sub-bursts in the sample, as shown in the left panel of Figure 4–1. To test for the significance of the correlation, I ran a Monte Carlo jackknife simulation that repeatedly resamples the peak flux and width values according to their uncertainties (treating the flux uncertainties as uniform and the width uncertainties as Gaussian) and then calculates correlation coefficients for the resulting sampled datasets. The result is a distribution of correlation coefficients and corresponding p-values. In linear space, I found a mean Spearman coefficient of $\sim -0.5$ with 93% of p-values indicating a greater than $2\sigma$ negative monotonic correlation. In log space, I found a mean Pearson coefficient of $\sim -0.5$ with 95% of p-values indicating a greater than $2\sigma$ log-log correlation. Both of these measurements provide evidence for a statistically marginal negative correlation, with the caveat that the peak flux values are lower limits that may be significantly underestimated. When RN2 was published, this analysis was repeated with the enlarged sample of 17 repeating sources with lower limit fluences, and the correlation disappeared.

To further explore this potential correlation, I also analyzed peak flux versus width for the larger sample of non-lower-bound fluxes from repeating source FRB 180916 presented in CHIME/FRB Collaboration et al. (2020b), as shown in the right panel of Figure 4–1 (see the next bullet for more information on how these more accurate fluxes were derived). Excluding the fluxes that are lower bounds, we complete the same analysis and find a mean Spearman coefficient of $\sim -0.5$ in linear space, with 97% of p-values indicating a greater than $2\sigma$ negative monotonic correlation, and a mean

Pearson coefficient of $\sim -0.5$ in log space, with 99% of p-values indicating a greater than $2\sigma$ log-log correlation.

Although the potential correlation has been restored with the availability of non-lower-bound flux values, we refrain from drawing any definite conclusions or interpretations until the instrumental selection function is quantified for fluence and flux values output from the calibration pipeline.

- **FRB 180916 Energy Distribution:** The most prolific new source of repeating FRBs presented in the RN1 paper was FRB 180916. Soon after its discovery with CHIME/FRB, follow-up observations with the EVN localized FRB 180916 to $\sim$2 milliarcsecond precision, placing it in the vicinity of a star-forming region on the outskirts of a host galaxy at a luminosity distance of $149.0 \pm 0.9$ Mpc (redshift $z \sim 0.0337 \pm 0.00021$; Marcote et al. 2020). After accruing 38 bursts from continued monitoring, CHIME/FRB discovered a 16.35 day periodicity in the clustering of burst arrival times from FRB 180916, with a $\sim$4 day active window each cycle (CHIME/FRB Collaboration et al., 2020b).

  This result was published in June 2020, after the CHIME/Cosmology primary beam model became available in the CHIME/FRB codebase. As described in §3.3.4, for the fluences and fluxes in this paper, we leverage the EVN localization and the robust composite primary and synthesized beam model to determine a per-frequency scaling between the location of the calibrator and the location of the FRB. By applying this scaling to the calibrated dynamic spectrum before generating the band-averaged time series, we can obtain more accurate fluences and fluxes (a test of this method using the injections system was described in §3.3.6). This scaling method was applied to the 25 bursts that were discovered while FRB 180916 was within the 600 MHz FWHM of the synthesized beam in which it was detected (otherwise, the burst could be completely attenuated to the noise floor in some parts of the band, see the top right panel of Figure 2–5). For the 13 other bursts in the CHIME/FRB sample, the standard calibration pipeline was used to obtain lower bound values. If multiple components were present in a given burst, fluence and flux values were measured for each sub-burst. From the 25 bursts detected within the FWHM of a synthesized beam, we derived a total of 33 sub-burst measurements.

The resulting fluence and flux values are listed in Extended Data Table 1 of CHIME/FRB Collaboration et al. (2020b). The left panel of Figure 4–2 shows the derived fluence and flux values and lower bounds for all of the sub-bursts in the CHIME/FRB sample as a function of activity phase (folding sub-burst arrival times at a period of 16.35 days), demonstrating that there is no clear trend in the 400−800 MHz fluence and flux versus phase. Using the 33 sub-bursts with accurate fluences, we also completed a first-pass attempt at analyzing the energy distribution of FRB 180916. The results of this analysis are presented in CHIME/FRB Collaboration et al. (2020b). The right panel of Figure 4–2 shows the resulting cumulative distribution function of sub-burst fluences, which we characterize by a power-law: $N(>F) \propto F^{\alpha+1}$ where $N$ is the number of bursts detected above

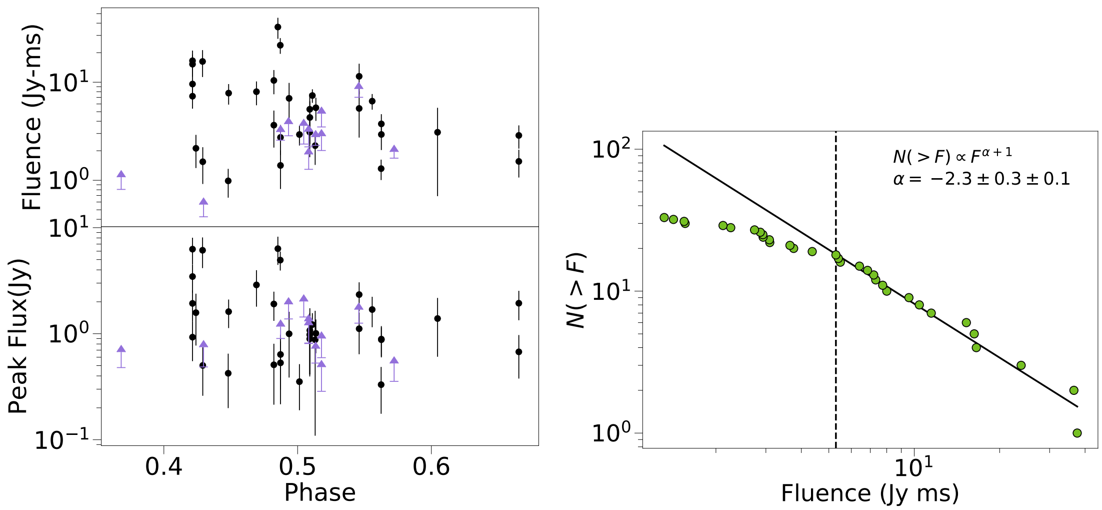


Figure 4–2: Fluence and flux results for FRB 180916 from the CHIME/FRB flux calibration pipeline. (Left) Fluence and flux values (black dots) and lower bounds (purple arrows) for all sub-bursts as a function of phase (assuming a 16.35-day activity period). No trend is seen. (Right) The cumulative distribution of sub-bursts in the CHIME/FRB sample, excluding burst detected beyond the 600 MHz FWHM of any CHIME/FRB synthesized beam. The black solid line represents the maximum-likelihood estimated power law described by the inset equation. The black dashed vertical line denotes the 5.3 Jy ms threshold below which fluence values are excluded from the maximum-likelihood estimation (see CHIME/FRB Collaboration et al. 2020b for more information).

a fluence of $F$. We determine $\alpha$, the power-law index of the differential distribution ($dN/dF \propto F^{\alpha}$), using the maximum-likelihood estimation methods laid out by Clauset et al. (2009) and Alstott et al. (2014). To avoid bias from the flattening of the distribution due to the telescope sensitivity limit, we exclude bursts below a threshold of 5.3 Jy ms in this calculation (see the "Burst characterization" section of CHIME/FRB Collaboration et al. 2020b for more information about how this threshold was chosen). To account for the large errors on the fluence values, we complete a Monte Carlo simulation that samples the fluences uniformly within their uncertainties, removes fluences below the 5.3 Jy ms threshold, and calculates $\alpha$ using a maximum-likelihood estimator for each resampling. The mean of the resulting distribution of power-law indices is $\alpha = -2.3 \pm 0.3 \pm 0.1$, where the first error is the statistical uncertainty from the maximum-likelihood estimator and the second error is the standard deviation of the $\alpha$ distribution.

This value stands as the first measurement of FRB 180916's power law index. In general, this value is consistent with the range of differential distribution indices measured for FRB 121102, which span from $\alpha \sim -1.7$ (derived from a heterogeneous sample of VLA, GBT, and Arecibo detections from 2−3 GHz; Law et al. 2017) to $\alpha = -2.8 \pm 0.3$ (derived from a sample of Arecibo detections at 1.4 GHz; Gourdji et al. 2019). However, none of these measurements has been robustly corrected for instrumental biases. Again, more definite analyses and conclusions about the intrinsic energy distribution will be completed once the CHIME/FRB selection function has been accounted for.

- **Catalog Fluences:** The automated flux calibration pipeline was used to obtain lower limit fluence and flux measurements for all 535 bursts in the first CHIME/FRB catalog (CHIME/FRB Collaboration et al., 2021).[1] Each

[1] The entire catalog is available for download at `https://www.chime-frb.ca/catalog`.

of the bursts was processed uniformly through the default pipeline, without any additional beam model scaling. Sub-burst structure was ignored for this analysis so that only one fluence and flux pairing was obtained for each burst.

As shown in the top left panel of Figure 4–3, the derived fluences generally correlate with detection SNR, as expected, with some spread determined by `bonsai`'s response to different burst properties and underestimation of the fluence values by calibration pipeline. For example, bursts with different temporal widths can result in detections at the same `bonsai` SNR, but those with larger widths will generally produce higher fluence measurements. This effect is demonstrated by the colouring of the points in the top left panel of Figure 4–3. The bottom left panel of Figure 4–3 shows the fluence and `bonsai` SNR values for a sample of synthetic bursts injected into and detected by the CHIME/FRB real-time pipeline using the injections infrastructure described in CHIME/FRB Collaboration et al. (2021). The fluences in this plot are the intrinsic values of the synthetic bursts, before beam attenuation (calibration for these injected bursts is completed using the upstream complex gain calibration as described in §3.3.6). The original injected population consisted of 19,000 bursts sampled from a Euclidean power-law fluence distribution ($N(>F) = F^{-1.5}$), distributed uniformly over CHIME's beam pattern, and synthesized with a range of realistic burst properties. Only the ∼1,500 bursts detected from this sample are shown in the bottom left panel of Figure 4–3. When compared to the correlation between the fluence and bonsai SNR of these injections, the correlation of the catalog values qualitatively agrees. However, the catalog is shifted lower in fluence and contains fewer high-fluence (>100 Jy ms) bursts, likely due to underestimation from the flux calibration pipeline.

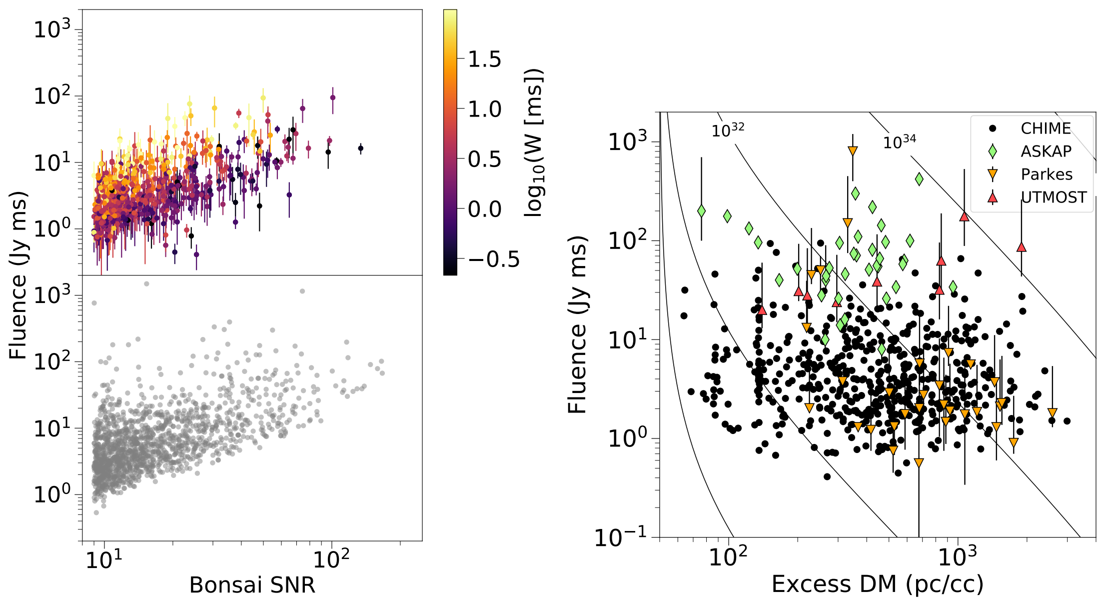


Figure 4–3: The distribution of catalog fluence values as a function of different variables. The top left panel shows the catalog fluences versus bonsai detection SNR. The points are coloured according to the broadened width of each burst. The bottom left panel shows the intrinsic fluence versus detected bonsai SNR for a sample of synthetic bursts injected into the CHIME/FRB real-time pipeline. The right panel shows the catalog fluences as a function of dispersion measure in excess of the Galactic contribution. The fluences of ASKAP- (1.4 GHz), Parkes- (1.3 GHz), and UTMOST-detected (843 MHz) bursts are plotted for comparison (these values are derived from http://frbcat.org/). Error bars are omitted for the CHIME/FRB bursts for the sake of plotting clarity. The solid black lines show contours of constant energy density in units of erg $Hz^{-1}$.

The right panel of Figure 4–3 shows the non-repeater catalog fluences plotted as a function of excess DM (burst DM subtracted by the Galactic contribution expected from the NE2001 model). Fluences from a sample of bursts detected at ASKAP (1.4 GHz), Parkes (1.3 GHz), and UTMOST (843 MHz) are also included for comparison. This plot is a recreation of Figure 2 in Shannon et al. (2018), which first identified a possible negative correlation between burst brightness and dispersion measure based on the fluence-DM distributions of the ASKAP and Parkes samples (see §1.4). Taking DM as a proxy for distance, Shannon et al. (2018) interpreted this negative correlation to mean that the high-fluence ASKAP bursts were nearby analogues of the more distant Parkes events. However, as demonstrated in the right panel of Figure 4–3, this negative correlation does not persist in the catalog fluences derived from the CHIME/FRB flux calibration pipeline, particularly for bursts with low excess DM ($<100\,\mathrm{pc\,cm^{-3}}$). The full implications of the CHIME/FRB fluence-DM distribution are still under investigation, with two primary prongs of inquiry: (1) Why don't we see the same trend in fluence and DM space, if we don't expect the underestimation of our fluences to be enhanced for low DM bursts? Could the original Shannon et al. (2018) correlation be caused by instrumental biases between the ASKAP and Parkes surveys? and (2) Why does CHIME/FRB not see as many FRBs at ASKAP-level brightnesses? Can this be solely accounted for by the beam attenuation that we are not correcting for in the flux calibration pipeline? This phase space will be particularly interesting to revisit once more accurate CHIME/FRB fluences are available.

## 4.2 CHIME/FRB Calibration Pipeline Future

The next iteration of CHIME/FRB flux calibration, bolstered with an accurate primary beam model, arcminute to sub-arcminute localizations from the intensity and baseband data pipelines, and a sophisticated injections system, will overcome many of the limitations of the first-pass pipeline described in this thesis. Here I outline two potential paths toward obtaining more accurate (non-lower-bound) fluence measurements from CHIME/FRB total-intensity data:

- **Beam Model Scaling:** Leveraging the higher precision localization methods, we could use the beam model to scale between the location of the calibrator and the location of the FRB. This method would be similar to the technique that we used in the FRB 180916 periodicity paper (CHIME/FRB Collaboration et al., 2020b), except using in-house CHIME/FRB baseband or intensity localizations rather than relying on precise localizations from other long-baseline telescopes. Unfortunately, this method would only work for burst that were detected approximately within the 600 MHz FWHM of a formed beam. Otherwise, the burst could be attenuated to the noise floor, so simply dividing by the beam model would amplify noise. One of the major advantages of this method is that it would be functional immediately upon the availability of intensity and baseband localizations, since the capability to scale between FRBs and their calibrators using the beam model already exists in the current flux calibration pipeline. This option could potentially act as a stepping stone until a more robust pipeline is fully tested and developed.

- **Forward Modeling:** The preferred, although more ambitious option, would be to develop the framework for forward modeling the intrinsic spectrum of each burst through the beam model in each of the detected beams. In this case, "forward modeling" means assuming some underlying spectral model for the FRB, multiplying that model by the beam model at the assumed burst location (either from independent localization or added as a parameter to the model), and then comparing the FRB data to the model and fitting parameters to deconstruct the intrinsic FRB spectrum excluding beam attenuation. This method was used to derive the fluence and flux of the SGR 1935+2154 burst detected in the far primary beam sidelobes of CHIME/FRB (see the "Estimate of burst fluence" section in CHIME/FRB Collaboration et al. 2020a), where we modeled the intrinsic spectrum as a power law with a spectral running parameter (as in §2.5.1). This method would require more code development and testing, but would reduce the noise introduced into the fluence calculation by all of the division and scaling involved in the first option. Additionally, this method would work even for events outside the 600 MHz FWHM of the formed beam.

Now that our data stream is calibrated in real-time up to the beam model and a static conversion factor (as described in §3.3.6), the next iteration of the flux calibration pipeline could also stop relying on total-intensity steady source transit observations, as they are no longer required to determine the BF/Jy conversion. This would save a significant amount of space on the archiver in the long-run.

Both of the options above would lead to significant improvement in our burst fluence measurements, as we would be calculating actual accurate fluence estimates rather than lower limits.

### 4.3 Closing Remarks

FRB science is a relatively young field that has recently been punctuated by rapid scientific and technological advances in tandem, from the discovery of periodicity in a repeating FRB (CHIME/FRB Collaboration et al., 2020b) to the localization of a population of bursts to their host galaxies[2] to the release of the first CHIME/FRB catalog (CHIME/FRB Collaboration et al., 2021). One of the most fundamental and prevailing mysteries about FRBs is their extreme energetics. With the exception of a single burst from Galactic magnetar SGR 1935+2154 (CHIME/FRB Collaboration et al., 2020a), FRBs are orders of magnitude more energetic than other radio transients of similar duration. Robust measurements of FRB brightness are crucial for FRB science, as they hold clues to possible progenitors through their energy distributions, maximum and minimum energy limits, magnitudes at different wavelengths, and correlations with other fundamental burst properties (§1.4). However, as described in §3.1.2, FRB flux calibration is a particularly difficult challenge that requires simultaneous knowledge of the time-variable telescope sensitivity, precise FRB localization, and characterization of the telescope beam pattern as a function of frequency and on-sky location. Due to the stringency of these requirements, most one-off FRB

[2] See `http://frbhosts.org`

fluence and flux measurements in the literature are lower bound estimates, and resulting fluence distributions are largely uncorrected for instrumental biases.

With the recent release of the first FRB catalog, the CHIME/FRB project has greatly increased the number of bursts detected to-date, ushering in a new era in the field focused on the detection of many FRBs for population studies rather than detailed study of individual FRB specimen. The past three years since the start of this thesis have been a particularly exciting whirlwind of technical development and scientific output for CHIME/FRB. In this thesis, I have described an automated pipeline that I developed in this time period for deriving lower-limit fluence and flux measurements for early CHIME/FRB detections. These measurements have been rendered particularly challenging by the novel nature of CHIME, however, the methodology in this pipeline has laid the groundwork for future accurate fluence measurements with CHIME/FRB. These measurements, integrated with the extensive injections system that quantifies CHIME/FRB's biases, will unlock CHIME as a powerful tool for population studies of FRB brightnesses.

References

Agarwal, D., Lorimer, D. R., Surnis, M. P., et al. 2020, MNRAS, 497, 352, doi: 10.1093/mnras/staa1927

Alstott, J., Bullmore, E., & Plenz, D. 2014, PLoS ONE, 9, e85777, doi: 10.1371/journal.pone.0085777

Baars, J. W. M., Genzel, R., Pauliny-Toth, I. I. K., & Witzel, A. 1977, A&A, 61, 99

Bailes, M., Jameson, A., Flynn, C., et al. 2017, PASA, 34, e045, doi: 10.1017/pasa.2017.39

Bandura, K., Cliche, J. F., Dobbs, M. A., et al. 2016a, Journal of Astronomical Instrumentation, 5, 1641004. https://arxiv.org/abs/1608.04347

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, 914522, doi: 10.1117/12.2054950

Bandura, K., Bender, A. N., Cliche, J. F., et al. 2016b, Journal of Astronomical Instrumentation, 5, 1641005. https://arxiv.org/abs/1608.06262

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

Bannister, K. W., Deller, A. T., Phillips, C., et al. 2019, Science, 365, 565, doi: 10.1126/science.aaw5903

Bassa, C., Tendulkar, S., Adams, E., et al. 2017, The Astrophysical Journal Letters, 843, L8

Bera, A., & Chengalur, J. N. 2019, MNRAS, 490, L12, doi: 10.1093/mnrasl/slz140

Berger, P., Newburgh, L. B., Amiri, M., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9906, Ground-based and Airborne Telescopes VI, 99060D, doi: 10.1117/12.2233782

Bhandari, S., Keane, E. F., Barr, E. D., et al. 2018, MNRAS, 475, 1427, doi: 10.1093/mnras/stx3074

Bhandari, S., Sadler, E. M., Prochaska, J. X., et al. 2020a, ApJL, 895, L37, doi: 10.3847/2041-8213/ab672e

Bhandari, S., Bannister, K. W., Lenc, E., et al. 2020b, ApJL, 901, L20, doi: 10.3847/2041-8213/abb462

Bochenek, C. D., Ravi, V., Belov, K. V., et al. 2020, arXiv e-prints, arXiv:2005.10828. https://arxiv.org/abs/2005.10828

Brandenberger, R., Cyr, B., & Varna Iyer, A. 2017, arXiv e-prints. https://arxiv.org/abs/1707.02397

Burgay, M., Esposito, P., Israel, G. L., et al. 2018, in IAU Symposium, Vol. 337, Pulsar Astrophysics the Next Fifty Years, ed. P. Weltevrede, B. B. P. Perera, L. L. Preston, & S. Sanidas, 319–321, doi: 10.1017/S1743921317008547

Caleb, M., Stappers, B. W., Rajwade, K., & Flynn, C. 2019, MNRAS, 484, 5500, doi: 10.1093/mnras/stz386

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

Chatterjee, S., Law, C. J., Wharton, R. S., et al. 2017, Nature, 541, 58. https://arxiv.org/abs/1701.01098

Chawla, P., Andersen, B. C., Bhardwaj, M., et al. 2020, ApJL, 896, L41, doi: 10.3847/2041-8213/ab96bf

Chen, G., Ravi, V., & Lu, W. 2020, ApJ, 897, 146, doi: 10.3847/1538-4357/ab982b

Cheng, Y., Zhang, G. Q., & Wang, F. Y. 2020, MNRAS, 491, 1498, doi: 10.1093/mnras/stz3085

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

—. 2019a, Nature, 566, 230, doi: 10.1038/s41586-018-0867-7

CHIME/FRB Collaboration, Andersen, B. C., Band ura, K., et al. 2019b, ApJL, 885, L24. https://arxiv.org/abs/1908.03507

CHIME/FRB Collaboration, Amiri, M., Bandura, K., et al. 2019c, Nature, 566, 235, doi: 10.1038/s41586-018-0864-x

CHIME/FRB Collaboration, Andersen, B. C., Bandura, K. M., et al. 2020a, arXiv e-prints, arXiv:2005.10324. https://arxiv.org/abs/2005.10324

CHIME/FRB Collaboration, Amiri, M., Andersen, B. C., et al. 2020b, Nature. https://arxiv.org/abs/2001.10275

CHIME/FRB Collaboration, :, Amiri, M., et al. 2021, arXiv e-prints, arXiv:2106.04352. https://arxiv.org/abs/2106.04352

Chittidi, J. S., Simha, S., Mannings, A., et al. 2020, arXiv e-prints, arXiv:2005.13158. https://arxiv.org/abs/2005.13158

Clauset, A., Shalizi, C. R., & Newman, M. E. J. 2009, SIAM Review, 51, 661, doi: 10.1137/070710111

Condon, J. J., & Ransom, S. M. 2016, Essential Radio Astronomy

Cordes, J. M., & Chatterjee, S. 2019, ARA&A, 57, 417, doi: 10.1146/annurev-astro-091918-104501

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

Cordes, J. M., & Wasserman, I. 2016, MNRAS, 457, 232, doi: 10.1093/mnras/stv2948

Cruces, M., Spitler, L. G., Scholz, P., et al. 2021a, MNRAS, 500, 448, doi: 10.1093/mnras/staa3223

—. 2021b, MNRAS, 500, 448, doi: 10.1093/mnras/staa3223

Deng, M., Campbell-Wilson, D., & CHIME Collaboration, T. 2014, "2014 16th International Symposium on Antenna Technology and Applied Electromagnetics (ANTEM)", 1. https://arxiv.org/abs/1708.08521

Deng, W., & Zhang, B. 2014, ApJL, 783, L35, doi: 10.1088/2041-8205/783/2/L35

Denman, N., Renard, A., Vanderlinde, K., et al. 2020, Journal of Astronomical Instrumentation, 9, 2050014, doi: 10.1142/S2251171720500142

Falcke, H., & Rezzolla, L. 2014, A&A, 562, A137, doi: 10.1051/0004-6361/201321996

Fanaroff, B. L., & Riley, J. M. 1974, MNRAS, 167, 31P, doi: 10.1093/mnras/167.1.31P

Farah, W., Flynn, C., Bailes, M., et al. 2019, MNRAS, 488, 2989, doi: 10.1093/mnras/stz1748

Fonseca, E., Andersen, B. C., Bhardwaj, M., et al. 2020, ApJL, arXiv:2001.03595. https://arxiv.org/abs/2001.03595

Foster, G., Hickish, J., Magro, A., Price, D., & Zarb Adami, K. 2014, MNRAS, 439, 3180, doi: 10.1093/mnras/stu188

Fuller, J., & Ott, C. D. 2015, MNRAS, 450, L71, doi: 10.1093/mnrasl/slv049

Gajjar, V., Siemion, A. P. V., Price, D. C., et al. 2018, ApJ, 863, 2, doi: 10.3847/1538-4357/aad005

Geng, J. J., & Huang, Y. F. 2015, ApJ, 809, 24, doi: 10.1088/0004-637X/809/1/24

Göğüş, E., Woods, P. M., Kouveliotou, C., et al. 1999, ApJL, 526, L93

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

Hashimoto, T., Goto, T., Wang, T.-W., et al. 2020, MNRAS, 494, 2886, doi: 10.1093/mnras/staa895

—. 2019, MNRAS, 488, 1908, doi: 10.1093/mnras/stz1715

Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, A&AS, 47, 1

Heintz, K. E., Prochaska, J. X., Simha, S., et al. 2020, ApJ, 903, 152, doi: 10.3847/1538-4357/abb6fb

Hessels, J. W. T., Spitler, L. G., Seymour, A. D., et al. 2019, ApJL, 876, L23, doi: 10.3847/2041-8213/ab13ae

Hilmarsson, G. H., Michilli, D., Spitler, L. G., et al. 2020, arXiv e-prints, arXiv:2009.12135. https://arxiv.org/abs/2009.12135

Hobbs, G., Manchester, R. N., Dunning, A., et al. 2020, PASA, 37, e012, doi: 10.1017/pasa.2020.2

Hogg, D. W. 1999, arXiv e-prints, astro. https://arxiv.org/abs/astro-ph/9905116

Inoue, S. 2004, MNRAS, 348, 999

Ioka, K. 2003, ApJL, 598, L79, doi: 10.1086/380598

Ioka, K., & Zhang, B. 2020, arXiv e-prints, arXiv:2002.08297. https://arxiv.org/abs/2002.08297

James, C. W., Ekers, R. D., Macquart, J. P., Bannister, K. W., & Shannon, R. M. 2019, MNRAS, 483, 1342, doi: 10.1093/mnras/sty3031

Jiang, P., Yue, Y., Gan, H., et al. 2019, Science China Physics, Mechanics, and Astronomy, 62, 959502, doi: 10.1007/s11433-018-9376-1

Jiang, P., Tang, N.-Y., Hou, L.-G., et al. 2020, Research in Astronomy and Astrophysics, 20, 064, doi: 10.1088/1674-4527/20/5/64

Josephy, A., Chawla, P., Fonseca, E., et al. 2019, arXiv e-prints, arXiv:1906.11305. https://arxiv.org/abs/1906.11305

Kashiyama, K., Ioka, K., & Mészáros, P. 2013, ApJL, 776, L39, doi: 10.1088/2041-8205/776/2/L39

Katz, J. I. 2018, Progress in Particle and Nuclear Physics, 103, 1, doi: 10.1016/j.ppnp.2018.07.001

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

Keane, E. F., Stappers, B. W., Kramer, M., & Lyne, A. G. 2012, MNRAS, 425, L71, doi: 10.1111/j.1745-3933.2012.01306.x

Kocz, J., Ravi, V., Catha, M., et al. 2019, MNRAS, 489, 919, doi: 10.1093/mnras/stz2219

Kumar, P., Lu, W., & Bhattacharya, M. 2017, MNRAS, 468, 2726, doi: 10.1093/mnras/stx665

Kunjaya, C., Mahasena, P., Vierdayanti, K., & Herlie, S. 2011, Ap&SS, 336, 455, doi: 10.1007/s10509-011-0790-y

Law, C. J., Bower, G. C., Burke-Spolaor, S., et al. 2015, ApJ, 807, 16, doi: 10.1088/0004-637X/807/1/16

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

Law, C. J., Omand, C. M. B., Kashiyama, K., et al. 2019, ApJ, 886, 24, doi: 10.3847/1538-4357/ab4adb

Law, C. J., Butler, B. J., Prochaska, J. X., et al. 2020, ApJ, 899, 161, doi: 10.3847/1538-4357/aba4ac

Levin, Y., Beloborodov, A. M., & Bransgrove, A. 2020, arXiv e-prints, arXiv:2002.04595. https://arxiv.org/abs/2002.04595

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

Lorimer, D. R., & Kramer, M. 2004, Handbook of Pulsar Astronomy, Vol. 4

Lu, W., & Kumar, P. 2019, MNRAS, 483, L93, doi: 10.1093/mnrasl/sly200

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

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

—. 2020, ApJ, 897, 1, doi: 10.3847/1538-4357/ab97b5

Lyutikov, M. 2020, ApJ, 889, 135, doi: 10.3847/1538-4357/ab55de

Lyutikov, M., Barkov, M. V., & Giannios, D. 2020, ApJL, 893, L39, doi: 10.3847/2041-8213/ab87a4

Lyutikov, M., Burzawa, L., & Popov, S. B. 2016, MNRAS, 462, 941, doi: 10.1093/mnras/stw1669

Macquart, J. P., & Ekers, R. 2018a, MNRAS, 480, 4211, doi: 10.1093/mnras/sty2083

—. 2018b, MNRAS, 480, 4211, doi: 10.1093/mnras/sty2083

Macquart, J. P., & Ekers, R. D. 2018c, MNRAS, 474, 1900, doi: 10.1093/mnras/stx2825

Macquart, J.-P., Shannon, R. M., Bannister, K. W., et al. 2019, ApJL, 872, L19, doi: 10.3847/2041-8213/ab03d6

Macquart, J.-P., Bailes, M., Bhat, N. D. R., et al. 2010, PASA, 27, 272, doi: 10.1071/AS09082

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

Manchester, R. N., Fan, G., Lyne, A. G., Kaspi, V. M., & Crawford, F. 2006, ApJ, 649, 235, doi: 10.1086/505461

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

Manchester, R. N., Lyne, A. G., Camilo, F., et al. 2001, MNRAS, 328, 17, doi: 10.1046/j.1365-8711.2001.04751.x

Maranda, B. 1989, The Journal of the Acoustical Society of America, 86, 1813, doi: 10.1121/1.398614

Marcote, B., Paragi, Z., Hessels, J. W. T., et al. 2017, ApJL, 834, L8. https://arxiv.org/abs/1701.01099

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

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

Masui, K., Lin, H.-H., Sievers, J., et al. 2015, Nature, 528, 523. https://arxiv.org/abs/1512.00529

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

McConnell, D., Allison, J. R., Bannister, K., et al. 2016, PASA, 33, e042, doi: 10.1017/pasa.2016.37

McKinnon, M. M. 2014, PASP, 126, 476, doi: 10.1086/676975

Mereghetti, S., Savchenko, V., Ferrigno, C., et al. 2020, ApJL, 898, L29, doi: 10.3847/2041-8213/aba2cf

Metzger, B. D., Berger, E., & Margalit, B. 2017, ApJ, 841, 14, doi: 10.3847/1538-4357/aa633d

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

Michilli, D., Seymour, A., Hessels, J. W. T., et al. 2018, Nature, 533, 132

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

Michilli, D., Masui, K. W., Mckinven, R., et al. 2020, arXiv e-prints, arXiv:2010.06748. https://arxiv.org/abs/2010.06748

Mingarelli, C. M. F., Levin, J., & Lazio, T. J. W. 2015, ApJL, 814, L20, doi: 10.1088/2041-8205/814/2/L20

Mucci, R. 1984, IEEE Transactions on Acoustics, Speech, and Signal Processing, 32, 548, doi: 10.1109/TASSP.1984.1164359

Newburgh, L. B., Addison, G. E., Amiri, M., et al. 2014, Proc. SPIE, 9145, 91454V. https://arxiv.org/abs/1406.2267

Ng, C., Vanderlinde, K., Paradise, A., et al. 2017, in Proceedings of XXXIInd General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS), 1–4, doi: 10.23919/URSIGASS.2017.8105318

Nicholl, M., Williams, P. K. G., Berger, E., et al. 2017, ApJ, 843, 84, doi: 10.3847/1538-4357/aa794d

Niino, Y. 2018, ApJ, 858, 4, doi: 10.3847/1538-4357/aab9a9

O'Neil, K. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 278, Single-Dish Radio Astronomy: Techniques and Applications, ed. S. Stanimirovic, D. Altschuler, P. Goldsmith, & C. Salter, 293–311. https://arxiv.org/abs/astro-ph/0203001

Oostrum, L. C., van Leeuwen, J., Attema, J., et al. 2017, The Astronomer's Telegram, 10693, 1

Oostrum, L. C., Maan, Y., van Leeuwen, J., et al. 2020, A&A, 635, A61, doi: 10.1051/0004-6361/201937422

Opperman, N., & Pen, U.-L. 2017, ArXiv e-prints. https://arxiv.org/abs/1705.04881

Otobe, E., Nakajima, J., Nishibori, K., et al. 1994, PASJ, 46, 503

Parent, E., Chawla, P., Kaspi, V. M., et al. 2020, ApJ, 904, 92, doi: 10.3847/1538-4357/abbdf6

Perley, R. A., & Butler, B. J. 2016, ArXiv e-prints. https://arxiv.org/abs/1609.05940

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

Petroff, E., Barr, E. D., Jameson, A., et al. 2016, PASA, 33, e045. https://arxiv.org/abs/1601.03547

Petroff, E., Burke-Spolaor, S., Keane, E. F., et al. 2017, MNRAS, 469, 4465, doi: 10.1093/mnras/stx1098

Petrosian, V. 1969, ApJ, 155, 1029, doi: 10.1086/149932

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

Pol, N., Lam, M. T., McLaughlin, M. A., Lazio, T. J. W., & Cordes, J. M. 2019, ApJ, 886, 135, doi: 10.3847/1538-4357/ab4c2f

Popov, M. V., & Stappers, B. 2007, A&A, 470, 1003, doi: 10.1051/0004-6361:20066589

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

Prochaska, J. X., Macquart, J.-P., McQuinn, M., et al. 2019, Science, 366, 231, doi: 10.1126/science.aay0073

Rajwade, K. M., Mickaliger, M. B., Stappers, B. W., et al. 2020, MNRAS, 495, 3551, doi: 10.1093/mnras/staa1237

Ravi, V. 2019, Nature Astronomy, 405, doi: 10.1038/s41550-019-0831-y

Ravi, V., Shannon, R. M., Bailes, M., et al. 2016, Science, 354, 1249. https://arxiv.org/abs/1611.05758

Ravi, V., Catha, M., D'Addario, L., et al. 2019, Nature, 572, 352, doi: 10.1038/s41586-019-1389-7

Remazeilles, M., Dickinson, C., Banday, A. J., Bigot-Sazy, M.-A., & Ghosh, T. 2015, MNRAS, 451, 4311

Rickett, B. J. 1977, ARA&A, 15, 479, doi: 10.1146/annurev.aa.15.090177.002403

Romero, G. E., del Valle, M. V., & Vieyro, F. L. 2016, PhRvD, 93, 023001, doi: 10.1103/PhysRevD.93.023001

Salpeter, E. E. 1969, Nature, 221, 31, doi: 10.1038/221031a0

Scheuer, P. A. G. 1968, Nature, 218, 920, doi: 10.1038/218920a0

Scholz, P., Spitler, L. G., Hessels, J. W. T., et al. 2016, ApJ, 833, 177. https://arxiv.org/abs/1603.08880

Shand, Z., Ouyed, A., Koning, N., & Ouyed, R. 2016, Research in Astronomy and Astrophysics, 16, 80, doi: 10.1088/1674-4527/16/5/080

Shannon, R. M., Macquart, J.-P., Bannister, K. W., et al. 2018, Nature, 562, 386, doi: 10.1038/s41586-018-0588-y

Spitler, L. G., Cordes, J. M., Hessels, J. W. T., et al. 2014, ApJ, 790, 101. https://arxiv.org/abs/1404.2934

Spitler, L. G., Scholz, P., Hessels, J. W. T., et al. 2016, Nature, 531, 202. https://arxiv.org/abs/1603.00581

Staveley-Smith, L., Wilson, W. E., Bird, T. S., et al. 1996, PASA, 13, 243

Taylor, J. H. 1974, AAPS, 15, 367

Tegmark, M., & Zaldarriaga, M. 2009, PhRvD, 79, 083530, doi: 10.1103/PhysRevD.79.083530

—. 2010, PhRvD, 82, 103501, doi: 10.1103/PhysRevD.82.103501

Tendulkar, S. P., Bassa, C. G., Cordes, J. M., et al. 2017, ApJL, 834, L7. https://arxiv.org/abs/1701.01100

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

Totani, T. 2013, PASJ, 65, L12, doi: 10.1093/pasj/65.5.L12

Venkatraman Krishnan, V., Flynn, C., Farah, W., et al. 2020, MNRAS, 492, 4752, doi: 10.1093/mnras/staa111

Vollmer, B., Gassmann, B., Derrière, S., et al. 2010, A&A, 511, A53, doi: 10.1051/0004-6361/200913460

Wadiasingh, Z., & Timokhin, A. 2019, ApJ, 879, 4, doi: 10.3847/1538-4357/ab2240

Yang, H., & Zou, Y.-C. 2020, ApJL, 893, L31, doi: 10.3847/2041-8213/ab800f

Yao, J. M., Manchester, R. N., & Wang, N. 2017, ApJ, 835, 29. https://arxiv.org/abs/1610.09448

Zackay, B., & Ofek, E. O. 2017, ApJ, 835, 11, doi: 10.3847/1538-4357/835/1/11

Zanazzi, J. J., & Lai, D. 2020, ApJL, 892, L15, doi: 10.3847/2041-8213/ab7cdd

Zhang, B. 2016, ApJL, 827, L31, doi: 10.3847/2041-8205/827/2/L31

Zhang, B. 2018, The Astrophysical Journal Letters, 867, L21

Zhang, B. 2018, ApJL, 854, L21, doi: 10.3847/2041-8213/aaadba

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

Zheng, Z., Ofek, E. O., Kulkarni, S. R., Neill, J. D., & Juric, M. 2014, ApJ, 797, 71, doi: 10.1088/0004-637X/797/1/71

Zhu, W., Li, D., Luo, R., et al. 2020, ApJL, 895, L6, doi: 10.3847/2041-8213/ab8e46